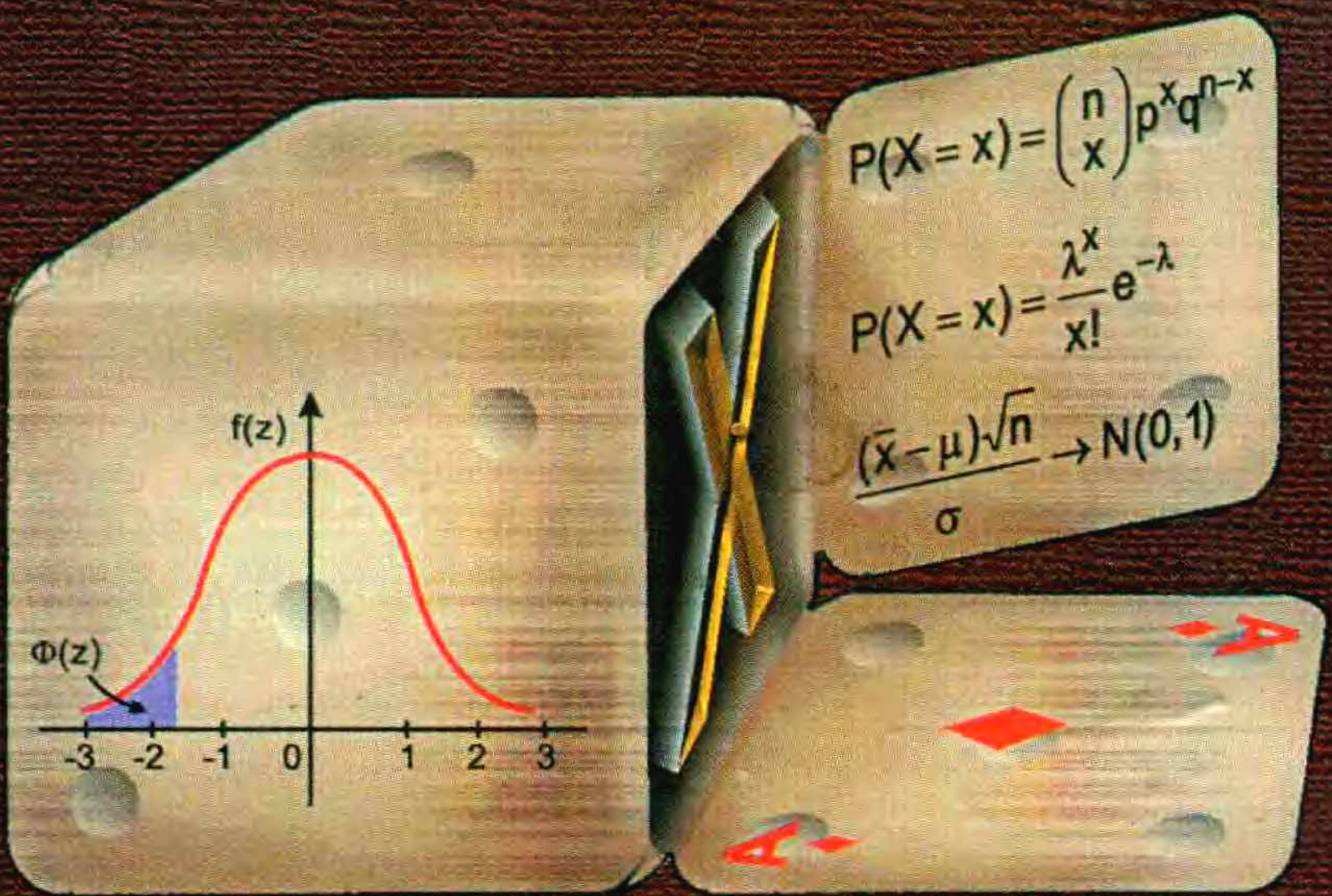
RUFINO MOYA C.
GREGORIO SARAVIA A.



PROBABILIDAD E INFERENCIA ESTADISTICA

EDITORIAL SM

PROBABILIDAD

E

INFERENCIA

ESTADÍSTICA

SEGUNDA EDICIÓN

RUFINO MOYA C.
Facultad de Ciencias Naturales y Matemática.
Universidad Nacional del Callao, Perú.

GREGORIO SARAVIA A.
Departamento de Estadística.
Universidad Federal de Minas Gerais, Brasil.

PROBABILIDAD E INFERENCIA ESTADÍSTICA

Rufino Moya C.
Gregorio Saravia A.

Impreso en Perú — Printed in Peru

Hecho el depósito legal, Ley Nº 26905

Composición y Diagramación: Rufino Moya C.
Montaje: **Editorial San Marcos**
RUC: 11029221

PRÓLOGO A LA SEGUNDA EDICIÓN

El deseo de mejorar el contenido y la exposición pedagógica de la primera edición nos ha llevado a preparar ésta. Los consejos de algunos colegas y la experiencia con la primera edición ha permitido escribir este, que esperamos constituya un mejor texto. Asimismo esperamos que las adiciones e innovaciones contribuya a conseguir estos objetivos.

Esta versión presenta el mismo esquema de la primera edición. Así, por la diversidad de sus ejemplos y problemas propuestos este libro será de gran ayuda a estudiantes de Estadística, Ingeniería, Economía, Biología, Ciencias, etc. Y consta de 9 capítulos, los 6 primeros tratan el cálculo de probabilidades. Éstos abarcan la probabilidad, aquí se ha introducido nociones de la teoría de confiabilidad; las variables aleatorias; las distribuciones de probabilidad discretas y continuas. En las distribuciones discretas se consideran también la multinomial y una generalización de la hipergeométrica. Los capítulos 7, 8 y 9 abordan la inferencia estadística.

Muchas personas han colaborado para la existencia de este libro. Destacan: VALENTÍN R. VEGA SALAS gerente general de Editorial Ciencias por per-

mitirnos usar libremente su taller y brindado todas las facilidades; NANCY PEÑA HUERTA secretaria de Editorial Ciencias en el mecanografiado del original; AMÉRICO CULQUE CARRILLO estudiante de la Facultad de Ciencias Económicas de la Universidad Nacional del Callao en la difícil tarea de corrección del mecanografiado y el montaje; Lic. en matemática ARMANDO VENERO B. quien estuvo siempre dispuesto a colaborar desinteresadamente en las gráficas y con sus consejos; ANÍBAL PAREDES GALVÁN por haber aceptado el reto de Editar este texto.

También expresamos nuestro agradecimiento a todos los colegas de las diferentes Universidades del país que nos han honrado al utilizar la primera Edición, especialmente a JUAN BAZÁN BACA, SERGIO LEYVA HARO, RICARDO POMALAYA Y ROLANDO M. CANALES DEL MAR profesores de la Universidad Nacional del Callao. Finalmente, en cuanto a los errores hemos tratado de corregirlos todos, sin embargo suponemos que se nos ha pasado algunos, así que cuando los encuentre escríbenos a la Facultad de Ciencias Naturales y Matemática de la Universidad arriba mencionada.

Rufino Moya C.
Gregorio Saravia A.

CONTENIDO

1

PROBABILIDAD

1.1 INTRODUCCION

Es difícil exagerar la importancia de la teoría de probabilidades; en muchos problemas de Ingeniería, Administración, Economía etc. necesitamos tomar decisiones frente a la incertidumbre. Para un Ingeniero posiblemente no tenga sentido el preguntarse: ¿Durante cuánto tiempo funcionará un determinado mecanismo?; pero si tendrá sentido el preguntarse y responder a la pregunta. ¿Cuál es la probabilidad que este mecanismo funcione más de 100 horas? o, ¿qué porcentaje de estos mecanismos funcionarán más de 100 horas?. Para un fabricante a gran escala tendrá sentido el preguntarse que porcentaje de sus productos serán aceptados en el mercado. A un candidato presidencial posiblemente no le interesa que Juan vote por él, pero si le interesará saber el porcentaje de electores que votarán por él.

En la mayoría de los problemas hay que tomar decisiones con base en experimentos. En este capítulo estudiaremos primero los experimentos aleatorios, luego a manera de pre-requisito recordar la teoría intuitiva de conjuntos, el análisis combinatorio; finalmente el concepto de probabilidad, sus propiedades y aplicaciones.

1.2 EXPERIMENTO ALEATORIO, ESPACIO MUESTRAL Y EVENTOS

1.2.1 EXPERIMENTO ALEATORIO

Los experimentos u operaciones reales o hipotéticos pueden dividirse - en dos clases: determinísticos y no determinísticos.

Un experimento es determinístico, si los resultados del experimento están completamente determinado y puede describirse por una ***fórmula matemática*** llamado tambien ***modelo determinístico***. Así, los siguientes ejemplos

(a) "Soltar una piedra en el aire"

(b) "Lanzar una pelota en un tanque de agua y ver si flota o se hunde". Son experimentos determinísticos, pues en el primer caso la piedra caerá,- aun más su movimiento se describe por las ecuaciones de caida libre,en el segundo caso la pelota flotará. También el siguiente experimento es determinístico .

(c) A un cuerpo de masa m en reposo, se somete a una fuerza constante F. El cuerpo se moverá con una aceleración constante

$$a = \frac{F}{m} \qquad \text{(segunda ley de Newton)}$$

Un experimento es no determinístico Si los resultados del experimento no puede predecirse con exactitud antes de realizar el experimento.

EJEMPLO 1

ε_1: Lanzar una moneda y observar la cara superior (C = cara , S = sello)

ε_2: Lanzar un dado y observar el número que aparece en la cara superior.

El lector habrá notado algunos aspectos comunes de los experimentos - descritos. Estos son :

(a) Cada experimento puede repetirse indefinidamente sin cambiar esencialmente las condiciones.

(b) Cada experimento es no determinístico.

(c) Cada experimento tiene varios resultados posibles que pueden describirse de antemano con precisión, por, ejemplo en ε_1 tal conjunto es $\{C,S\}$, en ε_2 es $\{1,2,3,4,5,6\}$.

EXPERIMENTOS ALEATORIOS Son experimentos que tienen las tres propiedades - (a), (b) y (c) antes mencionados.

EJEMPLO 2 Los experimentos ε_1 y ε_2 del ejemplo 1 y los siguientes, son experimentos aleatorios:

ε_3 : Extraer un artículo de un lote que contiene artículos defectuosos "D"- y no defectuosos "N".

ε_4 : Designar un delegado de un grupo de 50 personas.

ε_5 : Contar el número de automóviles que cruzan la intersección de dos calles, hasta, antes que ocurra un accidente.

ε_6 : Fabricar artículos, hasta producir 5 defectuosos y contar el número total de artículos fabricados.

ε_7 : Contar el número de vehículos que llegan a una estación de servicio en un día.

ε_8 : Elegir un punto del intervalo cerrado $[0,1]$

ε_9 : Observar el tiempo de vida de un artefacto eléctrico.

ε_{10} : De una urna que contiene bolas blancas y negras se escoge una y se anota su color.

ε_{11}: Verificar el estado de un transistor (0 = apagado, 1 = prendido)

1.2.2 ESPACIO MUESTRAL

Hemos dicho que cada experimento aleatorio tiene varios resultados posibles y que podemos describir con precisión el conjunto de estos resultados posibles. ***Llamaremos espacio muestral asociado a un experimento aleatorio ε, al conjunto de todos los resultados posibles de dicho experimento aleatorio*** y lo denotaremos por Ω. Así, por ejemplo, los espacios muestrales asociados a los respectivos experimentos del ejemplo 2, son:

Experimento :	Conjunto de resultados posibles = Espacio muestral
ε_1 :	$\Omega_1 = \{C, S\}$, C = Cara y S = Sello .
ε_2 :	$\Omega_2 = \{1,2,3,4,5,6\}$.
ε_3 :	$\Omega_3 = \{D, N\}$.
ε_4 :	$\Omega_4 = \{A_1, A_2, \ldots , A_{50}\}$; A_i representa una persona: Pedro, Juan , Rosa , Alberto etc.
ε_5 :	$\Omega_5 = \{0,1,2,3,4,\ldots \}$
ε_6 :	$\Omega_6 = \{5,6,7,8,9,10,\ldots \}$.
ε_7 :	$\Omega_7 = \{0,1,2,\ldots\}$.
ε_8 :	$\Omega_8 = \{x \in \mathbb{R} / 0 \leqslant x \leqslant 1\}$.
ε_9 :	$\Omega_9 = \{t \in \mathbb{R} / t \geqslant 0\}$.
ε_{10}:	$\Omega_{10} = \{b , n\}$, b = blanca, n = negra.
ε_{11}:	$\Omega_{11} = \{0,1\}$.

El experimento simple del lanzamiento de un dado, una moneda o la extracción de una bola de una urna, son modelos que nos permitirán una mejor exposición, pero es claro que cada uno de estos modelos tiene diferentes versiones; por ejemplo en el caso de la urna; podemos concebir la elección de una persona entre un grupo de hombres y mujeres que están en una sala; la urna es reemplazada por la sala; o puede referirse a un almacén en el cual se encuentran artículos procedentes de una línea de producción, la cual arroja el 5% de defectuosos D (bolas blancas) y el 95% de no defectuosas N (bolas negras); en este caso evidentemente el papel de la urna es desempeñado por el almacén. Igualmente el lanzamiento de una moneda nos servirá de modelo para luego aplicarlo, por ejemplo, al lanzamiento de un cohete, el funcionamiento de la válvula de un motor, etc.

Cuando el experimento aleatorio es *simple* como en el caso de los experimentos dados en el ejemplo 2, mayormente no hay dificultad en determinar el espacio muestral, pero cuando el experimento es compuesto, en algunos casos es un poco complicado.

Un experimento se dice que es compuesto, si consiste de dos o más experimentos simples sucesivos o simultáneos.

Consideremos dos tipos básicos de experimentos compuestos: aquellos, en que los experimentos simples están unidos por la partícula gramatical "o" en el sentido excluyente y aquellos donde los experimentos simples están unidos por la partícula gramatical "y".

1.2.3 EXPERIMENTOS UNIDOS POR LA O EXCLUYENTE

Un experimento compuesto ε, se dice que es una *o-combinación* de los experimentos simples, ε_1 y ε_2 si, sólo si el experimento ε ocurre, cuando el experimento ε_1 ó ε_2 ocurre (pero no ambos).

EJEMPLO 3 Considere el experimento, que consiste en lanzar un dado o una moneda. Hallar el espacio muestra para este experimento.

SOLUCION Observe que el experimento ε consiste de dos experimentos simples unidos por la "o" excluyente. Sean:

ε_1 : "Lanzar un dado"; luego, $\Omega_1 = \{1,2,3,4,5,6\}$

ε_2 : "Lanzar una moneda"; entonces, $\Omega_2 = \{C,S\}$

y $\varepsilon = \varepsilon_1$ ó ε_2. Por lo tanto, el espacio muestral Ω asociado a ε es la unión de Ω_1 y Ω_2. Es decir, $\Omega = \Omega_1 \cup \Omega_2 = \{1,2,3,4,5,6,C,S\}$

1.2.4 EXPERIMENTOS UNIDOS POR LA Y

Un experimento compuesto ε, se dice que es una *y-combinación* de los experimentos simples ε_1 y ε_2, sí y solamente si, el experimento ε ocurre, cuando ambos experimentos ε_1 y ε_2 ocurren.

CONSECUENCIA Si el experimento compuesto ε es una y-combinación de los experimentos ε_1 y ε_2, el espacio muestral Ω asociado a ε es el producto cartesiano de los espacios muestrales Ω_1 y Ω_2 correspondientes a ε_1 y ε_2 respectivamente. Es decir $\Omega = \Omega_1 \times \Omega_2$.

EJEMPLO 4 Se lanza una moneda tres veces. Hallar el espacio muestral asociado a este experimento.

SOLUCION Observe que el experimento ε consiste de tres experimentos simples unidos por la "y". En otras palabras el experimento ε ocurre, si los tres experimentos simples ocurren. Sea ε_i, i = 1,2,3, el experimento: "el i-ésimo lanzamiento de la moneda".

ε_1 : el primer lanzamiento; $\Omega_1 = \{C,S\}$.
ε_2 : el segundo lanzamiento; $\Omega_2 = \{C,S\}$.
ε_3 : el tercer lanzamiento ; $\Omega_3 = \{C,S\}$.

Así, ε consiste en efectuar primero ε_1, seguido de ε_2 y finalmente ε_3. Por lo tanto, el espacio muestral Ω asociado a ε es igual al producto cartesiano de los Ω_i, i = 1,2,3. Es decir,

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 = \{(x,y,z) \;/\; x,\ y,\ z = C,S\}$$
$$= \{CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS\}$$

Si el experimento compuesto consta de dos experimentos simples, una ***Tabla con dos entradas*** es de gran ayuda para construir el espacio muestral como se muestra en los ejemplos 5 , 6 y 7.

EJEMPLO 5 Se lanzan dos monedas simultáneamente y se observan las secuencias de caras y sellos.

Una tabla de dos entradas es la siguiente,

Primera moneda \ Segunda moneda	C	S
C	CC	CS
S	SC	SS

Entonces, el espacio muestral es, $\Omega = \{CC, CS, SS, SC\}$

Observe que el espacio muestral del experimento puede escribirse, como un conjunto de pares ordenados así,

$$\Omega = \{(x,y) \;/\; x = C \text{ ó } S, \quad y = C \text{ ó } S\}$$

Es decir, si $\Omega_1 = \{C,S\}$ y $\Omega_2 = \{C,S\}$. Entonces ,

$$\Omega = \Omega_1 \times \Omega_2 \quad \text{(Producto Cartesiano)}$$

EJEMPLO 6 Se lanzan una moneda y un dado simultáneamente y se observan las caras superiores .

La tabla de dos entradas es como sigue,

moneda \ dado	1	2	3	4	5	6
C	(C,1)	(C,2)	(C,3)	(C,4)	(C,5)	(C,6)
S	(S,1)	(S,2)	(S,3)	(S,4)	(S,5)	(S,6)

El espacio muestral de este experimento es,

$$\Omega = \{(C,1),(C,2),(C,3),(C,4),(C,5),(C,6),(S,1),(S,2),(S,3),(S,4),(S,5),(S,6)\}$$

EJEMPLO 7 Se lanzan dos dados simultáneamente y se observan las caras superiores.

La tabla siguiente presenta los resultados posibles del experimento.

Primer dado \ Segundo dado	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Entonces, el espacio muestral de este experimento es,

$$\Omega = \begin{Bmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6) \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6) \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6) \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6) \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6) \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{Bmatrix}$$

Observe, que Ω puede escribirse como un conjunto de pares ordenados, así

$$\Omega = \{(x,y) \ / \ x,y = 1,2,\ldots,6\}$$

esto es, si $\Omega_1 = \{1,2,3,4,5,6\}$ y $\Omega_2 = \{1,2,3,4,5,6\}$. Entonces,

$$\Omega = \Omega_1 \times \Omega_2 .$$

EJEMPLO 8 Se tiene una caja con 10 artículos diferentes. Se extraen 4 artículos, de uno en uno*, con reemplazamiento**. Describir el espacio muestral asociado a este experimento.

SOLUCION Sea $a_1, a_2, a_3, \ldots, a_{10}$ los diez artículos diferentes de la caja. Al hacer la primera extracción, puede salir cualquiera de los 10 artículos. Es decir,

$$\Omega_1 = \{a_1, a_2, \ldots, a_{10}\}.$$

Al hacer la segunda extracción, puede salir otra vez cualquiera de los 10 artículos, ya que el primero fue devuelto a la caja. Entonces,

$$\Omega_2 = \{a_1, a_2, \ldots, a_{10}\}.$$

Así, sucesivamente, es claro que en la tercera extracción obtenemos,

$$\Omega_3 = \{a_1, a_2, \ldots, a_{10}\}.$$

y en la cuarta

$$\Omega_4 = \{a_1, a_2, \ldots, a_{10}\}.$$

Es decir, $\Omega_1 = \Omega_2 = \Omega_3 = \Omega_4$. Entonces,

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 \times \Omega_4 = \Omega_1^4$$
$$= \{(u,v,x,y) \ / \ u,v,x,y = a_1, a_2, \ldots, a_{10}\}.$$

* Se dice también que hemos extraído una muestra de tamaño 4. (En general si se extrae n artículos se dice que hemos extraído una muestra de tamaño n).

** Se dice que la extracción se hace con reemplazamiento (o que la muestra se extrae con reemplazamiento), si después de cada extracción se registra el artículo y se devuelve a la caja.

EJEMPLO 9 Hay tres tiendas de víveres en una pequeña ciudad (numeradas 1,2, 3). Cuatro damas que viven en el poblado seleccionan al azar, y en forma independiente, una tienda para hacer sus compras sin salir de la ciudad. Dar un espacio muestral para el experimento que consiste en seleccionar las tiendas.

SOLUCION Desde que las tiendas estan numeradas con 1,2,3 y cada dama escoge una de estas tiendas al azar y en forma independiente, entonces, la primera dama escogerá al azar uno de los elementos del conjunto {1,2,3}. La segunda dama escogerá al azar uno de los elementos del mismo conjunto {1,2,3}. La tercera dama escogerá al azar uno de los elementos también del mismo conjunto {1,2,3} y finalmente la cuarta dama también escogerá al azar uno de los elementos del conjunto {1,2,3}. Por lo tanto, la elección de una tienda por las cuatro damas tendrá como espacio muestral

$$\Omega = \{1,2,3\} \times \{1,2,3\} \times \{1,2,3\} \times \{1,2,3\} = \{1,2,3\}^4$$

ó $\quad \Omega = \{(x_1, x_2, x_3, x_4) \,/\, x_i = 1,2,3,\ i = 1,2,3,4\}$.

EJEMPLO 10 Considere el experimento de verificar el estado (apagado, prendido) de cinco transistores iguales.

Utilizando los números 0 (cero) para "apagado" y 1 para "prendido", escribir los elementos del espacio muestral.

SOLUCION El resultado de verificar el primer transistor puede ser 0 ó 1;el resultado de verificar el segundo transistor también puede ser 0 ó 1; así sucesivamente. Entonces, el espacio muestral del experimento, verificar el estado de 5 transistores es

$$\Omega = \{0,1\}^5 = \{(x,y,z,w,v)/x,y,z,w,v = 0,1\}.$$

En muchos casos un diagrama, conocido con el nombre de ***diagrama del árbol***, es más sugerente para determinar el espacio muestral de un experimento aleatorio compuesto, por ejemplo construir el espacio muestral asociado al lanzamiento de una moneda cuatro veces y observar la secuencia de caras y sellos. El diagrama del árbol es como se muestra en la fig. 1.2.1.
Observe, que cada uno de los resultados posibles del experimento queda representado por una rama del árbol, así CCCC representa la primera rama,CCCS la segunda rama, así sucesivamente, SSSS la última rama. El espacio muestral es

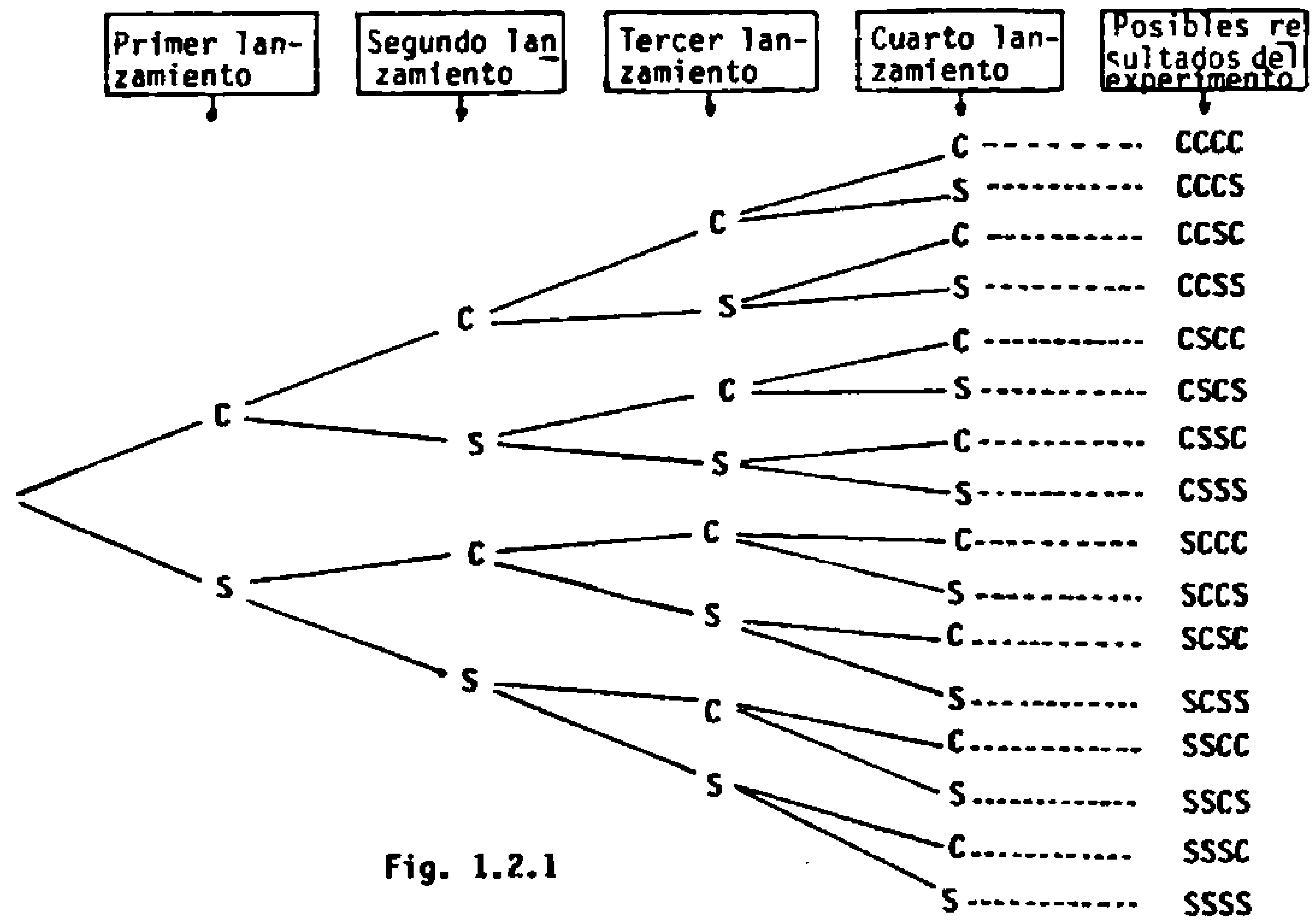


Fig. 1.2.1

$$\Omega = \begin{Bmatrix} CCCC , CCCS , CCSC , CCSS , CSCC , CSCS , CSSC , CSSS \\ SCCC , SCCS , SCSC , SCSS , SSCC , SSCS , SSSC , SSSS \end{Bmatrix}$$

o también

$$\Omega = \{(x_1, x_2, x_3, x_4) / x_1, x_2, x_3, x_4 = C \text{ ó } S\}$$

EJEMPLO 11 Los artículos provenientes de una línea de producción se clasifican en defectuosos "D" y no defectuosos "N", se observan los artículos y se anota su condición. Este proceso se contínua hasta observar dos defectuosos consecutivos o hasta que se observen tres artículos no defectuosos. Describir el espacio muestral asociado a este experimento.

SOLUCION El diagrama del árbol que representa los diversos resultados posibles del experimento está representado en la fig. 1.2.2.

El espacio muestral, entonces es

Ω = {DD , DNDD , DNDNDD , DNDNDN , DNDNN , DNNDD , DNNDN , DNNN , NDD , NDNDD , NDNDN , NDNN , NNDD , NNDN , NNN}

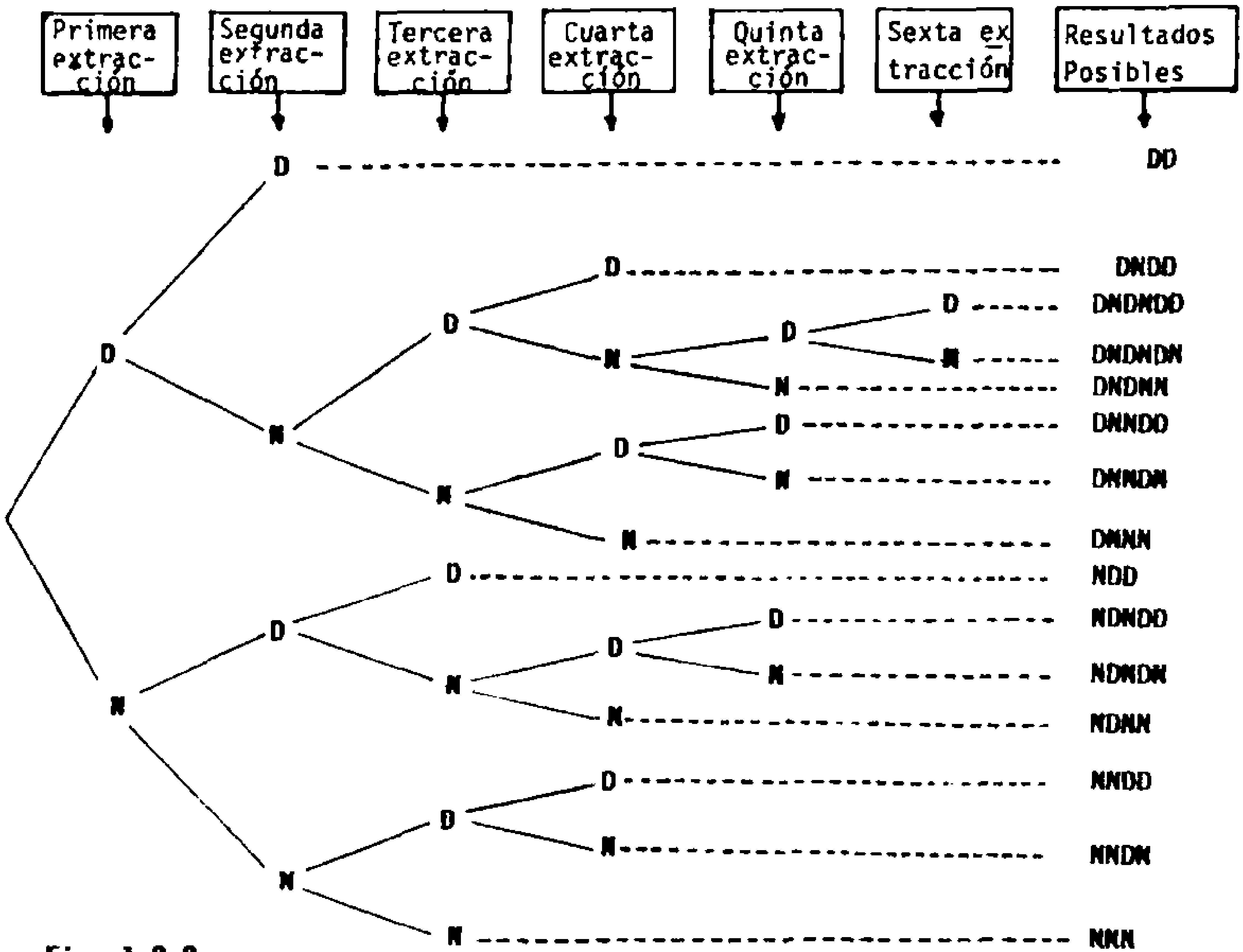


Fig. 1.2.2

EJEMPLO 12 Una línea de producción clasifica sus productos en defectuosos "D" y no defectuosos "N". De un almacén que guarda la producción diaria de esta línea se extraen artículos hasta observar dos defectuosos consecutivos o hasta que se hayan verificado cuatro artículos; construir el espacio muestral de este experimento.

SOLUCION El diagrama 1.2.3 muestra el diagrama del árbol que presenta los diversos resultados posibles del experimento.
El espacio muestral es

$$\Omega = \{DD , DNDD , DNDN , DNND , DNNN , NDD , NDND , NDNN , NNDD , NNDN , NNND , NNNN\}$$

Entre los espacios muestrales construídos para los experimentos dados en el ejemplos 2, podemos notar una diferencia entre $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ y $\Omega_5, \Omega_6, \Omega_7$ los primeros son finitos y los segundos no, más aún entre todos estos espacios muestrales y los conjuntos Ω_8 , Ω_9 que además de ser infinitos tienen - un número no numerable de elementos, es decir no podemos contar sus elementos. Estableceremos pues dos tipos de espacios muestrales: Discretos y Continuos.

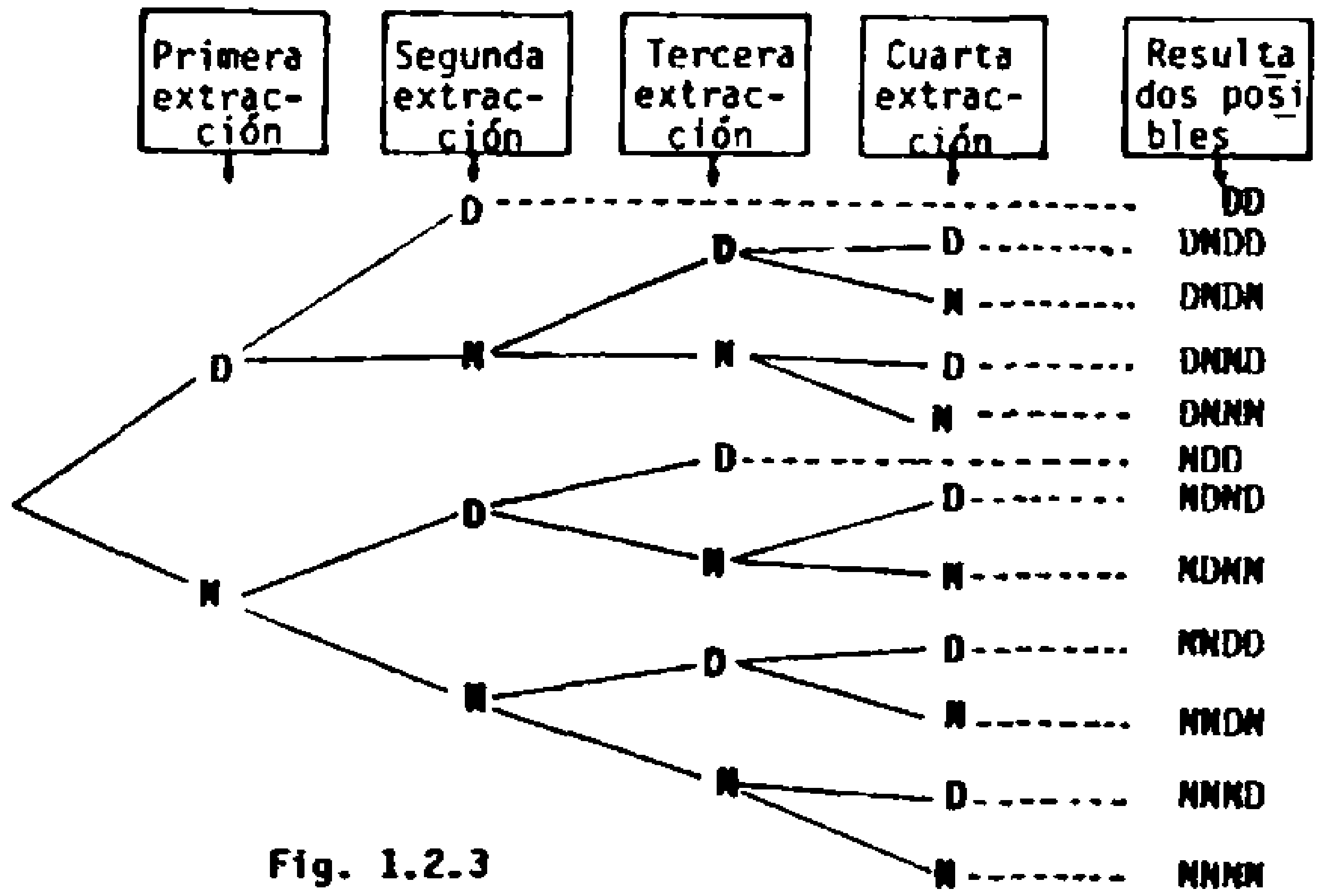


Fig. 1.2.3

1.2.5 ESPACIO MUESTRAL DISCRETO

Si tienen un número finito o infinito numerable de elementos.

(i) ***Espacios Muestrales Discretos Finitos*** Si el espacio muestral tiene un número finito de elementos por ejemplo Ω_1, Ω_2, Ω_3, Ω_4, Ω_{10}; los espacios muestrales dados por los ejemplos 3,4,5,6 y 7 también son espacios muestrales finitos. Otro ejemplo es el siguiente.

EJEMPLO 13 Un lote compuesto de n artículos provenientes de una línea de producción contiene m artículos defectuosos ($m \leq n$). Los artículos son extraídos uno por uno(sin reemplazamiento) hasta que el último artículo defectuoso sea extraido. Hallar el espacio muestral para este experimento.

SOLUCION El número de artículos extraídos será m o m+1 artículos, o m+2 etc. Entonces, el espacio muestral es

$$\Omega = \{m, m+1, m+2, \ldots, n\}.$$

(ii) ***Espacios Muestrales Discretos Infinitos*** Cuando puede establecerse una correspondencia uno a uno con el conjunto de los enteros positivos de modo que pueda ser enumerado como 1,2,3,... Por ejemplo Ω_5, Ω_6 y Ω_7. Otro ejemplo es el siguiente.

EJEMPLO 14 Lanzar una moneda hasta que ocurra cara.

El diagrama del árbol para este experimento da la figura 1.2.4.

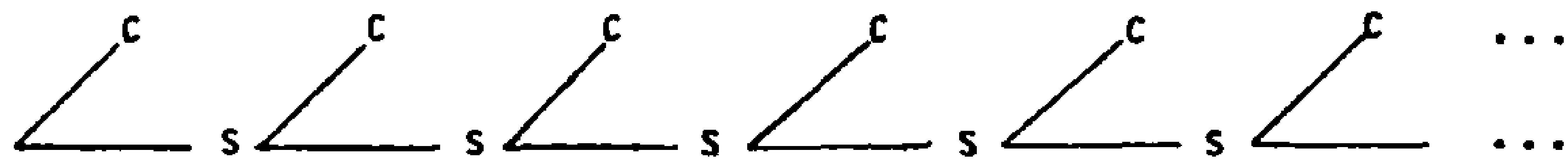


Fig. 1.2.4

Entonces, el espacio muestral es

$$\Omega = \{C, SC, SSC, SSSC, SSSSC, \ldots\}.$$

1.2.6 ESPACIO MUESTRAL CONTINUO

Un espacio muestral se dice que es continuo, si tiene un número no numerable de elementos. Es decir, cuyos elementos son todos los puntos de algún intervalo. Son espacios muestrales contínuos Ω_8 y Ω_9; también el ejemplo que sigue.

EJEMPLO 15 En un laboratorio químico, el volumen producido por día para un producto particular varía entre un valor mínimo, a, y un valor máximo, b, los cuales corresponden a la capacidad. Se escoge un día aleatoriamente y se observa la cantidad producida. Escribir el espacio muestral. El espacio muestral es entonces

$$\Omega = \{x \in \mathbb{R} \;/\; a \leq x \leq b\}.$$

Debe observarse también que a un experimento aleatorio podemos asociar más de un espacio muestral; es decir, de acuerdo a la característica del fenómeno que deseamos estudiar, un experimento aleatorio puede tener diferentes espacios muestrales. Por ejemplo, en el experimento lanzar tres monedas simultáneamente, si estamos interesados en la secuencia de caras y sellos que aparecen, el espacio muestral es

$$\Omega = \{CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS\}.$$

en cambio si estamos interesados en el número de caras que sale, el espacio muestral es

$$\Omega = \{0, 1, 2, 3\}.$$

1.2.7 EVENTOS

Hemos definido el espacio muestral como el conjunto de todos los resultados posibles de un experimento aleatorio. Es decir, podemos concebir al espacio muestral como un conjunto universal. Hablaremos, entonces en él, de subconjuntos y elementos. ***Se llama Evento a Cualquier Subconjunto del Espacio Muestral*** y lo denotaremos por A,B,C,D,E,F. etc. Así, si A es un evento entonces $A \subset \Omega$. Y llamaremos ***suceso a todo elemento de un espacio muestral*** y lo designaremos por ω, x, y, etc. Esto es, si x es un suceso, entonces $x \in \Omega$. Un evento con un sólo elemento es un evento elemental, así $A = \{\omega\}$ es un evento elemental.

Obsérvese que el evento $\{\omega\}$ y el suceso ω no son los mismos, esto lo sabemos de la teoría de conjuntos. En otras palabras, evento es cualquier elemento de $\mathcal{P}(\Omega)$, (así ϕ y Ω son eventos).

EJEMPLO 16 Consideremos los experimentos del ejemplo 2.

En ε_2 ; A : "ocurre un número par". Entonces

$$A = \{2,4,6\}.$$

En ε_3 ; A : "se extrae un artículo defectuoso". Luego

$$A = \{D\}.$$

En ε_5 ; A : "ocurre un accidente antes que crucen 100 automóviles". Entonces

$$A = \{0,1,2,3,\dots , 98\}.$$

En ε_6 ; A : " se fabrican más de 1000 artículos". Entonces

$$A = \{1001 , 1002,\dots\}.$$

En ε_8 ; A : "el punto se encuentra entre 0 y 1/2". Luego

$$A = \{x \in \mathbb{R} \ / \ 0 < x < \tfrac{1}{2}\}.$$

En ε_9 ; A : "el artefacto dura más de 2,000 horas". Entonces

$$A = \{t \in \mathbb{R} \ / \ t > 2{,}000\}.$$

EJEMPLO 17 En el experimento aleatorio lanzar cuatro monedas simultáneamente. Definimos los eventos siguientes:

A_1 : "Todas las monedas muestran el mismo lado".

A_2 : "Ocurren por lo menos dos caras".

A_3 : "Ocurre sello en el tercer lanzamiento". Entonces:

$A_1 = \{CCCC , SSSS\}$.

A_2 = {CCCC , CCCS , CCSC , CSCC , SCCC , CCSS , CSCS , CSSC , SCCS , SCSC , SSCC}.
A_3 = {CCSC , CSSC , CSSS , SCSC , SCSS , SSSS , CCSS , SSSC}.

EJEMPLO 18 Consideremos el experimento del ejemplo 7, y sea el evento A : "La suma total en los dados no excede a 6". Entonces

$$A = \{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,2),(4,1),(1,5),(2,4),(3,3),(4,2),(5,1)\}$$
$$= \{(x, y) \in \Omega \; / \; x + y \leqslant 6\}.$$

EJEMPLO 19 Consideremos el experimento aleatorio del ejemplo 15. Y sea A : "El volumen producido está entre c > a y b". Entonces,

$$A = \{x \in \mathbb{R} \; / \; a < c < x < b\}.$$

EJEMPLO 20 En el experimento aleatorio del ejemplo 14, sea A : "Se necesita más de cuatro lanzamientos". Entonces,

$$A = \{SSSSC , SSSSSC , SSSSSSC, \ldots\}.$$

Consideremos ahora los espacios muestrales del ejemplo 1. Asi tenemos que.
en Ω_1 : $\omega_1 = C$, $\omega_2 = S$, son sucesos;
En Ω_2 : $\omega_1 = 1$, $\omega_2 = 2$, $\omega_3 = 3$, $\omega_4 = 4$, $\omega_5 = 5$, $\omega_6 = 6$ son sucesos.
Finalmente diremos que un suceso ω es favorable a un evento A, si $\omega \in A$. Así SSSSC es un suceso favorable al evento A del ejemplo 20, igualmente los sucesos SSSSSSC , SSSSSSSC; son favorables al evento A, en cambio no lo son C, SC , SSC.

EJEMPLO 21 Sea el experimento, ε : "lanzar una moneda tres veces", y sea el suceso ω_3 = SCS. Hallar los posibles eventos en los cuales ω_3 sea un suceso favorable.

SOLUCION La Selección de ω_3, implica la ocurrencia de los siguientes eventos (y de muchos más). Los eventos que a continuación se dan son eventos diferentes y no son representaciones diferentes del mismo evento.
A_1 : "ocurre exactamente dos sellos"; A_2 : "ocurre a lo más una cara" y
A_3 : "en el segundo lanzamiento sale una cara".

EJEMPLO 22 En el espacio muestral del ejemplo 9, describir los siguientes eventos:
A : "Todas las damas escogen la tienda 1 ó todas escogen la tienda 2".
B : "Dos escogen la tienda 1 y las otras dos la tienda 2".

SOLUCION

$A = \{(1,1,1,1),(2,2,2,2)\}$.

$B = \{(x_1, x_2, x_3, x_4) \ / \ 2x_i \text{ son } 1,\ 2x_i \text{ son } 2\}$.

EJEMPLO 23 Considere el espacio muestral del ejemplo 10 y escriba los elementos de los siguientes eventos:

A : "Todos los transistores están apagados o todos están prendidos".

B : "Sólo el último transistor verificado está prendido".

SOLUCION

$A = \{(0,0,0,0,0),(1,1,1,1,1)\}$.

$B = \{(0,0,0,0,1)\}$.

PROBLEMAS 1.2

1. Dé un ejemplo de experimento aleatorio que es de interés para (a) un ingeniero electricista, (b) un economista, (c) el gerente de una compañía de automóviles, (d) un ingeniero de comunicaciones, (e) un médico, (f) un especialista en genética, (g) un biólogo, (h) un gerente de ventas.

2. Construir el espacio muestral apropiado para los siguientes experimentos aleatorios.

(a) Elegir una carta de una baraja de 52 cartas.

(b) Verificar el estado de dos transistores (apagado o prendido).

(c) Verificar el estado de 10 transistores iguales (apagado o prendido).

(d) Se lanzan n monedas y se observa el número de caras.

(e) Se pregunta a una persona por la fecha de su nacimiento (día del año).

(f) Inspeccionar las medidas de seguridad contra accidentes de una fábrica.

(g) Se pregunta a n personas por la fecha de su nacimiento (día del año).

(h) Un dardo se lanza en un blanco circular de radio r .

(i) Extraer una muestra de 5 bolas con reemplazamiento de una urna que contiene 12 bolas diferentes (esto es, las bolas se devuelve a la urna antes de extraer por segunda vez).

3. Un inversionista planea escoger dos de las cinco oportunidades de inversión que le han recomendado. Describa el espacio muestral que representa las opciones posibles.

4. Tres artículos son extraídos con reposición, de un lote de mercancías; cada artículo ha de ser identificado como defectuosos "D" y no defectuo-

so "N". Describa todos los puntos posibles del espacio muestral para este experimento.

5. Dos personas A y B se distribuyen al azar en tres oficinas numerada 1, 2 y 3. Si las dos personas pueden estar en la misma oficina, defina un espacio muestral adecuado.
6. Tres personas A, B y C se distribuyen al azar en dos oficinas numeradas con 1 y 2. Describa un espacio muestral adecuado a este experimento, (a) si los tres pueden estar en una misma oficina; (b) sí sólo se puede asignar una persona a cada oficina.
7. Durante el día, una máquina produce tres artículos cuya calidad individual, definida como defectuoso o no defectuoso, se determina al final del día. Describa el espacio muestral generado por la producción diaria.
8. El ala de un avión se ensambla con un número grande de remaches. Se inspecciona una sola unidad y el factor de importancia es el número de remaches defectuosos. Describa el espacio muestral.
9. Suponga que la demanda diaria de gasolina en una estación de servicio está acotada por 1,000 galones, que se lleva a un registro diario de venta. Describa el espacio muestral.
10. Se desea medir la resistencia al corte de dos puntos de soldadura. Suponiendo que el límite superior está dado por U, describa el espacio muestral.
11. De un grupo de transistores producidos bajo condiciones similares, se escoge una sola unidad, se coloca bajo prueba en un ambiente similar a su uso diseñado y luego se prueba hasta que falla. Describir el espacio - muestral
12. En el problema 11. (a) suponga que el experimento consiste en extraer dos transistores y se prueba hasta que fallan. Describir el espacio muestral (b) suponga que el experimento consiste en escoger 5 transistores y se - prueba hasta que fallan. Describir el espacio muestral.
13. Una urna contiene cuatro fichas numeradas: 2,4,6, y 8; una segunda urna contiene cinco fichas numeradas: 1,3,5,7, y 9. Sea un experimento aleatorio que consiste en extraer una ficha de la primera urna y luego una ficha de la segunda urna, describir el espacio muestral.
14. Una urna contiene tres fichas numeradas: 1,2,3; un experimento consiste en lanzar un dado y luego extraer una ficha de la urna. Describir el es-

pacio muestral.

15. Una línea de producción clasifica sus productos en defectuosos "D" y no defectuosos "N". De un almacén donde guardan la producción diaria de esta línea, se extraen artículos hasta observar tres defectuosos consecutivos o hasta que se hayan verificado cinco artículos. Construir el espacio muestral.

16. Lanzar un dado hasta que ocurra el número 4. Hallar el espacio muestral asociado a este experimento.

17. Una moneda se lanza tres veces. Describa los siguientes eventos:

    A : "ocurre por lo menos 2 caras".

    B : "ocurre sello en el tercer lanzamiento".

    C : "ocurre a lo más una cara".

18. En cierto sector de Lima, hay cuatro supermercados (numeradas 1,2,3,4). Seis damas que viven en ese sector seleccionan al azar y en forma independiente, un supermercado para hacer sus compras sin salir de su sector

(a) Dar un espacio muestral adecuado para este experimento.

(b) Describir los siguientes eventos:

    A : "Todas las damas escogen uno de los tres primeros supermercados"

    B : "Dos escogen el supermercado N° 2, dos el supermercado N° 3 y las otras dos el N° 4".

    C : "Dos escogen el supermercado N° 2 y las otras diferentes supermercados".

19. Tres máquinas idénticas que funcionan independientemente se mantienen funcionando hasta darle de baja y se anota el tiempo que duran. Suponer que ninguno dura más de 10 años.

(a) Definir un espacio muestral adecuado para este experimento.

(b) Describir los siguientes eventos:

    A : "Las tres máquinas duran más de 8 años".

    B : "El menor tiempo de duración de los tres es de 7 años".

    C : "Ninguna es dada de baja antes de los 9 años".

    D : "El mayor tiempo de duración de los tres es de 9 años".

20. En el espacio muestral del problema 4, describe los siguientes eventos:

    A : "Ocurre al menos 2 artículos no defectuosos".

    B : "Ocurre exactamente 2 artículos no defectuosos".

21. En el problema 16, describir el evento, "se necesitan por lo menos 5 lanzamientos".

22. El gerente general de una firma comercial, entrevista a 10 aspirantes a un puesto. Cada uno de los aspirantes es calificado como: Deficiente, Regular, Bueno, Excelente.
(a) Dar un espacio muestral adecuado para este experimento .
(b) Describir los siguientes eventos.
A : "Todos los aspirantes son calificados como deficientes o excelentes".
B : "Sólo la última persona extrevistada es calificado como excelente".
23. Considere el experimento de contar el número de carros que pasan por un punto de una autopista. Describa los siguientes eventos:
A : "Pasan un número par de carros".
B : "El número de carros que pasan es múltiplo de 6".
C : "Pasan por lo menos 20 carros".
D : "Pasan a lo más 15 carros".
24. En el problema 12. Describir los siguientes eventos. (1) en la parte (a).
A : "Los dos transistores duran a lo más 2,000 horas".
B : "El primero dura más de 2,000 horas, el otro menos de 3,000 horas".
(2) En la parte (b).
A : "Los cinco duran por lo menos 1,000 horas pero menos de 2,000 horas".
B : "El primero dura más de 2,000 horas, los demás a lo más 2,500 horas".

1.3 ALGEBRA DE EVENTOS

Hemos identificado el espacio muestral con el conjunto universal de la teoría de conjuntos, y los eventos como subconjunto del espacio muestral. - Identificaremos también el conjunto vacio (ϕ) de la teoría de conjuntos con el evento imposible, esto es, un evento que no ocurre. Por ejemplo, en el - experimento lanzar dos dados simultáneamente, el evento A : "obtener suma - 14", es un evento imposible. Al espacio muestral se llama también evento seguro.

En lo que sigue haremos una breve exposición a manera de revisión de la teoría de conjuntos en el lenguaje de eventos. Es decir, desde que los eventos son conjuntos, las operaciones de intersección "$\cap$", unión "U", inclusión - "$\subset$" serán definidos para eventos; las leyes y propiedades de la teoría de conjuntos son válidas.

1.3.1 OPERACIONES CON EVENTOS

SUB-EVENTOS Dado dos eventos, A y B se dice que A está ***contenido*** en B o que A es sub-evento de B y denotado por "$A \subset B$", si todo suceso favorable a A, - es favorable a B. En otras palabras, si ocurre el evento A, entonces ocurre el evento B. (Ver fig. 1.3.1a). En simbolos

$$A \subset B, \text{ si } \omega \in A \Longrightarrow \omega \in B.$$

EJEMPLO 1 Consideremos el experimento, lanzar una moneda hasta que ocurra - cara y contar el número de lanzamientos de la moneda. Definimos los siguientes eventos.

A : "Se necesita por lo menos 20 lanzamientos".

B : "Se necesita más de 5 lanzamientos".

En este experimento el espacio muestral es

$$\Omega = \{1, 2, 3, 4, 5, \ldots\}.$$

Entonces, $A = \{20, 21, 22, \ldots\}$

$B = \{6, 7, 8, 9, 10, \ldots\}$

es claro que $A \subset B$.

IGUALDAD DE EVENTOS Se dice que dos eventos A y B son ***iguales***, y se denota por "$A = B$", si $A \subset B$ y $B \subset A$.

EJEMPLO 2 En el experimento del ejemplo 1, consideremos los eventos,

A : "Se necesita a lo más 10 lanzamientos".

B : "Se necesita menos de 11 lanzamientos".

Entonces,

$$A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \quad \text{y}$$
$$B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

es claro que $A = B$.

UNION DE EVENTOS Dado dos eventos A y B, se llama ***unión de A con B*** y se designa por "$A \cup B$" al evento formado por los sucesos que pertenecen a A ó a B ó a ambos, es decir si ocurre el evento A ó B ó ambos. En símbolos (ver fig. 1.3.1a)

$$A \cup B = \{\omega \in \Omega \ / \ \omega \in A \ \vee \ \omega \in B\}$$

EJEMPLO 3 En el experimento del ejemplo 1, consideremos los siguientes eventos .

A : "Se necesita un número par de lanzamientos".

B : "Se necesita más de 10 lanzamientos".

Es decir,

$$A = \{2, 4, 6, 8, 10, \ldots\}$$
$$B = \{11, 12, 13, 14, \ldots\}$$

Entonces, es claro que $A \cup B = \{2, 4, 6, 8, 10, 11, 12, 13, 14, \ldots\}$.

En general, se dice que ocurre el evento $\bigcup_{i=1}^{n} A_i$, sí y sólo sí ocurre al menos uno de los eventos A_i .

INTERSECCION Dado los eventos A y B, se llama *intersección de* A *con* B y se denota "$A \cap B$", ó "AB" al evento formado por todos los sucesos favorables a A y a B. Es decir, ambos eventos A y B ocurren (la ocurrencia conjunta de A y B). En símbolos (ver fig. 1.3.1c)

$$AB = A \cap B = \{\omega \in \Omega \ / \ \omega \in A \wedge \omega \in B\}$$

EJEMPLO 4 En los eventos A y B del ejemplo 3. Se tiene que

$$A \cap B = \{12, 14, 16, 18, \ldots\}$$

En general se dice que ocurre el evento $\bigcap_{i=1}^{n} A_i$, si y sólo si ocurren todos los eventos A_i .

DIFERENCIA Dado los eventos A y B, se llama *diferencia de* A *con* B *y* se denota "A - B" al evento formado por los sucesos favorables a A que no son favorables a B. En símbolos (ver fig. 1.3.1d).

$$A - B = \{\omega \in \Omega \ / \ \omega \in A \wedge \omega \notin B\}$$

EJEMPLO 5 En los eventos A y B definidos en el ejemplo 3, se tiene que

$$A - B = \{2, 4, 6, 8, 10\} \quad \text{y}$$
$$B - A = \{11, 13, 15, 17, \ldots\}$$

COMPLEMENTO Si A es un evento del espacio muestral Ω, se llama *Complemento de* A, denotado por A' o $\bar{A}$ al evento formado por todos los sucesos que no pertenecen a A. Es decir, no ocurre el evento A. En símbolos. (Ver fig. 1.3.1e)

$$A' = \bar{A} = \Omega - A = \{\omega \in \Omega \ / \ \omega \notin A\}$$

EJEMPLO 6 Los complementos de los eventos definidos en el ejemplo 3, son - respectivamente:

$\bar{A} = \{1, 3, 5, 7, \ldots\} = \{\omega \in \Omega \,/\, \omega$ es impar$\}$

$\bar{B} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

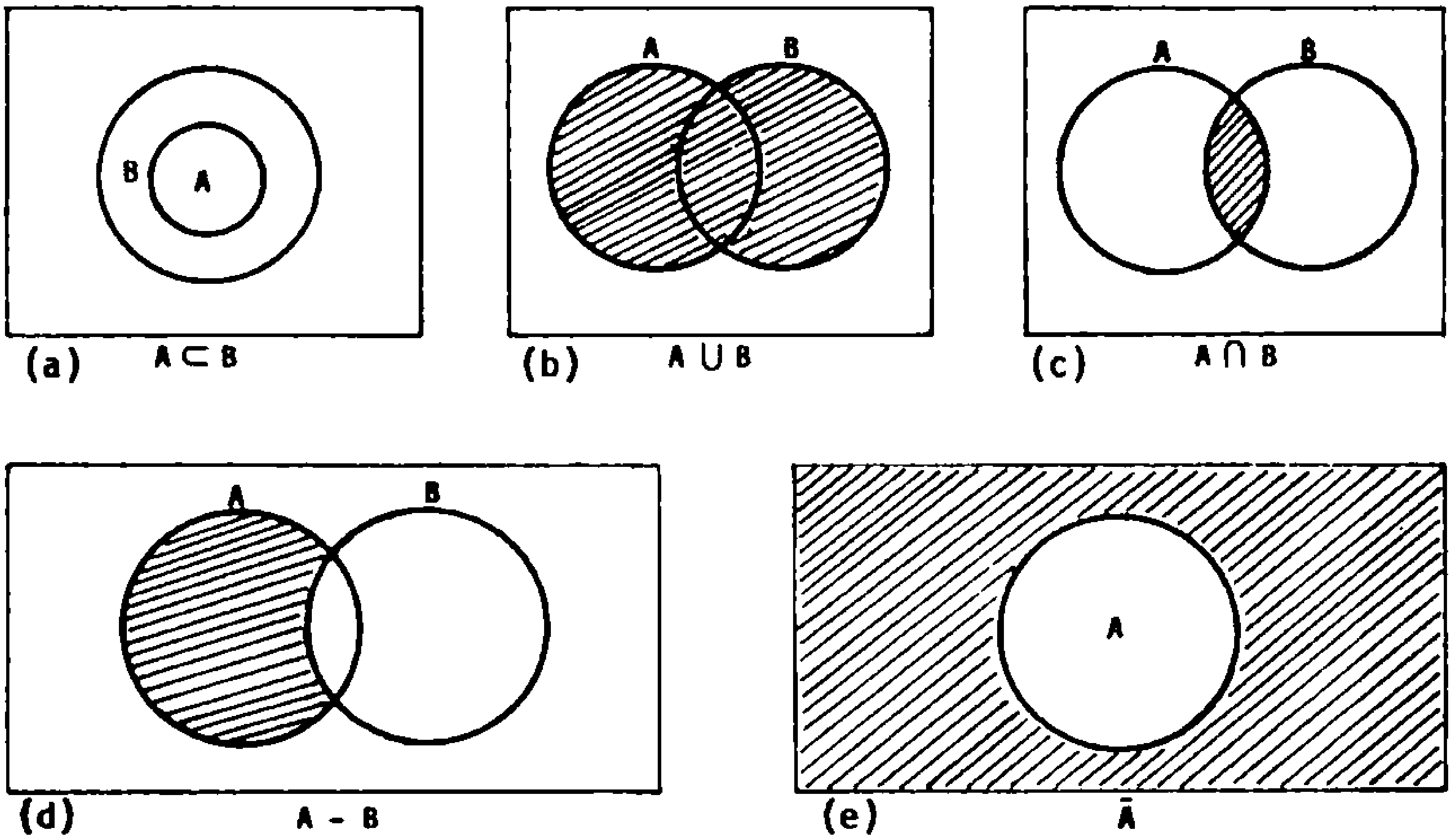


Fig. 1.3.1

LEYES DISTRIBUTIVAS Dado los eventos A , B y C, entonces

(D_1) : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(D_2) : $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

A manera de ilustración presentaremos la demostración de D_1 dejando D_2 para el lector.

D_1 : Demostraremos que

$$A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C) \quad (1)$$

y $$(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C) \quad (2)$$

Sea $\omega \in A \cap (B \cup C) \Longrightarrow \omega \in A$ y $(\omega \in B$ ó $\omega \in C)$

$\Longrightarrow (\omega \in A$ y $\omega \in B)$ ó $(\omega \in A$ y $\omega \in C)$

$\Longrightarrow (\omega \in A \cap B)$ ó $(\omega \in A \cap C)$

$\Longrightarrow \omega \in (A \cap B) \cup (A \cap C)$

Lo cual implica que, $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$ (3)

Sea ahora

$$\begin{aligned} \omega \in (A \cap B) \cup (A \cap C) &\Longrightarrow (\omega \in A \text{ y } \omega \in B) \text{ ó } (\omega \in A \text{ y } \omega \in C) \\ &\Longrightarrow \omega \in A \text{ y } (\omega \in B \text{ ó } \omega \in C) \\ &\Longrightarrow \omega \in A \text{ y } \omega \in (B \cup C) \\ &\Longrightarrow \omega \in A \cap (B \cup C) \end{aligned}$$

lo cual implica que

$$(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C) \qquad (4)$$

de (3) y (4) se tiene que

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

LEYES DE DE-MORGAN Sean los eventos A y B.

$$(DM_1) : \overline{A \cap B} = \bar{A} \cup \bar{B}$$

$$(DM_2) : \overline{A \cup B} = \bar{A} \cap \bar{B}$$

Demostraremos DM_1 a manera de ilustración y dejaremos DM_2 como ejercicio para el lector.

DM_1 : Ensayaremos una prueba más directa,

$$\omega \in \overline{A \cap B} \Longleftrightarrow \omega \notin A \cap B \Longleftrightarrow \omega \notin A \text{ ó } \omega \notin B \Longleftrightarrow \omega \in \bar{A} \text{ o}$$

$\omega \in \bar{B} \Longleftrightarrow \omega \in \bar{A} \cup \bar{B}$ y de este último concluímos que

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

PRODUCTO CARTESIANO Dado los eventos A y B, se llama ***producto Cartesiano*** de A con B, denotado "A x B", al conjunto de pares ordenados cuyos primeros elementos pertenecen a A y cuyos segundos elementos pertenecen a B, en símbolos.

$$A \times B = \{(\omega_1, \omega_2) \ / \ \omega_1 \in A \text{ y } \omega_2 \in B\}$$

Usaremos el producto cartesiano para construir espacio muestral asociado a un experimento compuesto. Es decir un experimento que consiste en realizar un experimento simple seguido de otro experimento simple o de experimentos simples simultáneos, como en el caso del lanzamiento de una moneda, dos, tres, cuatro veces o el lanzamiento de dos, tres, cuatro monedas simultáneamente; también en el lanzamiento de dos dados, o un dado y una moneda etc.como hemos visto en 1.2.2. Para terminar con esta revisión de la teoría de conjuntos generalizaremos las leyes distributivas y de De-Morgan. Sean A, $A_1, A_2, \ldots, A_n$ eventos, entonces:

$$(D_{1a})\ A \cap (A_1 \cup A_2 \cup \ \dots\ \cup A_n) = A \cap \left[\bigcup_{i=1}^{n} A_i\right] = \bigcup_{i=1}^{n} (A \cap A_i)$$

$$(D_{2a})\ A \cup (A_1 \cap A_2 \cap \ \dots\ \cap A_n) = A \cup \left[\bigcap_{i=1}^{n} A_i\right] = \bigcap_{i=1}^{n} (A \cup A_i)$$

$$(DM_{1a})\ \overline{A_1 \cap A_2 \cap \ \dots\ \cap A_n} = \overline{\bigcap_{i=1}^{n}} A_i = \bigcup_{i=1}^{n} \bar{A}_i = \bar{A}_1 \cup \bar{A}_2 \cup \ \dots\ \cup \bar{A}_n.$$

$$(DM_{2a})\ \overline{A_1 \cup A_2 \cup \ \dots\ \cup A_n} = \overline{\bigcup_{i=1}^{n}} A_i = \bigcap_{i=1}^{n} \bar{A}_i = \bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n.$$

NOTA Hemos visto que un evento puede definirse verbalmente, de manera que - es importante expresarlo en términos de operaciónes entre eventos. Considere mos por ejemplo los eventos A y B, entonces:

El evento de que ocurra por lo menos uno de ellos (esto es, uno o más - de ellos), se puede escribir como $A \cup B$.

El evento de que ninguno de los dos ocurra, se escribe como $\bar{A}\ \bar{B}$.
El evento de que ocurra exactamente uno de los dos eventos puede escribirse - así, $A\bar{B} \cup \bar{A}\ B$ donde "o" está en el sentido de exclusión en este caso.

1.3.2 EVENTOS MUTUAMENTE EXCLUYENTES Y COLECTIVAMENTE EXHAUSTIVOS

DEFINICION 1.3.1 Dos eventos A y B definidos en el mismo espacio muestral,- se dice que son *mutuamente excluyentes* si no pueden ocurrir juntos. Es decir la ocurrencia de uno excluye la ocurrencia del otro. En símbolos, si $A \cap B = \phi$.

EJEMPLO 7 Se lanza un dado dos veces. Sean los eventos:
A : "La suma de los puntos obtenidos en los dos lanzamientos es 7".
B : "En los dos dados se obtiene el mismo número".
A y B son eventos mutuamente excluyentes, pues el evento A es el conjunto

$\{(3,4), (4,3), (1,6), (6,1), (5,2), (2,5)\}$ y B es el conjunto

$\{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$, por lo tanto, $A \cap B = \phi$.

EJEMPLO 8 Sea el experimento, contar el número de personas atendidas por un banco en un período de tiempo. Sean los eventos:

A : "Se han atendido a menos de 20 personas".
B : "Se ha atendido a exactamente 25 personas".
C : "Se ha atendido exactamente 15 personas".

Los eventos A y B son mutuamente excluyentes, pues se tiene que

$$A = \{0, 1, 2, \ldots, 19\} \quad y \quad B = \{25\}$$

por lo tanto, $A \cap B = \phi$.
En cambio los eventos A y C no son mutuamente excluyentes, pues

$$A \cap C = C = \{15\} \neq \phi$$

La definición anterior se puede generalizar a tres o más eventos.

DEFINICION 1.3.2 Una colección de eventos $A_1, A_2, \ldots, A_n$ definidos sobre un mismo espacio muestral se dice que son mutuamente excluyentes, si la ocurrencia de uno de ellos excluye la ocurrencia de los otros. Es decir

$$A_i \cap A_j = \phi \ \forall\, i \neq j;\ i, j = 1, 2, \ldots, n.$$

EJEMPLO 9 Los elementos de un espacio muestral (los sucesos) son mutuamente excluyentes. En efecto, la ocurrencia de un resultado individual excluye la posibilidad de ocurrencia simultánea de los demás. Pues ocurre uno y sólo un resultado individual.

DEFINICION 1.3.3 Se dice que una colección de n eventos $A_1, A_2, \ldots, A_n$, definidos sobre el mismo espacio muestral son *Colectivamente exhaustivos* si la unión es igual al espacio muestral. Es decir,

$$A_1 \cup A_2 \cup \ldots \cup A_n = \bigcup_{i=1}^{n} A_i = \Omega$$

EJEMPLO 10 En el experimento, número de personas atendidas en un banco en un período determinado. Sean los eventos:
A : "menos de 11 personas han sido atendidas".
B : "de 11 a 20 personas han sido atendidas".
C : "más de 15 personas han sido atendidas".

Los eventos A, B y C son colectivamente exhaustivos, pues

$$A \cup B \cup C = \{0, 1, 2, \ldots\} = \Omega$$

EJEMPLO 11 Conteste justificando su respuesta cada una de las siguientes -

preguntas:

(a) Si A y B son eventos mutuamente excluyentes y colectivamente exhaustivos ¿son los eventos $\bar{A}$ y $\bar{B}$ mutuamente excluyente? ¿Son colectivamente exhaustivos?

(b) Si A y B son eventos mutuamente excluyentes pero no colectivamente exhaustivos, ¿son los eventos $\bar{A}$ y $\bar{B}$ colectivamente exhaustivos?

SOLUCION (a) Se tiene dos eventos A y B tales que $A \cap B = \phi$ y $A \cup B = \Omega$. Entonces, de la primera condición se obtiene

$\overline{A \cap B} = \bar{\phi} \Longrightarrow \bar{A} \cup \bar{B} = \Omega$, es decir $\bar{A}$ y $\bar{B}$ son colectivamente exhaustivos.

De la segunda condición se obtiene

$\overline{A \cup B} = \bar{\Omega} = \bar{A} \cap \bar{B} = \phi$, es decir $\bar{A}$ y $\bar{B}$ son mutuamente excluyentes.

(b) Se tiene dos eventos A y B tales que

$A \cap B = \phi$ y $A \cup B \neq \Omega$ ó $A \cup B \subset \Omega$. Entonces, existe un evento C - tal que $A \cup B \cup C = \Omega$ luego

$\bar{A} = B \cup C$ y $\bar{B} = A \cup C$. Por lo tanto

$\bar{A} \cup \bar{B} = (A \cup C) \cup (B \cup C) = A \cup B \cup C = \Omega$ ó

$\bar{A} \cup \bar{B} = \Omega$. Es decir, $\bar{A}$ y $\bar{B}$ son colectivamente exhaustivos.

PROBLEMAS 1.3

1. Sean A , B y C tres eventos cualesquiera en el espacio muestral Ω. Exprese cada uno de los siguientes eventos compuestos en términos de operaciones entre A , B y C.

(a) Ocurren exactamente dos de los eventos.

(b) Ocurren por lo menos uno de los eventos.

(c) Ocurren a lo más dos de los eventos.

(d) Ocurren todos los eventos.

(e) No ocurre ninguno de los eventos.

(f) No ocurre A, o no ocurre B o no ocurre C.

(g) Ocurre exactamente uno de los eventos.

(h) No ocurre más de uno de los eventos.

(*i*) Ocurre por lo menos uno de los tres eventos.

2. Determine el complemento para cada uno de los siguientes eventos:

(a) En el problema 1.2.9:"la demanda en un día determinado fué más de 8,000 galones".

(b) En el problema 1.2.11: "El transistor duró menos de 2,000 horas".

(c) En el problema 1.2.12b: "El primer transistor duró menos de 1,000 horas; el segundo, menos de 1,500 horas; y todos los demás más de 1,600 horas".

3. Verificar que :

(a) $A(B \cup C) = AB \cup AC$; (b) $A \cup BC = (A \cup B)(A \cup C)$

(c) $AB - ABC = AB\bar{C}$.

4. ¿Cuáles de las siguientes proposiciones son verdaderos para todos los eventos A,B,C y cuáles no?

(a) $(\bar{A} \cup \bar{B})C = \overline{AB} \cup C$; (b) $A \cap \overline{AC} = \phi$

(c) $(A - C)(B - C) = AB\bar{C}$; (d) $A(B - C) = AB - AC$

(e) $(\bar{A} \cup \bar{B}) \cap \overline{AB} = \phi$; (f) $(\bar{A} \cup \bar{B}) \subset \overline{AB}$

5. Sean A y B dos eventos de Ω, tales que no son mutuamente excluyentes:

(a) Exprese A(B) como la unión de dos eventos mutuamente excluyentes.

(b) Exprese $\bar{A}$ $(\bar{B})$ como la unión de dos eventos mutuamente excluyentes.

6. La tasa de desempleo para el siguiente período está pronosticado por un modelo económico. El pronóstico del modelo puede describirse como uno de los cinco eventos :

A_1 : "El desempleo será del 10% o más".

A_2 : "El desempleo será del 8% o más, pero menos del 10%".

A_3 : "El desempleo será del 6% o más, pero menos del 8%".

A_4 : "El desempleo será del 4% o más, pero menos del 6%".

A_5 : "El desempleo será menos del 4%".

Tome B_i para representar el desempleo actual de acuerdo a las mismas cinco clasificaciones (por ejemplo B_1 = El desempleo será del 10% o más).

(a) ¿Son mutuamente excluyentes los eventos $A_1,\dots,A_5$? ¿son colectivamente exhaustivos?

(b) ¿Qué indica los siguientes eventos, en palabras?

$$A_2 \cap B_3 ;\ A_3 \cup A_4 ;\ A_i \cap B_i ;\ A_i \cap B_j ;\ \text{donde } i > j$$

7. Un portafolio de acciones contiene cuatro acciones comunes. Durante un determinado día de negocios, tome:

A : "más de la mitad de las acciones subirán de precio".

B . "más de la mitad de las acciones bajarán de precio".

C : "más de la mitad de las acciones no cambiarán de precio".

(a) ¿Qué indican los siguientes eventos en palabras? $A \cup C$; $A \cap B$

(b) ¿Son los eventos A y B mutuamente excluyentes? ¿y A y C? ¿y B y C?

(c) ¿Son los eventos A,B y C colectivamente exhaustivos?

8. En la definición de los eventos A,B y C del problema 7, si se cambiara "más de la mitad" por la mitad o más, entonces:

(a) ¿Serían A y B mutuamente excluyentes? ¿A y C? ¿y B y C?

(b) ¿Son eventos A,B y C colectivamente exhaustivos?

9. Un blanco consiste de 10 círculos concéntricos de radio r_k, (k = 1,2,..., 10). El evento A_k significa acertar en el interior del círculo de radio r_k, (k = 1,2,..., 10). ¿Qué indica cada uno de los siguientes eventos?

(a) $B = \bigcup_{k=1}^{6} A_k$ (b) $C = \bigcap_{k=5}^{10} A_k$

10. Simplificar la expresión $A = (B \cup C) \cap (B \cup \bar{C}) \cap (\bar{B} \cup C)$

11. En que caso son verdaderas las siguientes igualdades:

(a) $A \cup B = \bar{A}$; (b) $AB = \bar{A}$; (c) $A \cup B = AB$

12. De la siguiente igualdad, hallar el evento X.

$$\overline{X \cup A} \cup \overline{X \cup \bar{A}} = B$$

13. ¿Son los eventos A, $\bar{A}$ B y $\overline{A \cup B}$ colectivamente exhaustivos.

14. Dos jugadores de ajedrez, juegan una partida. Sea A el evento, "el primer jugador gana" y B, "el segundo jugador gana". ¿Qué evento agregaría ud. a estos eventos para obtener una colección de eventos colectivamente exhaustivos?

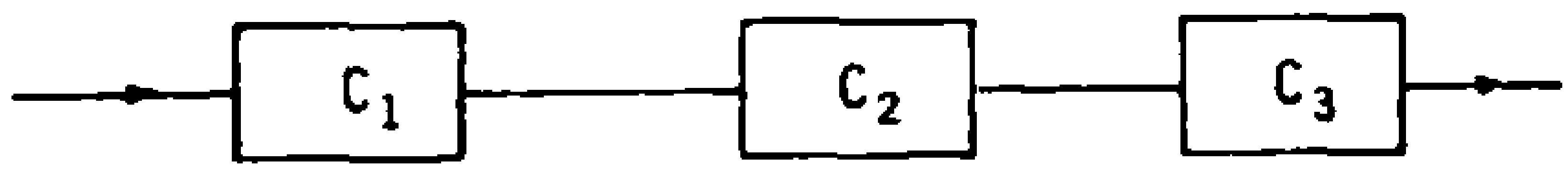
15. Una instalación consiste de dos calderos y un motor. Sea el evento A, "el motor está en buenas condiciones"; sea B_k (k = 1,2) el evento" el k-ésimo caldero está en buenas condiciones", y sea C, "la instalación puede funcionar, si el motor y al menos uno de los calderos estén en buenas condiciones". Exprese el evento C y C' en términos de A y B_k.

16. Considere el experimento de verificar la vida útil de dos tubos eléctricos iguales. Sea A_1, el evento "los dos tubos duran más de 10 años"; A_2 - el evento "el primer tubo observado dura más de 10 años y el segundo a lo más 15 años"; y A_3 el evento "ambos tubos duran a lo más 10 años"

(a) Escriba los elementos de Ω, A_1, A_2, A_3.

(b) ¿Cuáles de los sgtes. pares de eventos son mutuamente excluyentes A_1, A_2; A_1, A_3; A_2, A_3?

17. Considere el sistema en serie mostrado en la figura



Suponga que las componentes C_1, C_2 y C_3 pueden estar funcionando o - descompuesto. Verifique el estado de cada C_i, $i = 1,2,3$.

(a) Considerando 1 = "Funcionado" y 0 = "descompuesto". escriba el espacio muestral.

Defina los siguientes eventos :

E_1 : "todo el sistema está funcionando".

E_2 : "por lo menos dos delas componentes estan funcionando".

(b) ¿Son los eventos E_1 y E_2 mutuamente excluyentes?

(c) Interprete $\bar{E}_1$ y $\bar{E}_2$. (d) ¿Cuál es la relación entre $\bar{E}_1$ y $\bar{E}_2$?.

1.4 TECNICAS DE CONTEO

En muchas situaciones estaremos interesados sólo en el número de elementos que tiene un espacio muestral o un evento particular, en tales situaciones acudiremos a las técnicas de conteo, que serán de gran ayuda en estos casos.

1.4.1 PRINCIPIO DE MULTIPLICACION

El principio de multiplicación se enuncia como sigue:

TEOREMA 1.4.1 Si un experimento aleatorio (u operación) ε_1 ocurre de n_1 formas y si para cada una de estas, un experimento (u operación) ε_2 ocurre de n_2 formas, entonces los dos experimentos juntos ocurren de $n_1 n_2$ formas.

ACLARACION Para interpretar cabalmente el teorema anterior recuerde el concepto de experimento compuesto establecido en 1.2.2. Téngase tambien presente lo dicho en 1.2.4. Entonces, condición necesaria y suficiente para que se aplique el principio de multiplicación es que se realizan ambos experimentos (u operaciones) uno seguido del otro o simultaneamente.

EJEMPLO 1 ¿Cuántos elementos tiene el espacio muestral del experimento alea torio "lanzar una moneda y un dado simultaneamente"?

SOLUCION La moneda ocurre de 2 formas, es decir $n_1 = 2$, y para cada uno de estos, el dado ocurre de 6 formas, osea $n_2 = 6$. Por lo tanto, la moneda y el dado al ser lanzados simultaneamente ocurren de $n_1 \ n_2 = 2 \times 6 = 12$ formas diferentes. Es decir, el espacio muestral Ω tiene 12 elementos.
El espacio muestral asociado al experimento compuesto "lanzar una moneda y - un dado a la vez", se construye a partir de los sub-espacios muestrales Ω_1 , Ω_2 asociados a los experimentos "lanzar una moneda" y "lanzar un dado" respectivamente, como se ha visto en 1.2 . Así,

$$\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1 , \omega_2) \ / \ \omega_1 \in \Omega_1 \ \wedge \ \omega_2 \in \Omega_2\}$$

EJEMPLO 2 Una persona puede viajar de una ciudad A a otra ciudad B de 5 formas y de B a C de 6 formas. ¿De cuántas formas puede ir de A a C pasando por B? .

SOLUCION La persona puede ir de A a B de 5 formas, para cada una de estas,- hay 6 maneras de viajar de B a C. Por lo tanto, hay $5 \times 6 = 30$ formas de ir de A a C pasando por B. El diagrama de la fig. 1.4.1, muestra los caminos a seguir.

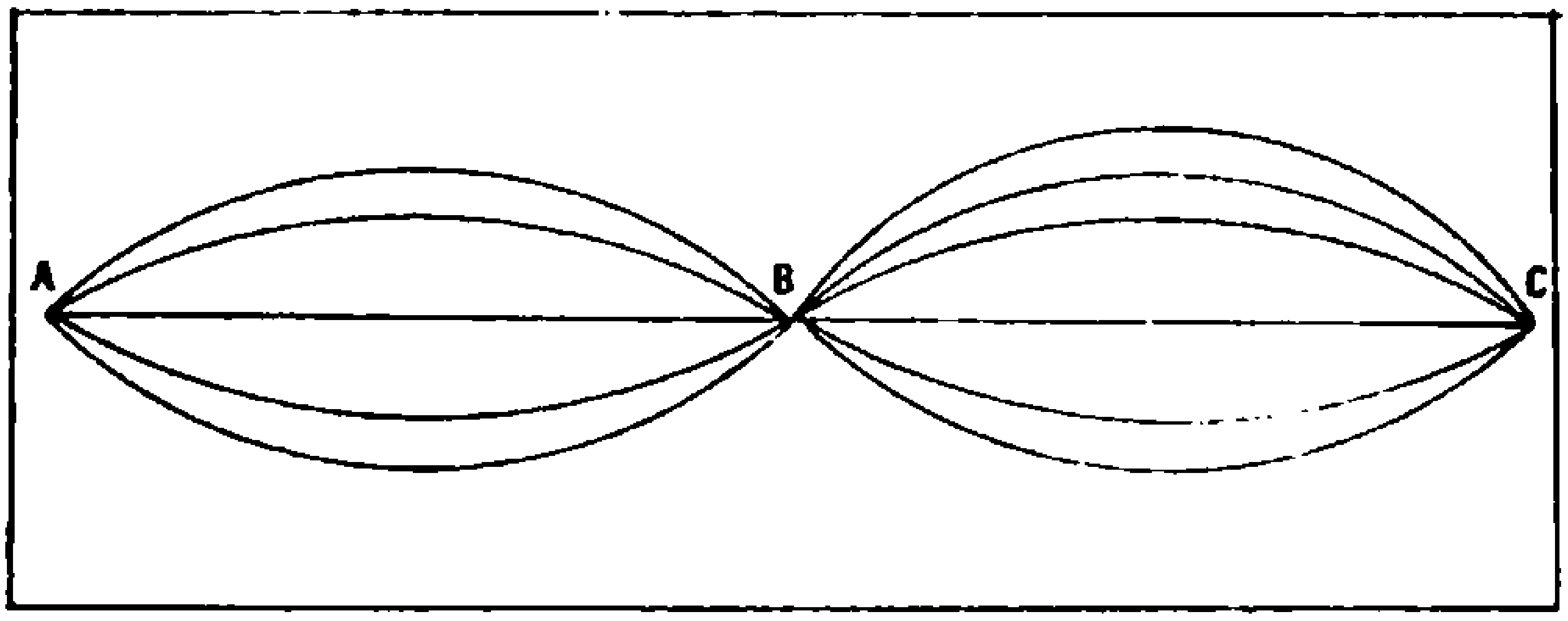


Fig. 1.4.1

La generalización del principio de multiplicación a cualquier número de experimentos u operaciones es como sigue.

TEOREMA 1.4 2 Si un experimento (u operación) ε_1 ocurre de n_1 formas, y para cada uno de éstos, un experimento (u operación) ε_2 ocurre n_2 formas, y para cada una de las dos primeras se puede efectuar un tercer experimento (u - operación) ε_3 de n_3 formas y así sucesivamente, la secuencia de k experimentos (u operaciones) se realizará de $n_1 \ n_2 \ \ldots \ n_k$ formas diferentes.

EJEMPLO 3 Un producto se arma en tres etapas, para la primera etapa se tienen disponibles 5 líneas de armado, para la segunda 4 y para la tercera 6 líneas de armado. ¿De cuántas maneras puede moverse el producto en el proceso de armado?

SOLUCION En la primera etapa, el producto puede moverse de 5 formas, para cada una de éstas en la segunda etapa, el producto puede moverse de 4 maneras y para cada una de las dos anteriores, en la tercera etapa puede moverse de 6 formas. Por lo tanto, el producto puede moverse de $5 \times 4 \times 6 = 120$ formas diferentes en el proceso de armado.

EJEMPLO 4 ¿Cuántos números pares de 3 dígitos se pueden formar con los dígitos 1,2,5,6,7,8,9, si cada dígito puede emplearse una sola vez?.

SOLUCION Puesto que los números por formarse deben ser pares, sólo hay 3 casos posibles para las unidades (cualquiera de los dígitos 2,6 ú 8). Para cada uno de éstos 3 casos hay 6 elecciones posibles para las decenas y 5 para las centenas. Luego, se puede formar un total de $3 \times 6 \times 5 = 90$ números pares diferentes.

CONSECUENCIAS Los siguientes resultados son consecuencias obvias del principio de multiplicación.

1. En el teorema 1.4.2, si $n = n_1 = n_2 = \dots = n_k$, entonces la secuencia de los k experimentos (u operaciones) ocurre de n^k formas.

EJEMPLO 5 Se lanza una moneda sucesivamente 6 veces, ¿de cuántas formas ocurre?

SOLUCION Aquí $n_1 = n_2 = \dots = n_6 = 2$. Por lo tanto, la secuencia de los 6 lanzamientos de la moneda ocurre de $2^6 = 64$ formas diferentes.

2. Sea $N(A)$ denota el número de elementos de un conjunto A. Supongamos que $N(A) = m$ y $N(B) = n$. Entonces, el número de formas de seleccionar un elemento de A y un elemento de B es igual a mn. En símbolos

$$N(A \times B) = N(A) \times N(B) = mn.$$

3. En general se cumple $N(A_1 \times A_2 \times \dots \times A_k) = N(A_1)N(A_2)\dots N(A_k)$.

1.4.2 PRINCIPIO DE ADICION

El principio de adición se enuncia de la siguiente manera:

TEOREMA 1.4.3 Si un experimento (u operación) ε_1 puede ocurrir de n_1 formas y un segundo experimento (u operación) ε_2 de n_2 formas, entonces el experimento (u operación) ε, que consiste en realizar ε_1 ó ε_2 ("o" en el sentido de exclusión, es decir ε_1 y ε_2 no pueden ocurrir juntos) ocurre de $n_1 + n_2$ formas, siempre que los espacios muestrales Ω_1 y Ω_2 asociados a ε_1 y ε_2 respectivamente sean disjuntos .

NOTA Para interpretar cabalmente el principio de adición téngase presente lo dicho en 1.2.3.

EJEMPLO 6 Consideremos el experimento de lanzar una moneda o un dado. ¿ De cuántas formas ocurre?

SOLUCION El experimento dado ε es compuesto; sean

ε_1 : lanzar una moneda ; $n_1 = 2$

ε_2 : lanzar un dado ; $n_2 = 6$

Luego,el experimento ε : "lanzar una moneda o un dado", ocurre de $n = n_1 + n_2 = 2 + 6 = 8$ formas. Aquí evidentemente la partícula gramatical "o" está en el sentido de exclusión y $\Omega_1 \cap \Omega_2 = \phi$.

EJEMPLO 7 Una persona puede viajar de A a B por vía aérea o por vía terrestre y tiene a su disposición 5 líneas aéreas, 6 líneas terrestres. ¿De cuántas formas puede hacer el viaje?

SOLUCION Aquí evidentemente, si la persona decide viajar por vía terrestre ya no viaja por vía aérea y viceversa. Luego, la partícula gramatical "o" está en el sentido de exclusión. Entonces, por el principio de adición se tiene que hay, $n_1 + n_2 = 5 + 6 = 11$ formas diferentes. La figura 1.4.2 da un diagrama de los caminos posibles.

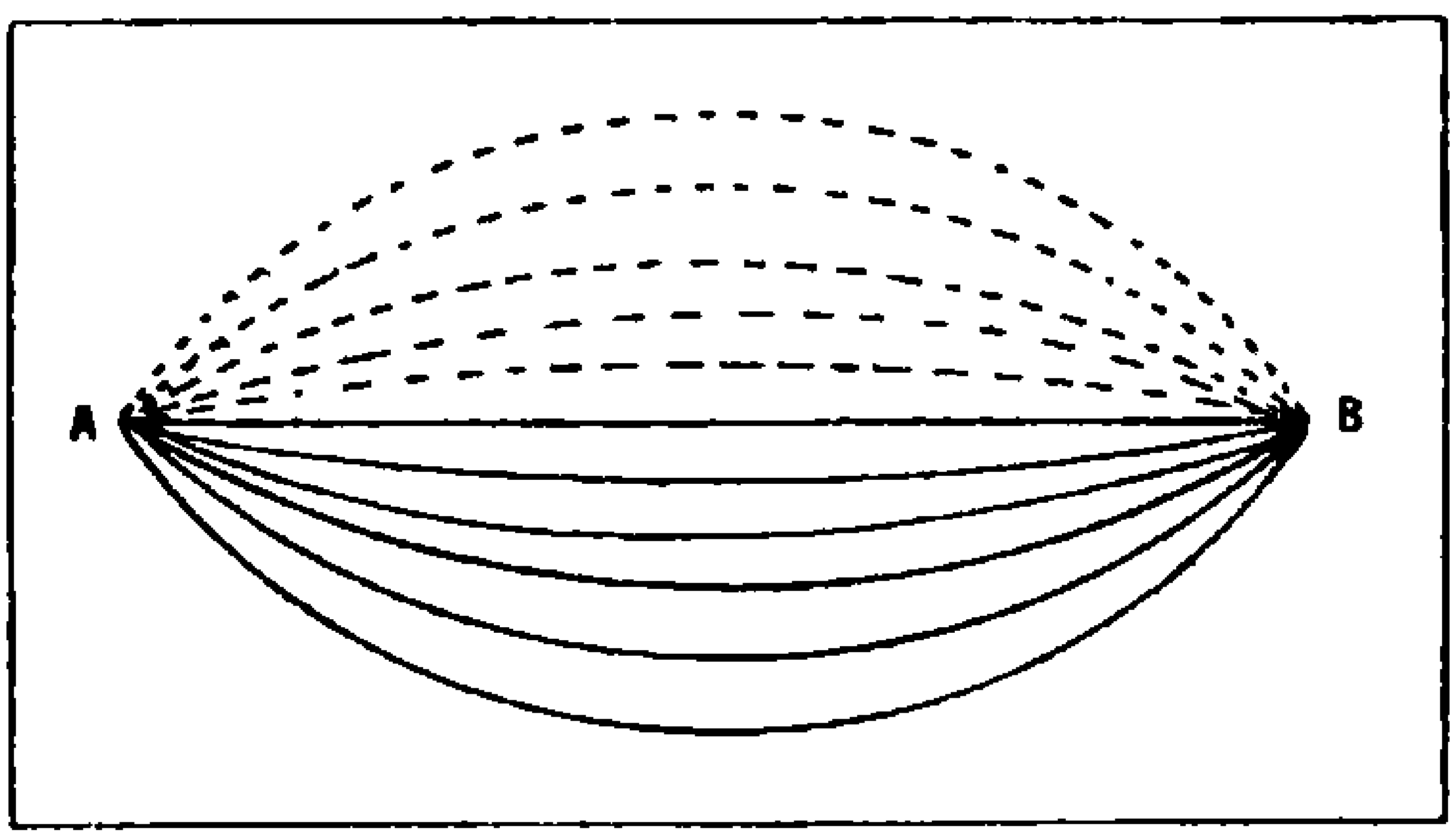


Fig. 1.4.2

La generalización del principio de adición es la siguiente:

TEOREMA 1.4.4 Si un experimento ε_1 ocurre de n_1 formas, ε_2 de n_2, ..., ε_k de n_k formas; entonces el experimento ε que consiste en realizar ε_1 ó ε_2, ó ... ó ε_k (la partícula ó, en el sentido excluyente. Es decir, los experimentos ε_1, ε_2 ; ...,ε_k no pueden realizarse juntos) ocurre de $n = n_1 + n_2 + ... + n_k$ formas, siempre que los correspondientes espacios muestrales sean disjuntos dos a dos; es decir,

$$\Omega_i \cap \Omega_j = \phi \quad \text{para} \quad i \neq j = 1,2,\ldots, k .$$

EJEMPLO 8 Un producto se vende en 3 mercados; en el primer mercado se tienen disponibles 5 tiendas, en el segundo 4 y en el tercer mercado, 6 tiendas. ¿De cuántas maneras puede venderse el producto?

SOLUCION Por el principio de adición tenemos que $n = n_1 + n_2 + n_3 = 5 + 4 + 6 = 15$ maneras diferentes.

Con este ejemplo y el ejemplo 3 trataremos de hacer notar la diferencia entre las situaciones en que usamos cada principio; es claro que si ud. desea comprar uno de los productos y decide comprar en una de las tiendas del mercado N° 1, por ejemplo, ya no comprará ni en el mercado N° 2 ni en el mercado N° 3 ; mientras que en el ejemplo 3, una vez que el producto, siguiendo una de las líneas de armado de la primera etapa, tiene que seguir inmediatamente una de las líneas para la segunda etapa y finalmente una para la tercera etapa.

Algo similar ocurre en los dos ejemplos paralelos dados con respecto a los viajes de una persona. En el primer caso, es claro que al llegar a B por uno de los 5 caminos sigue su viaje hasta llegar a C, por uno de los 6 caminos disponibles; en cambio en el segundo caso, si decide viajar por vía aérea ya no viaja por vía terrestre y viceversa.

CONSECUENCIA El resultado siguiente, es una consecuencia inmediata del principio de adición. Si A y B son conjuntos disjuntos, entonces

$$N(A \cup B) = N(A) + N(B).$$

En general, si A_1 , A_2, ...,A_k son conjuntos disjuntos dos a dos, entonces

$$N\left[\bigcup_{i=1}^{k} A_i\right] = \sum_{i=1}^{k} N(A_i).$$

1.4.3 PERMUTACION

FACTORIAL DE UN NUMERO Sea n un número entero positivo, el *factorial* de n,- se denota por "n!" o "$\underline{n|}$" y se define como el producto de todos los enteros consecutivos de 1 hasta n inclusive, es decir

$$n! = 1 \times 2 \times 3 \times \ldots \times (n-1)\, n = n(n-1) \times \ldots \times 2 \times 1.$$

Así, $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

$10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3628800$

observe que, $n! = n(n-1)!$

De este último, si n = 1, tenemos $1! = 1 \times 0!$. Luego definimos convencionalmente $0! = 1$.

Supongamos, ahora que estamos interesados sólo en el número de elementos que tiene un espacio muestral cuyos elementos son todas las ordenaciones (o arreglos) posibles de un conjunto de objetos. Por ejemplo, podemos estar interesados en saber, ¿de cuántas formas posibles pueden sentarse 8 personas alrededor de una mesa?, o podemos preguntarnos, ¿de cuántas formas pueden adjudicarse los lugares de partida a los 10 participantes de una carrera automovilística?. Los diferentes arreglos se llaman *permutaciones*.

DEFINICION 1.4.1 Una *permutación* es un arreglo de todos o parte de un conjunto de objetos.

Suponga que tenemos un conjunto de tres objetos, A = {a,b,c} y estamos interesados en el número de arreglos (las posibles permutaciones) con los elementos del conjunto A. Las posibles permutaciones son:

abc, acb, bac, bca, cab y cba

vemos que hay 6 permutaciones distintas.

Se puede llegar a la misma respuesta sin tener que escribir todas las ordenaciones posibles, de la siguiente manera: los arreglos de los 3 objetos es equivalente a disponerlos en 3 celdas, asi

3	2	1

Hay 3 formas posibles de llenar la primera celda, con cualquiera de los tres objetos a, b y c; para la segunda celda hay 2 formas posibles, cualquiera de los dos objetos restantes después de haber llenado la primera y solamente queda una forma de llenar la tercera. Aplicando el principio de multiplicación da un total de 3 x 2 x 1 = 6 formas (o permutaciones).

En general, supongamos que tenemos un conjunto de n objetos, y estamos interesados en el número de permutaciones que se puede hacer con los n objetos, generalizando el proceso anterior tendremos, n casillas.

n	n - 1	n-2	n-3	.	.	.	.	3	2	1

La primera casilla se puede llenar de n maneras; la segunda de n - 1 maneras la tercera de n - 2, etc. la última de solo una manera. Aplicando el principio de multiplicación se tiene

$n(n - 1)(n - 2)(n - 3) \ldots 2 \times 1 = n!$ permutaciones, que se denota por

$_nP_n$ o $P_{(n,n)}$ o P_n^n . Es decir $\quad _nP_n = P_{(n,n)} = P_n^n = n!$

la cual se lee así: "número de permutaciones de n objetos tomados de n a n" o simplemente número de permutaciones de n objetos diferentes. Para precisar mejor damos el siguiente teorema.

TEOREMA 1.4.5 *El número de permutaciones* de n objetos diferentes es

$$P_n^n = {}_nP_n = n!$$

EJEMPLO 9 Un inspector visita 6 máquinas diferentes durante el día. A fin de impedir a los operadores que sepan cuando inspeccionará, varía el orden de las visitas. ¿De cuántas maneras puede hacerlo?

SOLUCION Puesto que el inspector tiene que inspeccionar las 6 máquinas diferentes, el número de maneras, es el número de permutaciones de las 6 máquinas. Es decir

$$_6P_6 = 6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720 \text{ formas.}$$

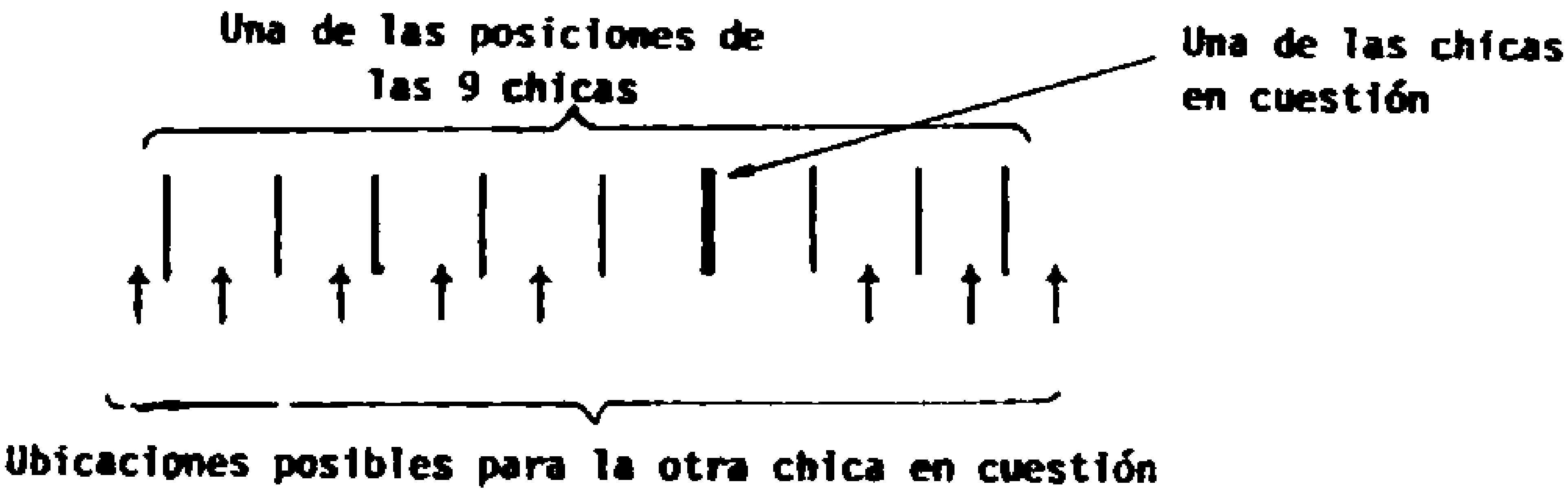
EJEMPLO 10 En una competencia automovilística intervienen 40 participantes. ¿De cuántas formas distintas se pueden adjudicar los lugares de largada a los 40 competidores de la competencia?

SOLUCION Se desea saber de cuántas formas posibles se pueden ordenar los 40 competidores. El número de tales ordenaciones posibles es

$$_{40}P_{40} = 40!$$

EJEMPLO 11 ¿De cuántas maneras se pueden colocar 10 chicas en una fila, de manera que dos chicas, en particular, no queden juntas?

SOLUCION Separemos del grupo una de las chicas en cuestión, de manera que consideremos sólo 9. Estas se podrán ordenar de 9! formas. Para cada ordenación de 9, la chica separada puede ubicarse en 8 lugares sin estar junto a la otra chica en cuestión. Por lo tanto, de acuerdo con el principio de multiplicación, el número total de formas diferentes es $8 \times 9!$.



EJEMPLO 12 ¿De cuantas maneras se pueden colocar 12 niños en una fila, de manera que cuatro niños, en particular, queden juntos?

SOLUCION Separemos los cuatro niños en cuestión, de manera que consideremos solamente 8. Estos se podrán ordenar de 8! formas. Para cada ordenación de 8 los 4 niños en cuestión pueden ubicarse en 9 lugares. Para cada posición, de los 4 niños, estos pueden ordenarse de 4! formas. Por lo tanto, de acuerdo con el principio de multiplicación, el número total de formas diferentes es $4! \times 9 \times 8! = 4! \times 9!$.

Supongamos ahora que tenemos un conjunto con n objetos diferentes y deseamos permutar r de estos n objetos $0 < r < n$. Tendremos entonces r celdas. La primera celda se puede llenar de n formas, la segunda de $n - 1$ formas, etc. y la r-ésima celda de $n - (r - 1) = n - r + 1$ formas.

n	n - 1	n - 2	n - 3	.	.	.	n - r + 1

Aplicando el principio de multiplicación, el número de permutaciones de los n objetos diferentes tomados r a r es

$$n(n - 1)(n - 2)(n - 3) \ldots (n - r + 1)$$

lo cual se puede escribir

$$n(n-1)(n-2)\ .\ .\ .\ (n-r+1)\ \frac{(n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}$$

Que se denota por $P(n,r)$, nPr o P_n^r. Usaremos P_n^r o nPr, indistintamente

TEOREMA 1.4.6 *El número de permutaciones de* n objetos diferentes tomados - r a r es

$$P_n^r = {}_nP_r = \frac{n!}{(n-r)!}$$

NOTA Hemos visto que una permutación es un arreglo de todos o parte de los - elementos de un conjunto que tiene objetos diferentes. Así, si A = {a, b, c} se vió que las diferentes permutaciones son:

abc, acb, bac, bca, cab, cba

Es decir, el orden de los elementos es importante.

Observe, que estos elementos son comparables con las ternas ordenadas - que se pueden formar con los elementos de dicho conjunto sin repetición de - sus elementos; (o sea no están las ternas (a,a,a),(b,b,b,) y (c,c,c)).

(a,b,c) , (a,c,b) , (b,a,c) , (b,c,a) , (c,a,b) , (c,b,a)

En general, si el conjunto tiene n elementos. Una permutación de los n elementos, *es una n-upla ordenada*; y el número de permutaciones de los n elementos es el número de n-uplas ordenadas que se forman con los n elementos - sin repetición.

Las permutaciones de los n-elementos tomados r a r, son *r-uplas ordenadas*, y el número de permutaciones de los n elementos tomados r a r, es el número de r-uplas ordenadas que se pueden formar con los n elementos sin repetición.

EJEMPLO 13 Un grupo está formado por 5 personas y desean formar una comisión integrada por presidente y secretario. ¿De cuántas maneras puede nombrarse esta comisión?

SOLUCION El cargo de presidente puede ser ocupada de 5 maneras diferentes ; y una vez ocupado el cargo de presidente, el cargo de secretario puede ser - ocupado de 4 maneras diferentes; entonces la elección de la comisión se puede hacer de 5 x 4 = 20 maneras diferentes o simplemente, número de permutaciones de 5 personas tomadas 2 a 2.

$${}_5P_2 = \frac{5!}{(5-2)!} = 5 \times 4 = 20 \text{ maneras.}$$

NOTA El lector puede dar nombres a las personas, digamos A,B,C,D,E. Entonces, se busca todos los pares ordenados que se pueden formar con dichas letras, sin repetición .

(A , B) , (A , C) , (A , D) , (A , E)
(B , A) , (B , C) , (B , D) , (B , E)
(C , A) , (C , B) , (C , D) , (C , E)
(D , A) , (D , B) , (D , C) , (D , E)
(E , A) , (E , B) , (E , C) , (E , D)

Donde cada letra de la primera componente indica la persona que ocupa el cargo de presidente, y la segunda, indica la persona que ocupa el cargo de secretario. Así (C,B) indica que C resultó elegido presidente y B secretario. Y es sin repetición, ya que el par (A , A) no está en la permutación, pues si estuviera significaría que la persona A ocupa el cargo de presidente y secretario; lo cuál no puede ser.

EJEMPLO 14 Encontrar el número total de enteros positivos que pueden formarse utilizando los dígitos {1,2,3,4}, si ningún dígito ha de repetirse cuando se forma un número.

SOLUCION El número total de enteros positivos es

$${}_4P_1 + {}_4P_2 + {}_4P_3 + {}_4P_4 = 4 + \frac{4!}{(4-2)!} + \frac{4!}{(4-3)!} + 4!$$

$$= 4 + 12 + 24 + 24 = 64 \text{ números diferentes.}$$

EJEMPLO 15 Se va a colorear un mapa de cuatro países, con colores diferentes para cada país. Si hay disponibles 6 colores diferentes. ¿De cuántas maneras puede colorear el mapa?

SOLUCION Se necesita permutaciones de cuatro de un conjunto de 6 elementos. Es decir,

$${}_6P_4 = \frac{6!}{(6-4)!} = 6 \times 5 \times 4 \times 3 = 360 \text{ maneras.}$$

EJEMPLO 16 En un omnibus que posee 37 asientos (en ocho filas de cuatro asientos cada una con un pasillo en el medio, y al final 5 asientos juntos), se desea ubicar 25 pasajeros.

(a) ¿De cuántas formas se pueden ubicar?

(b) ¿De cuántas formas se pueden ubicar si deciden no ocupar los últimos 5 - asientos?

(c) ¿De cuántas formas se pueden ubicar si viajan cinco amigos que deciden , viajar juntos en los últimos asientos?

(d) ¿De cuántas formas se pueden ubicar si ocupan los 18 asientos que poseen ventanilla?

(e) ¿De cuántas formas se pueden ubicar si 10 de los pasajeros están enfermos y deben viajar en asientos que poseen ventanilla?

SOLUCION

(a) En este caso se debe calcular el número de grupos de 25 asientos que se pueden formar de entre los 37 asientos considerando el orden, ya que en los asientos elegidos los pasajeros se pueden distribuir de diferentes formas y cada grupo de 25 asientos nos indicará cuales son los asientos elegidos por los pasajeros. El número buscado es

$$_{37}P_{25} = \frac{37!}{(37-25)!} .$$

(b) En este caso se considera sólo 32 asientos. El razonamiento es el mismo que en el caso anterior. El número buscado es

$$_{32}P_{25} = \frac{32!}{(32-25)!} .$$

(c) Como los cinco amigos viajan juntos en el último asiento, entonces los - restantes 20 pasajeros se ubicarán en los 32 asientos que quedan disponibles Y razonando como antes, este número es

$$_{32}P_{20} = \frac{32!}{(32-20)!} .$$

Ahora bien para cada una de estas ubicaciones disponibles de los 20 pasajeros los 5 amigos se ubicarán de $_5P_5 = 5!$ formas distintas. Luego, el número buscado es

$$_{32}P_{20} \cdot {}_5P_5 = \frac{32!}{(32-20)!} \cdot 5! .$$

(d) Primero veamos de cuántas formas pueden ocuparse los 18 asientos con ven tanilla. Se trata de hallar el número de grupos de 18 pasajeros que se pueden formar con los 25 y considerando el orden. El número es

$$_{25}P_{18} = \frac{25!}{(25 - 18)!} = \frac{25!}{7!}$$

Para cada una de estas formas de ocupar los asientos con ventanilla, ¿de cuántas formas se pueden ubicar los 7 pasajeros restantes en los 19 asientos aún libres?. Es claro que es de $_{19}P_7$ formas. Y el número buscado es

$$_{25}P_{18} \cdot {_{19}P_7} = \frac{25!}{7!} \cdot \frac{19!}{12!} .$$

(e) Veamos antes de cuántas formas posibles se pueden ubicar los 10 pasajeros enfermos en los 18 asientos con ventanilla. Se trata de hallar grupos de 10 asientos que se pueden formar de entre los 18 asientos, dichas formas posibles son

$$_{18}P_{10} = \frac{18!}{(18 - 10)!}$$

Para cada una de estas ubicaciones de los pasajeros enfermos; ¿De cuántas formas se pueden ubicar los 15 pasajeros restantes en los 27 asientos aún libres?. Es evidente que es $_{27}P_{15}$ formas. Por lo tanto, el número buscado es

$$_{18}P_{10} \cdot {_{27}P_{15}} = \frac{18!}{8!} \cdot \frac{27!}{12!} ,$$

Las permutaciones que ocurren por arreglos de objetos formando (o alrededor de un círculo) un círculo se llaman *permutaciones circulares*. En estas agrupaciones no hay primero ni último elemento, por hallarse todos en una línea cerrada. Para determinar el número de permutaciones circulares que pueden formarse con los n objetos distintos de un conjunto, basta observar que considerando fija la posición de uno cualquiera de ellos, los n - 1 restantes podrán cambiar de lugar, de (n - 1)! formas diferentes tomando todas las posiciones sobre la circunferencia relativa al primer punto. Si cambiamos ahora la posición de este, los de los demás respecto de él será seguro uno de los ya considerados. Por lo tanto, el número de permutaciones circulares será

$$(n - 1)!$$

La permutación circular se denota por P_c^n

TEOREMA 1.4.7 El número de permutaciones de **n** objetos distintos alrededor de un círculo es

$$P_c^n = (n - 1)!$$

EJEMPLO 17 ¿De cuántas formas diferentes pudieron sentarse, en la última cena, alrededor de la mesa, jesucristo y los 12 apóstoles?

SOLUCION (a) Si la mesa fuera circular, tendremos la permutación circular. El número de formas es

$$P_c^{13} = (13 - 1)! = 12! .$$

(b) Si la mesa no es circular, se tendrá una permutación de las 13 personas. El número de formas es

$$_{13}P_{13} = 13! .$$

1.4.4 PERMUTACION CON REPETICION

Hasta ahora hemos permutado objetos diferentes (esto, es,se pueden distiguir). Sin embargo, este no siempre es el caso. Supongamos por ejemplo, que estamos interesados en el número de permutaciones distinguibles uno de otro, que se pueden formar con las letras de la palabra "AMAR".

Si usáramos la palabra "AMOR" en vez de "AMAR", la permutación estudiada es aplicable directamente y daría el número de permutaciones distinguibles $_4P_4 = 4! = 24$. Sin embargo esoeramos que la respuesta al presente problema sea menor que 24, pués tenemos letras repetidas. Así, la permutación "MAOR" y "MOAR" corresponden a las permutaciones indistinguibles "MAAR" y "MAAR" en nuestro problema. Luego, a cada permutación de las letras de "AMAR" le corresponde $_2P_2$ permutaciones de "AMOR" que aparecen como permutaciones de las letras {A,O}, así

OMAR	AMOR	MAOR	MOAR	AOMR	OAMR	RAOM	ROAM
MROA	MRAO	MORA	MARO	ORMA	ARMO	AROM	ORAM
RMOA	RMAO	ROMA	RAMO	OMRA	AMRO	AORM	OARM

Reemplazando A por O vemos que cada uno de estos pares se vuelve indistinguible en el caso de la palabra "AMAR". Por lo tanto, hay

$\frac{_4P_4}{_2P_2} = \frac{4!}{2!} = 12$ permutaciones distinguibles que se pueden hacer con la palabra "AMAR".

Un ejemplo simple es el siguiente. Consideremos un conjunto con cuatro elementos diferentes $\{a,b,c,d\}$. Luego hay $4! = 24$ permutaciones distintas. Supongamos ahora que

$$a = b = x \quad y \quad c = d = y.$$

Entonces, se puede listar sólo las siguientes permutaciones distinguibles

XXYY , XYXY , YXXY , YXYX , XYYX , YYXX.

es decir, tenemos $\frac{_4P_4}{_2P_2\ _2P_2} = \frac{4!}{2!\ 2!} = 6$ permutaciones distinguibles.

Consideremos ahora el problema de contar el número de permutaciones de las 13 letras de la palabra "DIVISIBILIDAD". En este caso la letra I aparece 5 veces, la letra D aparece 3 veces, y cada una de las otras 5 aparecen exactamente una vez. $_{13}P_{13}$ es el número de permutaciones de los elementos de un conjunto que tiene 13 elementos tales como $\{D,A,V,E,S,I,B,O,L,U,T,N,R\}$. Denotemos por P el número (desconocido) de permutaciones diferentes de las letra de DIVISIBILIDAD. Correspondiendo a cada uno de estas permutaciones, hay $_5P_5 \cdot {}_3P_3$ permutaciones de las letras DAVESIBOLUTNR. Entonces

$$P \cdot {}_5P_5 \cdot {}_3P_3 = {}_{13}P_{13} \qquad ó \qquad P = \frac{_{13}P_{13}}{_5P_5 \cdot {}_3P_3} = \frac{13!}{5!3!}.$$

TEOREMA 1.4.8 El número de permutaciones distintas de n objetos de los cuales n_1 son de una clase, n_2 de una segunda clase, ..., n_k de una k-ésimo clase y todo los demás objetos de clase 1, se denota por $P_n^{n_1,n_2,\ldots,n_k}$ y está dado como

$$P_n^{n_1,n_2,\ldots,n_k} = \frac{n!}{n_1!n_2!\ldots n_k!}$$

EJEMPLO 18 Un estante de una librería tiene capacidad para 10 libros de Matemáticas que tiene pasta verde, 8 de Física de pasta roja y 7 de Química de pasta azúl. ¿De cuántas maneras pueden colocarse los libros según los colores?

SOLUCION Como interesa sólo los colores, entonces sea $n_1 = 10$, $n_2 = 8$, $n_3 = 7$. Luego, el número de permutaciones distinguibles es

$$P_{25}^{10,8,7} = \frac{25!}{10!\,8!\,7!} = 21,034,600 \ .$$

EJEMPLO 19 Suponga que un día oscuro nacen en cierto hospital cuatro pares de mellisos, idénticos, dos pares de mellizas, idénticas, nueve niños y once niñas. Se utiliza una tinta no indeleble para escribir sus nombres. El día siguiente (aún oscuro) la tinta desaparece. ¿De cuántas maneras es posible mezclar los niños?

SOLUCION Aquí $n_1 = 2$, $n_2 = 2$, $n_3 = 2$, $n_4 = 2$ para niños y $n_5 = 2$, $n_6 = 2$ para niñas. Y el total de niños es $n = 32$, luego

$$P_{32}^{2,2,2,2,2,2} = \frac{32!}{2!\ 2!\ 2!\ 2!\ 2!\ 2!} = \frac{32!}{(2!)^6} \ .$$

1.4.5 PARTICION DE UN CONJUNTO

A menudo se necesita contar el número de formas de particionar un conjunto de n objetos diferentes en r subconjunto llamados *celdas*. La partición se hace teniendo en cuenta los siguientes criterios:

1 Los r subconjuntos son disjuntos dos a dos.
2 La unión de todos los subconjuntos es igual al conjunto original.
3 El orden de los elementos dentro de cada celda no tiene importancia.

Empesaremos considerando un ejemplo simple. Sea el conjunto {a,b,c,d,e}. Las posibles particiones en dos celdas, en las que la primera contenga cuatro elementos y la segunda sólo un elemento, son:

{{a,b,c,d} , {e}} , {{a,b,c,e} , {d}} , {{a,b,d,e} , {c}}
{{a,c,d,e} , {b}} , {{b,c,d,e} , {a}}

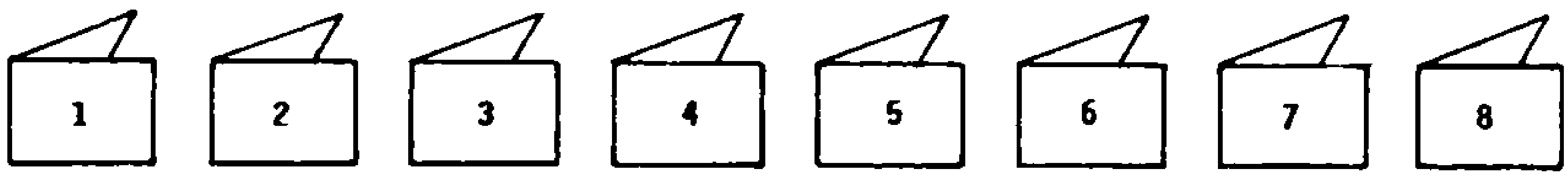
vemos que hay 5 formas distintas de hacer una partición de un conjunto de cinco elementos en dos celdas (o subconjuntos) que contenga cuatro elementos en la primera y uno en la segunda. El número de particiones para este ejemplo se puede escribir así,

$$\frac{5!}{4!\ 1!} = 5$$

Consideremos ahora el ejemplo siguiente

EJEMPLO 20 Suponga que un hombre tiene 8 bonos financieros diferentes de ocho compañías distintas, y que piensa regalarlos a sus hijos de la siguiente manera: a su hijo mayor, 3; a su segundo hijo, 3; y al menor, 2. ¿En cuántas formas puede repartir los bonos?

SOLUCION Una de las formas sería tomar los bonos, alineándolos y dar los tres primeros a su hijo mayor, los siguientes tres al segundo y los dos últimos al menor.



Como existen ${}_8P_8 = 8!$ formas de alinearlos en la mesa, también son $8!$, las formas en que pueden ser distribuídos. Sin embargo, no todas estas formas son diferentes. Hay $3!$ maneras de arreglar el orden de los bonos para su primer hijo, $3!$ para los de su segundo hijo, y $2!$ para los del tercero. Por lo tanto, los bonos pueden repartirse de

$$\frac{8!}{3!\ 3!\ 2!} = 560 \text{ formas}.$$

Este argumento puede generalizarse a un conjunto con n elementos y a r subconjuntos, tal que el primero tenga n_1 elementos, el segundo, n_2, y el r-ésimo subconjunto n_r elementos ($n_1 + n_2 + \ldots + n_r = n$). El número de divisiones posibles está dado por

$$\frac{n!}{n_1!n_2!\ \ldots\ n_r!}$$

que se denota así $\binom{n}{n_1,\ n_2,\ \ldots,\ n_r}$. Donde el número superior representa el número total de elementos del conjunto y los números inferiores representan al número de elementos asignados a cada celda.

TEOREMA 1.4.9 El número de formas posibles, que **n** objetos diferentes pueden dividirse en r grupos distinguibles conteniendo $n_1, n_2, \ldots, n_r$ objetos respectivamente donde $n_1 + n_2 + \ldots + n_r = n$, es

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1!\, n_2!\ \ldots n_r!}$$

1.4.6 COMBINACION

En muchos casos estaremos interesado en el número de formas de seleccionar r objetos de n. sin importar el orden. Estas selecciones se llaman *Combinaciones*.

DEFINICION 1.4.2 Un subconjunto de r elementos de un conjunto que tiene n - elementos diferentes. se llama una *combinación de los n elementos tomados* - r a r.

Determinaremos ahora. el número de combinaciones de r elementos que se pueden formar con los n objetos diferentes de un conjunto. Este número se denota por

$$C_n^r \quad \text{o} \quad C(n,r)$$

Comenzaremos nuestra discusión considerando un conjunto de cuatro elementos diferentes {a.b.c.d}. Tenemos los siguientes subconjuntos no vacíos -

{a} , {b} , {c} , {d}
{a,b} , {a,c} , {a,d} , {b,c} , {b,d} , {c,d}
{a,b,c} , {a,b,d} , {a,c,d} , {b,c,d}
{a,b,c,d}

Existen cuatro subconjuntos de un elemento. Es decir el número de combinaciones de un elemento es 4, o $C_4^1 = 4$, uno por cada elemento. También hay seis subconjuntos de 2 elementos; por lo tanto el número de combinaciones de 4 elementos; tomados 2 a 2 es 6 o sea $C_4^2 = 6$. Así mismo hay cuatro subconjuntos de 3 elementos; es decir el número de combinaciones de 4 elementos tomados 3 a 3 es 4, o sea
$C_4^3 = 4$; y hay sólo un subconjunto de cuatro elementos, por lo tanto el número de combinaciones de 4 elementos tomados 4 a 4 es uno o sea $C_4^4 = 1$.

En general con un conjunto que tiene n elementos se formará:

(a) n combinaciones de un elemento (combinaciones de los n elementos tomados uno a uno); es decir $C_n^1 = n$

(b) n combinaciones de n - 1 elementos (combinaciones de los n elementos tomados de n - 1 a n - 1); es decir

$$C_n^{n-1} = n$$

(c) Una combinación de n elementos (combinaciónes de los n elementos tomados n a n); es decir

$$C_n^n = 1$$

Luego, el problema será, calcular el número de combinaciones de los n elementos tomados r a r, o sea C_n^r , con

$$r \neq 1,\ n-1,\ n.$$

En el ejemplo discutido hemos obtenido $C_4^2 = 6$. Además se puede observar, que hay una conección entre el problema planteado y el problema de permutaciones considerado en la sección anterior. Para esto escribimos todas las combinaciones de 4 elementos tomados 2 a 2 en una columna como muestra la siguiente tabla y a su derecha sus respectivas permutaciones.

$$C_4^2 \left\{\begin{array}{c} \{a,b\} \\ \{a,c\} \\ \{a,d\} \\ \{b,c\} \\ \{b,d\} \\ \{c,d\} \end{array}\right. \quad \Bigg| \quad \overbrace{\begin{array}{ccc} (a,b) & , & (b,a) \\ (a,c) & , & (c,a) \\ (a,d) & , & (d,a) \\ (b,c) & , & (c,b) \\ (b,d) & , & (d,b) \\ (c,d) & , & (d,c) \end{array}}^{P_2^2} \longleftarrow P_4^2$$

Note que, cada par ordenado de la derecha de la recta es una permutación de 2 a 2 del conjunto situado a la izquierda de la recta, en la misma fila. Para cada subconjunto de dos elementos hay, pues P_2^2 permutaciones.
Entonces, P_2^2 es el número de columnas a la derecha de la recta y C_4^2 es el número de filas. Además todos los elementos que están a la derecha de la recta es el número total de permutaciones de 2 a 2 que se puede formar con los elementos del conjunto $\{a,b,c,d\}$. Es decir, es P_4^2 . Por lo tanto, se tiene que

$$C_4^2 \cdot P_2^2 = P_4^2$$

$$\text{ó} \quad C_4^2 = \frac{P_4^2}{P_2^2} = \frac{4!}{2!\,2!} = 6.$$

En general, consideremos ahora un conjunto con n elementos diferentes, C_n^r denotará el número (desconocido) de combinaciones de los n elementos tomados r a r. Imaginemos una tabla en la cual cada una de las combinaciones de r elementos determina una fila. En cada una de las filas escribiremos a su derecha las P_r^r , permutaciones de los r elementos del conjunto que identifica la fila.

$$
\begin{array}{cc|l}
 & & \overbrace{(\,.\;.\;.\;.\,)\quad(\,.\;.\;.\;.\,)\;.\;.\;.\;(\,.\;.\;.\;.\,)}^{P_r^r \text{ número de columnas}} \\
 & \underbrace{\{\qquad\qquad\}}_{\text{r-elementos}} & \text{r-uplas} \\
 & \{\ldots\ldots\ldots\} & \\
 & \{\ldots\ldots\ldots\} & \\
\underset{\text{número de filas}}{C_n^r} & \ldots\ldots\ldots\ldots & \\
 & \ldots\ldots\ldots\ldots & \\
 & \ldots\ldots\ldots\ldots & \\
 & \ldots\ldots\ldots\ldots & \\
 & \underbrace{\{\qquad\qquad\}}_{\text{r-elementos}} & \underset{\text{r-uplas}}{(\,.\;.\;.\,)}\;.\;.\;.\;.\qquad\qquad \underset{\text{r-uplas}}{(\,.\;.\;.\,)}
\end{array}
$$

El número total de permutaciones de n elementos diferentes tomados r a r que se pueden formar con los n elementos del conjunto es P_n^r . Es decir, el número total de elementos de la tabla. Luego,

$$C_n^r \;.\; P_r^r = P_n^r$$

de donde,
$$C_n^r = \frac{P_n^r}{P_r^r} = \frac{\dfrac{n!}{(n-1)!}}{r!} = \frac{n!}{r!\,(n-r)!}$$

$$\boxed{C_n^r = \frac{n!}{r!\,(n-r)!}}$$

Este número, en Matemática tiene un símbolo especial

$$\binom{n}{r} = \frac{n!}{r!\,(n-1)!}$$

llamado *Coeficiente Binomial*, porque aparece como coeficiente en el desarrollo del binomio $(a + b)^n$.

NOTA Una combinación de n elementos distintos tomados de a r, en realidad es una partición del conjunto en dos subconjuntos (o celdas), donde una de ellas contiene r objetos y la otra las n - r restantes. Por lo tanto, el número de tales combinaciones será el número de particiones, es decir

$$\binom{n}{r, n-1} = \frac{n!}{r!(n-r)!} = C_n^r$$

TEOREMA 1.4.10 El número de combinaciones de n objetos diferentes tomados - r a la vez, es

$$C_n^r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

EJEMPLO 21 Se extraen dos cartas de una baraja de 52 cartas. ¿De cuántas maneras se puede hacer esto?

SOLUCION Se necesita sólo subconjuntos de dos cartas, sin importar el orden entonces, en número de forma de seleccionar estas dos cartas es

$$C_{52}^2 = \frac{52!}{2!50!} = \frac{52 \times 51}{2} = 26 \times 51 = 1326$$

EJEMPLO 22 Un estudiante tiene que contestar 8 de 10 preguntas en un examen.

(a) ¿De cuántas maneras puede el estudiante escoger las 8 preguntas?

(b) Si las tres primeras son obligatorias, ¿de cuántas maneras puede escoger las preguntas?

(c) Si tiene que contestar 4 de las 5 primeras. ¿De cuántas formas puede - hacerlo?

SOLUCION

(a) Como interesa subconjuntos de 8 preguntas de un conjunto de 10 preguntas sin importar el orden, esto sería de

$$C_{10}^8 = \frac{10!}{8!2!} = 45 \quad \text{formas.}$$

(b) Puesto que las tres primeras preguntas son obligatorias; las 5 restantes tendrá que escoger de las 7 preguntas sobrantes. Luego, esto se hace de

$$1 \cdot C_7^5 = \frac{7!}{5!2!} = 21 \quad \text{formas .}$$

(c) Si tiene que contestar 4 de las 5 primeras, lo haría de $C_5^4 = 5$ maneras. Y las 4 preguntas restantes seleccionará de las 5 preguntas finales, lo cual lo haría de $C_5^4 = 5$ formas. Entonces, las 8 preguntas se selecionará de

$$C_5^4 \; . \; C_5^4 = 5 \; . \; 5 = 25 \text{ formas.}$$

EJEMPLO 23 ¿De cuántas maneras puede seleccionarse una partida de 4 o más - personas, si hay 10 personas disponibles?

SOLUCION Nos interesa los subconjuntos de 4,5,6,7,8,9, y 10 personas que se pueden formar con las 10 personas disponibles. Esto se hará de

$$C_{10}^4 + C_{10}^5 + C_{10}^6 + C_{10}^7 + C_{10}^8 + C_{10}^9 + C_{10}^{10} =$$
$$\frac{10!}{4!6!} + \frac{10!}{5!5!} + \frac{10!}{6!4!} + \frac{10!}{7!3!} + \frac{10!}{8!2!} + \frac{10!}{9!1!} + \frac{10!}{10!0!} = 848$$

EJEMPLO 24 Suponga que queremos formar comisiones de cuatro miembros de un grupo de 4 hombres R,S,T,U y 5 mujeres V,W,X,Y,Z. Si además se especifica - que R y S no pueden estar en la misma comisión a menos que la comisión esté formado por lo menos por una mujer. ¿Cuál es el número de comisiones que se puede formar?

SOLUCION Los diferentes casos que se presentan en la solución del problema, se visualiza esquemáticamente en el siguiente diagrama

RS' ⟶ C_7^3. R está en la comisión y S no está, entonces de las 9 personas quedan 7, y de estos 7 deben salir las 3 restantes de la comisión.

R'S ⟶ C_7^3. Igual que el caso anterior, S está y R no está.

R'S' ⟶ C_7^4. Ni R ni S están en la comisión, entonces de los 7 deben salir los 4 de la comisión

RS

1 M ⟶ $C_5^1 \; C_2^1$. R y S están en la comisión, entonces hay C_5^1 formas de escoger una mujer y C_2^1 es el número de formas de selecionar el otro hombre de {T,U}

2 M ⟶ C_5^2. Igual R y S están en la comisión y hay dos mujeres en la comisión lo cual se - puede hacer de C_5^2 formas.

Luego, el número total de comisiones aceptable es

$$2\,C_7^3 \ + \ C_7^4 \ + \ C_5^1 . C_2^1 + C_5^2 = 125$$

Observe que el problema puede resolverse así: Hay $C_9^4 = 126$ comisiones que se pueden formar con los 4 hombres y 5 mujeres. Y hay una comisión conformada por los cuatro hombres R,S,T,U no eceptable. Luego, el número de comisiones aceptables es 126 - 1 = 125.

EJEMPLO 25 ¿Cuántas comisiones integradas por un chico y una chica pueden formarse de cinco chicos y ocho chicas, si cierto chico rehusa trabajar con dos chicas?

SOLUCION Si llamamos R al chico que no quiere trabajar con dos chicas. Los diferentes casos se muestra en el diagrama siguiente

$R \longrightarrow C_6^1$. R está en la comisión, entonces de las 8 - chicas quedan solamente 6.

$R' \longrightarrow C_4^1 \, . \, C_8^1$. R no está en la comisión, entonces de los 4 restantes chicos sale 1 y otro - de las 8 chicas

Por lo tanto, el número total de comisiones aceptables será

$$C_6^1 \ + \ C_4^1 \times C_8^1 = 6 + 32 = 38.$$

Observe, que el problema puede resolverse así: hay $C_8^1\,C_5^1 = 40$ comisiones, - en los cuales hay dos que no es aceptable. Luego, hay 40-2=38 comisiones aceptables.

EJEMPLO 26 El asta de bandera de un barco tiene tres posiciones en las que puede colocarse una bandera. Suponiendo que el barco lleva cuatro banderas - (diferentes) para hacer señales.

(a) ¿Cuántas señales diferentes pueden hacerse con una bandera? (se supone - que la misma bandera colocada en posiciones diferentes indica diferentes señales).

(b) ¿Cuántas señales diferentes pueden hacerse con dos banderas?

(c) ¿Cuántas señales diferentes pueden hacerse con las banderas?

SOLUCION (a) Veamos primero de cuántas formas puede seleccionarse una bandera, lo cual se puede hacer de C_4^1 formas. Puesto que la misma bandera, colocada en posiciones diferentes indica señales diferentes, entonces por cada bandera seleccionada hay P_3^1 , señales diferentes. Por lo tanto, el número total de señales que puede hacerse con una bandera es

$$C_4^1 \ . \ P_3^1 = 4 \times 3 = 12.$$

(b) En este caso el número de maneras de escoger 2 banderas de entre los cuatro que lleva el barco es C_4^2 y por cada uno de estos subconjuntos de dos banderas hay P_3^2 señales diferentes que se pueden formar (es una permutación puesto que interesa la posición donde se pone cada bandera). Luego, el número de señales diferentes que pueden hacerse con 2 banderas es

$$C_4^2 \ . \ P_3^2 = \frac{4!}{2!2!} \ . \ \frac{3!}{(3-2)!} = 36.$$

(c) Puesto que se dispone sólo de 3 astas, entonces escogeremos subconjuntos de 3 banderas de las 4 existentes para señales. Esto se hace de C_4^3 formas ; y para cada uno de estas combinaciones hay P_3^3 señales diferentes que pueden hacerse. Entonces, el número de sañales diferentes que se pueden hacer con todas las banderas es

$$C_4^3 \ . \ P_3^3 = \frac{4!}{3!1!} \ . \ 3! = 24 \ .$$

EJEMPLO 27 Una joven tiene 15 amigos .

(a) ¿De cuántas maneras puede invitar a una cena a 6 de ellos?

(b) Si entre las 15 personas hay dos matrimonios y cada pareja asisten juntos a cualquier reunión. ¿De cuántas maneras puede invitar a 6 amigos?

(c) Si entre las 15 personas hay 2 que no pueden estar en la misma reunión. ¿De cuántas formas puede invitar a 6 amigos?

SOLUCION En cada caso se debe extraer un conjunto de 6 personas, entonces se trata de un problema de combinaciones.

(a) El número de maneras de elegir 6 personas de un grupo de 15 es

$$C_{15}^6 \ .$$

(b) Se puede presentar los siguientes casos:

(i) Ninguno de los dos matrimonios estan en la cena, entonces de los 11 restantes debe elegirse 6. El número de formas de elegir 6 de 11 es C_{11}^6

(*ii*) Uno de los matrimonios asiste a la reunión y los 4 restantes debe - elegirse de las otras 11 personas.
El número de formas de elegir un matrimonio de los dos es C_2^1 , y el número de maneras de escoger 4 personas de 11 es C_{11}^4.
Por lo tanto, el número de formas de elegir los 6, es $C_2^1 \times C_{11}^4$.

(*iii*) Los dos matrimonios asisten a la cena. El número de formas de elegir los dos matrimonios, es C_2^2 .
Las dos personas restantes se elige de los 11; y el número de maneras es C_{11}^2 . Por lo tanto, el número de formas de elegir los 6, es $C_2^2 \times C_{11}^2$.
Finalmente, el número de maneras de invitar a 6 amigos, es

$$C_{11}^6 + C_2^1 \times C_{11}^4 + C_2^2 \times C_{11}^2 .$$

(c) Se presentan los siguientes casos :

(*i*) Ninguna de las personas en cuestión están en la reunión, entonces - las 6 personas se eligen de las 13. El número de maneras de escoger 6 de 13 es

$$C_{13}^6 .$$

(*ii*) Una de las personas en cuestión está en la reunión, entonces las 5 personas restantes se eligen de las 13. El número de maneras de escoger una de las dos personas en cuestión de C_2^1 y el número de formas de elegir 5 de 13 es C_{13}^5 . Por lo tanto, el número de elegir las 6 personas, es $C_2^1 \times C_{13}^5$.
Finalmente el número de maneras de elegir 6 amigos en el caso (c), es

$$C_{13}^6 + C_2^1 \times C_{13}^5 .$$

1.4.7 NOTAS SOBRE MUESTREO CON Y SIN REEMPLAZO

Suponga un conjunto con **n** objetos. Considere el problema anterior de extraer **r** objetos de este conjunto. Puede no interesar el orden en que se extraen los objetos. También la extracción se puede hacer con o sin reemplazamiento. Sobre este último hemos dado una breve nota al pasar en 1.2. Aquí - formalizaremos estos conceptos.

DEFINICION 1.4.3 Si al extraer los r objetos del conjunto de n objetos, se considera el orden en que son seleccionados los objetos; el conjunto de los r objetos extraídos, se llama una *muestra ordenada de tamaño* r .

DEFINICION 1.4.4 Cuando un objeto se extrae y se reemplaza antes de extraer el siguiente objeto, se dice que el *muestreo es con reemplazamiento*.

Calculemos ahora, el número de formas de extraer una muestra ordenada de tamaño r de un conjunto de n objetos, si el muestreo es con reemplazamiento.

La primera extracción ocurre de n formas, uno por cada objeto; la segunda extracción también ocurre de n formas, ya que el muestreo es con reemplazamiento. Entonces, el número de formas de extraer dos objetos con reemplazamientos será $n.n = n^2$. Igualmente para la extracción de 3 objetos el número de formas es

$$n^2 . n = n^3$$

y para cualquier número r de extracciones, el número de formas será n^r .

DEFINICION 1.4.5 Si al extraer un objeto no se reemplaza, para extraer el siguiente, se dice que el *muestreo es sin reemplazamiento*.

El número de formas de extraer muestras ordenadas de tamaño r de un conjunto de n objetos, si el muestreo es sin reemplazamiento se obtiene así: la primera extracción ocurre de n formas, la segunda extracción ocurre de n - 1 formas. Entonces, el número de formas de extraer dos objetos sin reemplazamiento es n(n - 1). Similarmente para la tercera extracción hay n - 2 formas Luego, el número de formas de extraer 3 objetos es n(n - 1)(n - 2). Así sucesivamente, el número de formas de extraer una muestra de tamaño r, sin reemplazimiento es

$$n(n-1)(n-2)\ldots(n-r+1) = \frac{n!}{(n-r)!}$$

el cual es equivalente a P_n^r , número de permutaciones de n objetos tomados r a r. Si no interesa el orden en que se extraen los elementos de la muestra El número de maneras de escoger r objetos de los n, está dado por

$$C_n^r = \binom{n}{r} .$$

EJEMPLO 28 Considere las placas de automóviles que tiene tres letras seguidas de tres dígitos. Si pueden emplearse todas las combinaciones posibles, ¿cuántas placas diferentes pueden formarse?.

SOLUCION Las placas forman 6-uplas ordenadas; donde las tres primeras son letras y las tres siguientes dígitos. Como no hay restricciones respecto a las letras y números que se usan; es decir puede usarse letras y números iguales. Entonces, hay $(26)^3$ formas ordenadas de extraer las tres letras y $(10)^3$ formas de extraer los tres números. Luego, el número de formas de formar placas con tres letras seguidas de tres dígitos es

$$(26)^3 \cdot (10)^3 = 17576(10)^3$$

PROBLEMAS 1.4

1. ¿Cuántos números se pueden formar con los dígitos {1,2,3,4,5}. Suponiendo que no pueden repetirse estos?

2. ¿Cuántos números de tres cifras pueden formarse con los dígitos {0,1,2,3, 4}, si no pueden repetirse estos?

3. ¿Cuántos números de tres cifras distintos existen?

4. ¿De cuántas formas posibles pueden salir de una aula los 25 alumnos que estan en ella?. (Se sobreentiende que salen de uno por uno).

5. En un salón de clase se quiere sentar a 6 jóvenes y 5 chicas en una sóla fila, de manera que las chicas ocupen los lugares pares. ¿De cuántas maneras se puede hacer?

6. (a) ¿Cuántos números de tres cifras pueden formarse con los dígitos 0,1, 2,3,4 y 5, si cada dígito se utiliza una sóla vez?
   (b) ¿Cuántos de ellos son impares?
   (c) ¿Cuántos de ellos son mayores de 330?

7. Hay dos obras de 3 volúmenes cada una y otras dos de dos volúmenes cada una. ¿De cuántas maneras pueden colocarse los 10 libros en un estante, si deben quedar de tal manera que no se separen los volúmenes de la misma obra?

8. ¿Cuántas palabras distinguibles se pueden hacer con las letras de la palabra MISSISSIPPI?

9. ¿Cuántos números diferentes de 12 cifras pueden formarse si se dispone de

los dígitos: 2,2,2,2,4,4,4,5,5,5,5,5?

10. Dada una caja con los siguientes focos: 2 de 25 vatios, 3 de 50 vatios y 4 de 100 vatios.
    (a) ¿De cuántas maneras pueden escogerse 3 de ellos?
    (b) ¿Cuántas de estas selecciones de tres incluirán a los 2 de 25 vatios? ¿Cuántos no contendrán los de 25 vatios?
    (c) ¿Cuántas selecciones de tres focos incluirán exactamente uno de cada uno de las potencias?

11. ¿Cuántas cantidades diferentes de dinero pueden formarse con las monedas, siguientes, 1 de 50 centavos, 1 de un sol, 1 de 5 soles, 1 de 10 soles, 1 de 50 soles y 1 de 100 soles?

12. ¿Cuántos equipos de fútbol pueden formarse con 12 hombres que puedan ocupar cualquier posición delantera y 10 hombres que puedan ocupar cualquiera de las demás posiciones?

13. En cada caso determine el valor de n. Si,
    (a) $C_n^2 = C_n^8$; (b) $C_n^{11} = C_n^7$; (c) $C_{18}^n = C_{18}^{n-2}$.

14. Si $C_{(18,4)} - C_{(18,n+2)} = 0$, determine el valor de C_n^5 .

15. ¿De cuántas formas diferentes pueden arreglarse tres focos rojos, cuatro amarillos y tres azules en una serie navideña que contiene diez portafocos?

16. Un estudiante del primer año debe llenar un programa que consiste en un curso de idioma extranjero, uno de ciencias naturales, uno de ciencias sosiales y uno de español. Si hay cuatro posibilidades para escojer el idioma extranjero, seis para el curso de ciencias naturales, tres para el curso de ciencias sociales y dos para el curso de español, ¿De cuántas maneras puede llenar su programa el estudiante?

17. En una Biblioteca hay 8 libros de geometría, 14 de álgebra, 10 de física y 5 de química. ¿De cuántas maneras puede un estudiante seleccionar cuatro libros, de manera que sea uno de cada curso mencionado?

18. Un club tiene 15 miembros, 10 hombres y 5 mujeres, ¿cuántos comités de 8 miembros se pueden formar:
    (a) Si cada uno de ellos debe contener por lo menos 3 mujeres?

(b) Si en cada uno de ellos debe estar el presidente y la secretaria del club?

19. En 10 tubos de prueba se cultivan tres tipos de bacterias, tres tubos contienen bacterias del primer tipo, cuatro contienen bacterias del segundo tipo y tres bacterias del tercer tipo. De cuántas maneras distintas pueden ponerse en un porta-tubos, teniendo en cuenta solamente el orden del tipo de bacterias.

20. Un caballero entra a una tienda que tiene en exhibición 12 corbatas diferentes, a saber: 5 de tipo italiano, 4 de tipo inglés y 3 de tipo nacional ¿Cuántas compras diferentes puede hacer, si desea llevar como mínimo una corbata del tipo italiano y una del tipo inglés?

21. Un grupo de 14 viajeros, de los cuales 6 son mujeres y 8 son varones, deben ser alojado en un hotel que posee 7 habitaciones tales que puedan ser instalados dos viajeros en cada una de ellas.
    (a) ¿De cuántas formas posibles se pueden ubicar?
    (b) ¿De cuántas formas posibles se pueden ubicar, si en cada habitación debe estar personas de igual sexo?. Si en el grupo hay dos matrimonios. ¿De cuántas formas se pueden ubicar, si se desea que cada matrimonio ocupe una habitación y el resto de las habitaciones no sea ocupado por personas de distinto sexo?

22. En una clínica trabajan 18 enfermeras.
    (a) ¿Cuántas guardias diferentes de 3 enfermeras pueden formarse?
    (b) ¿En cuántas guardias de las formadas en (a) estará una enfermera determinada?

23. ¿Cuántas manos de poker de cinco cartas consisten de :
    (a) dos pares (dos cartas iguales)?
    (b) cuatro de la misma clase (iguales)?
    (c) full (3 cartas de la misma denominación y 2 de otra)?

24. Cinco amigos se encuentran en una fiesta. ¿Cuántos saludos de mano se intercambian si cada amigo estrecha la mano de todos los demás sólo una vez?

25. Un biólogo intenta clasificar 46,200 especies de insectos asignando a cada especie tres iniciales del alfabeto. ¿Será la clasificación completa? ¿sino, cuál es el número de iniciales que debería ser usado.

26. La mesa de sesiones del rectorado de cierta Universidad es rectangular; - er una sesión ordinaría, asiste el Rector, el secretario, nueve directores de programas académicos y dieciocho jefes de departamentos. El Rector y - el secretario ocupan permanentemente la cabecera y al frente de ellos están también permanentemente los dos directores más antiguos; el resto de los asistentes se sientan en las partes laterales de la mesa. Se pregunta ¿De cuántas maneras diferentes pueden sentarse en la mesa los asistentes (todos los asistentes) a la seción mencionada?

27. Para transmitir señales de día, se dispone de cuatro banderas triangulares distintas, y de tres juegos iguales, compuestos cada uno por nueve - banderas rectangulares, distintas entre si. Cada señal debe consistir en una bandera triangular, seguida de tres, dos, una o ninguna rectangular ; se desea saber que número de señales distintas pueden hacerse y qué sería mas perjudicial en cuanto a dicho número se refiere, si perder tres banderas rectangulares iguales o dos juegos de esta clase de banderas.

28. Para transmitir señales de una isla a la costa, se dispone de 6 luces - blancas y 6 rojas, colocados en los vértices de un exágono. En cada vértice no puede haber encendida más que una luz (blanca, roja) y el número mínimo de luces encendidas es tres. Hallar el número de señales distintas - que se pueden formar.

29. Una firma comercial tiene 10 vendedores. ¿De cuántas formas puede asignarse los vendedores en dos escritorios con cinco vendedores en cada escritorio? ¿Con siete en un escritorio y tres en la otra?.

30. Una firma comercial tiene diez vendedores. ¿De cuántas formas pueden los vendedores ser asignados a tres escritorios con tres en el primer escritorio, tres en el segundo y cuatro en el tercer escritorio?.

1.5 DEFINICION DE PROBABILIDAD

Aunque parece sorprendente, daremos aquí tres enfoques que dan lugar a tres definiciones diferentes de "probabilidad" lo cual puede ser chocante al que se inicia en estas cuestiones, al disponer de tres definiciones de probalidad, pues se preguntará, ¿Qué definición de probabilidad va a utilizar?. - Sin embargo, no lo es tanto ya que cualquiera que sea la definición de probabilidad que se utilice, ***las reglas de probabilidad son las mismas*** (Axiomas , Teoremas, etc); más aún las tres definiciones son complementarias y la definición adecuada de probabilidad que se use dependerá de la naturaleza del problema específico que se está tratando de resolver. Estas definiciones son

(1) Definición clásica o apriori.
(2) Definición de probabilidad por frecuencia relativa o aposteriori. Las definiciones (1) y (2) son probabilidades objetivas.
(3) Probabilidad subjetiva.

1.5.1 DEFINICION CLASICA

La definición clásica de probabilidad fue dada por Laplece en su obra "Teoría Analítica de las probabilidades" publicas en 1812. Esta definición se basa en el supuesto de que todos los resultados posibles de un experimento aleatorio son igualmente probables; es decir, cada uno de los elementos del espacio muestral tienen la misma posibilidad de salir. Así, por ejemplo, si lanzamos un dado no cargado (honesto), debe considerarse que hay igual posibilidad que salga cualquiera de los números del espacio muestral $\Omega = \{1, 2,3,4,5,6\}$, entonces la probabilidad que salga cualquier número será $\frac{1}{6}$.

En general si un experimento aleatorio tienen n resultados posibles, los n elementos del espacio muestral tendrán la misma posibilidad de salir. En consecuencia la probabilidad de que salga cualquiera de ellos es $\frac{1}{n}$. Las observaciones anteriores conducen a la definición clásica de probabilidad de un evento como la siguiente:

> La probabilidad de un evento es la razón entre el número de casos (sucesos) favorables y el número total de casos (sucesos) posibles, siempre que nada oblique a creer que algunos de estos sucesos debe tener preferencia a los demás, lo que hace que todos sean igualmente posibles.

En la definición anterior si, $N(\Omega) = n$, es el número de elementos del espacio muestral (número total de sucesos) y $N(A) = n_A$, es el número de elementos del evento A (o número de sucesos favorables); la probabilidad del evento A, denotado por "P[A]" es la razón de N(A) a N(Ω), o sea

$$\boxed{P[A] = \frac{N(A)}{N(\Omega)} = \frac{n_A}{n} = \frac{\text{número de casos favorables al evento A}}{\text{número de casos posibles}}}$$

OBSERVACIONES

1. La probabilidad de un evento cualquiera A está comprendido entre 0 y 1. - En efecto: n_A y n son enteros positivos y $0 \leq n_A \leq n$, Dividiendo por **n** se tiene

$$\frac{0}{n} \leq \frac{n_A}{n} \leq \frac{n}{n}, \text{ ó } 0 \leq P[A] \leq 1$$

2. $P[A] = 0$, si A es un evento imposible.
En efecto: $A = \phi$, entonces $n_A = 0$, luego $P[A] = \frac{0}{n} = 0$

3. $P[A] = 1$, si A es el evento seguro.
En efecto: $A = \Omega$, entonces $n_A = n$, luego $P[A] = \frac{n}{n} = 1$.

4. Puesto que todos los elementos de $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ son igualmente probables, se tiene $P\left[\{\omega_i\}\right] = \frac{1}{n}$, $i = 1, 2, \ldots, n$.

$$\text{y} \quad P[\Omega] = \sum_{i=1}^{n} P\left[\{\omega_i\}\right] = 1 .$$

Si A es un evento en Ω, entonces

$$P[A] = \sum_{\omega_i \in A} P\left[\{\omega_i\}\right] .$$

EJEMPLO 1 Si se lanza una moneda tres veces. Calcular la probabilidad que Ocurran: (a) dos caras ; (b) al menos dos caras; (c) a lo más dos caras.
SOLUCION El experimento aleatorio es, "lanzar una moneda tres veces". El espacio muestral asociado a este experimento es

Ω = {CCC , CCS , CSC , SCC , CSS , SCS , SSC , SSS}; luego $N(\Omega) = 8$

(a) Sea el evento A: "Ocurre dos caras".
Los sucesos favorables a **A** son {CCS , CSC , SCC}, o sea $N(A) = 3$
Por lo tanto, $P[A] = \frac{3}{8}$.

(b) Sea el evento B: "Ocurre al menos dos caras".
Los sucesos favorables al evento B son {CCS , CSC , SCC , CCC}, o sea $N(B) = 4$ .Por lo tanto, $P[B] = \frac{4}{8} = \frac{1}{2}$.

(c) Sea el evento C: "Ocurre a lo más dos caras".

Entonces, $C = \{SSS, SSC, SCS, CSS, CCS, CSC, SCC\}$ y $N(C) = 7$

Por lo tanto, $P[C] = \frac{7}{8}$.

Obsérvece, que decir ocurren dos caras no es lo mismo que decir ocurren al menos dos caras, pués en este último caso, 3 caras también es un suceso - favorable y si fuera posible la ocurrencia de 4,5, etc. caras también serían sucesos favorables. Algo similar ocurre cuando decimos ocurren a lo más 2 caras en cuyo caso; 0,1,2 caras son sucesos favorables. Aclararemos esto en - otros ejemplos posteriores.

EJEMPLO 2 Consideremos el lanzamiento de dos dados. Calcular la probabilidad de

(a) obtener suma 7 (b) obtener suma 6 ; (c) obtener suma mayor que 5;
(d) que el resultado del primer dado sea mayor que el resultado del segundo.

SOLUCION El experimento aleatorio es, "lanzar dos dados" .
El espacio muestral asociado a este experimento, es el conjunto de pares ordenados, en las que la primera componente es el resultado del primer dado y las segunda componente el resultado del segundo, como hemos visto en el ejemplo 7 de 1.2. y $N(\Omega) = 36$.

Sean los eventos siguientes :

$A = \{(\omega_1,\omega_2) \in \Omega \;/\; \omega_1 + \omega_2 = 7\}$ = obtener suma 7.

$B = \{(\omega_1,\omega_2) \in \Omega \;/\; \omega_1 + \omega_2 = 6\}$ = obtener suma 6.

$C = \{(\omega_1,\omega_2) \in \Omega \;/\; \omega_1 + \omega_2 > 5\}$ = obtener suma mayor que 5

$D = \{(\omega_1,\omega_2) \in \Omega \;/\; \omega_1 > \omega_2\}$ = el resultado del primer dado es mayor que del segundo.

$$\Omega = \left\{\begin{array}{l} (1,1), (1,2), (1,3), (1,4), (1,5), (1,6) \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6) \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6) \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6) \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \end{array}\right\}$$

Un simple conteo nos permite determinar $N(A) = 6$, $N(B) = 5$, $N(C) = 26$, $N(D) = 15$. Entonces :

(a) $P[A] = \frac{6}{36} = \frac{1}{6}$; (b) $P[B] = \frac{5}{36}$;

(c) $P[C] = \frac{26}{36} = \frac{13}{18}$; (d) $P[D] = \frac{15}{36} = \frac{2}{12}$.

El lector puede calcular la probabilidad de obtener suma 5, $\frac{4}{36}$; suma 4, $\frac{3}{36}$ etc. y si calcula la probabilidad de obtener suma 2,3,8,9,10,11,12, se dará cuenta porqué se apuesta a obtener suma 7.

Respecto a este mismo ejemplo, podemos concebir el resultado de nuestro experimento como la suma de los resultados individuales, (es decir, se lanza dos dados y se observa la suma obtenida de ambos), en este caso el espacio muestral sería $\Omega = \{2,3,4,5,6,7,8,9,10,11,12\}$. Es claro que ahora los sucesos no ocurren con la misma frecuencia, es decir no son igualmente posibles. Si suponemos que los sucesos son igualmente posibles tendríamos, por ejemplo que la probabilidad que ocurra 7 sería $\frac{1}{11}$, sin embargo hemos visto que esta probabilidad es $\frac{1}{6}$.

EJEMPLO 3 En una caja hay 20 bolas numeradas del 1 al 20. Se extrae al azar una bola. ¿Cuál es la probabilidad que el número de la bola extraída
(a) no exceda de 20? (b) sea el 32? (c) sea por lo menos 15?

SOLUCION El experimento aleatorio es "extraer una bola de la caja". Luego

$$\Omega = \{①, ②, \ldots, ⑳\} \quad \text{y} \quad N(\Omega) = 20$$

(a) Sea el evento A: "el número de la bola extraida no exceda de 20".
Como el número de cualquier bola que se halla en la caja no supera a 20, el evento A es igual al espacio muestral Ω, o sea $A = \Omega$, en este caso el evento A es el evento seguro. Por lo tanto.

$$N(A) = N(\Omega) = 20 \quad \text{y} \quad P[A] = 1 .$$

(b) Sea el evento B: "el número de la bola extraída es el 32".
Como en la caja no hay bola marcada con el número 32, B es el evento imposible, o sea $B = \phi$. Entonces

$$N(B) = 0, \quad \text{y} \quad P[B] = 0/20 = 0$$

(c) Sea el evento C: "el número de la bola extraída sea por lo menos 15".

$$C = \{⑮, ⑯, \ldots, ⑳\} , \quad N(C) = 6 \quad \text{y} \quad P[C] = 6/20 = 3/10.$$

EJEMPLO 4 Una lotería consta de 10000 billetes. Un billete se premia con 100 intis, cuatro billetes con 50 intis, diez billetes con 20 intis, veinte billetes con 10 intis, 165 billetes con 5 intis y 400 billetes con 1 inti cada uno. Los demás billetes no se premian. Se compra un billete, ¿cuál es la probabilidad de ganar

(a) por lo menos 10 intis? (b) a lo más 5 intis?

SOLUCION El experimento aleatorio es "elegir un billete".

$\Omega = \{B_1, B_2, \ldots, B_{10000}\}$ donde B_i representa el billete número i .

$N(\Omega) = 10000$

(a) Sea el evento A: "ganar almenos 10 intis".

Ganar almenos 10 intis significa que se puede ganar 10 intis, ó 20 intis ó 50 intis ó 100 intis. Es decir $N(A) = 20 + 10 + 4 + 1 = 35$, luego

$$P[A] = \frac{35}{10000} = 0.0035 \cdot$$

(b) Sea el evento B : "ganar a lo más 5 intis".

Ganar a lo maś 5 intis significa que se puede ganar 1 inti ó 5 intis; o sea $N(B) = 400 + 165 = 565$. Luego

$$P[B] = \frac{565}{10000} = 0.0565 .$$

EJEMPLO 5 Se elige una carta aleatoriamente de una baraja de 52 cartas. ¿Cuál es la probabilidad que sea un palo negro (espadas o tréboles)?. ¿Cuál que sea un diez?. ¿Cuál que sea una figura (Rey, Reyna, Sota)?. ¿Cuál es la probabilidad que sea un cuatro o menos?.

SOLUCION El experimento aleatorio es, "extraer una carta de 52".

El espacio muestral Ω tiene 52 elementos. O sea $N(\Omega) = 52$.

(a) Sea el evento A: "obtener un palo negro".

A tiene 26 elementos (13 espadas + 13 tréboles). Luego,

$$P[A] = \frac{26}{52} = \frac{1}{2} .$$

(b) Sea el evento B: "obtener un diez".

B tiene 4 elementos (pues hay 4 dieces). Entonces,

$$P[B] = \frac{4}{52} = \frac{1}{13}.$$

(c) Sea el evento C: "obtener una figura".
C tiene 12 elementos (4 Reyes + 4 Reinas y 4 Sotas). Luego,

$$P[C] = \frac{12}{52} = \frac{3}{13}.$$

(d) Sea al evento D: "obtener un cuatro o menos".
D tiene 16 elementos (4 Ases + 4 Dos + 4 tres + 4 Cuatros). Luego,

$$P[D] = \frac{16}{52} = \frac{4}{13}.$$

EJEMPLO 6 Dos personas A y B se distribuyen al azar en tres oficinas numeradas con 1,2, y 3 respectivamente, pudiendo estar ambos en una misma oficina. ¿Cuál es la probabilidad que

(a) la oficina 2 se quede vacía? (b) dos oficinas se queden vacias?

SOLUCION PRIMERA FORMA Definimos los siguientes eventos:

A_i : "la persona A está en la oficina i, (i = 1,2,3)".

B_j : "la persona B está en la oficina j, (j = 1,2,3)".

A_iB_j : "la persona A está en la oficina i y B en la oficina j,(i,j=1,2,3)".

E : "la oficina 2 se queda vacia".

F : "dos oficinas se quedan vacias".

El espacio muestral apropiado al experimento aleatorio "distribuir dos personas en tres oficinas", es

$$\Omega = \begin{Bmatrix} A_1 B_1 , A_1 B_2 , A_1 B_3 \\ A_2 B_1 , A_2 B_2 , A_2 B_3 \\ A_3 B_1 , A_3 B_2 , A_3 B_3 \end{Bmatrix}, \text{ o sea } N(\Omega) = 9.$$

(a) Los sucesos favorables al evento E son $\{A_1B_1 , A_3B_1 , A_1B_3 , A_3B_3\}$, o sea - $N(E) = 4$, luego

$$P[E] = \frac{4}{9}.$$

(b) Los sucesos favorables al evento F son $\{A_1B_1 , A_2B_2 , A_3B_3\}$, osea $N(F) = 3$ entonces

$$P[F] = \frac{3}{9} = \frac{1}{3} .$$

SEGUNDA FORMA El problema puede resolverse utilizando técnicas de conteo sin escribir el espacio muestral como sigue: Calculemos el número de elementos del espacio muestral. Ambas personas pueden estar en la misma oficina, entonces, la persona **A** puede distribuirse en cualquiera de las tres oficinas, lo que da 3 formas, la persona B puede también distribuirse en cualquiera de las tres oficinas; dando igualmente 3 formas. Por lo tanto, el número de formas de distribuir las dos personas es 3.3 = 9. Es decir $N(\Omega) = 9$.

(a) Hallaremos ahora el número de casos favorables al evento,
E : "la oficina 2 se queda vacía".

La oficina 2 se queda vacía es equivalente a que las dos personas se distribuyen en las dos oficinas restantes, y esto se puede hacer de $2^2 = 4$ formas Luego,

$$P[E] = \frac{4}{9} .$$

(b) Calculemos el número de casos favorables al evento,
F : "dos oficinas se quedan vacías".
Dos oficinas se quedan vacías es equivalente a que las 2 personas están en la misma oficina, y esto ocurre de 3 formas (la primera persona puede ubicarse en cualquiera de las 3 oficinas y la segunda de una sola forma, en la oficina del primero); entonces

$$P[F] = \frac{3}{9} = \frac{1}{3} .$$

EJEMPLO 7 En una compañía hay 6 varones y 4 damas que aspiran a ser miembros de un comité. Si se deben escoger 2 al azar escribiendo los nombres en hojas de papel y sacándolos de una urna. ¿Cuál es la probabilidad que los dos sean hombres? ¿Cuál que sean un hombre y una mujer o dos mujeres?

SOLUCION El experimento aleatorio es "extraer 2 nombres de los 10".
El espacio muestral tiene la forma $\Omega = \{\{A,B\},\{A,C\}, \ldots\}$. Como cada suceso es un conjunto de dos personas el espacio muestral Ω tiene C_{10}^2 elementos.
(a) Sea el evento A: "los dos sean hombres"
A tiene C_6^2 elementos. Por lo tanto

$$P[A] = \frac{C_6^2}{C_{10}^2} = \frac{1}{3} .$$

(b) Sea el evento B: "Sean un hombre y una mujer o dos mujeres".
B es un evento compuesto. La primera parte "sea un hombre y una mujer" ocurre de 6 x 4 formas, por el principio de multiplicación , la segunda parte "sean dos mujeres" ocurre de C_4^2 formas. Por el principio de adición se tiene N(B) = 6 x 4 + C_4^2 , luego

$$P[B] = \frac{24 + C_4^2}{C_{10}^2} = \frac{30}{C_{10}^2} = \frac{2}{3} .$$

EJEMPLO 8 Sobre una mesa hay 10 monedas con 4 caras y 6 sellos a la vista. Se separan 6 monedas al azar. ¿Cuál es la probabilidad que resultan 3 caras y 3 sellos?.

SOLUCION El experimento aleatorio, es "separar 6 monedas de las 10 que hay" Entonces el espacio muestral consta de subconjuntos de 6 elementos, o sea

$$\Omega = \{\{ Ⓒ,Ⓒ,Ⓢ,Ⓢ,Ⓢ,Ⓢ \} , \{......\},...\}$$
$$N(\Omega) = C_{10}^6 .$$

Sea el evento A: "resultan 3 caras y 3 sellos".
El evento A es compuesto. La primera parte; con cuatro caras tomadas de tres en tres, se forman C_4^3 grupos diferentes.
La segunda parte: con seis sellos, tomados de tres en tres se forman C_6^3 grupos distintos.
Como cada uno de los C_4^3 grupos de 3 caras puede combinarse con cada uno de los C_6^3 grupos de 3 sellos, por el principio de multiplicación, se pueden formar $C_4^3\, C_6^3$ grupos diferentes de 3 caras y tres sellos cada uno. Es decir, N(A) = $C_4^3\, C_6^3$. Por lo tanto,

$$P[A] = \frac{C_4^3\, C_6^3}{C_{10}^6} = \frac{4.20}{210} = \frac{8}{21} .$$

EJEMPLO 9 De 20 personas que contrajeron cierta enfermedad al mismo tiempo y que fueron llevados a una misma sala de un hospital, 15 se recuperan completamente en 3 días; al cabo del cual, se escogen aleatoriamente 5 personas para un chequeo. ¿Cuál es la probabilidad que los 5 sean dados de alta? ¿Cuál es la probabilidad que exactamente 4 sean dados de alta? ¿Cuál es la probabilidad que ninguno sea dado de alta?

SOLUCION El experimento aleatorio es escoger 5 personas de los 20. Luego, - el espacio muestral tiene la forma $\Omega = \{\{A, B, C, D, B\}, \{\ldots\ldots\}, \ldots\}$ por lo tanto, el espacio muestral Ω tiene $\binom{20}{5}$ elementos.

(a) Sea el evento A: "las 5 personas sean dado de alta".
(es decir, todos están sanos). Entonces, el evento A tiene $\binom{15}{5}$ elementos (ya que 15 de los 20 están sanos) . Luego,

$$P[A] = \frac{\binom{15}{5}}{\binom{20}{5}} = \frac{\dfrac{15!}{5! \times 10!}}{\dfrac{20!}{5! \times 15!}} = \frac{15! \times 15!}{10! \times 20!} = \frac{1001}{5168} .$$

(b) Sea el evento B: "exactamente 4 personas están sanos".
B es un evento compuesto. La primera parte: con 15 personas sanos tomados de a 4 se forman $\binom{15}{4}$ grupos diferentes. La segunda parte: con 5 personas enfermas tomados de a 1 se forman $\binom{5}{1}$ grupos diferentes. Luego ,

el número de elementos de B es $\binom{15}{4}\binom{5}{1}$. Por lo tanto,

$$P[B] = \frac{\binom{15}{4}\binom{5}{1}}{\binom{20}{5}} = \frac{5005}{5168} .$$

(c) Sea el evento C: "ninguna persona esta sana".
C tiene $\binom{5}{5} = 1$ elemento. Luego, $P[C] = 1/\binom{20}{5}$.

EJEMPLO 10 Se tiene cuatro urnas numeradas de 1 a 4 y cuatro bolas también numeradas de 1 a 4. Se coloca al azar una bola en cada urna. ¿Cuál es la probabilidad que la bola i sea colocada en la urna i ($i = 1,2,3,4$)? ¿Cuál es - la probabilidad que la bola 1 sea colocada en la urna 1 y la bola 2 en la urna 2?.

SOLUCION El experimento aleatorio es "colocar una bola en cada urna".
Puesto que cada bola se coloca en la urna al azar, entonces: la primera bola tendrá 4 formas de ser colocada en una urna; la segunda bola tendrá 3 formas de ser colocada en una de las urnas restantes. Por lo tanto, las dos primeras

bolas de colacan de 4 x 3 formas. La tercera bola se colocará de 2 formas - (en una de las 2 urnas restantes). Las tres primeras se colocarán de 4 x 3 x 2 formas. La última bola tendrá una sóla forma de ser colocada. Luego, el número de formas es 4 x 3 x 2 x 1 = 4!. Es decir

$$N(\Omega) = 4! = {}_4P_4$$

El espacio muestral tiene la forma $\Omega = \{\boxed{3}\boxed{2}\boxed{1}\boxed{4}, \boxed{1}\boxed{2}\boxed{4}\boxed{3}, \ldots\}$

Definimos los siguientes eventos :

B_i : "la bola i sea colocada en la urna i ($i = 1,2,3,4$)".

(a) $i = 1$; B_1: "la bola 1 sea colocada en la urna 1".

La bola 1 se colocará en la urna 1 de una sóla forma y las otras tres bolas se colocarán de 3! formas. O sea

$$N(B_1) = 1.\ 3! = 3!$$

En general, la bola i se colocará en la urna i de una sóla forma y las tres restantes se colocarán de 3! formas. O sea

$$N(B_i) = 1.\ 3! = 3! = {}_3P_3,\ (i = 1,2,3,4).$$

Entonces, $$P[B_i] = \frac{{}_3P_3}{{}_4P_4} = \frac{3!}{4!} = \frac{1}{4}, \quad i = 1,2,3,4.$$

(b) Sea A: "la bola 1 sea colocada en la urna 1 y la bola 2 en la urna 2"

Note que $A = B_1$ y B_2 .

La bola 1 se colocará en la urna 1 de una sóla forma, la bola 2 lo mismo, y las dos bolas restantes se colocarán de 2! formas. Es decir

$$N(A) = 1 \times 1 \times 2! = {}_2P_2 .$$

Luego, $$P[A] = \frac{{}_2P_2}{{}_4P_4} = \frac{2!}{4!} = \frac{1}{12} .$$

EJEMPLO 11 Un experimento aleatorio consiste en disponer los dígitos 2,3,4, 5,6,7,8,9, uno a continuación del otro; calcular la probabilidad de:

(a) Que el 3 aparece junto al cuatro y en ese orden.

(b) El número formado sea par;

(c) El número formado sea mayor que 6×10^7 ;

(d) El número formado sea múltiplo de 4;

(e) El número formado sea múltiplo de 3.

SOLUCION El experimento aleatorio es disponer uno a continuación del otro los dígitos 2,3,4,5,6,7,8,9. Entonces, $\Omega = \{42698753, 72934856, \ldots\}$; es decir cada elemento del espacio muestral son arreglos de los 8 dígitos. Por lo tanto, tiene 8! elementos. O sea $N(\Omega) = 8! = {}_8P_8$

(a) Sea el evento A: "El 3 aparece junto al cuatro y en ese orden".
Es decir, en las diferentes ordenaciones el 3 y el 4 aparecen juntos, así 34 y en ese orden, luego puede considerarse como uno solo. Por lo tanto, quedan solamente 7 número por permutar. Entonces, el evento A tiene 7! elementos, es decir $N(A) = 7! = {}_7P_7$. Luego,

$$P[A] = \frac{{}_7P_7}{{}_8P_8} = \frac{7!}{8!} = \frac{1}{8} .$$

(b) Sea el evento B: "el número formado sea par".
Para que el número formado sea par, debe terminar en 2,4,6, u 8. Las cuales se pueden escoger de 4 formas y para cada uno de estas habrá 7! formas de ordenar los 7 dígitos restantes. Entonces, el evento B tiene $4 \times 7!$ elementos.

Luego,
$$P[B] = \frac{4 \times 7!}{8!} = \frac{1}{2} .$$

(c) Sea el evento C: "el número formado sea mayor que 6×10^7".
Para que el número formado sea mayor que 6×10^7, tiene que empezar con los dígitos 6,7,8 ó 9; los cuales se pueden escoger de 4 formas y para cada uno de estos hay 7! formas de ordenar los 7 dígitos restantes. Es decir, el evento C tiene $4 \times 7!$ elementos. Por lo tanto,

$$P[C] = \frac{4 \times 7!}{8!} = \frac{1}{2} .$$

(d) Sea D, el evento: "el número formado es múltiplo de 4".
Un número es múltiplo de 4 cuando lo es, las dos últimas cifras, y ese conjunto es

32	24	36	28
52	64	56	48
72	84	76	68
92		96	

Es decir, hay 14 posibles terminaciones, y para cada una de estas hay, 6! formas de ordenar los dígitos restantes. Entonces, el evento D tiene $14 \times 6!$

elementos. Luego,

$$P[D] = \frac{14 \times 6!}{8!} = \frac{2 \times 7!}{8!} = \frac{1}{4} .$$

(e) Sea E, el evento: "el número formado sea múltiplo de 3".
Un número es múltiplo de 3, si la suma de los valores de sus cifras es múltiplo de 3. Pero, 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 44. Entonces, cualquier número que se forme no será múltiplo de 3. Por lo tanto, el evento E tiene 0 elementos. Luego,

$$P[E] = \frac{0}{8!} = 0 .$$

EJEMPLO 12 Diez libros se colocan aleatoriamente en un estante. Determinar la probabilidad que tres libros determinados, sean colocados juntos.

SOLUCION El experimento es "colocar 10 libros diferentes en un estante". Entonces, los elementos del espacio muestral son arreglos de los libros. Luego el número de elementos del espacio muestral Ω es 10! . Esto es,

$$n = N(\Omega) = 10! .$$

Sea el evento, **A**: "Tres libros determinados queden juntos".
Puesto que los 3 libros deben estar juntos, podemos considerar como si fuera uno sólo; luego, en vez de los diez habrá sólo 8 libros, los cuales pueden colocarse de 8! formas. Pero, los tres libros también pueden cambiar de posición entre ellas, la cual se hace de 3! formas. Por lo tanto, el evento **A** tiene 8! .3! elementos. Luego,

$$P[A] = \frac{8!.3!}{10!} = \frac{1}{15} .$$

EJEMPLO 13 Si se revuelven las 11 letras de la palabra "MISSISSIPPI" y se dispone en un orden arbitrario. ¿Cuál es la probabilidad que

(a) en la ordenación resultante las cuatro ies sean letras consecutivas?

(b) las cuatro ies sean consecutivas supuesto que la ordenación empieza por "M" y termina en "S"?

(c) las cuatro ies sean consecutivas, supuesto que la ordenación termina con las cuatro eses consecutivas?

SOLUCION (a) El experimento aleatorio es, disponer en orden arbitrario las letras de la palabra MISSISSIPPI. Luego, el espacio muestral Ω tiene $P_{11}^{1,4,4,2}$ elementos, que es una permutación de las 11 letras con repetición, de los -

cuales 2 letras "S" e "I" se repiten cuatro veces, la "P" dos veces y la "M" una vez.

Sea A, el evento: "en la ordenación resulta las cuatro ies consecutivas", - Una forma sería, por ejemplo

$$\begin{array}{cccccccccccc} \uparrow M & IIII & S & \uparrow P & \uparrow S & \uparrow S & \uparrow P & \uparrow S \\ & \uparrow & & & & & & \\ & 2 & & & & & & \end{array}$$

Para las cuatro ies hay 8 posiciones diferentes, y para cada uno de estas - las 7 letras restantes pueden ordenarse de $P_7^{1,4,2}$ maneras. Entonces, el evento A tiene $8 \times P_7^{1,4,2}$ elementos. Por lo tanto,

$$P[A] = \frac{8 \times P_7^{1,4,2}}{P_{11}^{1,4,4,2}} = \frac{\dfrac{8 \times 7!}{4!\ 2!}}{\dfrac{11!}{4!\ 4!2!}} = \frac{8 \times 7!4!}{11!} = \frac{4}{165}\ .$$

(b) Sea B, el evento: "los cuatro ies resultan consecutivas".

Supuesto que la ordenación empieza con "M" y termina en una "S".

Entonces, el espacio muestral en este caso se reduce a la variación de 9 letras ya que hay dos fijas. Es decir el nuevo espacio muestral Ω_1 tiene $P_9^{4,3,2}$ elementos; 4 ies y 3 eses y 2 pees.

Uno de los elementos del evento B tiene, por ejemplo, la siguiente forma

$$\begin{array}{cccccccc} M & S & IIII & P & S & S & P & S \\ \uparrow & & \uparrow & & \uparrow & \uparrow & \uparrow & \uparrow \\ 1 & & 2 & & 3 & 4 & 5 & 6 \end{array}\ .$$

Es decir, para las cuatro ies hay 6 posiciones, y para cada uno de estas, - las 5 letras restantes pueden ordenarse de $P_5^{3,2} = \binom{5}{2}$ formas. Luego, el - evento B tiene $6\binom{5}{2}$ elementos. Por lo tanto,

$$P[B] = \frac{6\binom{5}{2}}{P_9^{4,3,2}} = \frac{\dfrac{6 \times 5!}{3!2!}}{\dfrac{9!}{4!3!2!}} = \frac{6 \times 5! \times 4!}{9!} = \frac{1}{21}\ .$$

(c) En este caso la ordenación términa en las cuatro eses. Entonces, quedan sólo 7 letras que varían. Por lo tanto, el espacio muestral Ω_2 tiene $P_7^{4,2,1}$

elementos.

Sea C, el evento: "las cuatro ies resultan consecutivas".

Un elemento del evento C tiene, por ejemplo la siguiente forma

$$\underset{\substack{\uparrow\\1}}{M}\quad \underset{\substack{\uparrow\\2}}{P}\quad \underset{\substack{\uparrow\\3}}{IIII}\quad \underset{\substack{\uparrow\\4}}{P}\quad SSSS$$

Hay cuatro posiciones para las cuatro ies, y para una de estas posiciones - las 3 letras restantes pueden ordenar de $P_3^{2,1} = \binom{3}{1}$ maneras. Por lo tanto, el evento C tienen $4\binom{3}{1}$ elementos. Luego

$$P[C] = \frac{4\binom{3}{1}}{P_7^{4,2,1}} = \frac{\dfrac{4 \times 3!}{2!}}{\dfrac{7!}{4!2!1!}} = \frac{4 \times 3!4!}{7!} = \frac{4}{35} .$$

EJEMPLO 14 De una baraja de 52 cartas se extraen al azar 6 cartas. Determinar la probabilidad que tres de ellas sean de oro y dos de copas.

SOLUCION El experimento aleatorio es, "extraer 6 cartas de la baraja de 52". Entonces, los elementos del espacio muestral Ω son subconjuntos de seis cartas cada uno. Por lo tanto tiene C_{52}^{6} elementos. Es decir,

$$n = N(\Omega) = C_{52}^{6} = \binom{52}{6} .$$

Definimos el evento A de la siguiente manera:

A : "tres cartas sean de oro, dos de copas y uno diferente de oro y copa".

El número de formas de extraer 3 cartas de oro, de un total de 13, es

$$C_{13}^{3} = \binom{13}{3} .$$

El número de formas de extraer 2 cartas de copas, de un total de 13, es

$$C_{13}^{2} = \binom{13}{2} .$$

El número de formas de extraer una carta (que no sea oro ni copa), de las 26 cartas restantes, es

$$C_{26}^{1} = \binom{26}{1} .$$

Por lo tanto, el número de elementos del evento **A**, es $n_A = \binom{13}{3}\binom{13}{2}\binom{26}{1}$

Luego, $$P[A] = \frac{n_A}{n} = \frac{\binom{13}{3}\binom{13}{2}\binom{26}{1}}{\binom{52}{6}}$$

EJEMPLO 15 Se extraen 5 cartas al azar de una baraja de 52 cartas. ¿Cuál es la probabilidad de

(a) extraer exactamente 2 parejas (dos cuatros y dos 10 dieces, por ejemplo y uno diferente)?
(b) Extraer "full" (3 cartas iguales y 2 de otra también iguales)?
(c) Extraer "flor" (5 cartas del mismo color)?
(d) una corrida (5 cartas, comenzando por el as, o dos, o tres, ..., o con - el diez)?

SOLUCION El experimento aleatorio es, extraer 5 cartas de las 52. Entonces, el espacio muestral Ω tiene $\binom{52}{5}$ elementos.

(a) Sea **A**, el evento: "extraer dos parejas y uno diferente".
Un juego de naipes de 52 cartas contiene 13 grupos de números iguales así,

[A,A,A,A] , [2,2,2,2] , [3,3,3,3] , ...,[Q,Q,Q,Q] , [K,K,K,K]

Se necesita 2 grupos de cuatro de los 13 que hay, y esto se puede extraer de $\binom{13}{2}$ formas. Y de cada uno de estos grupos de cuatro, hay $\binom{4}{2}$ formas de extraer una pareja. Además de uno de los 11 grupos restantes debe extraerse una carta, esto se hace de 11 x$\binom{4}{1}$ formas (o también de las 44 cartas - restantes debe escogerse una y esto se hace de $\binom{44}{1}$ formas). Entonces, el - evento A tiene $\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{44}{1}$ elementos. Por lo tanto,

$$P[A] = \frac{\binom{13}{2}\binom{4}{2}^2\binom{44}{1}}{\binom{52}{5}} \quad \text{ó} \quad P[A] = \frac{\binom{13}{2}\binom{4}{2}^2 \times 11 \times \binom{4}{1}}{\binom{52}{5}}.$$

(b) Sea B, el evento: "extraer full".

Hay 13 formas de extraer un grupo de 4 y de este grupo hay $\binom{4}{3}$ forma. de extraer 3 cartas iguales. Existen 12 formas de escoger el segundo grupo de cuatro y de este hay $\binom{4}{2}$ formas de extraer 2 iguales. Entonces, el evento B tiene $13 \times 12 \times \binom{4}{3} \times \binom{4}{2}$ elementos. Por lo tanto,

$$P[B] = \frac{13 \times 12 \times \binom{4}{3}\binom{4}{2}}{\binom{52}{5}} .$$

(c) Sea C, el evento: "obtener flor".

En una baraja de 52 cartas, hay 4 grupos de 13 cartas cada uno de un mismo color. Entonces hay 4 formas de extraer uno de estos grupos de 13 cartas, y de este grupo hay $\binom{13}{5}$ formas de extraer 5 cartas del mismo color. Entonces, el evento C tiene $4\binom{13}{5}$ elementos. Luego,

$$P[C] = \frac{4\binom{13}{5}}{\binom{52}{5}} .$$

(d) Sea D, el evento: "obtener una corrida".

Hay $\binom{4}{1}$ formas de extraer un "as", también un dos, un tres, un 4 y un 5. Es decir hay $\binom{4}{1}^5$ formas de extraer una corrida que comienza con as.

También $\binom{4}{1}^5$ formas de extraer una corrida que comienza con dos y así sucesivamente, hay $\binom{4}{1}^5$ formas de extraer una corrida que comienza con 10 Luego, el evento D tiene $10. \binom{4}{1}^5$ elementos. Por lo tanto,

$$P[D] = \frac{10\binom{4}{1}^5}{\binom{52}{5}} .$$

EJEMPLO 16 De una urna que contiene doce bolas, de las cuales ocho son blancas y 4 negras, se extrae una muestra de tamaño 4 con reemplazo (sin reemplazo). Hallar la probabilidad que la muestra contenga exactamente tres bolas blancas.

SOLUCION (a) Si la muestra se extrae con reemplazo, se tiene

$$n = N(\Omega) = (12)^4.$$

Sea A, el evento: "la muestra contiene exactamente tres bolas blancas". Los elementos de A tiene la forma

$$A = \{\underbrace{bbbn}_{P_4^{3,1} = \binom{4}{3}}, \ldots\}, \text{ el orden interesa.}$$

la bola blanca sale de 8 formas en cada extracción (cuatro extracciones de las cuales 3 son blancas, lo que da 8^3 formas) y la negra de 4 formas. Luego el número de casos favorables es $N(A) = \binom{4}{3} 8^3.4$. Por lo tanto,

$$P[A] = \frac{\binom{4}{3} \cdot 8^3.4}{(12)^4} = \frac{32}{81}$$

(b) Si la muestra se extrae sin reemplazo, entonces el número de elementos del espacio muestral es

$\binom{12}{4}$, sin considerar el orden .

P_{12}^4, considerando el orden.

Sea B, el evento: "la muestra contiene exactamente tres bolas blancas". Entonces, $n_B = \binom{8}{3}\binom{4}{1}$ Sin considerar el orden.

Considerando el orden los elementos de B tiene la forma

$$B = \{\underbrace{bbbn}_{P_4^{3,1} = \binom{4}{3}}, \ldots\}$$

las bolas blancas pueden salir de $8 \times 7 \times 6 = \frac{8!}{5!} = P_8^3$ formas y la negra de 4 formas, luego

$$n_B = \binom{4}{3} P_8^3 \cdot 4 .$$

Por lo tanto, hay dos formas de determinar, la probabilidad del evento B.

(*i*) considerando el orden, $P[B] = \dfrac{\binom{4}{3} P_8^3 \cdot 4}{P_{12}^4}$.

(*ii*) Sin considerar el orden, $P[B] = \dfrac{\binom{8}{3}\binom{4}{1}}{\binom{12}{4}}$.

El lector puede verificar la igualdad de ambas soluciones.

EJEMPLO 17 Se distribuye al azar 6 bolas diferentes entre 3 cajas, ¿Cuál es la probabilidad que la primera caja contenga 3 bolas? ¿Cuál que hayan 2 bolas en cada caja?.

SOLUCION El experimento es "distribuir 6 bolas en tres cajas al azar"

Ω = {  , . . . }

El número de elementos del espacio muestral, se calcula así: la primera bola tiene 3 formas de distribuirse (cualquiera de las tres cajas). La segunda bola también puede distribuirse de 3 formas. Luego las dos bolas se distribuye de $3.3 = 3^2$ formas. La tercera bola, igualmente se distribuye de 3 formas. Por lo tanto las 3 bolas se distribuirán de $3^2.3 = 3^3$ formas. Así, sucesivamente, las 6 bolas se distribuirán de 3^6 formas. Por lo tanto

$$n = N(\Omega) = 3^6$$

(a) Sea A, el evento: "la primera caja contiene 3 bolas".

La primera caja contiene 3 bolas es equivalente a que las 3 bolas restantes se distribuye en las 2 cajas que quedan. El número de formas que se distribuye las 3 bolas en 2 cajas es 2^3. Y el número de formas de escoger grupos de 3 elementos es $\binom{6}{3}$. Por lo tanto, el número de casos favorables al evento A es

$$n_A = \binom{6}{3} 2^3$$

y

$$P[A] = \frac{\binom{6}{3} 2^3}{3^6} = \frac{5.2^5}{3^6} = 0.219 .$$

(b) B: "hayan 2 bolas en cada caja"

El número de formas que ocurra el evento B, es el número de particiones - de las 6 bolas en 3 subconjuntos con 2 bolas en cada uno. Es decir,

$$N(B) = \binom{6}{2,2,2}$$

Luego, $$P[B] = \frac{\binom{6}{2,2,2}}{3^6} = \frac{6!}{(2!)^3\,3^6}$$

EJEMPLO 18 Seis amigos desean viajar en el tren electrico suburbano compuesto por 3 vagones. Si cada uno escoge su vagón al azar (es decir todos tienen igual posibilidad de viajar en cualquiera de los vagones), ¿cuál es la probabilidad que

(a) todos viajan en un mismo vagón?

(b) esten distribuidos en los tres vagones?

SOLUCION El número de elementos del espacio muestral es $3^6 = N(\Omega)$.

(a) Sea el evento B: "los 6 amigos viajan en el mismo bagón"

$$N(B) = 3$$

En efecto: la primera persona se puede ubicar de 3 formas. Una vez ubicado - el primero, el segundo tiene una sola forma (el vagón donde se ubicó el primero), así sucesivamente los restantes tienen una sola forma de ubicarse. - Por lo tanto, los seis se ubicarán de

$$3 \times 1 \times 1 \times 1 \times 1 \times 1 = 3 \quad \text{formas}$$

Osea $$P[A] = \frac{3}{3^6} = \frac{1}{3^5}\,.$$

(b) B: "los seis amigos están distribuidos en los 3 vagones"

Esto se hace particionando las 6 personas en 3 grupos. Las diferentes particiones de los seis en tres grupos son:

1. Haya 2 personas en cada grupo, esto ocurre de $\binom{6}{2,2,2}$ formas. Esta partición, se ubica de una sóla forma en los vagones. Osea el número de formas de ubicar las 6 personas con 2 en cada vagón es $\binom{6}{2,2,2}$.

2. Haya 4 personas en un grupo y uno en cada grupo restante. Esto se hace de $\binom{6}{4,1,1}$ formas. Esta partición, se ubica en los vagones de $P_3^{2,1} = 3$ for

diferentes. Es decir, el número total de formas de ubicar a las personas en esta condición es $3\binom{6}{4,1,1}$

3. Haya 1,2 y 3 en cada vagón respectivamente, lo cual se hace de $\binom{6}{1,2,3}$ - formas. Esta partición, se ubica en los vagones de ${}_3P_3 = 3!$ formas diferentes. Por lo tanto, el número total de formas de ubicar a las 6 personas cumpliendo esta condición es $3!\binom{6}{1,2,3}$.

De (1), (2) y (3) el número de elementos del evento B es

$$N(B) = \binom{6}{2,2,2} + 3\binom{6}{4,1,1} + 3!\binom{6}{1,2,3}$$

luego,

$$P[B] = \frac{\binom{6}{2,2,2} + 3\binom{6}{4,1,1} + 3!\binom{6}{1,2,3}}{3^6}$$

1.5.2 DEFINICION POR FRECUENCIA RELATIVA

Planteamos antes las siguientes preguntas:

1. ¿Cuál es la probabilidad que un dado (lanzado sobre una superficie plana y liza) se pare sobre su borde?
2. ¿Cuál es la probabilidad que un vendedor de televisores en blanco y negro y a color vende uno a color en su próxima venta?
3. ¿Cuál es la probabilidad que la mitad o más de los alumnos que llevan el curso de estadística obtenga una nota aprobatoria en el curso?

El experimento hipotético empleado en la definición clásica de probabilidad, no nos ayuda a contestar estas preguntas. Para contestar cualquiera - de estas preguntas necesitamos más información. Así, pués en el primer caso, enumerar los resultados posibles de un sólo lanzamiento del dado no nos ayudará a determinar realmente la probabilidad que se quede parado sobre su borde. En el segundo caso; "vende un televisor a color", "vende un televisor en blanco y negro" y "vende uno a color y otro en blanco y negro" son tres - eventos que no son mutuamente excluyentes por lo tanto ¿qué probabilidad se asigna a cada uno de estos eventos?. Finalmente en la tercera pregunta "la - mitad o más obtendrán nota aprobatoria" y "menos de la mitad no aprobarán" - son dos eventos mutuamente excluyentes que dan todo el espacio muestral. Pero ¿Está ud. seguro al utilizar el principio de que todos los resultados posibles tienen la misma posibilidad de salir, para asignar una probabilidad de 1/2 a cada evento?.

Es decir, estas preguntas no se pueden responder utilizando la definición clásica de probabilidad. Sin embargo son preguntas razonables que una persona puede plantearse. Por lo tanto, son razones claras para fundamentar otra definición de probabilidad.

Lo expuesto anteriormente nos conduce a la siguiente interpretación de probabilidad en términos de frecuencia relativa.

> Si un experimento bien definido se repite n veces (n grande); sea $n_A < n$ el número de veces que el evento A ocurre en los n ensayos, entonces la frecuencia relativa de veces que ocurre el evento A "$\frac{n_A}{n}$", es la estimación de la probabilidad que ocurra el evento A, o sea
>
> $$P[A] = \frac{n_A}{n}$$

Además en la última parte del párrafo anterior, la estimación de la probabilidad por frecuencia relativa de un evento A, $\frac{n_A}{n}$ se acerca a la verdadera probabilidad de un evento, cuando n aumenta indefinidamente, es decir

$$P[A] = \lim_{n\to\infty} \frac{n_A}{n}$$

Naturalmente, en la práctica esto no es posible, sólo podemos buscar una estimación máxima de P[A] basada en n grande.

OBSERVACIONES

1. La frecuencia relativa de un evento, está comprendido entre 0 y 1. Por lo tanto $0 \leq P[A] \leq 1$.
   En efecto: Desde que $0 \leq n_A \leq n$, $\frac{0}{n} \leq \frac{n_A}{n} \leq 1$, se tiene que

   $0 \leq \frac{n_A}{n} \leq 1$. Luego, $0 \leq P[A] \leq 1$.

2. $\frac{n_A}{n} = 0$, si sólo si, en las n repeticiones del experimento el evento A -

no ocurre.

En efecto: si **A** no ocurre $n_A = 0$, $\frac{n_A}{n} = \frac{0}{n} = 0$, inversamente si $\frac{n_A}{n} = 0$, entonces $n_A = 0$.

Por lo tanto $P[A] = 0$ si, sólo si **A** no ocurre en las **n** repeticiones del experimento.

3. $\frac{n_A}{n} = 1$, si, sólo si el evento A ocurre en las **n** repeticiones del experimento. En particular $\frac{n_\Omega}{n} = 1$.

En efecto: Si **A** ocurre en las n veces que se repite el experimento, entonces

$n_A = n$; así $\frac{n_A}{n} = \frac{n}{n} = 1$. Inversamente, $\frac{n_A}{n} = 1$, si

$n_A = n$, lo que significa que el evento **A** ha ocurrido en las n repeticiones del experimento.

Por lo tanto $P[A] = 1$, si, sólo si **A** ocurre en todas las repeticiones del experimento.

EJEMPLO 19 Una muestra aleatoria de 10 fábricas que emplean un total de 10,000 personas, demostró que ocurriron 500 accidentes de trabajo durante un período reciente de 12 meses. Hallar la probabilidad de un accidente de trabajo en una industria determinada.

SOLUCION n = 10,000, número de veces que se repite el experimento.

Sea el evento A: "un accidente de trabajo en la industria determinada". Entonces $N(A) = 500$ y

$$P[A] = \frac{500}{10,000} = 0.05$$

Por definición de frecuencia relativa, ya que este valor de la probabilidad, se basa en una muestra, por lo tanto es una estimación del valor real desconocido. Observe, aquí se supone implícitamente que las formas de seguridad no han cambiado desde que se realizó el muestreo.

EJEMPLO 20 La distribución de los miembros de los partidos políticos es

Partido	A	B	C	D	E	F
Número Total de Militantes	105	100	70	45	40	15
Militantes Mujeres	15	20	5	10	3	2

¿Cuál es la probabilidad que un miembro seleccionado aleatoriamente,

(a) Sea una mujer?

(b) pertenece al partido B?

(c) Sea un hombre miembro del partido C?

SOLUCION n = total de socios = 105 + 100 + 70 + 45 + 40 + 15 = 375

(a) Sea A, el evento: "el socio seleccionando sea una mujer".

n_A = número de socios mujeres = 15 + 20 + 5 + 10 + 3 + 2 = 55.

Luego,

$$P[A] = \frac{55}{375} .$$

(b) Sea B, el evento: "el miembro seleccionado pertenece al partido B".

n_B = total de miembros del partido B = 100.

Entonces,

$$P[B] = \frac{100}{375} .$$

(c) Sea C, el evento: "la persona seleccionada es hombre y pertenece al partido C".

Y n_C = total de hombres que pertenecen al partido

$$n_C = 70 - 5 = 65$$

Luego,

$$P[A] = \frac{65}{375}$$

EJEMPLO 21 En una serie de observaciones sobre la longitud de vida de ratas machos, un biólogo ha encontrado que el 98% sobrevive aun 200 días después de nacer; 83% sobrevive 400 días; 40% sobreviven 600 días; 8% sobreviven 800 días; y no sobreviven después de 1000 días.

Estime las probabilidades de los eventos:

(a) "se mueren dentro los primeros 200 días";

(b) "muere entre 200 y 400 días" ; (c) "sobreviven a lo más 400 días" ;
(d) "mueren dentro los 1000 días" ;

SOLUCION Como no se conoce el número exacto de ratas; asumimos que hay 100x ratas. O sea $n = 100x$

(a) Si A, el evento: "la rata se muere dentro los primeros 200 días".
Como 98% de las ratas viven después de los 200 días de nacer, entonces - $\frac{98 \cdot 100x}{100} = 98x$ ratas viven 200 días después de nacer, luego hay - $100x - 98x = 2x$ ratas muertas. Por lo tanto,

$$P[A] = \frac{2x}{100x} = 0.02 \; .$$

(b) Sea B, el evento: "una rata se muere entre 200 y 400 días".
El número de ratas muertas entre 200 y 400 días es

$$98x - 83x = 15x$$

luego,

$$P[B] = \frac{15x}{100x} = 0.15 \; .$$

(c) Sea C, el evento: "una rata sobreviva a lo más 400 días".
Desde que 83x sobrevivan 400 días, tenemos

$$P[C] = \frac{83x}{100x} = 0.83 \; .$$

(d) Desde que todas las ratas mueren dentro de los 1000 días tenemos que la probabilidad pedida es 1.

1.5.3 PROBABILIDAD SUBJETIVA

Existen muchas situaciones donde el concepto de probabilidad basada en resultados igualmente posibles (definición clásica) y el de frecuencia relativa carece de significado. Por, ejemplo planteamos las siguientes preguntas ¿Cuál es la probabilidad que el presidente de la República vete cierta legislación? ¿Cuál es la probabilidad que una expedición desembarque en el planeta Marte en la próxima década? ¿Cuál es la probabilidad que en el próximo examen de estadística obtenga 15 de notas? Son eventos únicos, que no ha ocurrido antes. No hay forma que se pueda interpretar tales probabilidades como una frecuencia relativa o como una probabilidad clásica. *El enfoque sub*

jetivo de la probabilidad es pues adecuado en casos en que hay sólo una apor-tunidad de ocurrencia del evento y ocurrirá o no ocurrirá esa sola vez.
La probabilidad subjetiva se define así:

> Dado un experimento determinado, la probabilidad de un evento A es el - *grado de creencia* asignado a la ocurrencia de este evento por un individuo particular, basado en toda la evidencia a su disposición, con las siguientes exigencias:
>
> (1) $P[A] = 0$, representa la certeza que el evento A, no ocurrirá.
>
> (2) $P[A] = 1$, representa la certeza que el evento A, si ocurrirá.
>
> (3) $0 < P[A] < 1$, representa el grado de certeza que el evento A, ocurrirá.

Desde que, la probabilidad subjetiva de la ocurrencia de un evento A, - es un número asignado por un individuo de acuerdo a la evidencia que dispone Otra persona con otras evidencias, podría asignar a la ocurrencia del mismo evento A otra probabilidad diferente (un número diferente al anterior).

1.5.4 PROBABILIDAD FRENTE A APUESTAS

Haremos un estudio breve de la relación entre las *apuestas* a favor de - un evento y su probabilidad. Cuando se dice, por ejemplo, que las apuestas - son 3 a 1 que el boxeador A ganará la proxima pelea, significa que hay, 3 posibilidades en 4, que A, ganará la pelea. Las apuestas 3 a 1, se escribe : 3:1, se lee: "las apuestas son de 3 a 1 a favor de A", la cual se convierte - en una probabilidad, así

$$\frac{3}{3+1} = \frac{3}{4} = P[\text{ de que A ganará }]$$

DEFINICION Sea A un evento cualquiera, si las *apuestas* son a:b a favor del evento A, entonces la probabilidad que ocurra dicho evento es

$$\boxed{P[A] = \frac{a}{a+b}}$$

Además, decir que las apuestas son a:b a favor del evento A es lo mismo de-cir que las apuestas son b:a en contra del evento A. Entonces, la probabilidad que no ocurra A es

$$P[A] = \frac{b}{a+b}.$$

Puede también la probabilidad de un evento A convertirse en apuestas a favor de la ocurrencia del evento A; así, si P[A] es la probabilidad que ocurra el evento A; las apuestas a favor de la ocurrencia de **A** son

$$P[A] : [1 - P[A]]$$

Y las apuestas en contra de él son $[1 - P[A]]: P[A]$.

EJEMPLO 22 En una carrera de caballos, el caballo "claudio" tiene las apuestas 5:1 en su contra, mientras que el caballo "Royal" las tiene 9:1 en su contra. ¿Cuál es la probabilidad que cualquiera de estos caballos gane?

SOLUCION Sean los eventos:

C : "el caballo claudio gane la carrera",

R : "el caballo royal gane la carrera".

Entonces, $P[C] = \frac{1}{5+1} = \frac{1}{6}$, $P[R] = \frac{1}{9+1} = \frac{1}{10}$

Luego, $P[C \cup R] = P[C] + P[R] = \frac{1}{6} + \frac{1}{10} = \frac{4}{15}$

ya que C y R son eventos mutuamente excluyentes.

1.5.5 PROBABILIDAD EN ESPACIOS MUESTRALES FINITOS

Otro método para asignar probabilidades, en espacios muestrales finitos (es decir $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$), es como sigue: se asigna la probabilidad p_i a cada resultado ω_i, o sea $p_i = P[\{\omega_i\}]$ tal que

(1) $p_i > 0, \quad i = 1,2,3,\ldots, n$

(2) $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} P[\{\omega_i\}] = 1$

La suma de las probabilidades asignados a los puntos del espacio muestral es la unidad. (Teniendo en cuenta que los posibles resultados $\omega_1, \omega_2, \ldots, \omega_n$ son mutuamente excluyentes y colectivamente exahustivos).

La probabilidad de un evento **A** es la suma de las probabilidades asignadas de los puntos muestrales pertenecientes al evento **A**; esto es

$$P[A] = \sum_{i/\omega_i \in A} p_i = \sum_{\omega_i \in A} P[\{\omega_i\}] .$$

EJEMPLO 23 Ocho amigos juegan boliche una vez a la semana. Este grupo está formado por 2 parejas de casados, 3 jóvenes y una joven. Antes del juego cada uno coloca I/.100. en una bolsa cuyo contenido ganará el que obtenga mayor puntuación Si las mujeres tienen la mitad de la habilidad que los varones poseen para el juego. ¿Cuál es la probabilidad que un soltero gane?. ¿Cuál es la probabilidad que gane una mujer?. ¿Cuál es la probabilidad que gane un hombre casado?.

SOLUCION El espacio muestral Ω tiene 8 elementos, 5 hombres que tienen igual habilidad en el juego y 3 mujeres que tienen la mitad de habilidad que los hombres. Por lo tanto si p es la probabilidad de ganar de un hombre, entonces $\frac{1}{2}p$ es la probabilidad de ganar de una mujer.

Luego, se debe tener que $5p + \frac{3}{2}p = 1$, de donde

$$p = \frac{2}{13} \quad y \quad \frac{1}{2}p = \frac{1}{13} .$$

(a) Sea **A**, el evento: "gane un hombre soltero"

El evento **A** tiene 3 elementos con igual habilidad, por lo tanto

$$P[A] = 3\left(\frac{2}{13}\right) = \frac{6}{13}$$

(b) Sea **B**, el evento: "gane una mujer"

El evento **B** tiene 3 elementos con igual habilidad, es decir

$$P[B] = 3\left(\frac{1}{13}\right) = \frac{3}{13} .$$

(c) Sea **C**, el evento: "gane un hombre casado"

El evento **C** tiene 2 elementos con igual habilidad. Luego,

$$P[C] = 2\left(\frac{2}{13}\right) = \frac{4}{13} .$$

EJEMPLO 24 Un dado está cargado de tal modo que la probabilidad de obtener 1,2,3,4,5 ó 6 es proporcional a los números 1,2,3,4,5 y 6, respectivamente.. Si se lanza este dado, calcular la probabilidad del evento: "el resultado es un número par".

SOLUCION Como $P[\{k\}]$ es proporcional a k, para todo

$k = 1,2,3,4,5,6,$ se tiene que

$$P[\{k\}] = rk,$$ donde r es la constante de proporcionalidad. Por lo tanto, se debe tener que

$$\sum_{k=1}^{6} P[\{k\}] = r + 2r + 3r + 4r + 5r + 6r = 21r = 1$$

de donde $r = \frac{1}{21}$. Luego, $P[\{k\}] = \frac{1}{21}k$

Y si A, es el evento: "obtener un número par". Es decir $A = \{2,4,6\}$, tenemos

$$P[A] = \sum_{k\in A} P[\{k\}] = \sum_{k\in A} \frac{1}{21}k = 2(\frac{1}{21}) + 4(\frac{1}{21}) + 6(\frac{1}{21}) = \frac{12}{21} = \frac{4}{7}.$$

NOTA Cuando los resultados elementales son igualmente posibles, se tiene

$$p_1 = p_2 = \ldots = p_n, \quad y$$

$$P[A] = \frac{n_A}{n}$$

donde n_A representa el número de resultados contenidos en el evento A.

PROBLEMAS 1.5

1. Para cada una de las siguientes situaciones, indique ¿Cuál de los enfoques de probabilidad, sería más útil para determinar el valor de la probabilidad.
   (a) La probabilidad que se obtenga un "as" o un diez de "oro" en una sola extracción de una baraja de 52 cartas.
   (b) La probabilidad que una persona escogida aleatoriamente de entre las que entran a una tienda comercial haga una compra en dicha tienda.
   (c) La probabilidad que haya huelga de profesores en el próximo ciclo académico.
   (d) La probabilidad que un producto escogido aleatoriamente de una producción grande sea defectuoso.
   (e) La probabilidad que haya legislatura extraordinaria del congreso, después de la próxima legislatura.
2. Para cada uno de los siguientes experimentos aleatorios defina un espacio muestral adecuado, decir si son o no sus elementos igualmente posibles y

decir, si, se puede aplicar la definición clásica de probabilidad.

(a) Contar el número de pasas en un panetón hecho con pasas y frutillas.

(b) Contar el número de ases al extraer cinco cartas al azar de una baraja ordinaria de 52 cartas.

(c) Se lanzan 2 monedas y contar el número de caras obtenidas.

(d) Contar el número de niños nacidos en un día en cierto hospital.

(e) Una moneda correcta es lanzada hasta que aparece el mismo resultado dos veces sucesivas, contar el número de lanzamientos.

3. En cada uno de los casos siguientes, especificar un espacio muestral apropiado, asignar probabilidades a los sucesos (elementos de Ω) para luego hallar las probabilidades requeridas:

(a) Hallar la probabilidad que una caja quede vacia al distribuir al azar dos objetos distinguibles en dos cajas.

(b) Hallar la probabilidad de encontrar una familia sin niños (varones) entre las familias con tres hijos (ordenar empezando por el mayor).

(c) Calcular la probabilidad de obtener una figura, al extraer una carta aleatoriamente de una baraja de 52 cartas.

4. Clasifique los siguientes estimados de probabilidad por su tipo (clásica, frecuencia relativa, o subjetiva)

(a) La probabilidad que un consumidor demande a una compañía distribuidora de drogas es 0.005.

(b) La probabilidad de enviar por correo terrestre un despacho de Lima a Trujillo en 24 horas es 0.30.

(c) La probabilidad que las ventas en Diciembre sean mayores que en Julio es 0.75.

(d) La probabilidad de sacar una orden de pedido de un grupo de 10 sin mirar es 0.2.

5. El Instituto Nacional del Cancer está planteando mandar por correo un cuestionario sobre el cancer del seno. De experiencias pasadas con estos cuestionarios, el instituto sabe que sólo un 12% de las personas que reciben los cuestionarios los responden. Sin embargo, tambien saben que el 1% de los cuestionarios despachados tienen errores en la dirección y por lo tanto nunca serán puestos al correo, que un 3% se perderan o destruirán en la oficina de correos, que un 22% será remitido a personas que -

han cambiado de residencia y que sólo un 52 % de las personas que cambian de residencia dejan la nueva.

(a) ¿Los porcentajes dados en el problema representan estimados de probabilidad clásica, de frecuencia relativa o subjetiva?.

(b) ¿Cuál es la probabilidad que el instituto reciba la respuesta de un cuestionario?.

6. Se extraen 3 cartas, aleatoriamente, de una baraja de 52 cartas. ¿Cuál es la probabilidad que estas cartas sean; un tres, un siete y un as?

7. ¿Cuál es la probabilidad de obtener 2 caras y 2 sellos al lanzar una moneda cuatro veces?

8. ¿Cuál es la probabilidad que de 6 cartas tomadas de una baraja de 52, 3 sean rojas y 3 negras?

9. Se extraen 3 cartas, aleatoriamente, de una baraja de 36 cartas. Determinar la probabilidad que la suma de los puntos de estas cartas sea 21. Si la sota se cuenta como 2 puntos, el caballo como 3 puntos, el rey como 4 puntos, el as como 11 puntos y el resto como 6,7,8,9 y 10 puntos.

10. Una caja contiene nueve tikets numerados del 1 al 9. Se extraen 3 tikets al azar de la caja uno a uno sin reposición. Determinar la probabilidad que

(a) Sean alternativamente impar, par, impar ó par, impar, par.

(b) Los tres sean pares o impares.

11. Se colocan 6 bolas, aleatoriamente, en tres cajas inicialmente vacías. ¿Cuál es la probabilidad que la primera caja contenga exactamente 2 bolas?

12. Se lanzan 5 bolas en 4 cajas numeradas, de modo que cada bola tenga que caer en una de las cajas y tales que, todas tengan la misma probabilidad de caer en cualesquiera de las cajas. Determinar la probabilidad que en la primera caja caigan 2 bolas y 1 en la segunda.

13. Doce personas desean viajar en un tren que tiene 6 carretas. Cada pasajero selecciona con igual probabilidad cada uno de las carretas. Determinar la probabilidad que:

(a) hayan dos pasajeros en cada carreta.

(b) hayan; una carreta sin pasajeros, una con un pasajero, dos con dos pa sajeros cada uno y los dos restantes con tres y cuatro pasajeros, respectivamente.

14. Se colocan, aleatoriamente, 8 libros en un estante. Entre ellos hay una obra en cuatro tomos y otra en 3.
¿Cuál es la probabilidad que los tomos de cada obra esten juntos?.

15. De una baraja de 52 cartas se extraen, aleatoriamente, 5 cartas. ¿Cuál es la probabilidad que 3 sean de un mismo palo y los otros dos de palos diferentes?

16. Un dado está cargado de tal forma que los números pares tienen la misma probabilidad de salir, los números impares tienen la misma probabilidad de salir, y cada número par tiene probabilidad doble de salir que la de un número impar. Determinar la probabilidad que:
(a) Salga un número par.
(b) Salga un número mayor que 4.

17. Se colocan en un estante 10 obras al azar, entre las cuales hay una en 4 tomos, otra en 3 tomos y los restantes de un sólo tomo. ¿Cuál es la probabilidad que los tomos de cada obra estén juntos?

18. Se extraen tres bolas de tres urnas que contienen cada una nueve bolas numeradas del 1 al 9. Se forma un número cuyas unidades, decenas y centenas se sacan respectivamente, de la 1ra, 2da, y 3ra urna. ¿Cuál es la probabilidad que el número así formado sea múltiplo de 18?

19. Un experimento aleatorio consiste en disponer los dígitos: 1,2,3,4,5,6,7, 8; uno a continuación del otro. Calcular la probabilidad que:
(a) El 4 aparezca junto al 5 en ese orden; (b) El número formado sea par;
(c) El número formado sea mayor que 6×10^7 ;
(d) El número formado sea múltiplo de 4;
(e) El número formado sea múltiplo de 3.

20. Un experimento aleatorio consiste en disponer los dígitos: 1,2,3,4,5,6,7, 8,9; uno a continuación del otro. Calcular la probabilidad que:
(a) El 5 aparezca junto al 6 en ese orden (b) El 5 aparezca junto al 6;
(c) El 5,6 y 7 aparezca juntos y en ese orden;
(d) El 5 aparezca antes que el 6; (e) El número formado sea impar;
(f) El número formado sea menor que 5×10^8;

(g) El número formado sea múltiplo de 25;

(h) El número formado sea múltiplo de 6.

21. Se extraen seis cartas de una baraja ordinaria de 52 cartas, ¿Cuál es la probabilidad de:

(a) extraer una pareja (dos, cuatros por ejemplo) y cuatro que no forman - pareja?

(b) extraer dos parejas y dos que no forman pareja?

22. Se distribuye al azar 12 bolas diferentes entre tres cajas. ¿Cuál es la - probabilidad que la primera caja contenga 3 bolas?.

23. Se distribuye aleatoriamente n bolas diferentes entre N celdas. ¿Cuál es la probabilidad que una celda determinada contenga r bolas?.

24. Dada n celdas en las que se distribuye aleatoriamente n bolas, determinar la probabilidad que una celda quede vacía.

25. ¿Cuál es la probabilidad que una mano de cartas contenga 2 ó más ases?

26. Dado cinco segmentos de longitudes 1,3,5,7 y 9 unidades, hallar la probabilidad que tres segmentos seleccionados aleatoriamente de los cinco formen un triángulo.

27. De una urna que contiene 12 bolas , de las cuales 8 son blancas y 4 negras, se extraé una muestra de tamaño 4 con reemplazo (sin reemplazo) ¿Cuál es la probabilidad que la muestra contenga exactamente tres bolas blancas?

28. Nueve pasajeros abordan un tren de tres carretas. Cada pasajero escoge - aleatoriamente el carruaje para sentarse. ¿Cuál es la probabilidad que:

(a) haya tres personas en el primer carruaje?

(b) haya tres personas en cada carruaje?

(c) haya dos personas en un carro, tres en el otro, y cuatro en el carro restante?

29. El evento C tiene el doble de posibilidad que el evento A; el evento B - tiene igual posibilidad que la suma de posibilidades de A y C. Los eventos son mutuamente excluyentes y uno de ellos debe ocurrir. Hallar la probabilidad de cada uno de los eventos.

30. Un grupo de personas está formado por 6 hombres y 8 mujeres. Se desea for

mar una comisión integrada por cuatro delegados con igual representatividad; calcular.

(a) La probabilidad que la comisión sea mixta.

(b) La probabilidad que la comisión esté integrada por 3 hombres y una mujer.

31. Suponga que se ha cargado un dado de manera que la probabilidad que ocurra un número determinado es proporcional al mismo. Calcular la probabilidad que ocurra un número mayor que 4.

32. Suponga que se tiene un dado cargado de tal forma que la probabilidad del número que salga sea inversamente proporcional al mismo. Calcular la probabilidad de la ocurrencia (a) un número par (b) un número mayor que 4.

33. Suponga que ocho jugadores que tienen la misma capacidad participan en un torneo de eliminación sencilla (no se permiten los empates). ¿Cuál es la probabilidad de que cada uno sea el ganador del torneo?. ¿Cuál es la probabilidad que el jugador 1 gane sus primeros dos fuegos y pierde la final?

34. En la sección de control de calidad de una compañía, se encontró 5 artículos defectuosos, en una partida de 100 artículos tomados aleatoriamente de la producción de un día. Estime la probabilidad de producir un artículo defectuoso.

35. Cierto tipo de motor eléctrico falla por obstrucción de los cojinetes, por combustión del embobinado o por desgaste de las escobillas. Suponga que la probabilidad de obstrucción es el doble de la de combustión, la cual es cuatro veces más probable que la inutilización de las escobillas, ¿Cuál es la probabilidad que la falla sea por cada uno de los tres mecanismos?

36. En una zona de parqueo hay 10 lugares en fila. Una persona deja estacionado su carro en uno de estos lugares, que no es ninguno de los estremos. Al regresar encuentra que hay 4 carros estacionados (incluyendo el suyo). Calcular la probabilidad que los dos lugares vecinos al que ocupa su carro estén desocupados.

37. Cada uno de los coeficientes de la ecuación $ax^2 + bx + c = 0$ se determina por medio de una tirada de un dado prodinario. Calcular la probabili-

dad que la ecuación tenga (a) raíces reales, (b) raíces racionales.

38. Una urna contiene 5 bolas blancas, 4 rojas y 3 negras. Otra contiene 5 blancas, 6 rojas y 7 negras. Se elige una bola de cada urna. ¿Cuál es la probabilidad que todas sean del mismo color?

39. Ocho ejecutivos de una empresa llegan a su oficina diariamente en su automóvil y lo aparcan en uno de los tres aparcamientos. Si estos son escogidos al azar, ¿cuál es la probabilidad que en un día determinado se tenga 5 de los 8 automóviles en un aparcamiento, 2 en otro y 1 en otro?

40. Un fabricante de cereales desea cambiar el diseño de la caja de uno de sus productos, por lo que se muestra individualmente a cada una de 6 personas la caja anterior y la nueva, y se le pide que indique su preferencia. Suponiendo que cada uno de las personas no tenga verdadera preferencia por ninguna, ¿Cuál es la probabilidad que por lo menos 5 de las 6 personas prefieran el nuevo diseño?

41. Supongase que una persona está ubicada en el origen de un sistema de coordenadas cartesianas en el plano. Lanza una moneda.En cada lanzamiento, si obtiene cara avanza una unidad hacia arriba; si obtiene sello avanza una unidad hacia la derecha. Determine ud. la probabilidad que al cabo de 4 lanzamientos se encuentre en el punto (2,2).

42. Jaimito tiene 8 bolas blancas y 2 negras, las alinea al azar. ¿Cuál es la probabilidad que las 2 negras queden juntas? ¿De que las 2 negras ocupen posiciones de los extremos?

43. Se ordenan aleatoriamente 4 personas en un círculo, ¿Cuál es la probabilidad que 2 personas dadas estén contiguas?

44. Cuatro objetos se distribuyen al azar entre seis recipientes. ¿Cuál es la probabilidad que

(a) todos los objetos esten en el mismo recipientes?

(b) no haya dos objetos en un mismo recipiente?

45. El gerente regional de mercadeo de ELECTRONIC CALCULATORS está tratando de estimar su proyección de venta para el próximo año. El ha limitado sus estimados a 200,000, 250,000, 300,000, 350,000 ó 400,000 calculadoras. Más adelante estableció que estaba completamente indeciso entre la venta de 350,000 y 300,000 y que no podía decidir cuál era más probable. Sin em

bargo, cree que unas ventas de 350,000 son dos veces más probables que - 400,000 y que unas ventas de 300,000 son cuatro veces más probables que 200,000. Finalmente decidió que unas ventas de 250,000 son sólo un 50 % - más probables que las de 350,000.

(a) ¿Cuál es la probabilidad de vender 300,000 ó 350,000 calculadoras?

(b) ¿Cuál es la probabilidad de vender más de 400,000 o menos de 200,000 calculadoras?

46. Durante un período específico, el 80 % de las acciones ordinarias de una industria que incluye sólo 10 compañías han aumentado en valor comercial. Si un inversionista escoge aleatoriamente tres de esas acciones. Determine la probabilidad que

(a) Sólo una de las tres acciones aumente su cotización.

(b) Sólo dos acciones aumenten su cotización.

(c) Por lo menos dos acciones aumenten su cotización.

47. A es uno de los 6 caballos que van ha competir en una carrera, y lo va ha montar uno de los Jockeys B ó C. Hay 2 a 1 de que B monte a A, en cuyo caso todos los caballos tienen iguales probabilidades de ganar; si C monta a A su probabilidad de ganar se triplica. ¿Cuales son las apuestas en contra de la victoria de A?

48. En una urna hay 2 bolas azules, una blanca y 3 rojas. Se van a extraer al azar 2 bolas. Calcule ud. la probabilidad que las dos bolas sean rojas o una blanca y la otra azul.

49. Ocho personas compiten por un puesto público. Los cuatro primeros candidatos tienen la misma oportunidad de ganar, el quinto candidato tiene el doble oportunidad que los candidatos anteriores y los tres últimos tienen - el triple de oportunidad que el quinto candidato. ¿Cuál es la probabilidad que gane el quinto candidato?

50. Un edificio consta de 7 pisos con 4 departementos por piso. Determine la - probabilidad que dos jefes de familia, elegidos al azar, pertenescan a departamentos que por lo menos estén separados por 2 pisos.

51. Se lanzan 9 bolas en tres cajas inicialmente vacías. Cada una de las bolas tienen la misma probabilidad de caer en cualquiera de las cajas. Determinar la probabilidad que:

(a) hayan tres bolas en cada caja.

(b) hayan cuatro bolas en la primera caja, tres en la segunda, y dos en la tercera.

52. Un muchacho parado en una esquina lanza una moneda. Si cae cara, camina una cuadra al Este; si cae sello, camina una cuadra al Oeste. En cada esquina repite la operación. ¿Cuál es la probabilidad que después de 6 lanzamientos esté:

(a) en el punto de partida? (b) a dos cuadras del punto de partida?
(c) a cuatro cuadras del punto de partida?

53. De una baraja de 52 cartas, se extraen aleatoriamente 3 cartas, ¿Cuál es la probabilidad que sean del mismo palo?

1.6 AXIOMAS DE PROBABILIDAD Y PROPIEDADES

Independientemente de la forma como definimos probabilidad, esta cumple los axiomas siguientes, que son consecuencia inmediata de la definición (ver observaciones a la definición clásica y por frecuencia relativa)

Ax.1 $0 \leq P[A] \leq 1$, para cada evento A en (Ω)

Ax.2 $P[\Omega] = 1$

Ax.3 Para cualquier número finito k de eventos mutuamente excluyentes en Ω es

$$P\left[\bigcup_{i=1}^{k} A_i\right] = \sum_{i=1}^{k} P[A_i]$$

Una consecuencia inmediata de **Ax.3** es, si A y B son dos eventos mutuamente excluyentes en Ω, entonces

$$P[A \cup B] = P[A] + P[B]$$

Como todo axioma, *Ax.*1, *Ax.*2 y *Ax.*3 no se demuestran, sin embargo el lector habrá notado que estos son acordes con la definición de probabilidad como hemos hecho notar en las notas, después de cada definición.

En teoría más avanzada de probabilidad, en vez del tercer axioma se usa el siguiente

(A'x3) Si $A_1, A_2, A_3, \ldots$

es una secuencia numerable de eventos mutuamente excluyentes definidos en Ω, entonces

$$P[A_1 \cup A_2 \cup \ldots] = P[A_1] + P[A_2] + \ldots \quad \text{ó}$$

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$$

Debido a que la definición de probabilidad cumple los axiomas Ax.1, Ax.2 y Ax.3, se da una definición de probabilidad conocida como probabilidad axiomática o abstracta.

DEFINICION 1.6.1 Sea Ω un espacio muestral asociado a un experimento ε. La probabilidad P, es una función que asigna a cada evento $A(A \in \mathcal{P}(\Omega))$, un número P[A], llamado la *probabilidad del evento* A, tal que cumple los axiomas : Ax.1, Ax.2, Ax.3 .

NOTA En teoría mas avanzada de probabilidad, se considera una clase de subconjuntos de Ω, (no necesariamente todo $\mathcal{P}(\Omega)$), que cumplen ciertos axiomas, llamados σ-algebra

Los teoremas siguientes son consecuencia inmediata de los axiomas.

TEOREMA 1.6.1 Si ϕ es el evento imposible, entonces $P[\phi] = 0$

DEMOSTRACION Nóte que $\Omega = \Omega \cup \phi$

Ω y ϕ son mutuamente excluyentes, por lo tanto

$$P[\Omega] = P[\Omega] + P[\phi], \qquad \text{por Ax.3}$$
$$1 = 1 + P[\phi] \qquad \text{por Ax.2}$$

de donde $P[\phi] = 0.$

TEOREMA 1.6.2 Para cada evento A, se cumple que

$$P[\bar{A}] = 1 - P[A] \quad \text{ó} \quad P[A] = 1 - P[\bar{A}]$$

DEMOSTRACION Tenemos que $\Omega = A \cup \bar{A}$ y $A \cap \bar{A} = \phi$, es decir los eventos A y $\bar{A}$ son mutuamente excluyentes. Luego

$$P[\Omega] = P[A] + P[\bar{A}], \qquad \text{por Ax.3}$$
$$1 = P[A] + P[\bar{A}] \qquad \text{por Ax.2}$$

por lo tanto $P[\bar{A}] = 1 - P[A]$ ó $P[A] = 1 - P[\bar{A}]$.

TEOREMA 1.6.3 Si A y B son eventos tales que $A \subset B$, entonces

$$P[A] \leqslant P[B]$$

DEMOSTRACION Nótese que $B = A \cup (B \cap \bar{A})$ (ver figura 1.6.1) y $\quad A \cap (B \cap \bar{A}) = \phi$, es decir los eventos A y $B \cap \bar{A}$ son mutuamente excluyentes. Luego,

$$P[B] = P[A] + P[B \cap \bar{A}] \quad \text{por Ax.3}$$

Desde que $P[B \cap A] \geq 0$, se tiene que

$$P[B] \geq P[\bar{A}]$$

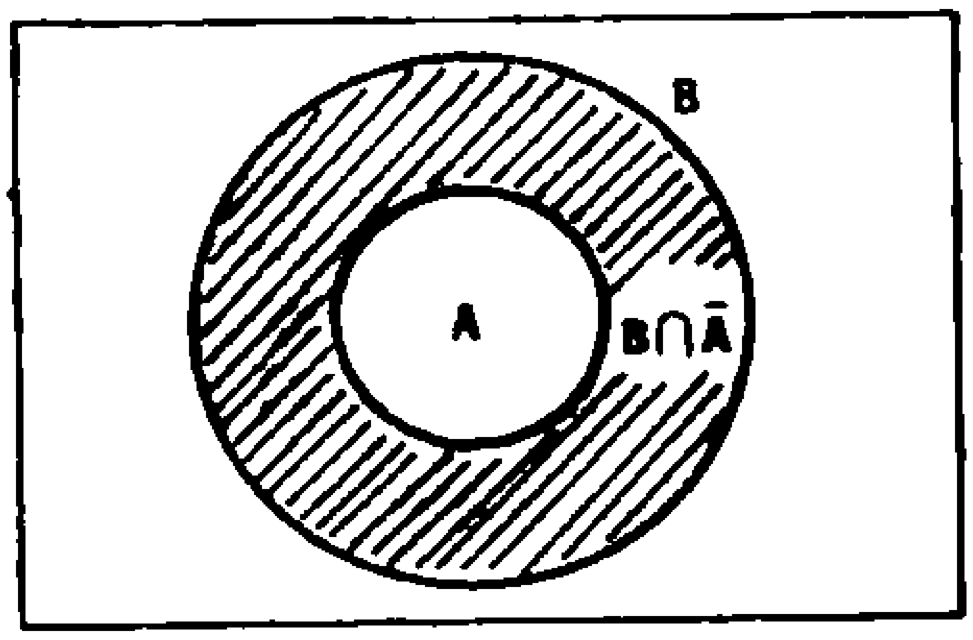


Fig. 1.6.1

TEOREMA 1.6.4 Si A y B son dos eventos cualesquiera en Ω entonces

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

DEMOSTRACION El evento $A \cup B$ puede representarse como la unión de los eventos A y $\bar{A} \cap B$, mutuamente excluyentes (ver fig. 1.6.2)

$$A \cup B = A \cup (\bar{A} \cap B)$$

Luego, $\quad P[A \cup B] = P[A] + P[\bar{A} \cap B] \quad$ por $\quad$ Ax.3 $\qquad$ (1)

El evento B tambien puede escribirse como la unión de los eventos mutuamente excluyentes $A \cap B$ y $\bar{A} \cap B$

$$B = (A \cap B) \cup (\bar{A} \cap B) \quad \text{y}$$

$$P[B] = P[A \cap B] + P[\bar{A} \cap B] \quad \text{por Ax.3}$$

ó $\quad P[\bar{A} \cap B] = P[B] - P[A \cap B]$

Sustituyendo este resultado en la ecuación (1), obtenemos

$$P[A \cup B] = P[A] + P[B] - P[A \cap B].$$

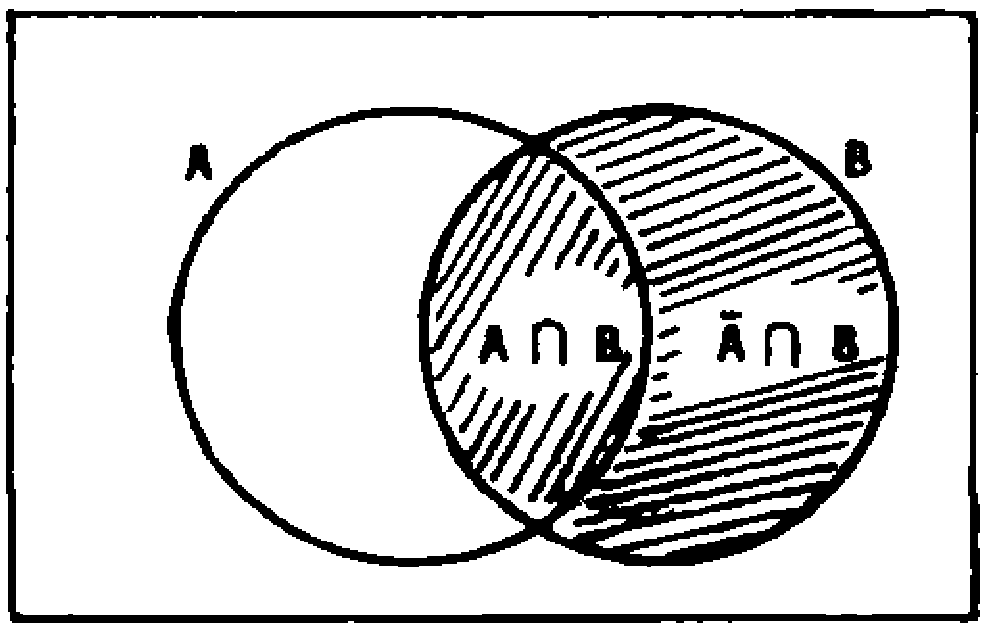


Fig. 1.6.2

CONSECUENCIA Una consecuencia importante del teorema 4, es la siguiente

$$P[A \cup B] \leq P[A] + P[B] \ , \qquad \text{ya que} \quad P[A \cap B] \geq 0$$

OBSERVACION En la demostración del teorema 4 se ha probado que

$$P[B] = P[AB] + P[\bar{A}B] \ \text{y} \ P[B - A] = P[\bar{A}B] = P[B] - P[AB]$$

TEOREMA 1.6.5 Si A, B y C son tres eventos cualesquiera en Ω, entonces

$$P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[C \cap B] + P[A \cap B \cap C]$$

DEMOSTRACION Podemos escribir $A \cup B \cup C = (A \cup B) \cup C$ y aplicar el teorema 1.6.4, desde que $A \cup B$ es un evento.

TEOREMA 1.6.6 Si $A_1, A_2, \ldots, A_k$ es una colección de eventos cualesquiera en Ω, entonces

$$P[A_1 \cup A_2 \cup \ldots \cup A_k] = \sum_{i=1}^{k} P[A_i] - \sum_{i<j=2}^{k} P[A_i \cap A_j] + \sum_{i<j<r=3}^{k} P[A_i \cap A_j \cap A_r]$$

$$+ \ldots + (-1)^{k+1} P[A_1 \cap A_2 \cap \ldots \cap A_k]$$

NOTA El lector puede probar por inducción.

EJEMPLO 1 La probabilidad que llueva en Huancayo el 12 de Octubre es 0.10 de que truene es 0.05 y que llueve y truene es 0.03. ¿Cuál es la probabilidad que llueve o truene ese día?

SOLUCION Definimos los siguientes eventos:

A : "llueva en Huancayo el 12 de Octubre"

B : "truene el 12 de Octubre" C : "llueva o truene ese día"

Entonces $P[A] = 0.10$, $P[B] = 0.05$ y $P[A \cap B] = 0.03$

El evento **C** se escribe $C = A \cup B$, y

$$P[C] = P[A \cup B] = P[A] + P[B] - P[A \cap B] \quad \text{teorema 1.6.4}$$
$$= 0.10 + 0.05 - 0.03 = 0.12.$$

EJEMPLO 2 La probabilidad que la señora hablantina reciba a lo más 5 llamadas telefónicas en un día es 0.20; y por lo menos 9 llamadas en un día es 0.50. ¿Cuál es la probabilidad que la señora hablantina reciba 6,7 u 8 llamadas en un día?

SOLUCION El espacio muestral Ω es $\{0,1,2,3, \ldots\}$ llamadas telefónicas en un día.

Sean los eventos:

A : "reciba a lo más 5 llamadas". Es decir **A** consta de 0,1,2,3,4,5 llamadas telefónicas en un día.

B : "Reciba por lo menos 9 llamadas". Es decir, **B** consta de 9,10,11,12, ... llamadas.

C : "reciba 6,7 u 8 llamadas". Es decir, $C = \{6,7,8\}$.

Además $P[A] = 0.20$, $P[B] = 0.50$

Los eventos A,B y C son mutuamente excluyentes y colectivamente exaustivos, entonces $\Omega = A \cup B \cup C$, y

$$P[\Omega] = P[A \cup B \cup C]$$

$$1 = P[A] + P[B] + P[C] \text{ , } \qquad \text{por Ax.3.}$$

$$= 0.20 + 0.50 + P[C]$$

De donde $P[C] = 0.30$.

EJEMPLO 3 Una caja contiene 100 tubos de televisión. La probabilidad que haya al menos un tubo defectuoso es 0.05 y de que haya al menos dos tubos defectuosos es 0.01. ¿Cuál es la probabilidad que la caja contenga,
(a) ningún tubo defectuoso? (b) exactamente un tubo defectuoso?
(c) a lo más un tubo defectuoso?

SOLUCION El espacio muestral Ω es $\{0,1,2, \ldots ,100\}$ tubos defectuosos.
Sean los eventos:
A : "haya al menos un tubo defectuoso ". Es decir,

$$A = \{1,2, \ldots , 100\}$$

B : "haya al menos dos tubos defectuosos". Es decir,

$$B = \{2,3, \ldots , 100\}$$

C : "ninguno de los tubos es defectuoso". Es decir

$$C = \{0\}$$

E : "exactamente 1 tubo defectuoso contiene la caja".

$$E = \{1\}$$

F : "hay a lo más un tubo defectuoso"

$$F = \{0,1\}$$

Ademas sabemos, $P[A] = 0.05$ y $P[B] = 0.01$

(a) Los eventos A y C son mutuamente excluyentes y colectivamente exhaustivo. Es decir $\Omega = A \cup C$. Por lo tanto,

$$P[C] = 1 - P[A] = 1 - 0.05 = 0.95 \text{ , por teorema 1.6.2}$$

(b) El evento **A** se puede escribir así, $A = E \cup B$, además E y B son eventos mutuamente excluyentes, entonces

$$P[E] = P[A] - P[B] = 0.05 - 0.01 = 0.04 \text{ .}$$

(c) El evento $F = C \cup E$, y además C y E son mutuamente excluyentes, entonces

$$P[F] = P[C] + P[E] = 0.95 + 0.04 = 0.99.$$

EJEMPLO 4 Una cadena de supermercados opera con 250 tiendas en toda la nación. Estas tiendas están ubicadas en ciudades de diferentes poblaciones como se indica en la tabla 2.2. Una tienda se selecciona aleatoriamente para hacer una encuesta sobre un producto nuevo importante. ¿Cuál es la probabilidad que la tienda esté ubicada en una ciudad con población que excede a los 200,000 personas? ¿Cuál es la probabilidad que la tienda esté ubicada en una ciudad con una población de 200,000 ó menos habitantes? ¿Cuál es la probabilidad que la tienda seleccionada aleatoriamente es una tienda de clase A o está ubicada en una ciudad con una población sobre los 350,000 habitantes?

Tabla 2.2 Tiendas propiedad de la cadena de supermercados ubicación y clases de tiendas.

Población de la ciudad de la cual está ubicado la tienda	*Clases de Tiendas*		*Número Total de Tiendas*
	Clase A	*Clase B*	
debajo de 50,000	37	35	72
50,000 a 100,000	44	20	64
100,001 a 200,000	30	24	54
200,001 a 350,000	24	16	40
sobre los 350,000	15	5	20
Número Total de tiendas	150	100	250

SOLUCION Sean los siguientes sucesos:

ω_1 : "la tienda está ubicada en una ciudad con población debajo de 50,000 habitantes".

ω_2 : "la tienda está ubicada en una ciudad con población de 50,000 a 100,000".

ω_3 : "la tienda está ubicada en una ciudad con población de 100,001 a 200,000".

ω_4 : "la tienda está ubicada en una ciudad con población de 200,001 a 350,000".

ω_5 : "la tienda está ubicada en una ciudad con población sobre los 350,000 habitantes".

(a) Sea A, el evento: "la tienda está ubicada en una ciudad con una población que excede a los 200,000", entonces

$$A = \{\omega_4 , \omega_5\} \qquad y$$

$$P[A] = P[\{\omega_4\}] + P[\{\omega_5\}] = \frac{40}{250} + \frac{20}{250} = \frac{60}{250} = \frac{6}{25} .$$

(b) Sea B, el evento: "la tienda está ubicada en una ciudad con una población de 200,000 o menos", entonces

$$B = \{\omega_1, \omega_2, \omega_3\} \qquad y$$

$$P[B] = P[\{\omega_1\}] + P[\{\omega_2\}] + P[\{\omega_3\}] = \frac{72}{250} + \frac{64}{250} + \frac{54}{250} = \frac{190}{250} = \frac{19}{25}$$

(c) Sea C, el evento: "la tienda es de clase A", entonces

$$P[C] = \frac{150}{250} = \frac{15}{25} .$$

Sea D, el evento: "la tienda está ubicada en una ciudad con población por encima de los 350,000 habitantes". Entonces

$$P[D] = \frac{20}{250} = \frac{2}{25} .$$

Se pide calcular

$$P[C \text{ o } D] = P[C \cup D] = P[C] + P[D] - P[C \cap D]$$

pués C y D no son mutuamente excluyentes. Y $P[C \cap D] = \frac{15}{250}$

por lo tanto $P[C \cup D] = \frac{150}{250} + \frac{20}{250} - \frac{15}{250} = \frac{155}{250} = \frac{31}{50}$.

EJEMPLO 5 De un grupo de personas, el 30% práctica fútbol y el 40% juega ajedrez. De los futbolístas el 50% juega ejedrez. Si se elige aleatoriamente una persona. ¿Cuál es la probabilidad que

(a) juega fútbol o ajedrez?

(b) practica sólo uno de estos deportes?

(c) no practique ni fútbol ni ajedrez?

SOLUCION (a) definimos los eventos:

A: "la persona elegida es fútbolista".

B: "la persona elegida juega ajedrez".

La probabilidad pedida es la de A ∪ B. Antes recordemos una observación respecto a la probabilidad y porcentaje; que están relacionados por una factor 100; es decir, por, ejemplo el hecho de que el 30% de las personas practican fútbol, entonces al elegir una persona de este grupo, la probabilidad de que sea futbolista es 0.3, además como el 50% de los futbolistas, o sea el 15% del total juega fútbol y ajedrez, tenemos

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] = 0.3 + 0.4 - 0.15 = 0.55$$

(b) Sea C, el evento: "la persona elegida practica sólo uno de estos deportes".

El evento C se puede escribir así,

$$C = (A \cap \bar{B}) \cup (B \cap \bar{A})$$

Los eventos $A \cap \bar{B}$ y $B \cap \bar{A}$ son mutuamente excluyentes, entonces

$P[C] = P[A \cap \bar{B}] + P[B \cap \bar{A}]$ por $Ax.3$

$= 0.15 + 0.25 = 0.40.$

Nótese que $C = A \Delta B$

(c) Sea D, el evento: "la persona elegida no practica ni fútbol ni ajedrez".
El evento D se escribe $D = \overline{A \cup B}$.
Luego

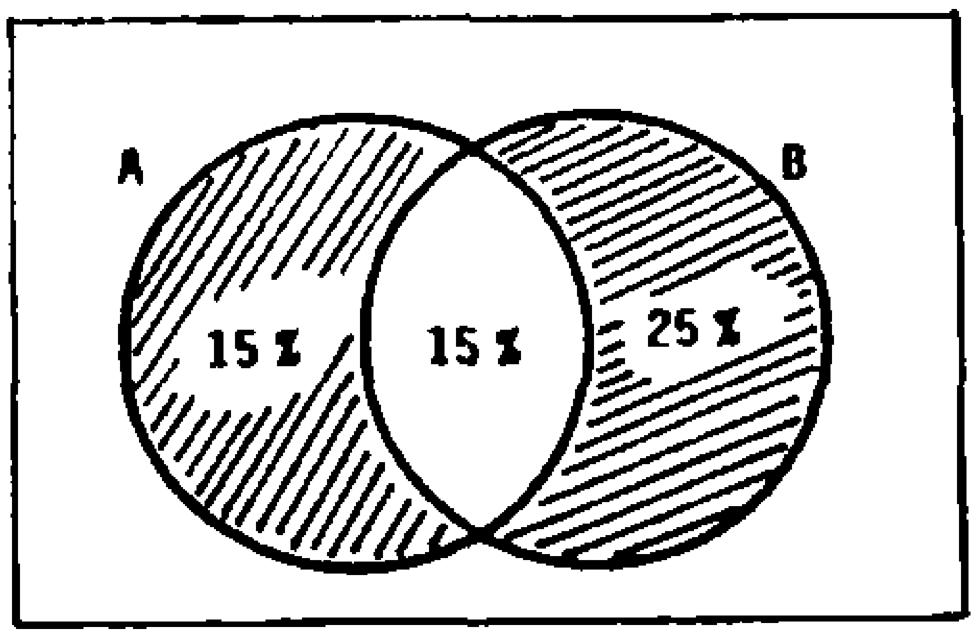


Fig. 1.6.3

$$P[D] = P[\overline{A \cup B}]$$
$$= 1 - P[A \cup B] \qquad \text{teorema 1.6.2}$$
$$= 1 - 0.55 = 0.45$$

EJEMPLO 6 Sean A y B dos eventos que no son mutuamente excluyentes tal que

$$P[A] = 0.20 \quad , \quad P[B] = 0.30 \quad y \quad P[A \cap B] = 0.10 \; .$$

Calcular: (a) $P[\bar{A} \cap \bar{B}]$; (b) $P[\bar{A} B]$; (c) $P[\bar{B}A]$; (d) $P[\bar{A} \cup B]$

SOLUCION

(a) $P[\bar{A} \cap \bar{B}] = P[\overline{A \cup B}]$, ley de De Morgan

$= 1 - P[A \cup B]$, teorema 1.6.2

$= 1 - [P[A] + P[B] - P[A \cap B]]$, teorema 1.6.4

$= 1 - 0.20 + 0.30 - 0.10 = 0.60.$

(b) $P[\bar{A} B] = P[B] - P[AB]$ Consecuencia de teorema 1.6.4

$= 0.30 - 0.10 = 0.20$

(c) $P[A\bar{B}] = P[A] - P[AB]$ Consecuencia del teorema 1.6.4

$= 0.20 - 0.10 = 0.10$.

(d) $P[\bar{A} \cup B] = P[\bar{A}] + P[B] - P[\bar{A}B]$ teorema 1.6.4

$= 0.80 + 0.30 - 0.20 = 0.90$.

EJEMPLO 7 De una urna que contiene 3 bolas blancas, 4 bolas rojas y 3 bolas verdes, se extraen al azar 2 bolas. ¿Cuál es la probabilidad de que ninguna sea verde o sean de distinto color?

SOLUCION El experimento aleatorio es "extraer dos bolas al azar"
El número de elementos del espacio muestral es $C_{10}^{2} = 45$
Definimos los siguientes eventos:

A: "de las 2 bolas extraídas ninguna sea verde o sean de distinto color".
B: "ninguna sea verde". C: "sean de distinto color".
El evento A se escribe $A = B \cup C$, y $B \cap C \neq \phi$. Luego

$$P[A] = P[B] + P[C] - P[B \cap C]$$

Cálculo de $P[B]$.

$$N(B) = C_7^2 = 21$$

$$P[B] = \frac{21}{45} = \frac{7}{15} \quad (1)$$

Cálculo de $P[C]$ El evento C se puede escribir así $C = C_1 \cup C_2 \cup C_3$, donde:
C_1: "una es blanca y la otra roja". C_2: "una es blanca y la otra verde".
C_3: "una es roja y la otra verde".
C_1, C_2, C_3 son mutuamente excluyentes dos a dos, entonces

$$P[C] = P[C_1] + P[C_2] + P[C_3]$$

$N(C_1) = 3 \times 4 = 12$, luego $P[C_1] = \frac{12}{45} = \frac{4}{15}$;

$N(C_2) = 3 \times 3 = 9$, luego $P[C_2] = \frac{9}{45} = \frac{3}{15}$;

$N(C_3) = 4 \times 3 = 12$, luego $P[C_3] = \frac{12}{45} = \frac{4}{15}$.

Por lo tanto,

$$P[C] = \frac{4}{15} + \frac{3}{15} + \frac{4}{15} = \frac{11}{15} . \qquad (2)$$

Calculemos ahora $P[B \cap C]$

Obseve que $B \cap C = C_1$, luego $P[B \cap C] = P[C_1] = \frac{4}{15}$ (3)

Finalmente, de (1) , (2) y (3) se tiene

$$P[A] = P[B] + P[C] - P[B \cap C]$$

$$= \frac{7}{15} + \frac{11}{15} - \frac{4}{15} = \frac{14}{15} .$$

Otra forma de resolver el problema anterior.

$A = B \cup C$, entonces $\bar{A} = \overline{B \cup C} = \bar{B}\,\bar{C}$, es decir

$\bar{A}$: "alguna sea verde y sean de igual color"

$$N(\bar{A}) = C_3^2 = 3$$

$$P[\bar{A}] = \frac{3}{45} = \frac{1}{15}$$

Por lo tanto ,

$$P[A] = 1 - P[\bar{A}] = 1 - \frac{1}{15} = \frac{14}{15} .$$

EJEMPLO 8 Se tiene en una caja 5 tikets de 100 intis cada uno, 3 tikets de 300 intis cada uno, y 2 tikets que valen 500 intis cada uno. Se escoge aleatoriamente 3 tikets. Determinar la probabilidad que:

(a) Al menos dos de ellas tenga el mismo precio.

(b) La suma de los precios de los tres tikets sea de 700 intis.

SOLUCION El experimento aleatorio es, escoger 3 tikets de los 10 que se tiene en la caja; entonces al espacio muestral C_{10}^3 elementos.

Sean los iguientes eventos:

A: "al menos dos de los tikets escogidos tengan el mismo precio".

B: "la suma de los precios de los tres tikets sea de 700 intis

(a) Calcularemos primero la probabilidad del evento $\bar{A}$, donde

$\bar{A}$: "tengan precio diferentes".

entonces, el número de elementos de $\bar{A}$, es $C_5^1\, C_3^1\, C_2^1$. Luego,

$$P[\bar{A}] = \frac{C_5^1\, C_3^1\, C_2^1}{C_{10}^3}$$

y $$P[A] = 1 - \frac{C_5^1\, C_3^1\, C_2^1}{C_{10}^3} = 0.75 .$$

(b) Sea los siguientes eventos:

B_1 : "un tikets de 500 intis y 2 de 100 intis".

B_2 : "2 tikets de 300 intis y 1 de 100 intis".

Entonces, el evento B se escribe $B = B_1 \cup B_2$, donde B_1 y B_2 son mutuamente excluyentes; por lo tanto,

$$P[B] = P[B_1] + P[B_2] .$$

El evento B_1 ocurre de $C_2^1\ C_5^2$ formas; luego ,

$$P[B_1] = \frac{C_2^1\, C_5^2}{C_{10}^3}$$

El evento B_2 ocurre de $C_3^2\, C_5^1$ formas, luego

$$P[B_2] = \frac{C_3^2\, C_5^1}{C_{10}^3} .$$

Finalmente,

$$P[B] = \frac{C_2^1\, C_5^1}{C_{10}^3} + \frac{C_3^2\, C_5^1}{C_{10}^3} = \frac{C_2^1\, C_5^2 + C_3^2\, C_5^1}{C_{10}^3} = \frac{7}{24} .$$

EJEMPLO 9 En una encuesta pública se determina que la probabilidad que una persona consuma el producto A es 0.50, que consuma el producto B es 0.37 que consuma el producto C es 0.30, que consuma A y B es 0.12, que consuma solamente A y C es 0.08, que consuma solamente B y C es 0.05 y que consuma solemente C es 0.15. Calcular la probabilidad que una persona consuma:

(a) A ó B pero no C.

(b) Solamente A.

SOLUCION Los datos del problema son los siguientes:

$$P[A] = 0.50; \quad P[B] = 0.37; \quad P[C] = 0.30; \quad P[AB] = 0.12 .$$

El evento solamente A y C, se escribe $AC\bar{B}$;

luego $P[AC\bar{B}] = 0.08$.

El evento solamente B y C, se escribe $\bar{A}BC$;

luego $P[\bar{A}BC] = 0.05$.

Y el evento solamente C, se escribe $\bar{A}\bar{B}C$;

luego $P[\bar{A}\bar{B}C] = 0.15$.

(a) Se pide calcular la probabilidad del evento $(A \cup B)\bar{C}$.

Observe que

$$P[(A \cup B)\bar{C}] = 1 - P[\bar{A}\bar{B} \cup C] . \qquad (1)$$

Por otro lado,

$$P[\bar{A}\bar{B} \cup C] = P[\bar{A}\bar{B}] + P[C] - P[\bar{A}\bar{B}C] . \qquad (2)$$

Cálculo de $P[\bar{A}\bar{B}]$

$$P[\bar{A}\bar{B}] = P[\overline{A \cup B}] = 1 - P[A \cup B] \qquad (3)$$

pero, $P[A \cup B] = P[A] + P[B] - P[AB]$

$$= 0.50 + 0.37 - 0.12$$

$$= 0.75 .$$

Reemplazando este valor en (3), obtenemos

$$P[\bar{A}\bar{B}] = 1 - 0.75 = 0.25$$

Reemplazando este último valor en (2), obtenemos

$$P[\bar{A}\bar{B} \cup C] = 0.25 + 0.30 - 0.15 = 0.40$$

Finalmente reemplazando este valor en (1) obtenemos la probabilidad pedida

$$P[(A \cup B)\bar{C}] = 1 - 0.40 = 0.60$$

(b) El evento solamente A, se escribe $A\bar{B}\bar{C}$. Note que el evento A puede escribirse

$A = AB \cup AC\bar{B} \cup A\bar{B}\bar{C}$, mutuamente excluyentes. Luego,

$$P[A] = P[AB] + P[AC\bar{B}] + P[A\bar{B}\bar{C}] ,$$

de donde

$$P[A\bar{B}\,\bar{C}] = P[A] - P[AB] - P[AC\bar{B}]$$
$$= 0.50 - 0.12 - 0.08$$
$$= 0.30\ .$$

NOTA Una forma práctica de resolver este problema es llevando los datos a un diagrama de Venn, como se observa en la fig. 1.6.4. Además, observe que las probabilidades indicadas en el diagrama corresponden a eventos mutuamente excluyentes. Luego,

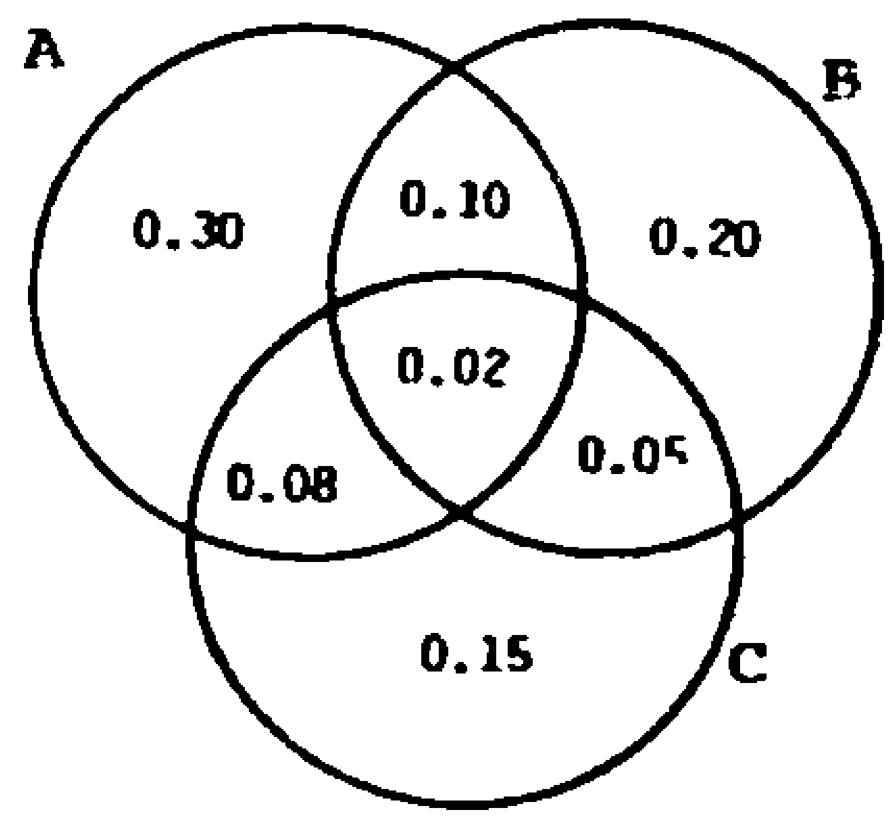


Fig. 1.6.4

(a) $P[(A \cup B)\bar{C}] = 0.30 + 0.10 + 0.20 = 0.60$

(b) $P[A\bar{B}\,\bar{C}] = 0.3$.

EJEMPLO 10 El cuadro siguiente contiene la clasificación de 321 obreros de un sindicato respecto a dos características:

1. El número de años de pertenencia de cada uno al sindicato ;
2. Su respuesta a la pregunta: "Desea ud. ir a la huelga para obtener un aumento de salarios".

Respuesta a la pregunta	Número de años en el sindicato				Total
	menos de 1	de 1 a 3	de 4 a 10	más de 10	
Si	27	54	137	28	246
No	14	18	34	3	69
No sé	3	2	1	0	6
Total	44	74	172	31	321 \ 321

Sean los eventos :

S : "obreros que contestaron si".
N : "obreros que contestaron no".
A : " obreros que pertenecen al sindicato menos de 1 año".
B : "obreros con 1 a 3 años en el sindicato".
C : "obreros con 4 a 10 años en el sindicato".

(a) Encuentre la probabilidad de cada uno de los siguientes eventos:

$$S \cap B;\quad S \cup B,\quad \overline{S \cup N} \cap A;\quad \overline{N \cap C}\,.$$

(b) Evalué la probabilidad de los siguientes eventos:

i) "Obreros que contestaron si y pertenecen por lo menos cuatro años al sindicato".

ii) "Obreros que contestaron si o no sé, y tienen más de 10 años en la empresa".

SOLUCION El espacio muestral tiene 321 elementos.

(a) i) El número de casos favorables al evento $S \cap B$ se obtiene de la tabla y es el número que aparece en la intersección de la fila que contestaron si, con la columna de obreros con 1 a 3 años en el sindicato. Este número es 54, luego,

$$P[S \cap B] = \frac{54}{321} = \frac{18}{107}\,.$$

ii) $P[S \cup B] = P[S] + P[B] - P[S \cap B]$ Teorema 1.6.4

$$= \frac{246}{321} + \frac{74}{321} - \frac{54}{321} = \frac{266}{321}\,.$$

iii) $P[\overline{S \cup N} \cap A] = P[\bar{S} \cap \bar{N} \cap A] = \dfrac{3}{321}\,.$

iv) $P[\overline{N \cap C}] = 1 - P[N \cap C] = 1 - \dfrac{34}{321} = \dfrac{287}{321}\,.$

(b) Definimos los siguientes eventos:

D: "obreros con más de 10 años en el sindicato".

E: "obreros que contestaron no sé".

i) Sea F: "obreros que contestaron si y pertenecen por lo menos cuatro años al sindicato".

Entonces $F = S \cap (C \cup D) = SC \cup SD)$;

los eventos SC y SD son mutuamente excluyentes, luego

$$P[F] = P[SC] + P[SD] = \frac{137}{321} + \frac{28}{321} = \frac{165}{321}\,.$$

ii) Sea G: "Obreros que contestaron si o no sé y tienen más de 10 años en el sindicato".

Entonces $G = (S \cup E) \cap D) = SD \cup ED$,
los eventos SD y ED son mutuamente excluyentes, luego

$$P[G] = P[SD] + P[ED] = \frac{28}{321} + \frac{0}{321} = \frac{28}{321} .$$

EJEMPLO 11 Una compañía que concierta citas por computadoras tiene en sus archivos los nombres y direcciones de 200 chicas. De estas 200, un total de 35 miden 1.65 mts. o menos de estatura; 60 son rubias; 12 de las rubias miden 1.65 mts. o menos. Pedro Carrillo envía su solicitud por correo, ¿Cuál es la probabilidad que
(a) reciba el nombre de una rubia?
(b) reciba el nombre de una chica rubia y estatura mayor de 1.65 mts?
(c) reciba el nombre de una rubia o estatura menor de 1.65 mts?
(d) reciba el nombre de una que no es rubia o estatura menor de 1.65 mts?

SOLUCION Construimos una tabla de dos entradas con los datos del problema.

Color del cabello \ Estatura	1.65 mts o menos	más de 1.65 mts.	Total
Rubia	12	48	60
no es Rubia	23	117	140
Total	35	165	200

Definimos los eventos:
R: "Recibe el nombre de una rubia".
E: "Estatura 1.65 mts o menos"

(a) $P[R] = \frac{60}{200} = \frac{3}{10}$.

(b) $$P[R\bar{E}] = \frac{48}{200} .$$

(c) $$P[R \cup E] = P[R] + P[E] - P[RE]$$
$$= \frac{60}{200} + \frac{35}{200} - \frac{12}{200} = \frac{83}{200} .$$

(d) $$P[\bar{R} \cup E] = P[\bar{R}] + P[E] - P[\bar{R}\, E]$$
$$= \frac{140}{200} + \frac{35}{200} - \frac{23}{200} = \frac{152}{200} .$$

PROBLEMAS 1.6

1. Sean A y B dos eventos en Ω tales que $P[A] = 0.2$, $P[\bar{B}] = 0.4$ y $P[\bar{A}\,\bar{B}] = 0.3$. Calcular

(a) $P[A \cup B]$; (b) $P[AB]$; (c) $P[A\bar{B}]$; (d) $P[\bar{A}B]$; (e) $P[\bar{A} \cup B]$

2. Dado $P[A] = 0.5$ y $P[A \cup B] = 0.6$. Determinar $P[B]$, si A y B son mutuamente excluyentes.

3. Si $P[A] = 0.4$, $P[B] = 0.5$, $P[C] = 0.7$, $P[AB] = 0.2$, $P[AC] = 0.2$, $P[BC] = 0.4$ y $P[ABC] = 0.1$. Hallar :

   (a) $P[A \cup B \cup C]$; (b) $P[A \cup B \cup \bar{C}]$

4. Si $P[A] = 0.4$, $P[B] = 0.5$ y $P[AB] = 0.3$. Hallar:

   $P[A \cup B]$, $P[A\bar{B}]$, $P[\bar{A}\bar{B}]$.

5. Si $P[A] = 0.2$, $P[\bar{A}B] = 0.4$, $P[\bar{A}C] = 0.2$, $P[\bar{A}BC] = 0.1$ y $P[BC] = 0.1$. Hallar :

   (a) $P[\bar{A}\bar{B}\bar{C}]$; (b) $P[A \cup B \cup C]$

6. Si los eventos A_j , $j = 1,2,3$ son tales que $A_1 \subset A_2 \subset A_3$ y $P[A_1] = 1/4$; $P[A_2] = 5/12$ y $P[A_3]' = 7/12$. Determinar la probabilidad de los siguientes eventos

   $\bar{A}_1 \cap A_2$; $\bar{A}_1 \cap A_3$; $\bar{A}_2 \cap A_3$; $A_1 \cap \bar{A}_2 \cap \bar{A}'_3$ y $\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3$.

7. Si $P[A] = \frac{1}{4}$ y $P[\bar{B}] = \frac{1}{5}$, ¿Son los eventos A y B mutuamente excluyentes? justifique su respuesta

8. ¿Si $P[A] = 1/3$ y $P[\bar{B}] = 1/4$, entonces A y B son mutuamente excluyentes?. Justifique su respuesta.

9. ¿Si $P[A] = P[\bar{B}]$, entonces $A = \bar{B}$? Justifique su respuesta.

10. ¿Cuáles de los siguientes representan tres eventos que son (1) colectivamente exhaustivos, (2) pares mutuamente excluyentes:

    (a) $P[A] = 0.6$, $P[B] = 0.2$, $P[C] = 0.1$ y $P[AB] = 0$

    (b) $P[A] = 0.1$, $P[B] = 0.4$, $P[C] = 0.5$, $P[A \cup B] = P[C]$, $P[A \cup C] = 0.6$ y $P[BC] = 0$

    (c) $P[A] = P[B] = 0.2$, $P[C] = 0.6$, $P[AB] = 0$ y

$$P[A \cup C] = P[B \cup C] = 0.8$$

(d) $P[A] = P[B] = P[C] = 0.35$ y $P[AB] = P[AC] = 0$

11. Demostrar que para cualquier par de eventos A y B en un espacio muestral se tiene que

$$P[A\bar{B} \cup B\bar{A}] = P[A] + P[B] - 2P[AB] \qquad (1)$$

donde el evento $A\bar{B} \cup B\bar{A}$ es el evento que ocurre exactamente uno de los eventos A y B. Compare (1) con la fórmula del teorema 1.6.4, de la cual podría decirse que es la fórmula de la probabilidad que ocurra por lo menos uno de los eventos.

12. Dado que $P[A] = 0.7$, $P[B] = 0.5$, $P[A\bar{B}] = 0.3$ y $P[A\bar{B}\bar{C}] = 0.1$.
Hallar :
(a) $P[A \cup \bar{A}B]$; (b) $P[\bar{A} \cup \bar{B}C]$; (c) $P[AB \cup \overline{A \cup \bar{A}\bar{B}}]$.

13. Si $P[B] = 0.1$, $P[C] = 0.2$, $P[BC] = 0.1$ y $P[A\bar{B}C] = 0.1$.
Hallar :
(a) $P[\bar{B}\,\overline{AC}]$; (b) $P[\bar{A}\bar{B}C \cup B\bar{C}]$.

14. Si $P[A] = 0.5$, $P[A \cup \overline{\bar{B}\bar{C}}] = 0.8$. Determinar $P[\bar{A}(B \cup C)]$.

15. Sean A,B, y C tres eventos en un espacio muestral. Exprese en términos de $P[A]$, $P[B]$, $P[C]$, $P[AB]$, $P[AC]$, $P[ABC]$, para $k = 0,1,2,3$ la probabilidad que :
(a) ocurran exactamente k de los eventos A,B,C;
(b) ocurran por lo menos k de los eventos A,B,C;
(c) ocurran cuando menos k de los eventos A,B,C.

16. A y B son dos eventos de un mismo espacio muestral. Las probabilidades $P[A]$, $P[B]$, y $P[AB]$ se dan. Encuéntrese una expresión en términos de estas probabilidades para las probabilidades de los siguientes eventos

(a) $\bar{A} \cup \bar{B}$; (b) $\bar{A}\bar{B}$; (c) $\bar{A} \cup B$, (d) $\bar{A}B$;

(e) $\overline{A \cup B}$; (f) $\overline{AB}$; (g) $\bar{A}(A \cup B)$; (h) $A \cup \bar{A}B$.

17. Se elige al azar un número entre los 200 primeros números enteros positivos. ¿Cúal es la probabilidad que el número elegido, sea divisible por 6 ó por 8?

18. Jaimito se presenta a dos universidades A y B. El estima la probabilidad que sea admitido en la universidad A en 0.8; a la universidad B en -

0.75, en al menos una de ellas en 0.95. ¿Cuál es la probabilidad que ingrese a ambas universidades.

19. Por cada 10,000 automóviles asegurados se roban 8,000 al año, se descomponen 2,500, y 6,300 de los autos robados resultan averiados. ¿Cuál es la probabilidad que un automóvil nuevo asegurado se pierda en el primer año? ¿Cuál es la probabilidad que lo roban o lo averien?

20. Un joyero produce 50,000 dijes en forma de corazón con motivos del "día de la madre". De los 50,000 dijes, 720 no están bien moldeados; 397 presentar rayaduras; 534 no tienen broche; 180 están rayadas y tienen defectos de moldura y 70 además de rayadas carecen de broche. ¿Cuál es la probabilidad de extraer un dije defectuoso de la caja en que están depositados todos?

21. Una caja contiene diez estampillas de 20 centavos, cinco estampillas de 15 centavos, y dos estampillas de 10 centavos. Se extrae aleatoriamente 6 estampillas; ¿Cuál es la probabilidad que su suma no exceda a 100 centavos?

22. En una urna hay 2 bolas azules, 1 blanca y 3 rojas. Se van a extraer al azar 2 bolas. Calcule ud. la probabilidad que las dos bolas sean rojas o una blanca y la otra azúl.

23. Un dado tiene 3 caras negras numeradas con 1,2,3; y las otras caras son blancas numeradas con: 4,5,6. Si se lanza este dado, ¿Cuál es la probabilidad de que aparezca un número par o una cara blanca?.

24. En una ciudad se publican tres revistas: A,B,C. Realizada una encuesta, se estima que de la población adulta: el 20% lee A, 30% lee B, 25% lee C; 10% lee A y B, 8% lee A y C y 12% lee B y C; además el 3% lee las tres revistas. Se elige aleatoriamente una persona adulta, calcular la probabilidad que lea al menos una de estas revistas.

25. Un banco tiene 50 cuentas de crédito, 8 de las cuales estan atrazadas en sus pagos. Si se selecciona al azar 5 cuentas de las 50, ¿Cuál es la probabilidad que por lo menos una cuenta de las cuentas escogidas corresponda a un cliente atrasado en sus pagos?

26. Se extrae una carta de una baraja de 52 cartas. Se gana I/. 10.00, si el resultado es par o divivible por 3. ¿Cuál es la probabilidad de ganar?

27. Un vendedor está tratando de vender artefacto a tres clientes. Sea A, B, C los eventos de hacer una venta al primero, segundo y tercer cliente respectivamente. La probabilidad que el primer cliente o el segundo, pero no el tercero comprarían es 0.65. La probabilidad que el primero y el segundo cliente comprarían es 0.20. La probabilidad de que haga primera venta pero no la tercera es 0.25, la probabilidad que ni el primer, ni el segundo cliente comprarían es 0.25. La probabilidad que el segundo no compra pero el tercero si es 0.30. ¿Cuál es la probabilidad que sólo uno de los dos primeros, compran pero no el tercero?.

28. En una urna existen 3 bolas rojas; 6 blancas; 4 verdes y 2 negras. Determine ud. la probabilidad que al elegir 3 bolas al azar:
(a) Ellas no resultan del mismo color.
(b) Ellas resultan de colores diferentes.

29. En cierta ciudad, la probabilidad que una familia tenga televisor es 0.85, un refrigerador es 0.60 y que tengan ambos es 0.50. ¿Cuál es la probabilidad que una familia tenga al menos uno de estos artefactos?.

30. Una compañía comercial tiene 130 sucursales localizadas en las tres regiones del país y se dedican a la venta de diversos artículos tal como aparece en el cuadro.

Regiones	Carros	Repuestos	Art. eléctricos	Total
Costa	50	20	30	100
Sierra	10	5	10	25
Selva	1	0	4	5
Total	61	25	44	130

Se selecciona, al azar, una sucursal para colocar en el mercado un nuevo producto que pueda ser vendido por cualquiera de las sucursales. Determine la probabilidad que:
(a) la sucursal seleccionada no esté localizada en la selva o venda repuestos.
(b) no venda carros o artefactos eléctricos y esté localizada en la costa o en la selva.

31. De una baraja de 52 cartas, se extrae al azar una de ellas. Determinar la

probabilidad que:

(a) Sea figura o copa; (b) Sea figura; pero no espada.

32. En una ciudad se publican tres revistas: A, B y C. El 30% de la población lee A, el 20% lee B, el 15% lee C, el 12% lee A y B, el 9% A y C, el 6% B y C, y el 3% leen A, B y C. Determinar el porcentaje de personas que:

(a) leen al menos uno de las tres revistas. (b) lee solamente A.

(c) leen B o C; pero no A. (d) leen A o no lee B ni C.

33. El gerente de una planta química situada en el puerto del callao sabe que en un pleito que se le avecina en la corte, la compañía puede ser culpable de contaminar el mar. Mas aún, él sabe que si la encuentran culpable, la compañía tendrá que instalar un sistema de purificación del agua, pagar una multa, o ambos. Hasta ahora, solo un 10% de las compañías en casos similares han tenido que pagar la multa e instalar un sistema de purificación. Adicionalmente, cuando la decisión de la corte no obliga a las dos penas, una compañía ha tenido tres veces más la probabilidad de ser multada que de ser requerida para instalar el sistema de purificación. Si el 28% de las compañías han sido culpables. ¿Cuál es la probabilidad de que esta compañía sea requerida para instalar un sistema de purificación?

34. Los 500 clientes de crédito de creditos S.A., estan categorizados según el número de años que han tenido cuenta de crédito con Créditos S.A. , y por su promedio de saldo de crédito. De estos clientes, 210 han tenido saldos menores a I/. 1000; otros 260 han tenido cuenta de Crédito cuando menos cinco años, y 80 han tenido saldos mayores de I/. 1000 y cuenta de crédito por menos de cinco años. Si se selecciona al azar un cliente, ¿cuál es la probabilidad que tenga

(a) un saldo de crédito mayor a I/. 1000 ?

(b) un saldo de crédito menor a I/. 1000 o haya tenido cuenta de crédito por lo menos cinco años?

(c) un saldo de crédito menor a I/. 1000 y haya tenido cuenta de crédito por lo menos cinco años?

35. De los 250 empleados de una compañía, 130 fuman cigarrillos. Hay 150 hombres que trabajan en esta compañía, de los cuales 85 fuman cigarrillos. ¿Cuál es la probabilidad que un empleado seleccionado al azar,

(a) no fume cigarrillos? (b) sea mujer y fume cigarrilos?

(c) sea hombre o fume cigarrillos?

1.7 PROBABILIDAD CONDICIONAL, REGLA DE MULTIPLICACION

La definición de probabilidad, en sus diversas formas discutidas en las secciones anteriores, relaciona todo el espacio muestral Ω y hemos utilizado el símbolo $P[A]$ para denotar la probabilidad de estos eventos; podríamos haber usado el símbolo $P[A|\Omega]$, que se lee: "la probabilidad del evento A dado que ha ocurrido Ω". Frecuentemente estaremos interesados en obtener la probabilidad de un evento, donde dicho evento está *condicionado* a la ocurrencia de un subconjunto del espacio muestral. Es decir, se da que el evento B ha ocurrido, y se quiere saber la probabilidad que ocurra el evento A. Evidentemente, la probabilidad de un evento A es diferente cuando tenemos la información que ha ocurrido ya un subconjunto B de Ω.

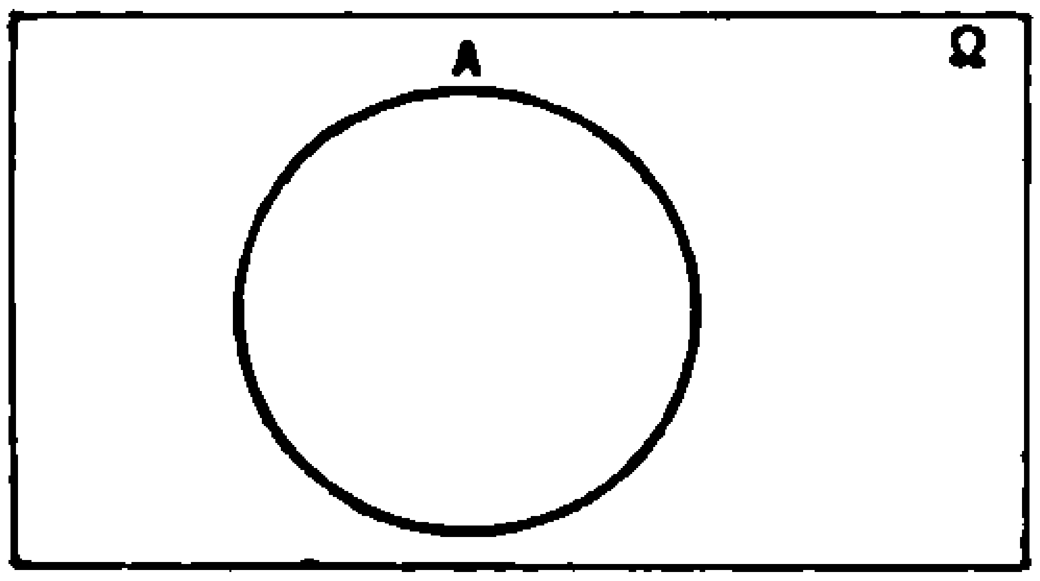


Fig. 1.7.1

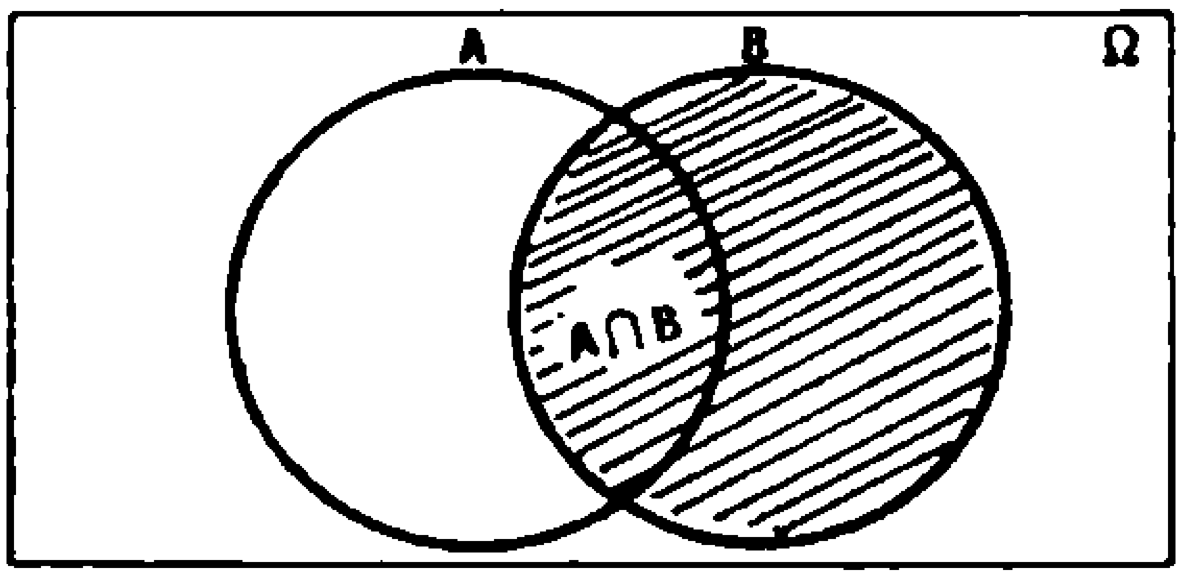


Fig. 1.7.2

Se dice que ya ha ocurrido B, entonces, se tiene que el espacio muestral Ω se ha restringido al subconjunto B. Pués se sabe que no ha ocurrido todo suceso ω que pertenece a $\bar{B}$. Por lo tanto, sería razonable definir la *probabilidad del evento A dado que ha ocurrido B*, denotado por $P[A|B]$, tal como sugiere la figura 1.7.2 (parte sombreada), igual a la razón del área $A \cap B$ al área B visualizados como probabilidades

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]}, \quad \text{si} \quad P[B] > 0$$

Así, $P[A \mid \Omega]$ se puede escribir ahora,

$$P[A \mid \Omega] = \frac{P[A \cap \Omega]}{P[\Omega]} = P[A]$$

Daremos algunas ilustraciones para aclarar mejor esta idea intuitiva. Consideremos el "lanzamiento de un dado y observar el número que resulta en la cara superior". El espacio muestral es $\Omega = \{1,2,3,4,5,6\}$, sea los siguientes eventos:

A: "Se observa un número impar". B: "Se observa un número mayor que 3".

Desde que $A = \{1,3,5\}$, $B = \{4,5,6\}$ se tiene $A \cap B = \{5\}$ entonces, $P[A \cap B] = \frac{1}{6}$ y $P[B] = \frac{3}{6}$. Por lo tanto, la probabilidad de que el evento A ocurra, dado que el evento B ha ocurrido es

$$P[A \mid B] = \frac{1/6}{3/6} = \frac{1}{3} .$$

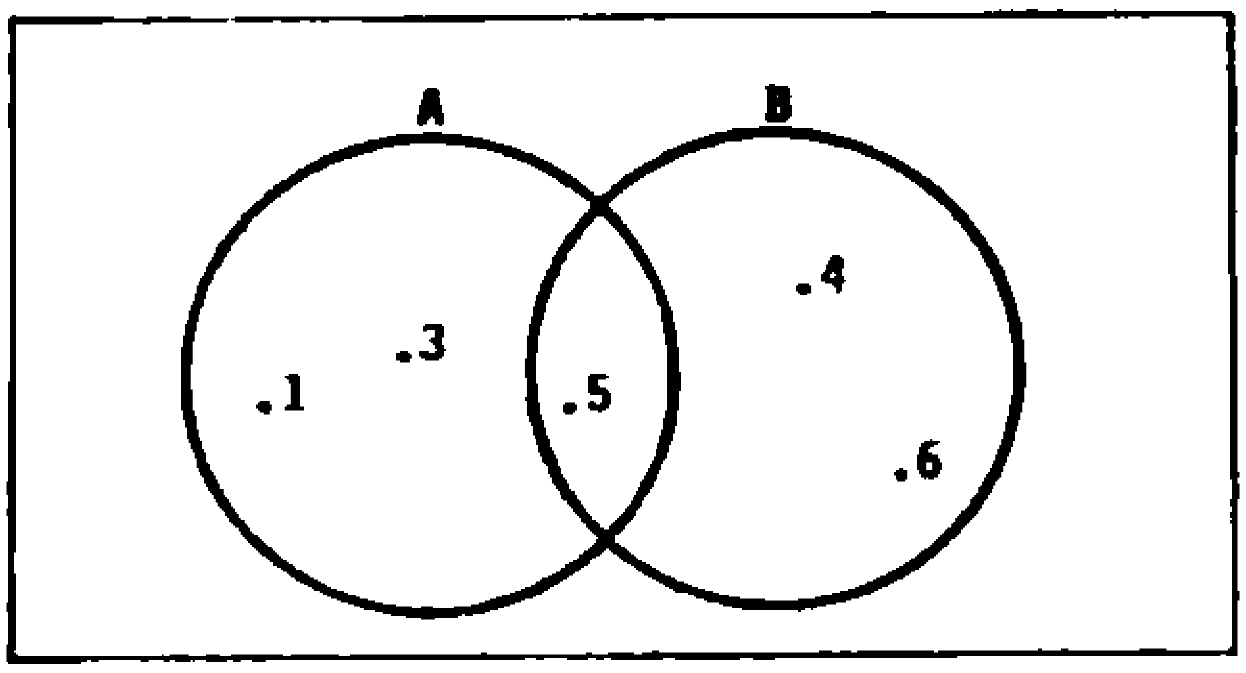


Fig. 1.7.3

Note que $N(A \cap B) = 1$ y $N(B) = 3$, entonces

$$P[A \mid B] = \frac{N(A \cap B)}{N(B)} = \frac{1}{3} .$$

consideremos ahora el lanzamiento de un par de dados. Y suponiendo que se nos informa haber obtenido suma mayor que 6. ¿Cuál es la probabilidad de obtener suma 7?

Obsérvese que la información proporcionada descarta, por ejemplo, la ocurrencia del par (2,3) y descarta la ocurrencia de todos los sucesos fuera del "trapecio" de la figura 1.7.4. Entonces si, **A**, es el evento: "obtener suma 7" y B : "se obtubo suma mayor que 6". Se tiene $N(A \cap B) = 6$ y - $N(B) = 21$. Luego,

$$P[A \mid B] = \frac{N(A \cap B)}{N(B)} = \frac{6}{21}$$

(1 , 1)	(1 , 2)	(1 , 3)	(1 , 4)	(1 , 5)	(1 , 6)
(2 , 1)	(2 , 2)	(2 , 3)	(2 , 4)	(2 , 5)	(1 , 6)
(3 , 1)	(3 , 2)	(3 , 3)	(3 , 4)	(3 , 5)	(3 , 6)
(4 , 1)	(4 , 2)	(4 , 3)	(4 , 4)	(4 , 5)	(4 , 6)
(5 , 1)	(5 , 2)	(5 , 3)	(5 , 4)	(5 , 5)	(5 , 6)
(6 , 1)	(6 , 2)	(6 , 3)	(6 , 4)	(6 , 5)	(6 , 6)

Fig. 1.7.4

o aplicando nuestra definición intuitiva, tenemos

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{6/36}{21/36} = \frac{6}{21} .$$

Formalizaremos ahora la definición de probabilidad condicional.

DEFINICION 1.7.1 Sea un evento B con $P[B] > 0$, la probabilidad ***condicional*** de que ocurra el evento A, dado que ha ocurrido B, denotado $P[A \mid B]$ se define como sigue

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} .$$

NOTA 1 $P[A \mid B]$ no está definida, si $P[B] = 0$

NOTA 2 Como hemos visto intuitivamente en los dos ejemplos dados.

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{N(A \cap B)/n}{N(B)/n} = \frac{N(A \cap B)}{N(B)}$$

Es decir, la probabilidad condicional es una probabilidad calculada en un espacio muestral reducido, B; pués a partir de la información sabemos con probabilidad 1 que el evento B ya ocurrió. En la práctica podemos resolver el problema usando la definición, esto es calculado $P[A \cap B]$ y $P[B]$ con respecto al espacio muestral original o considerando la probabilidad del evento A con respecto al espacio muestral reducido B (como indica la nota 2 y los ejemplos que hemos dado).

La definición de probabilidad condicional $P[\,.\mid.\,]$ dado como resultado de la noción intuitiva presentada en la discusión de la definición, es una probabilidad definida en un espacio muestral reducido y es de esperar que puedan establecerse los axiomas y resultados establecidos para $P[\,.\,]$. Es decir, si B es un evento tal que $P[B] > 0$, $P[\,. \mid B]$ satisface los tres axiomas

Ax1. $0 \leqslant P[A \mid B] \leqslant 1$

Ax2. $P[\Omega \mid B] = 1$

Ax3. $P[A_1 \cup A_2 \cup \ldots \cup A_n \mid B] = P[A_1 \mid B] + P[A_2 \mid B] + \ldots + P[A_n \mid B]$

ó $$P\left[\bigcup_{i=1}^{n} A_i \mid B\right] = \sum_{i=1}^{n} P[A_i \mid B]$$

para una secuencia de n eventos $A_1, A_2, \ldots, A_n$, mutuamente excluyentes $(A_i \cap A_j = \phi \;\forall\; i \neq j)$.

En general, para una secuencia numerable de eventos $A_1, A_2, \ldots$ mutuamente excluyentes, se tiene

$$P[\bigcup_{i=1}^{\infty} A_i \mid B] = \sum_{i=1}^{\infty} P[A_i \mid B]$$

De los axiomas **Ax1, Ax2, Ax3** de la probabilidad condicional, se demuestra que los teoremas dados en 1.6 siguen siendo válidos para la probabilidad - condicional. Es decir, si **B** es un evento tal que **P[B]** > 0, **P[. | B]** tiene las siguientes propiedades:

TEOREMA 1.7.1 $P[\phi \mid B] = 0$

TEOREMA 1.7.2 $P[\bar{A} \mid B] = 1 - P[A \mid B]$ o $P[A \mid B] = 1 - P[\bar{A} \mid B]$

TEOREMA 1.7.3 Si $A \subset C$ entonces $P[A \mid B] < P[C \mid B]$

TEOREMA 1.7.4 $P[A \cup C \mid B] = P[A \mid B] + P[C \mid B] - P[A \cap C \mid B]$

EJEMPLO 1 En Lima, Perú la probabilidad que llueva el día primero de Julio es 0.50 y la probabilidad que llueva los dos primeros dias de Julio es 0.40. Dado que llovió el día primero, ¿cuál es la probabilidad que - llueva el día siguiente?

SOLUCION Definimos los siguientes eventos

A: "llovió el primer día de Julio" B: "llueve el segundo día de Julio"

Entonces $P[A] = 0.50$ y $P[A \cap B] = 0.40$.

Usando la definición de probabilidad condicional se tiene

$$P[B \mid A] = \frac{P[A \cap B]}{P[A]} = \frac{0.40}{0.50} = 0.80.$$

EJEMPLO 2 En el ejemplo 11 de 1.6 suponga que él llama a la chica y concierta una cita. Convienen en encontrarse en el café **AMARGO**. Cuando llega, ella esta sentada en la barra y ve que tiene el pelo rubio. ¿Cuál es la probabilidad que su estatura sea mayor que 1.65 mts.? ¿Cuál de que sea menor que 1.65 mts?

SOLUCION Por la definición de probabilidad condicional se tiene

$$P[\bar{E} \mid R] = \frac{P[\bar{E}\ R]}{P[R]} = \frac{48/200}{60/200} = \frac{4}{5}$$

por el teorema 1.7.2 se obtiene

$$P[E \mid R] = 1 - P[\bar{E} \mid R] = 1 - \frac{4}{5} = \frac{1}{5} .$$

EJEMPLO 3 En una Universidad de 10000 estudiantes y 1000 profesores, el 10% de los profesores son de izquierda y 90% de derecha, mientras que en los estudiantes este porcentaje es al contrario. Se selecciona al azar un miembro de la Universidad y se encuentra que es de derecha, ¿cuál es la probabilidad que se haya seleccionado un estudiante? ¿un profesor?

SOLUCION Note que el espacio muestral es la comunidad Universitaria con $10000 + 1000 = 11000$ integrantes. Es decir $N(\Omega) = 11000$.

Definimos los siguientes eventos:

D: "un integrante de derecha"

E: "el integrante es un estudiante"

(a) Debemos calcular $P[E \mid D] = \dfrac{P[ED]}{P[D]}$,

$$N(ED) = 1000 \quad \text{y} \quad N(D) = 1900$$

Luego, $$P[E \mid D] = \frac{1000/_{11000}}{1900/_{11000}} = \frac{10}{19}.$$

(b) $$P[\bar{E} \mid D] = \frac{P[\bar{E}D]}{P[D]}$$

$N(\bar{E}D) = 900$. Entonces

$$P[\bar{E} \mid D] = \frac{900/_{11000}}{1900/_{11000}} = \frac{9}{19}.$$

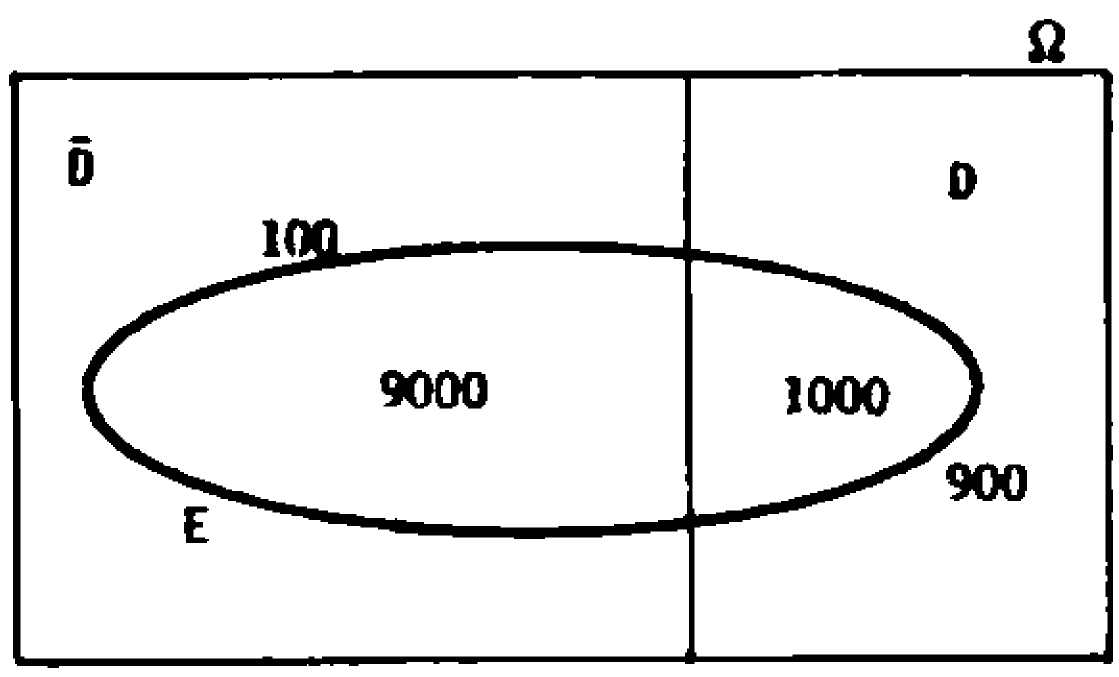


Fig. 1.7.5 Diagrama de Venn para el ejemplo 3.

o tambien usando el teorema 1.7.2

$$P[\bar{E} \mid D] = 1 - P[E \mid D] = 1 - \frac{10}{19} = \frac{9}{19}.$$

EJEMPLO 4 Cierta universidad en formación en su primer año de funcionamiento tiene tres curricula: Ciencia, Administración e Ingeniería. La clasificación de los alumnos por su sexo, es como sigue

	Ciencia	Administración	Ingeniería	Total
Hombres	250	350	200	800
Mujeres	100	50	50	200
Total	350	400	250	1,000

Tabla 1.7.1

Se selecciona un estudiante aleatoriamente del grupo. Si se sabe que el estudiante es hombre. ¿Cuál es la probabilidad que esté en Ciencias? ¿Cuál que esté en Ingeniería? ¿Cuál que el estudiante está matriculado en Administración?. Si el estudiante es una mujer. ¿Cuál que esté en Ciencias? ¿Cuál que esté matriculado en Ingeniería? ¿Cuál que esté en Administración?

SOLUCION Definimos los siguientes eventos :

B_1: "El estudiante seleccionado es hombre".

B_2: "El estudiante seleccionado es mujer".

A_1: "El estudiante sigue ciencias".

A_2: "El estudiante está matriculado en Administración"

A_3: "El estudiante está matriculado en Ingeniría".

$$P[B_1] = \frac{800}{1000} = 0.80 \ ; \qquad P[B_2] = \frac{200}{1000} = 0.20,$$

estas probabilidades se llaman algunas veces ***Probabilidades marginales***. De la misma manera :

$$P[A_1] = \frac{350}{1000} = 0.35 \ ; \qquad P[A_2] = \frac{400}{1000} = 0.40$$

$$P[A_3] = \frac{250}{1000} = 0.25, \quad \text{son probabilidades marginales.}$$

$P[B_i \cap A_j]$, $i = 1,2$; $j = 1,2,3$. Se llaman también ***probabilidades conjuntas***, están dadas en la tabla 1.7.2

	A_1	A_2	A_3	Total
B_1	0.25	0.35	0.20	0.80
B_2	0.10	0.05	0.05	0.20
Total	0.35	0.40	0.25	1.00

Tabla 1.7.2

donde, por ejemplo $P[B_1 \cap A_1] = \dfrac{250}{1000} = 0.25$.

Cálculo de las probabilidades pedidas. Es decir

$$P[A_j \mid B_i] = \frac{P[A_j \cap B_i]}{P[B_i]}\,, \quad i = 1,2; \quad j = 1,2,3.$$

por ejemplo :

$$P[A_1 \mid B_1] = \frac{250}{800} = \frac{0.25}{0.80} - 0.3125 .$$

$$P[A_1 \mid B_2] = \frac{100}{200} = \frac{0.10}{0.20} = 0.5 .$$

$$P[A_2 \mid B_1] = \frac{350}{800} = \frac{0.35}{0.80} = 0.4375 .$$

$$P[A_2 \mid B_2] = \frac{50}{200} = \frac{0.05}{0.20} = 0.25 .$$

$$P[A_3 \mid B_1] = \frac{200}{800} = \frac{0.20}{0.80} = 0.25 .$$

$$P[A_3 \mid B_2] = \frac{50}{200} = \frac{0.05}{0.20} = 0.25 .$$

EJEMPLO 5 En una universidad el 70% de los estudiantes son de Ciencias y el 30% de letras; de los estudiantes de Ciencias - el 60% son varones y los de letras son varones el 40%. Si se elige aleatoriamente un estudiante, calcular la probabilidad que:

(a) sea un estudiante varón.

(b) Sea un estudiante varón, si es de Ciencias.

(c) Sea un estudiante de Ciencias, si es varón.

(d) Sea un estudiante de Ciencias y varón.

SOLUCION La información contenida en el enunciado, la resumimos en la tabla 1.7.3 con dos entradas. Es claro que el 42% de los estudiantes son varones y estudian ciencias, este porcentaje lo obtenemos a partir del hecho de que el 60% de los de Ciencias son varones; es decir el 60% del 70% - del total. En forma análoga se obtuvieron los otros porcentajes. Definimos ahora los eventos:

A: "El estudiante elegido es de Ciencias".

B: "El estudiante elegido es varón".

Esp. \ Sexo	Varones	Mujeres	Total
Ciencias	42 %	28 %	70 %
Letras	12 %	18 %	30 %
Total	54 %	46 %	100 %

Tabla 1.7.3

(a) $P[B] = 0.54$.

(b) $P[B \mid A] = \dfrac{P[A \cap B]}{P[A]} = \dfrac{0.42}{0.70} = 0.6$.

(c) $P[A \mid B] = \dfrac{P[A \cap B]}{P[B]} = \dfrac{0.42}{0.54} = 0.778.$

(d) $P[A \cap B] = 0.42$.

EJEMPLO 6 Cinco cartas, numeradas de 1 a 5, son puestas en una caja y revoleadas completamente. Se selecciona tres cartas aleatoriamente y sin restitución, y se poðen en una mesa mostrando el número. Sea A_i el evento el número i $(1 \leqslant i \leqslant 5)$ está entre estos seleccionados (Así, $A_1 = (1,2,4)$, ó $(1,3,5)$ etc).Suponga que cada combinación de tres cartas son igualmente probables. Calcular $P[A_i \mid A_j]$.

SOLUCION El número total de combinaciones de tres cartas seleccionadas de las cinco de la caja es

$$N(\Omega) = C(5,3) = 10 .$$

El número total de combinaciones que tienen un número especifico i es

$$N(A_i) = C(4,2) = 6. \qquad \text{Entonces}$$

$$P[A_i] = \frac{C(4,2)}{C(5,3)} = \frac{6}{10} ; \qquad i = 1,2,3,4,5 .$$

El número total de combinaciones con los números especificados i, y j, - $i \neq j$ es $C(3,1) = 3 = N(A_i \cap A_j)$. Entonces

$$P[A_i \cap A_j] = \frac{3}{10} .$$

Luego, por la definición de probabilidad condicional

$$P[A_i \mid A_j] = \frac{P[A_i \cap A_j]}{P[A_j]} = \frac{3/10}{6/10} = \frac{1}{2} .$$

EJEMPLO 7 Demostremos que $P[A \mid B] + P[\bar{A} \mid B] = 1$. Si $P[B] > 0$

DEMOSTRACION Por la definición de probabilidad condicional

$$P[A \mid B] + P[\bar{A} \mid B] = \frac{P[A \cap B]}{P[B]} + \frac{P[\bar{A} \cap B]}{P[B]} = \frac{1}{P[B]} \, P[A \cap B] +$$

$$+ P[\bar{A} \cap B] = \frac{P[B]}{P[B]} = 1.$$

EJEMPLO 8 Demostrar que $P[A \mid \bar{B}] + P[\bar{A} \mid \bar{B}] = 1$.

DEMOSTRACION $\bar{B} = (A \cap \bar{B}) \cup (\bar{B} \cap \bar{A})$

$$P[\bar{B}] = P[A \cap \bar{B}] + P[\bar{A} \cap \bar{B}]$$

ya que $A \cap \bar{B}$ y $\bar{A} \cap \bar{B}$ son eventos mutuamente excluyentes.
Luego,

$$1 = \frac{P[A \cap \bar{B}]}{P[\bar{B}]} + \frac{P[\bar{A} \cap \bar{B}]}{P[\bar{B}]}$$

$$1 = P[A \mid \bar{B}] + P[\bar{A} \mid \bar{B}] .$$

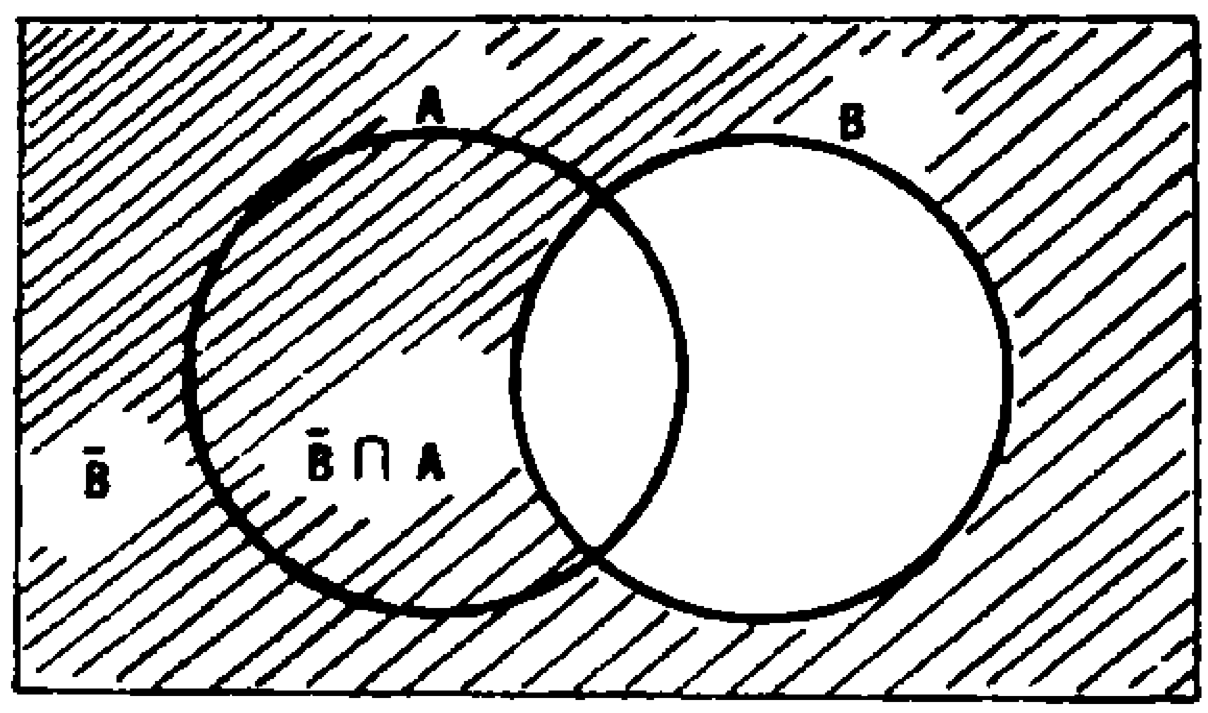


Fig. 1.7.6

EJEMPLO 9 Un hombre tiene dos carros viejos, *a* y *b*; ellos tienen problemas para arrancar en las mañanas frías. La probabilidad que ambos arrancan es 0.1; la probabilidad que arranca b y a no es 0.2; la probabilidad - que ninguno de ellos arranca es 0.4. Hallar la probabilidad que

(a) el carro a arranca.

(b) arranca a, dado que arrancó b.

(c) arranca b, dado que a no arrancó .

SOLUCION Sean los siguientes eventos:

A: "El carro a arranca". B: "El carro b arranca" .

Entonces, según el enunciado se tiene

$$P[A \cap B] = 0.1, \qquad P[\bar{A} \cap B] = 0.2, \qquad P[\bar{A} \cap \bar{B}] = 0.4$$

(a) $\bar{A} = (\bar{A} \cap B) \cup (\bar{A} \cap \bar{B})$

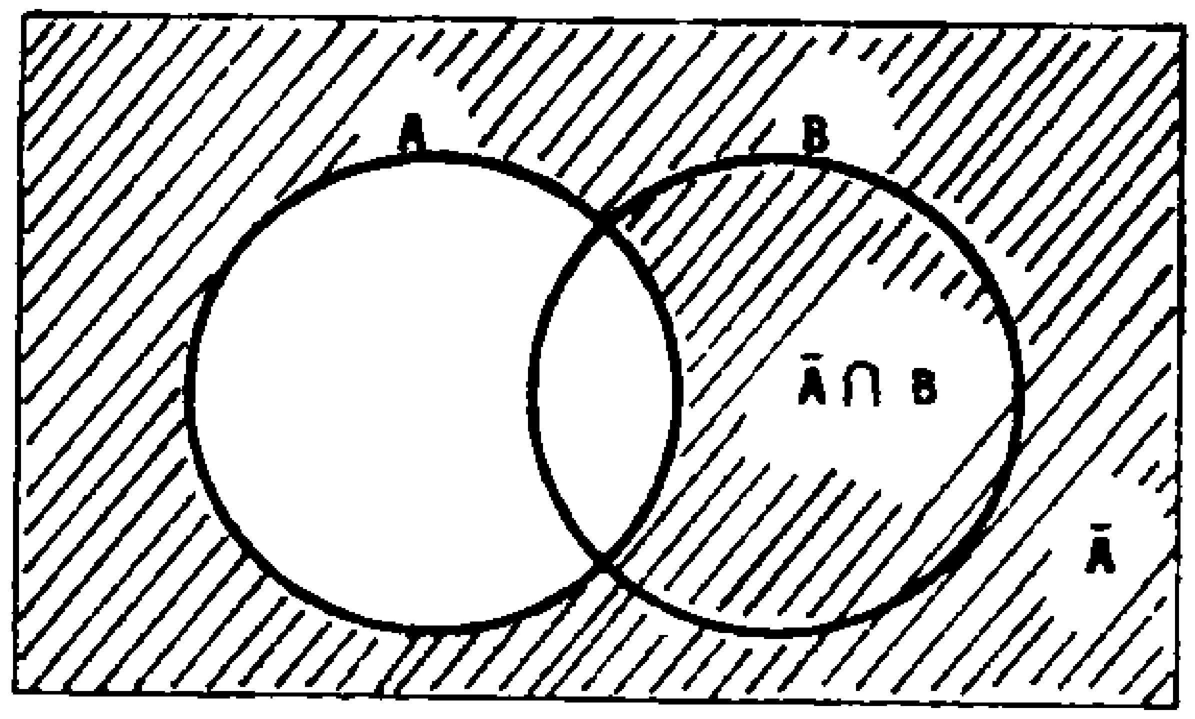


Fig. 1.7.7

$$P[\bar{A}] = P[(\bar{A} \cap B) \cup (\bar{A} \cap \bar{B})]$$

$$= P[\bar{A} \cap B] + P[\bar{A} \cap \bar{B}]$$

$$= 0.2 + 0.4 = 0.6 \; .$$

Ya que los eventos $\bar{A} \cap B$ y $\bar{A} \cap \bar{B}$ son mutuamente excluyentes. Luego,

$$P[A] = 1 - P[\bar{A}] = 1 - 0.6 = 0.4$$

(b) Debemos calcular antes $P[B]$. Observe que $B = (A \cap B) \cup (\bar{A} \cap B)$

$$P[B] = P[(A \cap B) \cup (\bar{A} \cap B)] = P[A \cap B] + P[\bar{A} \cap B] = 0.1 + 0.2 = 0.3$$

ya que los eventos $A \cap B$ y $\bar{A} \cap B$ son mutuamente excluyentes. Entonces,

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{0.1}{0.3} = \frac{1}{3} \; .$$

(c) $$P[B \mid \bar{A}] = \frac{P[\bar{A} \cap B]}{P[\bar{A}]} = \frac{0.2}{0.6} = \frac{1}{3} \; .$$

EJEMPLO 10 De una urna que contiene 12 bolas, de las cuales ocho son blancas, se extrae una muestra de tamaño 4, con reemplazo (sin reemplazo). Encuentre la probabilidad que la bola observada en la tercera extracción haya sido blanca, dado que la muestra contiene exactamente tres bolas blancas.

SOLUCION Sean los eventos :

A: "la muestra contiene exactamente tres bolas blancas".

B: "la bola extraída en la tercera extracción haya sido blanca".

El problema consiste en calcular

$$P[B \mid A] = \frac{P[A \cap B]}{P[A]}$$

(a) En el caso de muestreo con reemplazo .

$n = N(\Omega) = (12)^4$ y $n_A = \binom{4}{3} \times 8^3 \times 4$, entonces

$$P[A] = \frac{\binom{4}{3}8^3 \times 4}{(12)^4} .$$

Los elementos de $A \cap B$ son de la forma $b\ b\ \underset{\substack{\uparrow \\ \text{fijo}}}{b}\ b'$,

entonces varían las 2 blancas y la de diferente color que ocurre de $\binom{3}{2}$ - formas. Luego, $n_{A \cap B} = \binom{3}{2}\ 8^3 \times 4$, entonces

$$P[A \cap B] = \frac{\binom{3}{2} \times 8^3 \times 4}{(12)^4} .$$

Por lo tanto, $P[B \mid A] = \dfrac{P[A \cap B]}{P[A]} = \dfrac{\dfrac{\binom{3}{2} \times 8^3 \times 4}{(12)^4}}{\dfrac{\binom{4}{3}8^3 \times 4}{(12)^4}} = \dfrac{\binom{3}{2}}{\binom{4}{3}} = \dfrac{3}{4}$

(**b**) En el caso del muestreo sin reemplazo.

$n = N(\Omega) = P^4_{12}$, $n_A = \binom{4}{3}\ P^3_8 \times 4$, entonces

$$P[A] = \frac{\binom{4}{3} P^3_8 \times 4}{P^4_{12}} .$$

Los elementos de $A \cap B$, tienen la forma $b\ b\ \underset{\substack{\uparrow \\ \text{fijo}}}{b}\ b'$, es decir $\binom{3}{2}$

formas de variar. Es decir $n_{A \cap B} = \binom{3}{2}\ P^3_8 \times 4$

Luego, $P[A \cap B] = \dfrac{\binom{3}{2}\ P^3_8 \times 4}{P^4_{12}}$.

por lo tanto, $P[B \mid A] = \dfrac{\dfrac{\binom{3}{2} P^3_8 \times 4}{P^4_{12}}}{\dfrac{\binom{4}{3} P^3_8 \times 4}{P^4_{12}}} = \dfrac{\binom{3}{2}}{\binom{4}{3}} = \dfrac{3}{4}$.

Compare con la respuesta anterior, explique

EJEMPLO 11 La probabilidad que la construcción de un edificio termine a tiempo es 17/20, la probabilidad que no haya huelga es 3/4, y la probabilidad que la construcción se termine a tiempo dado que no hubo huelga es 14/15; la probabilidad que haya huelga y no se termine la construcción a tiempo es 1/10. ¿Cuál es la probabilidad que
(a) la construcción se termina a tiempo y no haya huelga?
(b) No haya huelga dado que la construcción se terminó a tiempo?
(c) la construcción no se termina a tiempo si hubo huelga?
(d) la construcción no se termina a tiempo si no hubo huelga?

SOLUCION Definimos los siguientes eventos :
A: "La construcción se termina a tiempo".
B: "No haya huelga".
del enunciado del problema tenemos :

$$P[A] = \frac{17}{20}, \quad P[B] = \frac{3}{4}, \quad P[A \mid B] = \frac{14}{15}, \quad P[\bar{A} \cap \bar{B}] = \frac{1}{10}$$

(a) $P[A \cap B] = ?$

Sabemos que $P[A \mid B] = \dfrac{P[A \cap B]}{P[B]}$, de donde $\dfrac{14}{15} = \dfrac{P[A \cap B]}{3/4}$

luego, $P[A \cap B] = \dfrac{3}{4} \times \dfrac{14}{15} = \dfrac{7}{10} = 0.7$.

(b) $$P[B \mid A] = \frac{P[A \cap B]}{P[A]} = \frac{7/10}{17/20} = \frac{14}{17} .$$

(c) $$P[\bar{A} \mid \bar{B}] = \frac{P[\bar{A} \cap \bar{B}]}{P[\bar{B}]} = \frac{1/10}{1 - \frac{3}{4}} = \frac{4}{10} = \frac{2}{5} .$$

(d) $$P[\bar{A} \mid B] = \frac{P[\bar{A} \cap B]}{P[B]} = \frac{P[B] - P[A \cap B]}{P[B]} = 1 - \frac{P[A \cap B]}{P[B]}$$
$$= 1 - P[A \mid B] = 1 - \frac{14}{15} = \frac{1}{15} .$$

De paso el lector habrá notado que hemos "demostrado" el teorema 1.7.2 .

$$P[\bar{A} \mid B] = 1 - P[A \mid B] .$$

EJEMPLO 12 Se retira una carta roja de una baraja de 52 cartas (26 rojas y 26 negras); se extraen 13 cartas. Calcular la probabilidad que si todas son de un mismo color, ellas son negras.

SOLUCION Como se retira una carta roja, entonces queda 51 cartas, (25 rojas y 26 negras) por lo tanto, el espacio tiene $\binom{51}{13}$ elementos.

Definimos los siguientes eventos:

R: "obtener las 13 cartas rojas".

N: "obtener las 13 cartas negras".

E: "las 13 cartas son del mismo color".

Observe que $E = R \cup N$, Se pide calcular

$$P[N \mid E] = \frac{P[E \cap N]}{P[E]} = \frac{P[N \cap (R \cup N)]}{P[E]} = \frac{P[N]}{P[E]},$$

además $P[N] = \dfrac{\binom{26}{13}}{\binom{51}{13}}$, entonces

$$P[E] = P[R] + P[N] = \frac{\binom{25}{13}}{\binom{51}{13}} + \frac{\binom{26}{13}}{\binom{51}{13}} = \frac{\binom{25}{13}+\binom{26}{13}}{\binom{51}{13}}.$$

Luego,

$$P[N \mid E] = \frac{\dfrac{\binom{26}{13}}{\binom{51}{13}}}{\dfrac{\binom{25}{13}+\binom{26}{13}}{\binom{51}{13}}} = \frac{\binom{26}{13}}{\binom{25}{13}+\binom{26}{13}} = \frac{\dfrac{26!}{13!\,13!}}{\dfrac{25!}{13!\,12!} + \dfrac{26!}{13!\,13!}}$$

$$= \frac{\dfrac{26!}{13!\,13!}}{\dfrac{25!\,13! + 26!\,12!}{13!\,13!\,12!}} = \frac{26}{13 + 26} = \frac{2}{3}.$$

1.7.1 REGLA DE MULTIPLICACION

De la definición de probabilidad condicional, obtenemos una fórmula - para hallar la probabilidad de la intersección (o producto) de los eventos A y B, esto es, de

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} \,, \qquad P[B] > 0 \qquad (1)$$

$$P[B \mid A] = \frac{P[A \cap B]}{P[A]} \,, \qquad P[A] > 0 \qquad (2)$$

Multiplicando ambos miembros de la expresión (1) por $P[B]$ y por - $P[A]$ la expresión (2), obtenemos las ecuaciones

$$P[A \cap B] = P[B]\,P[A \mid B] = P[AB] \quad \text{y}$$

$$P[A \cap B] = P[A]\,P[B \mid A] = P[AB] \,,$$

este resultado, en teoría de probabilidad, se denomina ***REGLA DE MULTIPLICACION*** o probabilidad de la intersección, (también probabilidad conjunta); expresa la probabilidad de que ocurra los eventos A y B es igual a la probabilidad de la ocurrencia de uno de ellos multiplicado por la probabilidad condicional que ocurra el segundo, dado que el primero ha ocurrido.

EJEMPLO 13 Una urna contiene 5 bolas blancas y 6 negras; se extrae al azar sucesivamente y sin reposición (con reposición) dos bolas, ¿cuál es la probabilidad que las dos resultan blancas?

SOLUCION (a) Sin reposición .

PRIMERA FORMA Sean los eventos:

A_1: "La primera bola resultó blanca".

A_2: "La segunda bola resultó blanca".

E : "Las dos bolas resultan blancas".

Es claro que la probabilidad pedida es la del evento

$$E = A_1 \cap A_2 = A_1A_2$$

Es decir, E es la intersección de los dos eventos y la ocurrencia de A_1 - influye en la de A_2. O sea

$$P[E] = P[A_1A_2] = P[A_1]\,P[A_2 \mid A_1]$$

En la urna hay 11 bolas de los cuales 5 son blancas, entonces

$$P[A_1] = \frac{5}{11}$$

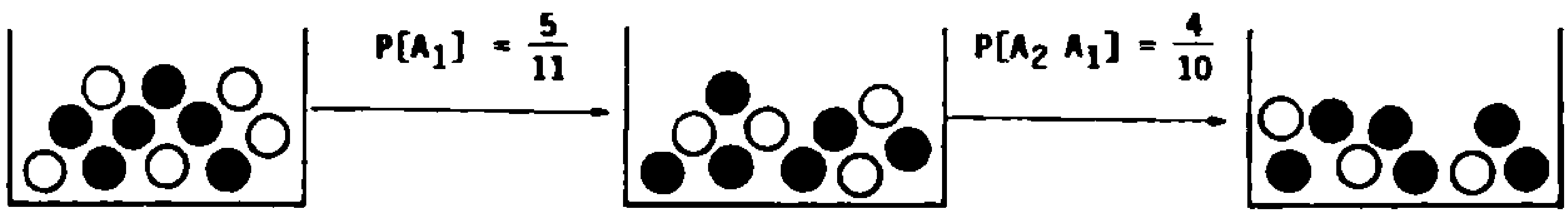
Después de la ocurrencia del evento A_1, queda 10 bolas de las cuales 4 son blancas, luego

$$P[A_2 \mid A_1] = \frac{4}{10} .$$

Por lo tanto

$$P[E] = P[A_1]\,P[A_2 \mid A_1] = \frac{5}{11} \times \frac{4}{10} = \frac{2}{11} .$$

Un esquema que ayuda a visualizar la solución de este problema es



SEGUNDA FORMA El problema se puede enfocar de otra manera: Como existe - 5 + 6 = 11 bolas, el número de formas de seleccionar dos de ellas esta dado por $\binom{11}{2}$ y cada una de estas formas tienen igual probabilidad de ocurrir. El número de sucesos favorables al evento E, la obtenemos de la siguiente manera: como existen 5 bolas blancas, podemos eligir dos de ellas de $\binom{5}{2}$ formas diferentes, luego aplicando la definición clásica es

$$P[E] = \frac{\binom{5}{2}}{\binom{11}{2}} = \frac{\frac{5!}{2!3!}}{\frac{11!}{2!9!}} = \frac{2}{11} .$$

(**b**) Con reemplazo. En este caso también el evento E se escribe como la intersección de los eventos A_1 y A_2 . O sea $E = A_1A_2$. Y la ocurrencia - de A_1 no afecta a la ocurrencia de A_2. En efecto:

$$P[A_1] = \frac{5}{11}$$

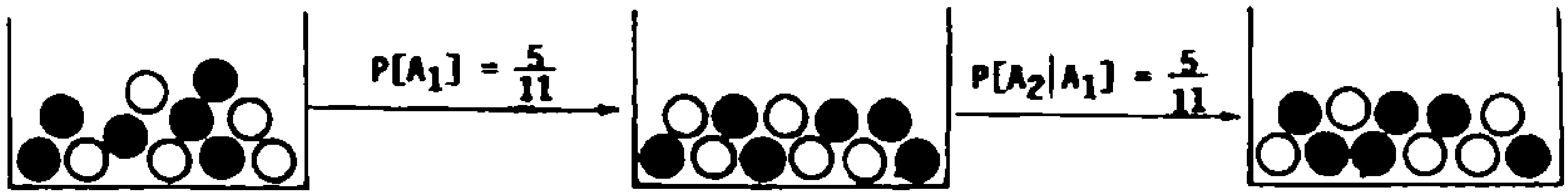
Puesto que después de la ocurrencia de A_1 se devuelve la bola a la urna, - también

$$P[A_2 \mid A_1] = \frac{5}{11}$$

Por lo tanto,

$$P[A_1A_2] = P[A_1]\ P[A_2 \mid A_1] = \frac{5}{11} \times \frac{5}{11} = \frac{25}{121} .$$

El esquema siguiente ayuda a visualizar la solución del problema



EJEMPLO 14 En el ejemplo 13. ¿Cuál es la probabilidad de obtener una bola de cada color?

SOLUCION (a) Sin restitución.

PRIMERA FORMA Sean los siguientes eventos:

A_1: "La primera bola resultó blanca".
A_2: "La sequnda bola resultó blanca".

B_1 "La primera bola resultó negra".
B_2 "La segunda bola resultó negra".
E : "obtener una bola de cada color".

En este caso la probabilidad pedida, es, la del evento E formado por la unión de los eventos A_1B_2 y B_1A_2, es decir, $E = A_1B_2 \cup B_1A_2$ y como dichos eventos son mutuamente excluyentes, tenemos

$$P[E] = P[A_1B_2] + P[B_1A_2]$$
$$= P[A_1]\,P[B_2 \mid A_1] + P[B_1]\,P[A_2 \mid B_1] \qquad (1)$$

La ocurrencia del primer evento influye en la ocurrencia del segundo. Puesto que en la urna hay 11 bolas de las cuales 5 son blancas, se tiene

$$P[A_1] = \frac{5}{11}$$

Similarmente $P[B_1] = 6/11$

Después de la ocurrencia de A_1, queda 10 bolas de las cuales 6 son negras, entonces

$$P[B_2 \mid A_1] = 6/10$$

Similarmente $P[A_2 \mid B_1] = 5/10$. Por lo tanto, reemplazando estos resultados en (1) se obtiene

$$P[E] = \frac{5}{11} \times \frac{6}{10} + \frac{6}{11} \times \frac{5}{10} = \frac{6}{11}.$$

Un diagrama que ayuda a visualizar estos problemas es el llamado "*ARBOL DE PROBABILIDAD PARA EXPERIMENTOS SUCESIVOS*" y es como indica la figura 1.7.8.

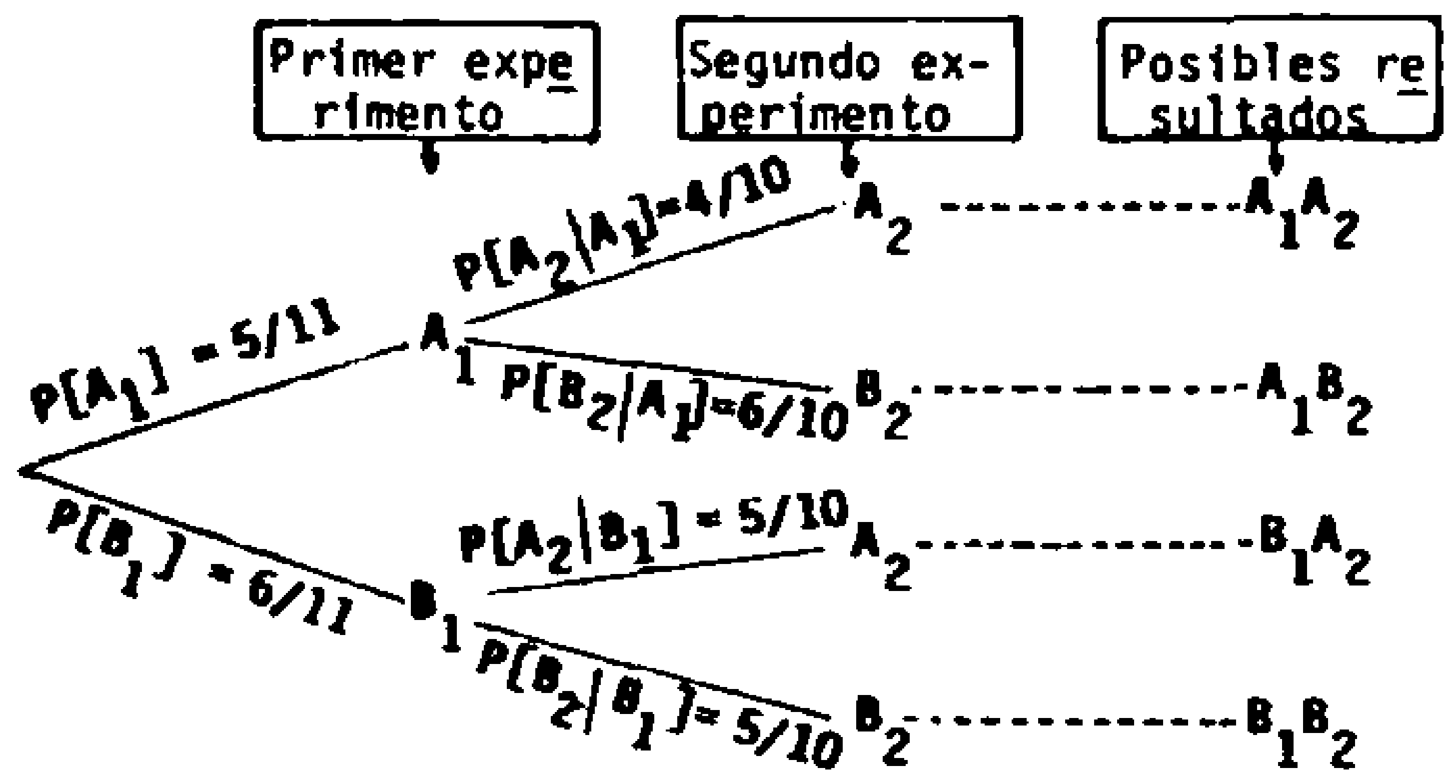


Fig. 1.7.8

Cada rama completa del diagrama del arból, se llama una *trayectoria* y representa un posible resultado del experimento. En cada segmento que une la secuencia de experimentos se pone sus respectivas probabilidades.

SEGUNDA FORMA El número de formas de extraer dos bolas de 11 es $\binom{11}{2}$. Cada uno de estos son igualmente posibles.

Número de sucesos favorables al evento E : Existen $\binom{5}{1}$ formas de elegir - una bola blanca y $\binom{6}{1}$ formas de eligir una bola negra, entonces aplicando el principio de multiplicación, el número de sucesos favorables a E está dado $\binom{5}{1}\binom{6}{1}$ = 5 x 6, y aplicando la definición clásica es

$$P[E] = \frac{\binom{5}{1}\binom{6}{1}}{\binom{11}{2}} = \frac{5 \times 6}{\frac{11!}{2!9!}} = \frac{5 \times 6}{11 \times 5} = \frac{6}{11} .$$

(b) Con restitución. Como en el caso anterior se pide calcular la probabilidad del evento $E = A_1B_2 \cup B_1A_2$. Es evidente que

$$\begin{aligned} P[E] &= P[A_1B_2] + P[B_1A_2] \\ &= P[A_1]\, P[B_2 \mid A_1] + P[B_1]\, P[A_2 \mid B_1] \\ &= \frac{5}{11} \times \frac{6}{11} + \frac{6}{11} \times \frac{5}{11} = \frac{60}{121} . \end{aligned}$$

EJEMPLO 15 En un sistema de alarma, la probabilidad que se produzca un peligro es 0.10. Si éste se produce, la probabilidad que la alarma funcione es de 0.95. La probabilidad que funcione la alarma sin haber habido peligro es 0.03. Determinar la probabilidad que haya un peligro y la alarma no funcione.

SOLUCION Definimos los siguientes eventos:

P : "hay peligro",

F : "la alarma funciona". Entonces, $\bar{F}$: "la alarma no funciona".

Luego, debemos determinar la probabilidad del evento:

$P\bar{F}$: "haya peligro y la alarma no funciona"

$$P[P\bar{F}] = P[P]\,P[\bar{F} \mid P]$$

$P[P] = 0.10$. Si ocurre el evento **P**, $P[F \mid P] = 0.95$ pero

$P[\bar{F} \mid P] = 1 - P[F \mid P] = 1 - 0.95 = 0.05.$ (teorema 1.7.2)

Por lo tanto, $P[P\bar{F}] = (0.10)\,(0.05) = 0.005.$

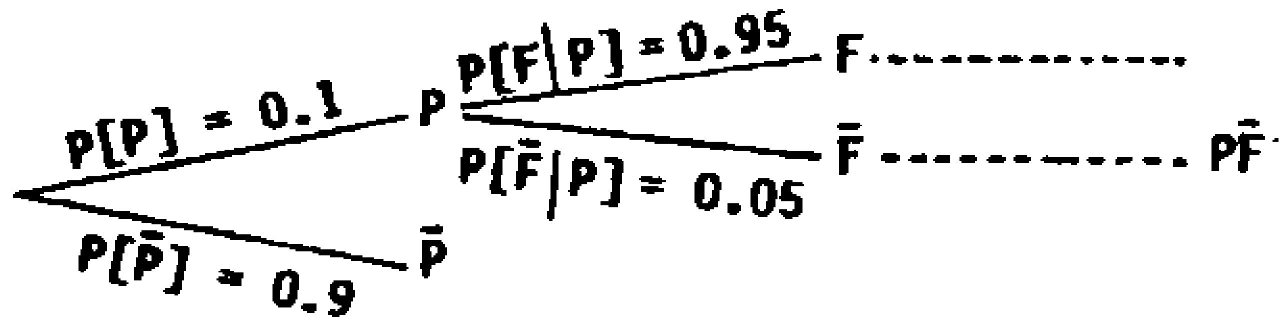


Fig. 1.7.9 Arbol de probabilidades para el problema 15.

TEOREMA 1.7.5 Si A, B y C son eventos de Ω, tales que $P[A] \neq 0$ y $P[A \cap B] \neq 0$, entonces

$$P[A \cap B \cap C] = P[A]\,P[B \mid A]\,P[C \mid A \cap B]$$

DEMOSTRACION Consideremos dos eventos $A \cap B$ y C; de la definición de probabilidad condicional

$$P[C \mid A \cap B] = \frac{P[C \cap A \cap B]}{P[A \cap B]} \quad \text{y} \quad P[B \mid A] = \frac{P[A \cap B]}{P[A]}$$

tenemos

$$P[A]\,P[B \mid A]\,P[C \mid A \cap B] = P[A]\,\frac{P[A \cap B]}{P[A]} \times \frac{P[A \cap B \cap C]}{P[A \cap B]}$$

$$= P[A \cap B \cap C].$$

El siguiente teorema es una generalización del teorema anterior.

TEOREMA 1.7.6 Si $A_1, A_2, \ldots, A_n$ son eventos de un espacio muestral finito y $P[A_1 \cap A_2 \cap \ldots \cap A_{n-1}] \neq 0$, entonces

$$P[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]\,P[A_2 \mid A_1]\,P[A_3 \mid A_1 \cap A_2] \ldots P[A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}]$$

la demostración queda como ejercicio para el lector interesado.

EJEMPLO 16 Dos establos A y B tienen 1,000 cabezas de vacuno cada uno. Existe una epidemia que afecta a los cascos y la boca del ganado. La proporción de ganados afectados con $\frac{1}{5}$ y $\frac{1}{4}$ respectivamente (por establo). Se escoge un ganado al azar.

(a) ¿Cuál es la probabilidad que el ganado escogido viene del rancho **A** y tiene afección a los cascos y la boca?.

(b) Si el 70% de los ganados afectados tienen edad menor que un año, ¿cuál es la probabilidad que el ganado escogido venga del rancho B, tiene afección y es mayor que un año de edad?

SOLUCION Definimos los siguientes eventos:

A: "el ganado escogido es del rancho A"

B: "el ganado escogido es del rancho B"

E: "el ganado tiene afección al casco y la boca".

(a) Debemos calcular

$$P[A \cap E] = P[A]\ P[E \mid A] = \frac{1000}{2000} \times \frac{1}{5} = \frac{1}{10} .$$

(b) Definimos el evento F: la vaca escogida tenga edad mayor que un año. - Entonces,

$$P[B \cap E \cap F] = P[B]\ P[E \mid B]\ P[F \mid (B \cap E)]$$

$$= \frac{1000}{2000} \times \frac{1}{4} \times \frac{75}{250} = \frac{1}{2} \times \frac{1}{4} \times \frac{3}{10} = \frac{3}{80} .$$

ya que en el rancho B hay 250 ganados afectados y de estos el 30% son mamayores que un año de edad.

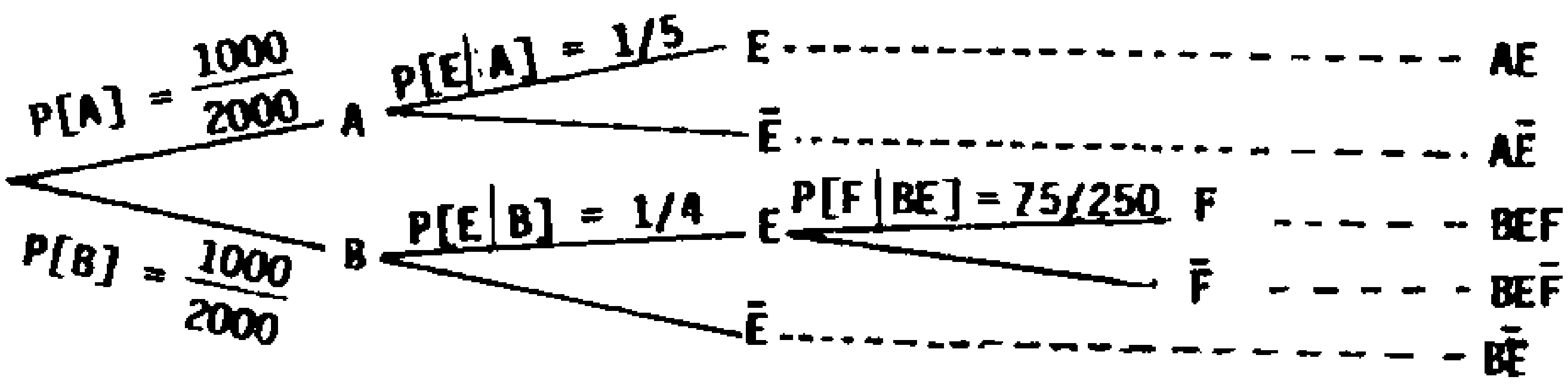


Fig. 1.7.10 Arbol de probabilidad para el problema 16

EJEMPLO 17 Un lote de 100 fusibles contiene 2 fusibles defectuosos. Si se prueban los fusibles uno por uno, ¿cuál es la probabilidad que el último fusible defectuoso sea detectado en la tercera prueba?

SOLUCION Definimos los siguientes eventos:

D_{ij} : "el i-ésimo defectuoso se obtuvo en la j-ésima extracción, $i = 1,2$; $j = 1,2,3$".

B_{ij} : "el i-ésimo bueno se obtuvo en la j-ésima extracción $i = 1,2$, ; $j = 1,2,3$".

E : "el último fusible defectuoso es detectado en la tercera prueba"

En el diagrama del árbol de probabilidades, podemos seguir sólo por las ramas que cumplen las condiciones requeridas por el evento cuya probabilidad se quiere calcular (en nuestro caso el evento E) y llegar solamente a los resultados favorables a dicho evento. Para el problena en cuestión, siguiendo este proceso se obtiene el árbol de probabilidades de la fig. 1.7.12. - De donde el evento E se escribe

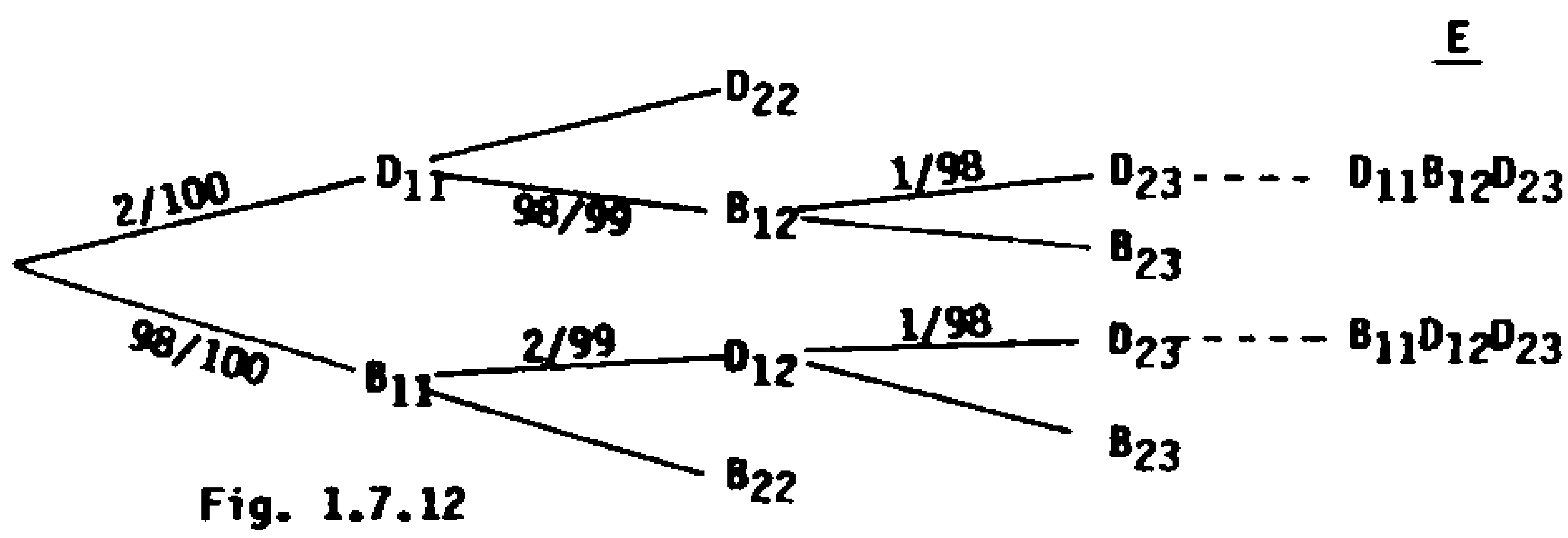


Fig. 1.7.12

$$E = D_{11}\,B_{12}\,D_{23} \cup B_{11}\,D_{12}\,D_{23}$$

Luego,

$$\begin{aligned} P[E] &= P[D_{11}\,B_{12}\,D_{23}] + P[B_{11}\,D_{12}\,D_{23}] \\ &= P[D_{11}]\,P[B_{12} \mid D_{11}]\,P[D_{23} \mid D_{11}B_{12}] + P[B_{11}] \\ &\quad P[D_{12} \mid B_{11}]\,P[D_{23} \mid B_{11}D_{12}] \\ &= \frac{2}{100} \cdot \frac{98}{99} \cdot \frac{1}{98} + \frac{98}{100} \cdot \frac{2}{99} \cdot \frac{1}{98} \\ &= \frac{2}{9900} + \frac{2}{9900} = \frac{1}{2475}\ . \end{aligned}$$

EJEMPLO 18 Un lote de 100 lámparas contiene 10 piezas defectuosas. Si se selecciona 3 lámparas aleatoriamente, ¿cuál es la probabilidad que sólo una sea defectuosa?

PRIMER METODO Definimos los siguientes eventos:

D : "se selecciona una lámpara defectuosa".

N : "se selecciona una lámpara no defectuosa".

A : "sólo una sea defectuosa de las tres extraidas"

Siguiendo el mismo proceso del ejemplo anterior se obtiene el diagrama del árbol de probabilidades de la fig. 1.7.13. De donde

$$A = NND \cup NDN \cup DNN$$

Eventos Favorables a A

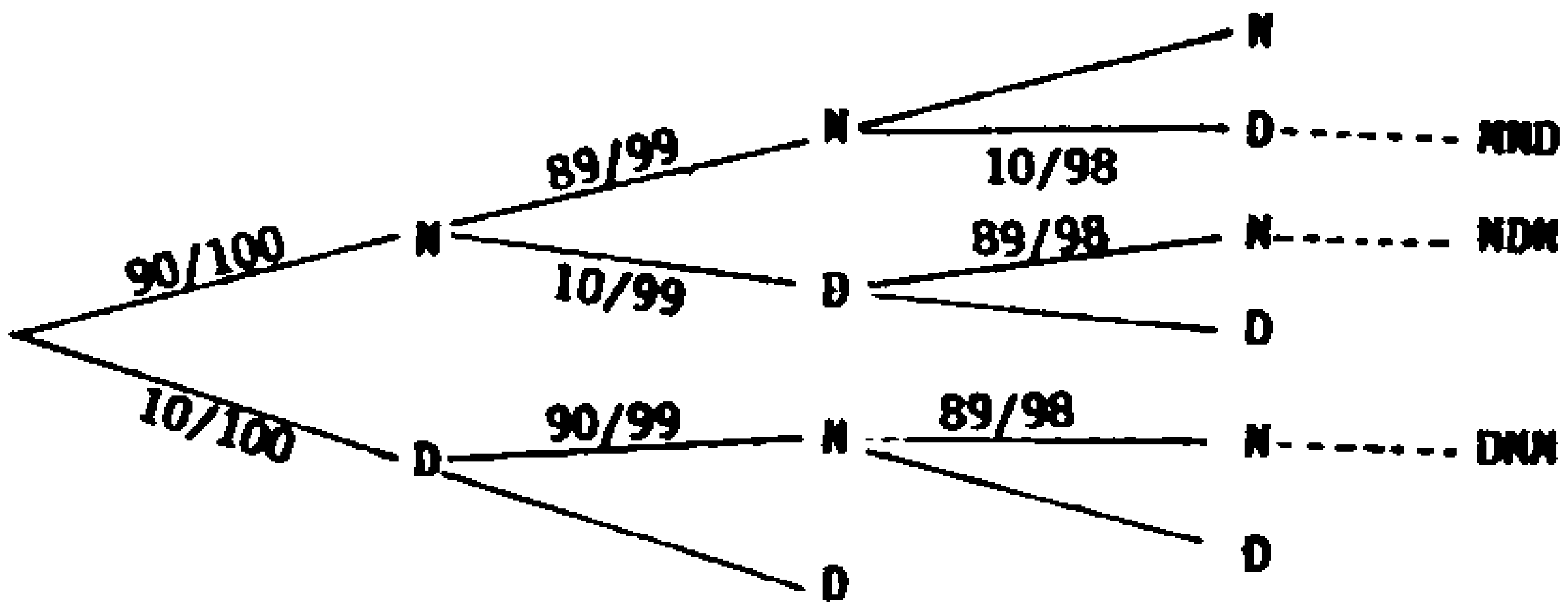


Fig. 1.7.13

$$P[A] = P[NND] + P[NDN] + P[DNN]$$

$$= \frac{90}{100}\cdot\frac{89}{99}\cdot\frac{10}{98} + \frac{90}{100}\cdot\frac{10}{99}\cdot\frac{89}{98} + \frac{10}{100}\cdot\frac{90}{99}\cdot\frac{89}{98}$$

$$= \frac{89}{11 \times 98} + \frac{89}{11 \times 98} + \frac{89}{11 \times 98} = \frac{267}{1078} = 0.2476\ .$$

SEGUNDO METODO El número de elementos del espacio muestral es $\binom{100}{3}$. Como el evento A contiene 1 defectuoso y 2 no defectuoso, entonces A tiene - $\binom{10}{1}\binom{90}{2}$ elementos. Luego, por la definición clásica es

$$P[A] = \frac{\binom{10}{1}\binom{90}{2}}{\binom{100}{3}} = \frac{10 \times \dfrac{90!}{88!\,2!}}{\dfrac{100!}{97!\ 3!}} = \frac{3 \times 89}{11 \times 98} = \frac{267}{1078} = 0.2476\ .$$

EJEMPLO 19 Un grupo que consta de 5 hombres y 10 mujeres se divide al azar en cinco grupos de tres personas cada uno. Calcular la probabilidad que en cada grupo haya un hombre.

SOLUCION Definimos los siguientes eventos:

A_i : "en el grupo i haya un hombre (i = 1,2,3,4,5)".

A : "en cada grupo de 3 personas haya un hombre".

El evento A se escribe así, $A = A_1A_2A_3A_4A_5$. Luego,

$$P[A] = P[A_1]\,P[A_2 \mid A_1]\,P[A_3 \mid A_1A_2]\,P[A_4 \mid A_1A_2A_3]\,P[A_5 \mid A_1A_2A_3A_4]$$

$$= \frac{\binom{5}{1}\binom{10}{2}}{\binom{15}{3}} \cdot \frac{\binom{4}{1}\binom{8}{2}}{\binom{12}{3}} \cdot \frac{\binom{3}{1}\binom{6}{2}}{\binom{9}{3}} \cdot \frac{\binom{2}{1}\binom{4}{2}}{\binom{6}{3}} \cdot \frac{1}{1}$$

$$= \frac{\frac{5!}{4!} \times \frac{10!}{8!2!}}{\frac{15!}{3!12!}} \times \frac{\frac{4!}{3!} \times \frac{8!}{2!6!}}{\frac{12!}{3!9!}} \times \frac{\frac{3!}{2!} \times \frac{6!}{2!4!}}{\frac{9!}{3!6!}} \times \frac{2! \times \frac{4!}{2!2!}}{\frac{6!}{3!3!}} \times 1$$

$$= \frac{5!10!3^5}{15!} = 0.081$$

EJEMPLO 20 En un cajón hay 80 tubos buenos y 20 malos; en un segundo cajón el 30% son malos y en un tercer cajón, el 25% son malos. Se sabe que el número de tubos del tercer cajón es el triple de los que hay en el segundo y en total hay 260 tubos. Se mezclan los tubos de las tres cajas.

(a) Al extraer, al azar, un tubo; calcule la probabilidad que sea malo, si se sabe que pertenece al segundo cajón.

(b) Al extraer, al azar, 2 tubos; calcule la probabilidad que el primero y el segundo sean malos.

SOLUCION Sea x el número de tubos en el segundo cajón, o sea

$$100 + x + 3x = 260 ,$$

de donde $x = 40$, y $3x = 120$.

Entonces, en la primera caja hay 100 tubos de los cuales 80 son buenos y 20 malos; en el segundo cajón hay 40 tubos de los cuales 28 son buenos y 12 malos; y en el tercer cajón hay 120 tubos de los cuales 90 son buenos y 30 malos.

(a) Sea D, el evento: "obtener un tubo defectuoso" y C : "el tubo pertenece al segundo cajón".

Luego,

$$P[D \mid C] = \frac{P[A\,C]}{P[C]} = \frac{12/260}{40/260} = \frac{12}{40} = 0.3 .$$

(b) Sea D, el evento: "obtener el primer y el segundo tubo defectuoso." Entonces,

$$D = D_1 D_2$$

$$P[D] = P[D_1]\,P[D_2 \mid D_1] = \frac{62}{260} \times \frac{61}{259} = \frac{3782}{67340} = 0.056 .$$

EJEMPLO 21 Una caja contiene 7 tarjetas marcadas "sin premio" y 5 con "premio mayor". En un concurso, dos personas A y B, extraen tarjetas de la caja en forma alternada hasta que una de ellas saca una marcada con el "premio mayor". Si A selecciona la tarjeta en primer lugar, ¿cuál es la probabilidad que extraiga una con "premio mayor"?

SOLUCION Definimos los siguientes eventos:

A_i : "el jugador A obtiene la tarjeta con premio mayor en su i-ésima jugada".

B_j : "el jugador B obtiene la tarjeta con premio mayor en su j-ésima jugada".

A_p : "el jugador A extrae una tarjeta con premio mayor".

El evento A_p se escribe, $A_p = A_1 \cup \bar{A}_1\bar{B}_1A_2 \cup \bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2A_3 \cup \bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2\bar{A}_3\bar{B}_3A_4$

Luego,

$$P[A_p] = P[A_1] + P[\bar{A}_1\bar{B}_1A_2] + P[\bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2A_3] + P[\bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2\bar{A}_3\bar{B}_3A_4]$$

$$= \frac{5}{12} + \frac{7}{12}\cdot\frac{6}{11}\cdot\frac{5}{10} + \frac{7}{12}\cdot\frac{6}{11}\cdot\frac{5}{10}\cdot\frac{4}{9}\cdot\frac{5}{8} + \frac{7}{12}\cdot\frac{6}{11}\cdot\frac{5}{10}\cdot\frac{4}{9}\cdot\frac{3}{8}\cdot\frac{2}{7}\cdot\frac{5}{6} = \frac{248}{12\times 11\times 3} = \frac{62}{99} = 0.63 \ .$$

El diagrama del árbol de probabilidad para este problema se muestra en la fig. 1.7.14 .

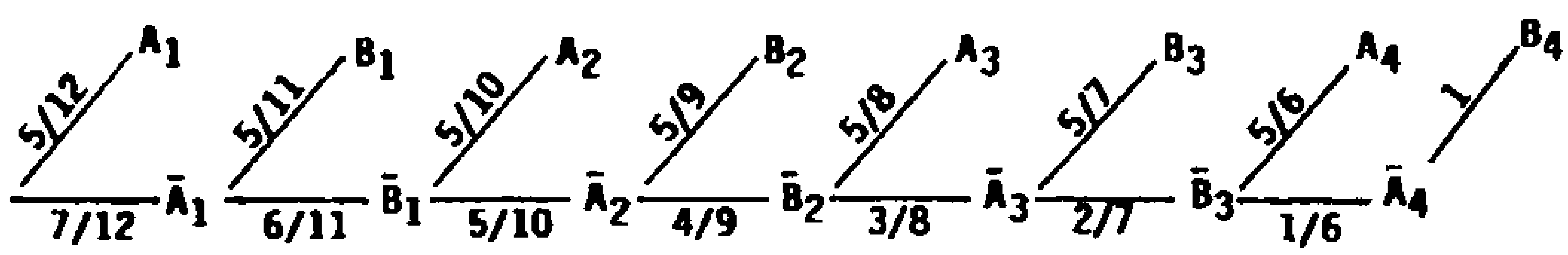


Fig. 1.7.14

EJEMPLO 22 Una urna contiene 10 bolas, 5 marcadas con la letra A y 5 bolas marcadas con la letra B, Dos jugadores, A y B juegan de la siguiente forma: comienza el jugador A extrayendo una bola y a continuación B realiza también una extracción, y asi alternadamente. Las extracciones se hacen sin reposición. Gana el primer jugador que extraiga una bola con su letra (A una bola A y B una bola B).

(a) ¿Cuál es la probabilidad que gane el jugador A? ¿Cuál de B?

(b) ¿Cuál es la probabilidad que no gane ninguno de los dos?

SOLUCION Definimos los siguientes eventos;

G_A "gana el jugador A"

G_B "gana el jugador B"

G_N "no gana ninguno de los jugadores"

A_K $(K = 1,2,\ldots)$: "En la K-ésima extracción se obtiene una bola marcada con A"

B_K $(K = 1,2,\ldots)$: "En la K-ésima extracción se obtiene una bola marcada con B"

Consideremos el diagrama del árbol de probabilidad de la fig. 1.7.15

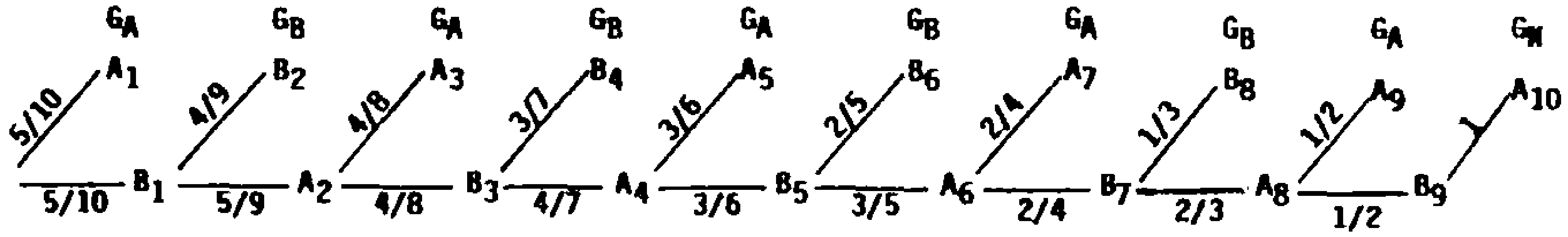


Fig. 1.7.15

(a) De este diagrama obtenemos,

$$P[G_A] = \frac{5}{10} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7} \times \frac{3}{6} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7} \times \frac{3}{6}$$

$$\times \frac{3}{5} \times \frac{2}{4} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{175}{252} .$$

$$P[G_B] = \frac{5}{10} \times \frac{4}{9} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7} \times \frac{3}{6} \times \frac{2}{5} + \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7}$$

$$\frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{76}{252} .$$

(b) No gana ninguno de los jugadores, cuando A saque todas las bolas marcadas con B y B todas las bolas marcadas con A, o sea

$$P[G_N] = \frac{5}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2} \times 1 = \frac{1}{252} .$$

PROBLEMAS 1.7

1. Dado, $P[A] = 0.5$ y $P[A \cup B] = 0.7$. Hallar $P[B]$, si $P[A \mid B] = 0.5$.

2. Si $P[B] = P[A \mid B] = P[C \mid AB] = \frac{1}{2}$. Hallar, $P[ABC]$.

3. Si, $P[A] = \frac{1}{2}$, $P[B] = \frac{1}{3}$, $P[AB] = \frac{1}{4}$. Hallar: $P[A \cup B]$, $P[A \mid B]$, $P[B \mid A]$, $P[A \cup B \mid B]$.

4. Si, $P[A \cup B] = \frac{11}{12}$, $P[A] = \frac{3}{4}$, $P[AB] = \frac{1}{2}$; hallar : $P[B]$, $P[A \mid B]$, $P[B \mid A]$.

5. Considere dos eventos A y B, tales que :

$$P[A] = \frac{1}{4}, \quad P[B \mid A] = \frac{1}{2}, \quad P[A \mid B] = \frac{1}{4}.$$ Diga, si cada uno de los incisos son verdaderos o falsos:

   (a) Los eventos A y B son mutuamente excluyentes.

   (b) A es subevento de B.

   (c) $P[\bar{A} \mid \bar{B}] = \frac{3}{4}$

   (d) $P[A \mid B] + P[A \mid \bar{B}] = 1$

6. Sea B un evento con probabilidad mayor que cero. Demostrar que para cualquier evento A.

   (a) $A \subset B$ implica $P[A \mid B] = \frac{P[A]}{P[B]}$

   (b) $B \subset A$ implica $P[A \mid B] = 1$

7. El evento A puede suceder, sólo si uno de dos eventos mutuamente excluyentes B_1 ó B_2 ocurre; esto es $A \subset B_1 \cup B_2$. Demuestre que $A = AB_1 \cup AB_2$ y exprese $P[A]$.

8. Si $P[A] > 0$ y $P[B] > 0$ ¿Son verdaderas las siguientes proposiciones? justifique su respuesta.

(a) Si $P[A] = P[B]$, entonces $P[A \mid B] = P[B \mid A]$

(b) Si $P[A \mid B] = P[B \mid A]$, entonces $P[A] = P[B]$.

9. Un aparato electrónico consta de dos partes. La probabilidad que falle la primera es 0.20, que fallen las dos partes es 0.15 y de que falle sólo la segunda parte es 0.45. Calcular la probabilidad que:
   (a) falle sólo la primera parte .
   (b) falle la primera parte cuando se sabe que falló la segunda.

10. Una urna contiene 7 bolas rojas y 3 blancas. Se extrae aleatoriamente tres bolas de la urna, sucesivamente sin reposición. Determinar la probabilidad que las dos primeras sean rojas y la tercera blanca.

11. En un lote de 20 televisores se sabe que hay 5 defectuosos. Se extrae al azar una muestra de tres televisores sin reposición. Hallar la probabilidad que la muestra contenga:
   (a) 0 defectuosos ; (b) 1 defectuoso
   (c) 2 defectuosos ; (d) 3 defectuosos .

12. Suponga que dos artefactos eléctricos defectuosos han sido incluidos en un embarque de seis artefactos eléctricos. El departamento de recepción de la compañia compradora empieza a probar los seis artefactos uno a uno. ¿Cuál es la probabilidad que:
   (a) el último artefacto defectuoso sea encontrado en la cuarta prueba?
   (b) a lo más, cuatro artefactos necesitan probarse para localizar los dos defectuosos?

13. Una caja contiene 5 fichas negras y 5 blancas. Si se extrae 5 fichas sin reposición. ¿Cuál es la probabilidad que la tercera ficha sea negra?

14. Se tienen 10 cartas numeradas de 1 a 10, las que se ponen en una caja. Se selecciona cuatro cartas aleatoriamente, sin reposición y se colocan en una mesa mostrando el número. Sea A_i el evento: "el número i ($1 < i < 10$) está entre estos seleccionados" (por ejemplo, A_1 {1,2,4,5} , {3,1,6,8}). Calcular $P[A_i \mid A_j]$.

15. Un cazador trata de matar un oso. La probabilidad que aparezca un oso en un radio menor que R_1 es de 0.1, en un radio entre R_1 y R_2 es de 0.3, y un radio mayor que R_2 es 0.2. Si aparece un oso en un radio menor que R_1, el cazador será capaz de matarlo con una probabilidad de

0.7; con una probabilidad de 0.5 si aparece en un radio entre R_1 y R_2 ; con una probabilidad de 0.2, si el radio es mayor de R_2. ¿Cuál es la probabilidad que el cazador mate un oso?

16. Se mezclan dos válvulas defectuosas con dos buenas. Se comienza a probar las válvulas una a una hasta que se descubren las defectuosas; ¿Cuál es la probabilidad que la segunda válvula defectuosa sea la segunda, la tercera, la cuarta probada?

17. Cuando se acerque el tren, un operador de la estación apretará un botón con una probabilidad de 0.95; si aprieta el botón, el interruptor opera con una probabilidad de 0.99; si el interruptor opera, sonará una alarma con una probabilidad de 0.9. ¿Cuál es la probabilidad que la alarma suene?

18. Considere tres urnas; la urna I contiene dos bolas blancas y cuatro rojas, la II contiene ocho bolas blancas y cuatro rojas, y la III contiene una bola blanca y tres rojas. Se selecciona una bola de cada urna. ¿Cuál es la probabilidad que la bola seleccionada de la urna II sea blanca, dado que la muestra contiene exactamente dos bolas blancas?

19. Suponga que cierta factoría produce tres productos designados por A, B y C. ¿Cuál es la probabilidad de seleccionar un producto defectuoso A, si se sabe que el 30% de los productos producidos en la factoría son productos A y 5% de los productos A son defectuosos?

20. Suponga que tiene tres colecciones de eventos

$\{A_1, A_2, \ldots, A_k\}$, $\{B_1, B_2, \ldots, B_\ell\}$, $\{C_1, C_2, \ldots, C_m\}$, tales que los eventos $A_1, A_2, \ldots, A_k$, son mutuamente excluyentes y colectivamente exhaustivos.

Los eventos $B_1, B_2, \ldots, B_\ell$, son mutuamente excluyentes y colectivamente exhaustivos.

Los eventos $C_1, C_2, \ldots, C_m$. son mutuamente excluyentes y $\bigcup_{i=1}^{m} C_i \neq \Omega$ (es decir no son colectivamente exhaustivos).

Evalúe numéricamente cada una de las siguientes cantidades. Si no puede evaluar numéricamente especifique una cota superior e inferior numéricamente.

(a) $\sum_{i=1}^{m} P[C_i]$; (b) $\sum_{i=1}^{k} P[A_2 \mid A_j]$

(c) $\sum_{i=1}^{k} \sum_{j=1}^{k} P[A_i \cap A_j]$; (d) $\sum_{j=1}^{k} P[C_2 \mid A_j C_2]$

(e) $\sum_{j=1}^{k} P[\bar{A}_j]$; (f) $\sum_{i=1}^{k} \sum_{j=1}^{\ell} P[A_i] P[B_j \mid A_i]$

21. Un experimento consiste en lanzar dos dados uno o dos veces. Un jugador gana si consigue la suma 7 en el primer lanzamiento; pierde si saca 2 ó 12; si consigue otras sumas no pierde nada, en este último caso tiene opción para un segundo lanzamiento y si este segundo lanzamiento consigue la suma 7 pierde, en caso contrario gana y termina el juego. Cuál es la probabilidad que el jugador pierda?

22. Un jugador gana si saca 7 u 11 en el primer lanzamiento de un par de dados buenos, pierde si saca 2, 3, ó 12 en el primer lanzamiento. Sin embargo, si en el primer lanzamiento saca un 4,5,6,8,9 ó 10 continúa tirando el dado hasta obtener el número que obtuvo en el primer lanzamiento o hasta obtener un 7 . Si obtiene su primer número antes de obtener un 7, gana; en otro caso pierde. Calcular la probabilidad que el jugador gana en dos o menos lanzamientos.

23. En una ciudad, el 70% de los adultos escuchan radio; el 40% lee el periódico y el 10% ve televisión, entre los que escuchan radio el 30% lee los periódicos y el 4% ve T.V. El 90% de los que ven T.V. lee el periódico, y sólo el 2% de la población total lee el periódico, ve televisión y escucha radio. Si se elige una persona al azar, calcule ud. la probabilidad.
(a) de que lea el periódico; escuche radio o vea televisión.
(b) sabiendo que lee el periódico, de que vea televisión.

24. En una urna hay 8 bolas numeradas de 1 a 8, se extraen al azar y sucesivamente 3 bolas.
(a) ¿Cuál es la probabilidad que sean las tres pares?
(b) ¿Cuál es la probabilidad que sean tres números consecutivos?

25. Una urna contiene 3 bolas blancas y 5 bolas negras. Se extraen al azar y sucesivamente dos bolas. Sabiendo que la primera bola extraída es blanca; ¿cuál es la probabilidad que la segunda bola estraída también sea blanca?

26. Una urna contiene 4 bolas marcadas con la letra A y 4 bolas marcadas con la letra B. Dos jugadores, A y B juegan de la siguiente forma: comienza A extrayendo una bola continúa B realizando dos extracciones de a una por vez y así alternativamente (A una extracción, B dos extracciones) hasta agotar las bolas de la urna. Gana el primer jugador que extraiga una bola con su letra (A una bola A; B una bola B).
(a) ¿Cuál es la probabilidad que gane el jugador B?
(b) ¿Cuál es la probabilidad que gane el jugador A?

27. Se tiene cinco urnas numeradas de 1 a 5 y cinco bolas numeradas de 1 a 5. Se coloca al azar una bolilla en cada urna. ¿Cuál es la probabilidad de que.
(a) al menos una bola sea colocado en la urna que tiene su número?
(b) Exactamente una bola sea colocado en la urna que tiene su mismo número?.

28. Tres jugadores A, B y C extraen (en ese orden) bolas de una urna, en donde hay 10 bolas blancas y 10 negras, tomando una cada vez. El primero que obtiene una bola blanca gana el juego. Hallar la probabilidad de ganar de cada uno de los jugadores.

29. Al tirar tres dados, sucesivamente, determine la probabilidad que aparezca una sucesión creciente de números consecutivos.

30. Con referencia al problema 34 de 1.6. Suponga que se sabe que el cliente ha tenido cuenta de crédito cuando menos cinco años, ¿cuál es la probabilidad que tenga un saldo inferior a I/.1000?

31. En el problema 35 de 1.6. Suponga que se encuentra con una empleada de la compañía, ¿cuál es la probabilidad que no fume cigarrillos?

32. Tres jugadores A, B y C extraen aleatoriamente cada uno una bola de una urna que contiene doce, de las cuales ocho son negras y cuatro son blancas, hasta que uno de ellos (que será el ganador) saque la primera bola blanca. Hallar la probabilidad de ganar de cada jugador, sabiendo que -

empieza A, que los otros siguen en el orden indicado y que la extracción es sin reposición.

33. El departamento de crédito de la Cooperativa la Tacaña sabe por experiencia que la probabilidad de que un acreedor deje de pagar un préstamo es de 0.04. También se encontró que dado un incumplimiento de pago de préstamo hay una probabilidad de 0.40 de que se pidiera el préstamo para salir de vacaciones. Además, la Cooperativa sabe que la probabilidad de incumplimiento es la misma para empleados estatales que para el resto de la población.
    (a) ¿Cuál es la probabilidad que un prestatario pida prestado para financiar sus vacaciones y luego no cumpla?
    (b) Si la probabilidad de que se haga un préstamo a un empleado es de 0.02, ¿cuál es la probabilidad que un prestatario sea empleado estatal y no cumpla con el pago?

34. En el otoño, la probabilidad que un día lluvioso seguirá a otro día lluvioso es 0.80 y la probabilidad que un día soleado seguirá a otro día soleado es 0.40. Suponer que cada día es clasificado como lluvioso o soleado y que el tiempo de cualquier día depende sólo del tiempo del día anterior, encontrar la probabilidad que un día lluvioso es seguido por tres días lluviosos, después por dos días soleados y, finalmente por otro día lluvioso.

1.8 TEOREMA DE BAYES

Estableceremos en esta sección, el teorema más importante de este capítulo y el camino para llegar a él exige la definición de partición de un espacio muestral y el teorema de probabilidad total.

1.8.1 PARTICION DE UN ESPACIO MUESTRAL

DEFINICION 1.8.1 Se dice que la colección de eventos $B_1, B_2, \ldots, B_k$ del espacio muestral Ω representa una partición del espacio muestral Ω, si cumple las siguientes condiciones:

(a) los eventos $B_1, B_2, \ldots, B_k$ son mutuamente excluyentes.

En símbolos $B_i \cap B_j = \phi, \quad i \neq j, \quad i,j = 1,2,3, \ldots, k$

(b) los eventos $B_1, B_2, \ldots, B_k$ son colectivamente exhaustivos. En símbolo

$$\bigcup_{i=1}^{k} B_i = \Omega$$

(c) $P[B_i] > 0$,

$i = 1,2, \ldots , k$

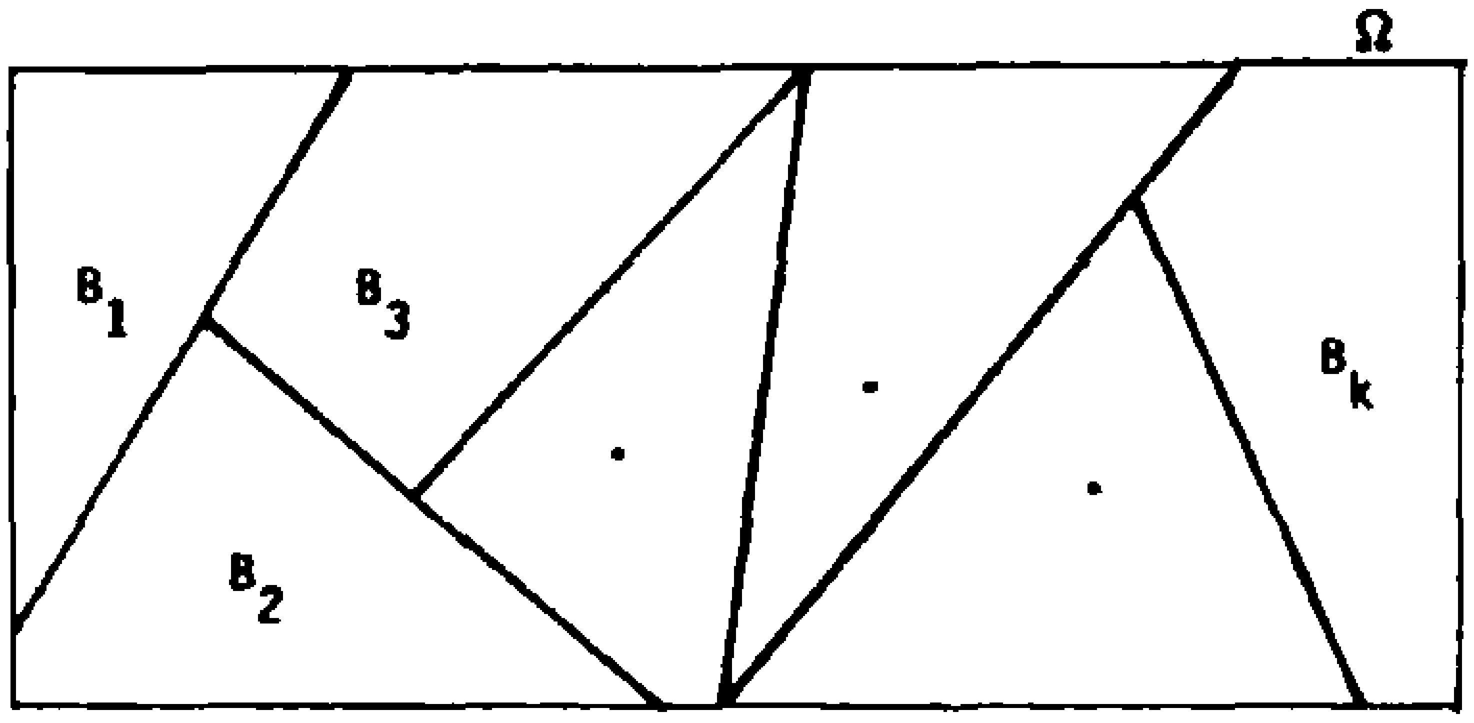


Fig. 1.8.1. Partición del espacio muestral

EJEMPLO 1 En el lanzamiento de un dado, $\Omega = \{1,2,3,4,5,6\}$ Si $B_1 = \{1,2\}$, $B_2 = \{3,4,5\}$, $B_3 = \{6\}$

B_1, B_2, B_3 representa una partición del espacio muestral Ω.

En cambio, si :

$$E_1 = \{1,2\}, \quad E_2 = \{4\}, \quad E_3 = \{3,6\}$$

$E_1 \cup E_2 \cup E_3 \neq \Omega$, por lo tanto E_1, E_2 y E_3 no representan una partición de Ω .

1.8.2 TEOREMA DE PROBABILIDAD TOTAL

TEOREMA 1.8.1 Sea $B_1, B_2, \ldots, B_k$ una partición del espacio muestral Ω, entonces para cualquier evento A en Ω , se cumple

$$P[A] = \sum_{i=1}^{k} P[B_i] P[A \mid B_i] = P[B_1] P[A \mid B_1] + P[B_2] P[A \mid B_2] + \ldots + P[B_k] P[A \mid B_k] .$$

DEMOSTRACION 1 Sabemos que

$\Omega = B_1 \cup B_2 \cup \ldots \cup B_k$, hipotesis

2. Para cualquier evento A en Ω se tiene

$A = A \cap \Omega$

$A = A \cap (B_1 \cup B_2 \cup \ldots \cup B_k)$ (ver fig. 1.8.2)

$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)$

3. Los eventos, $A \cap B_i$ y $A \cap B_j$, $i \neq j$ son mutuamente excluyentes, - pués

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = \phi, \; i \neq j, \; i, j = 1,2, \ldots, k$$

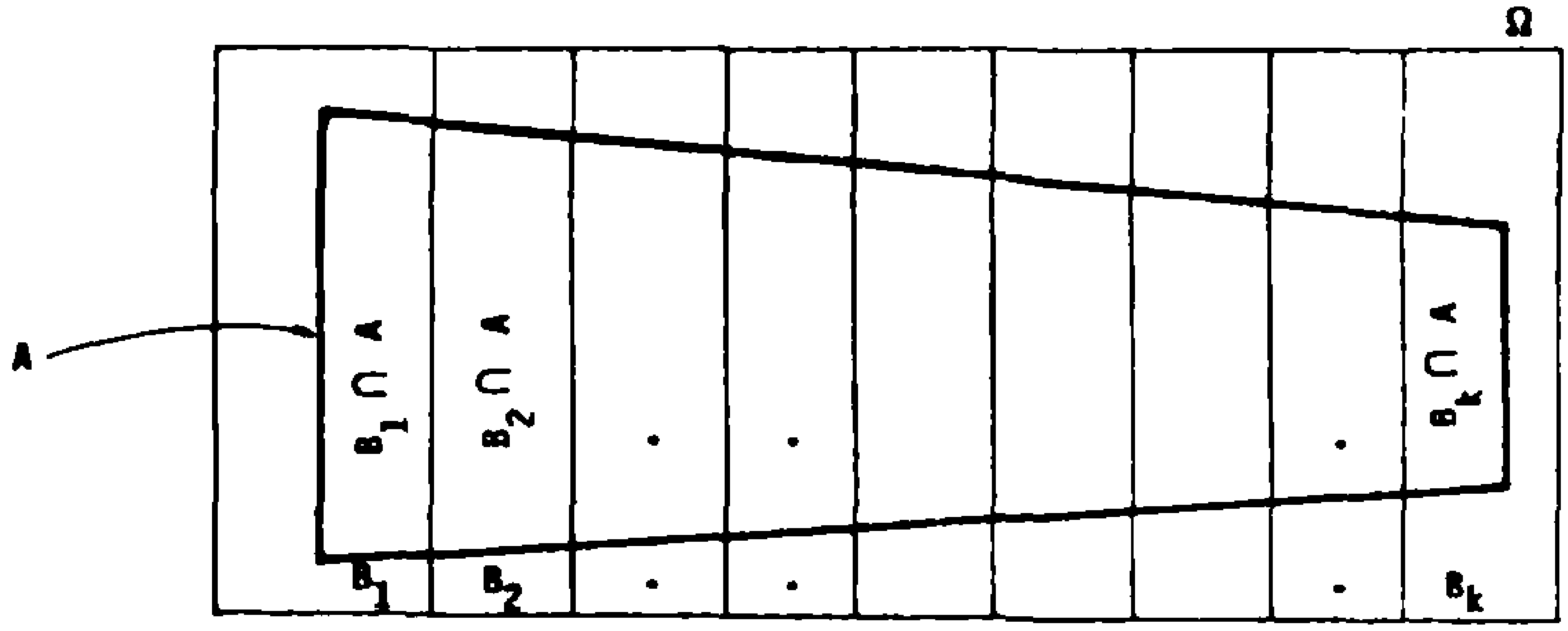


Fig. 1.8.2. Relación entre el evento A en Ω y la partición de Ω .

4. Tomando probabilidades a ambos miembros de la igualdad del paso (2) tenemos,

$$P[A] = P[A \cap B_1] + P[A \cap B_2] + \dots + P[A \cap B_k]$$

$$= P[B_1]\, P[A \mid B_1] + P[B_2]\, P[A \mid B_2] + \dots + P[B_k]\, P[A \mid B_k]$$

$$P[A] = \sum_{i=1}^{k} P[B_i]\, P[A \mid B_i] \ .$$

El diagrama del árbol de probabilidad de la fig. 1.8.3 para experimentos sucesivos da una visión esquemática del teorema de probabilidad total.

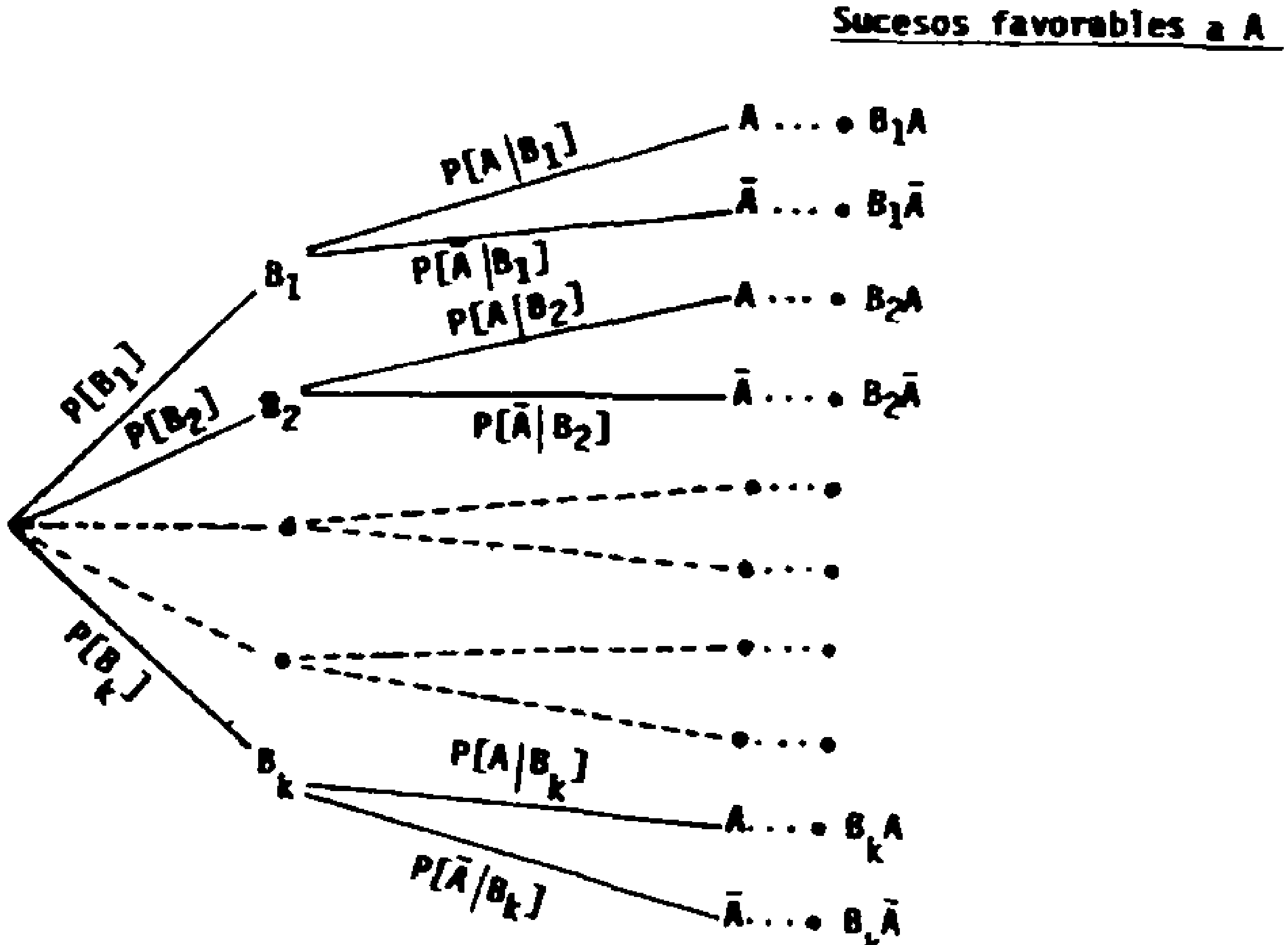


Fig. 1.8.3

de donde

$$P[A] = P[\bigcup_{i=1}^{k} B_i A] = \sum_{i=1}^{k} P[B_i]P[A \mid B_i] .$$

COROLARIO 1 Si B es un evento en Ω tal que $0 < P[B] < 1$, entonces para cualquier evento A en Ω

$$P[A] = P[B]\,P[A \mid B] + P[\bar{B}]\,P[A \mid \bar{B}]$$

DEMOSTRACION 1 Los eventos B y $\bar{B}$ forman una portación de Ω y $\Omega = B \cup \bar{B}$ (por hipótesis)

2. Para cualquier evento A en Ω se tiene

$$A = A \cap \Omega = A\Omega$$

ó $\quad A = A[B \cup \bar{B}] = AB \cup A\bar{B}$ (ver fig. 1.8.4)

3. $AB \cap A\bar{B} = A(B \cap \bar{B}) = \phi$. Es decir, los eventos A B y $A\bar{B}$ son mutuamente excluyentes.

4. Aplicando probabilidad a ambos miembros de la igualdad obtenida en el paso (2) se tiene

$$P[A] = P[AB \cup A\bar{B}]$$
$$= P[AB] + P[A\bar{B}]$$
$$= P[B]\,P[A \mid B] + P[\bar{B}]\,P[A \mid \bar{B}]$$

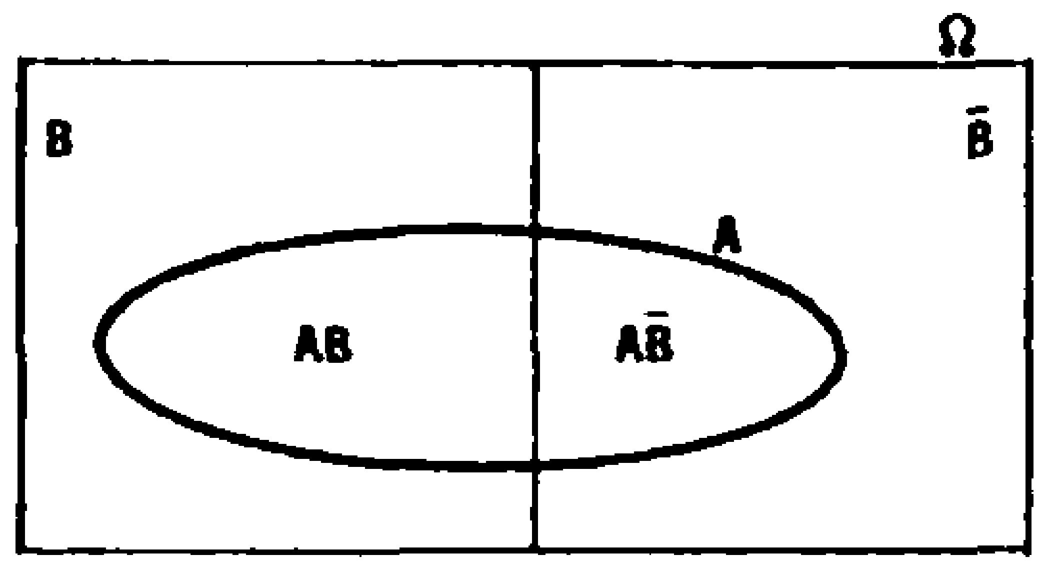


Fig. 1.8.4

El diagrama del árbol de probabilidad de la fig. 1.8.5 para experimentos sucesivos, muestra una visión esquemática del corolario 1.

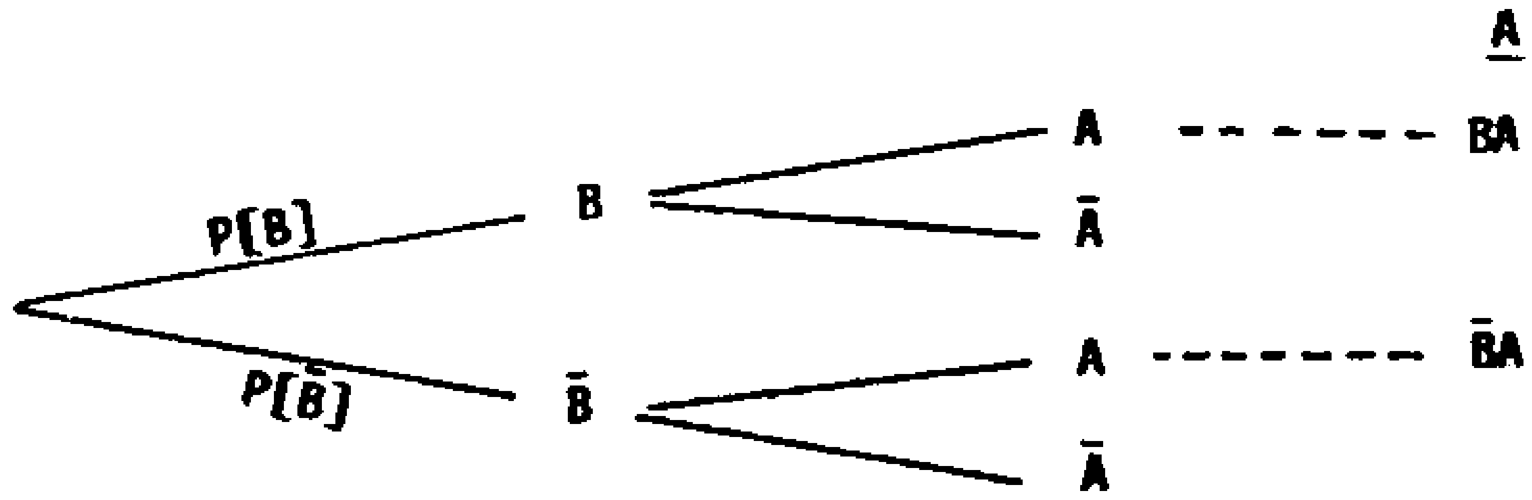


Fig. 1.8.5

EJEMPLO 2 En un laboratorio hay tres jaulas. En la jaula I hay dos conejos pardos y tres blancos, la jaula II tiene 4 conejos pardos y dos blancos y la jaula III contiene 5 conejos pardos y 5 blancos. Se selecciona al azar una jaula y se saca un conejo al azar de esta jaula. ¿Cuál es la probabilidad que el conejo escogido sea blanco?

SOLUCION 1 Definimos los siguientes eventos:

I: "la jaula I es seleccionada"

II: "la jaula II es seleccionada"

III: "la jaula III es seleccionada"

A: "el conejo escogido es blanco".

2. El espacio muestral está constituida por los conejos de las tres jaulas y estas formas una partición del espacio muestral, es decir

$$\Omega = I \cup II \cup III$$

3. $A \subset \Omega$ y se escribe $A = IA \cup IIA \cup IIIA$. Luego, por el teorema - probabilidad total

$$P[A] = P[I]\,P[A|I] + P[II]\,P[A \mid II] + P[III]\,P[A \mid III]$$

4. Puesto que, se escoge una jaula al azar, las tres son igualmente posibles, o sea

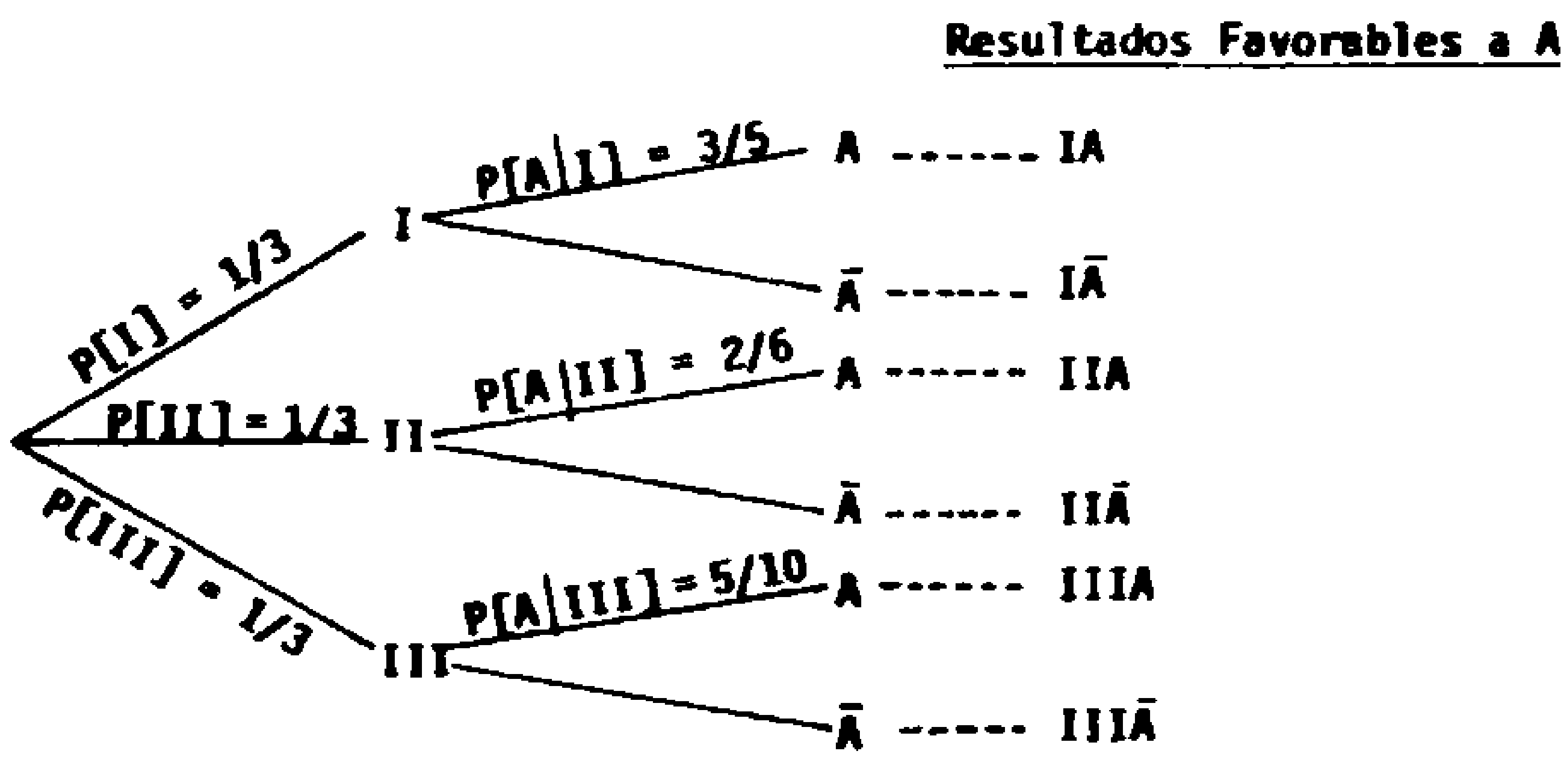


Fig. 1.8.6 Diagrama del árbol de probabilidad para experimentos sucesivos del ejemplo 2.

$$P[I] = P[II] = P[III] = \frac{1}{3}$$

5. Si se selecciona la jaula I, $P[A \mid I] = \frac{3}{5}$. Si se sáca la jaula II , $P[A \mid II] = \frac{2}{6}$. Finalmente si se selecciona la jaula III, $P[A \mid III]$ - $= \frac{5}{10}$. Reemplazando estos valores en el paso (3) se tiene

$$P[A] = \frac{1}{3} \cdot \frac{3}{5} + \frac{1}{3} \cdot \frac{2}{6} + \frac{1}{3} \cdot \frac{5}{10}$$

$$= \frac{1}{5} + \frac{1}{9} + \frac{5}{30} = \frac{43}{90} .$$

EJEMPLO 3 Del record pasado, se conoce que cierta máquina que produce tornillos trabaja correctamente el 90% del tiempo. Si la máquina no está trabajando correctamente, el 5% de los tornillos producidos son defectuosos. Cuando está trabajando bien solamente el 0.5% de los tornillos son defectuosos. Si se escoge un tornillo aleatoriamente, ¿cuál es la probabilidad que sea defectuoso?

SOLUCION Sean los siguientes eventos:

D: "el tornillo es defectuoso".

M: "la máquina está trabajando correctamente".

$$\Omega = M \cup \bar{M} \qquad y \qquad D = MD \cup \bar{M}D$$

Aplicando el corolario del teorema de probabilidad total

$$P[D] = P[M]\, P[D \mid M] + P[\bar{M}]\, P[D \mid \bar{M}]$$

$$= (0.9)(0.005) + (0.01)(0.05) = 0.0095.$$

El esquema del árbol de probabilidad que visualiza este problema se muestra en la fig. 1.8.7 .

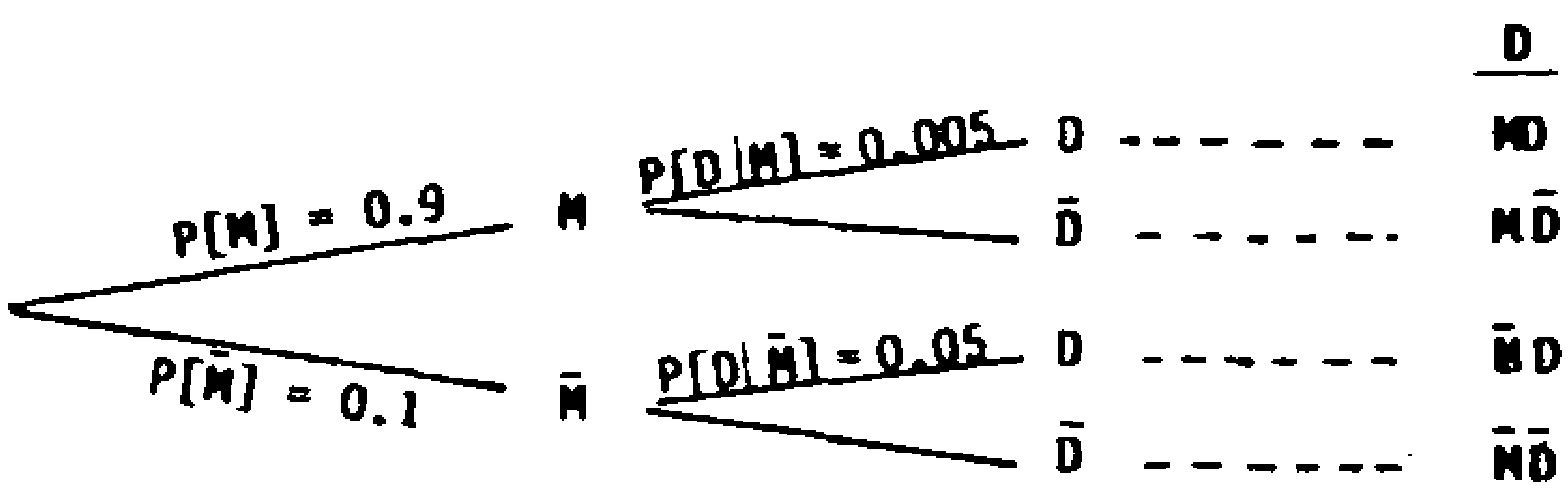


Fig. 1.8.7

EJEMPLO 4 Una urna contiene 3 bolas rojas y x blancas. Se extrae una bola de la urna y se reemplaza por una del otro color. Se saca de la urna una segunda bola. Sabiendo que la probabilidad de que la segunda bola sea roja es $\frac{17}{50}$. Determinar el número de bolas blancas .

SOLUCION Sean los siguientes eventos

R: "la segunda bola es roja"

r: "la bola extraída es roja"

b: "la bola extraída es blanca"

Los experimentos sucesivos del problema se lleva a un diagrama del árbol - de probabilidad obtenida la fig. 1.8.8, donde vemos que el evento R se escribe $R = rr \cup br$, por lo tanto

$$P[R] = P[r] \times P[r \mid r] + P[b] \times P[r \mid b]$$

$$= \frac{3}{x+3} \times \frac{2}{x+3} + \frac{x}{x+3} \times \frac{4}{x+3} = \frac{4x+6}{(x+3)^2} = \frac{17}{50}$$

de donde se obtiene la ecuación de segundo grado

$$17x^2 - 98x - 147 = 0$$

resolviendo la ecuación se obtiene $x = 7$ bolas blancas.

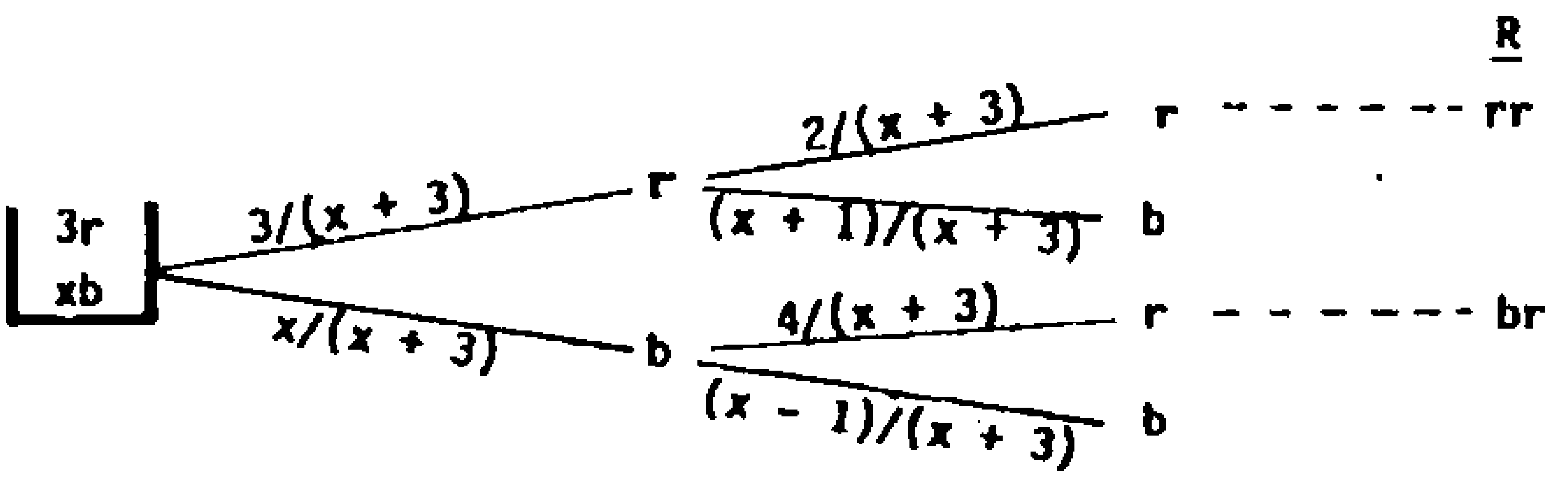


Fig. 1.8.8

EJEMPLO 5 Juan escoge al azar uno de los enteros 1,2,3 y luego lanza un - dado tantas veces como indica el número que escogió. Calcular la probabilidad que el puntaje total obtenido en los lanzamientos sea igual a 5.

SOLUCION PRIMER METODO Llevando directamente los experimentos sucesivos - del enunciado del problema al diagrama del árbol de probabilidades se obtiene la fig. 1.8.9

Definimos los siguientes eventos

i : obtener el entero i , i = 1,2,3". $P[\{i\}] = \frac{1}{3}$, $i = 1,2,3$.

A : "Obtener la suma 5 al lanzar el dado".

El evento **A** se escribe como la unión de las intersecciones de los i con suma 5.

Luego, por el teorema de probabilidad total se obtiene

$$P[A] = \frac{1}{3} \times \frac{1}{6} + \frac{1}{3} \times \frac{1}{36} + \frac{1}{3} \times \frac{1}{36} + \frac{1}{3} \times \frac{1}{36} + \frac{1}{3} \times \frac{1}{36} + \frac{1}{3} \times \frac{1}{6^3} + \frac{1}{3} \times \frac{1}{6^3}$$

$$+ \frac{1}{3} \times \frac{1}{6^3} + \frac{1}{3} \times \frac{1}{6^3} + \frac{1}{3} \times \frac{1}{6^3} + \frac{1}{3} \times \frac{1}{6^3}$$

$$= \frac{1}{3} \times \frac{6}{36} + \frac{1}{3} \times \frac{4}{36} + \frac{1}{3} \times \frac{1}{36} = \frac{11}{108} .$$

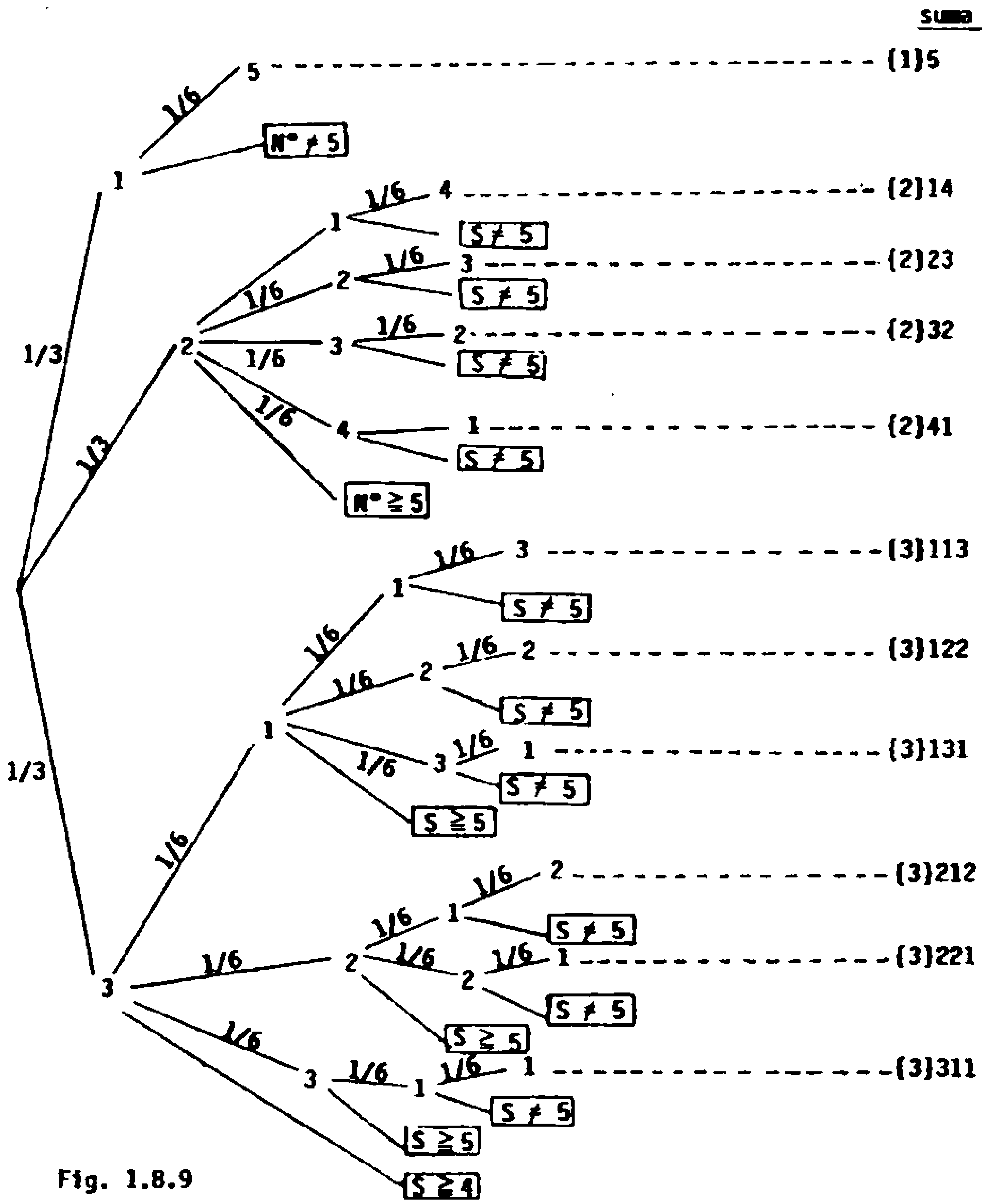


Fig. 1.8.9

SEGUNDO METODO Observe que estamos frente a un experimento ε que consiste en realizar ε_1 ó ε_2 ó ε_3 en el sentido excluyente, donde ε_i consiste en lanzar i dados (i = 1,2,3). El espacio muestral asociado a ε es como sigue

$$\Omega = \left\{ [1, 2, 3, 4, 5, 6], \begin{bmatrix} 11 & 12 & 13 & 14 & 15 & 16 \\ 21 & 22 & 23 & 24 & 25 & 26 \\ 31 & 32 & 33 & 34 & 35 & 36 \\ 41 & 42 & 43 & 44 & 45 & 46 \\ 51 & 52 & 53 & 54 & 55 & 56 \\ 61 & 62 & 63 & 64 & 65 & 66 \end{bmatrix}, \begin{bmatrix} 111 & 112 & 121 & 122 & 113 \\ 211 & 221 & 212 & \ldots & 266 \\ 311 & 131 & 321 & \ldots & \ldots \\ . & & & & \\ . & & & & \\ . & \ldots & \ldots & \ldots & 666 \end{bmatrix} \right\}$$

Los tres bloques definen los eventos E_1, E_2, E_3 que representan una partición de Ω. Definimos el evento,

A : "obtener suma 5 con el dado".

Observe que $A = E_1A \cup E_2A \cup E_3A$. Luego, por el teorema de probabilidad total

$$P[A] = P[E_1]\,P[A \mid E_1] + P[E_2]\,P[A \mid E_2] + P[E_3]\,P[A \mid E_3]$$

$$= \frac{1}{3} \times \frac{1}{6} + \frac{1}{3} \times \frac{4}{36} + \frac{1}{3} \times \frac{6}{216}$$

$$= \frac{6}{108} + \frac{4}{108} + \frac{1}{108} = \frac{11}{108}\,.$$

EJEMPLO 6 Se tiene una caja con doce tarros de conserva de las cuales ocho son de durazno. Se extrae una muestra con reemplazo (sin reemplazo) de tamaño 4. Después se selecciona una conserva de la muestra. Determine la probabilidad que sea de durazno.

SOLUCION Definimos los siguientes eventos :

M_i : "la muestra contiene i tarros de conserva de duraznos $i = 0,1,2,3,4$".

D : "la conserva es de durazno"

Para que pueda obtenerse de la muestra un tarro de durazno, está debe contar por lo menos una conserva de durazno. Entonces, el evento D se escribe

$$D = \bigcup_{i=1}^{4} M_iD \quad . \qquad \text{Luego}$$

$$P[D] = \sum_{i=1}^{4} P[M_i]\,P[D \mid M_i] \qquad (*)$$

(a) **Extracción con reemplazamiento**

De las 12 conservas se extrae una a una con reemplazo 4 tarros, entonces $N(\Omega) = 12^4$.

Cálculo de $P[M_i]$ y $P[D \mid M_i]$, $i = 1,2,3,4$.

1. M_1 : "la muestra contiene 1 tarro de durazno"

Los elementos de M_1 tienen la forma $\underbrace{D\,\bar{D}\,\bar{D}\,\bar{D}}_{P_4^{3,1}}$

D ocurre de 8 formas y para cada uno de estos, los tres $\bar{D}$ ocurren de 4^3 formas, luego $N(M_1) = P_4^{3,1} \times 8 \times 4^3$. Por lo tanto

$$P[M_1] = \frac{P_4^{3,1} \cdot 8 \times 4^3}{12^4}.$$

Si ocurrió M_1, $P[D \mid M_1] = \frac{1}{4}$

2. M_2: "la muestra contiene 2 tarros de durazno".
Los elementos de M_2 tienen la forma $\underbrace{D\ D\ \bar{D}\ \bar{D}}_{P_4^{2,2}}$

Las dos D ocurren de 8^2 formas y para cada una de estas, las dos $\bar{D}$ - ocurren de 4^2 formas, es decir $N(M_2) = P_4^{2,2} \times 8^2 \times 4^2$, por lo tanto

$$P[M_2] = \frac{P_4^{2,2} \times 8^2 \times 4^2}{12^4}$$

Ahora, si ocurrió M_2, $P[D \mid M_2] = \frac{1}{2}$.

3. M_3: "la muestra contiene 3 tarros de durazno"
Los elementos de M_2 tienen la forma $\underbrace{D\ D\ D\ \bar{D}}_{P_4^{3,1}}$

Las tres D ocurren de 8^3 formas y para cada una de estas, la $\bar{D}$ ocurre de 4 formas, osea $N(M_3) = P_4^{3,1} \times 8^3 \times 4$. Luego

$$P[M_3] = \frac{P_4^{3,1} \times 8^3 \times 4}{12^4}$$

Si ocurrió M_3 , $P[D \mid M_3] = \frac{3}{4}$.

4. M_4: "la muestra contiene 4 tarros de durazno". Es decir, los cuatro elementos de la muestra son tarros de durazno (D D D D).
Estas cuatro D ocurren de 8^4 formas. Osea $N(M_4) = 8^4$. Por lo tanto

$$P[M_4] = \frac{8^4}{12^4}$$

Si ocurrió M_4, $P[D \mid M_4] = 1$.
El digrama del árbol de probabilidades que resume los pases (1), (2), (3), y (4) se muestra en la fig. 1.8.10.

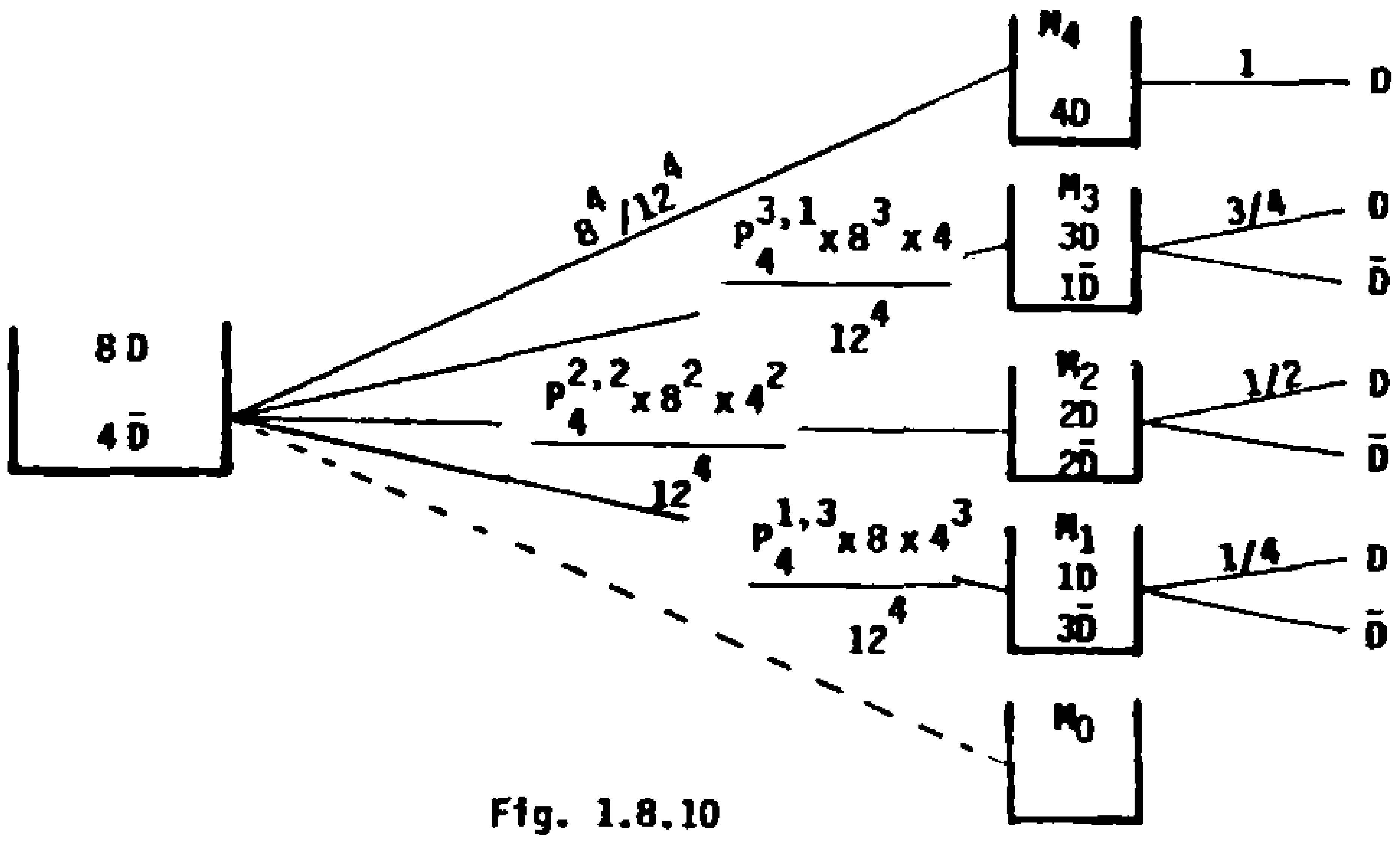


Fig. 1.8.10

Reemplazando (1), (2), (3) y (4) en (*) obtenemos

$$P[D] = \frac{P_4^{1,3} \times 8 \times 4^3}{12^4} \times \frac{1}{4} + \frac{P_4^{2,2} \times 8^2 \times 4^2}{12^4} \times \frac{1}{2} + \frac{P_4^{3,1} \times 8^3 \times 4}{12^4} \times \frac{3}{4} + \frac{8^4}{12^4} \times 1$$

$$= \frac{8^3}{12^4} + \frac{6 \times 8^3}{12^4} + \frac{12 \times 8^3}{12^4} + \frac{8^4}{12^4}$$

$$= \frac{8^3}{12^4}\,[\,1 + 6 + 12 + 8\,] = \left(\frac{8}{12}\right)^3 \times \frac{27}{12}$$

$$= \frac{2}{3}\,.$$

(b) **Extracción sin reemplazamiento.**

De las 12 conservas de la caja se extrae una muestra sin reemplazo de 4 tarros, entonces $N(\Omega) = C_{12}^4$.

Cálculo de $P[M_i]$ y $P[D \mid M_i]$. $i = 1,2,3,4$.

1. M_1: "la muestra contiene 1 tarro de durazno".

Una D ocurre de $C_8^1 = 8$ formas y para cada una de éstas, las tres $\bar{D}$ ocurren de C_4^3 . Osea $N(M_1) = C_8^1 C_4^3$. Por lo tanto

$$P[M_1] = \frac{C_8^1 C_4^3}{C_{12}^4}\,.$$

Si ocurrió M_1, $P[D \mid M_1] = \frac{1}{4}$.

2. M_2: "la muestra contiene 2 tarros de durazno".

Las dos D ocurren de C_8^2 y para cada una de éstas, las dos $\bar{D}$ ocurren de de C_4^2. Es decir $N(M_2) = C_8^2 C_4^2$. Entonces

$P[M_2] = \dfrac{C_8^2 C_4^2}{C_{12}^4}$. Si ocurrió M_2 , $P[D \mid M_2] = \dfrac{1}{2}$.

3. M_3: "la muestra contiene 3 tarros de durazno"
   Las tres D ocurren de C_8^3 y para cada una de éstas, las $\bar{D}$ ocurre de - $C_4^1 = 4$ formas. Osea $N(M_3) = C_8^3 C_4^1$. Luego $P[M_3] = \dfrac{C_8^3 C_4^1}{C_{12}^4}$

   Si ocurrió M_3 , $P[D \mid M_3] = \dfrac{3}{4}$.

4. M_4: "la muestra contiene 3 tarros de durazno".
   Las cuatro D ocurren de C_8^4 formas. Es decir $N(M_4) = C_8^4$.
   Por lo tanto

$$P[M_4] = \frac{C_{12}^4}{C_{12}^4}$$

   Si ocurrió M_4, $P[D \mid M_4] = 1$.

   Reemplazando (1) , (2) , (3) y (4) en (*) obtenemos,

$$P[D] = \frac{C_8^1 C_4^3}{C_{12}^4} \times \frac{1}{4} + \frac{C_8^2 C_4^2}{C_{12}^4} \times \frac{1}{2} + \frac{C_8^3 C_4^1}{C_{12}^4} \times \frac{3}{4} + \frac{C_8^4}{C_{12}^4}$$

$$= \frac{70 + 168 + 84 + 8}{495} = \frac{330}{945} = \frac{2}{3} .$$

El diagrama del árbol de probabilidad que resume los pasos (1),(2),(3) y (4) se muestra en la fig. 1.8.11.

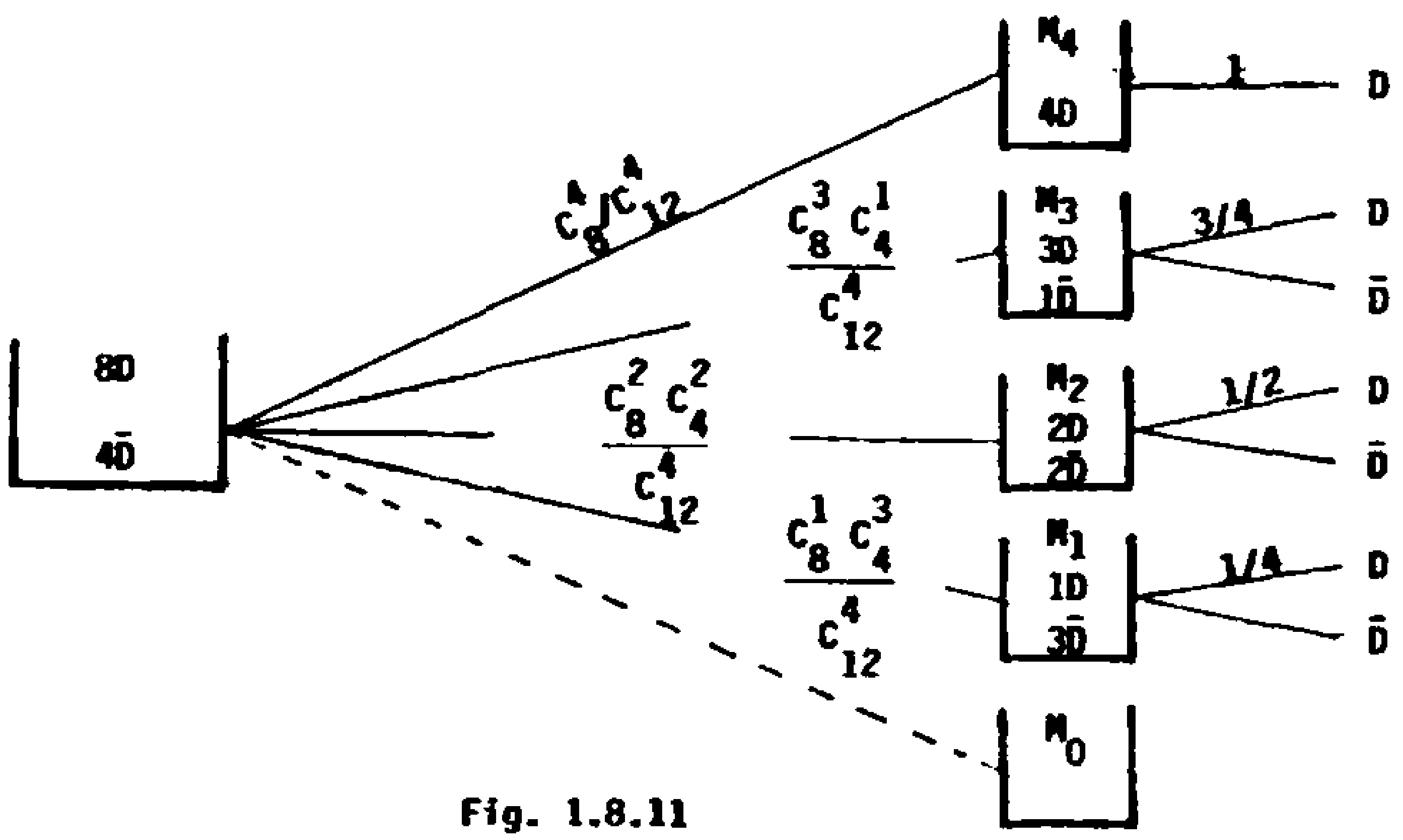


Fig. 1.8.11

EJEMPLO 7 En el bolsillo derecho de su casaca Ud. tiene 3 monedas de 100 Intis y cuatro de 50 intis; en el bolsillo izquierdo tiene 6 monedas de 100 intis y 3 de 50 intis. Tome aleatoriamente 5 monedas del bolsillo derecho y pase al izquierdo. Luego extraiga al azar una moneda del bolsillo izquierdo. - Determinar la probabilidad que sea de 100 intis.

SOLUCION El primer experimento que se realiza es extraer 5 monedas del bolsillo derecho, la cual se hace de C_7^5 formas.

Definimos ahora los siguientes eventos:

B_i: "Se pasan i (i = 1,2,3) monedas de 100 del bolsillo derecho al izquierdo".

M: "Obtener una moneda de 100 intis al extraer una del bolsillo izquierdo".

Entonces el evento M se escribe

$$M = B_1M \cup B_2M \cup B_3M$$

por lo tanto

$$P[M] = P[B_1]\, P[M \mid B_1] + P[B_2]\, P[M \mid B_2] + P[B_3]\, P[M \mid B_3]$$

Calculemos ahora $P[B_i]$ y $P[M \mid B_i]$ $i = 1,2,3$.

B_1: "se pasan 1 moneda de 100 intis y 4 de 50" ,
entonces el número de elemento de B_1 es $C_3^1\, C_4^4 = C_3^1$ y

$$P[B_1] = \frac{C_3^1}{C_7^5} .$$

B_2: "se pasan 2 monedas de 100 intis y 3 de 50",
entonces el número de elementos de B_2, es $C_3^2\, C_4^3$; y

$$P[B_2] = \frac{C_3^2\, C_4^3}{C_7^5}$$

B_3: "se pasan 3 monedas de 100 intis y 2 de 50"
luego, el número de elementos de B_3, es $C_3^3\, C_4^2 = C_4^2$ y

$$P[B_3] = \frac{C_4^2}{C_7^5} .$$

Si ocurre B_1, en el bolsillo izquierdo hay 7 monedas de 100 intis y 7 de 50 entonces,

$$P[M \mid B_1] = \frac{7}{14} .$$

Si ocurre B_2, en el bolsillo izquierdo hay, 8 monedas de 100 y 6 de 50, - entonces,

$$P[M \mid B_2] = \frac{8}{14} .$$

si ocurre B_3, en el bolsillo izquierdo hay, 9 monedas de 100 y 5 de 50, luego,

$$P[M \mid B_3] = \frac{9}{14} .$$

Por lo tanto,

$$P[M] = \frac{C_3^1}{C_7^5} \times \frac{7}{14} + \frac{C_3^2\, C_4^3}{C_7^5} \times \frac{8}{14} + \frac{C_4^2}{C_7^5} \times \frac{9}{14}$$

$$= \frac{1}{14\, C_7^5} \left[7\, C_3^1 + 8\, C_3^2 C_4^3 + 9\, C_4^2 \right]$$

$$= 0.58 .$$

Un esquema que ayuda a visualizar, la solución de este problema, es el diagrama del árbol de probabilidad de la fig. 1.8.12

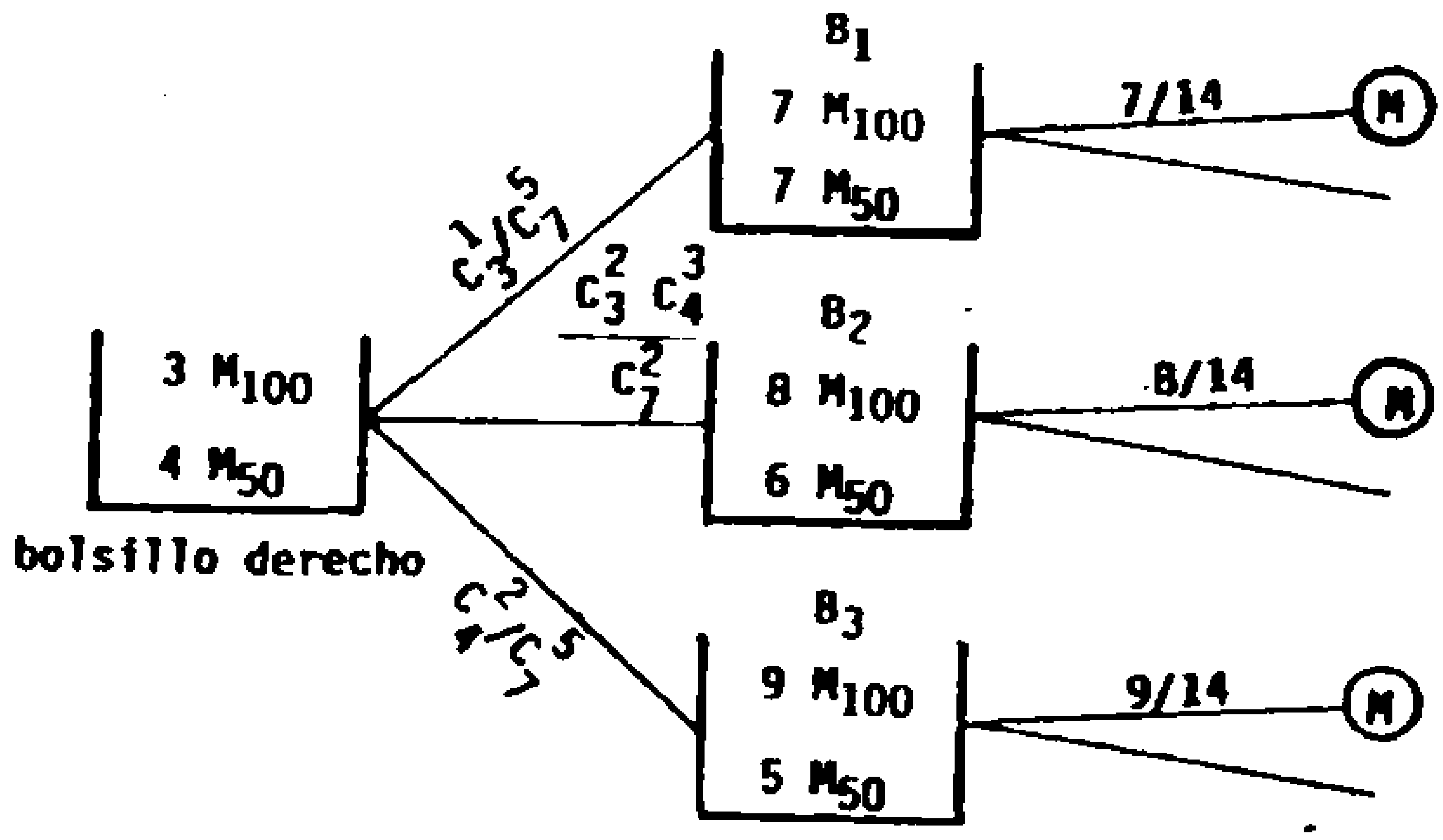


Fig. 1.8.12 bolsillo izquierdo

EJEMPLO 8 Una urna contiene 10 bolas rojas y 8 blancas. Se lanza un dado y se extrae de la urna tantas bolas como indica el número obtenido en el dado. Calcular la probabilidad que todas las bolas sean blancas.

SOLUCION Definimos los siguientes eventos

B_i: "obtener el número i en el dado $(i = 1,2, \ldots, 6)$"

A : "todas las bolas extraidas son blancas" .

El evento A se escribe, $A = \bigcup_{i=1}^{6} B_i A$. Por el teorema de probabilidad total se tiene

$$P[A] = \sum_{i=1}^{6} P[B_i]\, P[A \mid B_i] = \sum_{i=1}^{6} \frac{1}{6}\, P[A \mid B_i] = \frac{1}{6} \sum_{i=1}^{6} P[A \mid B_i] \ ,$$

Pues, $P[B_i] = \frac{1}{6}$, $i = 1,2, \ldots, 6$.

Cálculo de $P[A \mid B_i]$ $i = 1, \ldots, 6$

$$P[A \mid B_1] = \frac{\binom{8}{1}\binom{10}{0}}{\binom{18}{1}} = \frac{8}{18} \ .$$

$$P[A \mid B_2] = \frac{\binom{8}{2}\binom{10}{0}}{\binom{18}{2}} = \frac{8}{18} \times \frac{7}{17} \ .$$

$$P[A \mid B_3] = \frac{\binom{8}{3}\binom{10}{0}}{\binom{18}{3}} = \frac{8}{18} \times \frac{7}{17} \times \frac{6}{16} \ .$$

$$P[A \mid B_4] = \frac{\binom{8}{4}\binom{10}{0}}{\binom{18}{4}} = \frac{8}{18} \times \frac{7}{17} \times \frac{6}{16} \times \frac{5}{15} \ .$$

$$P[A \mid B_5] = \frac{\binom{8}{5}}{\binom{18}{5}} = \frac{8}{18} \times \frac{7}{17} \times \frac{6}{16} \times \frac{5}{15} \times \frac{4}{14} \ .$$

$$P[A \mid B_6] = \frac{\binom{8}{6}}{\binom{10}{6}} = \frac{8}{18} \times \frac{7}{17} \times \frac{6}{16} \times \frac{5}{15} \times \frac{4}{14} \times \frac{3}{13} \ .$$

Luego,

$$P[A] = \frac{1}{6} \times \frac{8}{18} \left[1 + \frac{7}{17} + \frac{7}{17} \times \frac{6}{16} + \frac{7}{17} \times \frac{6}{16} \times \frac{5}{15} + \frac{7}{17} \times \frac{6}{16} \times \frac{5}{15} \times \frac{4}{14} + \frac{7}{17} \times \right.$$

$$\times \frac{6}{16} \times \frac{5}{15} \times \frac{4}{14} \times \frac{3}{13}\Bigg]$$

$$= \frac{2}{3 \times 9}\left[1 + \frac{7}{17}\left[1 + \frac{4}{8} + \frac{1}{28} + \frac{3}{28 \times 13}\right]\right]$$

$$= \frac{2}{27}\left[1 + \frac{7}{17}\left[\frac{3}{2} + \frac{1}{28} \times \frac{16}{13}\right]\right]$$

$$= \frac{2}{27}\left[1 + \frac{7}{17} \times \frac{251}{2 \times 7 \times 13}\right] = \frac{2}{27} \times \frac{693}{2 \times 17 \times 13}$$

$$= \frac{77}{221 \times 3} = \frac{77}{663} = 0.116\ .$$

EJEMPLO 9 Una companía, por experiencias anteriores sabe que de un determinado número de lotes adquiridos, el 60% de ellos no tiene defectuosos, el 25% tiene sólo un defectuoso, el 10% tiene 2 defectuosos y el 5% tiene 3 defectuosos. Dicha campañía realiza un plan de muestreo de aceptación de lotes, que consiste en extraer una muestra de 3 artículos (uno después de otro) de cada lote que desea inspeccionar, se acepta dicho lote si a lo más encuentra un defectuoso en la muestra. Cada lote contiene 50 artículos. ¿Cuál es la probabilidad de aceptar el lote?.

SOLUCION Sea los siguientes eventos:

d : "artículo defectuoso".

b : "artículo no defectuoso".

B_0: "El lote no tiene artículos defectuosos".

B_1: "El lote tiene un artículo defectuoso".

B_2: "El lote tiene 2 artículos defectuosos".

B_3: "El lote tiene 3 artículos defectuosos".

A: "aceptar el lote".

$$P[A] = 1 - P[\bar{A}]$$

$$= 1 - P[B_2ddb \cup B_2dbd \cup B_2bdd \cup B_3ddd \cup B_3ddb \cup B_3dbd \cup B_3bdd]$$

$$= 1 - \left[(0.1)(3)\frac{2}{50 \times 49} + (0.05)\left(\frac{1}{50 \times 49 \times 8} + \frac{3 \times 47}{50 \times 49 \times 8}\right)\right]$$

$$= 1 - \frac{1}{50 \times 49}\left[\frac{6}{10} + \frac{142 \times 5}{8 \times 100}\right]$$

$$= 1 - \frac{595}{50 \times 49 \times 400} = 1 - \frac{17}{28000} = 1 - 0.0006 = 0.9994\ .$$

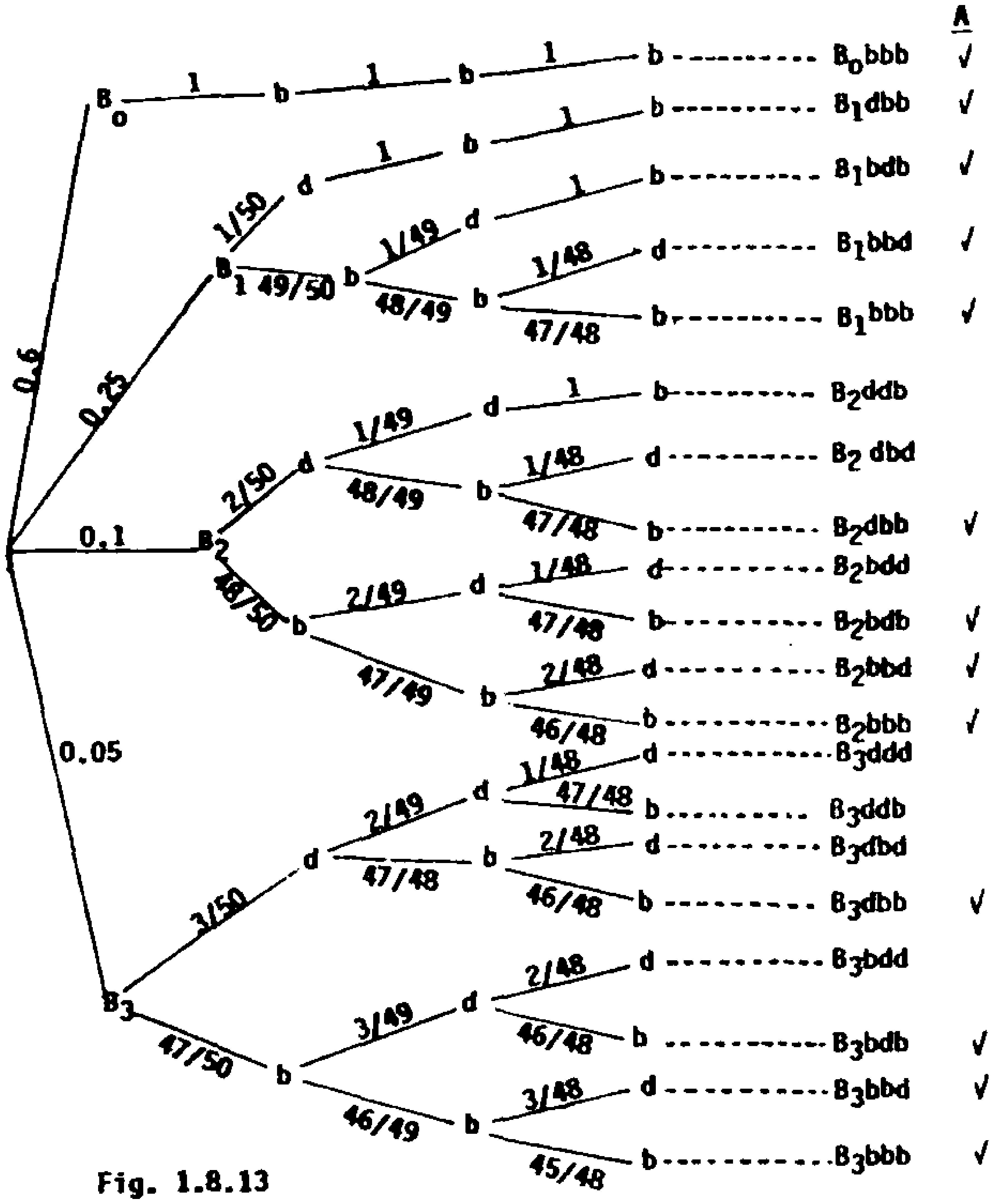


Fig. 1.8.13

EJEMPLO 10 Se lanza un dado 2 veces. El número que muestra la primera vez es el número de bolas blancas que se colocan en una urna, y el número obtenido en la segunda vez indica el número de bolas negras que se colocarán en la misma urna. Luego de la urna, se extraerá al azar una bola. Determinar la probabilidad que ésta sea blanca.

SOLUCION Sean los siguientes eventos:

A_{ij} : "en la urna hay i bolas blancas y j-negras", con i , j = 1,2,3,4,5,6 (Así, por ejemplo: "A_{25} es el evento en la urna hay 2 bolas blancas y 5 negras").

B: "la bola obtenida es blanca".

Los eventos A_{ij} son mutuamente excluyentes, también

$$\sum_{i=1}^{6} \sum_{j=1}^{6} A_{ij} = \Omega, \quad y \quad P[A_{ij}] = \frac{1}{36} > 0$$

Luego, dichos eventos forman una partición del espacio muestral. Por lo tanto,

$$P[B] = \sum_{i=1}^{6} \sum_{j=1}^{6} P[A_{ij}]\, P[B \mid A_{ij}] \; ; \quad y \quad P[B \mid A_{ij}] = \frac{i}{i+j} \,, \; \forall \; i,j=1,2,3,4,5,6$$

$$= \frac{1}{36} \sum_{i=1}^{6} \sum_{j=1}^{6} \frac{i}{i+j} = \frac{1}{36} \sum_{i=1}^{6} \left[\frac{i}{i+1} + \frac{i}{i+2} + \frac{i}{i+3} + \frac{i}{i+4} + \frac{i}{i+5} + \frac{i}{i+6} \right]$$

$$= \frac{1}{36}(18) = \frac{1}{2} \,.$$

TEOREMA 1.8.2 TEOREMA DE PROBABILIDAD TOTAL, DE LA PROB. COND

Sean B_1, B_2, ..., B_k eventos que forman una partición del espacio muestral Ω, y sean A y E eventos, entonces

$$P[E \mid A] = P[B_1 \mid A]\, P[E \mid B_1A] + P[B_2 \mid A]\, P[E \mid B_2A] + \dots + P[B_k \mid A]\, P[E \mid B_kA]$$

PRUEBA

$$P[B_1 \mid A]\, P[E \mid B_1A] + P[B_2 \mid A]\, P[E \mid B_2A] + \dots + P[B_k \mid A]\, P[E \mid B_kA]$$

$$= \frac{P[B_1A]}{P[A]} \, \frac{P[EB_1A]}{P[B_1A]} + \frac{P[B_2A]}{P[A]} \, \frac{P[EB_2A]}{P[B_2A]} + \dots + \frac{P[B_kA]}{P[A]} \, \frac{P[EB_kA]}{P[B_kA]}$$

$$= \frac{P[EB_1A]}{P[A]} + \frac{P[EB_2A]}{P[A]} + \dots + \frac{P[EB_kA]}{P[A]}$$

$$= \frac{P[EB_1A] + P[EB_2A] + \dots + P[EB_kA]}{P[A]}$$

$$= \frac{P[EA(B_1 \cup B_2 \cup \dots \cup B_k)]}{P[A]} = \frac{P[EA\Omega]}{P[A]} = \frac{P[EA]}{P[A]} = P[E \mid A]$$

De donde

$$\boxed{P[E \mid A] = \sum_{k=1}^{k} P[B_i \mid A]\, P[E \mid AB_i]}$$

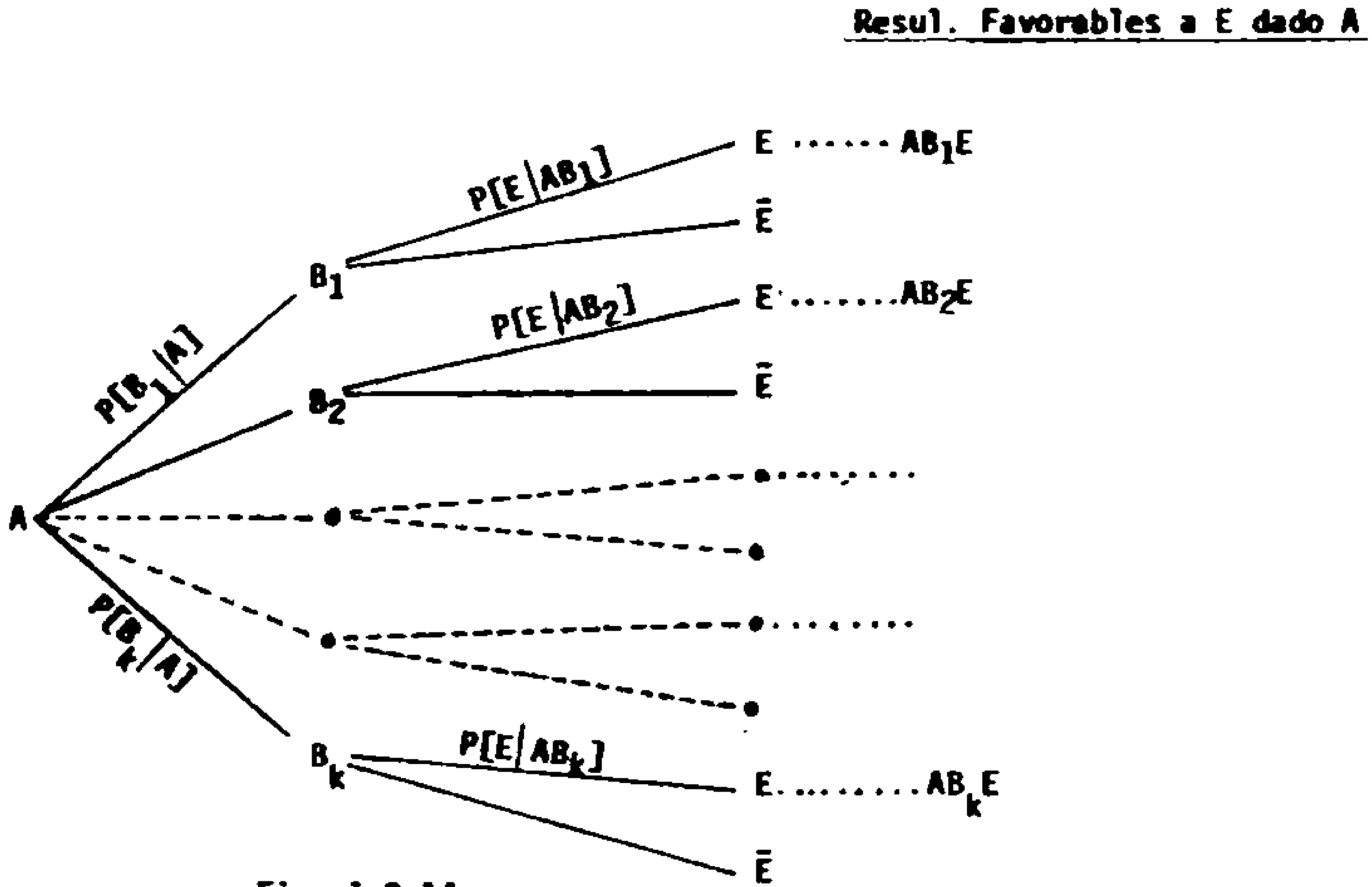


Fig. 1.8.14

EJEMPLO 11 Todos los miembros de un club son médicos o abogados, 40 % de los miembros son médicos mientras que el 30 % de las mujeres son médicos. El 50 % de los médicos y el 30 % de los abogados ganan más de \$ 60,000 por año. Sin embargo solamente el 20 % de las mujeres médicos y el 10 % de las mujeres abogados ganan más de \$ 60,000, por año.

(a) Si se escoge aleatoriamente un miembro del club. ¿cuál es la probabilidad que gane más de \$ 60,000 por año?

(b) Si se escoge aleatoriamente una mujer. ¿Cuál es la probabilidad que ella gane más de \$ 60,000 por año?

SOLUCION Sean los siguientes eventos:

M: "El miembro del club es un médico".

A: "El miembro del club es un abogado".

F: "El miembro del club es femenino".

G: "El miembro del club gana más de \$ 60,000 por año".

(a) Debemos calcular $P[G]$.

$\Omega = A \cup M$. Los eventos A y M forman una partición del espacio muestral Ω (el club). Además $G \subset \Omega$.

Aplicando el teorema de probabilidad total tenemos,

$$P[G] = P[A]\,P[G \mid A] + P[M]\,P[G \mid M]$$

$$= (0.6)(0.3) + (0.4)(0.5) = 0.38 .$$

(b) Se debe calcular $P[G \mid F]$. El diagrama siguiente, muestra el árbol de probabilidades para este caso.

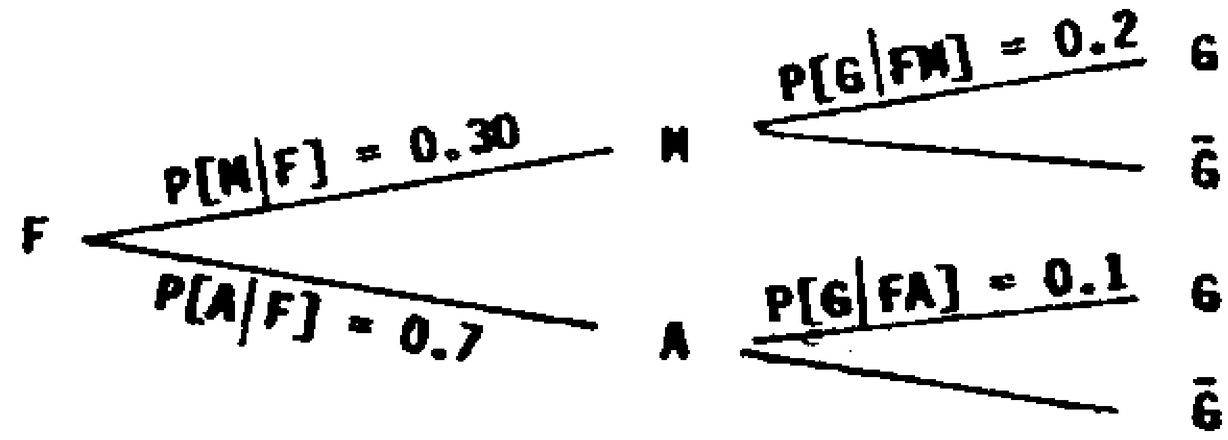


Fig. 1.8.15

Luego,

$$P[G \mid F] = P[M \mid F]\,P[G \mid FM] + P[A \mid F]\,P[G \mid FA]$$

$$= (0.3)(0.2) + (0.7)(0.1) = 0.13 .$$

1.8.3 TEOREMA DE BAYES

En el ejemplo 2. Supongamos ahora que el conejo escogido aleatoriamente se ve que es blanco. ¿Cuál es la probabilidad que provenga de la jaula I?

Usando la notación del ejemplo 2, debemos calcular $P[I \mid A]$. Por la definición de probabilidad condicional y el teorema de multiplicación tenemos

$$P[I \mid A] = \frac{P[IA]}{P[A]} = \frac{P[I]P[I \mid A]}{P[A]}$$

$P[A]$ se ha calculado por el teorema de probabilidad total y es $\frac{43}{90}$. Luego,

$$P[I \mid A] = \frac{1/3 \times 3/5}{\frac{43}{90}} = \frac{90}{43 \times 5} = \frac{18}{43} .$$

Este ejemplo se generaliza por el siguiente teorema.

TEOREMA 1.8.3 TEOREMA DE BAYES Si los eventos $B_1, B_2, \ldots, B_k$ forman una partición del espacio muestral Ω y A es un evento cualquiera de Ω, entonces

$$P[B_r \mid A] = \frac{P[B_r]\,P[A \mid B_r]}{\sum_{i=1}^{k} P[B_i]\,P[A \mid B_i]}$$

$$= \frac{P[B_r]\,P[A \mid B_r]}{P[B_1]\,P[A \mid B_1] + P[B_2]\,P[A \mid B_2] + \ldots + P[B_k]\,P[A \mid B_k]}$$

para $r = 1, 2, \ldots, k$.

Este teorema resulta como consecuencia inmediata del teorema de probabilidad total. En efecto

$$P[B_r \mid A] = \frac{P[B_r A]}{P[A]} = \frac{P[B_r]\, P[A \mid B_r]}{\sum_{i=1}^{k} P[B_i]\, P[A \mid B_i]}$$

El numerador resulta del teorema de multiplicación y el denominador del teorema de probabilidad total.

COROLARIO 2 Si A y B son eventos en Ω tales que $P[A] > 0$ y $0 < P[B] < 1$ entonces

$$P[B \mid A] = \frac{P[B]\, P[A \mid B]}{P[B]\, P[A \mid B] + P[\bar{B}]\, P[A \mid \bar{B}]}$$

Este corolario es una consecuencia inmediata del corolario 1. En efecto:

$$P[B \mid A] = \frac{P[B\, A]}{P[A]} = \frac{P[B]\, P[A \mid B]}{P[B]\, P[A \mid B] + P[\bar{B}]\, P[A \mid \bar{B}]}$$

EJEMPLO 12 Una compañía de desarrollo urbano está considerando la posibilidad de construir un centro comercial en un sector de Lima metropolitana, Perú. Un elemento vital en esta consideración es un proyecto de una autopista que une este sector con el centro de la ciudad. Si el consejo municipal aprueba esta autopista, hay una probabilidad de 0.90 de que la compañía construya el centro comercial en tanto que si la autopista no es aprobada la probabilidad es de sólo 0.20. Basándose en la información disponible, el presidente de la compañía estima que hay una probabilidad de 0.60 que la autopista sea aprobada.

(a) ¿Cuál es la probabilidad que la compañía construya el centro comercial?

(b) Dado que el centro comercial fué construido. ¿Cuál es la probabilidad de que la autopista haya sido aprobada?.

SOLUCION Definimos los eventos:

A: "la autopista es aprobada".

B: "el centro comercial es construído".

(a) Debemos calcular $P[B]$, aplicando el corolario del teorema de probabilidad total. El evento B se escribe $B = AB \cup \bar{A}B$ y

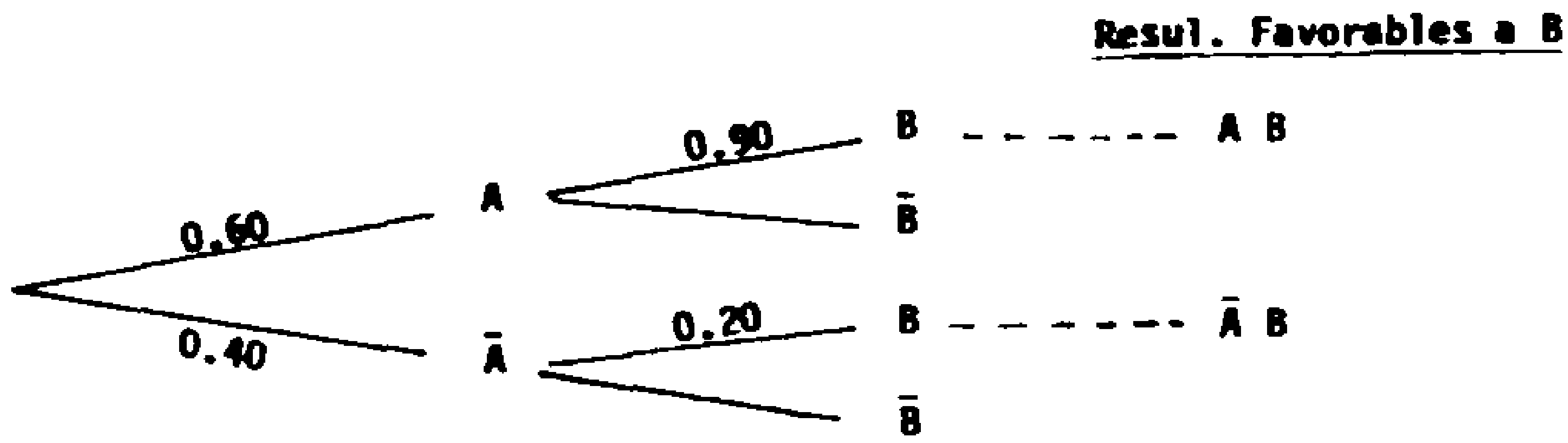


Fig. 1.8.16

$$P[B] = P[AB] + P[\bar{A}\,B] = P[A]\,P[B\,|\,A] + P[\bar{A}]\,P[B\,|\,\bar{A}]$$

$$= (0.60)(0.90) + (0.40)(0.20) = 0.54 + 0.08$$

$$= 0.62\ .$$

(b) Por el corolario del teorema de Bayes tenemos

$$P[A\,|\,B] = \frac{P[A]\,P[B\,|\,A]}{P[B]} = \frac{(0.60)(0.90)}{0.62} = \frac{54}{62} = 0.87\ .$$

EJEMPLO 13 En una línea de producción hay dos procesos, **A** y B. En el proceso **A** hay un 20 % de defectuosos y en **B** hay 25 % . En una muestra de 300 productos hay 200 del proceso **A** y 100 del B.

(a) Si se extrae un producto al azar, hallar la probabilidad que sea defectuoso.

(b) Si al extraer el producto resultó defectuoso, halle la probabilidad de que sea del proceso **A**.

SOLUCION Sean los siguientes eventos:

A: "el producto es del proceso A".

B: "el producto es del proceso B".

D: "el producto es defectuoso".

$\bar{D}$: "el producto es no defectuoso".

$\Omega = A \cup B$. Es decir, A y B forman una partición de Ω .

(a) Debemos calcular **P[D]** . Este evento se escribe $D = AD \cup BD$ y por el teorema de probabilidad total es

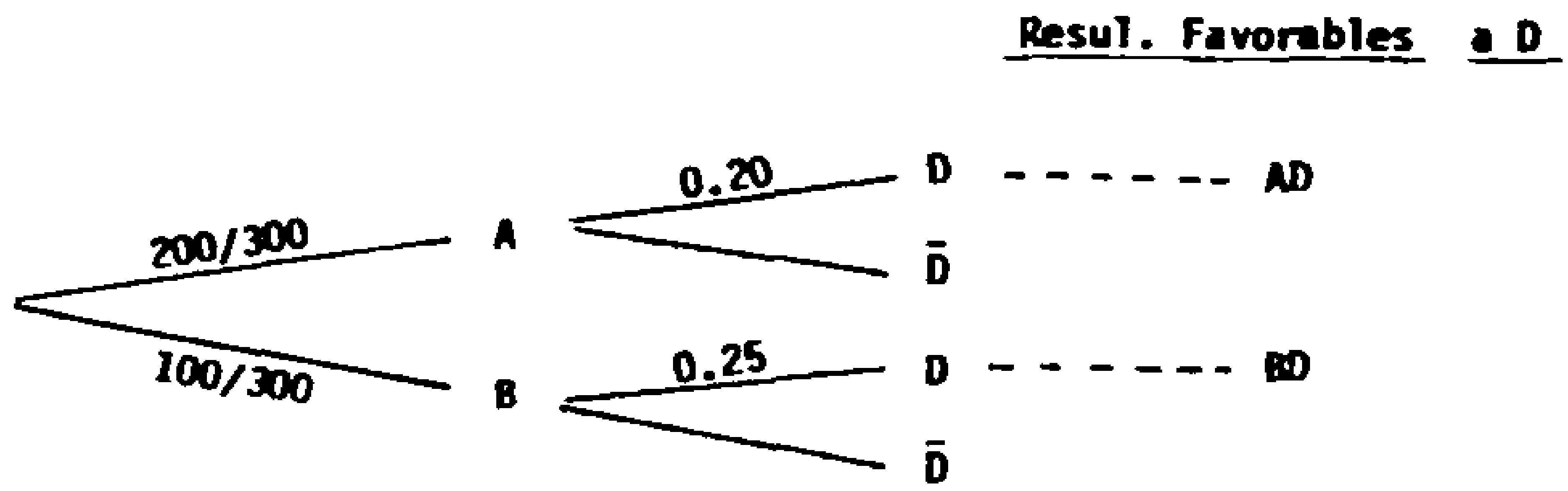


Fig. 1.8.17

$$P[D] = P[AD] + P[BD] = P[A]\,P[D \mid A] + P[B]\,P[D \mid B]$$

$$= \frac{2}{3}(0.20) + \frac{1}{3}(0.25) = \frac{65}{300} \ .$$

(b) Por el teorema de Bayes se tiene

$$P[A \mid D] = \frac{P[A]\,P[D|A]}{P[D]} = \frac{(2/3)(0.20)}{65/300} = 0.615$$

EJEMPLO 14 La compañía ensambladora de automóviles CAR-PERU, se ha presentado a una licitación, para ensamblar un nuevo modelo de automóvil. La probabilidad que CAR-PERU gane la licitación es de 0.90 si una firma competidora MOTOR ANDINO no se presente a ella en tanto que es de sólo 0.20 Sí MOTOR-ANDINO se presenta. El gerente general de CAR-PERU estima que hay una probabidad de 0.80 que MOTOR-ANDINO se presente.

(a) ¿Cuál es la probabilidad que CAR-PERU gane la licitación?.

(b) Dado que CAR-PERU ganó la licitación, ¿cuál es la probabilidad que MOTOR-ANDINO se haya presentado a ella?

SOLUCION Sean los siguientes eventos:

E: "la compañía MOTOR-ANDINO se presenta a la licitación".

G: "la compañía CAR-PERU gana la licitación".

El evento G se escribe $G = EG \cup \bar{E}G$ y por el corolario del teorema de probabilidad total es

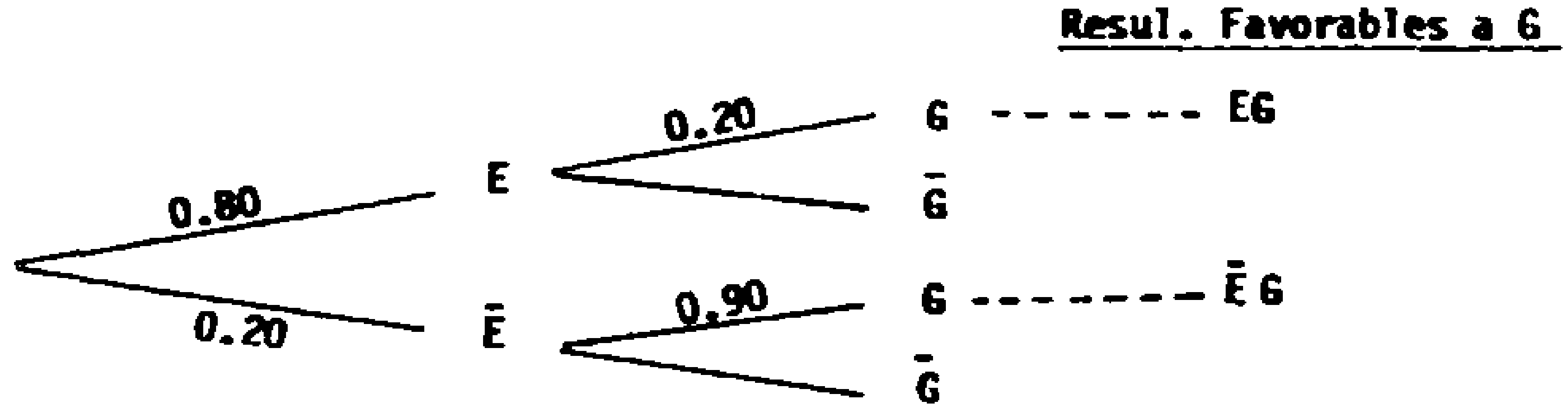


Fig. 1.8.18

$$P[G] = P[EG] + P[\bar{E}G]$$

$$P[G] = P[E]\,P[G \mid E] + P[\bar{E}]\,P[G \mid \bar{E}]$$

$$= (0.80)(0.20) + (0.20)(0.90) = 0.34 \ .$$

(b) Aplicando el corolario del teorema de Bayes, se tiene

$$P[E \mid G] = \frac{P[E]\,P[G|E]}{P[G]} = \frac{(0.80)(0.20)}{0.34} = \frac{8}{17}$$

EJEMPLO 15 La compañía "compre ahora" efectúa una encuesta del mercado para evaluar la lucratividad de cada uno de sus nuevos productos. Encuestas anteriores indican que el 90% de los nuevos productos debieran resultar lucrativos; sin embargo, un análisis posterior de la confiabilidad de las encuestas ha demostrado que sólo el 70% de los productos que se pronosticaban como lucrativos, lo fueron efectivamente. En contraste, de los productos pronosticados como no lucrativos por las encuestas, el 20% resultó ser lucrativo. La compañía ha comercializado un nuevo producto llamado X. Dado que X resultó lucrativo. ¿Cuál es la probabilidad que la encuesta haya pronosticado a X como no lucrativo?.

SOLUCION Sean los eventos:

P: "El producto X fue pronosticado como lucrativo".

L: "El producto X resulto lucrativo".

Observe que $\Omega = P \cup \bar{P}$. Es decir, P y $\bar{P}$ forman una partición de Ω.

Se pregunta $P[\bar{P} \mid L]$. El evento $L \subset \Omega$, se escribe $L = PL \cup \bar{P}L$. Luego,

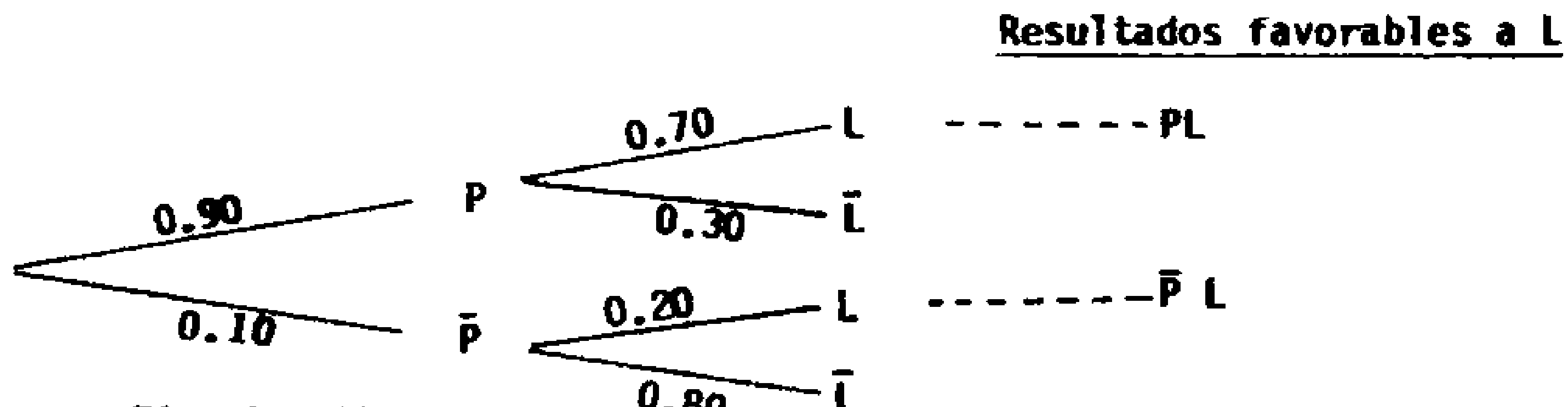


Fig. 1.8.19

$$P[L] = P[P]\,P[L \mid P] + P[\bar{P}]\,P[L \mid \bar{P}]$$

$$= (0.90)(0.70) + (0.10)(0.20) = 0.65 .$$

Entonces, según el corolario del teorema de Bayes, se tiene que

$$P[\bar{P} \mid L] = \frac{P[\bar{P}]\,P[L \mid \bar{P}]}{P[L]} = \frac{(0.10)(0.20)}{0.65} = \frac{2}{65} .$$

EJEMPLO 16 Una vacuna produce inmunidad contra el sarampión en un 95% de los casos, supongamos que en una población, el 30% de las personas se ha vacunado. Supongamos además que una persona vacunada sin inmunidad tiene la misma probabilidad de contraer sarampión que la persona que no está vacunada ¿Cuál es la probabilidad que una persona que contrajo sarampión esté vacunada?.

SOLUCION 1 Definimos los eventos :

V: "la persona está vacunada".

$\bar{V}$: "la persona no está vacunada".

S: "persona con sarampión".

2. El espacio muestral Ω, toda la población, se compone de personas vacunadas y no vacunadas, $\Omega = V \cup \bar{V}$. Osea V y $\bar{V}$ forman una partición de Ω

3. Debemos calcular $P[V \mid S]$, la probabilidad que la persona esté vacunada dado que contrajo sarampión.

4. El evento $S \subset \Omega$, se escribe - $S = VS \cup \bar{V}S$. Luego

   $P[S] = P[V]\, P[S \mid V] + P[\bar{V}]\, P[S \mid \bar{V}]$

   Teorema de probabilidad total.

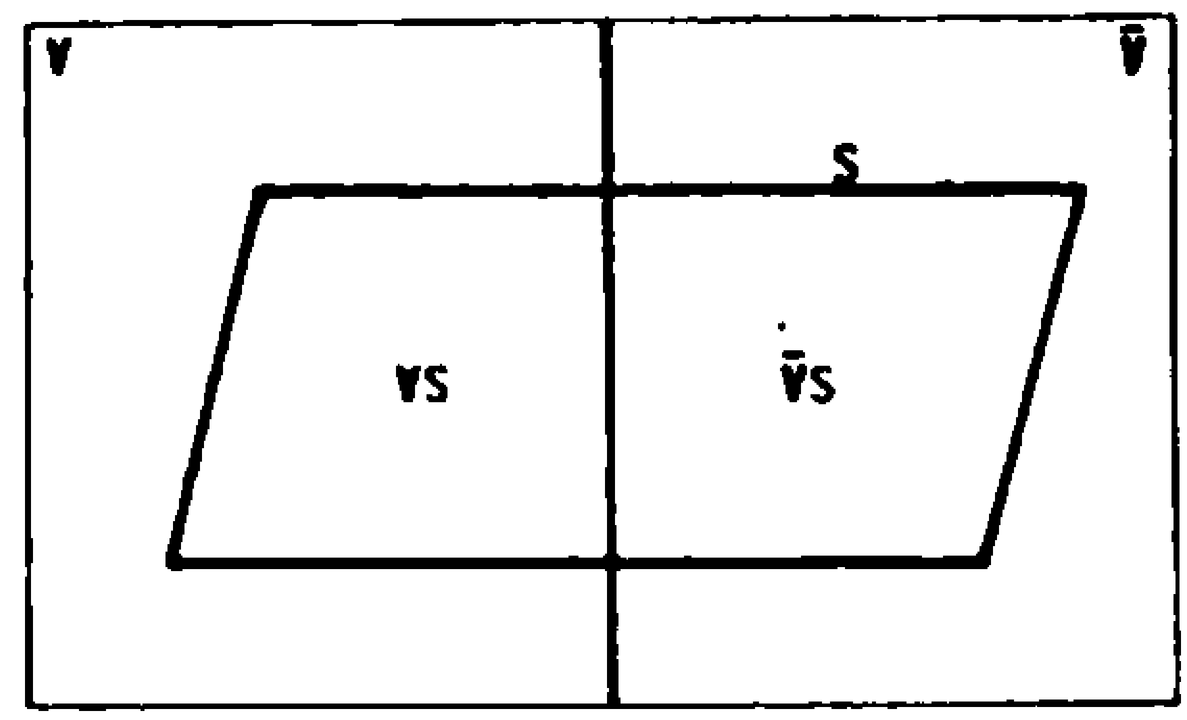


Fig. 1.8.20

$$P[S] = \frac{30}{100} \times \frac{\frac{30 \times 5}{100}}{30} + \frac{70}{100} \times 1 = 0.015 + 0.7 = 0.715.$$

Por lo tanto, según el corolario del teorema de Bayes

$$P[V \mid S] = \frac{P[V]\, P[S \mid V]}{P[S]} = \frac{0.015}{0.715} = \frac{3}{143} \ .$$

EJEMPLO 17 Supongamos que la ciencia médica ha desarrollado una prueba, para el diagnóstico del cáncer, que tiene el 95% de exactitud tanto en los que tienen cáncer como entre los que no tienen. Si 0.005 de la población realmente tiene cáncer, calcular la probabilidad que determinado individuo tenga cáncer, si la prueba dice que tiene.

SOLUCION 1 Definimos los eventos siguientes:

C: "la persona examinada realmente tenga cáncer".

D: "el diagnostico dice la persona examinada tiene cáncer".

2. Ω: "Todas las personas", se escribe $\Omega = C \cup \bar{C}$. Es decir C y $\bar{C}$ forman una partición del espacio muestral.

3. Del enunciado: $P[C] = 0.005$, $P[D \mid C] = 0.95$, $P[\bar{D} \mid \bar{C}] = 0.95$. Y debemos calcular $P[C \mid D]$.

4. El evento D se escribe $D = CD \cup \bar{C}D$ y por el corolario del teorema de - probabilidad total

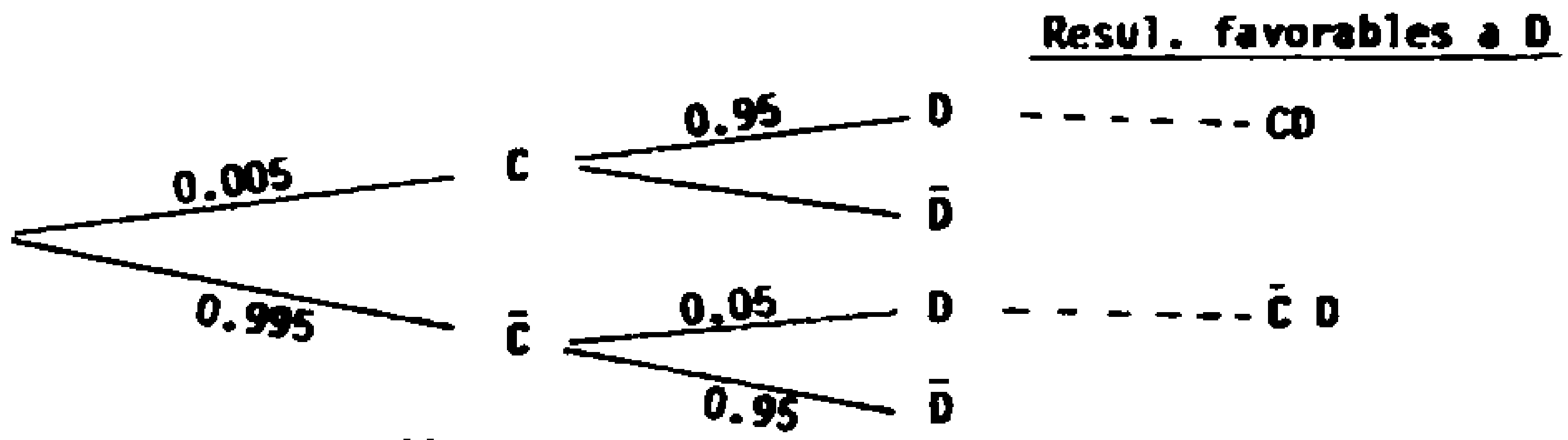


Fig. 1.8.21

$$P[D] = P[C]\, P[D \mid C] + P[\bar{C}]\, P[D \mid \bar{C}]$$

$$= (0.005)(0.95) + (0.995)(0.05) = 0.0545\ .$$

5. Por el corolario del teorema de Bayes se obtiene

$$P[C \mid D] = \frac{P[C]\ P[D \mid C]}{P[D]} = \frac{(0.005)(0.95)}{0.05450} = \frac{475}{5450} = 0.087\ ,$$

EJEMPLO 18 Dos proveedores, A Y B, entregan la misma pieza a un fabricante, que guarda las existencias de esta pieza en un mismo lugar. Los antecedentes demuestran que 5 % de las piezas entregadas por A son defectuosas, y que 9 % de las piezas entregadas por B también son defectuosas. Además A entrega cuatro veces más piezas que B. Si se extrae al azar una pieza y se encuentra que no es defectuosa, ¿cuál es la probabilidad de que haya fabricado A?.

SOLUCION Sean los siguientes eventos:

A: "extraer una pieza entregada por A".

B: "extraer una pieza entregada por B".

N: "extraer una pieza no defectuosa".

Debemos calcular $P[A \mid N]$. $\Omega = A \cup B$, o sea A y B forman una partición de Ω. Puesto que $N \subset \Omega$, se escribe $N = AN \cup BN$.

Del enunciado $P[A] = \dfrac{4}{5}$ y $P[B] = \dfrac{1}{5}$ ¿por qué?.

Luego, por el teorema de probabilidad total es

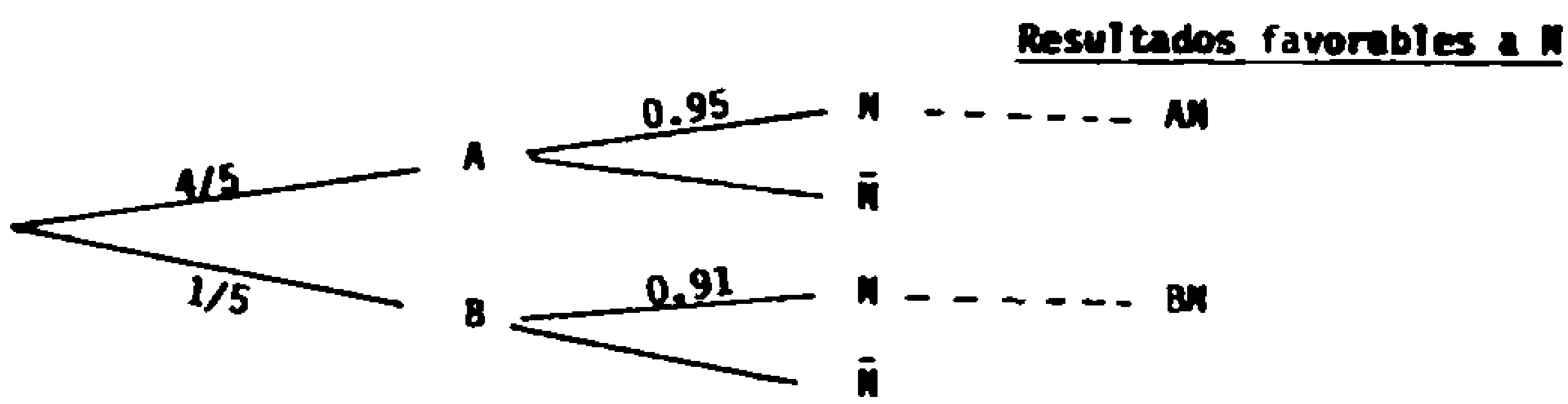


Fig. 1.8.22

$$P[N] = P[A]\, P[N \mid A] + P[B]\, P[N \mid B]$$

$$= \frac{4}{5}(0.95) + \frac{1}{5}(0.91) = \frac{471}{500}.$$

Y según el teorema de Bayes. Se tiene

$$P[A \mid N] = \frac{P[A]\,P[N \mid A]}{P[A]} = \frac{4/5\ (0.95)}{47/500} = \frac{380/500}{471/500} = 0.807\ ,$$

EJEMPLO 19 El gerente general de una cadena sudamericana de supermercados - estima la proporción de sus establecimientos que alcanzarán la meta de una - venta anual equivalente a doce millones de dólares en la forma siguiente:

PROPORCION DE ESTABLECIMIENTOS (π)	PROBABILIDAD P (π)
$\pi_1 = 0.60$	$P(\pi_1) = 0.20$
$\pi_2 = 0.70$	$P(\pi_2) = 0.50$
$\pi_3 = 0.80$	$P(\pi_3) = 0.30$

Es decir, el gerente general, bazándose en experiencias anteriores estima que hay una probabilidad de 0.20 de que 60 % de las tiendas alcanzarán - los doce millones de venta anual; una probabilidad de 0.50 que alcancen el - 70 % y finalmente una probabilidad de 0.30 de que 80 % alcancen la meta. Se - selecciona al azar uno de los negocios.

(a) ¿Cuál es la probabilidad que este haya alcanzado la meta considerada?

(b) Dado que este negocio alcanzó la meta, ¿cuál es la probabilidad que - el 80 % de los negocios haya vendido doce millones de dólares?

SOLUCION Sea el evento A: "obtener un negocio que alcanza la meta considerada".

Las formas diferentes de obtener un negocio que alcanza la meta fijada, se - observa en el diagrama del árbol de probabilidades.

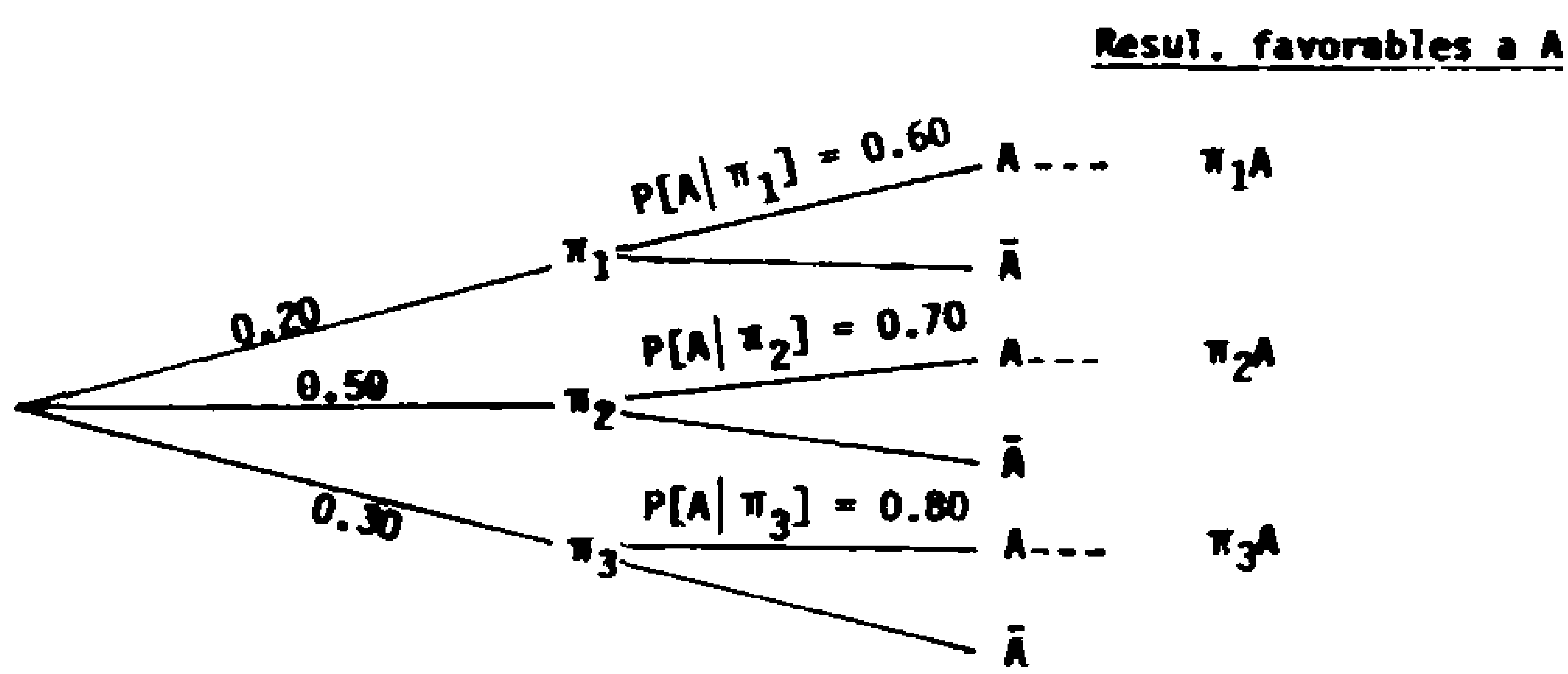


Fig. 1.8.23

(a) El evento se escribe, $A = \bigcup_{i=1}^{3} \pi_i A$, y por el teorema de probabilidad total es

$$P[A] = P[\pi_1]\,P[A \mid \pi_1] + P[\pi_2]\,P[A \mid \pi_2] + P[\pi_3]\,P[A \mid \pi_3]$$
$$= (0.20)(0.60) + (0.50)(0.70) + (0.30)(0.80)$$
$$= 0.12 + 0.35 + 0.24 = 0.71 .$$

(b) De acuerdo con el teorema de Bayes se tiene

$$P[\pi_3 \mid A] = \frac{P[\pi_3]\,P[A \mid \pi_3]}{P[A]} = \frac{(0.30)(0.80)}{0.71} = \frac{0.24}{0.71} = 0.338$$

EJEMPLO 20 En el almacén de una firma comercial distribuidora de fusibles se encuentra 80 cajas con 100 fusibles cada una. 20 cajas contienen fusibles producidos por una empresa A, 30 cajas contienen fusibles producidos por una empresa B, el resto de cajas contienen fusibles producidos por una compañía C. A produce el 3% de artículos defectuosos, B el 5% y C el 4% de artículos defectuosos. Si se selecciona una de estas cajas al azar, se toma una de sus fusibles y se encuentra que es defectuoso, ¿cuál es la probabilidad de que haya sido producido por B?

SOLUCION Definimos los siguientes eventos:

A: "Elegir una caja de fusibles producidos por la empresa A".

B: "Elegir una caja de fusibles producidos por la empresa B".

C: "Elegir una caja de fusibles producidos por la empresa C".

D: "obtener un artículo defectuoso".

Debemos calcular $P[B \mid D]$.

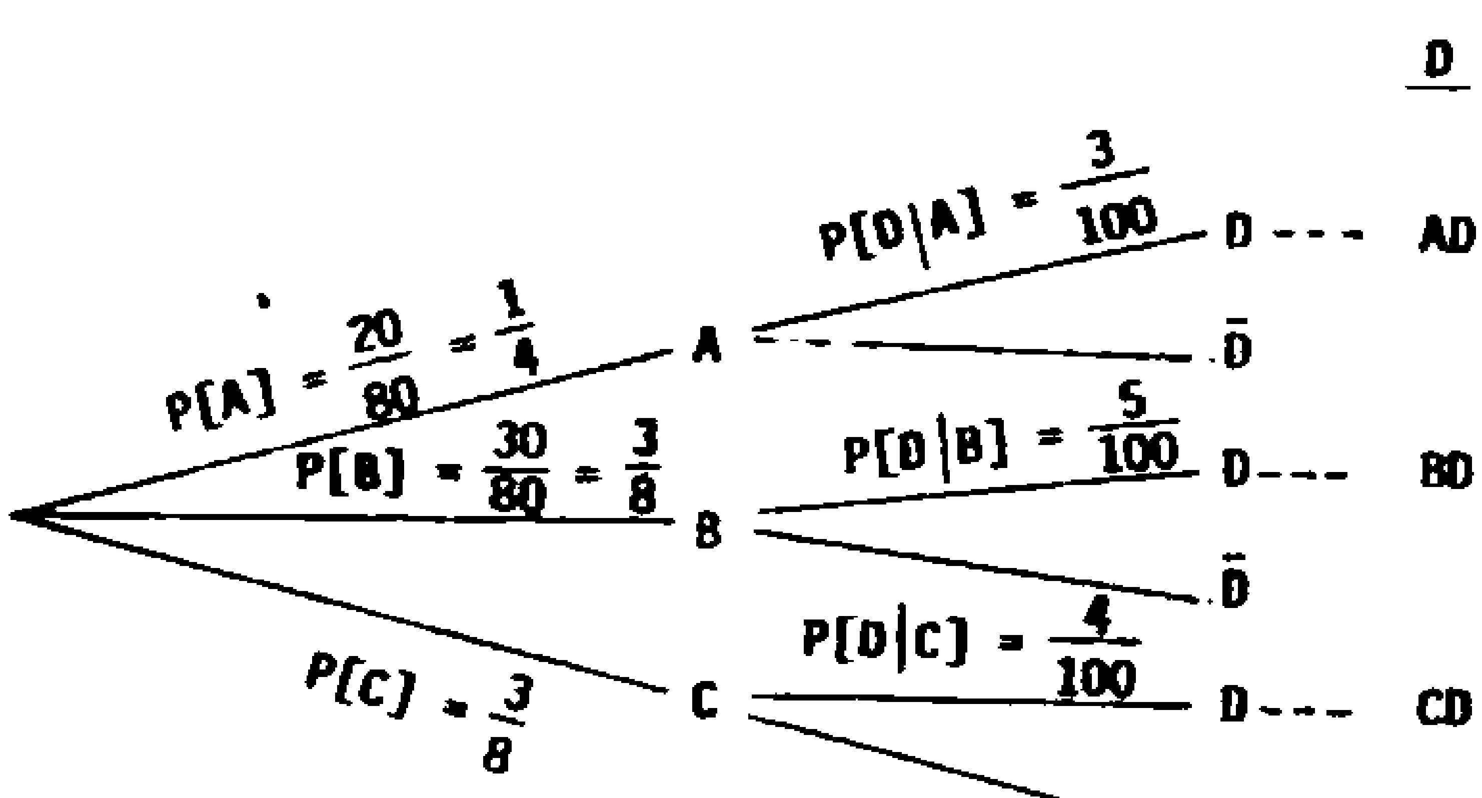


Fig. 1.8.24

$$P[D] = P[AD] + P[BD] + P[CD]$$

$$= P[A]\,P[D\mid A] + P[B]\,P[D\mid B] + P[C]\,P[D\mid C]$$

$$= \frac{2}{8} \times \frac{3}{100} + \frac{3}{8} \times \frac{5}{100} + \frac{3}{8} \times \frac{4}{100} = \frac{33}{800}.$$

Según el teorema de Bayes tenemos

$$P[B\mid D] = \frac{P[B]\,P[D\mid B]}{P[D]} = \frac{3/8 \times 5/100}{33/800} = \frac{5}{11} = 0.45$$

EJEMPLO 21 Considere 18 tiradores clasificados en 4 grupos. En el primer grupo hay 5 tiradores con probabilidades 0.8 de dar en el blanco, en el segundo hay 7 con probabilidad 0.7, en el tercero hay 4 con probabilidad de 0.6 y en el último 2 con probabilidad 0.5 de dar en el blanco. Se elige aleatoriamente un tirador, dispara y no da en el blanco. ¿A qué grupo es más probable que pertenezca?

SOLUCION Definimos los siguientes eventos :

E_i : "el tirador elegido pertenece al grupo i ($i = 1,2,3,4$)"

F: "al disparar no da en el blanco";

por lo tanto, el evento F podemos escribir como sigue,

$$F = \bigcup_{i=1}^{4} E_i F$$

$$P[F] = \sum_{i=1}^{4} P[E_i]\,P[F\mid E_i]\,, \quad \text{(Teor. de prob. total)}$$

las probabilidades de los eventos E_i son :

$$P[E_1] = \frac{5}{18}\,,\ P[E_2] = \frac{7}{18}\,,\ P[E_3] = \frac{4}{18}\,,\ P[E_4] = \frac{2}{18}$$

Si ocurre E_1, $P[F\mid E_1] = 0.2$

Si ocurre E_2, $P[F\mid E_2] = 0.3$

Si ocurre E_3, $P[F\mid E_3] = 0.4$

Si ocurre E_4, $P[F\mid E_4] = 0.5$

luego, $$P[F] = \frac{5}{18}(0.2) + \frac{7}{18}(0.3) + \frac{4}{18}(0.4) + \frac{2}{18}(0.5)$$

$$= \frac{1}{18}\left[1 + 2.1 + 1.6 + 1\right] = \frac{5.7}{18}.$$

Calcularemos ahora $P[E_i \mid F]$ $(i = 1,2,3,4)$, según teorema de Bayes

$$P[E_1 \mid F] = \frac{P[E_1]\,P[F \mid E_1]}{P[F]} \quad \frac{\frac{5}{18}(0.2)}{\frac{5.7}{18}} = \frac{1}{5.7} \;;$$

$$P[E_2 \mid F] = \frac{P[E_2]\,P[F \mid E_2]}{P[F]} = \frac{\frac{7}{18}(0.3)}{\frac{5.7}{18}} = \frac{2.1}{5.7} \;;$$

$$P[E_3 \mid F] = \frac{P[E_3]\,P[F \mid E_3]}{P[F]} = \frac{\frac{4}{18}(0.4)}{\frac{5.7}{18}} = \frac{1.6}{5.7} \;;$$

$$P[E_4 \mid F] = \frac{P[E_4]\,P[F \mid E_4]}{P[F]} = \frac{\frac{2}{18}(0.5)}{\frac{5.7}{18}} = \frac{1}{5.7}$$

por lo tanto, el jugador elegido, es más probable que pertenezca al segundo grupo.

El diagrama del árbol de probabilidad de la fig.1.8.25 da una visión esquemática de la solución de este problema.

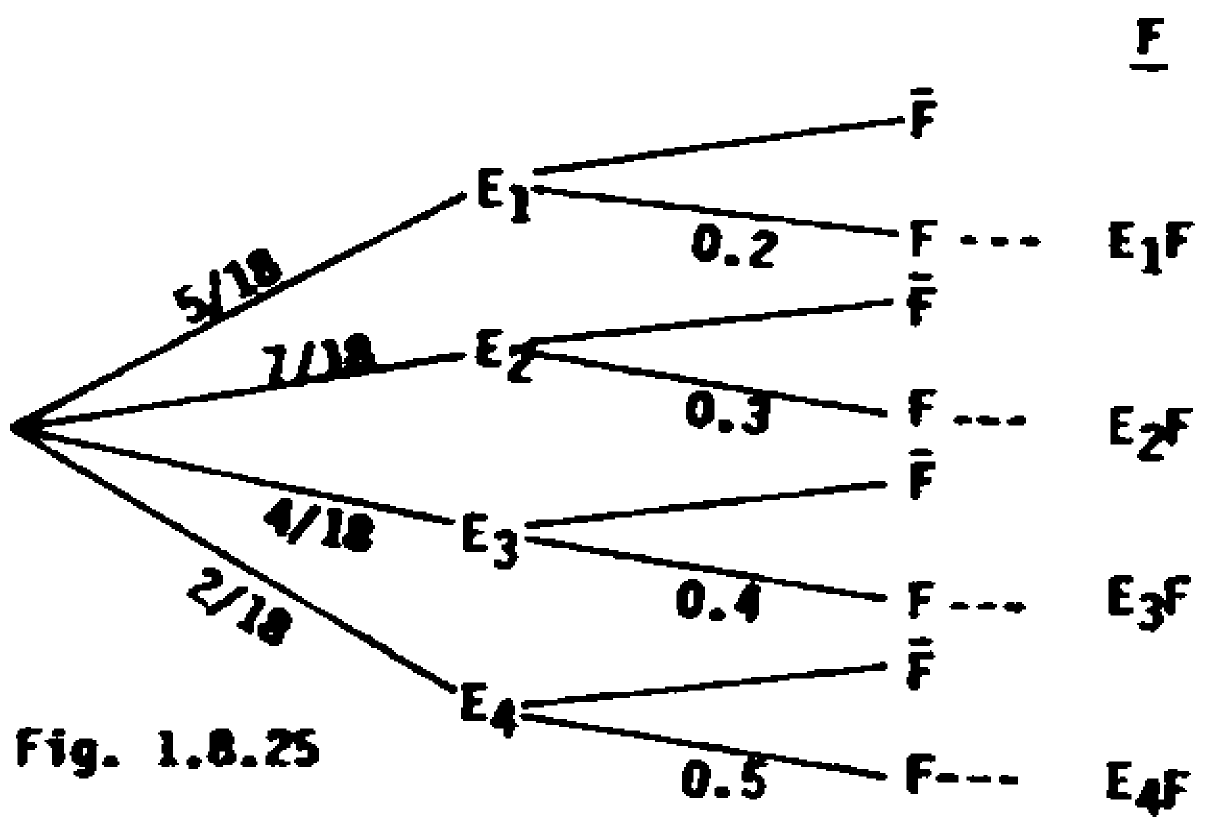


Fig. 1.8.25

EJEMPLO 22 Consideremos una muestra de tamaño 3, extraída de la siguiente manera. Se empieza con un grupo de 12 cartas: 7 espadas y 5 diamantes. En cada ensayo se extrae una carta, se ve el palo y se devuelve, juntos con otra carta adicional del mismo palo. ¿Cuál es la probabilidad que el número de espadas en el grupo de cartas antes de la tercera extracción sea 8, dado que - la muestra contiene 2 espadas y un diamante?.

SOLUCION Sean los siguientes eventos :

E: "espada" ; D: "diamante".

A: "La muestra contiene 2 espadas y 1 diamante".

B: "El número de espadas antes de la tercera extracción es 8".

Se pide determinar $P[B \mid A]$.

El diagrama del árbol de probabilidades que da una visión esquemática del problema es

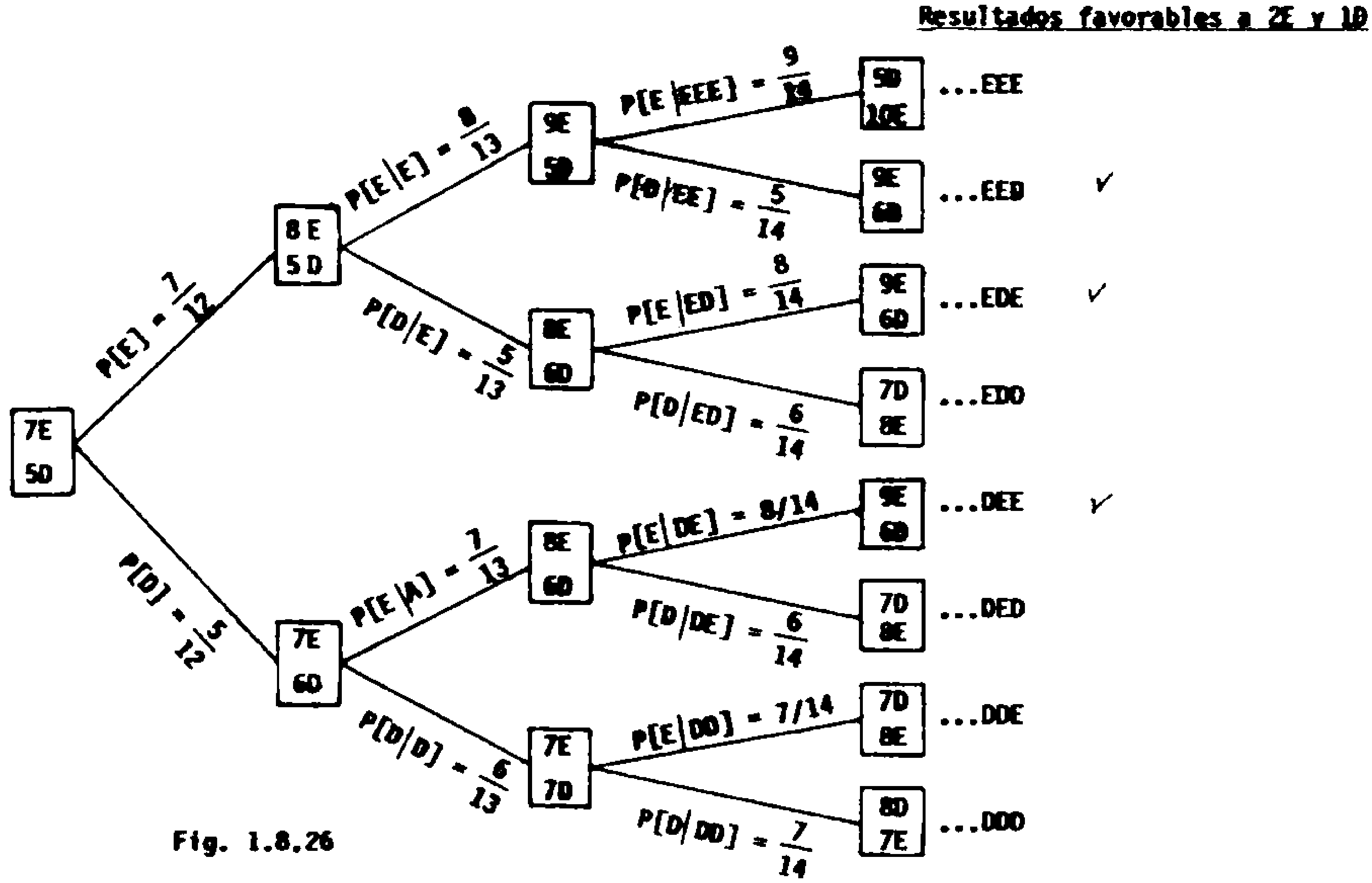


Fig. 1.8.26

El evento **A** se escribe $A = EED \cup EDE \cup DEE$ y por el teorema de probabilidad total se tiene

$$P[A] = P[E]\,P[E|E]\,P[D|EE] + P[E]\,P[D|E]\,P[E|ED] + P[D]\,P[E|D]\,P[E|DE]$$

$$= \frac{7}{12} \times \frac{8}{13} \times \frac{5}{14} + \frac{7}{12} \times \frac{5}{13} \times \frac{8}{14} + \frac{5}{12} \times \frac{7}{13} \times \frac{8}{14} = \frac{15}{39}$$

y por el teorema de Bayes tenemos,

$$P[B \mid A] = \frac{\frac{7}{12} \times \frac{5}{13} \times \frac{8}{14} + \frac{5}{12} \times \frac{7}{13} \times \frac{8}{14}}{\frac{15}{39}} = \frac{\frac{10}{39}}{\frac{15}{39}} = \frac{2}{3} .$$

EJEMPLO 23 Una caja contiene 3 monedas; una corriente, otra de dos caras, y la tercera cargada tal que la probabilidad que se obtenga cara al lanzarla es 2/3. Se escoge una moneda al azar y se lanza. Si aparece cara se lanza la moneda de nuevo. Si aparece sello se escoge otra moneda entre las dos que quedan y se lanza.

Sea C el evento: "se escoge la moneda cargada".
Sea X el evento: "sale primero sello y después cara".
Se pide calcular:

(a) $P[C \mid X]$; (b) $P[\bar{X} \mid C]$.

SOLUCION Sean los eventos:

M_1 : "La moneda correcta"; $P[C] = P[S] = \frac{1}{2}$

M_2 : "La moneda de dos caras", $P[C] = 1$

M_3 : "La moneda cargada"; $P[C] = \frac{2}{3}$, $P[S] = \frac{1}{3}$

Consideremos el árbol de probabilidad siguiente

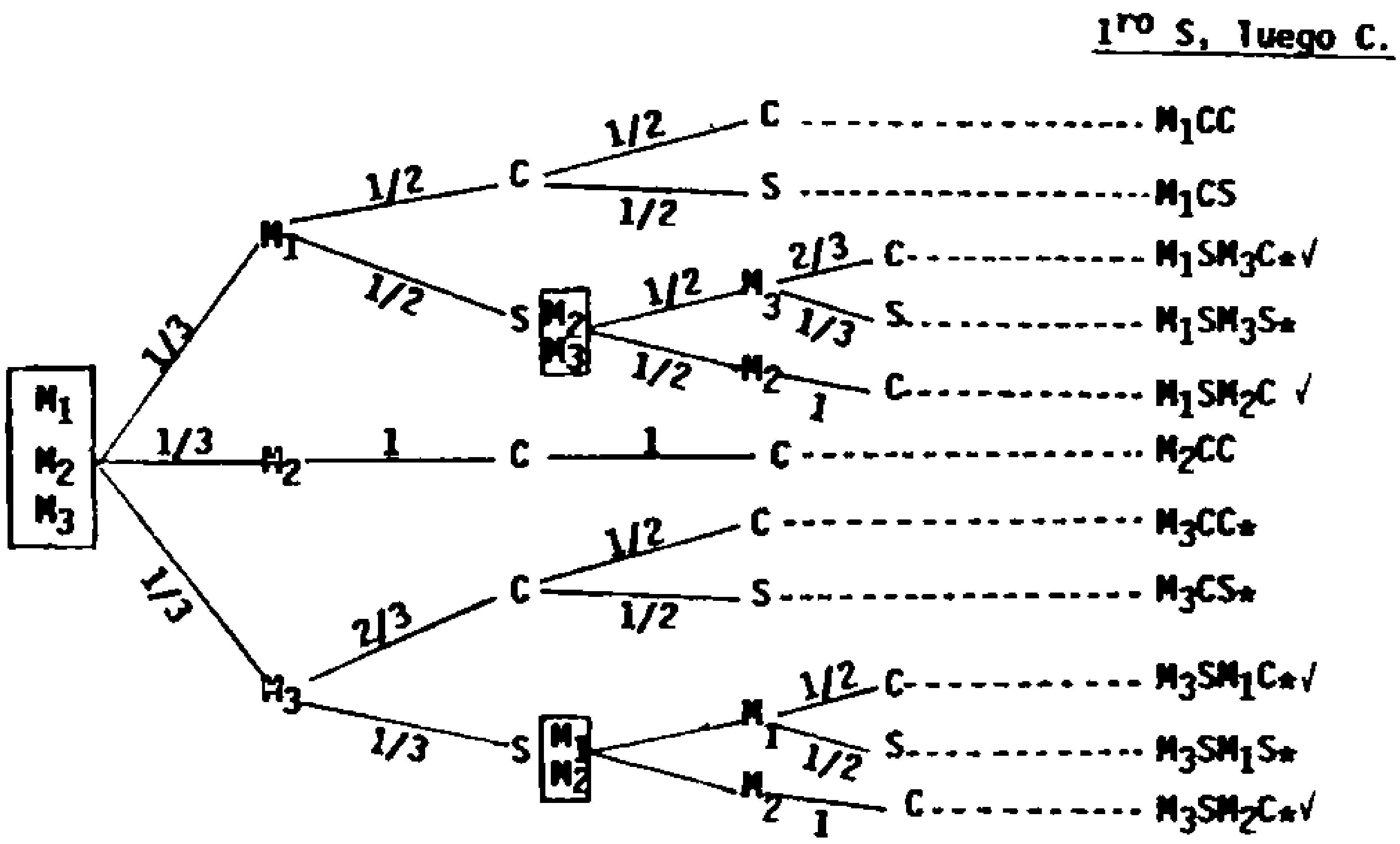


Fig. 1.8.27

Entonces se tiene que, los sucesos favorables al evento C son los que están marcados con * , ya que la moneda cargada puede escogerse la primera vez o la segunda.

Los sucesos favorables al evento X son los que están marcados con ✓ . Y los sucesos favorables al evento $C \cap X$ con los que tienen ambas marcas *✓ . - Por lo tanto:

(a) $$P[X] = \frac{1}{3} \times \frac{1}{2} \times \frac{1}{2} \times \frac{2}{3} + \frac{1}{3} \times \frac{1}{2} \times \frac{1}{2} \times 1 + \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} \times 1$$

$$= \frac{2}{36} + \frac{3}{36} + \frac{1}{36} + \frac{2}{36} = \frac{2}{9} .$$

y $$P[C \mid X] = \frac{P[C \cap X]}{P[X]}$$

$$= \frac{\frac{1}{3} \times \frac{1}{2} \times \frac{1}{2} \times \frac{2}{3} + \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} \times 1}{2/9}$$

$$= \frac{5/56}{2/9} = \frac{5}{8} .$$

(b) $$P[C] = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} \times 1 + 2\left[\frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} \times \frac{1}{2}\right] + \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} + \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} + \frac{1}{3}$$

$$\times \frac{1}{2} \times \frac{1}{2} \times \frac{2}{3} + \frac{1}{3} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{3}$$

$$= \frac{2}{36} + \frac{2}{36} + \frac{4}{27} + \frac{2}{27} + \frac{2}{36} + \frac{1}{36} = \frac{5}{12} .$$

y $P[\bar{X} \mid C] = 1 - P[X \mid C]$ prop. de probab. condicional.

$$= 1 - \frac{P[X \cap C]}{P[C]} = 1 - \frac{5/36}{5/12} = \frac{2}{3} .$$

EJEMPLO 24 En una caja hay 4 bolas, las que han sido colocados lanzando una moneda 4 veces. Si salió cara se colocó una bola blanca y si salió sello se colocó una bola roja. De la caja se extrae una bola y resultó ser blanca. - ¿Cuál es la probabilidad que en la caja queden al menos 2 bolas blancas?

SOLUCION Definimos los siguientes eventos:

B_i : "en la caja hay i bolas blancas ($i = 0,1,2,3,4$)".

B : "obtener una bola blanca al extraer una de la caja".

A : "en la caja quedan al menos dos bolas blancas"

El evento B se escribe así,

$$B = \bigcup_{i=1}^{4} B_i B$$

de donde, $P[B] = \sum_{i=1}^{4} P[B_i]\, P[B \mid B_i]$

y $A = B_3 \cup B_4$

debemos calcular $P[A \mid B] = \dfrac{P[B_3]\, P[B \mid B_3] + P[B_4]\, P[B \mid B_4]}{P[B]}$

Cálculo de $P[B_i]$ $i = 1,2,3,4$ y $P[B \mid B_i]$

B_1 : "en la caja hay 1 bola blanca y 3 rojas"

$P[B_1]$ = P[obtener en los lanzamientos de la moneda 1C y 3S] $= \dfrac{P_4^{1,3}}{2^4} = \dfrac{4}{2^4}$

B_2 : " en la caja hay 2 bolas blancas y 2 rojas"

$P[B_2]$ = P[obtener en los lanzamientos de la moneda 2C y 2S] $= \dfrac{P_4^{2,2}}{2^4} = \dfrac{6}{2^4}$

B_3 : " en la caja hay 3 bolas blancas y una roja"

$P[B_3]$ = P[obtener en los lanzamientos de la moneda 3C y 1S] $= \dfrac{P_4^{3,1}}{2^4} = \dfrac{4}{2^4}$

B_4 : "en la caja hay 4 bolas blancas"

$P[B_4]$ = P[obtener en los lanzamientos de la moneda 4C] $= \dfrac{1}{2^4}$.

Si ocurre B_1, $P[B \mid B_1] = \dfrac{1}{4}$;

Si ocurre B_2, $P[B \mid B_2] = \dfrac{1}{2}$;

Si ocurre B_3, $P[B \mid B_3] = \dfrac{3}{4}$;

Si ocurre B_4, $P[B \mid B_4] = 1$.

Luego, $P[B] = \dfrac{4}{2^4}\times\dfrac{1}{4}+\dfrac{6}{2^4}\times\dfrac{1}{2}+\dfrac{4}{2^4}\times\dfrac{3}{4}+\dfrac{1}{2^4}\times 1 = \dfrac{1}{2}$

por lo tanto, $P[A \mid B] = \dfrac{\dfrac{4}{2^4} \times \dfrac{3}{4} + \dfrac{1}{2^4} \times 1}{\dfrac{1}{2}} = \dfrac{1}{2}$, teorema de Bayes

La fig. 1.8.28 es el diagrama del árbol de probabilidades que ilustra la solución de este problema.

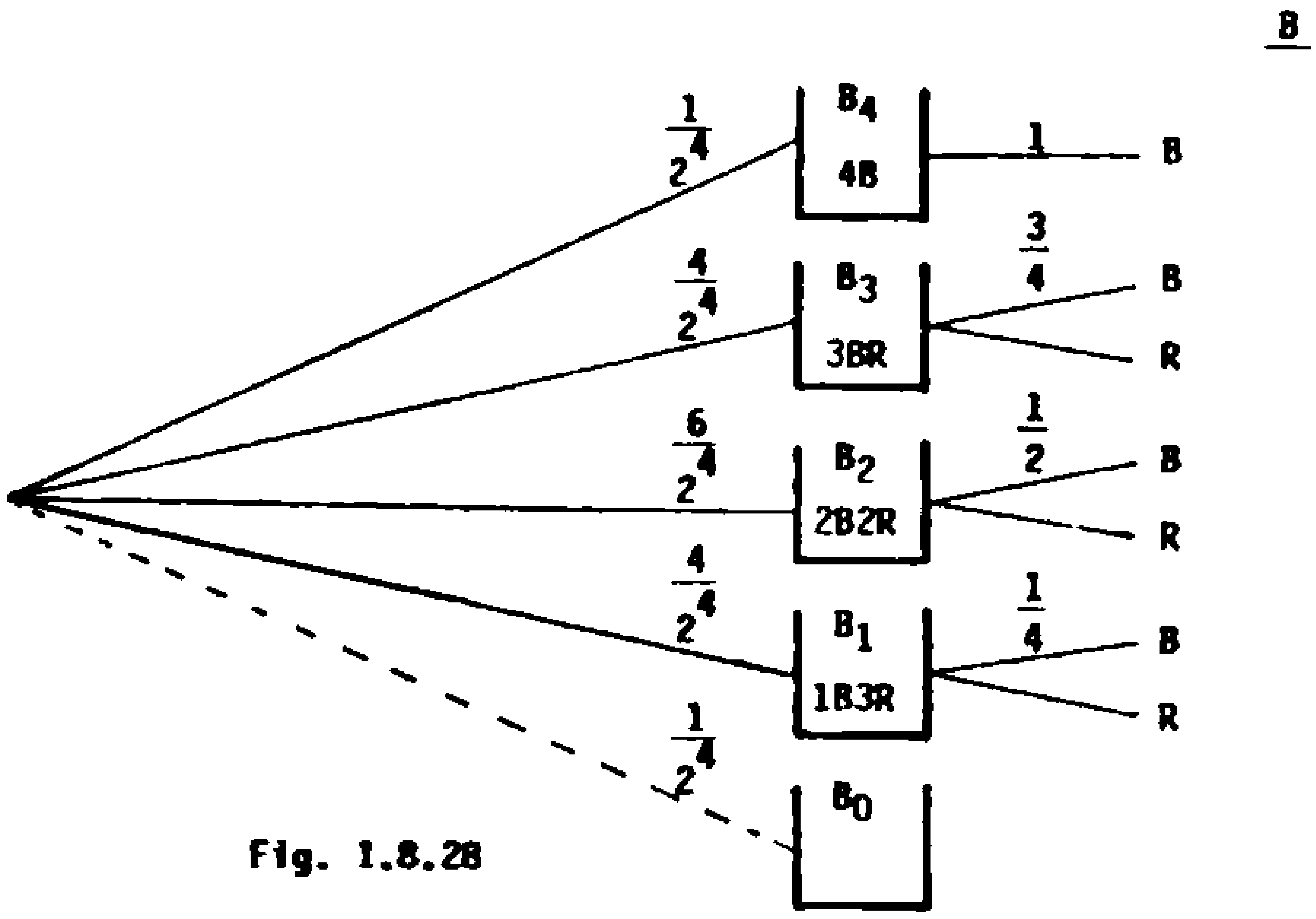


Fig. 1.8.28

PROBLEMAS 1.8

1. Una urna contiene 5 bolas blancas y 7 bolas negras. Se extrae una bola al azar; se pone fuera de la urna y su color no es visto; después se extrae otra bola aleatoriamente. ¿Cuál es la probabilidad que esta última sea - blanca?.

2. De un conjunto de fotos una persona A escoge 3 fotos de hombre y 2 de mujer. Otra persona B escoge 2 fotos de hombre y una de mujer. La persona A extrae de su grupo, al azar, 2 fotografías, que sin verlas, las hecha al grupo que tiene la persona B. Si de este nuevo grupo que ahora tiene la - persona B, se extrae al azar, una fotografía, determinar la probabilidad de que :
   (a) ésta sea una mujer.
   (b) éste sea un hombre.

3. Se tiene 5 urnas idénticas; dos de ellas de igual contenido, con 3 bolas blancas y 5 negras, y las otras tres teniendo cada una, 4 blancas y 3 negras. Se elige una urna aleatoriamente. De la urna elegida se extrae al azar una bola, ¿cuál es la probabilidad que sea blanca?.

4. Una urna contiene 3 bolas blancas, 4 negras y 5 rojas. Se extraen 3 bolas sin reposición. ¿Cuál es la probabilidad que sean del mismo color?

5. Se tiene 2 urnas A y B, con 5 bolas rojas, 3 blancas y una roja, 2 blancas respectivamente. Se lanza un dado, si se obtiene 3 ó 6, se pasa una bola de B a A, y luego se extrae una bola de A. En cualquier otro caso se pasa una bola de A a B, y luego se extrae una bola de B.
Determinar la probabilidad que:
(a) ambas bolas sean rojas.
(b) ambas bolas sean blancas.

6. Un restaurante ofrece, a elección, menús a base de carne, de pollo o de jamón. Si el cliente lo desea puede pedir vino tinto o vino blanco. Se sabe que las probabilidades que un cliente elija carne, pollo o jamón son, respectivamente, 0.60, 0.30 y 0.10. También se sabe que: las probabilidades que un cliente pida vino tinto, vino blanco o ningún vino, después de haber elegido carne, son 0.40, 0.10 y 0.50 respectivamente; las probabilidades, después que el cliente eligió pollo, son 0.05, 0.25 y 0.70 respectivamente; y las correspondientes probabilidades, después que el cliente eligió jamón, son 0.15, 0.20 y 0.65. Por último se sabe que la probabilidad que un cliente deje una buena propina es: 0.80 sí a pedido carne o vino tinto, 0.30 si ha pedido carne y vino blanco, 0.60 si ha pedido carne sin vino; 0.40 si ha pedido pollo y vino tinto, 0.80 si ha pedido pollo y vino blanco, 0.70 si ha pedido pollo pero no vino; o 0.70 si ha pedido jamón y vino tinto, 0.70 si ha pedido jamón y vino blanco, 0.50 si ha pedido jamón pero no vino. ¿Cuál es la probabilidad que un cliente deje una buena propina?

7. Tres urnas contienen respectivamente una bola blanca, dos negras; 2 blancas, una bola negra; 2 blancas y 2 bolas negras. De la primera urna se transfiere una bola a la segunda; después se transfiere una a la tercera urna y en seguida se extrae una bola de la tercera urna. ¿Cuál es la probabilidad que este sea blanca?

8. La urna U_1 contiene seis bolas blancas y cuatro negras. La urna U_2 contiene dos blancas y dos negras. Se transfieren dos bolas de la urna U_1 a la urna U_2. A continuación, se extrae sin reemplazo una muestra de tamaño 2 de la urna U_2. ¿Cuál es la probabilidad que la muestra contenga exactamente una bola blanca?

9. Considere una urna con 10 bolas de las cuales 5 son negras. Se elige al azar un entero n del conjunto $\{1,2,3,4,5,6\}$, luego se selecciona una muestra de tamaño n sin reemplazamiento de la urna Determinar la probabilidad que todas las bolas de la muestra sean negras.

10. Ud. escoge al azar uno de los números enteros del conjunto $\{1,2,3,4\}$. Luego lanza un tetraedro regular cuyas caras estan marcadas con los números 1,2,3,4, tantas veces indica el número que escogió.
Calcular la probabilidad que el puntaje total obtenido en los lanzamientos sea igual a 4.

11. En una sociedad hay 15 miembros, 5 mujeres y 10 hombres. A una reunión asistieron 13 miembros para elegir un presidente. ¿Cuál es la probabilidad de que sea una mujer la elegida?

12. Una urna contiene 5 bolas negras y x blancas. Si al sacar dos bolas, la probabilidad que ambas sean negras es 5/14, calcule x.

13. Se tiene 5 fusiles que, cuando son apuntados apropidamente y disparados, tienen probabilidades de dar en el blanco respectivamente como sigue: 0.5 0.6, 0.7, 0.8 y 0.9. Se escoge aleatoriamente uno de los fusiles, se apunta y dispara. ¿Cuál es la probabilidad de dar en el blanco?

14. Una compañía que fabrica tornillos, tiene tres factorías, F_1, F_2 y F_3. Las factorías F_2 y F_3 producen el mismo número de tornillos, mientras que F_1 produce el doble de las de F_2 . Por experiencia pasada se sabe, que el 2% de los tornillos producidos por F_1 y F_2 respectivamente son defectuosos, en tanto que el 4% de los fabricados por F_3 son defectuosos. Los tornillos producidos por las tres factorías se guardan en un mismo almacén. Si se escoge aleatoriamente un tornillo del almacén. ¿Cuál es la probabilidad que sea defectuosa?

15. Una caja contiene 3 monedas, dos normales y una con dos caras. Se elige una moneda al azar y se efectúa una tirada. Si sale cara se tira otra vez Si sale sello, entonces se elige una de las otras dos de la caja y se tira.
   (a) Determinar la probabilidad que salgan dos caras
   (b) Si se tira la misma moneda las dos veces, hallar la probabilidad de que sea la moneda con dos caras.

(c) Hallar la probabilidad que aparezca sello las dos veces.

16. Suponga que B_1, B_2 y B_3 son eventos mutuamente excluyentes. Si

$$P[B_j] = \frac{1}{3} \quad y \quad P[A \mid B_j] = \frac{j}{6}, \; j = 1,2,3. \quad \text{Hallar } P[A]$$

17. Una urna I contiene dos bolas blancas y tres azules, una segunda urna II contiene tres bolas blancas y cuatro azules, se extrae aleatoriamente una bola de la urna I y se coloca en la segunda, luego se extrae aleatoriamente una bola de esta; calcular la probabilidad que ésta última sea blanca.

18. Una caja contiene 5 monedas; cuatro de ellas son normales y la quinta con cara en ambos lados. Se elige aleatoriamente una moneda y luego se lanza dos veces. ¿Cuál es la probabilidad de obtener dos caras?.

19. Por la noche dos coches se aproximan uno al otro en una autopista. Si ninguno de los dos choferes esta borracho, ambos pasarán a salvo con una probabilidad de 0.999. Cada uno puede estar borracho con una probabilidad de 0.1, la probabilidad que ambos están borrachos es de 0.01. Si sólo el chofer A está borracho, pasarán a salvo con una probabilidad de 0.7. Si sólo el chofer B está borracho, pasarán a salvo con una probabilidad de 0.8. Si ambos están borrachos, pasarán sin peligro con una probabilidad de 0.4. ¿Cuál es la probabilidad que pasen a salvo?

20. En una urna A, hay 5 bolillas rojas y 5 negras; en otra urna B, hay 3 bolillas rojas y 6 negras, y en una C, hay 4 bolillas rojas y 2 negras. Se extrae una bolilla de A y se coloca en B, y de esta se extrae una bolilla que se coloca en C. De C se extrae una bolilla. Calcular la probabilidad de sacar una bolilla negra de C.

21. Se ha lanzado un dado 2 veces. El número obtenido en la primera vez es el número de bolas blancas colocadas en la urna y el número obtenido en la segunda vez, es el número de bolas negras colocadas en la misma urna. Se sabe también que el número total de bolas en la urna es 8.
¿Cuál es la probabilidad que la urna contenga exactamente 5 bolas blancas?

22. En una caja hay 3 monedas correctas y una no correcta tal que: **P**[C] = 1/3 **P**[S] = 2/3. Una persona elige al azar dos monedas de la caja y las coloca en otra caja; luego extrae al azar una moneda de esta segunda caja y la arroja, resultándoles sello. Determinar la probabilidad que esta moneda sea la no correcta.

23. Una moneda es tal que: $P[C] = 2/3$ y $P[S] = 1/3$. Se lanza dicha moneda Si aparece, cara se selecciona al azar un número del 1 al 9; si sale sello, se selecciona al azar un número del 1 al 5. Determinar:
   (a) la probabilidad de obtener un número par.
   (b) Si se selecciona un número par, ¿cuál es la probabilidad que haya salido sello?

24. Tres máquinas I, II y III manufacturan el 30%, 30% y 40% de la producción total de un cierto artículo. Las máquinas producen 4%, 3% y 2% de productos defectuosos, respectivamente. Se toma un artículo al azar, se prueba y resulta ser defectuoso. ¿Cuál es la probabilidad que haya sido manufacturado por la máquina I? ¿La máquina II? ¿y la III?.

25. De los alumnos del primer año de un determinado programa académico, se sabe que el 40% asistió a centros secundarios privados y el 60% asistió a centros estatales. El registro de matrícula señala que al final del curso alcanzaron una nota media A el 30% de los alumnos que asistieron a centros secundarios privados y sólo el 20% de los que asistieron a centros estatales. Al final del ciclo, se elige al azar un alumno de dicho curso y tiene nota media A. ¿Cuál es la probabilidad que el alumno hubiera asistido a un centro estatal?

26. (Problema de Bertrand). Se dan tres cajas cada una con dos cajones. Cada uno de los cajones de una caja contiene una moneda de oro; cada uno de los de la otra, una moneda de plata. En cuanto a los de la tercera caja, uno de sus cajones contiene una moneda de plata y el otro una moneda de oro. Escogemos una caja al azar, abrimos un cajón, y la moneda que en él encontramos es de oro. ¿Cuál es la probabilidad que la moneda en el otro cajón de esa caja sea de plata?

27. Una urna contiene dos bolas. La urna se llenó lanzando al aire una moneda dos veces y poniendo en la urna una bola blanca por cada cara y una bola negra por cada sello. Sacamos una de las dos bolas de la urna y nos encontramos que es blanca. Encuéntrese la probabilidad que la otra bola sea también blanca.

28. Una estación meteorológica suele acertar el 60% de las veces que pronostico día lluvioso, también la probabilidad de acertar en su pronóstico dado que indica tiempo no lluvioso es de 0.8. Se sabe que la probabilidad

que llueva en un día cualquiera es de 0.25.

(a) ¿Cuál es la probabilidad que la estación dé un pronóstico correcto en un día cualquiera?

(b) ¿Cuál es la probabilidad que llueva dado que se sabe que el pronóstico está correcto?

29. La caja A contiene nueve cartas numeradas de 1 a 9. La caja B contiene cinco cartas numeradas de 1 a 5. Se elige una caja al azar y de ella se extrae una carta. Si la carta extraída es par. Calcular la probabilidad de que la carta provenga de la caja A.

30. La urna U_1 contiene una bola blanca y una negra. La urna U_2 sólo contiene bolas blancas. Se elige una urna al azar de la que se saca una primera bola que se repone, y se saca luego otra bola. Ambas resultan blancas.

(a) ¿Cuál es la probabilidad que ambas bolas fueron extraídas de la urna U_2 ?

(b) Si de la urna elegida de hacen tres extracciones con reposición. ¿Cuál es la probabilidad que las tres bolas fueron extraídas de la urna U_1 ?

31. La urna I tiene 3 bolas blancas y 7 negras. La urna II tiene 20 bolas de las cuales algunas son blancas y las demás son negras. Se escoge una urna al azar y de ella se extrae una bola, encontrándose que es blanca. Si la probabilidad que esta bola blanca provenga de la urna I es igual a 1/3 determinar el número de bolas blancas que existían originalmente en la urna II.

32. Una caja contiene 6 monedas de las cuales 4 son normales y las dos restantes con cara en ambos lados. Se elige aleatoriamente una moneda y luego se lanza dos veces. ¿Cuál es la probabilidad que se haya elegido una moneda normal, si se obtuvieron dos caras en los lanzamientos?

33. Un fabricante recibe el 60% de sus materias primas del distribuidor A y el resto del B. Por experiencia sabe que el 92% de los costales de materia prima llegan en buenas condiciones. ¿Cuál es la probabilidad que un costal dañado sea del distribuidor A? ¿Que sea del B?. Un estudio posterior revela que el 95% de los costales del distribuidor A llega en buen estado, mientras que sólo el 87.5% de B es de buena calidad. Con esta información, ¿cuál es la probabilidad que un costal dañado proceda de A?

34. Un fabricante recibe el 45% de su material para un transistor, de la compañía A, el 35% de la compañía B y el resto de la compañía C. Sabe por experiencia que el 1% del material de la compañía A será defectuoso, y que el 2% del material de la compañía B y C estará en malas condiciones también... En un lote que contenga material de las tres compañías hay mil transistores. ¿Cuál es la probabilidad que un transistor esté defectuoso?. ¿Cuál es la probabilidad que si un transistor estuviese defectuoso contase con material de la compañía B?

35. Un aparato especial para medir el contenido alcohólico en la sangre de una persona arrojó el siguiente resultado: de 500 voluntarios, 240 estaban borrachos (el nivel de alcohol en la sangre era de 0.0015 o más). Los mismos 500 voluntarios se sometieron a una prueba sanguínea inmediatamente después, encontrándose 280 personas con un nivel de 0.0015 o más. Después se determinó que 180 personas resultaron estar borrachos en ambas pruebas ¿Qué porcentaje de personas resultaron estar ebrios sin que lo indicara el aparato?. Supóngase que una persona realmente estuviera borracha y que pasara la prueba en el aparato. Según la información dada anteriormente, ¿cuál es la probabilidad que la prueba resultara positiva?.

36. Una urna contiene inicialmente una bola blanca y una bola negra. Una bola es extraída al azar y es reemplazada por dos bolas de su mismo color. Luego una de las tres bolas que ahora hay en la urna es axtraida al azar y es reemplazada por dos de su mismo color. Después una de las cuatro bolas que ahora hay en la urna es extraída al azar y es reemplazada por dos de su mismo color, etc. ¿Cuál es la probabilidad que:
(a) Las tres primeras bolas son blancas?
(b) La tercera bola es blanca?
(c) Hay 3 bolas blancas en las cuatro primeras extracciones?.

37. Todas las noches, el señor Pérez llega tarde a su casa. La señora Pérez que es una buena esposa, le deja encendida la luz de la entrada a la casa La probabilidad que el señor Pérez llegue borracho es 0.60. Si llega borracho, hay una probabilidad de 0.90 que olvide apagar la luz en tanto que ésta es sólo de 0.05 si llega sobrio.
(a) ¿Cuál es la probabilidad que el señor Pérez apague la luz en una nocne cualquiera?

(b) Dado que el señor Pérez apagó la luz una cierta noche. ¿Cuál es la probabilidad que haya llegado borracho?

38. El profesor López dicta un curso de Estadística y quiere tomar una prueba en cada clase. Sabedor que a veces se olvida de ir a hacer su clase, ha dado instrucciones a su jefe de prácticas que se haga cargo de la clase cuando él está ausente. Si el profesor López hace la clase, la probabilidad es 0.70 que tome la prueba en tanto que si el jefe de práctica hace la clase, ésta probabilidad es de sólo 0.10. Si el profesor López falta el 80% de la clases.
    (a) ¿Cuál es la probabilidad que haya una prueba en una clase dada?
    (b) Suponiendo que hubo prueba en una clase determinada, ¿Cuál es la probabilidad que el profesor López haya estado ausente?.

39. La compañía COMPUT-PERU está considerado comercializar una calculadora electrónica. De acuerdo con una investigación hecha en el mercado, la probabilidad que el producto tenga éxito es 0.80 si una firma competidora no introduce un producto similar en el mercado, en tanto que la probabilidad de éxito es 0.30 si la firma competidora comercializa el producto similar. Además, la compañía estima que hay una probabilidad de 0.40 que la firma competidora comercialice el producto. Dado que el producto de la compañía COMPUT-PERU tuvo éxito, ¿Cuál es la probabilidad que la firma competidora haya comercializado su producto?

40. Uno de dos peritos mercantiles verifica el estandar de un artículo. La probabilidad que el artículo caiga en manos del primer perito es igual a 0.55 y al segundo, 0.45. La probabilidad que el artículo estandarizado sea reconocido como tal por el primer perito es igual a 0.9, y por el segundo, 0.98. Durante la verificación el artículo fue reconocido como estandarizado. Hallar la probabilidad que el artículo lo haya examinado el segundo perito.

41. Juan tiene dos bolsas; la bolsa I, con tres bolas rojas y dos blancas; la bolsa II, con una bola roja y cuatro blancas. Juan escogió una bola de la bolsa I al azar y lo coloca en la bolsa II, luego escoge una bola de la bolsa II y encuentra que es roja, ¿Cuál es la probabilidad de que la bola transferida de la bolsa I a la II haya sido roja?.

42. Una persona tiene tres bolsas azules y dos verdes. Cada bolsa azul contiene

bolas rojas y blancas en una razón de 4 a 1, y cada bolsa verde contiene bolas rojas y blancas en una razón de 1 a 4. Tal persona escoge una bolsa al azar, y también al azar una bola de la bolsa elegida, y ve que es blanca, ¿Cuál es la probabilidad que haya sacado una bolsa verde?.

43. En el problema 17. Se sabe que de la urna II se extrajo una bola blanca, ¿Cuál es la probabilidad que de la urna I se halla pasado una bola azul?

44. En el ejemplo 8 de 1.8. Si se sabe que todas las bolas que se extrajeron fueron blancas. ¿Cuál es la probabilidad de haber obtenido 2 en el dado?.

45. La probabilidad que un accidente aéreo, el cual es debido a fallas estructurales, sea diagnosticada correctamente es 0.85 y la probabilidad que un accidente aéreo el cual no es debido a fallas estructurales sea diagnosticado incorrectamente como debido a fallas estructurales es 0.35. Si el 30% de todos los accidentes aéreos son debidos a fallas estructurales, - encontrar la probabilidad, que un accidente aéreo es debido a fallas estructurales, dado que ha sido diagnosticado como debido a fallas estructurales.

46. Una compañía de petróleo, debe decidir, si taladra o no, un lugar determinado, que la compañía tiene bajo contrato. Por investigaciones geológicas practicadas, se sabe que existe una probabilidad de 0.45 que una formación tipo I se extiende debajo del lugar prefijado para taladrar: 0.30 de probabilidad que exista una formación de tipo II y de 0.25 de tipo III. Estudios anteriores indican que el petróleo se encuentra en un 30% de las veces en las formaciones de tipo I y un 40% en las del tipo II y 20% en - las de tipo III. Determinar la probabilidad que si no se encontro petróleo, la perforación fue hecha en la formación de tipo I.

47. Se sabe que la probabilidad que una persona tuberculosa sometida a un examen radiológico se le diagnostique como tal es 0.95 y que la probabilidad que una persona no tuberculosa sometida a un examen radiológico se le diagnostique erróneamente como tuberculoso es 0.002. Se sabe, además, que el 0.1% de los adultos residentes en cierta ciudad son tuberculosos. Si a uno de los residentes, seleccionado al azar, se le diagnostica tuberculosis. ¿Cuál es la probabilidad que realmente sea tuberculoso?

48. Se tienen dos urnas U_1 y U_2 . La primera tiene 2 bolas blancas y 3 negras la segunda 2 bolas blancas y 3 rojas. Se extrae al azar una bola de U_1 y se pasa a U_2 . Luego se extrae una bola de U_2 y se pasa a U_1 . Finalmente se extrae al azar 2 bolas de U_1 y resultan ser blanca y negra. Determinar la probabilidad que U_1 no tenga ninguna bola roja.

49. Una fábrica de unidades de aire acondicionado recibe 70 % de sus termostatos de la compañía A, 20 % de la compañía B y el resto de sus termostatos de la compañía C. Por experiencia pasado se sabe que la compañía A produce 1/2 % de termostatos defectuosos; la compañía B, 1 % y la compañía C, 1.5 % . Se selecciona al azar una unidad de aire acondicionado de la línea de producción y resulta defectuosa. ¿Cuál es la probabilidad que el termostato haya sido producido por la compañía B?

50. Se sabe que una empresa industrial utiliza cuatro máquinas en la fabricación de cierto producto y que la producción diaria de cada una de ellas es, respectivamente 1000, 1200, 1800 y 2000 piezas. Se sabe además que en promedio, el 1 % de la producción de la primera máquina es "defectuosa", el 1/2 % de la producción de la segunda es defectuosa, el 1/2 % de la producción de la tercera es defectuosa y el 1 % de la producción de la cuarta es defectuosa. Si de la producción de cierto día se extrae, al azar, una pieza que resulta ser defectuosa. ¿Cuál es la probabilidad que dicha pieza proceda: de la cuarta máquina? ¿de la tercera máquina? ¿de la segunda máquina? ¿de la primera máquina?

51. Se lanza al aire una moneda. Si sale cara se introduce una bola blanca en una urna; si sale sello, la bola que se introduce será negra. Esto se hace cuatro veces. Finalmente se extrae dos bolas y resultan ser blancas, ¿Cuál es la probabilidad que las otros dos bolas sean blanca y negra?

52. De una urna que contiene seis bolas blancas y cuatro negras, se transfieren cinco de ellas a una segunda urna que se encuentra vacía. Se toman de ella tres bolas y se ponen en una caja vacía. Se extrae una bola de la caja y resulta ser blanca . ¿Cuál es la probabilidad que exactamente cuatro de las bolas transferidas de la primera urna a la segunda hayan sido blancas?

53. La probabilidad que un accidente de aviación debido a fallas mecánicas sea diagnosticado correctamente es 0.72 y la probabilidad que un accidente de aviación que no se debe a fallas mecánicas sea diagnosticado in-

correctamente como que se debió a fallas mecánicas es 0.12. Si 40% de todos los accidentes de aviación se deben a fallas mecánicas, ¿cuál es la probabilidad que un accidente de aviación que se diagnosticó como debido a fallas mecánicas sea realmente a esta causa?

54. Tres empaquetadoras se emplearon en una juguetería durante el período de Navidades. María, que empaqueta 40% de todos los juguetes, se olvida de quitar la etiqueta con el precio 1 vez en 50; Juana, que empaqueta el 30% de todos los juguetes, se olvida de quetar la etiqueta con el precio una vez en 10, y Elena, que empaqueta el resto de los juguetes, se olvida de quitar la etiqueta con el precio 1 vez en 20. Dado que un cliente se quejó de que una etiqueta con el precio no fue quitada de un regalo antes de haber sido empaquetado, ¿cuál es la probabilidad que María cometiera el error?

1.9 EVENTOS INDEPENDIENTES Y SECUENCIA DE EXPERIMENTOS INDEPENDIENTES

Hemos visto que, si los eventos A y B son mutuamente excluyentes como indica la figura 1.9.1; O sea $AB = \phi$, y si $P[A] > 0$, $P[B] > 0$, se tiene

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = 0 \qquad y \qquad P[B \mid A] = \frac{P[A \cap B]}{P[A]} = 0$$

por otro lado tenemos, si $B \subset A$ como muestra la figura 1.9.2, entonces

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B]}{P[B]} = 1$$

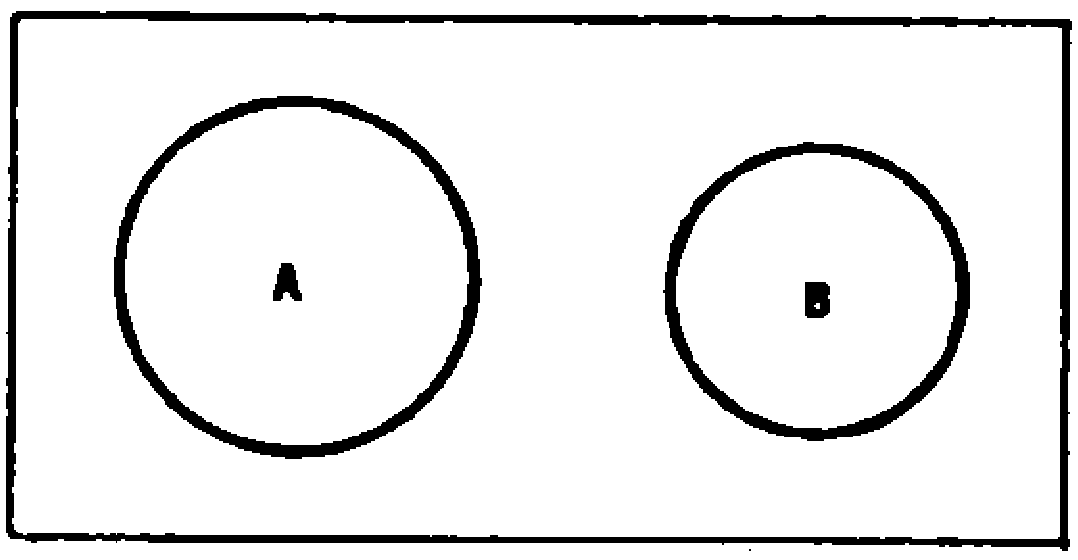


Fig. 1.9.1. Eventos mutuamente excluyentes

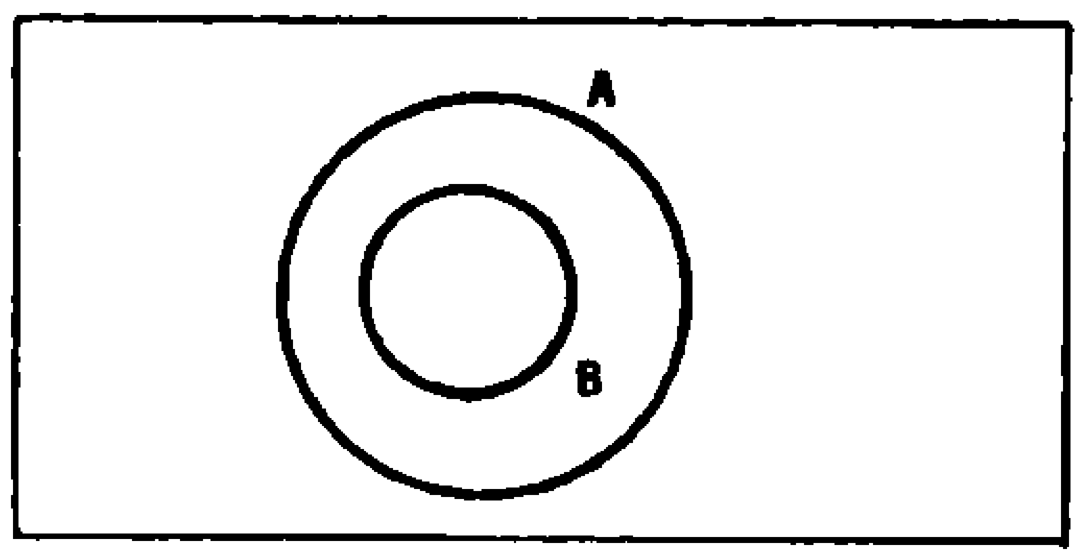


Fig. 1.9.2. evento B subconjunto de A

En el primer caso, los eventos A y B no pueden ocurrir simultáneamente, así que el conocimiento de la ocurrencia de B nos dice que A no ocurre (o viciversa). En el segundo caso si ocurre B, debe ocurrir A. Y en general hemos visto al definir probabilidad condicional que la ocurrencia de un evento con diciona la probabilidad de la ocurrencia de un segundo evento. Sin embargo hay muchos casos donde los eventos están totalmente sin conexión, y la ocurrencia de uno de ellos no cambia la probabilidad de la ocurrencia del otro, en este caso se dice que son *EVENTOS INDEPENDIENTES*. No obstante antes de dar la definición formal de independencia, consideraremos algunos ejemplos.

Consideraremos por ejemplo, el experimento de lanzar una moneda dos veces y observar la secuencia de caras y sellos. Entonces el espacio muestral es

$$\Omega = \{CC, CS, SC, SS\}$$

Consideremos los siguientes eventos:

A : "El primer resultado una cara".

B : "El segundo lanzamiento resulta una cara".

Entonces, $A = \{CC, CS\}$ y $B = \{CC, SC\}$. Por lo tanto

$$P[A] = \frac{1}{2}, \qquad P[B] = \frac{1}{2}$$

Supongamos ahora que ha ocurrido el evento B, entonces

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{1/4}{1/2} = \frac{1}{2} = P[A]$$

Es decir, la ocurrencia del evento B no afecta la probabilidad de la ocurrencia de A. Es decir, la probabilidad de A no depende del conocimiento de la ocurrencia o no ocurrencia del evento B.

Así informalmente hablando, dos eventos A y B son *independientes*, si la ocurrencia de uno de ellos np efecta la probabilidad de la ocurrencia del otro. En símbolos significa que:

$$P[A \mid B] = P[A] \qquad \text{Si} \qquad P[B] > 0 \qquad (1)$$

$$P[B \mid A] = P[B] \qquad \text{si} \qquad P[A] > 0 \qquad (2)$$

Por ejemplo, si queremos hallar la probabilidad de la ocurrencia de A y B, esto es $P[A \cap B]$, tenemos

$$P[A \cap B] = P[B]\, P[A \mid B] = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

En general, de la primera igualdad y la regla de la multiplicación se tiene

$$P[A \cap B] = P[B]\, P[A \mid B] = P[B]\, P[A]$$

y de la segunda igualdad y la regla de la multiplicación se tiene el mismo resultado.

$$P[A \cap B] = P[A]\,P[B \mid A] = P[A]\,P[B]$$

Esto nos lleva a la siguiente definición.

DEFINICION 1.9.1 Los eventos A y B en Ω son ***independientes*** si, y solamente si se cumple una de las siguientes condiciones

(*i*) $P[A \cap B] = P[A]P[B]$

(*ii*) $P[A \mid B] = P[A]$, si $P[B] > 0$. (*iii*) $P[B \mid A] = P[B]$, si $P[A] > 0$

En otro caso se dice que los eventos A y B ***son dependientes***. Los eventos independientes son llamados algunas veces, ***Estadísticamente independientes, estocásticamente independientes o independientes en el sentido probabilístico***

EJEMPLO 1 Considere el lanzamiento simultáneo de una moneda y un dado. Sean los eventos; A: "Se obtiene cara en la moneda", y B: "en el dado sale 6". Verificar que A y B son independientes.

SOLUCION
$$\Omega = \left\{ \begin{matrix} (C,1) & (C,2) & \ldots & (C,6) \\ (S,1) & (S,2) & \ldots & (S,6) \end{matrix} \right\}$$

entonces, $A = \{(C,1), (C,2) \ldots (C,6)\}$

$B = \{(C,6), (S,6)\}$

$$A \cap B = \{C, 6\}, \quad \text{luego} \quad P[A \cap B] = \frac{1}{12} \qquad (1)$$

$$P[A] = \frac{1}{2}, \quad P[B] = \frac{1}{6}, \quad \text{por lo tanto}$$

$$P[A]\,P[B] = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12} \qquad (2)$$

de (1) y (2), $P[A \cap B] = P[A]\,P[B]$

lo que indica que dichos eventos son independientes.

EJEMPLO 2 En un estudio de una enfermedad al pulmón se examinan 10000 personas mayores de 60 años. Se halla que 4,000 personas de este grupo son fumadores. Entre los fumadores 1800 padecen de desórdenes pulmonares. Entre los que no fuman 1500 tienen desórdenes pulmonares. ¿Son los eventos "fumadores" y "desórdenes pulmonares" independientes?

SOLUCION Sea los eventos :

A: "la persona elegida al azar es fumador".
B: "la persona tiene desórdenes pulmonares".

Entonces,

$$P[A] = \frac{4000}{10000} = 0.4 \quad y \quad P[B] = \frac{3300}{10000} = 0.33$$

La probabilidad condicional de fumadores dado un desorden pulmonar es,

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{\frac{1800}{10000}}{\frac{3300}{10000}} = \frac{1800}{3300} = 0.55 \neq P[A]$$

por lo tanto, los eventos A y B no son independientes.

Una consecuencia inmediata de la definición es el teorema siguiente.

TEOREMA 1.9.1 Si A y B en Ω son eventos independientes, entonces

(*i*) A y $\bar{B}$ son independientes;
(*ii*) $\bar{A}$ y B son independientes;
(*iii*) $\bar{A}$ y $\bar{B}$ son independientes.

DEMOSTRACION (*i*) $P[A \cap \bar{B}] = P[A]\,P[\bar{B} \mid A]$

$= P[A]\,[1 - P[B \mid A]]$ teorema 1.7.2
$= P[A]\,[1 - P[B]]$ Hipotesis
$= P[A]\,P[\bar{B}]$.

Lo que demuestra que A y $\bar{B}$ son independientes.

(*ii*) se demuestra similarmente.

(*iii*) $P[\bar{A}\,\bar{B}] = P[\bar{A}]\,P[\bar{B} \mid \bar{A}] = P[\bar{A}]\,[1 - P[B \mid \bar{A}]]$ teorema 1.7.2

$= P[\bar{A}]\,[1 - P[B]] = P[\bar{A}]\,P[\bar{B}]$ por (ii)

EJEMPLO 3 Sean A y B dos eventos independientes, tales que la probabilidad que ocurran simultáneamente es de 1/6 y la probabilidad que ninguno ocurra es de $\frac{1}{3}$. Encuentre $P[A]$ y $P[B]$

SOLUCION Del enunciado se tiene

1. $P[AB] = \frac{1}{6}$, $P[\bar{A}\,\bar{B}] = \frac{1}{3}$

2. A y B eventos independientes; entonces

$$P[A\ B] = P[A]\ P[B] = \frac{1}{6}$$

3. A y B independientes, entonces $\bar{A}$ y $\bar{B}$ independientes, teorema 1.9.1 (*iii*) . O sea

$$\frac{1}{3} = P[\bar{A}\ \bar{B}] = P[\bar{A}]\ P[\bar{B}] = [1 - P[A]][1 - P[B]] \quad \text{teorema 1.6.2}$$

$$= 1 - P[A] - P[B] + P[A]\ P[B]$$

$$= 1 - P[A] - P[B] + \frac{1}{6} \qquad \text{por paso 1}$$

de donde $P[A] + P[B] = \frac{5}{6}$ ó $P[B] = \frac{5}{6} - P[A]$

4. Reemplazando (3) en (2) obtenemos

$$P[A]\left[\frac{5}{6} - P[A]\right] = \frac{1}{6}$$

$$\left[P[A]\right]^2 - \frac{5}{6}P[A] + \frac{1}{6} = 0\ .$$

Resolviendo la ecuación de segundo grado se obtiene

$$P[A] = \frac{1}{2}\ , \qquad y \qquad P[A] = \frac{1}{3}$$

5. El conjunto solución será $\{ P[A] = \frac{1}{2}, \ P[B] = \frac{1}{3} \}$

o $\{ P[A] = \frac{1}{3}, \ P[B] = \frac{1}{2} \}$

NOTA En la práctica existen muchas situaciones donde no se puede con facilidad determinar si dos eventos son independientes; sin embargo en muchos casos los requerimientos puede justificarse intuitivamente de la consideración física del experimento, como veremos en ejemplos posteriores. Por ejemplo, l. falla de uno de los motores de un avión de cuatro motores es independiente d la falla de los otros; en un concurso de tiro al blanco, la probabilidad de acertar por uno de los competidores es independiente de que acierten o no los otros concursantes, etc.

EJEMPLO 4 Para señalar las emergencias que pueden presentarse por la avería de alguno de los equipos en una sala de operaciones, se han instalado dos instrumentos indicadores que funcionan independientemente. La probabilidad de que el indicador accione durante la avería es igual a 0.95 para el primero de ellos y 0.9 para el segundo. Hallar la probabilidad que durante una avería accione sólo un indicador.

SOLUCION 1 Sea los siguientes eventos:

A: " el primer indicador acciona".

B: "el segundo indicador acciona".

$E = A\bar{B}$: "sólo acciona el indicador A"

$F = \bar{A}B$: "sólo acciona el indicador B".

2. Para hallar la probabilidad que accione sólo un indicador, debemos calcular la probabilidad de E o F. Es decir, $P(E \cup F)$.

3. Puesto que E y F son eventos mutuamente excluyentes

$$P[E \cup F] = P[E] + P[F]$$

$$P[E \cup F] = P[A\bar{B}] + P[\bar{A}B]$$

4. Los eventos A Y B son independientes, entonces por el teorema 1.9.1 partes (i) y (ii) los eventos A y $\bar{B}$, así acomo $\bar{A}$ y B son independientes respectivamente. Luego

$$\begin{aligned} P[E \cup F] &= P[A]P[\bar{B}] + P[\bar{A}]P[B] \\ &= (0.95)(0.1) + (0.05)(0.9) \\ &= 0.095 + 0.045 = 0.14. \end{aligned}$$

CONSECUENCIA 1 Si A y B son eventos cualesquiera en Ω independientes, el teorema 1.6.4 , se escribe

$$P[A \cup B] = P[A] + P[B] - P[A]P[B].$$

TEOREMA 1.9.2 'Si A y B son eventos cualesquiera en Ω independientes, entonces

$$P[A \cup B] = 1 - P[\bar{A}]P[\bar{B}] = 1 - [1 - P[A]][1 - P[B]]$$

DEMOSTRACION

$$\begin{aligned} P[A \cup B] &= 1 - P[\overline{A \cup B}] && \text{teorema 1.6.2;} \\ &= 1 - P[\bar{A}\bar{B}] && \text{ley de DE-Morgan ;} \\ &= 1 - P[\bar{A}]P[\bar{B}] && \text{teorema 1.9.1 } (iii) \\ &= 1 - [1 - P[A]][1 - P[B]] && \text{teorema 1.6.2.} \end{aligned}$$

EJEMPLO 5 Un tirador acierta el 80 % de sus disparos y otro (en las mismas condiciones de tiro), el 70 % . ¿Cuál es la probabilidad de dar en el blanco cuando ambos tiradores disparan sobre el simultaneamente? Se considera que se ha dado en el blanco cuando por lo menos, una de las dos balas ha hecho -

impacto en el .

SOLUCION Definimos los siguientes eventos

B_i : "el tirador i acierta en el blanco, $i = 1,2$"

$P[B_1] = 0.80$, $P[B_2] = 0.70$

PRIMERA FORMA observe que los eventos B_1 y B_2 son independientes,

$$\begin{aligned} P[B_1 \cup B_2] &= P[B_1] + P[B_2] - P[B_1 B_2] \\ &= P[B_1] + P[B_2] - P[B_1] P[B_2] , \quad \text{Consecuencia 1} \\ &= 0.80 + 0.70 - (0.80)(0.70) \\ &= 0.94 . \end{aligned}$$

SEGUNDA FORMA

$$\begin{aligned} P[B_1 \cup B_2] &= 1 - [1 - P[B_1]][1 - P[B_2]] \qquad \text{Teorema 1.9.2} \\ &= 1 - [1 - 0.80][1 - 0.70] \\ &= 0.94 . \end{aligned}$$

El concepto de eventos independientes puede extenderse a más de dos eventos .

DEFINICION 1.9.2 Tres eventos A, B y C en Ω se dice que son mutuamente independientes si y solamente si, cumplen las siguientes condiciones

(i) Ellos son independientes por pares. Es decir :

$$P[A B] = P[A] P[B] \quad ; \quad P[A C] = P[A] P[C] \text{ y}$$

$$P[B C] = P[B] P[C] .$$

(ii) $P[A B C] = P[A] P[B] P[C]$

DEFINICION 1.9.3 Los eventos $A_1, A_2, \ldots , A_n$ en Ω son independientes - si sólo si

$$P[A_i A_j] = P[A_i] P[A_j] , \quad i \neq j$$

$$P[A_i A_j A_k] = P[A_i] P[A_j] P[A_k] , \quad i \neq j , \ j \neq k , \ i \neq k$$

$$\vdots$$

$$P\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n} P[A_i] .$$

EJEMPLO 6 Se lanza una moneda tres veces. Sean los eventos,

A_i : "se obtiene cara en el i-ésimo lanzamiento". El espacio muestral es

$$\Omega = \{CCC, CCS, CSC, SCC, CSS, SSC, SCS, SSS\}$$

Así, A_1 : "cara en el primer lanzamiento".

A_2 : "Cara en el segundo lanzamiento".

A_3 : "Cara en el tercer lanzamiento".

Es decir

$$A_1 = \{CCC, CCS, CSC, CSS\}$$

$$A_2 = \{CCC, CCS, SCC, SCS\}$$

$$A_3 = \{CCC, CSC, SCC, SSC\}$$

Entonces, $P[A_i] = \dfrac{1}{2}, \quad i = 1,2,3$

$$P[A_i \cap A_j] = \frac{1}{4}, \quad i \neq j \tag{1}$$

también $P[A_i]P[A_j] = \dfrac{1}{2} \times \dfrac{1}{2} = \dfrac{1}{4}, \quad i \neq j$ (2)

De (1) y (2) $P[A_i \cap A_j] = P[A_i]\,P[A_j], \quad i \neq j$ (3)

Por otro lado se tiene

$$P[A_1 \cap A_2 \cap A_3] = \frac{1}{8} \quad \text{y} \tag{4}$$

$$P[A_1]\,P[A_2]\,P[A_3] = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8} \tag{5}$$

De (4) y (5) $P[A_1 \cap A_2 \cap A_3] = P[A_1]\,P[A_2]\,P[A_3]$ (6)

De (3) y (6); los eventos A_1, A_2 y A_3 son mutuamente independientes.

EJEMPLO 7 Dado el espacio muestral Ω (generado por algún fenómeno aleatorio)

$\Omega = \{(1,1,1), (1,1,0), (1,0,1), (0,1,1), (0,0,0)\}$. Sea

$\omega = (x, y, z) \in \Omega$ tal que $P[\{\omega\}]$ está dada por la siguiente tabla

ω	(1 , 1 , 1)	(1 , 1 , 0)	(1 , 0 , 1)	(0 , 1 , 1)	(0 , 0 , 0)
$P[\{\omega\}]$	1/8	3/16	3/16	3/16	5/16

Considere los eventos asociados a las proposiciones que se indican :

A_1 : "La primera coordenada es 1".

A_2 : "La segunda coordenada es 1".

A_3 : "La tercera coordenada es 1".

(a) Calcular $P[(A_1 \cup A_2) \mid (A_2 \cap A_3)]$

(b) ¿Son A_1, A_2, A_3 eventos mutuamente independientes? justifique su respuesta.

SOLUCION (a) $A_1 = \{(1,1,1), (1,1,0), (1,0,1)\}$
$A_2 \quad \{(1,1,1), (1,1,0), (0,1,1)\}$ y
$A_3 = \{(1,1,1), (1,0,1), (0,1,1)\}$

Luego, $A_1 \cup A_2 = \{(1,1,1), (1,1,0), (1,0,1), (0,1,1)\}$ y

$A_2 \cap A_3 = \{(1,1,1), (0,1,1)\}$

Observe que $A_2 \cap A_3 \subset A_1 \cup A_2$, por lo tanto

$$P[(A_1 \cup A_2) \mid (A_2 \cap A_3)] = \frac{P[A_2 \cap A_3]}{P[A_2 \cap A_3]} = 1 .$$

(b) $$P[A_1] = \frac{1}{8} + \frac{3}{16} + \frac{3}{16} = \frac{8}{16} = \frac{1}{2} . \qquad (1)$$

$$P[A_2] = \frac{1}{8} + \frac{3}{16} + \frac{3}{16} = \frac{8}{16} = \frac{1}{2} . \qquad (2)$$

$$P[A_3] = \frac{1}{8} + \frac{3}{16} + \frac{3}{16} = \frac{8}{16} = \frac{1}{2} . \qquad (3)$$

De (1) y (2), $$P[A_1]\, P[A_2] = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} . \qquad (4)$$

Por otro lado $A_1 \cap A_2 = \{(1,1,1), (1,1,0)\}$, luego

$$P[A_1 \cap A_2] = \frac{1}{8} + \frac{3}{16} = \frac{5}{6} . \qquad (5)$$

De (4) y (5) se observa que

$$P[A_1 \cap A_2] \neq P[A_1]\, P[A_2]$$

lo cual es suficiente para decir que A_1, A_2 y A_3 no son mutuamente independiente.

EJEMPLO 8 En tres establos A , B y C hay una epidemia que afecta a los cascos y a la boca del ganado, la proporción de ganados afectados : $\frac{1}{6}$, $\frac{1}{4}$ y $\frac{1}{3}$ respectivamente. Se escoge aleatoriamente una cabeza de ganado de cada establo .

(a) ¿Cuál es la probabilidad que exactamente uno de estos esté afectado por la epidemia?

(b) Si exactamente uno de estos está afectado, ¿cuál es la probabilidad de que venga del establo A?

SOLUCION Definimos los eventos siguientes:

A: "El ganado que proviene del establo A está afectado".

B: "El ganado que proviene del establo B está afectado".

C: "El ganado que proviene del establo C está afectado".

E: "Exactamente uno de los tres está afectado por la epidemia".

$A\bar{B}\bar{C}$: "sólo el ganado que proviene del establo A está afectado".

$\bar{A}B\bar{C}$: "sólo el ganado que proviene del establo B está afectado".

$\bar{A}\bar{B}C$: "sólo el ganado que proviene del establo C está afectado".

Luego, $E = A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C$.

Los eventos $A\bar{B}\bar{C}$, $\bar{A}B\bar{C}$ y $\bar{A}\bar{B}C$ son mutuamente excluyentes. También los eventos A, B y C son mutuamente independientes, entonces también son los eventos $A, \bar{B}$ y $\bar{C}$; $\bar{A}$, B y $\bar{C}$; $\bar{A}$, $\bar{B}$ y C .

(a) $P[E] = P[A\bar{B}\bar{C}] + P[\bar{A}B\bar{C}] + P[\bar{A}\bar{B}C]$

$$= P[A]\,P[\bar{B}]\,P[\bar{C}] + P[\bar{A}]\,P[B]\,P[\bar{C}] + P[\bar{A}]\,P[\bar{B}]\,P[C]$$

$$= \frac{1}{6}\cdot\frac{3}{4}\cdot\frac{2}{3} + \frac{5}{6}\cdot\frac{1}{4}\cdot\frac{2}{3} + \frac{5}{6}\cdot\frac{3}{4}\cdot\frac{1}{3} = \frac{31}{72}$$

(b) Según el teorema de Bayes tenemos

$$P[A \mid E] = \frac{P[A \cap E]}{P[E]} = \frac{P[A \cap A\bar{B}\bar{C}]}{P[E]} = \frac{P[A\bar{B}\bar{C}]}{P[E]} = \frac{6/72}{31/72} = \frac{6}{31}$$

EJEMPLO 9 En el ejemplo 19 de 1.8. Supongamos que se seleccionan dos negocios al azar.

(a) ¿Cuál es la probabilidad que ambos negocios hayan alcanzado la meta considerada?

(b) Dado que ambos negocios alcanzaron la meta. ¿Cuál es la probabilidad de que el 80% de los negocios haya vendido doce millones de dólares?

SOLUCION Definimos los eventos siguientes:

A_1: "el primer negocio alcanzó la meta".

A_2: "el segundo negocio alcanzó la meta".

E : "ambos negocios alcanzaron la meta".

(a) Debemos calcular $P[E] = P[A_1A_2]$. Además A_1 y A_2 son eventos independientes. En el diagrama del árbol de probabilidad tenemos

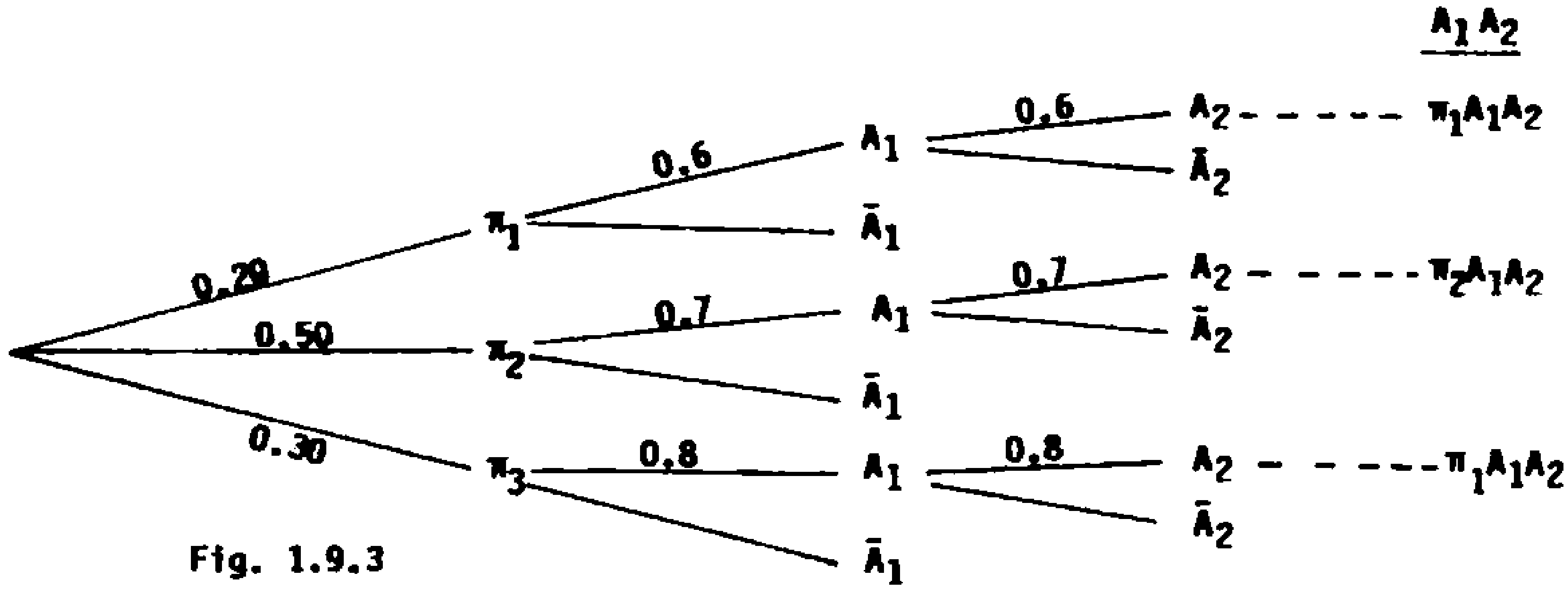


Fig. 1.9.3

$$P[E] = P[A_1A_2] = P[\pi_1A_1A_2] + P[\pi_2A_1A_2] + P[\pi_3A_1A_2]$$

$$= P[\pi_1 \; P[A_1 \mid \pi_1] \, P[A_2 \mid \pi_1A_1] + P[\pi_2] \, P[A_1 \mid \pi_2] \, P[A_2 \mid \pi_2A_1]$$

$$+ P[\pi_3] \, P[A_1 \mid \pi_3] \, P[A_2 \mid \pi_3A_1]$$

$$= (0.20)(0.60)^2 + (0.50)(0.70)^2 + (0.30)(0.80)^2$$

$$= 0.072 + 0.245 + 0.192 = 0.509$$

(b) Según el teorema de Bayes tenemos

$$P[\pi_3 \mid A_1A_2] = \frac{P[\pi_3] \, P[A_1|\pi_3] \, P[A_2|\pi_3A_1]}{P[A_1A_2]} = \frac{0.192}{0.509} = \frac{192}{509} .$$

EJEMPLO 10 Cuando una máquina que produce engranajes está trabajando apropiadamente, el 92% de las piezas satisfacen las especificaciones. Cuando la máquina no trabaja bien sólo el 60% de los engranajes satisfacen los requerimientos. La máquina está en buen estado el 90% del tiempo. Se seleccionan dos engranajes y ambos resultan de calidad aceptable. ¿Cuál es la probabilidad que la máquina no haya estado trabajando bien?

SOLUCION Sean los eventos siguientes:

A_1: "el primer engranaje satisface las especificaciones".
A_2: "el segundo engranaje satisface las especificaciones".
B : "la máquina está trabajando apropiadamente".

Debemos calcular $P[\bar{B} \mid A_1A_2]$. Además, A_1 y A_2 son eventos independientes.

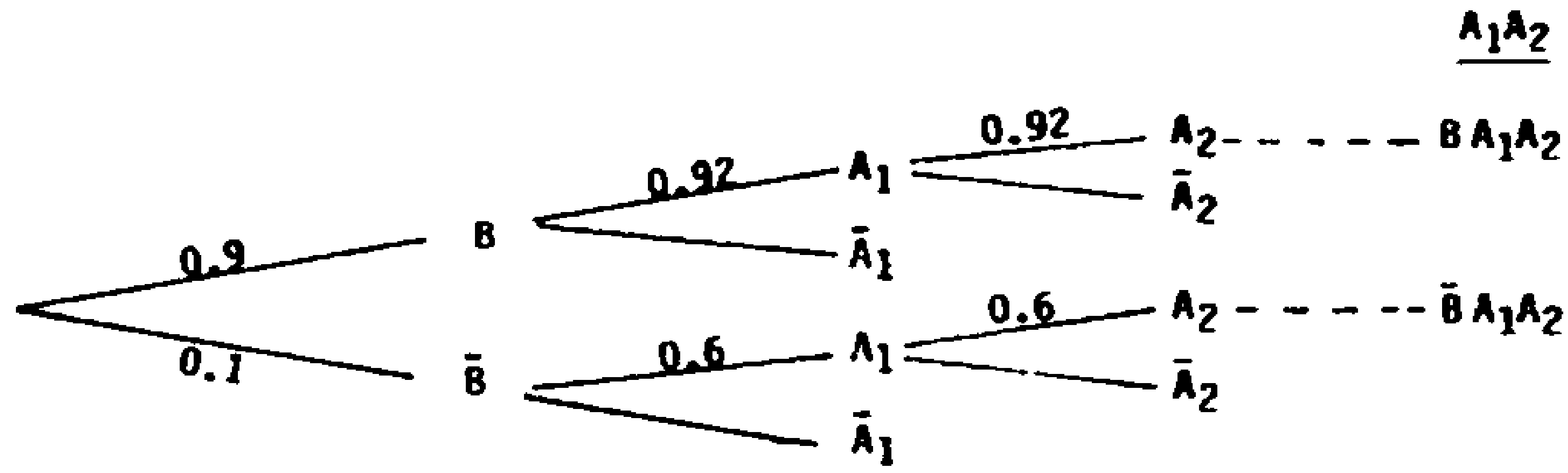


Fig. 1.9.4

Entonces, $P[A_1A_2] = P[BA_1A_2] + P[\bar{B}\ A_1A_2]$

$$= P[B]\ P[A_1|B]\ P[A_2|BA_1] + P[\bar{B}]\ P[A_1|\bar{B}]\ P[A_2|\bar{B}\ A_1]$$

$$= (0.90)(0.92)^2 + (0.10)(0.60)^2$$

$$= 0.76176 + 0.036 = 0.79776.$$

y según el teorema de Bayes tenemos

$$P[\bar{B} \mid A_1A_2] = \frac{P[\bar{B}\ A_1A_2]}{P[A_1A_2]} = \frac{(0.1)(0.60)^2}{0.79776} = \frac{3600}{79776} = 0.045\ .$$

EJEMPLO 11 Un gerente está a la espera de las llamadas telefónicas de sus clientes para efectuar un negocio, la probabilidad de que lo llame cualquiera de sus clientes es de 0.2. (Las llamadas de los clientes son eventos independientes). La probabilidad de efectuar el negocio es de 0.10 si recibe la llamada de un cliente; es de 0.3 si recibe la llamada de dos clientes y de 0.7 si recibe la llamada de tres clientes. Si no recibe llamada no realiza negocio. ¿Cuántas llamadas de clientes es más probable que haya recibido el gerente sabiendo que se realizó el negocio?

SOLUCION Definimos los eventos siguientes:

N : "se realizó el negocio".

C_i: "llama i-clientes , $i = 0,1,2,3$".

Así , C_1: "llama un cliente", C_2: "llama 2 clientes". etc. y

$$P[C_1] = 0.2\ ; \qquad P[C_2] = (0.2)^2\ ; \qquad P[C_3] = (0.2)^3$$

Tenemos que calcular $P[C_i \mid N]$, $i = 1,2,3$. La mayor de estas probabilidades será la probabilidad pedida.

El diagrama del árbol de probabilidades se muestra en la fig. 1.9.5

$$P[N] = P[C_1N] + P[C_2N] + P[C_3N]$$

$$= P[C_1]\, P[N|C_1] + P[C_2]\, P[N|C_2] + P[C_3]\, P[N|C_3]$$

$$= (0.2)(0.1) + (0.04)(0.3) + (0.008)(0.7)$$

$$= 0.0376 \,.$$

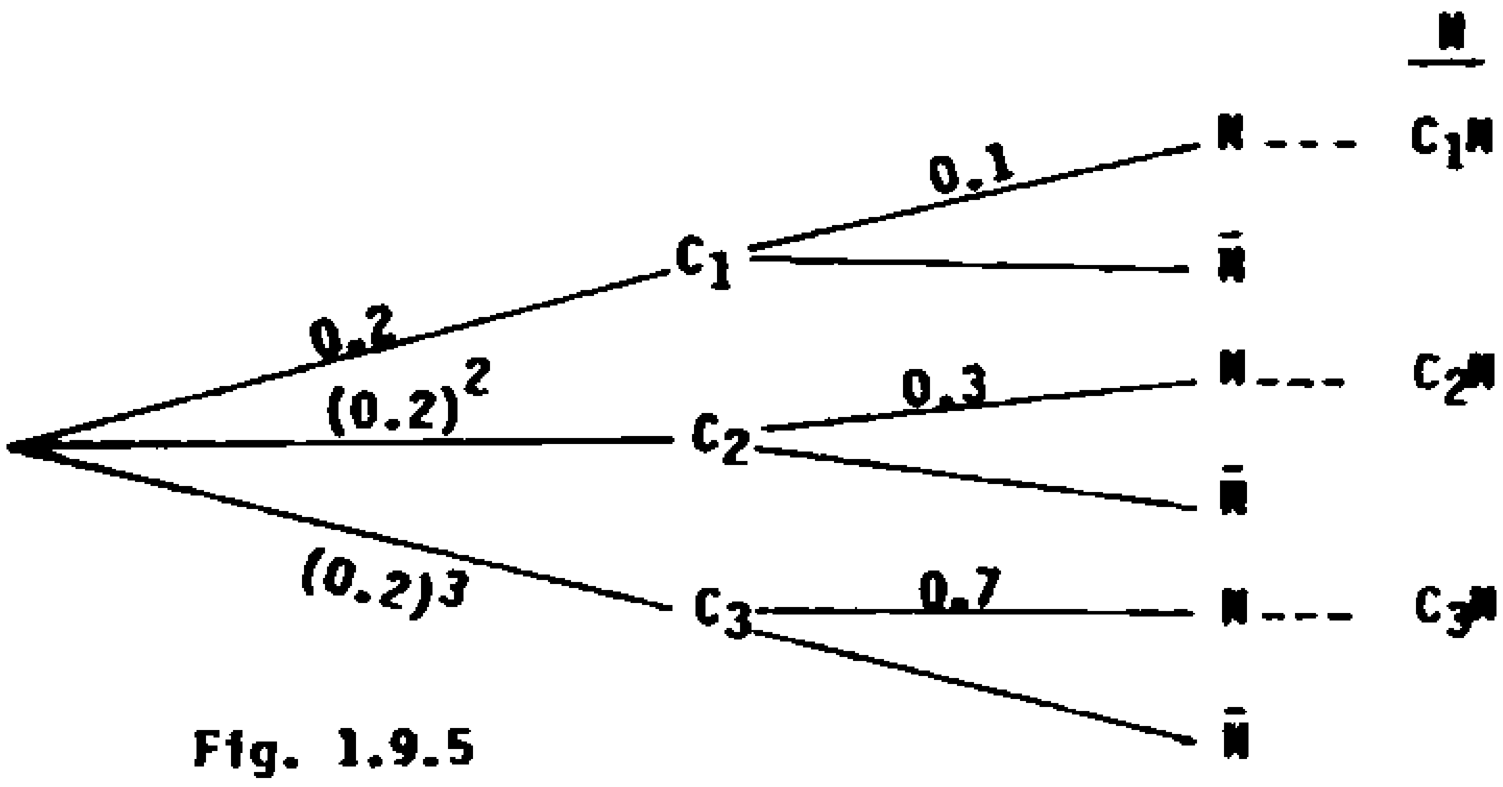


Fig. 1.9.5

Aplicando el teorema de Bayes en cada caso tenemos :

$$P[C_1|N] = \frac{P[C_1]\, P[N \mid C_1]}{P[N]} = \frac{(0.2)(0.1)}{0.0376} = \frac{200}{376} \,. \qquad (1)$$

$$P[C_2|N] = \frac{P[C_2]\, P[N \mid C_2]}{P[N]} = \frac{(0.04)(0.3)}{0.0376} = \frac{120}{376} \,. \qquad (2)$$

$$P[C_3|N] = \frac{P[C_3]\, P[N \mid C_3]}{P[N]} = \frac{(0.008)(0.7)}{0.0376} = \frac{56}{376} \,. \qquad (3)$$

de (1) , (2) y (3), es más probable que el gerente haya recibido la llamada de un cliente.

EJEMPLO 12 Se tiene 10 urnas : $A_1, A_2, \ldots, A_{10}$ con bolillas.

Sea B el evento: "bolilla blanca" y $P[B \mid A_i] = \frac{i}{10}$, $i = 1,2, \ldots 10$.

Se elige aleatoriamente una urna de la que se extraen 2 bolillas una a una, con reposición. Sabiendo que resultaron blancas, hallar la probabilidad de que la urna elegida fuese A_{10}.

SOLUCION 1 Sean los eventos B_i: "la i-ésima bola es blanca, $i = 1,2$".

2. Del diagrama del árbol de probabilidades y teniendo en cuenta que los eventos B_1 y B_2 son independientes ¿por qué?, tenemos

$$P[B_1B_2] = \sum_{i=1}^{10} P[A_i]\, P[B \mid A_i]^2$$

$$= \frac{1}{10}\left(\frac{1}{10}\right)^2 + \frac{1}{10}\left(\frac{2}{10}\right)^2 + \frac{1}{10}\left(\frac{3}{10}\right)^2 + \ldots + \frac{1}{10}\left(\frac{10}{10}\right)^2$$

$$= \frac{1}{10}\left[\frac{1}{10^2} + \frac{4}{10^2} + \frac{9}{10^2} + \frac{16}{10^2} + \frac{25}{10^2} + \frac{36}{10^2} + \frac{49}{10^2} + \frac{64}{10^2} + \frac{81}{10^2} + 1\right]$$

$$P[B_1B_2] = \frac{1}{10}\left[\frac{385}{100}\right],$$

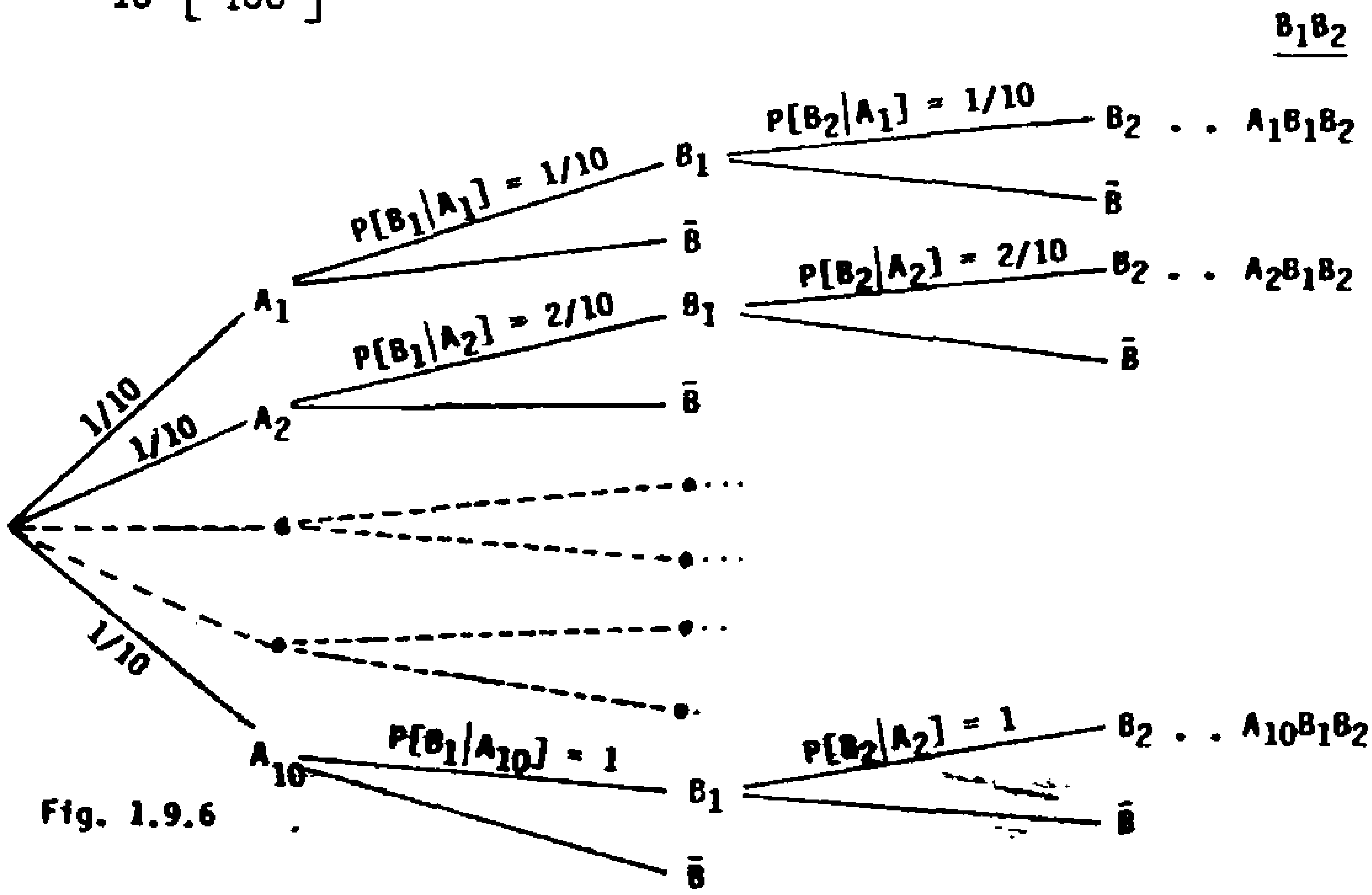


Fig. 1.9.6

y según el teorema de Bayes se tiene

$$P[A_{10} \mid B_1B_2] = \frac{P[A_{10}\ B_1B_2]}{P[B_1B_2]} = \frac{1/10 \times 1^2}{385/1000} = \frac{100}{385}.$$

EJEMPLO 13 El dado A tiene 4 caras rojas y 2 blancas, y el dado B tiene 2 caras rojas y 4 blancas. Se elige uno de estos dados al azar, se lanza n veces, y en todas ellas sale roja. En estas condiciones, la probabilidad de que se haya lanzado el dado A es 32/33. ¿Cuántas veces se lanzó el dado?.

SOLUCION Del enunciado , se tiene :

$$\text{Dado A, tiene} \begin{cases} \text{4 caras rojas} & (4r) \\ \text{2 caras blancas} & (2b) \end{cases}$$

$$\text{Dado B, tiene} \begin{cases} \text{2 caras rojas} & (2r) \\ \text{4 caras blancas} & (4b) \end{cases}$$

El diagrama del árbol de probabilidades que visualiza el problema, de la - fig. 1.9.7 .

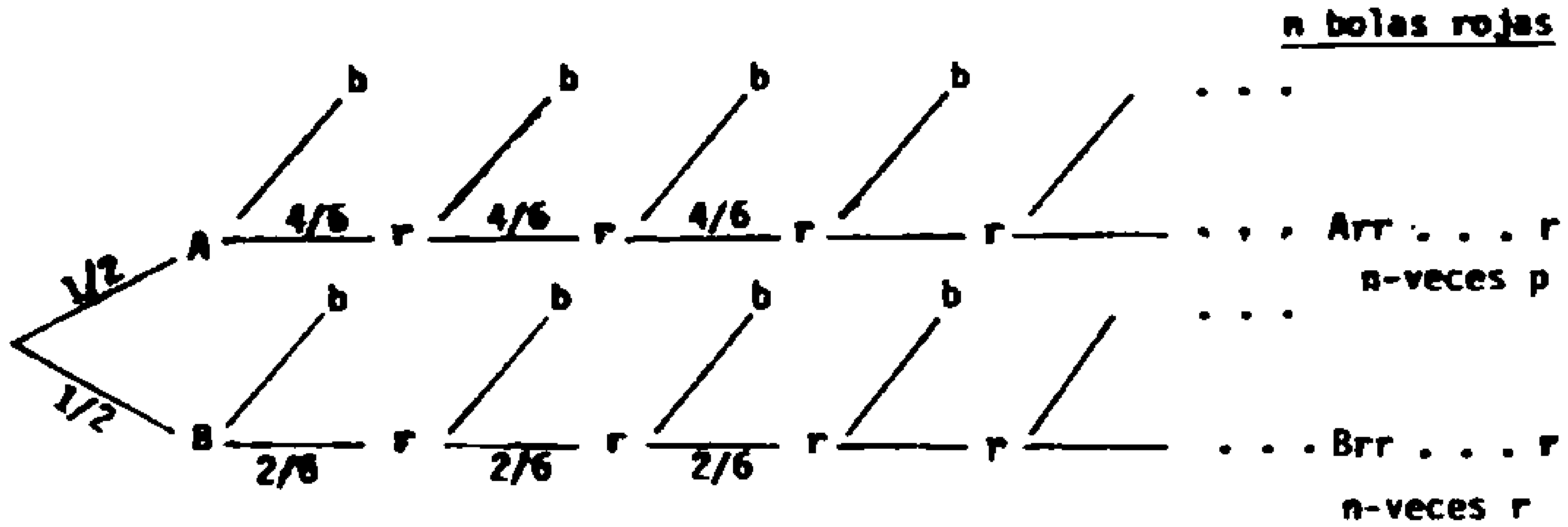


Fig. 1.9.7

Si R es el evento: "los n lanzamientos resultan rojas". Entonces,

$R = Arr \ldots r \cup Brr \ldots r$, y

$$P[R] = P[Arr \ldots r] + P[Brr \ldots r]$$

$$= \frac{1}{2}\left(\frac{4}{6}\right)^n + \frac{1}{2}\left(\frac{2}{6}\right)^n = \frac{1}{2}\left[\left(\frac{2}{3}\right)^n + \left(\frac{1}{3}\right)^n\right]$$

Luego ,

$$P[A \mid R] = \frac{\frac{1}{2}\left(\frac{2}{3}\right)^n}{\frac{1}{2}\left[\left(\frac{2}{3}\right)^n + \left(\frac{1}{3}\right)^n\right]} = \frac{32}{33}$$

$$\left(\frac{2}{3}\right)^n + \left(\frac{1}{3}\right)^n = \frac{33}{32}\left(\frac{2}{3}\right)^n$$

$$\left(\frac{1}{3}\right)^n = \frac{33}{32}\left(\frac{2}{3}\right)^n - \left(\frac{2}{3}\right)^n = \frac{1}{32}\left(\frac{2}{3}\right)^n$$

$$2^n = 32, \text{ de donde } n = 5.$$

Por lo tanto, se lanzó el dado cinco veces.

EJEMPLO 14 Un sólo misil de cierta variedad tiene una probabilidad de $\frac{1}{4}$ de derribar un bombardero a reacción, una probabilidad de $\frac{1}{4}$ de dañarlo y una - probabilidad de $\frac{1}{2}$ de errar el blanco. Igualmente, dos disparos que produzcan daños derribarán el avión. Si se lanzan cuatro de tales misiles, ¿cuál es la probabilidad de derribar un bombardero?

SOLUCION Sean los eventos :

D: "derribar un bombardero".

A: "un misíl derriba el bombardero".

B: "un misíl daña el bombardero".

E: "un misíl yerra el blanco".

Los elementos del espacio muestral son palabras de 4 letras tomadas del conjunto {A, B, E}. O sea ,

$$\Omega = \{AAAA, \overbrace{AAAB}^{P_4^{3,1}}, \ldots , \overbrace{BEEE}^{P_4^{3,1}} , EEEE\}$$

Es claro que el bombardero no es derribado si ocurre los sucesos, $\{BEEE , EEEE\} = \bar{D}$, en todos los demás casos el avión es derribado. Por lo tanto,

$$P[D] = 1 - P[\bar{D}] = 1 - [\, P_4^{3,1}\, P[BEEE] + P[EEEE]\,]$$

$$= 1 - [\binom{4}{3} (\tfrac{1}{4})(\tfrac{1}{2})^3 + (\tfrac{1}{2})^4]$$

$$= 1 - [\, (\tfrac{1}{2})^3 + (\tfrac{1}{2})^4\,] = 1 - \frac{3}{16} = \frac{13}{16} = 0.8125 .$$

CONSECUENCIA 2 Si A, B y C son eventos cualesquiera en Ω independientes, entoces el teorema 1.6.5 se escribe

$$P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A]\,P[B] - P[B]\,P[C] - P[A]\,P[C] + P[A]\,P[B]\,P[C]$$

TEOREMA 1.9.3 Si $A_1 , A_2 , \ldots , A_n$ son n eventos en Ω independientes, entonces

$$P[\bigcup_{i=1}^{n} A_i] = 1 - \prod_{i=1}^{n} P[\bar{A}_i]$$

$$= 1 - [1 - P[A_1]]\,[1 - P[A_2]] \ldots [1 - P[A_n]].$$

EJEMPLO 15 Durante el primer año de uso un amplificador de radio puede requerir tres tipos de reparaciones y las probabilidades correspondientes son: 0.05 , 0.04 y 0.12. ¿Cuál es la probabilidad que un amplificador seleccionado al azar requiera reparacion durante su primer año de uso? Cada tipo de reparación es independiente de los otros dos.

SOLUCION 1 Definimos los siguientes eventos :

E_i: "el amplificador seleccionado requiere reparación del tipo

i (i = 1,2,3)".

E : "el amplificador seleccionado requiere reparación".

2. El evento E se escribe $E = \bigcup_{i=1}^{3} E_i$

PRIMERA FORMA Usando la consecuencia 2

$$P[\bigcup_{i=1}^{3} E_i] = P[E_1] + P[E_2] + P[E_3] - P[E_1]\,P[E_2] - P[E_2]\,P[E_3] - P[E_1]\,P[E_3] + P[E_1]\,P[E_2]\,P[E_3]$$

$$= 0.05 + 0.04 + 0.02 - (0.05)(0.04) - (0.04)(0.02) - (0.05)(0.02) + (0.05)(0.04)(0.02)$$

$$= 0.10624.$$

SEGUNDA FORMA Usando el teorema 3

$$P[\bigcup_{i=1}^{3} E_i] = 1 - [1 - P[E_1]]\,[1 - P[E_2]][1 - P[E_3]]$$

$$= 1 - (1 - 0.05)\,(1 - 0.04)\,(1 - 0.02)$$

$$= 0.10624\ .$$

EJEMPLO 16 Un cazador dispara 7 balas a un león enfurecido. Si la probabilidad de que una bala acierte es 0.6. ¿Cuál es la probabilidad que el cazador esté todavía vivo?

SOLUCION 1 Definimos los eventos siguientes:

V : "el cazador esta vivo".

E_i: "en el i-ésimo tiro acierte , i = 1,2, ..., 7".

2.Los eventos E_i , i = 1,2,..., 7, son mutuamente independientes, y

$$P[E_i] = 0.6,\quad P[\bar{E}_i] = 0.4 = 1 - P[E_i]\ ,\quad i = 1,2,\ldots,7\ .$$

3. El cazador estará vivo, si el león ha muerto, es decir, si al menos uno de los 7 disparos ha acertado en el blanco. Entonces, el evento V se escribe así

$$V = \bigcup_{i=1}^{7} E_i\ .$$

$$P[V] = P[\bigcup_{i=1}^{7} E_i] = 1 - [1 - P[E_1]]\,[1 - P[E_2]]\ldots[1 - P[E_7]],\text{teor. } 1.9.3$$

$$= 1 - (0.4)^7 = 1 - 0.0016384 = 0.9983616.$$

El problema podría ser abordado de la siguiente manera. El cazador está muerto si el león está vivo y esto sucede si los 7 disparos fueron errados Es decir,

$$\bar{V} = \bar{E}_1\bar{E}_2 \ldots \bar{E}_7 .$$

Luego, $P[V] = 1 - P[\bar{V}] = 1 - P[\bar{E}_1] \ldots P[\bar{E}_7]$

EJEMPLO 17 El gerente de la compañía ABC viaja en un avión de 6 motores para asistir a una reunión importante en Brasil. La probabilidad que un motor falle es 0.10 y cada uno funciona independientemente de los otros. Si se necesita al menos un motor en cada lado del avión. ¿Cuál es la probabilidad que el gerente esté ausente de la reunión a causa de un accidente de su avión?.

SOLUCION 1 Definimos los siguientes eventos :

M_i: "El motor i-ésimo funciona perfectamente, ($i = 1,2, \ldots, 6$)".
A : "El gerente esté ausente de la reunión a causa de un accidente".
$\bar{A}$: "El gerente no esté ausente de la reunión".

$$P[M_i] = 0.9 \quad , \quad i = 1,2, \ldots, 6.$$

2. $P[A] = 1 - P[\bar{A}]$

3. Supongamos que los motores M_1, M_2 y M_3 esten a un lado y los motores M_4, M_5 y M_6 al otro lado. Además los M_i son independientes.

4. El evento $\bar{A}$ es equivalente a la ocurrencia conjunta de los eventos,
E :"al menos uno de los motores M_i funcionan perfectamente, $i = 1,2,3$"
F :"al menos uno de los motores M_i funcionan perfectamente, $i = 4,5,6$".

o sea $E = \bigcup_{i=1}^{3} M_i$ y $F = \bigcup_{i=4}^{6} M_i$. Por lo tanto $\bar{A} = EF$.

5. Los eventos E y F son independientes. Entonces

$$P[\bar{A}] = P[E]\,P[F] = P[\bigcup_{i=1}^{3} M_i]\;P[\bigcup_{i=4}^{6} M_i]$$

$$= \{1 - [1 - P[M_1]][1 - P[M_2]][1 - P[M_3]]\}\{1 - [1 - P[M_3]][1 - P[M_5]]$$
$$[1 - P[M_6]]\}$$

$= [1 - (0.1)^3]^2 = [1 - 0.001]^2 = (0.999)^2 = 0.998001.$

Reemplazando en (2) este resultado obtenemos

$P[A] = 1 - 0.998001 = 0.001999$.

EJEMPLO 18 La probabilidad que falle un motor en un avión es 0.10. ¿Con cuántos motores debe estar equipado un avión para tener una seguridad de 0.999 de que el avión vuele? (supóngase que es suficiente que un motor funcione para que el avión se mantenga en vuelo).

SOLUCION 1 Definimos los siguientes eventos:

M_i: "el motor i funciona perfectamente, ($i = 1,2, \ldots , n$)"
A : " el avión se mantiene en vuelo"

2. Los eventos M_i son independientes, $i = 1,2, \ldots, n$; y $P[M_i] = 0.9$, $i = 1,2, \ldots , n$

3. El avión se mantiene en vuelo si al menos uno de los motores funciona. Es decir,

$$A = \bigcup_{i=1}^{n} M_i \quad . \text{ Luego}$$

$$0.999 = P[A] = P[\bigcup_{i=1}^{n} M_i] = 1 - [1 - P[M_1]][1 - P[M_2]] \ldots [1 - P[M_n]] \text{ teor. } 1.9.3$$

$$= 1 - (0.1)^n$$

de donde $(0.1)^n = 0.001$

4. Tomando logaritmo a ambos miembros de la expresión anterior

$$n \log (0.1) = \log (0.001)$$

$$n [- \log 10] = - \log 10^3$$

$$- n = - 3 \qquad \text{osea} \qquad n = 3.$$

El avión debe estar equipado con 3 motores.

EJEMPLO 19 El sistema de números binarios tiene un papel muy importante en la operación de los computadoras electrónicas. Este sistema implica el uso de dos dígitos únicamente, 0 y 1. Si la probabilidad que aparezca un dígito incorrecto es p, y los errores en los dígitos se presentan en forma independiente uno de otros, ¿cuál es la probabilidad de que un número de n-dígitos sea incorrecto?

SOLUCION 1 Definimos los siguientes eventos :

D_i: "el dígito i-aparesca incorrecto, ($i = 1,2,\ldots, n$)"

D : "el número de n-dígitos es incorrecto"

2. Los eventos D_i son independientes $i = 1,2, \ldots, n$; y $P[D_i] = p$
3. El número de n-dígitos es incorrecto si, al menos uno de los dígitos aparece incorrecto; es decir $D = \bigcup_{i=1}^{n} D_i$. Luego ,

$$P[D] = P[\bigcup_{i=1}^{n} D_i] = 1 - [1 - P[D_1]][1 - P[D_2]] \ldots [1 - P[D_n]]$$

$$= 1 - (1 - p)^n .$$

EJEMPLO 20 Un generador tiene 6 componentes disipadores de corriente eléctrica. La probabilidad que ocurra una avería que desconecte el primer disipador es 0.6; para el segundo, 0.2; y 0.3 para cada uno de los cuatro restantes. Determinar la probabilidad que el generador esté completamente desconectado, si:

(a) Todos los disipadores están conectados en serie.

(b) Los disipadores están conectados en serie-paralelo, como se observa en la fig. 1.9.8

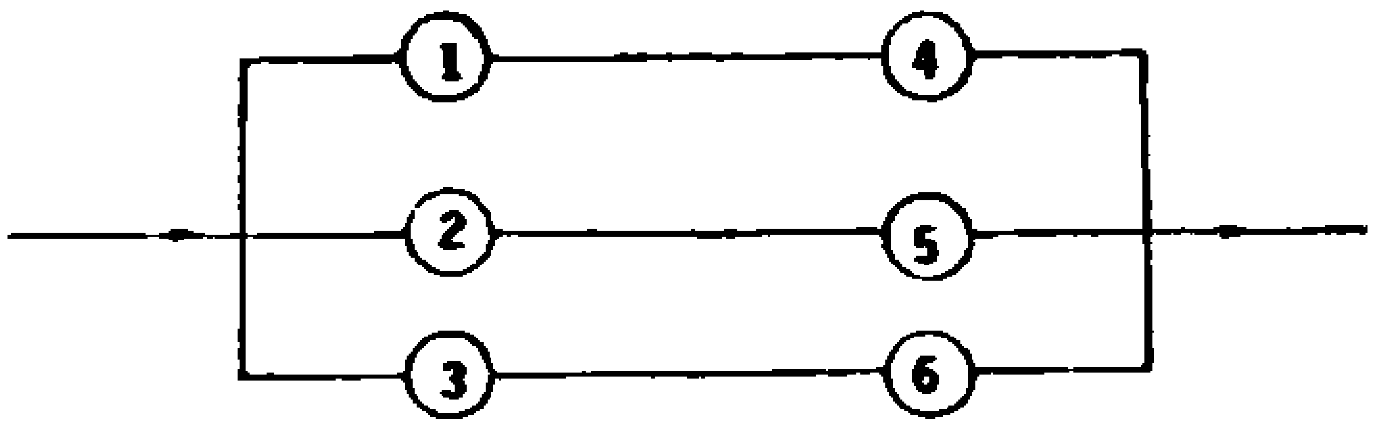


Fig. 1.9.8

SOLUCION Definimos los siguientes eventos :

D_i: "el disipador i (i = 1,2,3,4,5,6) está desconectado"

$\bar{D}_i$: "el disipador i (i = 1,2,3,4,5,6) está conectado:

D : "el generador está desconectado"

$\bar{D}$: "el generador está conectado" .

Sus respectivas probabilidades son :

$$P[D_1] = 0.6 , \quad P[D_2] = 0.2 , \quad P[D_i] = 0.3 , \quad i = 3,4,5,6$$

(a) $P[D] = 1 - P[\bar{D}]$

$P[\bar{D}] = P[\bar{D}_1\bar{D}_2\bar{D}_3\bar{D}_4\bar{D}_5\bar{D}_6]$

$= P[\bar{D}_1]\, P[\bar{D}_2]\, P[\bar{D}_3]\, P[\bar{D}_4]\, P[\bar{D}_5]\, P[\bar{D}_6]$ (suponiendo que los -

$\bar{D}_i$ son independientes, $i = 1,2,3,4,5,6$)

$= (0.4)(0.8)(0.7)^4$

$\approx 0.077.$

Luego, $P[D] = 1 - 0.077 = 0.923.$

(b) El generador está desconectado, si cada uno de las conecciones en paralelo están desconectados; es decir, si :

F_i: "la conección i en paralelo ($i = 1,2,3$) está desconectado".

$\bar{F}_i$: "la conección i en paralelo ($i = 1,2,3$) esta conectado"; entonces,

$$D = F_1F_2F_3 \text{ . Por lo tanto}$$

$$P[D] = P[F_1F_2F_3]$$

$$= P[F_1]\,P[F_2]\,P[F_3] \quad (\text{Los } F_i \text{ independientes})$$

pero, $P[F_1] = 1 - P[\bar{F}_1]$, donde $\bar{F}_1 = \bar{D}_1\bar{D}_2$ entonces,

$$P[F_1] = 1 - P[\bar{D}_1]\,P[\bar{D}_2]$$

$$= 1 - (0.3)(0.8) \tag{1}$$

$$P[F_2] = 1 - P[\bar{F}_2] \text{ , donde } \bar{F}_2 = \bar{D}_3\bar{D}_4 \text{ ;}$$

entonces, $P[F_2] = 1 - P[\bar{D}_3]\,P[\bar{D}_4]$

$$= 1 - (0.7)^2 \tag{2}$$

$$P[F_3] = 1 - P[\bar{F}_3] \text{ , donde } \bar{F}_3 = \bar{D}_5\bar{D}_6 \text{ ;}$$

entonces, $P[F_3] = 1 - P[\bar{D}_5]\,P[\bar{D}_6]$

$$= 1 - (0.7)^2 \tag{3}$$

De (1) , (2) y (3) obtenemos,

$$P[D] = [1 - (0.3)(0.8)] \times [1 - (0.7)^2]^2$$

$$\approx 0.1977.$$

EJEMPLO 21 En la figura 1.9.9, suponga que la probabilidad que cada relé esté cerrado es p y cada relé se abre o se cierra independientemente de cualquier otro. Encontrar la probabilidad que la corriente pase de R a S.

SOLUCION 1 Sean los eventos :

E : "la corriente pasa por I"

F : "la corriente pasa por II".

G : "la corriente pasa por III".

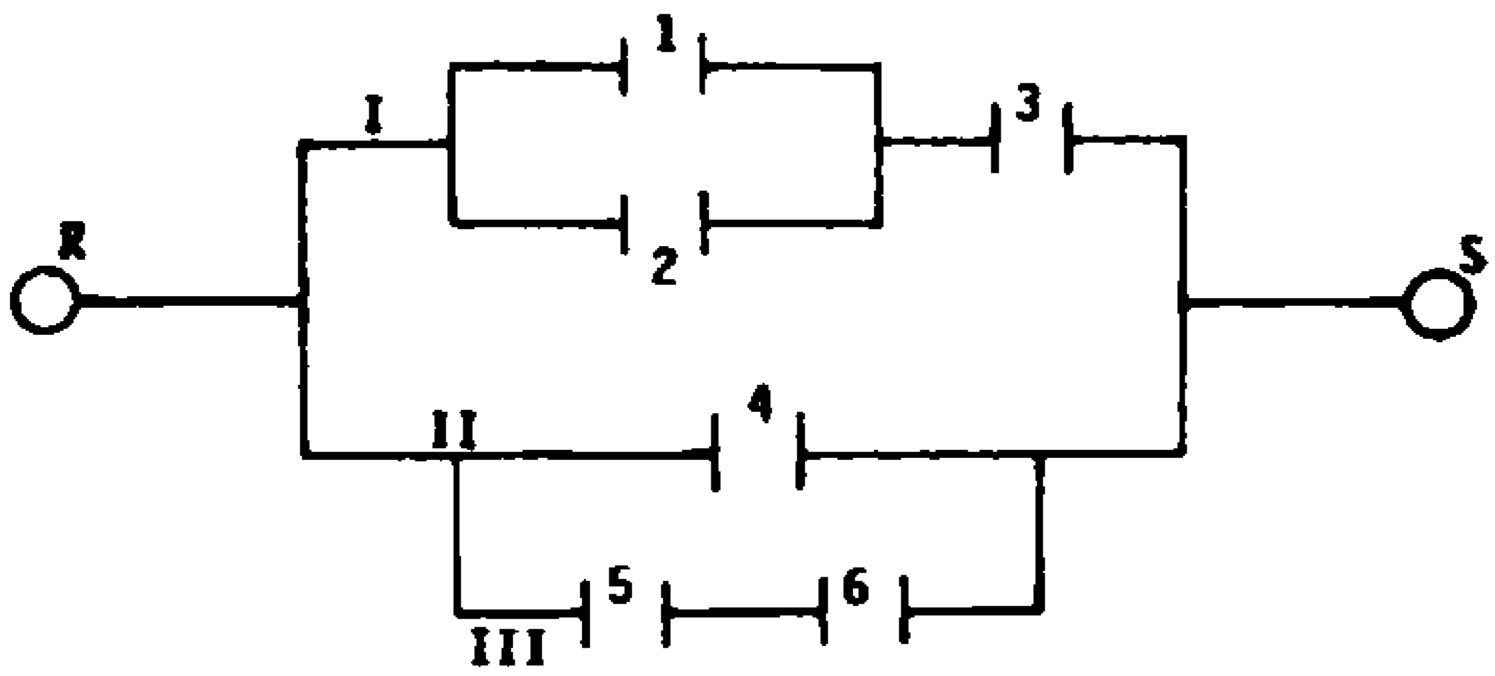


Fig. 1.9.9

2. Si E_1, E_2, E_3 y G_5, G_6 indican que la corriente pasa por 1,2,3,5 y 6 - respectivamente, se tiene que

$$E = (E_1 \cup E_2) \cap E_3 \qquad y$$

$$G = G_5 \cap G_6 .$$

3. Cálculo de las probabilidades de los eventos E , G y F respectivamente

$$P[E] = P[E_1 \cup E_2]P[E_3] = \left[P[E_1] + P[E_2] - P[E_1E_2]\right]P[E_3]$$

$$= [P[E_1] + P[E_2] - P[E_1]P[E_2]]P[E_3]$$

$$= (p + p - p^2)p = 2p^2 - p^3 .$$

$$P[G] = P[G_5 \cap G_6] = P[G_5]P[G_6] = p^2 .$$

$$P[F] = p .$$

4. La corriente pasará de R a S si pasa por I, II ó III ; es decir, si ocu- rre E ó F ó G ; entonces

$$P[E \cup F \cup G] = P[E] + P[F] + P[G] - P[E \cap F] - P[E \cap G] - P[F \cap G] + $$
$$+ P[E \cap F \cap G]$$

$$= 2p^2 - p^3 + p + p^2 - (2p^2 - p^3)p - (2p^2 - p^3)p^2 - pp^2$$
$$+ (2p^2 - p^3)pp^2$$

$$= p + 3p^2 - 4p^3 - p^4 + 3p^5 - p^6 .$$

CONFIABILIDAD La confiabilidad de un sistema " C_s " se define como la probabilidad que el sistema funciona satisfactoriamente para un intervalo de - tiempo especificado en las mismas condiciones.

(a) Para un sistema en serie, tal como se muestra en la fig. 1.9.10. consideremos los siguientes eventos

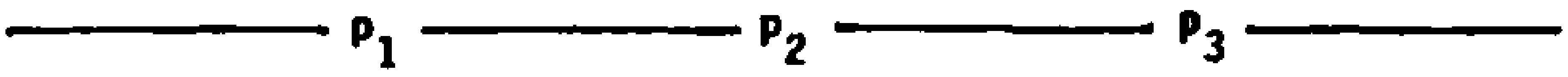


Fig. 1.9.10. Sistema en serie

E : "El sistema funciona satisfactoriamente".

E_i: "la componente p_i del sistema funciona satisfactoriamente,

$i = 1,2,3$".

Entonces el evento E es la intersección de los eventos E_1, E_2 y E_3. - Osea la confiabilidad del sistema en serie de la fig. 1.9.10 es

$$C_s = P[E] = P[E_1 \cap E_2 \cap E_3],$$

Asumimos que el funcionamiento satisfactorio o insatisfactorio de cada componente es independiente del funcionamiento de las otras componentes es decir,

$$C_s = P[E] = P[E_1]\, P[E_2]\, P[E_3].$$

Si C_i representa la confiabilidad de la respectiva componente p_i, - $i = 1,2,3$. Entonces

$$C_s = P[E_1]\, P[E_2]\, P[E_3] = C_1 C_2 C_3 .$$

El lector puede obtener la confiabilidad de un sistema con n componentes

(b) En un sistema en paralelo, tal como se muestra en la fig. 1.9.11, la confiabilidad C_s puede calcularse de dos maneras.

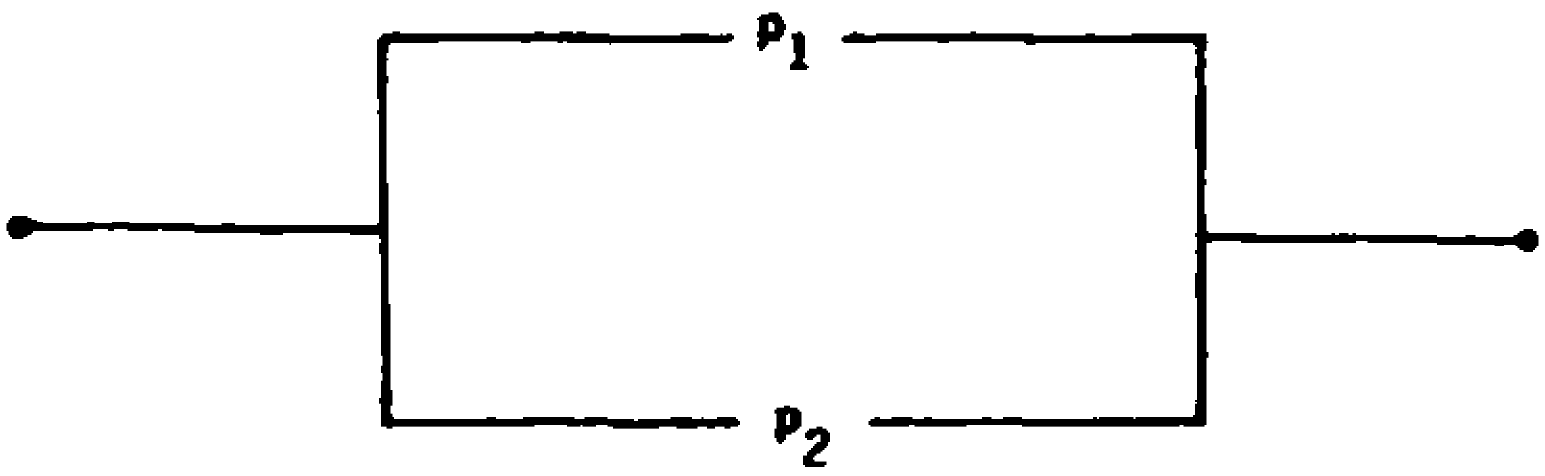


Fig. 1.9.11. Sistema en paralelo

PRIMER METODO Definimos los siguientes eventos :

E : "el sistema funciona satisfactoriamente".

E_i: "la componente p_i del sistema funciona satisfactoriamente, $i = 1,2$"

El sistema funciona correctamente, si al menos una de las componentes funciona correctamente, entonces el evento E, es la unión de los eventos E_1 y E_2. Es decir, la confiabilidad C_s del sistema en paralelo de la fig. 1.10.11 es

$$C_s = P[E] = P[E_1 \cup E_2]$$
$$= P[E_1] + P[E_2] - P[E_1 \cap E_2] \text{ , por teorema 1.6.4}$$

Asumiendo que el funcionamiento correcto o incorrecto de cada componente es independiente del funcionamiento del otro, tenemos

$$C_s = P[E_1] + P[E_2] - P[E_1]P[E_2] .$$

SEGUNDO METODO Un segundo método para calcular C_s es, usando el teorema 1.9.2

$$C_s = P[E_1 \cup E_2] = 1 - [1 - P[E_1]][1 - P[E_2]] = 1 - P[\bar{E}_1] P[\bar{E}_2]$$

también se llega a la mismo considerando:

1. Los siguientes eventos :

   F : "El sistema no funciona satisfactoriamente" ;

   F_i: "la componente p_i del sistema no funciona satisfactoriamente, $i = 1,2$"

   $\bar{F}$: "El sistema funciona satisfactoriamente"

2. $C_s = P[\bar{F}] = 1 - P[F]$ por teorema 1.6.2

3. El sistema falla, si los dos componentes fallan. Es decir $F = F_1F_2$. Esto es

$$P[F] = P[F_1]P[F_2]$$
$$= [1 - P[\bar{F}_1]][1 - P[\bar{F}_2]] \qquad \text{por teorema 1.6.2}$$

4. Sustituyendo el resultado anterior en el paso (2) se obtiene

$$C_s = 1 - [1 - P[\bar{F}_1]][1 - P[\bar{F}_2]]$$

El segundo método se generaliza usando el teorema 1.9.3 para un sistema en paralelo de n componentes

$$C_s = 1 - [1 - P[E_1]][1 - P[E_2]] \ldots [1 - P[E_n]].$$
$$= 1 - [1 - P[\bar{F}_1]][1 - P[\bar{F}_2]] \ldots [1 - P[\bar{F}_n]]$$

EJEMPLO 22 Una máquina presenta un sistema de dos componentes A Y B dispuestos en serie, las confiabilidades de que las componentes trabajan correctamente son 0.70 y 0.80, respectivamente. Suponga que A y B funcionan independientemente, y ambas componentes del sistema deben funcionar correctamente - para que la máquina lo haga. Para incrementar la confiabilidad del sistema - se emplea una componente similar, en paralelo, a fin de formar el sistema S que se observa en la fig. 1.9.12. La máquina funcionará siempre que, por lo menos uno de las componentes (sub-sistemas) trabajen correctamente, Calcular la confiabilidad del sistema S.

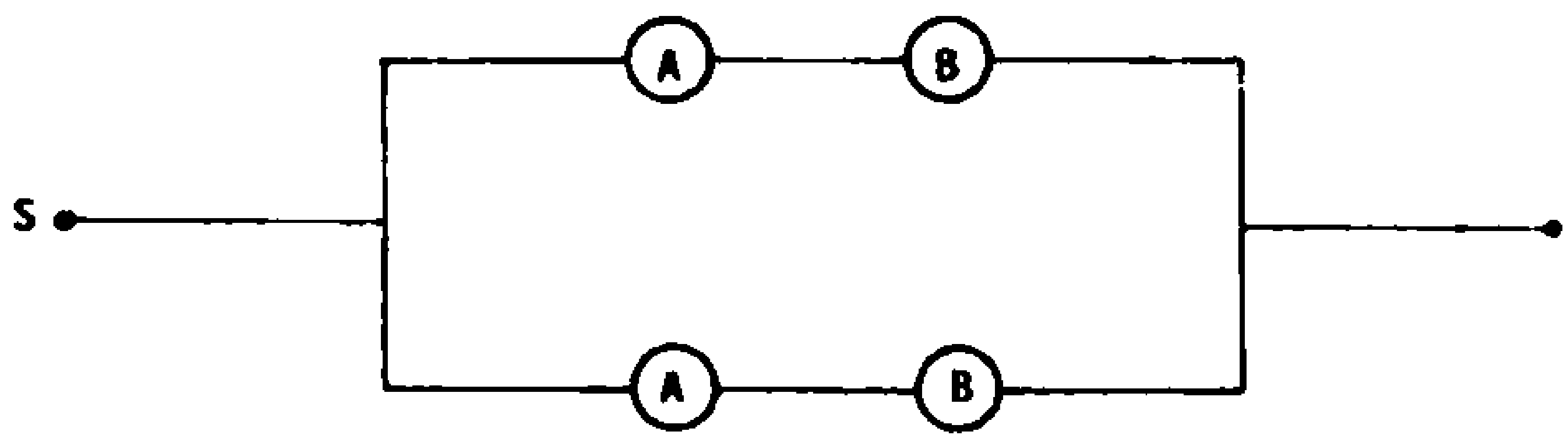


Fig. 1.9.12

SOLUCION PRIMER METODO Definimos los siguientes eventos:

E_1: "la componente A funciona correctamente",
E_2: "la componente B funciona correctamente",
E : "el sistema S funciona correctamente",
C_S: "confiabilidad del sistema".

El evento E se escribe, $E = E_1E_2 \cup E_1E_2$: entonces,

$$
\begin{aligned}
C_S = P[E] &= P[E_1E_2 \cup E_1E_2] \\
&= P[E_1E_2] + P[E_1E_2] - P[E_1E_2E_1E_2] \quad \text{, teorema 1.6.4} \\
&= 2P[E_1E_2] - [P[E_1E_2]]^2 \\
&= P[E_1E_2][2 - P[E_1E_2]] \\
&= P[E_1]\,P[E_2][2 - P[E_1]\,P[E_2]] \quad \text{, } E_1 \text{ y } E_2 \text{ independientes} \\
&= (0.70)(0.80)\,[2 - (0.70)(0.80)] \\
&= 0.8064 .
\end{aligned}
$$

SEGUNDO METODO Definimos los siguientes eventos

F : "el sistema S no funciona correctamente" , entonces,
$\bar{F}$: "el sistema S funciona correctamente" ; luego,

$$C_S = P[\bar{F}] = 1 - P[F] .$$

El sistema S falla, si las dos componentes en paralelo fallan; es decir, si

F_1: "falla la primera componente AB en serie" ,

F_2: "falla la segunda componente AB en serie" ;

entonces, $F = F_1F_2$. Luego

$$\begin{aligned} P[F] &= P[F_1]\,P[F_2] \\ &= [1 - P[\bar{F}_1]] \times [1 - P[\bar{F}_2]] \\ &= [1 - (0.70) \times (0.80)] \times [1 - (0.70) \times (0.80)] \\ &= [1 - (0.70) \times (0.80)]^2 \\ &= 0.1936 . \end{aligned}$$

Por lo tanto,

$$C_S = P[\bar{F}] = 1 - 0.1936 = 0.8064 .$$

1.9.1 EXPERIMENTOS INDEPENDIENTES

Para terminar esta sección introduciremos el concepto de experimentos independientes, para hacer más plausible la aceptación de la solución intuitiva que se ha dado a algunos ejemplos anteriores de esta sección.

DEFINICION 1.9.4 Sea ε un experimento que consiste de una secuencia de n ensayos, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. *Los ensayos son independientes* si el resultado de cualquier ensayo no afecta la probabilidad de los resultados de los otros ensayos . Además si los A_i , $i = 1, \ldots, n$ son eventos cualesquiera de Ω_i , $i = 1,2, \ldots, n$, respectivamente, donde los Ω_i son los espacios muestrales asociados a los experimentos ε_i respectivos. La probabilidad de la ocurrencia de los n eventos es

$$P[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]\,P[A_2] \ldots P[A_n]$$

Los ensayos que no son independientes se dice que son dependientes.

Consideremos el experimento de lanzar una moneda y un dado. Sea A el evento: "obtener una cara y un seis".

$$\Omega = \Omega_1 \times \Omega_2 \text{ , donde : } \quad \Omega_1 = \{C , S\}, \quad \Omega_2 = \{1,2,3,4,5,6\}$$

$$A = A_1 \cap A_2 = A_1A_2, \quad \text{donde} \quad A_1 = \{C\}, \quad A_2 = \{6\}$$

Entonces,

$$P[A] = P[A_1]\,P[A_2] = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}.$$

EJEMPLO 23 Se lanza una moneda hasta que aparezca la primera cara. ¿Cuál es la probabilidad que la primera cara aparezca

(a) en el segundo lanzamiento?
(b) en el tercer lanzamiento?

SOLUCION **PRIMERA FORMA** El espacio muestral es

$$\Omega = \{C, SC, SSC, \ldots\}$$

(a) $A = \{SC\}, \quad A_1 = \{S\}, \quad A_2 = \{C\},$

$$P[A] = P[A_1]\,P[A_2] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

(b) $$P[SSC] = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

pues el resultado del segundo lanzamiento no es afectado por lo que ocurrió en el primer lanzamiento y el resultado del tercer lanzamiento no es afectado por lo que ocurrió en el segundo.

SEGUNDA FORMA

(a) El evento $A = \{SC\}$ se considera como un resultado del lanzamiento de dos monedas y $\Omega = \Omega_1 \times \Omega_2 = \{CC, CS, SC, SS\}$.

Luego, $P[A] = \frac{1}{4}$.

(b) También podemos considerar el evento SSC como un resultado de lanzamiento de 3 monedas. Entonces el espacio muestral es

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 = \{CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS\}.$$

Luego, $P[SSC] = \frac{1}{8}$.

EJEMPLO 24 Un dado sesgado es tal que la probabilidad de obtener cada número par es 2/9 y la probabilidad de obtener cada número impar es 1/9. Suponga que el dado se lanza 3 veces. Si Ud. gana cada vez que aparece un 2 ó un 4. ¿Cuál es la probabilidad que Ud. gane

(a) Todas las veces?.

(b) exactamente 2 veces?.

SOLUCION Sean los siguientes eventos:

A: "ganar todas las veces".

B: "ganar exactamente 2 veces".

C: "aparece un 2 ó un 4".

F: "aparece un número diferente de 2 y 4".

y $$P[C] = P[\{2,4\}] = P[\{2\}] + P[\{4\}] = \frac{2}{9} + \frac{2}{9} = \frac{4}{9}$$

(a) El evento $A = CCC$, por lo tanto

$$P[A] = P[C]\,P[C]\,P[C] = \left(\frac{4}{9}\right)^3$$, ya que son independientes.

(b) El evento, $B = CCF \cup CFC \cup FCC$, por lo tanto

$$P[B] = 3\left(\frac{4}{9}\right)^2\left(\frac{5}{9}\right)$$

pues,

$$P[F] = P[\{1,3,5,6\}] = P[\{1\}] + P[\{3\}] + P[\{5\}] + P[\{6\}] = \frac{5}{9}.$$

PROBLEMAS 1.9

1. Si $P[A] = \frac{1}{6}$, $P[AB] = \frac{1}{18}$, $P[B] = \frac{1}{3}$

   ¿Son A y B independientes?.

2. Una urna contiene 4 bolas blancas y 5 negras. Se extraen sucesivamente y sin reposición dos bolas, sean los eventos.

   A: "la primera bola extraída es negra".

   B: "la segunda bola extraída es blanca".

   ¿Son los eventos A y B independientes?

3. De una baraja ordinaria de 52 cartas se extraen sucesivamente dos cartas, restituyendo la primera antes de extraer la segunda. Sea A el evento (suceso) "la primera carta es una pica", B el evento "la segunda carta es as o rey" y C el evento "la primera carta es as o rey. De los tres pares de eventos: A y B; A y C; B y C, determine cuales (si los hay) son independientes.

4. Si A y B son independientes, $P[A] = 1/3$ y $P[B] = 1/4$.
   Hallar, $P[A \cup B]$.

5. Si A y B son independientes, y $P[A] = P[B] = 1/2$. Calcular $P[A\bar{B} \cup \bar{A}B]$.

6. Dado $P[A] = 0.5$ y $P[A \cup B] = 0.7$. Hallar $P[B]$, si A y B son independientes.

7. Si A y B son independientes, y $P[A] = P[B \mid A] = 1/2$.
   Hallar $P[A \cup B]$.

8. Si A y B son eventos independientes con $P[A] = 0.2$, $P[B] = 0.3$. ¿Cuál es la probabilidad que
   (a) al menos uno ocurra ; (b) exactamente uno ocurra ;
   (c) ninguno ocurra ; (d) ambos ocurran?

9. Sean A y B dos eventos independientes, se sabe que la probabilidad de que ocurra al menos uno de dichos eventos es 0.6 y que la probabilidad de que ocurra A es 0.4. Calcular la probabilidad que ocurra B.

10. Si un conejo es inyectado con una droga A la probabilidad que muera dentro de las 24 horas siguientes es de 0.63 y si es inyectado con una droga B dicha probabilidad es de 0.45. ¿Cuál es la probabilidad que un conejo sobreviva más de 24 horas después de haber sido inyectado simultaneamente con las drogas A y B, si se supone que la acción de las mismas son independientes?

11. Cierto insecticida mata en la primera aplicación al 90% de los mosquitos pero desarrolla cierta resistencia entre los que sobreviven, de manera que el porcentaje que muere en una aplicación posterior del insecticida es la tercera parte del porcentaje que muere en la aplicación inmediatamente anterior. ¿Cuál es la probabilidad que un mosquito sobreviva:
    (a) tres aplicaciones de insecticida?
    (b) tres aplicaciones de insecticida, sabiendo que sobrevivió las dos primeras? .

12. Las probabilidades que tres tiradores den en el blanco son, respectivamente, iguales a 4/5, 3/4 y 2/3. Si en un disparo simultáneo por los tres tiradores, exactamente dos dan en el blanco; hallar la probabilidad de que el tercer tirador haya fallado.

13. Obtener la probabilidad que en 6 lanzamientos independientes de un da-

do correcto, aparezca el número 3 almenos una vez.

14. La probabilidad que un misil disparado contra una blanco no sea interceptado es 2/3. Dado que el misil no ha sido interceptado su probabilidad de dar en el blanco es 3/4. Si se dispara 4 misiles, independientemente, ¿cuál es la probabilidad que
(a) todos den en el blanco?
(b) al menos uno de en el blanco?
¿Cuántos misiles deben dispararse para que,
(c) al menos uno, no sea interceptado con probabilidad 0.95?

15. Cuatro hombres lanzan cada uno un dado. ¿Cuál es la probabilidad que:
(a) cada uno obtenga un 4 ;
(b) cada uno obtenga un número par de puntos ;
(c) todos obtengan el mismo número ?

16. Cada uno de n individuos lanzan una moneda al aire. Exprese en términos de n, la probabilidad que:
(a) ninguno obtenga cara;
(b) todos obtengan cara ;
(c) al menos uno obtenga una cara.

17. Ocho boletos numeradas, 111, 121, 122, 122, 211, 212, 212,221 estan colocados en una bolsa, revueltas. Se va a escoger uno al azar. Se definen los siguientes eventos:
A: "el primer dígito del boleto escogido es 1"
B: "el segundo dígito en el boleto escogido es 1"
C: "el tercer dígito del boleto escogido es 1"
(a) ¿Son los eventos A, B y C mutuamente independientes?
(b) Calcular $P[A \cup B \mid B \cap C]$

18. Suponga que un misil tiene la probabilidad 1/2 de destruir su blanco y la probabilidad de 1/2 de errarlo. Suponiendo que los lanzamientos de los misiles forman pruebas independientes, determínese el número de misiles que deben lanzarse para conseguir que la probabilidad de destruir el blanco sea por lo menos 0.99.

19. ¿Cuántas personas deben escoger una carta, cada una de diferente baraja para tener una probabilidad mínima de 0.9 de que por lo menos se escoja un as?

20. Se dispara cada uno de los fusiles A, B y C, la probabilidad de dar en el blanco es 0.15, 0.25 y 0.35, respectivamente. Calcular la probabilidad.
    (a) De que al menos uno de los tres dé en el blanco
    (b) de que acierte uno solo.

21. En un club el 60% de las personas fuman; 10 personas son seleccionadas sucesivamente al azar con reemplazamiento; ¿Cuál es la probabilidad del evento: "de las 10 personas seleccionadas 3 fuman"? ¿Y la del evento: "de las 10 personas seleccionadas 3 fuman", ¿y la del evento: "de las 10 personas seleccionadas por lo menos 3 fuman"?.

22. Un teatro tiene sólo un proyector. La bombilla del proyector funciona; la probabilidad que se queme antes de terminar la película es 0.40. De las 20 lámparas de reserva, una tiene un defecto no-visible. De las restantes lámparas de reserva, la probabilidad que se quemen es 0.20 antes de terminar la película.
    (a) ¿Cuál es la probabilidad que se queme la lámpara en funcionamiento y seleccionado al azar un extra, se escoja la lámpara defectuosa?
    (b) ¿Cuál es la probabilidad que se queme la lámpara es funcionamiento y seleccionada una perfecta para reemplazarla, se queme a su vez, antes de terminar la película?.

23. La probabilidad de que un hombre viva 10 años es 1/4, y la probabilidad de que su esposa viva 10 años es 1/3. Suponiendo que estos eventos son independientes, hallar la probabilidad que:
    (a) Por lo menos uno de ellos esté vivo entre los 10 años,
    (b) ninguno esté vivo entre los 10 años
    (c) solamente la esposa esté viva entre los 10 años
    (d) solamente el esposo esté vivo entre los 10 años

24. Una persona que tiene 35 años de edad, padece de cierta enfermedad; consultados los médicos las opiniones están en la relación 9 a 7 en contra de que la persona viva hasta los 40 años. Otra persona tiene 45 años y las opiniones están en la relación 3 a 2 en contra de que viva hasta los 50 años. Hallar la probabilidad que cuando menos una de estas personas viva 5 años más.

25. En una urna hay 15 bolas, de las cuales 5 son blancas. Se extraen al azar cinco bolas con reemplazamiento. ¿Cuál es la probabilidad que se selec

cionen x bolas blancas con $x = 0,1,2,3,4,5$.

26. Una pieza de equipo electrónico tiene 3 partes esenciales. Anteriormente, la parte A ha fallado el 20% del tiempo; la parte B, 40% del tiempo y - parte C, 30% del tiempo. La parte A opera independientemente de las partes B y C. Las partes B y C están interconectadas de tal manera que la falla de cualquiera afecta a la otra, por eso, cuando falla la parte C dos de - cada 3 veces puede fallar también la parte B.
Suponiendo que por lo menos dos de las 3 partes deben operar para permitir el funcionamiento del equipo. ¿Cuál es la probabilidad que el equipo - funcione?.

27. Un sistema consiste de 4 componentes: A, B, C_1, C_2. La probabilidad de falla es 0.01 para A, 0.02 para B, 0.10 para C_1 y 0.10 para C_2. Si para el funcionamiento del sistema son necesarios los componentes A y B y al menos uno de los C, ¿Cuál es la probabilidad de que el sistema funcione?

28. La probabilidad de que un cazador dé en el blanco con un tiro es 0.40.
   (a) ¿Cuál es la probabilidad que falle 4 tiros consecutivos?
   (b) ¿Cuál es la probabilidad que dé en el blanco por lo menos una vez en 4 tiros consecutivos?
   (c) ¿Cuántos tiros debe disparar para tener una seguridad aproximadamente de 0.96 de dar en el blanco por lo menos una vez?

29. Considere tres urnas; la urna I contiene una bola blanca y dos negras, la urna II contiene tres bolas blancas y dos negras y la urna III contiene dos bolas blancas y tres negras. Se extrae una bola de cada una. ¿Cuál es la probabilidad de que entre las bolas extraídas haya
   (*i*) una blanca y dos negras; (*ii*) por lo menos dos negras;
   (*iii*) más negras que blancas.?

30. Una urna contiene 12 bolas, de las cuales 5 son blancas y 7 negras se sacan dos bolas y se vuelven a la urna. Se saca otra vez dos bolas y se - vuelven a la urna, y así continúa hasta hacer 5 etracciones.
   (a) ¿Cuál es la probabilidad de sacar dos bolas negras en cada uno de los tres primeros experimentos y una pareja de una blanca y una negra en cada una de las otras dos extracciones.?
   (b) ¿Cuál es la probabilidad de sacar dos bolas negras tres veces y las - otras dos veces, dos blancas?

31. La producción diaria de una máquina que produce una pieza muy complicada de las siguientes probabilidades para el número de piezas producidas:

$$P[\{1\}] = 0.10 \,, \qquad P[\{2\}] = 0.30 \,, \qquad P[\{3\}] = 0.60$$

Además, la probabilidad de producir piezas defectuosas es 0.03. Las defectuosas aparecen independientemente. Hallar la probabilidad de no producir piezas defectuosas en un día.

32. Se lanza 6 dados. ¿Cuál es la probabilidad que aparezcan cada uno de los números posibles?

33. Se lanzan 7 dados. ¿Cuál es la probabilidad que aparezcan cada uno de los números posibles?

34. Si una máquina que produce engranajes está trabajando correctamente, el 92% de las piezas satisfacen las especificaciones. Si la máquina no trabaja bien, sólo el 60% de los engranajes producidos satisfacen las especificaciones. La máquina trabaja correctamente el 90% del tiempo. Se seleccionan cuatro engranajes y todos satisfacen los requerimientos. ¿Cuál es la probabilidad que la máquina no haya estado trabajando bien?

35. Un fabricante está considerando comprar un lote grande de piezas de un proveedor. El fabricante estima la proporción de piezas defectuosas en el lote en la forma siguiente:

Proporción de piezas defectuosas (π)	Probabilidad de la proporción $P(\pi)$
$\pi_1 = 0.10$	$P(\pi_1) = 0.20$
$\pi_2 = 0.15$	$P(\pi_2) = 0.30$
$\pi_3 = 0.25$	$P(\pi_3) = 0.50$

Suponga que se elige 3 piezas al azar del lote:

(a) ¿Cuál es la probabilidad que los tres sean de calidad aceptable?

(b) Si las tres piezas resultaron de calidad aceptable. ¿Cuál es la probabilidad de que el lote contenga 10% de piezas defectuosas?

36. En el ejemplo 24 de 1.9. Suponga que el dado se lanza 5 veces. ¿Cuál es la probabilidad de ganar por lo menos 4 veces?

37. De tres sucesos A_1, A_2 y A_3 se sabe que son mutuamente independientes, -

que la probabilidad del primero es el doble de la del segundo, que la probabilidad de la ocurrencia simultánea de los 2 primeros sucesos es 0.02 y que la probabilidad de que ocurra al menos uno de ellos es 0.64. Calcular la probabilidad de cada uno de los eventos.

38. Un aparato tiene 4 válvulas que funcionan independientemente, sus probabilidades de falla son: 0.1, 0.2, 0.3, 0.4 respectivamente , para la primera, segunda, tercera y cuarta válvula. Dos de estas válvulas han fallado. Hallar la probabilidad de que hayan fallado la primera y segunda.

39. Un circuito eléctrico consta de 4 interruptores en serie. Suponga que el funcionamiento de los interruptores son estadisticamente independientes. Si la probabilidad de falla (esto es, que quede abierto) de cada interruptor es de 0.2. ¿Cuál es la probabilidad de falla del circuito?.

40. Resuelva el problema anterior cuando el circuito consta de 4 interruptores en paralelo.

41. Las probabilidades de que tres tubos se quemen son respectivamente, 0.1 , 0.2 y 0.3. Las probabilidades de que un proyector se pare; si uno, dos o tres tubos se queman son: 0.25, 0.6 y 0.9, respectivamente. Hallar la probabilidad de que el proyector se pare.

42. La compañía constructora "La amiga" debe tener cuando menos dos obras dentro de una semana para mantener el empleo de su personal básico. La compañía ha sometido proyectos para cada una de las licitaciones de tres obras de tipo A y dos obras de tipo B. Las firmas ganadoras serán comunicadas - dentro de la semana crucial. Suponga que la compañía tiene probabilidad - 1/2 de que se le otorgue una obra de tipo A y probabilidad 3/4 de que se le otorgue una obra de tipo B. Si las decisiones serán hechas independientemente, ¿Cuál es la probabilidad de que dicha firma esté en condiciones de continuar el empleo de su personal básico?

43. Dos de tres elementos de una calculadora, que funcionan independientemente, fallaron. Hallar la probabilidad que hayan fallado los elementos - primero y segundo; si las probabilidades de falla de los elementos primero, segundo y tercero son respectivamente iguales a $p_1 = 0.2$; $p_2 = 0.4$; $p_3 = 0.3$.

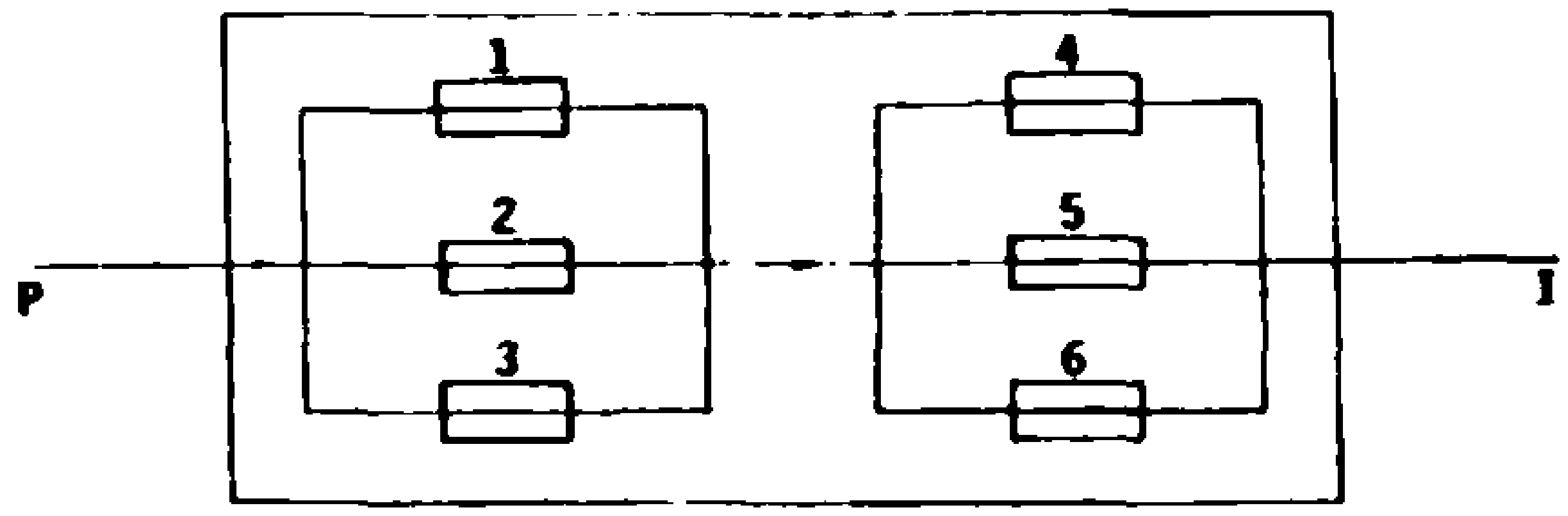
44. Las probabilidades de que tres tiradores A, B y C den en el blanco son - respectivamente, 1/3, 1/4, y 1/5. Cada uno dispara una vez al blanco, se

pide:

(i) ¿Cuál es la probabilidad que hayan sido B y C?, si exactamente dos - dan en el blanco.

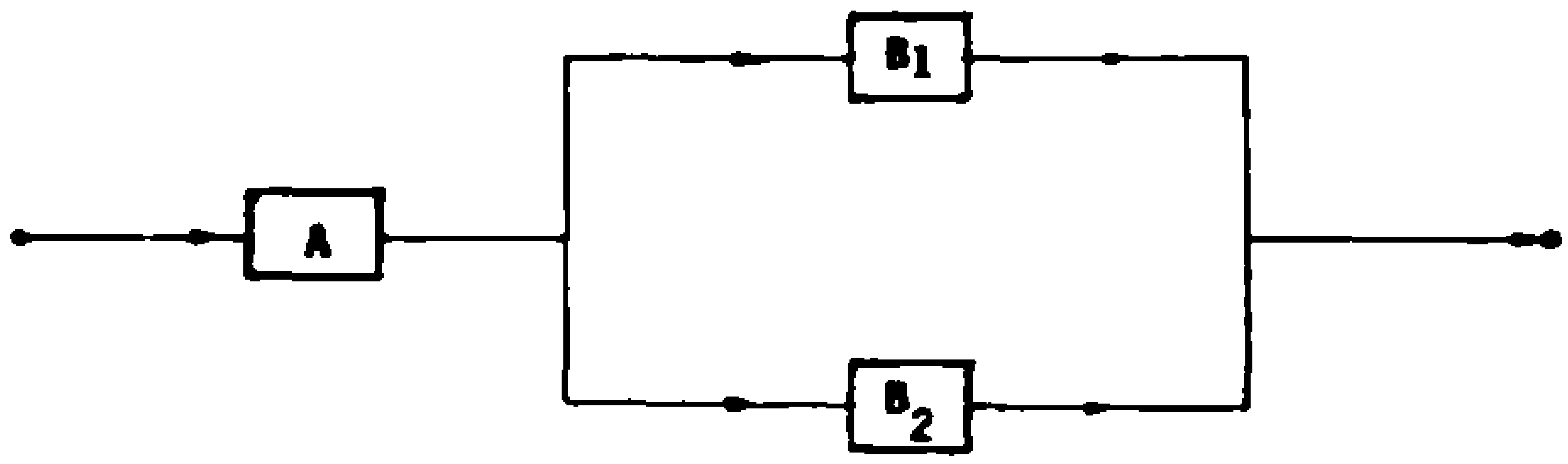
(ii) ¿Cuál es la probabilidad que hayan sido B ó C?, si se ha dado en el blanco.

45. Se tiene el siguiente circuito del diagrama.



Halle la probabilidad que el circuito falle (no pase la corriente de P a I); siendo 0.3, la probabiliadd de falla de cualquiera de los 6 componentes del circuito.

46. Para que funcione adecuadamente, un equipo electrónico debe tener las dos componentes conectadas que aparecen en el diagrama en correcto funcionamiento. El diagrama muestra que A debe funcionar y lo mismo alguno de los dos B. Suponga que las componentes B funcionan independientemente de A e independientemente una de otra, y que la confiabilidad de A es 0.9 y la - de B_1 y B_2 es 0.8. Calcular la confiabilidad del equipo.



47. Una componente juega un papel esencial en el funcionamiento de un determinado equipo. Si en lugar de instalar un componente, se utiliza un sistema identico con varios en paralelo, aumenta la confiabilidad del equipo ya - que seguirá funcionando siempre y cuando uno de los componentes esté funcionando. El correcto funcionamiento de una nave espacial depende de un - mecanismo cuya confiabilidad es de 95% . ¿Cuántos de estos mecanismos deben incorporarse al sistema para que la seguridad que el mecanismo funcione en forma satisfactoria sea (a) 99% ; (b) 99.9% ; (c) 99.99% ?

48. Las series mundiales de béisbol termina cuando uno de los equipos gana su cuarto juego. Suponga que los dos equipos tienen igual habilidad. ¿Cuál - es la probabilidad que la serie termina al final del cuarto juego? ¿Al quinto juego? ¿Al sexto juego?

49. Se lanzan dos dados simultaneamente y se repite tres veces el experimento ¿Cuál es la probabilidad que salgan por lo menos una vez la suma 7 y la suma 9?.

50. Se lanzan simultaneamente dos dados honestos y se repite la experiencia - tres veces. ¿Cuál es la probabilidad de obtener por lo menos una vez la suma 7 y 11?

51. Dos personas juegan a "cara" o "sello" con una sola moneda, y convienen - en continuar el juego, hasta que tanto las caras como los sellos hayan - aparecido cuando menos dos veces. Hállese la probabilidad que el juego no termine cuando se ha lanzado la moneda cinco veces.

52. Cada uno de tres tiradores hace un disparo contra tres blancos fugaces, escogiendo un blanco al azar e independientes de los demás tiradores. La probabilidad es p_1 ,p_2 y p_3 respectivamente. Determine las probabilidades de los eventos siguientes:

A: "Exactamente en uno de los blancos no habrá ningún agujero".

B: "En cada uno de los blancos habrá al menos un agujero".

C: "Los tres agujeros aparece en el mismo blanco".

53. La probabilidad de obtener cría con un huevo fértil es 0.95. Se sacan para incubar, tres huevos de una caja que contenía 12, de los cuales 4 huevos eran fértiles y 8 infértiles. ¿Cuál es la probabilidad de obtener alguna cría de los tres huevos?

1.10 PROBABILIDAD EN ESPACIO MUESTRAL INFINITO NUMERABLE Y CONTINUO

Aunque no lo hemos mencionado, el lector habrá notado que todos los ejemplos anteriores de probabilidad han sido referidos a espacios muestrales finitos. Así, la definición clásica de probabilidad se ha definido para espacios muestrales finitos, pues cuando se habla *de espacios muestrales infinitos*, no tienen sentido hablar del cociente n_A/n, ya que el espacio muestral tiene infinitos elementos. Sin embargo la definición de probabilidad en un - espacio muestral finito dado en 1.5.5, puede ser modificado, asignando probabilidades p_i a todos los posibles sucesos ω_i, $i = 1,2, \ldots$; es decir $p_i = P[\{\omega_i\}]$, tal que:

(1) $p_i > 0$, $i = 1,2, \ldots$

(2) $\sum_{i=1}^{\infty} p_i = 1$

Luego, definimos la probabilidad de un evento A en Ω de la siguiente manera

$$P[A] = \sum_{\omega_i \in A} P[\{\omega_i\}] = \sum_{i/\omega_i \in A} p_i$$

Obsérvese que la sumatoria está tomada sobre todos los sucesos favorables a A.

EJEMPLO 1 Se lanza una moneda hasta que ocurra cara. Calcular la probabilidad de lanzarla a lo maś 3 veces.

SOLUCION El espacio muestral es

$$\Omega = \{C , SC , SSC , SSSC, \ldots \}$$

es claro que los sucesos no tienen la misma probabilidad así

$$P[\{\omega_1\}] = P[C] = \frac{1}{2}$$

$$P[\{\omega_2\}] = P[SC] = \frac{1}{2} \times \frac{1}{2} = \frac{1}{2^2}$$ como ya hemos visto.

Y así sucesivamente, en general obtenemos

$$P[\{\omega_i\}] = \frac{1}{2^i} , \quad i = 1,2,3, \ldots$$

Ahora, sea E el evento: "lanzarla a lo más 3 veces".
Entonces,

$$E = \{\omega_1, \omega_2, \omega_3\} = \{C, SC, SSC\} \quad \text{y}$$

$$P[E] = P[\{\omega_1\}] + P[\{\omega_2\}] + P[\{\omega_3\}]$$
$$= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8} .$$

EJEMPLO 2 Se lanza un dado hasta que ocurra un 4; calcular la probabilidad de lanzar

(a) 3 veces. (b) A lo más 3 veces.

SOLUCION El espacio muestral se puede escribir así,

$$\Omega = \{4, {}_*4 , {}_{**}4 , {}_{***}4, \ldots\} = \{\omega_1, \omega_2, \omega_3, \ldots \}$$

donde cada * representa un resultado diferente de 4.

$$P[\{\omega_1\}] = \frac{1}{6}\,; \qquad P[\{\omega_2\}] = \frac{5}{6}\times\frac{1}{6}\,; \qquad P[\{\omega_3\}] = (\frac{5}{6})^2\,\frac{1}{6}$$

en general

$$P[\{\omega_i\}] = (\frac{5}{6})^{i-1}.\,\frac{1}{6}\,, \qquad i = 1,2,\ \ldots$$

Entonces,

(a) $P[\{\omega_3\}] = P[\{**4\}] = (\frac{5}{6})^2\,\frac{1}{6} = \frac{25}{216}$

(b) Si **A**, es el evento: "lanzar a lo más 3 veces".

$$A = \{4,\ *4,\ **4\} = \{\omega_1,\ \omega_2,\ \omega_3\}$$

$$P[A] = \frac{1}{6} + \frac{5}{6}\times\frac{1}{6} + (\frac{5}{36})^2\,\frac{1}{6} = \frac{91}{216}\,.$$

El diagrama del árbol de probabilidades que ilustra este ejemplo es

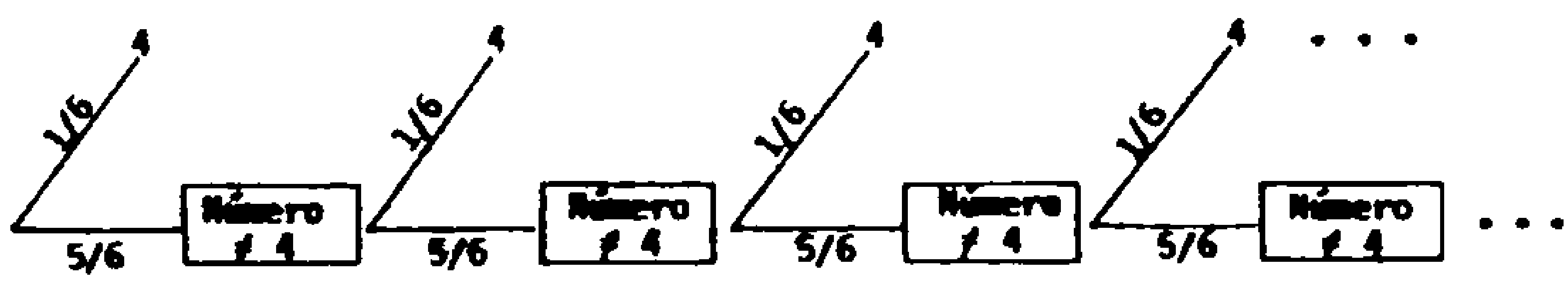


Fig. 1.10.1

EJEMPLO 3 Ud. lanza alternativamente un dado y una moneda hasta obtener 6 en el dado o cara en la moneda; en el primer caso gana y en el segundo pierde. Calcular la probabilidad de ganar.

SOLUCION Este es un típico ejemplo de experimentos "truncados" obsérvese que la ocurrencia de 6 en el dado, cara en la moneda detiene el experimento. Enumeremos a continuación los resultados.

$$\omega_1 = 6$$
$$\omega_2 = *\,C$$
$$\omega_3 = *\,S6$$
$$\omega_4 = *\,S\,*\,C$$
$$\omega_5 = *\,S\,*\,S\,6$$
....
....

Entonces, $\Omega = \{6,\ C,\ *\,S6,\ *\,S*C,\ *S*S6, *\,S*S*C,\ \ldots.\}$ donde cada * representa un número diferente de 6 con probabilidad 5/6 y la probabilidad que salga 6 es 1/6. Sea G el evento: "ganar". Los sucesos en los cuales se gana son

$$G = \{\omega_1, \omega_3, \omega_5, \omega_7, \omega_9, \ldots\} .$$

$$P[G] = P[\{\omega_1\}] + P[\{\omega_3\}] + P[\{\omega_5\}] + P[\{\omega_7\}] + \ldots$$

$$= \frac{1}{6} + \frac{5}{6} \cdot \frac{1}{2} \cdot \frac{1}{6} + (\frac{5}{6})^2 \cdot (\frac{1}{2})^2 \cdot \frac{1}{6} + (\frac{5}{6})^3 (\frac{1}{2})^3 \cdot \frac{1}{6} + \ldots$$

$$= \frac{1}{6}\left[1 + \frac{5}{12} + (\frac{5}{12})^2 + (\frac{5}{12})^3 + \ldots\right]$$

$$= \frac{1}{6}\left[\frac{1}{1 - \frac{5}{12}}\right] = \frac{1}{6} \cdot \frac{12}{7} = \frac{2}{7} .$$

NOTA Para el lector no familiarizado con sumas de infinitos términos.

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r} , \qquad \text{si } |r| < 1$$

Una forma no muy rigurosa, pero conveniente de demostrar esto es la siguiente

$$\sum_{k=0}^{\infty} r^k = 1 + r + r^2 + r^3 + \ldots \tag{1}$$

$$r \sum_{k=0}^{\infty} r^k = r + r^2 + r^3 + r^4 + \ldots \tag{2}$$

restando (2) de (1)

$$\sum_{k=0}^{\infty} r^k - r \sum_{k=0}^{\infty} r^k = 1$$

Factorizando

$\sum_{k=0}^{\infty} r^k (1 - r) = 1$, y a partir de esto obtenemos finalmente

$$\sum_{k=0}^{\infty} r^k = 1 + r + r^2 + r^3 + \ldots = \frac{1}{1 - r}$$

en el ejemplo 3, $r = \frac{5}{12}$

EJEMPLO 4 Se dispara un rifle hasta acertar en el blanco. Suponer que la probabilidad que se acierte es de 0.9 para cada tiro y que los tiros son independientes. Calcular la probabilidad

(a) de que se necesitan más de dos disparos.

(b) de que el número de disparos requeridos sea múltiplo de 3.

SOLUCION Sea el evento A: "acertar en el blanco", entonces

$$\Omega = \{A, \bar{A}A, \bar{A}\bar{A}A, \bar{A}\bar{A}\bar{A}A, \bar{A}\bar{A}\bar{A}\bar{A}A, \ldots\}$$

(a) Sea E, el evento: "se necesita más de dos tiros", entonces $\bar{E}$; es el evento: "se necesita a lo más dos tiros". Luego,

$$P[E] = 1 - P[\bar{E}] = 1 - [P[A] + P[\bar{A}A]]$$

$$= 1 - [(0.9 + (0.1)(0.9)] = 1 - 0.99 = 0.01$$

(b) Sea F el evento: "se necesita un número múltiplo de 3".

Entonces, $F = \{\bar{A}\bar{A}A, \bar{A}\bar{A}\bar{A}\bar{A}\bar{A}A, \ldots\}$. Luego,

$$P[F] = (0.1)^2(0.9) + (0.1)^5(0.9) + (0.1)^8(0.9) + \ldots$$

$$= (0.1)^2(0.9)[1 + (0.1)^3 + (0.1)^6 + \ldots]$$

$$= \frac{1}{100} \times \frac{9}{10}\left[\frac{1}{1 - \frac{1}{1000}}\right] = \frac{9}{10^3}\ \frac{10^3}{999} = \frac{1}{111} .$$

EJEMPLO 5 Tres jugadores A, B y C arrojan al aire una moneda en ese orden, hasta que aparezca una "cara" (gana quien la obtiene primero). ¿Cuáles son las respectivas probabilidades de ganar?

SOLUCION Sean los eventos :

G_A : "gana el jugador A".

G_B : "gana el jugador B".

G_C : "gana el jugador C".

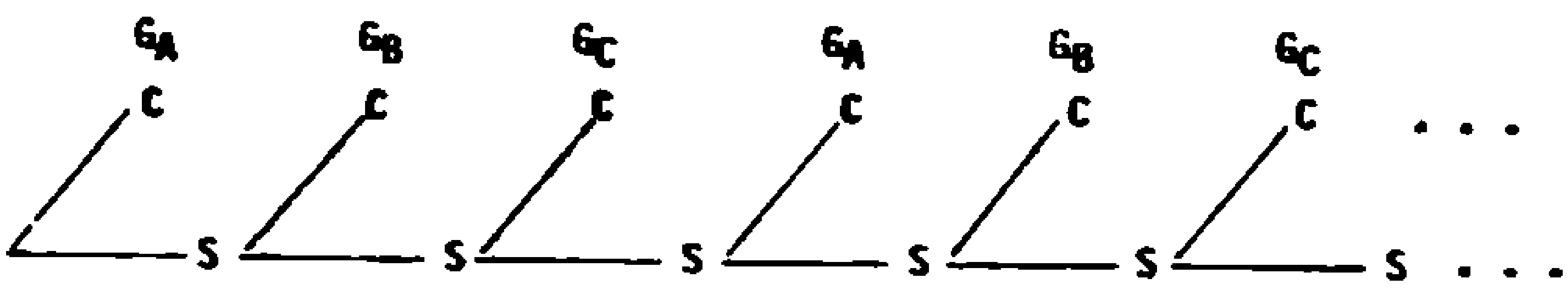


Fig. 1.10.2

El espacio muestral es

$$\Omega = \{C, SC, SSC, SSSC, SSSSC, SSSSSC, \ldots\}$$

Los sucesos favorables a cada evento son respectivamente,

$$G_A = \{C,\ SSSC,\ SSSSSSC,\ .\ .\ .\ \}$$

$$G_B = \{SC,\ SSSSC,\ SSSSSSSC,\ .\ .\ .\}$$

$$G_C = \{SSC,\ SSSSSC,\ SSSSSSSSC,\ .\ .\ .\}$$

$$P[G_A] = \frac{1}{2} + (\frac{1}{2})^4 + (\frac{1}{2})^7 + (\frac{1}{2})^{10} + \ .\ .\ .$$

$$= \frac{1}{2}\left[1 + (\frac{1}{2})^3 + (\frac{1}{2})^6 + (\frac{1}{2})^9 + \ .\ .\ .\right]$$

$$= \frac{1}{2}\left[\frac{1}{1-\frac{1}{8}}\right] = \frac{1}{2}\cdot\frac{8}{7} = \frac{4}{7}\ .$$

$$P[G_B] = (\frac{1}{2})^2 + (\frac{1}{2})^5 + (\frac{1}{2})^8 + (\frac{1}{2})^{11} + \ .\ .\ .$$

$$= (\frac{1}{2})^2\left[1 + (\frac{1}{2})^3 + (\frac{1}{2})^6 + (\frac{1}{2})^9 + \ .\ .\ .\right]$$

$$= \frac{1}{4} \times \frac{8}{7} = \frac{2}{7}\ .$$

$$P[G_C] = (\frac{1}{2})^3 + (\frac{1}{2})^6 + (\frac{1}{2})^9 + (\frac{1}{2})^{12} + \ .\ .\ .$$

$$= (\frac{1}{2})^3\left[1 + (\frac{1}{2})^3 + (\frac{1}{2})^6 + (\frac{1}{2})^9 + \ .\ .\ .\right]$$

$$= \frac{1}{8} \times \frac{8}{7} = \frac{1}{7}\ .$$

EJEMPLO 6 Una persona lanza repetidas veces dos dados y gana si obtiene 8 - puntos antes de obtener 7. Calcular la probabilidad de ganar.

SOLUCION Los sucesos posibles en los cuales la persona gana es

$$\omega_1 = 8$$

$$\omega_2 = *\ 8$$

$$\omega_3 = *\ *\ 8$$

$$\omega_4 = ***8$$

$$\omega_5 = ****8$$

. . . .

. . . .

donde cada * representa un resultado diferente de 7, 8 y tiene una probabili_ dad igual a

$$1 - P[\{7\}] - P[\{8\}] = 1 - \frac{6}{36} - \frac{5}{36} = \frac{25}{36}$$

Se considera diferente de 8 pues en caso contrario termina el juego y diferente de 7 pues en caso contrario pierde. Es decir,

$$\Omega = \{8,\ 7,\ *\ 8,\ *\ 7,\ **\ 8,\ **7,\ .\ .\ .\}$$

Sea G el evento: "ganar el juego".

$$P[G] = P[\{\omega_1\}] + P[\{\omega_2\}] + P[\{\omega_3\}] + \ldots$$

$$= \frac{5}{36} + \frac{25}{36} \cdot \frac{5}{36} + \left(\frac{25}{36}\right)^2 \cdot \frac{5}{36} + \left(\frac{25}{36}\right)^3 \cdot \frac{5}{36} + \ldots$$

$$= \frac{5}{36}\left[1 + \frac{25}{36} + \left(\frac{25}{36}\right)^2 + \left(\frac{25}{36}\right)^3 + \ldots\right]$$

$$= \frac{5}{36}\left[\frac{1}{1 - \frac{25}{36}}\right] = \frac{5}{36} \cdot \frac{36}{11} = \frac{5}{11} .$$

EJEMPLO 7 Tres personas participan en un juego llamado disparejo, en el cual cada uno lanza al aire simultáneamente una moneda; si uno de los resultados es diferente a los otros dos, la persona que obtiene el resultado diferente pierde.

(a) ¿Cuál es la probabilidad que uno de ellos pierda, en una tirada, si las tres monedas no estan cargadas?

(b) Si ninguno pierde en la primera vuelta, se lanza al aire las monedas nuevamente, hasta que alguno pierda.
¿Cuál es la probabilidad que se necesite un número par de tiradas para que alguien pierda?

SOLUCION (a) Sea E, el evento: "perder en una jugada". Entonces, los sucesos favorables a E son,

$$E = \{\underbrace{CCS}_{P_3^{2,1}},\ \underbrace{SSC}_{P_3^{2,1}}\} = \{CCS,\ CSC,\ SCC,\ SSC,\ SCS,\ CSS\},$$

Luego,

$$P[E] = \binom{3}{2}\left(\frac{1}{2}\right)^3 + \binom{3}{2}\left(\frac{1}{2}\right)^3 = 6\left(\frac{1}{2}\right)^3 = \frac{3}{4} .$$

(b) Sea F, el evento: "se necesita un número par de lanzamientos para que alguien pierda".

El lector puede escribir el espacio muestral y verificar que se obtiene

$$P[F] = 12 \times \frac{1}{8} \times \frac{1}{8} + 48 \times \frac{1}{8} \times \frac{1}{8} \times \frac{1}{8} \times \frac{1}{8} + 192 \times \frac{1}{8} \times \frac{1}{8} \times \frac{1}{8} \times \frac{1}{8} \times \frac{1}{8} \times \frac{1}{8} + \ldots$$

$$= \frac{12}{8^2} + \frac{48}{8^4} + \frac{192}{8^6} + \ldots$$

$$= \frac{12}{8^2}\left[1 + \frac{4}{8^2} + \frac{16}{8^4} + \ldots\right] = \frac{12}{8^2}\left[\frac{1}{1 - \frac{4}{8^2}}\right] = \frac{1}{5} .$$

EJEMPLO 8 Tres jugadores A, B y C, extraen (en ese orden) una carta con reposición de una baraja de 52 cartas. El primero que obtiene corazón gana. Calcular la probabilidad que gane A .

SOLUCION Sean los siguientes eventos :

G_A: "gana el jugador A".

C : "carta corazón", y $\bar{C}$: "carta diferente de corazón".

$$P[C] = \frac{1}{4} \qquad \text{y} \qquad P[\bar{C}] = \frac{3}{4}$$

Construimos el árbol de probabilidades siguiente

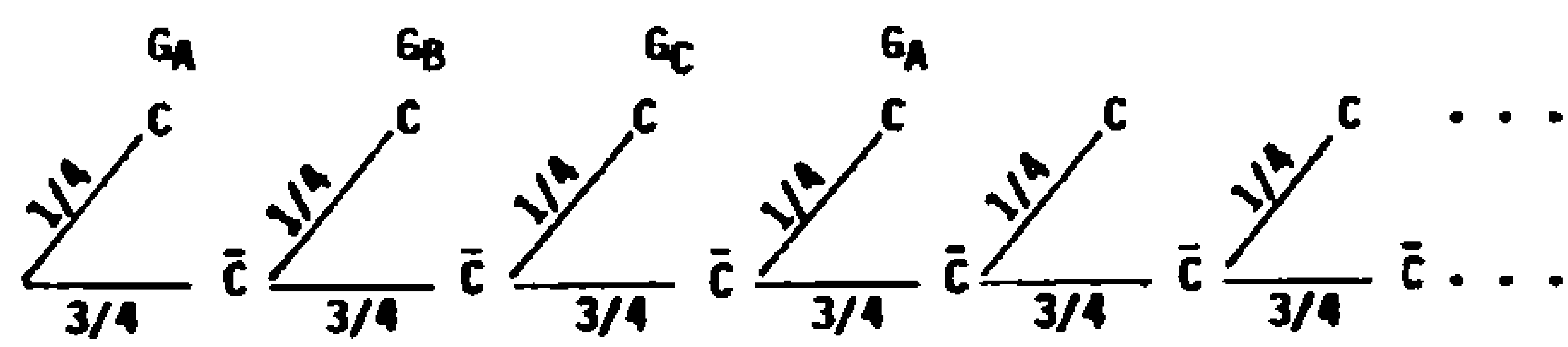


Fig. 1.10.3.

$$\Omega = \{C, \bar{C}\,C, \bar{C}\,\bar{C}\,C, \bar{C}\,\bar{C}\,\bar{C}\,C, \bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,C, \bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,C, \ldots\}$$

$$G_A = \{C, \bar{C}\,\bar{C}\,\bar{C}\,C, \bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,C, \bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,\bar{C}\,C, \ldots\}$$

Por lo tanto,

$$P[G_A] = \frac{1}{4} + \left(\frac{3}{4}\right)^3 \frac{1}{4} + \left(\frac{3}{4}\right)^6 \frac{1}{4} + \left(\frac{3}{4}\right)^9 \frac{1}{4} + \ldots$$

$$= \frac{1}{4}\left[1 + \left(\frac{3}{4}\right)^3 + \left(\frac{3}{4}\right)^6 + \left(\frac{3}{4}\right)^9 + \ldots\right]$$

$$= \frac{1}{4}\left[\frac{1}{1 - \frac{27}{64}}\right] = \frac{1}{4} \cdot \frac{64}{37} = \frac{16}{37} .$$

EJEMPLO 9 Dos jugadores que tienen la misma habilidad juegan una secuencia de partidas, hasta que uno de ellos gane dos juegos consecutivos. Determinar

(a) la probabilidad que se necesiten un número par de jugadas para terminar el juego.

(b) la probabilidad que se necesite un número impar de jugadas para terminar el juego;

(c) la probabilidad de ganar de cada jugador.

SOLUCION Llamaremos A y B a los jugadores, y G_A, G_B los eventos, que gana una partida el jugador A y B respectivamente.
Entonces algunos de los resultados posibles de las partidas se observan en el diagrama de la fig. 1.10.4 .

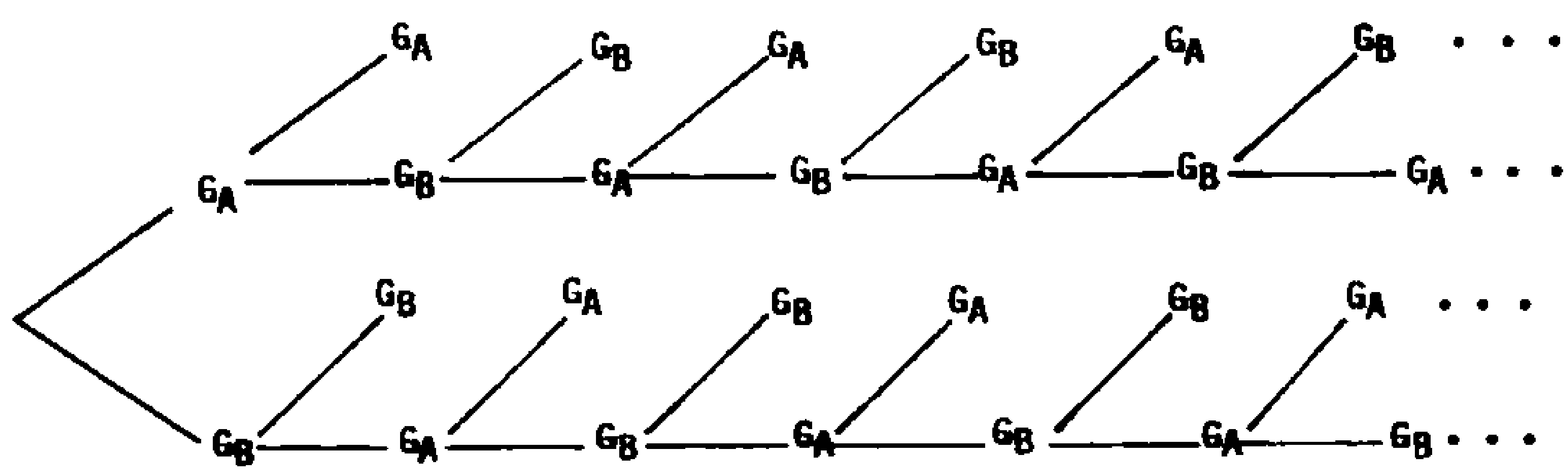


Fig. 1.10.4

Es decir, el espacio muestral tiene la siguiente forma

$$\Omega = \{G_AG_A,\ G_BG_B,\ G_AG_BG_B,\ G_BG_AG_A,\ G_AG_BG_AG_A,\ G_BG_AG_BG_B,\ \ldots\}$$

puesto que ambos jugadores tienen la misma habilidad, entonces

$$P[G_A] = P[G_B] = \frac{1}{2}\ .$$

Además los eventos G_A y G_B son independientes.

(a) Sea E: "se necesita un número par de jugadas para terminar el juego". Entonces, los elementos del evento E son,

$$E = \{G_AG_A,\ G_BG_B,\ G_AG_BG_AG_A,\ G_BG_AG_BG_B,\ \ldots\}$$

Luego

$$P[E] = \underbrace{(\frac{1}{2})^2 + (\frac{1}{2})^2}_{(\frac{1}{2})} + \underbrace{(\frac{1}{2})^4 + (\frac{1}{2})^4}_{(\frac{1}{2})^3} + \underbrace{(\frac{1}{2})^6 + (\frac{1}{2})^6}_{(\frac{1}{2})^5} + \ldots$$

$$P[E] = \frac{1}{2} + (\frac{1}{2})^3 + (\frac{1}{2})^5 + \ldots$$

$$= \frac{1}{2}\left[1 + (\frac{1}{2})^2 + (\frac{1}{2})^4 + \ldots\right]$$

La expresión dentro del corchete, es una serie geométrica de razón $(1/2)^2$. - Por lo tanto

$$P[E] = \frac{1}{2}\left[\frac{1}{1-(\frac{1}{2})^2}\right] = \frac{2}{3}\ .$$

(b) Sea F: "se necesita un número impar de jugadas para terminar el juego". Los elementos del evento F son,

$$F = \{G_AG_BG_B,\ G_BG_AG_A,\ G_AG_BG_AG_BG_B,\ G_BG_AG_BG_AG_A,\ \ldots\}$$

$$P[F] = \underbrace{(\frac{1}{2})^3 + (\frac{1}{2})^3}_{(\frac{1}{2})^2} + \underbrace{(\frac{1}{2})^5 + (\frac{1}{2})^5}_{(\frac{1}{2})^4} + \underbrace{(\frac{1}{2})^7 + (\frac{1}{2})^7}_{(\frac{1}{2})^6} + \ldots$$

$$= (\frac{1}{2})^2 + (\frac{1}{2})^4 + (\frac{1}{2})^6 + \ldots$$

$$= (\frac{1}{2})^2\left[1 + (\frac{1}{2})^2 + (\frac{1}{2})^4 + \ldots\right]$$

$$= (\frac{1}{2})^2\left[\frac{1}{1-(\frac{1}{2})^2}\right] = \frac{1}{3}\ .$$

(c) Sea A : "gana el juego el jugador A"

$$A = \{G_AG_A,\ G_BG_AG_A,\ G_AG_BG_AG_A,\ \ldots\}$$

$$P[A] = (\frac{1}{2})^2 + (\frac{1}{2})^3 + (\frac{1}{2})^4 + (\frac{1}{2})^5 + \ldots\}$$

$$= (\frac{1}{2})^2\left[1 + (\frac{1}{2}) + (\frac{1}{2})^2 + (\frac{1}{2})^3 + \ldots\right]$$

$$= (\frac{1}{2})^2\left[\frac{1}{1-\frac{1}{2}}\right] = \frac{1}{2}\ .$$

Sea B : "gana el juego el jugador B"

En forma completamente similar para el caso del jugador A, se obtiene

$$P[B] = \frac{1}{2} .$$

1.10.1 ESPACIO MUESTRAL CONTINUO

El último tipo de espacios muestrales considerando, es el del tipo continuo. Ninguno de las definiciones dadas de probabilidad es aplicable es este caso. También se presenta otra dificultad. En los espacios muestrales discretos, todos los subconjuntos se llaman eventos y se les puede asignar probabilidades. Pero pueden construirse subconjuntos de un espacio muestral - continuo que no son eventos por lo tanto cualquier asignación de probabilidad que se les haga es inconsistente con los axiomas de probabilidad. Sin embargo en los problemas prácticos que se estudien en este libro, los subconjuntos de Ω serán eventos. Y por ahora estudiaremos el tipo de espacio muestral continuo que tiene sus elementos de la misma verosimilitud, esto significa que la probabilidad que un punto ocurre en un subconjunto de Ω es proporcional a la longitud del subintervalo. Así definimos la probabilidad en un espacio muestral contínuo: como la razón entre la longitud del evento "ℓ_A" y la longitud del espacio muestral "ℓ_Ω" o "$\ell(\Omega)$"

$$P[A] = \frac{\ell_A}{\ell_\Omega}$$

Aquí el concepto de longitud representa un concepto más amplio; según el caso puede ser: longitud misma, área, volumen, etc; más apropiadamente se puede hablar de medida del evento $m(\Omega)$ y $m(A)$, luego

$$P[A] = \frac{m(A)}{m(\Omega)}$$

EJEMPLO 10 Se elige aleatoriamente un punto dentro del segmento determinado por el intervalo $[2, 10]$. ¿Calcular la probabilidad que pertenezca al - segmento $[3, 5]$?.

SOLUCION

$$\Omega = \{x \mid x \in [2, 10]\} \quad y \quad \ell_\Omega = 10 - 2 = 8,$$

$$A = \{x \mid x \in [3, 5]\} \quad y \quad \ell_A = 5 - 3 = 2 .$$

Luego, $P[A] = \frac{2}{8} = \frac{1}{4}$.

EJEMPLO 11 El espacio muestral asociado a un experimento aleatorio es el conjunto

$\Omega = \{(x, y) \in \mathbb{R}^2 \,/\, x^2 + y^2 \leqslant 10\}$. Calcular la probabilidad del evento

$$E = \{(x, y) \in \Omega \,/\, x^2 + y^2 \leqslant 5\}$$

SOLUCION Es claro que Ω es un círculo de radio $R = 10$ y el evento E es un círculo de radio $r = 5$. Luego,

$$P[E] = \frac{A_E}{A_\Omega} = \frac{\pi r^2}{\pi R^2} = \frac{25\pi}{100\pi} = \frac{1}{4},$$

donde A_E = área del círculo radio 5

A_Ω = área del círculo de radio 10.

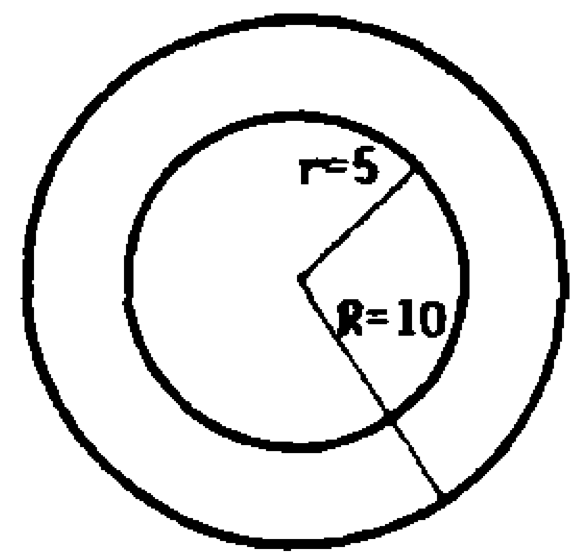


Fig. 1.10.5

EJEMPLO 12 Se elige un punto del cuadrado con vértices opuestos (0,0) y (1 , 1) ; sea E el *evento* la *suma* de las *coordenadas* del *punto* es *menor* que 3/4. Hallar la probabilidad de E.

SOLUCION El experimento aleatorio es "elegir un punto del cuadrado con vértice opuestos (0, 1) y (1, 1)". Entonces

$$\Omega = \{(x, y) \in \mathbb{R}^2 \,/\, 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1\}$$

Sea el evento E : " la suma de las coordenadas del punto elegido es menor que 3/4" .

$$E = \{(x, y) \in \Omega \,/\, x + y < \tfrac{3}{4}\}$$

A_Ω = Area del cuadrado = 1

A_E = Area del triángulo sombreado.

$$P[E] = \frac{A_E}{A_\Omega} = \frac{(3/4)(3/4)\,/\,2}{1}$$

$$= \frac{9}{32}.$$

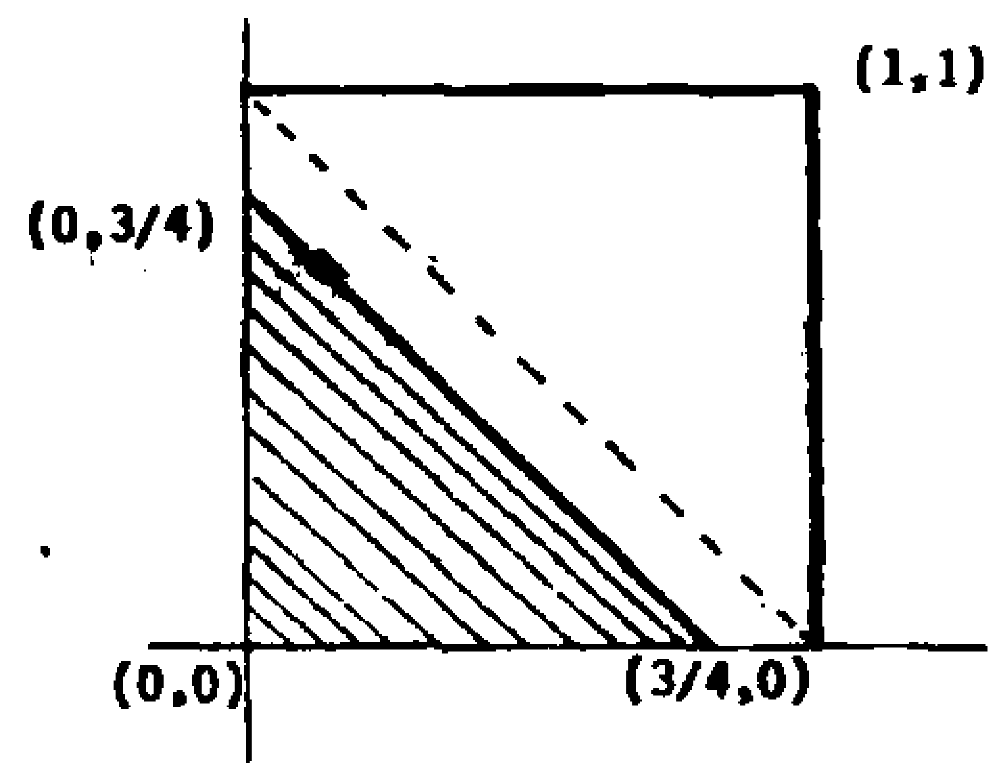


Fig. 1.10.6

EJEMPLO 13 Sea el intervalo $[-r, r]$ la base de un semicírculo, si se elige un punto aleatoriamente de este intervalo, calcular la probabilidad de que la longituddel segmento perpendicular de este punto al semicírculo es menor que $r/2$.

SOLUCION El experimento aleatorio es "elegir un punto en el intervalo $[-r,r]$

$$\Omega = \{x \ / \ x \in [-r, r]\}$$
$$\ell_\Omega = 2r$$

Sea **A**, el evento: "la longitud del segmento perpendicular del punto elegido al semi-círculo es menor que $r/2$" .

$\bar{A}$: "longitud del segmento perpendicular del punto elegido al semicírculo es mayor o igual a $r/2$ "

$$d = 2\sqrt{r^2 - (r/2)^2}$$
$$= 2\sqrt{3r^2/4} = \sqrt{3}r$$

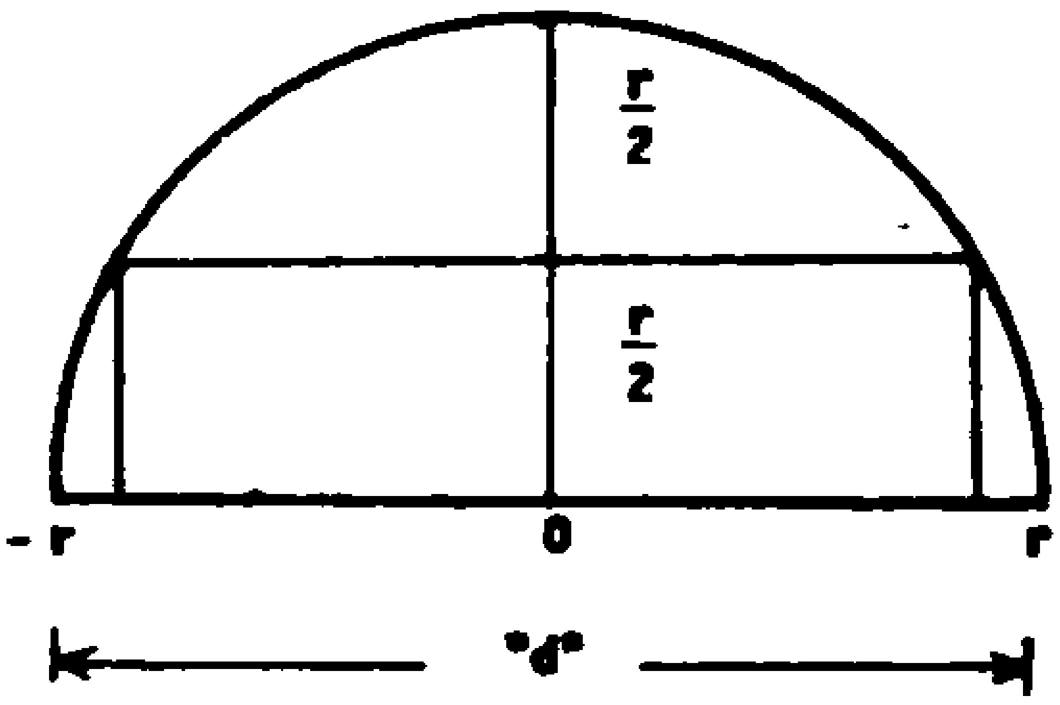


Fig. 1.10.7

Entonces,

$$P[\bar{A}] = \frac{\ell_{\bar{A}}}{\ell_\Omega} = \frac{d}{2r} = \frac{\sqrt{3}r}{2r} = \frac{\sqrt{3}}{2}.$$

Por lo tanto, $P[A] = 1 - P[\bar{A}] = 1 - \frac{\sqrt{3}}{2}$.

EJEMPLO 14 En un segmento AB de longitud "ℓ" se escoge al azar dos puntos L y M Calcular la probabilidad que el punto L esté más cercano a M que a A.

SOLUCION El experimento aleatorio es "elegir dos puntos del segmento AB"

Sea $x = AL$ e $y = AM$, entonces

$$\Omega = \{(x, y) \ / \ 0 \leq x \leq \ell, \ 0 \leq y \leq \ell\}$$

Por lo tanto

$$m(\Omega) = \ell \times \ell = \ell^2$$

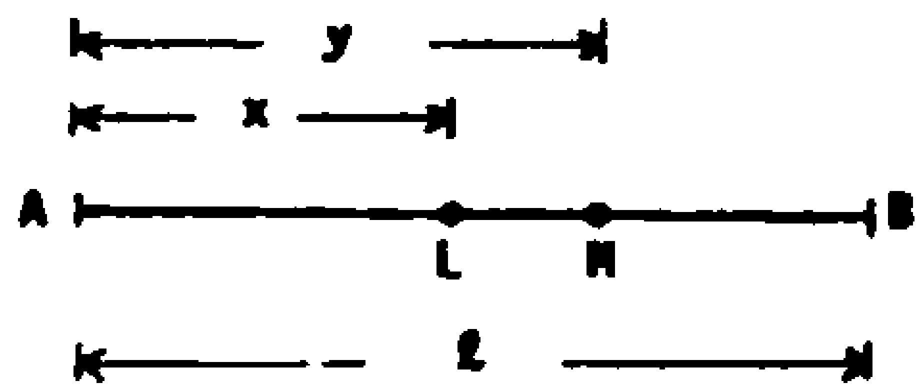


Fig. 1.10.8

Sea C, el evento: "el punto L está más cercano a M que a A".

$LM < AL = |y - x| < x \Longleftrightarrow 0 < y < 2x$, estos puntos están dados en la fig. 1.10.9, cuya área da los casos favorables al evento C. Es decir

$$C = \{(x, y) \ / \ 0 < y < 2x\}$$

$$m(C) = \frac{3\ell^2}{4}$$

luego, $$P[C] = \frac{3\ell^2/4}{\ell^2} = \frac{3}{4}.$$

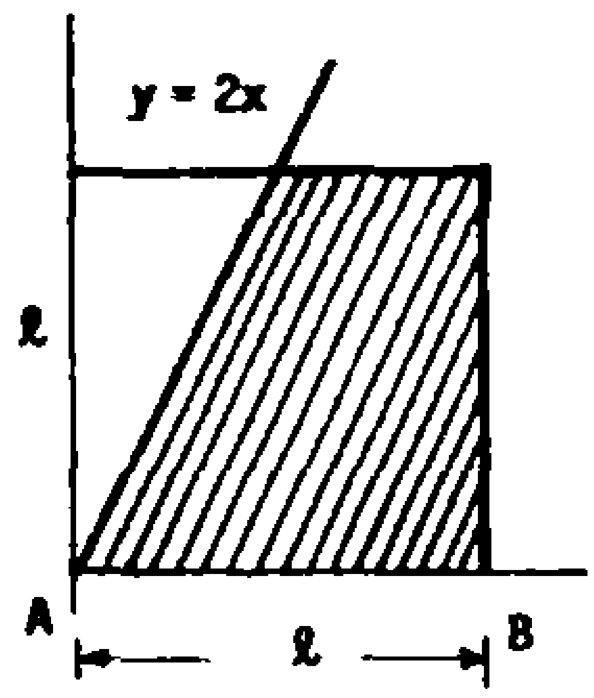


Fig. 1.10.9

EJEMPLO 15 La luz de un semáforo aparece cada 4 minutos y dura un minuto para luego cambiar a verde (por lo tanto es verde durante 3 minutos, roja 1 minuto, etc). Cada hora en punto, la luz del semáforo cambia a roja primeramente.

(a) Si se llega al semáforo en un instante al azar entre las 7.55 y las 8.05 a.m., ¿Cuál es la probabilidad que usted tenga que detenerse ante el semáforo?.

(b) Si se llega al semáforo en un instante al azar entre las 7.54 y las 8.04 a.m., ¿Cuál es la probabilidad que usted debe detenerse ante el semáforo?.

SOLUCION 1 (Ud. va en automóvil) $\Omega = \{t/t \in [7.55, 8.05]\}$

(a) Sea el evento A: "ud. tenga que detenerse ante el semáforo si llega entre las 7.55 y 8.05".

R: " Tiempo que dura la luz roja" = 1 minuto.

V: "Tiempo que dura la luz verde" = 3 minutos.

Según el diagrama de la fig. 1.10.10, se tiene

Fig. 1.10.10

$$\ell_\Omega = 8.05 - 7.55 = 10 \text{ mint.}$$

$$\ell_A = 3R = 3 \text{ mint.}$$

Por lo tanto, $$P[A] = \frac{\ell(A)}{\ell(\Omega)} = \frac{3}{10}$$

(b) Sea el evento B: "Ud tenga que detenerse, si llega al semáforo entre 7.54 y 8.04"; de la fig. 1.10.11

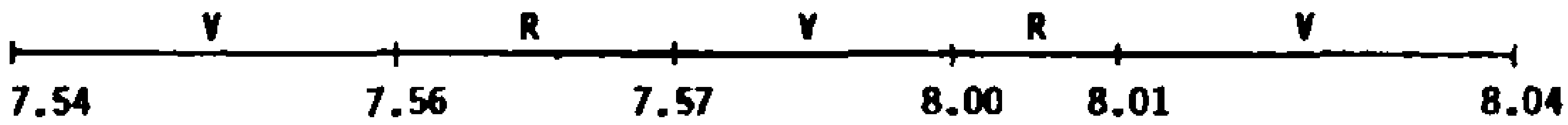


Fig. 1.10.11

$$\ell_\Omega = 8.04 - 7.54 = 10 \text{ mint.}$$

$$\ell_B = 2R = 2 \text{ mint.}$$

Luego, $P[B] = \dfrac{2}{10} = \dfrac{1}{5}$.

2. (Ud. va a pie). Queda como ejercicio para el lector.

PROBLEMAS 1.10

1. Dos jugadores lanzan alternativamente una moneda. El primero que obtiene cara gana. Hallar la probabilidad de ganar de cada jugador.

2. Dos jugadores A y B lanzan dos dados; A comienza el juego. Gana el jugador A, si obtiene 6 puntos, antes de que B obtenga 7 puntos, y gana B si saca 7 puntos antes de que A saque 6 puntos. ¿Cuál es la probabilidad de ganar de cada jugador?

3. Al lanzar un par de dados Ud. gana, si obtiene 8 ó 10 puntos antes de obtener 7 ó 9; calcular la probabilidad de ganar.

4. Tres monedas se lanzan simultáneamente hasta que las tres muestren los mismos resultados. ¿Calcular la probabilidad de
   (a) Realizar 3 lanzamientos;
   (b) Realizar 3 ó más lanzamientos;
   (c) Realizar 3 ó menos lanzamientos;
   (d) Realizar un número par de lanzamientos;
   (e) que el número de lanzamientos sea múltiplo de 3?

5. Cuatro personas juegan disparejos, para el cual cada una lanza al aire simultáneamente una moneda; si una cara es diferente de las otras tres, la persona que obtiene la cara diferente pierde.
   (a) ¿Cuál es la probabilidad que uno de ellos pierda en la primera tirada?
   (b) Si ninguno pierde en la primera tirada, se lanzan al aire las monedas nuevamente, hasta que alguno pierda.
   ¿Cuál es la probabilidad que se necesite un número par de tiradas

para que alguien pierda?

6. Se aplican alternativamente una dosis de veneno a un ratón blanco y luego a un ratón negro. La probabilidad que un ratón blanco muera por efectos del veneno es 2/5 y la probabilidad que un ratón negro muera por efectos del veneno es de 3/4. ¿Cuál es la probabilidad que muera primero un ratón blanco?

7. Un jugador arroja dos dados. Si en su primera jugada hace un total de 7 ó 11 puntos, gana el juego. Si en su primera jugada hace un total de 2,3 ó 12, punto pierde el juego. Si en su primera jugada hace un total de 4,5,6 8,9 ó 10, puntos, contínua arrojando los dados hasta que obtenga el puntaje que obtuvo en la primera tirada o hace 7. En el primer caso, gana; en el segundo pierde. ¿Cuál es la probabilidad que gane?

8. Jaime se presenta a un examen de manejo varias veces hasta que lo aprueba Suponga que la probabilidad que lo pruebe en cualquier examen sea de 0.1 y que las pruebas son independientes. Calcular la probabilidad que (a) le tome más de 4 intentos; (b) le tome más de 10 intentos.

9. Dos jugadores A y B extraen (en ese orden) alternativamente una carta con reposición, de una baraja de 52, hasta obtener un as que es la carta ganadora. Determinar la probabilidad de ganar de cada jugador.

10. La probabilidad que un estudiante de aviación apruebe el examen escrito para obtener su licencia de piloto es 0.7. Calcular la probabilidad de que un estudiante apruebe el examen,
    (a) antes del cuarto intento
    (b) después del segundo intento
    (c) en un número par de intentos.

11. Tres jugadores A, B y C, de igual habilidad en el juego, juegan de la siguiente manera; juegan A y B mientras que C descansa, el ganador de este partido se enfrenta a C mientras descansa el perdedor. El juego continúa hasta que un jugador gane dos partidos consecutivas. Determinar la probabilidad de ganar el juego cada jugador. ¿Cuál es la probabilidad que se necesita un número par de jugadas para terminar el juego?

12. Dos jugadores A y B juegan un match. Sus probabilidades respectivas de ganar una partida son entre si como 2 : 3. Determinar la probabilidad de ganar el match de cada jugador, si para ganarlo hay que ganar dos partidas

seguidas.

13. Tres jugadores A, B y C extraen aleatoriamente cada uno, una bola de una urna que contiene 12, de las cuales ocho son negras y cuatro blancas, hasta que uno de ellos saque la primera bola blanca que será el ganador. Determinar la probabilidades de ganar de cada jugador, sabiendo que empieza A, los otros siguen en el orden indicado, y que la extracción se hace con reposición.

14. En el problema 2. Suponga que los jugadores acuerdan realizar n lanzamientos. Determinar:
(a) la probabilidad de ganar de cada jugador;
(b) la probabilidad de que queden empates.

15. Un amigo y ud. hacen turno para tirar un dado hasta que uno de ustedes llegue a obtener un tres o un cuatro. Si su amigo tiro primero, ¿cuál es la probabilidad que lo obtenga ud.?

16. Tres inspectores hacen turno comprobando los componentes electrónicos tal y como salen de una cadena de montaje. Si el 10 por 100 de todos los componentes producidos en la cadena de montaje son defectuosos, ¿cuál es la probabilidad de que el inspector que compruebe el primer componente sea el mismo que encuentre el primer componente defectuoso?

17. Una regla de longitud de 20 cm. se rompe al azar en dos partes. ¿Cuál es la probabilidad que la longitud de la parte más larga sea al menos el doble de la más corta?

18. Se escoje al azar un punto entre el 0 y el 1 en el eje de las X del plano XY. A continuación se dibuja un círculo con centro en el origen, y radio determinado por el punto escogido. Calcular la probabilidad que el área del círculo sea menor que $\pi/2$.

19. Sobre el segmento AB, se toma al azar dos puntos X_1, X_2. ¿Cuál es la probabilidad que AX_1, X_1X_2, X_2B formen un triángulo?

20. El punto medio del segmento AB es M. Se elige al azar un punto X en dicho segmento. ¿Calcular la probabilidad de que pueda formarse un triángulo con los segmentos AX, BX y AM.?

21. Calcular la probabilidad de que, elegido un punto al azar en el interior de un cuadrado, ninguno de los segmentos que lo unen a los cuatro vértices sea mayor que el lado del cuadrado.

2

VARIABLES ALEATORIAS

2.1 DEFINICION Y EJEMPLOS

El lector habrá notado en el capítulo anterior, que no todo espacio muestral Ω asociado a un experimento aleatorio está constituído por elementos numéricos, si no que en muchos casos son entes abstractos; así, al hablar del lanzamiento de una o más monedas teníamos resultados tales como: C, CS, CSS, etc. Al hablar de la prueba de dos resistencias los elementos del espacio muestral son {BB, BD, DB, DD}. (B bueno, D defectuoso).

También habrá notado que muchas veces no es fácil describir el espacio muestral, Ω, asociado a un experimento aleatorio, cuando sus elementos no son números. Por otro lado la probabilidad **P** es una función cuyo dominio es $\mathcal{P}(\Omega)$ y rango el intervalo de números reales $[0, 1]$. (Es decir, $P:\mathcal{P}(\Omega) \to [0,1]$), y si los elementos de $\mathcal{P}(\Omega)$ son entes abstractos, no podemos aplicar el cálculo metemático; por lo tanto es conveniente, que el dominio de la función **P** sea también un conjunto de números reales. El objeto de la presente sección es justamente asignar un valor numérico $x \in \mathbb{R}$ a cada suceso $\omega \in \Omega$ (si no lo es), es decir "cuantificar" los sucesos. Comenzaremos nuestra discusión con un ejemplo simple.

Consideremos, el experimento aleatorio de "lanzar una moneda tres veces". El espacio muestral Ω es el conjunto formado por los siguientes 8 puntos,

Ω = {CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS}

Suponga ahora que sólo nos interesa el número de caras que salen, de manera que los sucesos CCS, CSC, SCC pueden considerarse equivalentes, también los sucesos, CSS, SCS, SSC se consideran equivalentes, podemos introducir una función X definida sobre Ω de tal manera que:

$$X(CCC) = 3,$$
$$X(CCS) = X(CSC) = X(SCC) = 2$$
$$X(CSS) = X(SCS) = X(SSC) = 1 \qquad y$$
$$X(SSS) = 0$$

Es decir, la función X en Ω definida por $X(\omega)$ = "número de caras obtenidas al lanzar una moneda tres veces" es una función a valores reales, que tiene como dominio el espacio muestral Ω y el subconjunto de números reales, $R_X = \{x/x = 0,1,2,3\}$ como rango. En símbolos

$$X: \Omega \longrightarrow \{0,1,2,3\}$$
$$\omega \rightarrow X(\omega)$$

La fig. 2.1.1 da una idea intuitiva de lo expresado en el párrafo anterior.

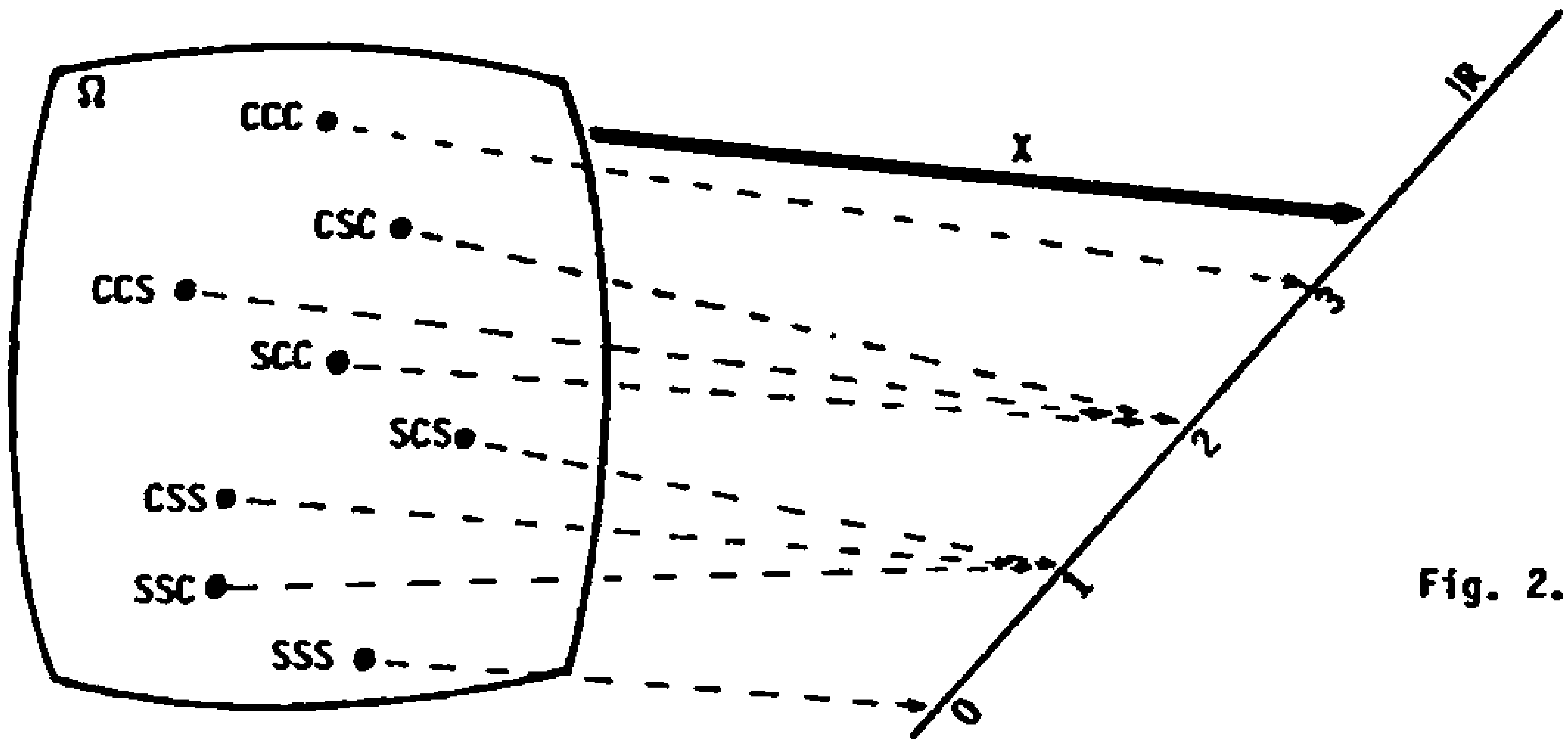


Fig. 2.1.1

tenemos así, un nuevo conjunto {0,1,2,3}, ahora formado por números reales, a cada uno de los cuales le corresponde una probabilidad de la siguiente manera:

$$P[\{3\}] = P[CCC] = \frac{1}{8}$$

$$P[\{2\}] = P[CCS] + P[CSC] + P[SCC] = \frac{3}{8}$$

$$P[\{1\}] = P[SSC] + P[SCS] + P[CSS] = \frac{3}{8} \qquad (*)$$

$$P[\{0\}] = P[SSS] = \frac{1}{8}$$

Vemos pues, que la función X hace corresponder a cada elemento ω de Ω un número real x, y además, el conjunto de elementos de Ω, cuya imagen es uno de estos números reales, es un elemento de $\mathcal{P}(\Omega)$, o sea un evento, y tiene por lo tanto, una determinada probabilidad. La función X que cumplen estas condiciones se *llama variables aleatorias*.

DEFINICION 2.1.1 Dado un experimento aleatorio ε y Ω el espacio muestral asociado a ε. Una función X que asigna a cada elemento ω en Ω uno y solamente un número real $x = X(\omega)$, se llama *variable aleatoria*. Es decir, X es una función real , $X : \Omega \longrightarrow \mathbb{R}$

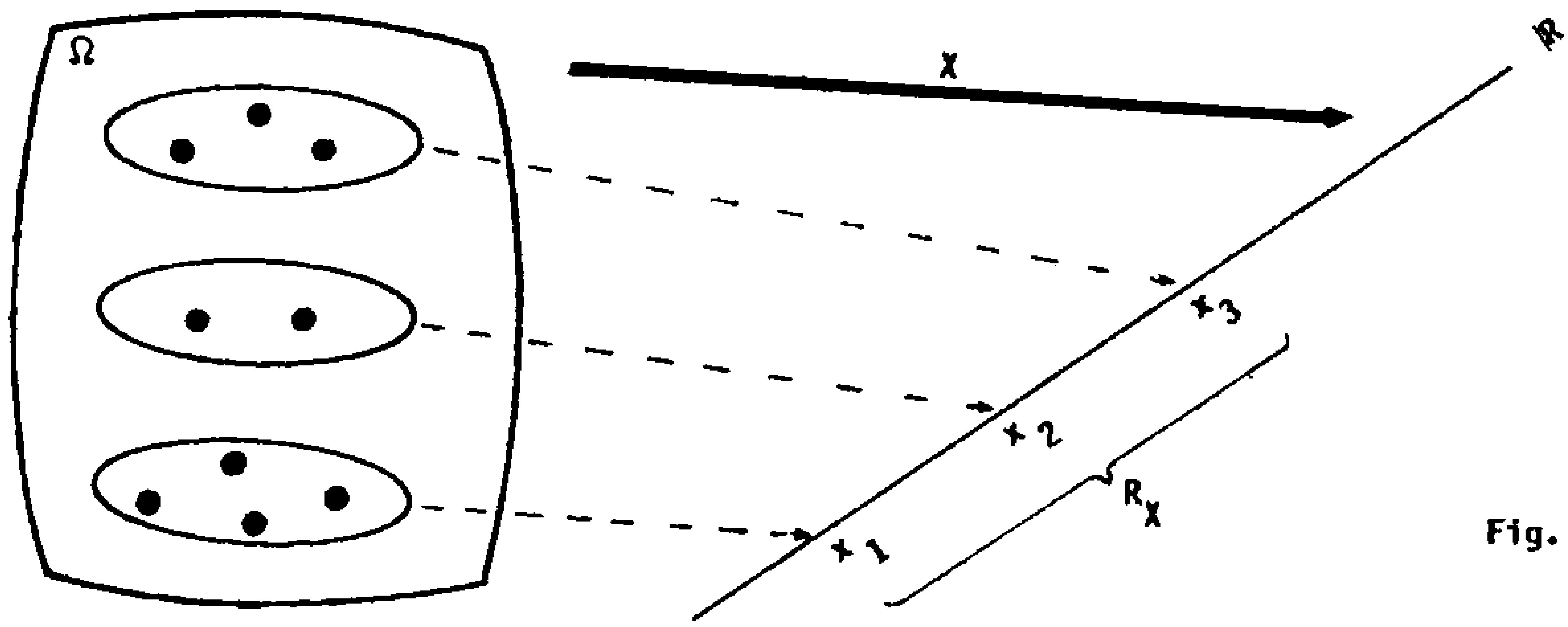


Fig. 2.1.2

El dominio de la variable aleatoria X es Ω y el rango es un subconjunto de $\mathbb{R}$ que lo denotaremos por "R_X". Rigurosamente, al hablar de función asociamos a ella el conjunto de partida y el conjunto de llegada, mas en nuestro caso vamos a trabajar siempre con Ω como dominio que a su vez lo vamos a tomar como conjunto de partida. El rango R_X de la variable aleatoria X está dado por el siguiente conjunto de números reales.

$$R_X = \{x \in \mathbb{R} \ / \ X(\omega) = x, \ \omega \in \Omega\} = X(\Omega)$$

Cuando hubiera cierta duda sobre el rango de una variable aleatoria, vamos a tomar como $\mathbb{R}$, o, como un conjunto que razonablemente contenga a R_X (ver. ejemplo 2).

EJEMPLO 2 Pensemos en los estudiantes de una universidad A, cada estudiante vamos a concebirlo como un suceso, a este suceso vamos a asignarle su altura así diremos que Juan mide 1.72 mts, Pedro mide 1.66 mts. Es decir, estamos cuantificado a los estudiantes X(Juan) = 1.72 mts, X(Pedro) = 1.66 mts. Aho ra el lector se estará preguntando, ¿y el rango de la variable aleatoria X?; en principio podemos decir que $\mathbb{R}$ contiene todo número que define la altura de un estudiante; luego podemos decir que el conjunto $R = \{x \in \mathbb{R} \ / \ x > 0\}$ - contiene a R_X. Después de razonar un poco en este problema podemos decir sin temor a equivocarnos que el conjunto $[0.5, 3]$ contiene a R_X; y llegamos a esta conclusión, pues con probabilidad cero vamos a encontrar un estudiante con altura mayor de 3 metros o menor que 0.5 metros (¿Ud. ya vió alguno?).

Es oportuno, antes de pasar a otro ejemplo, hacer, un comentario respecto al ejemplo anterior; este es típico en estadística y revela toda la libertad que tiene un estadístico al empezar un trabajo, más esta libertad tiene un precio, "ser cuidadoso".

EJEMPLO 3 Se lanza una moneda tres veces, sea X una función definida por - $X(\omega) = n_C - n_S$, donde n_C representa el número de caras y n_S, número de sellos obtenidos; X, así definido es una variable aleatoria. Sabemos que el espacio muestral (dominio de X) es

$$\Omega = \{CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS\}$$

Los correspondientes valores de X (imágenes de X) son:

$$\begin{aligned}
X(CCC) &= 3 - 0 = 3 \\
X(CCS) &= X(CSC) = X(SCC) = 2 - 1 = 1 \\
X(SSC) &= X(SCS) = X(CSS) = 1 - 2 = -1 \\
X(SSS) &= 0 - 3 = -3
\end{aligned}$$

Luego, es claro que R_X está exactamente definido por el conjunto,

$$R_X = \{-3, -1, 1, 3\} .$$

EJEMPLO 4 Un lote de artículos grande contiene artículos defectuosos D, y - no defectuosos N. Se extrae sucesivamente 4 artículos. Definimos X como número de artículos defectuosos obtenidos. La X, así definida es una variable - aleatoria; su dominio es

$$\Omega = \{DDDD, NDDD, DNDD, DDND, DDDN, DDNN, DNDN, DNND, NDND, NNDD, NDDN, DNNN, NDNN, NNDN, NNND, NNNN\}$$

Los valores de X son:

$$\begin{aligned}
&X(DDDD) = 4 \\
&X(NDDD) = X(DNDD) = X(DDND) = X(DDDN) = 3 \\
&X(DDNN) = X(DNDN) = X(DNND) = X(NDND) = X(NNDD) = X(NDDN) = 2
\end{aligned}$$

$X(NNND) = X(NNDN) = X(NDNN) = X(DNNN) = 1$

$X(NNNN) = 0$

Luego, $R_X = \{0,1,2,3,4\}$.

EJEMPLO 5 Bajo la misma suposición del ejemplo anterior, consideremos la extracción de artículos hasta lograr un artículo defectuoso y definimos **X** como el número necesario de extracciones. El dominio de **X** es

$$\Omega = \{D, ND, NND, NNND, \ldots\}$$

Las imágenes de **X** son :

$$X(D) = 1, \quad X(ND) = 2, \quad X(NND) = 3, \quad X(NNND) = 4, \ldots$$

Luego, $R_X = \{1,2,3,4, \ldots\}$.

EJEMPLO 6 Sea **X** una variable aleatoria que se considera como el beneficio de un jugador, en un juego en el que se tira un dado y el jugador gana 100 soles, si sale los números 1 ó 3, no gana ni pierde si sale los números 2 ó 5, pierde 100 soles si sale 4 ó 6.

El dominio de **X** es $\Omega = \{1,2,3,4,5,6\}$.
La imágenes de **X** son :

$$X(1) = X(3) = 100; \quad X(2) = X(5) = 0; \quad X(4) = X(6) = -100$$

Luego; $R_X = \{-100, 0, 100\}$.

Hemos visto que cada elemento del rango R_X de la variable aleatoria **X** tiene una probabilidad que han sido inducidas por las probabilidades asignadas a los posibles resultados del espacio muestral Ω a través de la función **X**, ver (*). Esto nos indica que podemos usar nuestra teoría de probabilidades desarrollada en el capítulo anterior, para calcular probabilidades en R_X; entonces así, como hablábamos de eventos, como subconjuntos de Ω, en R_X también hablaremos de eventos como subconjuntos de él. Y usaremos paréntesis para denotar eventos en R_X, así $(X(\omega) = x)$ o $[X(\omega) = x]$ o simplemente $[X = x]$ que se lee, "la variable aleatoria toma el valor x" y $P[X = x]$, denotará la probabilidad que la variable aleatoria toma el valor x. Formalizaremos esto con las siguientes definiciones.

DEFINICION 2.1.2 EVENTOS EQUIVALENTES Sea Ω un espacio muestral asociado a un experimento aleatorio ε, y **X** una variable aleatoria con rango R_X definida sobre Ω. Un evento A en Ω y un evento E_X en R_X se dice que son *eventos equivalentes*, Si,

$$A = \{\omega \in \Omega \ / \ X(\omega) \in E_X\}$$

La figura 2.1.3, ilustra este concepto.

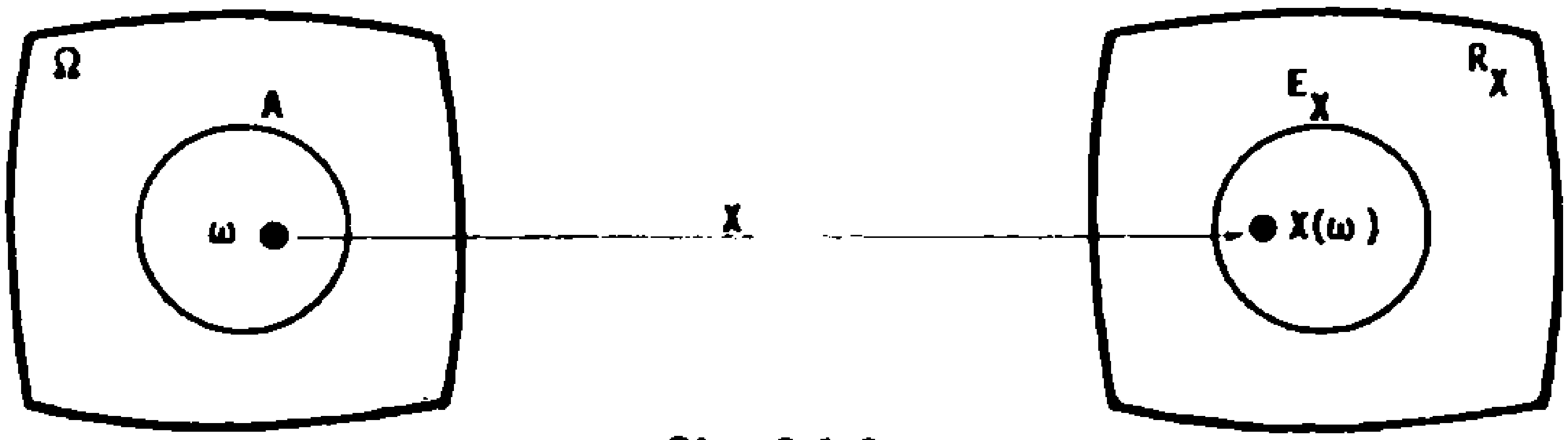


Fig. 2.1.3

Simplemente, si A es un evento en Ω que consiste de todos los resultados posibles para el cual $X(\omega) \in E_X$, entonces A y E_X son equivalentes. - Ahora, es claro que la ocurrencia de A implica la ocurrencia de E_X y viceversa, la ocurrencia de E_X implica la ocurrencia de A. Entonces, es natural definir la probabilidad de E_X como la probabilidad de A. Antes de formalizar - esto daremos algunas notaciones.

NOTA 1 Nótese que A y E_X son eventos asociados a diferentes espacios.

Si A es un evento en Ω tal que $A = \{\omega \in \Omega \ / \ X(\omega) = a\}$ su evento equivalente en R_X es $E_X = \{a\}$, lo cual se denota por $[X = a]$. O sea, el evento - $[X = a]$ es el conjunto de puntos en el espacio Ω que son aplicados en el número real a por la función X. Por ejemplo, en el experimento aleatorio "lanzar una moneda tres veces", y $X(\omega)$ = números de caras obtenidas.

Sea $A = \{\omega \in \Omega \ / \ X(\omega) = 2\} = \{CCS, CSC, SCC\} \equiv [X = 2]$, por otro lado si $E_X = \{1,0\}$, tenemos que,

$A = \{CSS, SCS, SSC, SSS\}$ ya que $X(CSS) = X(SCS) = X(SSC) = 1$ y $X(SSS)=0$.

Luego,

$A = \{\omega \in \Omega \ / \ X(\omega) = 1,0\} = \{CSS, SCS, SSC, SSS\}$ lo denotaremos por - $[X = 1 \text{ ó } 0]$.

En general si,

$A = \{\omega \in \Omega \ / \ X(\omega) = a \text{ ó } b\}$ se denotará por $[X = a, b]$ Similarmente, Si $A = \{\omega \in \Omega \ / \ a < X(\omega) < b\}$, se denotará por $[a < X < b]$.

Ahora si queremos hallar la probabilidad de los eventos asociados a R_X

tales como $[X = a]$, $[X = a, b]$, $[a < X < b]$ etc. usaremos las probabilidades de estos eventos en el espacio original Ω, es decir, pondremos,

$$P[X = a] = P[\{\omega \in \Omega \ / \ X(\omega) = a\}] = P[A]$$

que se lee: "la probabilidad que la variable aleatoria toma el valor a".

$$P[X = a, b] = P[\{\omega \in \Omega \ / \ X(\omega) = a \text{ ó } b\}] = P[A]$$

la probabilidad que la variable aleatoria toma el valor a ó b.

$$P[a < X < b] = P[\{\omega \in \Omega \ / \ a < X(\omega) < b\}] = P[A]$$

la probabilidad que la variable aleatoria toma valores entre a y b.
También se puede considerar eventos de la forma

$$[a \leq X \leq b] \qquad ; \qquad [a < X \leq b] \quad ; \quad \text{etc.}$$

DEFINICION 2.1.3 Si A es un evento en el espacio muestral Ω y E_X un evento en el rango R_X de la variable aleatoria X, entonces definimos la probabilidad E_X como

$$\boxed{P[E_X] = P[A] \text{, donde } A = \{\omega \in \Omega \ / \ X(\omega) \in E_X\}}$$

EJEMPLO 7 En el ejemplo 3 consideremos.

(a) $E_X = \{3\}$, entonces el evento A en Ω equivalente a E_X es, $A = \{CCC\}$.

$$A = \{\omega \in \Omega \ / \ X(\omega) = 3\} = \{CCC\} \equiv [X = 3].$$

Luego, $P[E_X] = P[X = 3] = \dfrac{1}{8}$

(b) $E_X = \{1\}$, aquí el evento equivalente es,

$$A = \{\omega \in \Omega \ / \ X(\omega) = 1\} = \{CCS, CSC, SCC\} = [X = 1]$$

Luego, $P[E_X] = P[X = 1] = P[A] = \dfrac{3}{8}$

(c) Consideremos el evento $E_X = \{2\}$. En este caso el evento equivalente a $\{2\}$ es ϕ, por lo tanto :

$$P[\{2\}] = P[X = 2] = P[\phi] = 0$$

(d) Consideremos el evento $E_X = \{-1, -3\}$. El evento equivalente es,

$$A = \{\omega \in \Omega \ / \ X(\omega) = -1, -3\} = \{SSC, SCS, CSS, SSS\} = [X = -1, -3]$$

Luego, $P[X = -1, -3] = P[A] = \dfrac{4}{8} = \dfrac{1}{2}$

(e) Consideremos $E_X = \{1,2\}$. El evento equivalente es,

$$A = \{\omega \in \Omega \;/\; X(\omega) = 1,2\} = \{CCS, CSC, SCC, \phi\} \equiv [X = 1,2]$$

Por lo tanto,

$$P[X = 1,2] = P[A] = \frac{3}{8}, \qquad \text{(ver parte c)}$$

EJEMPLO 8 Referimos al ejemplo 5, consideremos un lote grande de artículos que contiene el $100p\%$ de artículos defectuosos, calcular la probabilidad de extraer más de 3 artículos hasta obtener el primer defectuoso.
En este caso, $E_X = \{4,5,6, \ldots\}$ y el evento equivalente en Ω es,

$$A = \{NNND, NNNND, NNNNND, \ldots\}$$

Luego,

$$\begin{aligned} P[E_X] &= P[X > 3] = P[A] \\ &= 1 - P[\bar{A}] = 1 - P[X \leqslant 3] \\ &= 1 - [P[D] + P[ND] + P[NND]] \\ &= 1 - [p + (1-p)p + (1-p)^2 p] = 1 - p - p + p^2 - p \\ &\quad + 2p^2 - p^3 \\ &= (1-p)^3 . \end{aligned}$$

PROBLEMAS 2.1

1. Una urna contiene 12 bolas numeradas de 1 a 12. Se saca una bola y defina la variable aleatoria X tal que $X(\omega)$ = número de divisores del número obtenido. Hallar :
   (a) el dominio de X; (b) evaluar $X(\omega)$ para cada $\omega \in \Omega$;
   (c) escriba el rango de X ;
   (d) el evento equivalente en R_X a cada uno de los siguientes eventos en Ω

   $A = \{2,3,5,7,11\}$, $B = \{4,9\}$, $C = \{6,8,10\}$, $D = \{1,12\}$.

2. Se lanzan dos dados y sean i, j los números obtenidos ($i, j = 1,2,3,4,5,6$), se define la variable aleatoria $X(\omega) = m.c.d\ (i.j)$. Hallar :
   (a) el dominio de X ; (b) evaluar $X(\omega)$ para cada $\omega \in \Omega$;
   (c) el rango de la variable aleatoria X
   (d) el evento equivalente en Ω a cada uno de los siguientes eventos en R_X

   $E_X = \{2\}$, $F_X = \{2,4,5\}$, $G_X = \{1,3,6\}$

   (e) La probabilidad de los eventos E_X, F_X y G_X

3. Una factoría produce 10 paracaídas diario. Sea X, el número de paracaídas defectuosas.
   (a) ¿Es X una variable aleatoria?
   (b) Si su respuesta en (a) es, Sí. Halle el dominio y rango de la variable aleatoria X.

4. Una urna contiene 5 bolas numeradas de 1 a 5. Se extraen dos bolas sin reemplazamiento. Se define X como la suma de los números obtenidos. Determinar :
   (a) el dominio de X
   (b) el rango de la variable aleatoria X ;
   (c) $P[X = 3]$, $P[X = 5]$, $P[6 \leq X \leq 8]$.

5. Una caja contiene 5 transistores de radio, de las cuales dos son defectuosos. Los transistores se prueban uno a uno hasta encontrar el segundo transistor defectuoso. Sea X el número de pruebas efectuadas.
   (a) Describa el dominio de X ;
   (b) Describa el rango de la variable aleatoria X ;
   (c) ¿Cuál es el evento equivalente en Ω al evento $[X = 4]$?
   (d) ¿Cuál es el evento equivalente en Ω al evento $[X = 3]$?

6. Se venden 1,000 números para un sorteo en el que hay un premio mayor de I/. 500.00 , cuatro premios de I/. 100.00 y cinco premios de I/. 10.00 . El número cuesta I/. 1.00. Si X es el beneficio neto al comprar un número Hallar ;
   (a) el dominio de X ; (b) el rango de X ;
   (c) la probabilidad de cada uno de los elementos del rango de X.

2.2 VARIABLES ALEATORIAS DISCRETAS

DEFINICION 2.2.1 Si el rango de la variable aleatoria X, es un conjunto finito o infinito numerable, se llama *variable aleatoria discreta*. En este caso

$$R_X = \{x_1, x_2, x_3, \ldots\}$$

EJEMPLO 1 Suponga que el número de días de trabajo en un año particular es 280 y los records de los empleados se marcan cada día que ellos están ausente del trabajo. Se selecciona aleatoriamente un records y se observa los días marcados. La variable aleatoria X se define como el número de días ausentes del trabajo, entonces $R_X = \{0,1,2, \ldots , 280\}$. Luego, X es una variable -

aleatoria discreta con un número finito de posibles valores.

EJEMPLO 2 La variable aleatoria definida en el ejemplo 5 de 2.1, es una variable aleatoria discreta con un número infinito numerable de posibles valores.

2.2.1 FUNCION O LEY DE PROBABILIDAD

DEFINICION 2.2.2 Sea X una variable aleatoria discreta con rango R_X. Una función definida por

$$p(x) = P[X = x] = \sum_{\{\omega \in \Omega / X(\omega) = x\}} P[\{\omega\}]$$

donde la suma es sobre los sucesos $\omega \in \Omega$ tal que $X(\omega) = x$ y satisface las siguientes condiciones

1. $p(x) > 0 , \ \forall \ x \in R_X$; 2. $\sum_{x \in R_X} p(x) = \sum_{x \in R_X} P[X = x] = 1$

se llama *función de probabilidad o ley de probabilidad* (también llamada función de cuantía) de la variable aleatoria X.

La colección de pares $[(x, p(x)), \ \forall x \in R_X]$ se llama *distribución de probabilidad de* X.

Si $x \notin R_X$, $[X = x]$ es un evento imposible, por lo tanto $p(x) =$

$P[X = x] = 0$. Por esta razón cuando definimos una función de probabilidad $p(x)$, para $x \in R_X$, no diremos nada sobre la probabilidad en las $x \notin R_X$, pués entenderemos tácitamente que la función $p(x)$ está bién definida para todo $x \in R_X$ y asumiendo para los eventos imposibles $p(x) = 0$.

Con lo expresado en el párrafo anterior el dominio de la función p puede considerarse como el conjunto de los números, reales y su rango el conjunto $\langle 0, 1] \cup \{0\}$. Es decir,

$$p: \mathbb{R} \longrightarrow [0,1] .$$

La distribución de probabilidad se representa usualmente en una tabla, ver tabla 2.2.1. También se representa gráficamente como muestra la fig. 2.2.1

Representación tabular de la distribución de probabilidad

Tabla 2.2.1

x	x_1	x_2	x_3	.	.	.
$p(x) = P[X = x]$	$p(x_1)$	$p(x_2)$	$p(x_3)$	.	.	.

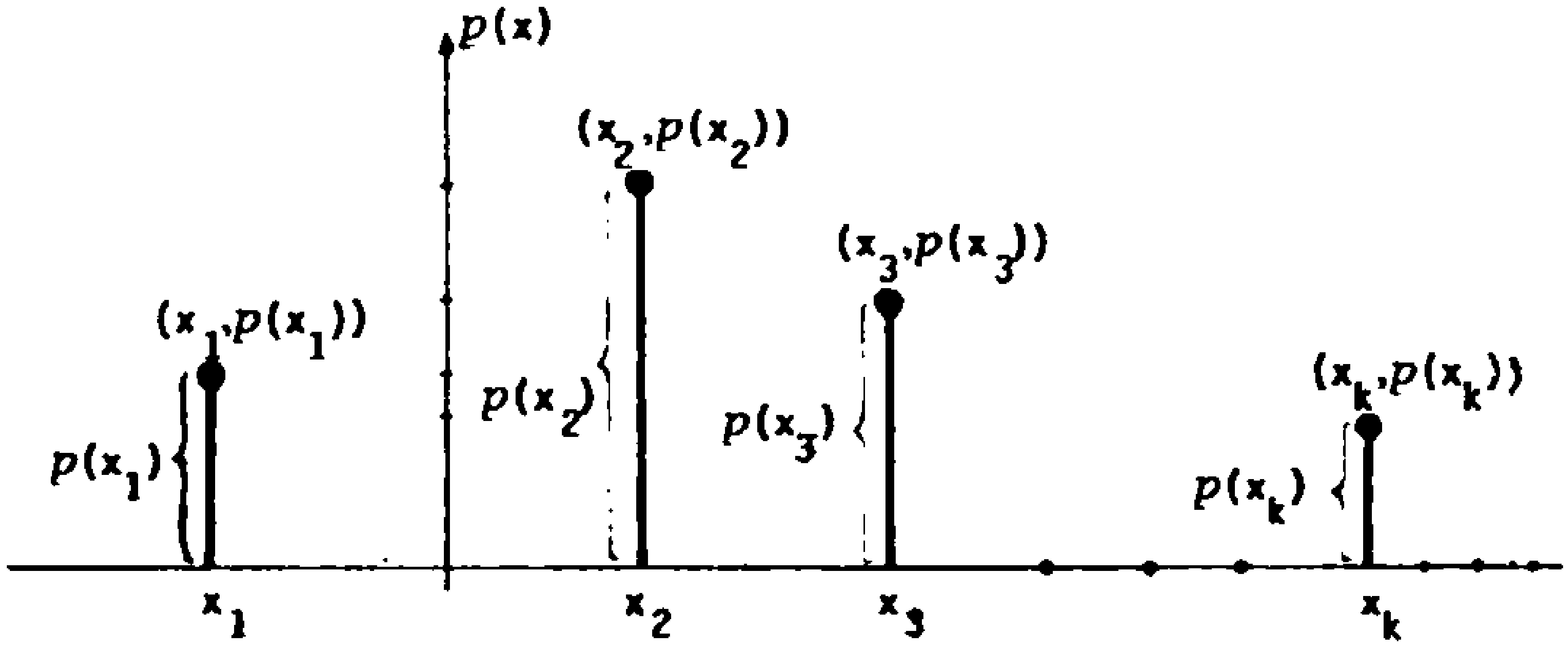


Fig. 2.2.1. Representación Gráfica de la distribución de probabilidad

Hemos denotado a los eventos en R_X por E_X, pero no habrá confusión alguna si se conviene en representar a estos eventos por A, B, C etc. Y la probabilidad de un evento A con R_X, se define de la siguiente manera

$$\boxed{P[A] = \sum_{x \in A} p(x) = \sum_{x \in A} P[X = x]} \qquad (I)$$

EJEMPLO 3 En el ejemplo 3 de 2.1, hemos considerado el lanzamiento de una moneda tres veces y definimos $X(\omega) = n_C - n_S$. Hallar la distribución de probabilidad en forma tabular y gráfica.

SOLUCION Recordamos que $R_X = \{-3, -1, 1, 3\}$, pués el espacio muestral es

$$\Omega = \{CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS\}$$

La distribución de probabilidad se obtiene, calculando $p(x)$ para cada $x \in R_X$

$$p(3) = P[X = 3] = P[CCC] = \frac{1}{8}$$

$$p(1) = P[X = 1] = P[CCS] + P[CSC] + P[SCC] = \frac{3}{8}$$

$$p(-1) = P[X = -1] = P[SSC] + P[SCS] + P[CSS] = \frac{3}{8}$$

$$p(-3) = P[X = -3] = P[SSS] = \frac{1}{8}$$

Luego, cada x con su respectiva $p(x)$ se lleva a una tabla similar a la tabla 2.2.1

Representación Tabular

x	-3	-1	1	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

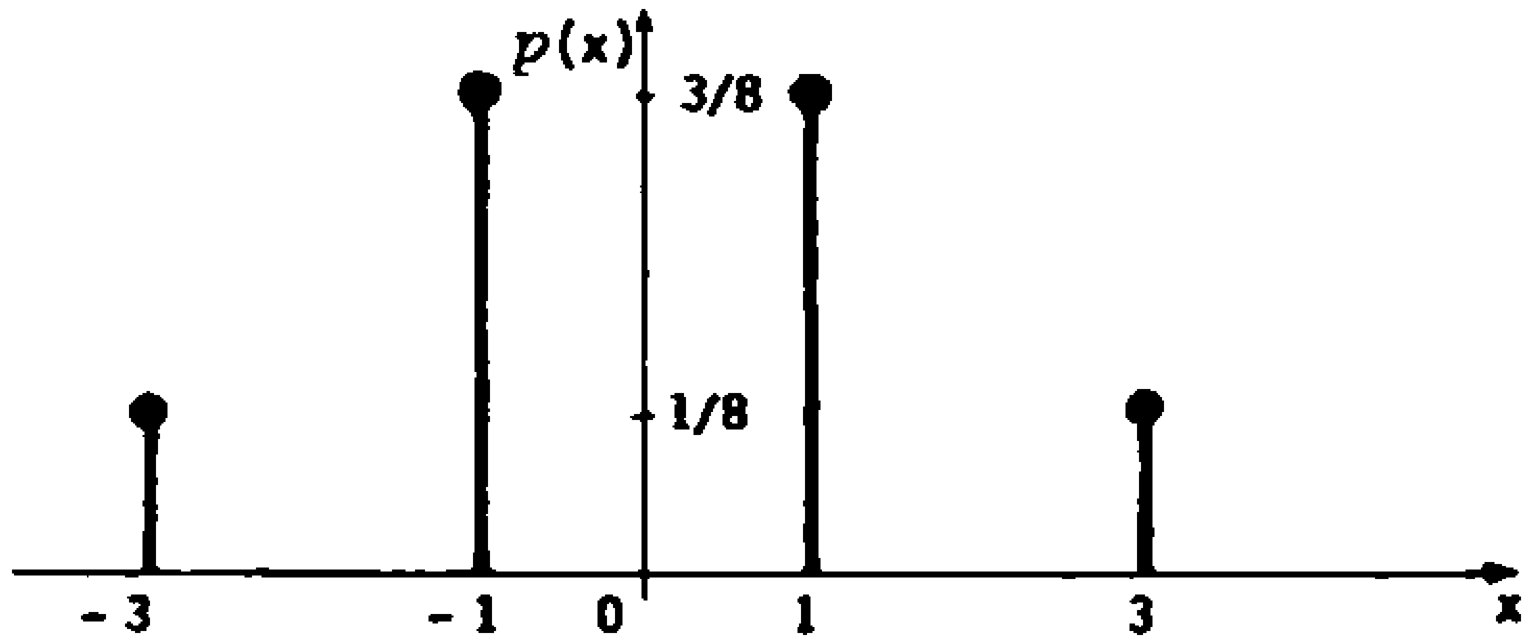


Fig. 2.2.2. Gráfico de la distribución de probabilidad para el experimento, lanzamiento de una moneda tres veces y $X = n_C - n_S$.

Es claro que ,

1. $p(x) > 0, \quad \forall\ x \in R_X$; 2. $\sum_{x \in R_X} p(x) = 1$

Pero para $x = -2$, tenemos $p(-2) = P[X = -2] = 0$, pues es imposible que la diferencia $n_C - n_S$ sea -2 en tres lanzamientos; similarmente, por ejemplo,

$$p(\tfrac{1}{2}) = P[X = \tfrac{1}{2}] = 0. \quad \text{En general si,}$$

$$x \notin \{-3, -1, 1, 3\} \text{ es } p(x) = P[X = x] = 0 .$$

EJEMPLO 4 Para cada uno de las siguientes funciones, determíne la constante k para que $f(x)$ sea una función de probabilidad de una variable aleatoria X .

(a) $f(x) = xk, \quad x = 1,2,3, \ldots, 10$

(b) $f(x) = k(\frac{1}{3})^x , \quad x = 1,2,3, \ldots$

SOLUCION (a) para que $f(x)$ sea una función de probabilidad debe cumplir la definición de ésta

1. $f(x) = k\,x > 0, \quad \forall\ x = 1,2, \ldots, 10$, si, solo si $k > 0$

2. $\sum_{x=1}^{10} kx = k [1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10] = 1$

de donde $k = 1/55$. Entonces,

$$f(x) = \frac{x}{55}, \quad x = 1,2, \ldots, 10$$

es una función de probabilidad .

(b) 1. $f(x) = k(\frac{1}{3})^x > 0, \ \forall\ x = 1,2, \ldots$, si, sólo si $k > 0$.

2. $$\sum_{x=1}^{\infty} k\left(\frac{1}{3}\right)^x = \frac{k}{3}\left[1 + \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \ldots\right]$$

$$= \frac{k}{3}\left[\frac{1}{1-\frac{1}{3}}\right] = k\left(\frac{1}{3}\right)\left(\frac{3}{2}\right) = 1$$

de donde, $k = 2$. Luego ,

$$f(x) = 2\left(\frac{1}{3}\right)^x , \quad x = 1,2,3, \ldots$$

es una función de probabilidad .

EJEMPLO 5 En un lote de 10 artículos, hay 3 artículos defectuosos. Del lote se toma al azar una muestra de cuatro artículos sin reposición. Sea **X** la variable aleatoria que representa el número de artículos defectuosos en la muestra.

(a) Describir el dominio de **X** .

(b) Evaluar $X(\omega)$ para cada $\omega \in \Omega$.

(c) Evaluar la función de probabilidad, la representación tabular y gráfica de la distribución de probabilidad.

SOLUCION La variable aleatoria **X** está definida por $X(\omega)$ = número de artículos defectuosos en la muestra de tamaño 4.

(a) Ω = {NNNN, NNND, NNDN, NDNN, DNNN, NNDD, NDND, NDDN, DNDN, DDNN, DNND, NDDD, DNDD, DDND, DDDN}

(b) $X(\text{NNNN}) = 0$

$X(\text{NNND}) = X(\text{NNND}) = X(\text{NNDN}) = X(\text{NDNN}) = X(\text{DNNN}) = 1$

$X(\text{NNDD}) = \cdot \quad \cdot \quad \cdot = X(\text{DNND}) = 2$

$X(\text{NDDD}) = \cdot \quad \cdot \quad \cdot = X(\text{DDDN}) = 3$

$R_X = \{0,1,2,3\}$

(c) La función de probabilidad se obtiene calculando $p(x)$ para cada $x \in R_X$

$$p(0) = P[X = 0] = \frac{\binom{3}{0}\binom{7}{4-0}}{\binom{10}{4}}$$

$$p(1) = P[X = 1] = \frac{\binom{3}{1}\binom{7}{4-1}}{\binom{10}{4}}$$

Representación Tabular

x	0	1	2	3
$p(x)$	$\frac{5}{30}$	$\frac{15}{30}$	$\frac{9}{30}$	$\frac{1}{30}$

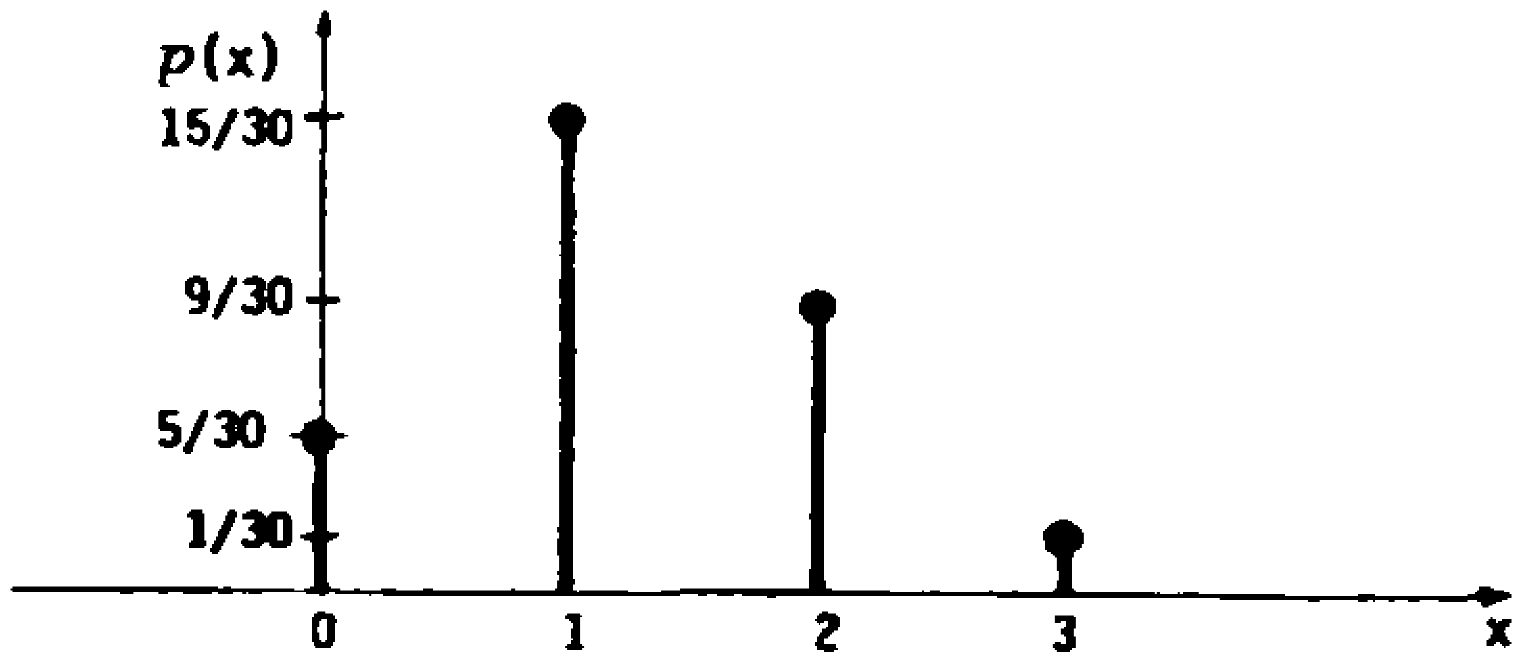


Fig. 2.2.3. Gráfico de la distribución de probabilidad para la muestra sin reposición de tamaño 4.

En general, la función de probabilidad de X es ,

$$p(x) = P[X = x] = \frac{\binom{3}{x}\binom{7}{4-x}}{\binom{10}{4}}, \quad x = 0,1,2,3.$$

Es claro que :

1. $\quad p(x) > 0 , \qquad \forall\ x \in R_X$

2. $\quad \sum_{x \in R_X} p(x) = \frac{5}{30} + \frac{15}{30} + \frac{9}{30} + \frac{1}{30} = 1$

Este ejemplo se puede generalizar de la siguiente manera.

EJEMPLO 6 Supóngase que se tiene n artículos (n finito) de los cuales r son defectuosos. Se extrae una muestra aleatorio de tamaño m sin reemplazamiento. Sea X el número de artículos defectuosos en la muestra. Hallar la función de probabilidad de la variable aleatoria **X**.

SOLUCION La variable aleatoria **X** está definida por

$X(\omega)$ = número de artículos defectuosos en la muestra de tamaño m

$$R_X = \{0,1,2,3, \ldots , \min(m, r)\}$$

$$p(x) = P[X = x] = \frac{\binom{r}{x}\binom{n-r}{m-x}}{\binom{n}{m}}, \quad x = 0,1,2, \ldots, \min(m,r)$$

es la función de probabilidad de X. Esta distribución se conoce como la ***distribución hipergeométrica***.

EJEMPLO 7 Se pone un ratoncito en un laberinto. Hay cinco caminos posibles, de los cuales sólo uno lleva fuera del laberinto. Supongamos que el ratoncito escoge un camino aleatoriamente hasta escoger el camino correcto; supongamos además que un camino incorrecto no se escoge dos veces. Sea **X** definido como el número de caminos incorrectos. Hallar :

(a) El dominio de **X** .

(b) El rango de la variable aleatoria **X**.

(c) La función de probabilidad asociado a **X** y su gráfica.

SOLUCION $X(\omega)$ = número de caminos incorrectos hasta encontrar el correcto.

Sea E = "Se escoge un camino correcto"

F = "Se escoge un camino incorrecto"

(a) El espacio muestral es

$$\Omega = \{E,\ FE,\ FFE,\ FFFE,\ FFFFE\}$$

(b) $R_X = \{0,1,2,3,4\}$.

(c) $p(0) = P[X = 0] = P[E] = \frac{1}{5}$

$p(1) = P[X = 1] = P[FE] = P[F]\,P[E\,|\,F] = \frac{4}{5} \times \frac{1}{4} = \frac{1}{5}$

$p(2) = P[X = 2] = P[FFE] = P[F]\,P[F\,|\,F]\,P[E|FF] = \frac{4}{5} \times \frac{3}{4} \times \frac{1}{3} = \frac{1}{5}$

$p(3) = P[X = 3] = P[FFFE] = P[F]\,P[F|F]\,P[F|FF]\,P[E|FFF]$

$= \frac{4}{5} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{5}$

$p(4) = P[X = 4] = P[FFFFE] = P[F]\,P[F|F]\,P[F|FF]\,P[F|FFF]$

$P[E|FFFF] = \frac{4}{5} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} \times 1 = \frac{1}{5}$

Luego, $p(x) = \frac{1}{5}$, $x = 0,1,2,3,4$.

es la función de probabilidad, llamado *distribución uniforme*

Representación Tabular

x	0	1	2	3	4
$p(x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

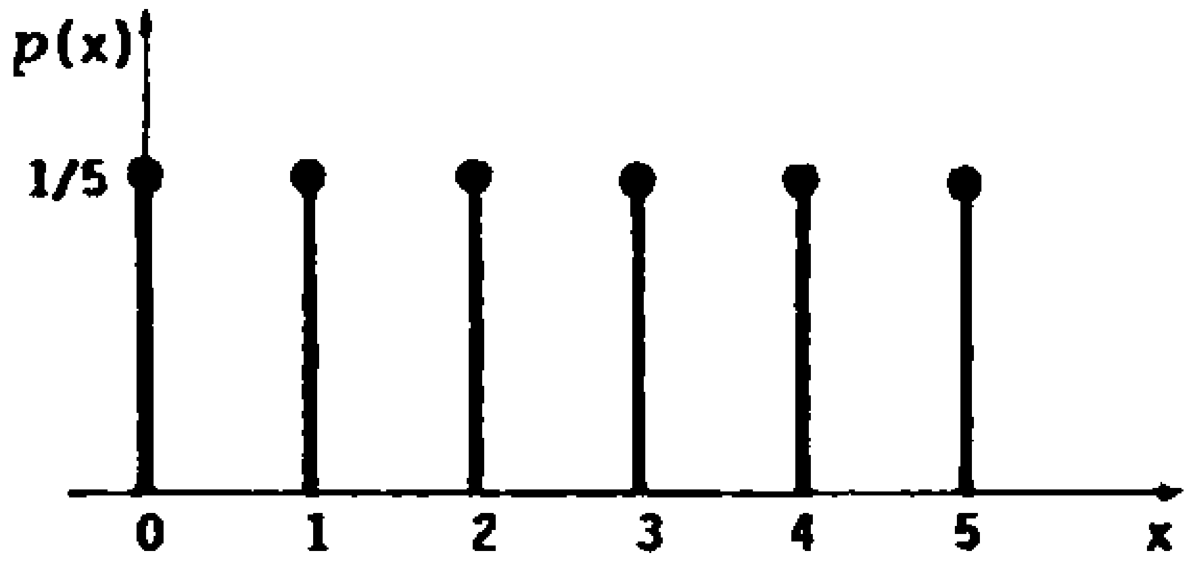


Fig. 2.2.4. Gráfico de la distribución de probabilidad para el experimento del ratoncito.

EJEMPLO 8 Se tiene 6 cajas numeradas 1,2,3,4,5 y 6; se tiene también 6 cartas numeradas 1,2,3,4,5 y 6. Se coloca al azar, (aleatoriamente) una carta en cada caja. Sea X la variable aleatoria que indica el lugar que ocupa la primera carta con número par. Determine Ud. la distribución de probabilidad de X. Represente sus respuestas en una tabla de la siguiente forma

x	$p(x)$

SOLUCION La variable aleatoria X está definida por

$X(\omega)$ = número de la caja que indica el lugar que ocupa la primera carta con número par.

Cartas {1,2,3,4,5,6} y {2,4,6} son pares.

$R_X = \{1,2,3,4\}$ $\quad \boxed{1}\ \boxed{2}\ \boxed{3}\ \boxed{4}\ \boxed{5}\ \boxed{6}$

Sea F_i el suceso : "en la caja i se colocó una carta impar, $i = 1,2,3$"

E_i el suceso : "en la caja i se colocó una carta par, $i = 1,2,3,4$".

Entonces el dominio de X es

$$\Omega = \{E_1,\ F_1E_2,\ F_1F_2E_3,\ F_1F_2F_3E_4\}\ .$$

$$p(1) = P[X = 1] = P[E_1] = \frac{3}{6} = \frac{1}{2}$$

$$p(2) = P[X = 2] = P[F_1E_2] = P[F_1]\,P[E_2|F_1] = \frac{3}{6} \times \frac{3}{5} = \frac{3}{10}$$

$$p(3) = P[X = 3] = P[F_1F_2F_3] = P[F_1]\,P[F_2|F_1]\,P[E_3|F_1F_2]$$

$$= \frac{3}{6} \times \frac{2}{5} \times \frac{3}{4} = \frac{3}{20}$$

$$p(4) = P[X = 4] = P[F_1F_2F_3E_4] = P[F_1]\,P[F_2|F_1]\,P[F_3|F_1F_2]\,P[E_4|F_1F_2F_3]$$

$$= \frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} \times \frac{3}{3} = \frac{1}{20}$$

La representación tabular es,

x	1	2	3	4
p(x)	1/2	3/10	3/20	1/20

EJEMPLO 9 Hallar la función de probabilidad de la variable aleatoria **X**, definida como el máximo de los números anotados en dos bolas extraídas con reemplazo (sin reemplazo) de una urna que contiene seis bolas numeradas del 1 al 6.

SOLUCION (a) Experimento con reemplazo, X está definida por

X(ω) = El máximo de los números anotados en dos bolas extraídas de la urna.

$$R_X = \{1,2,3,4,5,6\}$$

Como las extracciones es con reemplazo, entonces Ω tiene 6 x 6 = 36 elementos. Luego,

$$p(1) = P[X = 1] = \frac{1}{36}$$

$$p(2) = P[X = 2] = \frac{3}{36}$$

$$p(3) = P[X = 3] = \frac{5}{36}$$

....

	1	2	3	4	5	6
1	11	12	13	14	15	16
2	21	22	23	24	25	26
3	31	32	33	34	35	36
4	41	42	43	44	45	46
5	51	52	53	54	55	56
6	61	62	63	64	65	66

En general

$$p(x) = P[X = x] = \frac{2x - 1}{36}, \qquad x = 1,2,3,4,5,6$$

es la función de probabilidad de la variable aleatoria **X**.

(**b**) Sin reemplazamiento . Es decir, la variable aleatoria está definida

X(ω) = el máximo de los números anotados en dos bolas extraídas de la urna sin reemplazo.

$$R_X = \{2,3,4,5,6\}$$

En este caso Ω tiene 6 x 5 = 30 elementos. Luego,

$$p(2) = P[X = 2] = \frac{2}{30} = \frac{2(2 - 1)}{30} = \frac{2 - 1}{15}$$

$$p(3) = P[X = 3] = \frac{4}{30} = \frac{2(3 - 1)}{30} = \frac{3 - 1}{15}$$

$$p(4) = P[X = 4] = \frac{6}{30} = \frac{2(4 - 1)}{30} = \frac{4 - 1}{15}$$

....

En general se tiene que

$$p(x) = P[X = x] = \frac{x - 1}{15}, \qquad x = 2,3,4,5,6$$

es la función de probabilidad de la variable aleatoria X .

EJEMPLO 10 Se lanza un dado hasta que ocurra un número mayor que 2, encontrar la función de probabilidad del número necesario de lanzamientos.

SOLUCION Aquí la variable aleatoria X está definida por

$X(\omega)$ = el número necesario de lanzamientos hasta obtener un número mayor que 2.

$$R_X = \{1,2,3,4,5, \ldots\}$$

Ahora si E es el evento: "obtener un número mayor que 2 en un lanzamiento". El espacio muestral sería

$$\Omega = \{E, \bar{E}\ E, \bar{E}\ \bar{E}\ E, \bar{E}\bar{E}\bar{E}E, \ldots\}. \quad Y \quad P[E] = \frac{4}{6} = \frac{2}{3}, \quad \text{pues}$$

$$E = \{3,4,5,6\}$$

$$p(1) = P[X = 1] = P[E] = \frac{4}{6} = \frac{2}{3}$$

$$p(2) = P[X = 2] = P[\bar{E}\ E] = (\frac{2}{6})(\frac{2}{3}) = (\frac{1}{3})(\frac{2}{3})$$

$$p(3) = P[X = 3] = P[\bar{E}\ \bar{E}\ E] = (\frac{1}{3})^2 (\frac{2}{3})$$

En general

$$p(x) = P[X = x] = (\frac{1}{3})^{x-1} (\frac{2}{3}), \qquad x = 1,2,3,4, \ldots$$

la cual define la función de probabilidad de la variable aleatoria X; pues es evidente que:

(1) $$p(x) = \frac{2}{3} (\frac{1}{3})^{x-1} > 0 \qquad \forall\, x \in R_X$$

(2) $$\sum_{x-1}^{\infty} \frac{2}{3} (\frac{1}{3})^{x-1} = \frac{2}{3} \sum_{x=1}^{\infty} (\frac{1}{3})^{x-1} = \frac{2}{3} \sum_{x=0}^{\infty} (\frac{1}{3})^{x}$$

$$= \frac{2}{3} \left[1 + \frac{1}{3} + (\frac{1}{3})^2 + \ldots \right]$$

$$= \frac{2}{3} \left[\frac{1}{1 - \frac{1}{3}} \right] = \frac{2}{3} \times \frac{3}{2} = 1 .$$

EJEMPLO 11 Dos bombarderos lanzan alternativamente bombas al blanco hasta el primer impacto. La probabilidad de impacto en el blanco por el primer bombardero es igual a 0.7, y la del segundo, 0.8. La primera bomba la lanza el primer bombardero. Hallar la función de probabilidad de la variable aleatoria X, número de bombas lanzadas por ambos bombarderos.

SOLUCION $X(\omega)$ = número de bombas lanzadas por ambos bombarderos.

$R_X = \{1,2,3,4,\ .\ .\ .\}$. Sean los sucesos:

B_1 : "La bomba lanzada por el primer bombardero da en el blanco"

B_2 : la bomba lanzada por el segundo bombardero da en el blanco".

El espacio muestral Ω es

$$\Omega = \{B_1,\ \bar{B}_1 B_2,\ \bar{B}_1 \bar{B}_2 B_1, \bar{B}_1\bar{B}_2\bar{B}_1 B_2,\ \bar{B}_1\bar{B}_2\bar{B}_1\bar{B}_2 B_1, \bar{B}_1\bar{B}_2\bar{B}_1\bar{B}_2\bar{B}_1 B_2,\ \ldots\ \}$$

donde : $P[B_1] = 0.7$, $P[\bar{B}_1] = 0.3$

$P[B_2] = 0.8$, $P[\bar{B}_2] = 0.2$

$$\begin{aligned}
p(1) &= P[X = 1] = 0.7\\
p(2) &= P[X = 2] = P[\bar{B}_1 B_2] = (0.3)(0.8)\\
p(3) &= P[X = 3] = P[\bar{B}_1\bar{B}_2 B_1] = (0.3)(0.2)(0.7)\\
p(4) &= P[X = 4] = P[\bar{B}_1\bar{B}_2\bar{B}_1 B_2] = (0.3)^2(0.2)(0.8)\\
p(5) &= P[X = 5] = P[\bar{B}_1\bar{B}_2\bar{B}_1\bar{B}_2 B_1] = (0.3)^2(0.2)^2(0.7)\\
p(6) &= P[X = 6] = P[\bar{B}_1\bar{B}_2\bar{B}_1\bar{B}_2\bar{B}_1 B_2] = (0.3)^3(0.2)^2(0.8)\\
p(7) &= P[X = 7] = P[\bar{B}_1\bar{B}_2\bar{B}_1\bar{B}_2\bar{B}_1\bar{B}_2 B_1] = (0.3)^3(0.2)^3(0.7)
\end{aligned}$$

....

En general

$$p(x) = P[X = x] = \begin{cases} (0.3)^{\frac{x-1}{2}}\ (0.2)^{\frac{x-1}{2}}\ (0.7), & x = 1,3,5,7,\ .\ .\ .\\ (0.3)^{\frac{x}{2}}\ (0.2)^{\frac{x}{2}-1}\ (0.8)\ , & x = 2,4,6,8,\ .\ .\ . \end{cases}$$

es la función de probabilidad de la variable aleatoria X.

EJEMPLO 12 Hallar la función de probabilidad de la variable aleatoria X, número de cartas extraídas sin reemplazo, de una baraja de 52 cartas, hasta que (a) aparezca un trébol (b) aparezca un as.

SOLUCION (a) $X(\omega)$ = número de cartas extraídas sin reemplazo de una baraja de 52 hasta obtener un trébol.

$R_X = \{1,2,3,4,\ .\ .\ .\ ,\ 40\}$

Obtención de la función de probabilidad de X.

$$p(1) = P[X = 1] = \frac{13}{52} = \frac{\binom{39}{0}\binom{13}{1}}{\binom{52}{1}}$$

$$p(2) = P[X = 2] = \frac{39}{52}\cdot\frac{13}{51} = \frac{\binom{39}{2-1}\binom{13}{1}}{52 \times 51} = \frac{\binom{39}{2-1}}{\binom{52}{2-1}}\cdot\frac{13}{52-(2-1)}$$

$$p(3) = P[X = 3] = \frac{39}{52}\cdot\frac{38}{51}\cdot\frac{13}{50} = \frac{\binom{39}{3-1}}{\binom{52}{3-1}}\cdot\frac{13}{52-(3-1)}$$

$$p(4) = P[X = 4] = \frac{39}{52}\cdot\frac{38}{51}\cdot\frac{37}{50}\cdot\frac{13}{49} = \frac{\binom{39}{4-1}}{\binom{52}{4-1}}\cdot\frac{13}{52-(4-1)}$$

En general

$$p(x) = P[X = x] = \frac{\binom{39}{x-1}}{\binom{52}{x-1}}\left[\frac{13}{52-(x-1)}\right] \; ; \quad x = 1,2,3,\ldots,40.$$

es la función de probabilidad de la variable aleatoria X.

(b) $X(\omega)$ = número de cartas extraídas, hasta que aparezca un as.

$$R_X = \{1,2,3,4,\ldots,49\}.$$

Calculemos ahora la función de probabilidad de la variable aleatoria X .

$$p(1) = P[X = 1] = \frac{\binom{48}{0}}{\binom{52}{0}}\cdot\frac{4}{52}$$

$$p(2) = P[X = 2] = \frac{\binom{48}{2-1}}{\binom{52}{2-1}}\cdot\frac{4}{52-(2-1)}$$

$$p(3) = P[X = 3] = \frac{\binom{48}{3-1}}{\binom{52}{3-1}}\cdot\frac{4}{52-(3-1)}$$

En general

$$p(x) = P[X = x] = \frac{\binom{48}{x-1}}{\binom{52}{x-1}} \cdot \frac{4}{52 - (x-1)} \quad ; \; x = 1,2,3,4,\ldots,49$$

es la probabilidad de la variable aleatoria X .

EJEMPLO 13 La variable aleatoria X toma los valores 0,1,2,3, ... con probabilidad .

$$P[X = x_i] = \frac{c}{3^{x_i}} \quad ; \quad x_i = 0,1,2,3, \ldots$$

(a) Calcular el valor de la constante c .

(b) Cálcular la probabilidad que X tome un valor impar .

SOLUCION (a) por la segunda condición de la definición de función de probabilidad

$$\sum_{x_i = 0}^{\infty} \frac{c}{3^{x_i}} = c\left[1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \ldots\right] = c\left[\frac{1}{1 - \frac{1}{3}}\right] = 1$$

de donde $c = \frac{2}{3}$.

(b) Sea A, el evento: "la variable aleatoria X toma un valor impar".
La probabilidad del evento A se determina por la fórmula (I) .

$$\begin{aligned} P[A] = \sum_{x_i \text{ es impar}} P[X = x_i] &= \frac{2}{3}\left[\frac{1}{3} + \frac{1}{3^3} + \frac{1}{3^5} + \frac{1}{3^7} + \ldots\right] \\ &= \frac{2}{9}\left[1 + \frac{1}{3^2} + \frac{1}{3^4} + \frac{1}{3^6} + \ldots\right] \\ &= \frac{2}{9}\left[\frac{1}{1 - \frac{1}{9}}\right] = \frac{1}{4} . \end{aligned}$$

EJEMPLO 14 Una urna contiene 8 bolas blancas y 12 negras. Se extraen las bolas una a una, sin reposición, hasta que hayan aparecido 5 blancas. Hallar - función de probabilidad del número de extracciones.

SOLUCION Definimos la variable aleatoria X por

$X(\omega)$ = número de extracciones hasta obtener 5 blancas.

$$R_X = \{5,6,7,8,9,10,11, \ldots, 17\}$$

un esboso del espacio muestral es el siguiente

$$\Omega = \{BBBBB,\ \underbrace{BBBBN}_{P_5^{4,1}}\underbrace{B}_{\text{fijo}},\ \underbrace{BBBBNN}_{P_6^{4,2}}\underbrace{B}_{\text{fijo}},\ \dots\}$$

$$p(5) = P[X = 5] = P[BBBBB] = \frac{\binom{8}{5}}{\binom{20}{5}}$$

$p(6) = P[X = 6] = P[$4B en las 5 primeras extracciones y la última una B $] = P[$4B en las 5 primeras extracciones $]\ P[$una B en la sexta dado 4B$]$

$$= \frac{\binom{8}{4}\binom{12}{1}}{\binom{20}{5}} \cdot \frac{4}{15} = \frac{\binom{8}{4}\binom{12}{(6-1)-4}}{\binom{20}{6-1}} \cdot \frac{4}{20-(6-1)}$$

$p(7) = P[X = 7] = P[$4B en las 6 primeras extracciones y la última una B$] = P[$4B en los 6 primeras extracciones$]\ P[$una B en la septima dado 4B$]$

$$= \frac{\binom{8}{4}\binom{12}{2}}{\binom{20}{6}} \cdot \frac{4}{14} = \frac{\binom{8}{4}\binom{12}{(7-1)-4}}{\binom{20}{7-1}} \cdot \frac{4}{20-(7-1)}$$

En general se tiene que,

$p(x) = P[X = x] = P[$4B en las x-1 primeras extracciones, las demás negras y la última una B $] = P[$4B en las x-1 primeras extracciones, las demás negras$]\ P[$x-ésima una B dado que se tienen 4B$]$.

$$p(x) = P[X = x] = \frac{\binom{8}{4}\binom{12}{(x-1)-4}}{\binom{20}{x-1}} \left[\frac{4}{20-(x-1)}\right],\quad x = 5,6,\dots,17\,.$$

es la función de probabilidad de la variable aleatoria X .

EJEMPLO 15 Hallar la función de probabilidad de la variable aleatoria **X**, el número de veces que se lanza una moneda hasta obtener cara:

(*i*) por primera vez.

(*ii*) por segunda vez.

SOLUCION En cada caso la variable aleatoria X se define respectivamente por

(i) $X(\omega)$ = número de veces que se lanza una moneda hasta obtener una cara.

$$R_X = \{1,2,3, \ldots\}, \qquad \Omega = \{C, SC, SSC, \ldots\}$$

Obtención de la función de probabilidad de X .

$$p(1) = P[X = 1] = \frac{1}{2}$$

$$p(2) = P[X = 2] = \frac{1}{2} \times \frac{1}{2} = (\tfrac{1}{2})^2$$

$$p(3) = P[X = 3] = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = (\tfrac{1}{2})^3$$

En general

$$p(x) = P[X = x] = (\tfrac{1}{2})^x \quad , \quad x = 1,2,3, \ldots$$

es la función de probabilidad de X .

(ii) $X(\omega)$ = número de veces que se lanza una moneda hasta obtener 2 caras.

$$R_X = \{2,3,4,5, \ldots\}$$

$$\Omega = \{CC, \underbrace{SC}_{P_2^{1,1}}\ \underset{\text{fijo}}{\underset{\uparrow}{C}}, \underbrace{SSC}_{P_3^{2,1}}\ \underset{\text{fijo}}{\underset{\uparrow}{C}}, \underbrace{SSSC}_{P_4^{3,1}}\ \underset{\text{fijo}}{\underset{\uparrow}{C}}, \ldots\}$$

Obtendremos ahora la función de probabilidad de X .

$$p(2) = P[X = 2] = \frac{1}{2} \times \frac{1}{2} = (\tfrac{1}{2})^2 = \binom{1}{1}(\tfrac{1}{2})^2$$

$$p(3) = P[X = 3] = \binom{2}{1}(\tfrac{1}{2})^3$$

$$p(4) = P[X = 4] = \binom{3}{1}(\tfrac{1}{2})^4 \qquad \text{recuerde que} \quad \binom{3}{1} = \binom{3}{2}$$

$$p(5) = P[X = 5] = \binom{4}{1}(\tfrac{1}{2})^5 \qquad \text{y en general} \quad \binom{n}{1} = \binom{n}{n-1}$$

por lo tanto, en general se tiene

$$p(x) = P[X = x] = \binom{x-1}{1}(\tfrac{1}{2})^x \quad , \qquad x = 2,3,4, \ldots$$

es la función de probabilidad de X .

2.2.2 FUNCION DE DISTRIBUCION DE UNA VARIABLE ALEATORIA DISCRETA

Otro concepto importante en el desarrollo de los siguientes capítulos - es el de *función de distribución, o función de distribución acumulativa*, como se conoce algunas veces, debido a que se considera eventos de la forma:

$$[X \leq x] \text{ y su probabilidad inducida } P[X \leq x].$$

DEFINICION 2.2.3 Sea X una variable aleatoria discreta con rango $R_X = \{x_1, x_2, \ldots\}$ y función de probabilidad, $p(x_i) = P[X = x_i]$, sea x un número real cualquiera, *la Función de Distribución* de X se denota por - "F(x)" y se define como

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} p(x_i) = \sum_{x_i \leq x} P[X = x_i]$$

EJEMPLO 16 En el ejemplo 3. El lanzamiento de una moneda tres veces y - $X(\omega) = n_C - n_S$. Hemos calculado la distribución de probabilidad.

$$p(-3) = \frac{1}{8}, \quad p(-1) = \frac{3}{8}, \quad p(1) = \frac{3}{8} \quad \text{y} \quad p(3) = \frac{1}{8}$$

Calculemos ahora la función de distribución para esta variable aleatoria. En efecto, desde que F(x) está definida para todo $x \in \mathbb{R}$, consideremos los siguientes casos :

1. Si $x < -3$, es claro que

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} p(x_i) = 0 .$$

2. Si $x = -3$, $F(-3) = P[X \leq -3] = \sum_{x_i \leq -3} p(x_i) = p(-3) = \frac{1}{8}$.

3. Si $x \in [-3, -1\rangle$, $F(x) = P[X \leq x] = \sum_{x_i \leq x} p(x_i) = p(-3) = \frac{1}{8}$

4. Si $x = -1$, $F(-1) = P[X \leq -1] = \sum_{x_i \leq -1} p(x_i)$

$$= p(-3) + p(-1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8} .$$

5. Si $x \in [-1, 1\rangle$, $F(x) = P[X \leq x] = \sum_{x_i \leq x} p(x_i) = p(-3) + p(-1) = \frac{4}{8}$

6. Si x=1, $F(1) = P[X \leqslant 1] = \sum_{x_i \leqslant 1} p(x_i) = p(-3) + p(-1) + p(1) = \frac{4}{8} + \frac{3}{8} = \frac{7}{8}$.

7. Si $x \in [1,3\rangle$, $F(x) = P[X \leqslant x] = \sum_{x_i \leqslant x} p(x_i) = p(-3) + p(-1) + p(1) = \frac{7}{8}$.

8. Si x = 3, $F(3) = P[X \leqslant 3] = \sum_{x_i \leqslant 3} p(x_i) = p(-3) + p(-1) + p(1) + p(3)$

$$= \frac{7}{8} + \frac{1}{8} = 1 .$$

9. Si x > 3, $F(x) = P[X \leqslant x] = \sum_{x_i \leqslant x} p(x_i) = p(-3) + p(-1) + p(1) + p(3) = 1.$

NOTA 1 El lector habrá observado, que si $x \in [-3,-1\rangle$, entonces $F(x) = F(-3)$; Si $x \in [-1, 1\rangle$, $F(x) = F(-1)$. etc. En general, si

$$x \in [x_k, x_k + 1\rangle$$

se verifica que $F(x) = F(x_k)$, donde x_k, y x_{k+1} son elementos del rango - de X. (ver 2.2.3).

Luego, la función de distribución podemos escribir así,

$$F(x) = \begin{cases} 0 & , \text{ Si } \quad x < -3 \\ 1/8 & , \text{ Si } \quad -3 \leqslant x < -1 \\ 4/8 & , \text{ Si } \quad -1 \leqslant x < 1 \\ 7/8 & , \text{ Si } \quad 1 \leqslant x < 3 \\ 1 & , \text{ Si } \quad x \geqslant 3 \end{cases}$$

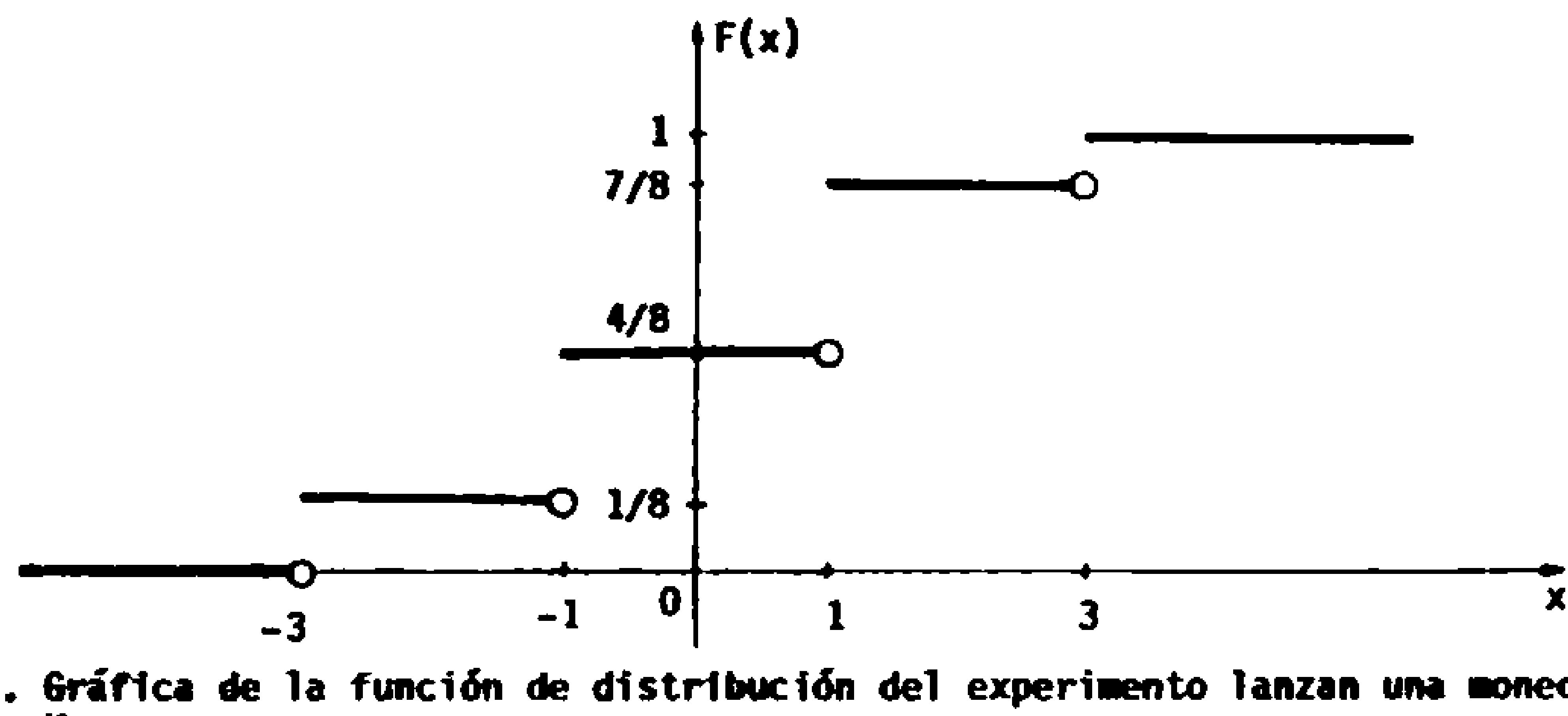


Fig. 2.2.5. Gráfica de la función de distribución del experimento lanzan una moneda tres veces, $X = n_c - n_s$

La función de distribución se representa también en una tabla como la siguiente,

Representación Tabular

x	-3	-1	1	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$F(x)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	1

EJEMPLO 17 En un lote de 8 artículos hay dos defectuosos. Del lote se toma al azar una muestra (sin restitución) de cuatro artículos. Sea X el número de artículos defectuosos en la muestra.

(a) Determine la función de probabilidad de X, construya su gráfica.

(b) Determine la función de distribución y construya su gráfica.

SOLUCION (a) X = número de artículos defectuosos en la muestra de 4.

$$R_X = \{0,1,2\} .$$

El espacio muestral tiene $\binom{8}{4}$ elementos . Por lo tanto

$$p(x) = \frac{\binom{2}{x}\binom{6}{4-x}}{\binom{8}{4}} , \quad x = 0,1,2 .$$

Representación Tabular

x	0	1	2
$p(x)$	$\frac{15}{70}$	$\frac{40}{70}$	$\frac{15}{70}$

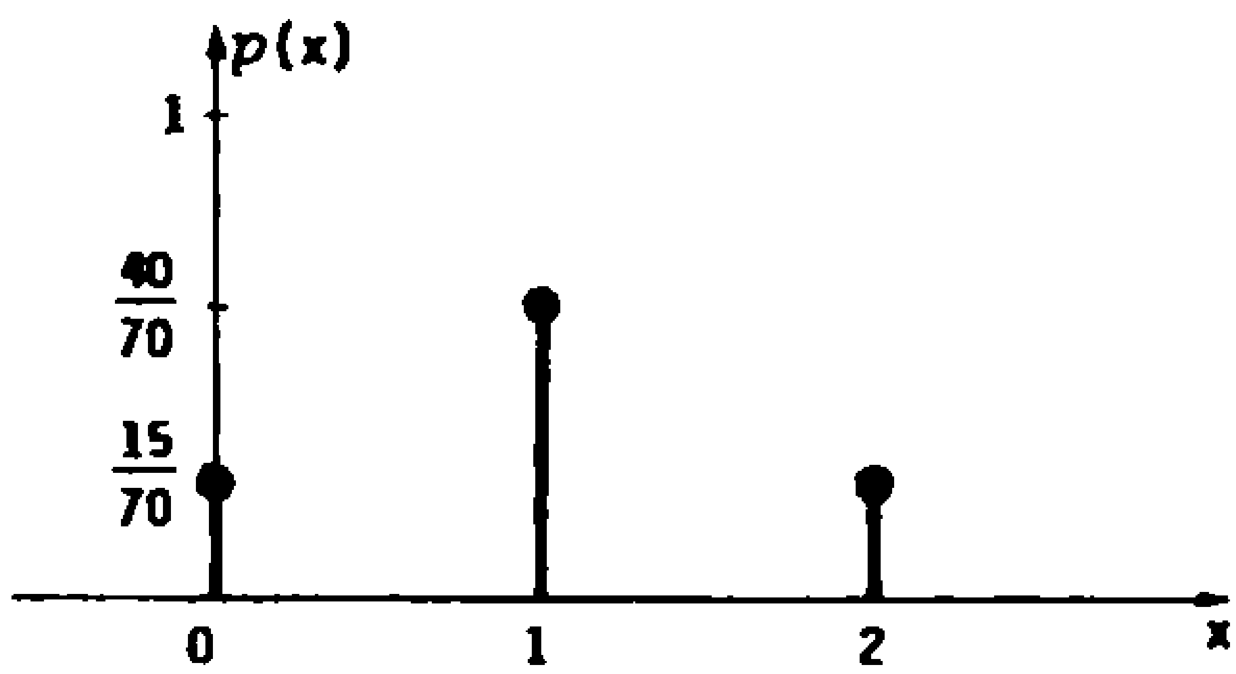


Fig. 2.2.6. Gráfica de la función de probabilidades.

(b) La función de distribución es,

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{15}{70} & , \quad 0 \leq x < 1 \\ \frac{55}{70} & , \quad 1 \leq x < 2 \\ 1 & , \quad 2 \geq x \end{cases}$$

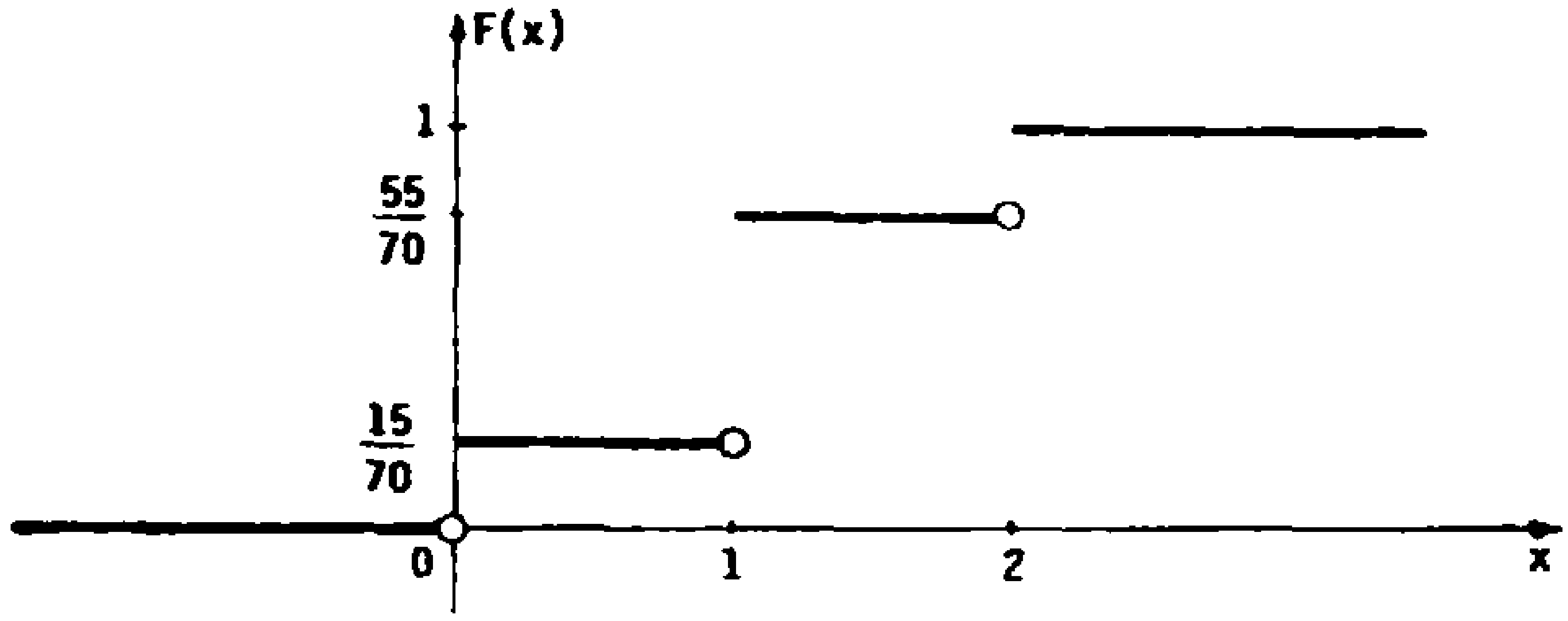


Fig. 2.2.7. Gráfica de la función de distribución

EJEMPLO 18 Sea $p(x) = P[X = x] = \frac{2}{3}(\frac{1}{3})^{x-1}$, $x = 1,2,3,\ldots$

Hallar la función de distribución $F(x)$.

SOLUCION

1. Si $x < 1$, es claro que $F(x) = 0$.

2. Si x es un número real mayor o igual que uno $(x \geq 1)$

$$F(x) = P[X \leq x] = \sum_{k=1}^{[\![x]\!]} \frac{2}{3}(\frac{1}{3})^{k-1}$$

$$= \frac{2}{3}\left[1 + \frac{1}{3} + \frac{1}{3^2} + \ldots + (\frac{1}{3})^{[\![x]\!]-1}\right]$$

$$= \frac{2}{3}\left[\frac{1 - (\frac{1}{3})^{[\![x]\!]-1}(\frac{1}{3})}{1 - \frac{1}{3}}\right]$$

$$= \frac{2}{3}\left[\frac{1 - (\frac{1}{3})^{[\![x]\!]}}{\frac{2}{3}}\right] = 1 - (\frac{1}{3})^{[\![x]\!]}$$

De (1) y (2) Se tiene que la función de distribución es,

$$F(x) = \begin{cases} 1 - (\frac{1}{3})^{[\![x]\!]} & , \quad x \geq 1 \\ 0 & , \quad x < 1 \end{cases}$$

EJEMPLO 19 Dada la función de probabilidad de la variable aleatoria X ; $p(x) = 2kx$, donde k es una constante y $x = 1,2,3,\ldots, n$. Hallar :

(a) el valor de k; (b) la función de distribución de **X**.

SOLUCION (a) Por la definición de función de probabilidad

$$\sum_{x=1}^{n} 2kx = 2k(1+2+3+\ldots+n) = 2k\left[\frac{n(n+1)}{2}\right] = 1$$

de donde $k = \dfrac{1}{n(n+1)}$

Por lo tanto, la función de probabilidad de X es

$$p(x) = \frac{2x}{n(n+1)}, \quad x = 1,2,3,\ldots,n.$$

(b) Desde que $F(x)$ está definida para todo $x \in \mathbb{R}$, consideremos :

1. Si $x < 1$, es claro que $F(x) = 0$.
2. Si $1 \le x < n$, entonces

$$F(x) = P[X \le x] = \sum_{x_i \le x < n} p(x_i)$$

$$= \sum_{x_i=1}^{[\![x]\!]} \frac{2x_i}{n(n+1)} = \frac{2}{n(n+1)}\left[1+2+\ldots+[\![x]\!]\right]$$

$$= \frac{2}{n(n+1)} \cdot \frac{[\![x]\!]\left([\![x]\!]+1\right)}{2} = \frac{[\![x]\!]\left([\![x]\!]+1\right)}{n(n+1)}.$$

3. Si $x \ge n$, es claro que $F(x) = 1$.

De (1), (2) y (3) obtenemos ,

$$F(x) = \begin{cases} 0 & , \quad x < 1 \\ \dfrac{[\![x]\!]\left([\![x]\!]+1\right)}{n(n+1)} & , \quad 1 \le x < n \\ 1 & , \quad x \ge n \end{cases}$$

EJEMPLO 20 Se lanza un dado dos veces, llamamos x al resultado del primer lanzamiento e y del 2do lanzamiento. Definimos la variable aleatoria **X** de la siguiente forma :

$$X(x,y) = \begin{cases} x - y & \text{Si} \quad x \geq y \\ x + y & \text{Si} \quad x < y \end{cases}$$

Determinar el rango de la variable aleatoria, las probabilidades asociadas, y la función de distribución asociada.

SOLUCION Para hallar el rango de la variable aleatoria, por facilidad construimos una tabla de doble entrada como sigue. El espacio muestral del experimento "lanzar un dado dos veces" es

$$\Omega = \{(x, y) \ / \ x, y = 1,2 \ldots , 6\}$$

La variable aleatoria

$$X : \Omega \longrightarrow \mathbb{R}$$

esta definida por

$$X(x,y) = \begin{cases} x - y \ , & x \geq y \\ x + y \ , & x < y \end{cases}$$

Luego,

$$X(1, 1) = 1 - 1 = 0$$
$$X(1,2) = 1 + 2 = 3$$

etc.

$y \backslash x$	1	2	3	4	5	6
1	0	1	2	3	4	5
2	3	0	1	2	3	4
3	4	5	0	1	2	3
4	5	6	7	0	1	2
5	6	7	8	9	0	1
6	7	8	9	10	11	0

De la tabla vemos que $R_X = X(\Omega) = \{0,1,2,3,4,5,6,7,8,9,10,11\}$.
El espacio muestral del experimento tiene 6 x 6 = 36 elementos. Entonces la probabilidad asociada a cada valor de X se obtiene contando los casos favorables en la tabla y dividiendo por 36. Los probabilidades asociadas y la fundición de distribución para cada uno de estos valores se da en la tabla siguiente

x	0	1	2	3	4	5	6	7	8	9	10	11
$P[X = x]$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
$F_X(x)$	$\frac{6}{36}$	$\frac{11}{36}$	$\frac{15}{36}$	$\frac{19}{36}$	$\frac{22}{36}$	$\frac{25}{36}$	$\frac{27}{36}$	$\frac{30}{36}$	$\frac{32}{36}$	$\frac{34}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

EJEMPLO 21 Se tiene una urna con 2 bolas negras, 3 bolas blancas y 4 rojas. Se extrae sucesivamente una bola sin reposición hasta que salga una roja. Hallar la distribución (función de cuantía), del número de extracciones que hay que realizar y la función de distribución (función de probabilidad acumulada) - correspondiente. Graficarlas.

SOLUCION La variable aleatoria **X** está definida por

$X(\omega)$ = número de extracciones hasta obtener una bola roja.

$$R_X = \{1,2,3,4,5,6\}$$

El espacio muestral asociado al experimento es

$$\Omega = \{R, \bar{R}\,R,\ \bar{R}\,\bar{R}\,R,\ \bar{R}\,\bar{R}\,\bar{R}\,R,\ \bar{R}\,\bar{R}\,\bar{R}\,\bar{R}\,R,\ \bar{R}\,\bar{R}\,\bar{R}\,\bar{R}\,\bar{R}\,R\}.$$

Entonces;

$$p(1) = P[X = 1] = P[R] = \frac{4}{9}$$

$$p(2) = P[X = 2] = P[\bar{R}\,R] = \frac{5}{9}\cdot\frac{4}{8} = \frac{5}{18}$$

$$p(3) = P[X = 3] = P[\bar{R}\,\bar{R}\,R] = \frac{5}{9}\cdot\frac{4}{8}\cdot\frac{4}{7} = \frac{10}{63}$$

$$p(4) = P[X = 4] = P[\bar{R}\,\bar{R}\,\bar{R}\,R] = \frac{5}{9}\cdot\frac{4}{8}\cdot\frac{3}{7}\cdot\frac{4}{6} = \frac{5}{63}$$

$$p(5) = P[X = 5] = P[\bar{R}\,\bar{R}\,\bar{R}\,\bar{R}\,R] = \frac{5}{9}\cdot\frac{4}{8}\cdot\frac{3}{7}\cdot\frac{2}{6}\cdot\frac{4}{5} = \frac{2}{63}$$

$$p(6) = P[X = 6] = P[\bar{R}\,\bar{R}\,\bar{R}\,\bar{R}\,\bar{R}\,R] = \frac{5}{9}\cdot\frac{4}{8}\cdot\frac{3}{7}\cdot\frac{2}{6}\cdot\frac{1}{5}\cdot 1 = \frac{1}{126}$$

La representación Tabular de la función de cuantía y la función de distribución es

x	1	2	3	4	5	6
$p(x)$	$\frac{4}{9}$	$\frac{5}{18}$	$\frac{10}{63}$	$\frac{5}{63}$	$\frac{2}{63}$	$\frac{1}{126}$
$F(x)$	$\frac{4}{9}$	$\frac{13}{18}$	$\frac{111}{126}$	$\frac{121}{126}$	$\frac{125}{126}$	1

F(x) se escribe también

$$F(x) = \begin{cases} 0 & , \quad x < 1 \\ 4/9 & , \quad 1 \leq x < 2 \\ 13/18 & , \quad 2 \leq x < 3 \\ 111/126 & , \quad 3 \leq x < 4 \\ 121/126 & , \quad 4 \leq x < 5 \\ 125/126 & , \quad 5 \leq x < 6 \\ 1 & , \quad x \geq 6 \end{cases}$$

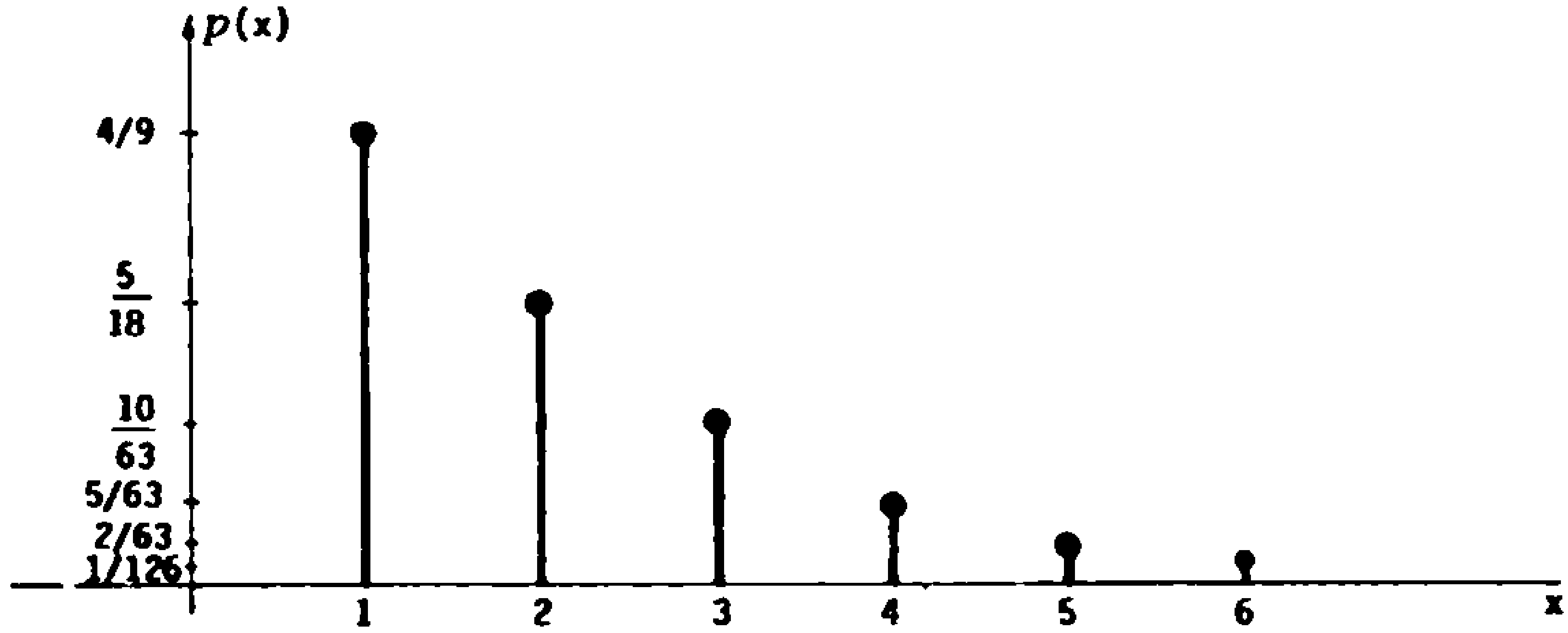


Fig. 2.2.8

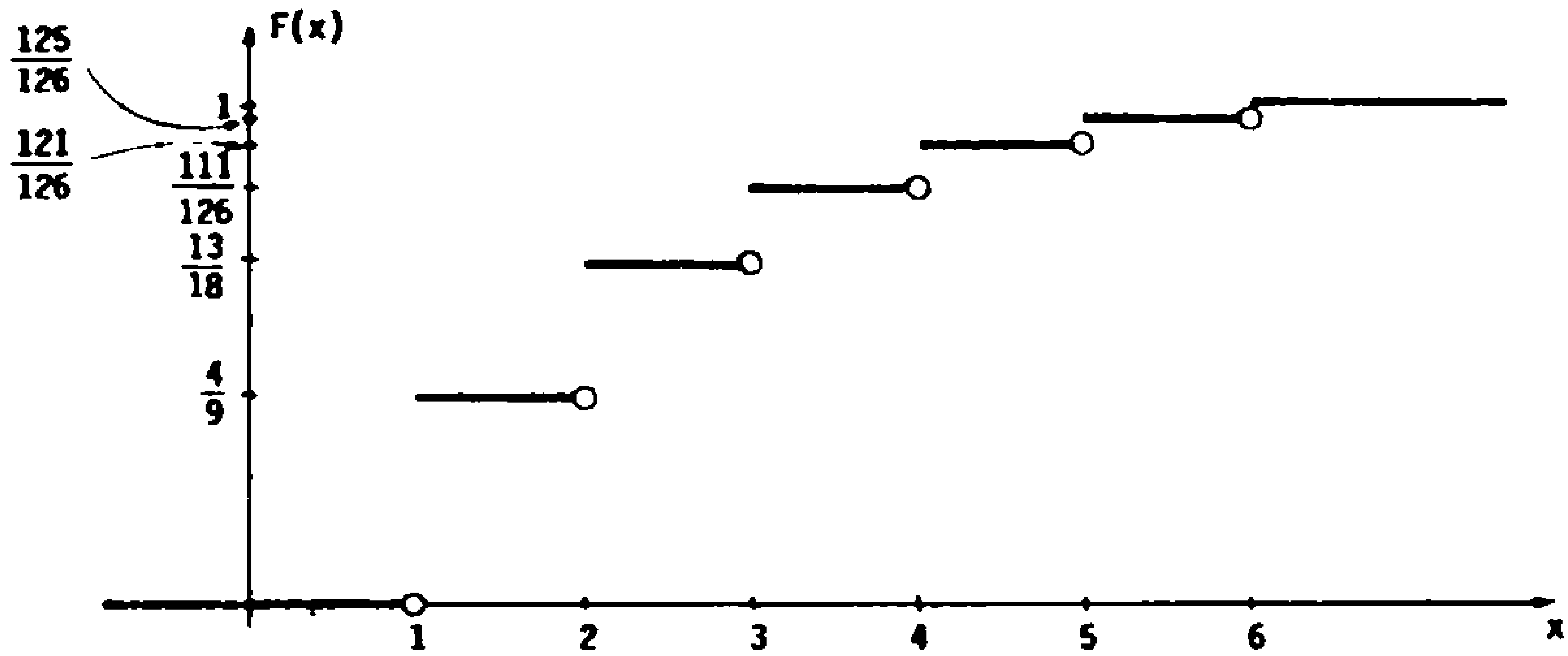


Fig. 2.2.9

2.2.3 PROPIEDADES DE LA FUNCION DE DISTRIBUCION

NOTACION Usaremos las siguientes notaciones :

$F(\infty)$ en vez de $\lim\limits_{x \to \infty} F(x)$ y $F(-\infty)$ en vez de $\lim\limits_{x \to -\infty} F(x)$

PROPIEDAD 1 $0 \leq F(x) \leq 1$, $\forall x \in \mathbb{R}$, pués F(x) es una probabilidad para cualquier x real y las probabilidades están limitadas por 0 y 1.

PROPIEDAD 2 $F(x)$ es una función no decreciente.

En efecto, sean $x_1, x_2 \in \mathbb{R}$ tales que $x_1 \leq x_2$, entonces se tiene

$$\{x \,/\, X \leq x_1\} \subset \{x \,/\, X \leq x_2\}$$

Aplicando probabilidades a ambos eventos, por el teorema 1.6.3

$$F(x_1) = P[X \leq x_1] \leq P[X \leq x_2] = F(x_2)$$

obtenemos

$$F(x_1) \leq F(x_2)$$

PROPIEDAD 3

(a) $F(\infty) = P[\{x \,/\, X < \infty\}] = P[X < \infty] = 1$, pués el evento $\{x \,/\, x < \infty\}$ es el conjunto de todos los números reales.

(b) $F(-\infty) = P[\{x \,/\, X < -\infty\}] = P[X < -\infty] = 0$, pués el evento $\{x \,/\, X < -\infty\}$, es el conjunto nulo.

PROPIEDAD 4 Sea $x_k, x_{k+1} \in R_X$, si x es tal que

$$x_k \leq x < x_{k+1}\,,\quad \text{entonces} \quad F(x) = F(x_k)$$

Es decir, la función $F(x)$ es constante e igual a $F(x_k)$ para todo $x \in [x_k, x_{k+1}\rangle$. Esto implica que si X es una variable discreta, $F(x)$ es una función "escalonada" (escalera), y la altura de un escalón en $x_k (x_k \in R_X)$ es igual a la $P[X = x_k]$ Ver figura 2.2.10.

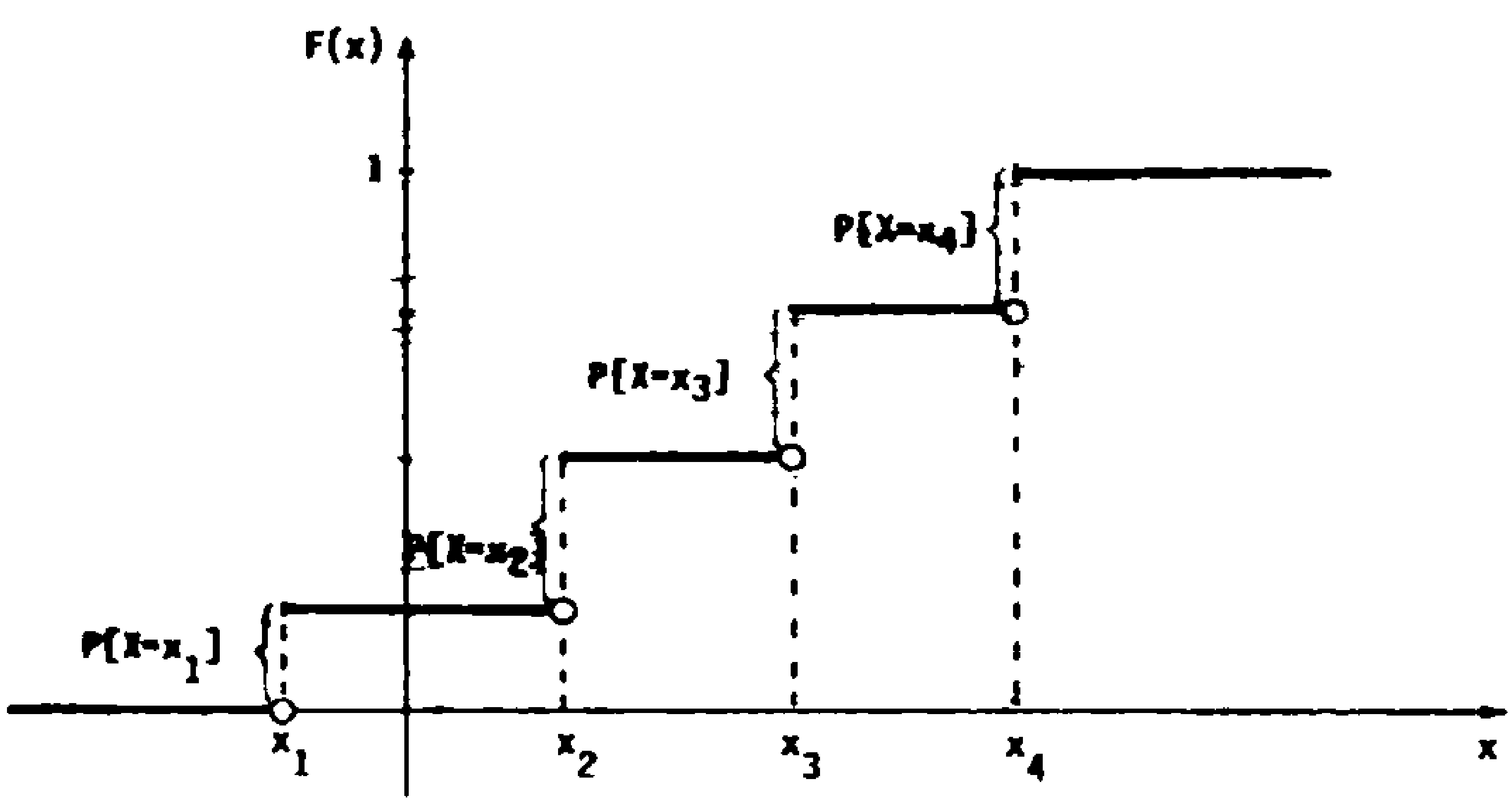


Fig. 2.2.10. Función de distribución

NOTA 2 (a) La función de distribución dá directamente

$$P[X \leq a] = F(a)$$

(b) $P[X \geq a] = 1 - P[X < a] = 1 - F(a - 1)$, Si a es entero y

$$= 1 - F(\llbracket a \rrbracket), \text{ si } a \text{ no es entero}$$

(c) La función de distribución F(x) se puede usar para determinar probabilidades de cualquier clase con relación a X; en particular consideremos el evento

$$(a < X \leq b), \quad a,b \in \mathbb{R} \text{ y } a < b$$

consideremos los siguientes eventos :

$$A = \{\omega \in \Omega \,/\, X(\omega) \leq a\} = (X \leq a) \quad .... \quad (1)$$

$$B = \{\omega \in \Omega \,/\, a < X(\omega) \leq b\} = (a < X \leq b)$$

desde que a y b son dos números reales cualesquiera tales, que $a < b$, entonces es claro que $A \cap B = \phi$.

Luego, $$P[A \cup B] = P[A] + P[B] \quad (2)$$

pero , $A \cup B = (X \leq a \text{ ó } a < X \leq b) = (X \leq b)$

por lo tanto (2) , se escribe

$$P[X \leq b] = P[X \leq a] + P[a < x \leq b] \quad \text{por (1)}$$

$$F(b) = F(a) + P[a < X \leq b]$$

de donde $$\boxed{P[a < X \leq b] = F(b) - F(a)} \quad (I)$$

También podemos calcular $P[a \leq X \leq b]$

$$P[a \leq X \leq b] = P[X = a \text{ ó } a < X \leq b], \quad \text{Si } a \in R_X .$$

desde que los eventos $(X = a)$ y $(a < X \leq b)$ son excluyentes, se tiene

$$P[a \leq X \leq b] = P[X = a] + P[a < x \leq b]$$

$$= P[X = a] + F(b) - F(a) \quad \text{por (I)}$$

por lo tanto,

$$\boxed{P[a \leq X \leq b] = F(b) - F(a) + P[X = a]} \quad (II)$$

Supongamos que estamos interesados en $P[a < X < b]$.

En efecto,

$$P[a < X < b] = P[a < X < b] - P[X = b], \quad \text{si } b \in R_X$$
$$= F(b) - F(a) - P[X = b]$$

Luego,

$$\boxed{P[a < X < b] = F(b) - F(a) - P[X = b]} \qquad \text{(III)}$$

También

$$\boxed{p(x_i) = P[X = x_i] = F(x_i) - F(x_i - 1)}$$

El lector puede fácilmente calcular la $P[a \leq X < b]$.

EJEMPLO 22 En el ejemplo 19. Calcular

(a) $P[9 \leq X \leq 20]$; (b) $P[X > 10.8]$.

SOLUCION

(a) $$P[9 \leq X \leq 20] = P[9 < X \leq 20] + P[X = 9]$$
$$= F(20) - F(9) + P[X = 9]$$
$$= \frac{20(21)}{n(n+1)} - \frac{9(10)}{n(n+1)} + \frac{2 \times 9}{n(n+1)}$$
$$= \frac{348}{n(n+1)} .$$

(b) $$P[X > 10.8] = 1 - P[X \leq 10.8] = 1 - F(10.8)$$
$$= 1 - \frac{[\![10.8]\!]\,([\![10.8]\!] + 1)}{n(n+1)}$$
$$= 1 - \frac{(10)(11)}{n(n+1)} = \frac{n(n+1) - 110}{n(n+1)} .$$

EJEMPLO 23 La variable aleatoria X tiene la siguiente función de distribución acumulativa

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 1/2 & 1 \leq x < 2 \\ 5/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Calcular : (a) $P[X(\omega) < 0.5] + P[X(\omega) \geq 2]$.

(b) la distribución de probabilidad de **X** .

SOLUCION

(a) $P[X(\omega) < 0.5] + P[X(\omega) \geq 2] = F(0.5) + 1 - P[X(\omega) < 2]$

$$= \frac{1}{8} + [1 - P[X(\omega) \leq 1]] = \frac{1}{8} + [1 - F(1)]$$

$$= \frac{1}{8} + [1 - \frac{1}{2}] = \frac{1}{8} + \frac{1}{2} = \frac{5}{8} .$$

(b) Teniendo en cuenta la propiedad 4 de la función de distribución, se tiene que el rango de la variable aleatoria **X** es el conjunto {0,1,2,3} y la distribución de probabilidad es

x	0	1	2	3
$p(x)$	1/8	3/8	1/8	3/8

PROBLEMAS 2.2

1. Para cada uno de las siguientes funciones, determine la constante k para que $p(x)$ satisfaga las condiciones de una función de probabilidad para una variable aleatoria **X**

   (a) $p(x) = \frac{x}{k}$, $x = 1,2,3,4$;

   (b) $p(x) = xk$, $x = 1,2,3, \ldots , 10,11,12$;

   (c) $p(x) = k(\frac{1}{5})^x$, $x = 0,1,2,3, \ldots$;

   (d) $p(x) = k(x+1)^2$, $x = 1,2,3.$

2. Consideremos el experimento de escoger un entero de 1 a 20 aleatoriamente Definimos la variable aleatoria **X** por: $X(\omega) = 0$ Si el entero escogido es divisible por 2, $X(\omega) = 1$. Si es divisible por 3 pero no por 2 y $X(\omega) = 5$ en otros casos. Determinar :

   (a) la función de probabilidad asociado a la variable aleatoria **X** y su gráfica.

   (b) la función de distribución asociado a **X** y su gráfica

   (c) la probabilidad de que $X \geq 3$ y $X > 0$

3. La variable aleatoria X representa el número de bolas blancas de una muestra de tamaño 2, extraída, sucesivamente y sin reposición de una urna que contiene 6 bolas de las cuales 4 son blancas.
   (a) Describir el dominio de X .
   (b) Evaluar $X(\omega)$ para cada $\omega \in \Omega$.
   (c) Evaluar y graficar la función de cuantía de X .
   (d) Evaluar y graficar la función de distribución de X.

4. En un lote de ocho artículos hay dos defectuosos. Del lote se toma al azar una muestra (sin restitución) de cuatro artículos. Sea X el número de artículos defectuosos de la muestra. Determinar :
   (a) el dominio de rango de X.
   (b) La función probabilidad de X y su gráfica.
   (c) La función de distribución de X y su gráfica.

5. Sea X, la variable aleatoria que describe el lanzamiento de un dado cargado. La distribución de probabilidad de X está dado por :

$$p(1) = p(2) = \frac{1}{6} \; ; \quad p(3) = \frac{1}{12} \; ; \quad p(4) = p(5) = \frac{1}{4} \; ; \quad p(6) = d$$

   (a) Determine el valor de d .
   (b) Calcule la función de distribución en $x = 3.6$, o sea $F(3.6)$
   (c) Hallar $P[3 \leq X < 5]$.

6. Determine la distribución de probabilidad de la variable aleatoria X, definida en el problema 2.1.6 .

7. Durante el curso de un día, una máquina produce tres artículos, cuya calidad individual, definida como defectuosos o no defectuoso, se determina al final del día. Sea X la variable aleatoria que representa al número de unidades defectuosos. Suponga que cada punto del espacio muestral tiene igual probabilidad. Determinar :
   (a) el dominio de X .
   (b) la distribución de probabilidad de X y su gráfica.
   (c) la función de distribución de X y su gráfica.

8. En el problema anterior (7), suponga que un artículo defectuoso representa una pérdida de \$, 250.00. Sea X la variable aleatoria que representa la utilidad total diaria. Suponiendo que cada punto del espacio muestral tiene igual probabilidad, hallar la distribución la probabilidad para X . Un artículo no defectuoso representa una ganancia de \$ 1000.

9. Una urna contiene 10 bolas numeradas 1,2,3, . . . , 10 respectivamente. - Sea X el número que se obtiene al extraer al azar una bola de la urna. Hallar :
   (a) la función de probabilidad de X y su gráfica
   (b) la función de distribución de X y su gráfica
   (c) generalice el problema a una urna con n bolas numeradas de $1,2,\ldots,n$.

10. Sea la urna definida en el problema 9. Se extraen dos bolas sin reemplazamiento, y sea X la suma de los números que se obtienen. Hallar la distribución de probabilidad de X. Bosquejar su gráfica. La función de distribución - de X.

11. Suponer que en el problema 10, se extraen las bolas con reemplazo y define - la variable aleatoria X como la suma de los números que se obtienen. Hallar la función de probabilidad de X y bosquejar su gráfica.

12. Se tiene una urna con 3 fichas negras y 2 rojas. Se extraen sucesivamente una ficha sin reposición hasta que salga una roja. Sea X el número de extracciónes que hay que realizar. Determinar :
   (a) Los valores que puede tomar la variable y sus probabilidades asociadas
   (b) la función de distribución acumulativa y diseñar su gráfica.

13. Hay 10 estudiantes inscritas en una clase de estadística, de entre los - cuales 3 tiene 19 años, 4 tienen 20, 1 tiene 21, 1 tiene 24 y 1 tiene 26 años. De esta clase se seleccionan dos estudiantes al azar sin reemplazamiento. Sea X la edad promedio de los 2 estudiantes seleccionados. Hallar la distribución de probabilidad de X y su función de distribución.

14. Un hombre tiene cuatro llaves en su llavero. Como está oscuro, no puede - ver cuál es la llave de la puerta de su casa, por lo que prueba cada uno, hasta encontrar la correcta. Sea X el número de llaves que se prueba (incluyendo la correcta) para abrir la puerta. ¿Cuál es la distribución de - probabilidad de X? ¿Cuál es la función de distribución de X? .

15. Se tiene una urna con 3 fichas, numeradas de 1 a 3. Se extrae al azar una ficha, luego se lanza una moneda tantas veces como indica el número de la ficha obtenida. Si X representa el número de caras, Determinar:
   (a) el dominio de X
   (b) el rango y las probabilidades asociadas.
   (c) la distribución de probabilidad y su gráfica.
   (d) la función de distribución y su gráfica.

16. Una urna contiene 8 bolas númeradas de 1 a 8. Se extrae dos bolas sucesivamente y con reposición, X representa el mínimo entre los dos números anotados en las bolas extraídas. Determinar :
(a) el dominio de X
(b) el rango y la función de cuantía de X.
(c) la función de distribución de X.
(d) los elementos de los eventos $[X \leq 2.5]$, $[X \leq 5.75]$, $[X > 6.90]$
(e) $P[X \leq 5.75]$, $P[1.75 \leq X \leq 6.75]$

17. En determinada ciudad 1/3 de las familias no tienen automóvil, 1/3 tiene uno, 1/6 tiene dos, 1/12 tiene tres y 1/12 tiene cuatro automóviles. Cada automóvil tiene 5 llantas. Sea X la variable aleatoria que representa el número de llantas por familia. ¿Cuál es
(a) el dominio de X ?
(b) el rango de X ?
(c) la distribución de probabilidad de X ? gráficar
(d) la función de distribución de X ? gráficar

18. Suponga que al pasar un neutrón a través del plutonio pueda con igual probabilidad dejar libre 1,2 ó 3 neutrones y que esta segunda generación de neutrones pueda a su vez, con igual probabilidad dejar libre 1,2,ó 3 neutrones de la tercera generación. ¿Cuál es la función de probabilidad del número de neutrones de la tercera generación?

19. La variable aleatoria X tiene función de probabilidad definida por

$$P[X = x] = \frac{k}{5^x} , \quad x = 0,1,2, \ldots$$

(a) Calcular el valor de la constante k
(b) calcular la probabilidad que X tome un valor par.

20. Se reparten cinco naipes de una baraja de 52 cartas. Sea X el número de naipes de color rojo repartidos. Determinar :
(a) La función de probabilidad de X .
(b) la función de distribución para X.

21. Determine la función de probabilidad de la variable aleatoria X, el número de ases en una mano de 13 cartas extraídas sin reemplazo de una baraja de 52 cartas.

22. Hallar la función de probabilidad de la variable aleatoria X definida co-

mo el máximo de los números anotados en dos bolas extraídas con reemplazo (sin reemplazo) de una urna que contiene 8 bolas numeradas de 1 al 8.

23. En un lote grande de artículos, el 10 % son defectuosos. Se escogen al azar 4 artículos. Escribir la función de probabilidad de la variable aleatoria X, número de artículos defectuosos entre los cuatro escogidos y construir su gráfica.

24. Dos cañones tiran al blanco alternativamente hasta el primer impacto por uno de los cañones. La probabilidad de impacto en el blanco por el primer cañón es igual a 0.30, y por el segundo, 0.70. Comienza a tirar el primer cañón. Hallar la función de probabilidad de la variable aleatoria X e Y, números de proyectiles lanzados por los cañones, primero y segundo respectivamente.

25. Se lanza una serie de cohetes hasta que el primer lanzamiento exitoso tiene lugar. Si esto no ocurre en 5 ensayos, el experimento se detiene y se inspecciona el equipo. Supóngase que hay una probabilidad constante de 0.80 de tener un lanzamiento exitoso y que los ensayos son independientes Además, el costo del primer lanzamiento es k dólares, mientras que los lanzamientos que siguen cuestan k/3 dólares. Cada vez que hay un lanzamiento exitoso, se obtiene cierta cantidad de información que puede expresarse como una ganancia financiera de C dólares. Si X es costoso neto del experimento, hallar la distribución de probabilidad de X.

26. La urna I contiene 1 ficha blanca y 2 negras, la urna II contiene 3 fichas blancas y 2 negras, la urna III contiene 2 fichas blancas y 3 negras Extraemos una ficha de cada urna y llamamos X a la variable aleatoria que representa el número de fichas blancas extraídas. Determinar la función de cuantía de la variable aleatoria X(los valores que toma y las probabilidades asociadas).

27. En una caja hay 8 fusibles buenos y 15 defectuosos. Se necesitan 5 fusibles buenos. Se extraen los fusibles uno a uno sin reposición y se va probando, hasta que se obtenga los 5 fusibles buenos. Determinar la función de probabilidad del número de extracciones.

28. En el problema 1c, Calcular la probabilidad que la variable aleatoria tome un valor impar.

29. La variable aleatoria X tiene la siguiente función de distribución acumulada

$$F(x) = \begin{cases} 0 & x < 10 \\ 1/4 \quad , & 10 \leq x < 15 \\ 3/4 \quad , & 15 \leq x < 20 \\ 1 \quad , & x \geq 20 \end{cases}$$

1. (a) Calcular $P[X \leq 10.5] + P[X \geq 15.5]$.

   (b) Calcular $P[10.2 \leq X \leq 15.5]$.

   (c) Calcular la distribución de probabilidad de la variable aleatoria.

30. La función de distribución de la variable aleatoria X ésta dado por

$$F(x) = \begin{cases} 0 \quad , & x < -1 \\ 0.5 \quad , & -1 \leq x < 0 \\ 0.8 \quad , & 0 \leq x < 2 \\ 1 \quad , & x \geq 2 \end{cases}$$

Determinar :

(a) $P[X \leq -0.5] + P[X \geq 1.5]$

(b) la distribución de probabilidad (función de cuantia) de X .

31. Hallar la función de probabilidad de la variable aleatoria X, número de veces que se lanza una moneda hasta obtener cara por tercera vez.

32. Un jugador que tiene I/. 70, juega con un dado, con la siguiente regla: En la primera tirada, apuesta I/. 10 a los números pares y desiste si gana. Si pierde, apuesta I/. 20 a los números pares en la segunda tirada y desiste si gana. Si pierde de nuevo, apuesta sus I/. 40 finales a los números pares en la tercera tirada. El juego con un dado duplica la apuesta al ganador. Hallar la distribución de probabilidad de la ganancia neta del jugador.

2.3 VARIABLES ALEATORIAS CONTINUAS

DEFINICION 2.3.1 Si el rango R_X, de una variable aleatoria X es un intervalos sobre la recta de los números reales, se llama *variable aleatoria continua*.

EJEMPLO 1 La variable aleatoria definida en el ejemplo 2 de 2.1 es una variable aleatoria continua, pues toma todos los valores de un intervalo, por ejemplo [0.5,3] .

EJEMPLO 2 Sea X la variable aleatoria que representa el número de kilogramos que pierde una persona al seguir una dieta específica durante cierto período. Es una variable aleatoria contínua, pues su rango (los valores que puede tomar)son todos los puntos de un intervalo, por ejemplo [1,3] .

2.3.1 FUNCION DE DENSIDAD DE PROBABILIDAD

DEFINICION 2.3.2 Sea X una variable aleatoria contínua con rango R_X. *La función de densidad de probabilidad* asociado a la variable aleatoria, es una función $f(x)$ integrable que satisface las siguientes condiciones :

1. $f(x) \geq 0$, para todo $x \in \mathbb{R}$ (o $f(x) > 0$, $x \in R_X$)

2. $$\int_{R_X} f(x)dx = 1 \qquad (\text{ó} \int_{-\infty}^{+\infty} f(x)dx = 1)$$

Esta definición indica, la existencia de una función real f definida sobre R_X. La condición (1) establece que la gráfica de la función de densidad está "arriba" del eje X. La condición (2) indica que el área acotada por la curva $f(x)$, el eje X y las rectas verticales que pasan por los puntos extremos de R_X es 1, como se indica en la figura 2.3.1 .

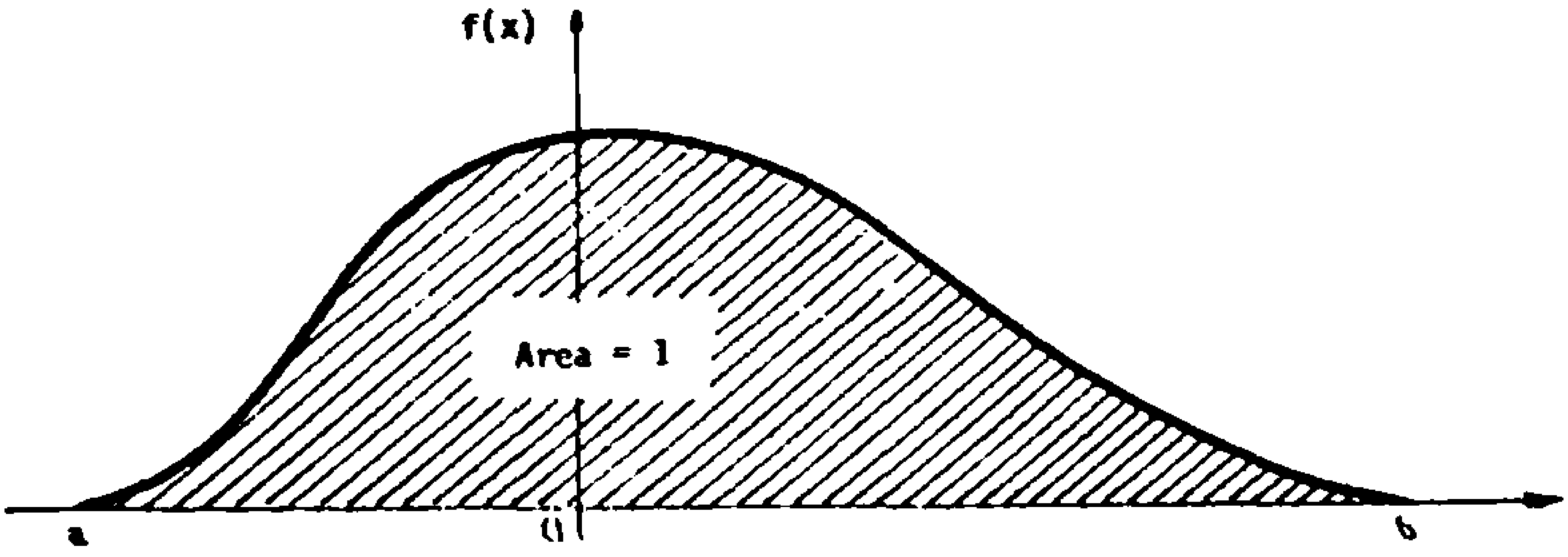


Fig. 2.3.1

Supongamos ahora que estamos interesados en calcular la probabilidad de que la variable aleatoria tome los valores entre a y b, donde el intervalo $[a,b] \subset R_X$. Es decir, queremos calcular la $P[a \leq X \leq b]$; puesto que todo área vale 1, bien podemos definir esta probabilidad como el área acotado por la gráfica de f, las rectas $x = a$, $x = b$ y el eje X. Por lo tanto, la probabilidad del evento

$$A = \{x \,/\, a \leq X \leq b\}$$

se define como sigue,

$$\boxed{P[A] = P[a \leq X \leq b] = \int_a^b f(x)dx} \qquad (*)$$

Este concepto está ilustrado en la figura 2.3.2.

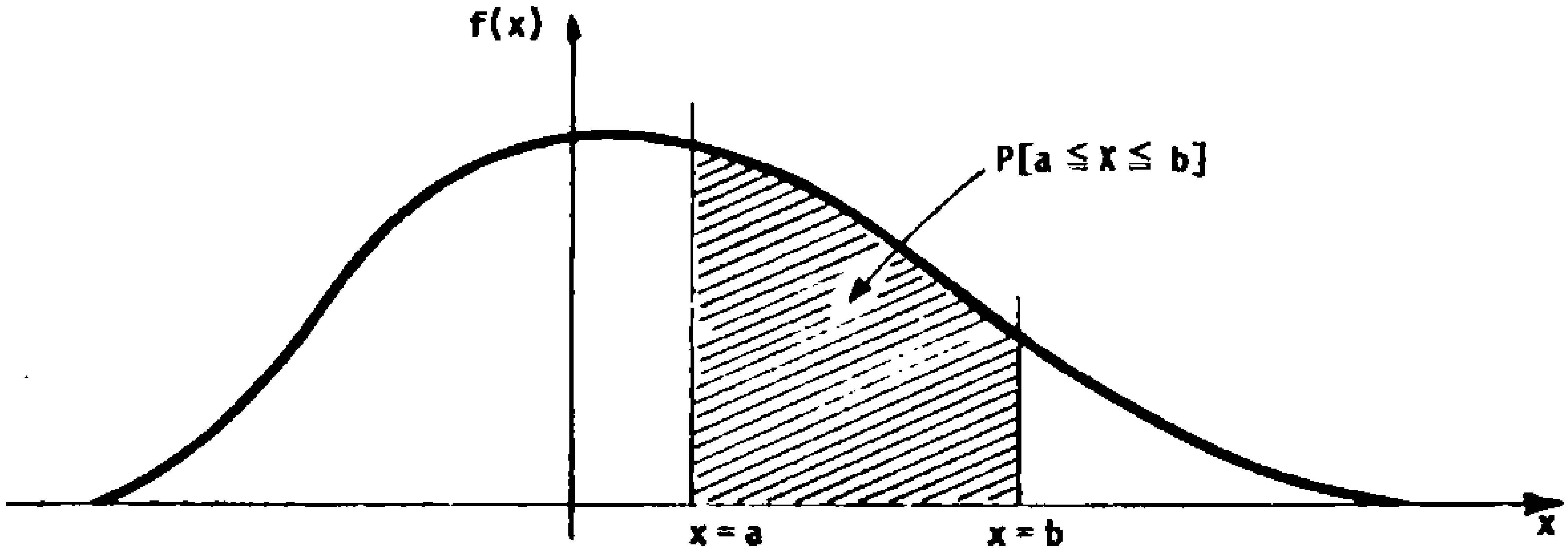


Fig. 2.3.2

OBSERVACIONES

1. Es importante darse cuenta que $f(x)$ no representa la probabilidad de algo y que solamente cuando la función se integra entre dos puntos produce una probabilidad.

2. Una consecuencia de la definición de probabilidad de un evento (ecuación $*$) es que para cualquier valor específico de X, digamos x_o, $P[X = x_o] = 0$, - puesto que

$$P[X = x_o] = P[x_o < X \leq x_o] = \int_{x_o}^{x_o} f(x)dx = 0.$$

Este resultado puede parecer contradictorio a la intuición, pués, si X es una variable aleatoria contínua se admite que X asume todos los valores - de algún intervalo, entonces $P[X = x_o] = 0$ no es equivalente a decir - que el evento $[X = x_o]$ en R_X es imposible. Recordemos que si $A = \phi$, entonces $P[A] = 0$, sin embargo el hecho que $P[X = x_o] = 0$ y que el - evento $A = \{x \,/\, X = x_o\}$ no es vacío indica que la inversa no es verdadera.

3. Una consecuencia inmediata de (2) es el siguiente resultado,

$$P[a \leq X \leq b] = P[a < X \leq b] = P[a < X < b] = P[a \leq X < b]$$

cuando X es una variable aleatoria contínua.

EJEMPLO 3 Sea X una variable aleatoria con función de densidad·

$$f(x) = \begin{cases} a(3x - x^2) & , \quad \text{si } 0 \leq x \leq 3 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar el coeficiente a ;

(b) construir la gráfica de la función de densidad de probabilidad ;

(c) Calcular la probabilidad que X se encuentre en el intervalo [1,2] .

(d) Hallar la probabilidad que X se encuentre en el intervalo $\langle -1 , \frac{3}{2} \rangle$

SOLUCION (a) Como todos los valores de variable aleatoria X se hallan comprendidos en el intervalo [0,3]. Es decir $R_X = [0,3]$. Entonces, por la condición (2) de la definición de función de densidad

$$\int_0^3 a(3x - x^2)dx = a\left[\frac{3}{2}x^2 - \frac{1}{3}x^3 \right]_0^3 = a(\frac{27}{2} - 9) = 1$$

de donde $a = \frac{2}{9}$.

(b) La gráfica de la función $f(x)$ en el intervalo [0,3] es la parábola

$y = \frac{2}{3}x - \frac{2}{9}x^2$ y fuera de éste intervalo, el gráfico es el propio eje de las abscisas

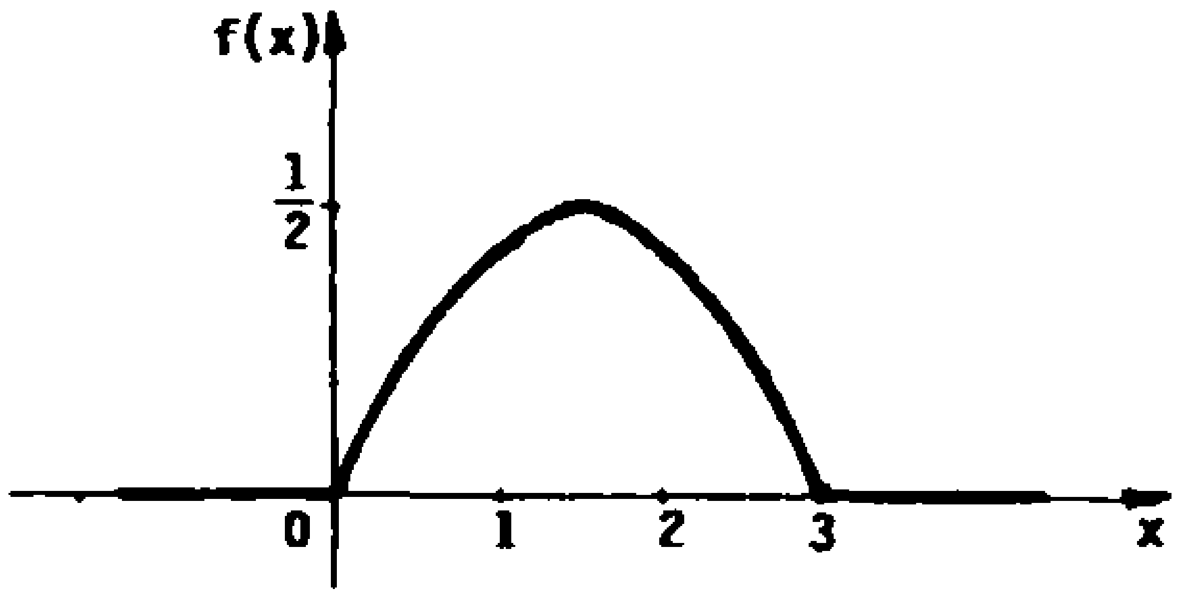


Fig. 2.3.3

(c) La probabilidad que la variable X se encuentre en el intervalo [1,2] se determina por la ecuación (*) .

$$P[1 \leq X \leq 2] = \int_1^2 (\frac{2}{3}x - \frac{2}{9}x^2)\,dx = \left[\frac{x^2}{3} - \frac{2x^3}{27} \right]_1^2$$

$$= \frac{4}{3} - \frac{16}{27} - \frac{1}{3} + \frac{2}{27} = \frac{13}{27} .$$

(d) Por la ecuación (*)

$$P[-1 < X < \frac{3}{2}] = \int_{-1}^0 0\,dx + \int_0^{3/2} (\frac{2}{3}x - \frac{2}{9}x^2)dx$$

$$= 0 + \left[\frac{x^2}{3} - \frac{2x^3}{27}\right]_0^{3/2} = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}.$$

EJEMPLO 4 Sea

$$f(x) = \begin{cases} kx^2 & , \quad \text{Si} \quad |x - 2| \leq 1 \\ 0 & , \quad \text{Si} \quad |x - 2| > 1 \end{cases}$$

(a) Hallar el valor de k tal que f sea una función de densidad de una variable aleatoria continua X ;

(b) Calcular la probabilidad que la variable aleatoria X se encuentre en el intervalo $\langle 2,3 \rangle$

SOLUCION Por la propiedad de valor absoluto

$|x - 2| \leq 1$, si sólo si $-1 \leq x - 2 \leq 1$, de donde $1 \leq x \leq 3$.

$|x - 2| > 1$, si sólo si $x - 2 > 1 \vee x - 2 < -1$ o sea $x > 3 \vee x < 1$

Por lo tanto la función $f(x)$ se escribe

$$f(x) = \begin{cases} 0 & \quad x < 1 \\ k\,x^2 & , \quad 1 \leq x \leq 3 \\ 0 & , \quad x > 3 \end{cases}$$

(a) Como todos los valores de la variable aleatoria se hallan en el intervalo $[1,3]$. Por la condición (2) de la definición de función de densidad

$$\int_1^3 kx^2 dx = k\frac{x^3}{3}\Big|_1^3 = k\left[9 - \frac{1}{3}\right] = k\,\frac{26}{3} = 1$$

de donde $\quad k = \dfrac{3}{26}$.

(b) Por la ecuación $(*)$

$$P[2 < X < 3] = \int_2^3 \frac{3x^2}{26}\,dx = \frac{x^3}{26}\Big|_2^3 = \frac{27}{26} - \frac{8}{26} = \frac{19}{26}.$$

EJEMPLO 5 La longitud de vida de una especie de plantas en cierto medio ambiente es una variable aleatoria contínua. Supongamos que la función de densidad de X es

$$f(x) = \frac{1}{120}\, e^{-x/120} \quad , \qquad x \geq 0$$

(a) ¿Qué proporción de plantas de esta especie mueren antes de los 100 días.

(b) Si una planta individual vive durante 100 días, ¿Cuál es la probabilidad que viva otros 100 días más?

SOLUCION El rango de la variable aleatoria es

$$R_X = \{x \,/\, 0 \leq x < \infty\}$$

(a) La proporción de plantas que mueren antes de los 100 días está dado por $P[X < 100]$. Aplicando la fórmula de la ecuación (*)

$$P[X < 100] = \int_0^{100} \frac{1}{120}\, e^{-\frac{x}{120}}\, dx = -\, e^{-\frac{x}{120}}\Big|_0^{100} = 1 - e^{-\frac{5}{6}} .$$

(b) Aplicando el concepto de probabilidad condicional dado en el capítulo 1

$$P[X > 200 \mid x > 100] = \frac{P[(X > 200) \cap (X > 100)]}{P[X > 100]}$$

$$= \frac{P[X > 200]}{P[X > 100]} = \frac{e^{-20/12}}{e^{-10/12}} =$$

$$= e^{-10/12} = e^{-5/6} .$$

EJEMPLO 6 Para qué valor de k la función

$$f(x) = \frac{k}{(1 + |x|)^2} \quad , \quad -\infty < x < \infty$$

es una función de densidad de probabilidad.

SOLUCION

$$\int_{-\infty}^{\infty} \frac{k\, dx}{(1 + |x|)^2} = \int_{-\infty}^{0} \frac{k\, dx}{(1 - x)^2} + \int_{0}^{\infty} \frac{k\, dx}{(1 + x)^2}$$

$$= \lim_{t \to -\infty} \frac{k}{1 - x}\Big|_t^0 + \lim_{t \to \infty} - \frac{k}{1 + x}\Big|_0^t = 1$$

de donde $\qquad k = \frac{1}{2}$.

EJEMPLO 7 La gráfica de la función de densidad de una variable aleatoria continua X, es como se muestra en la fig. 2.3.4 . Determinar el número m tal que

$$P[X > m] = \frac{1}{2} .$$

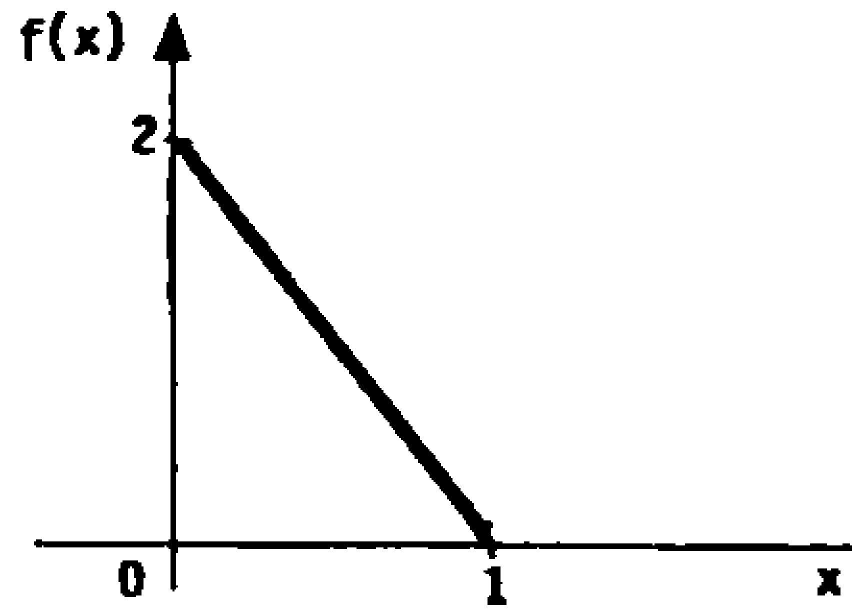


Fig. 2.3.4

SOLUCION De la figura observamos que la recta pasa por los puntos (1,0) y (0,2). Sea $y = ax + b$ la ecuación de la recta, entonces

$$\text{Si } x = 1, \text{ se tiene } 0 = a + b \quad (1)$$
$$\text{Si } x = 0, \text{ se tiene } 2 = 0 + b \quad (2)$$

de (1) y (2) $a = -2$ y $b = 2$

Es decir $y = 2(1 - x)$

Por lo tanto, la función de densidad de la variable aleatoria X, es

$$f(x) = \begin{cases} 2(1-x) & , \quad 0 \leq x \leq 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Cálculo de m . Por la ecuación (*)

$$P[X > m] = \int_m^{\infty} f(x)\,dx = \int_m^1 2(1-x)dx + \int_1^{\infty} 0dx = (m-1)^2 = \frac{1}{2}$$

de donde $m = 1 \pm \frac{1}{\sqrt{2}}$

De las dos soluciones escogemos el que se encuentra en el intervalo [0,1]

o se $m = 1 - \frac{1}{\sqrt{2}} = \frac{2 - \sqrt{2}}{2}$

EJEMPLO 8 Un agricultor encuentra que el peso en kilogramos de una papaya es una variable aleatoria X con función de densidad

$$f(x) = \begin{cases} \frac{3}{32}(4x - x^2) & , \quad 0 \leq x \leq 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) ¿Cuál es la probabilidad que una papaya pese menos de un kilogramo?

(b) Si el agricultor cosecha 20,000 papayas, ¿cuántas de ellas pesa menos -

de un kilogramo?

(c) El agricultor escoge al azar cuatro papayas para regalar a un amigo, - ¿cuál es la probabilidad que las cuatro pesen menos de un kilogramo?

SOLUCION La variable aletoria **X** está definido por

$$X(\omega) = \text{peso de una papaya (en kg.)}$$

(a) $$P[X < 1] = \int_0^1 \frac{3}{32}(4x - x^2)dx = \frac{3}{32}(2x^2 - \frac{x^3}{3}) \Big|_0^1 = \frac{3}{32}(2 - \frac{1}{3}) = \frac{5}{32}$$

(b) El número de papayas que pesan menos de un kilogramo es igual a

$$20000\, P[X < 1] = 20000 \times \frac{5}{32} = 3{,}125 \text{ papayas}$$

(c) Sean X_1, X_2, X_3 y X_4 los pesos de las papayas escogidas, entonces - se tiene

$$P[X_1 < 1,\, X_2 < 1,\, X_3 < 1,\, X_4 < 1] = (P[X < 1])^4 = (\frac{5}{32})^4$$, pués - los X_i son independientes .

EJEMPLO 9 La variable aleatoria **Y**. tiene por función de densidad

$$f(y) = \begin{cases} 1/5 & , \quad 0 \le y \le 5 \\ 0 & , \quad \text{en otro lugar} \end{cases}$$

Calcular la probabilidad que sean reales las raíces de la ecuación

$$4x^2 + 4xy + (y + 2) = 0$$

SOLUCION La solución de la ecuación dada es

$$x = \frac{-4y \pm \sqrt{16y^2 - 4(4)(y+2)}}{8} = \frac{-y \pm \sqrt{y^2 - y - 2}}{2}$$

Las raíces serán reales, esto es $x \in \mathbb{R}$ sólo, si $y^2 - y - 2 \ge 0$

$$y^2 - y - 2 = (y-2)(y+1) \ge 0 \iff \begin{cases} y - 2 \ge 0 \wedge y + 1 \ge 0 \\ \vee \\ y - 2 \le 0 \wedge y + 1 \le 0 \end{cases}$$

$$\iff \begin{cases} y \ge 2 \wedge y \ge -1 & (1) \\ \vee \\ y \le 2 \wedge y \le -1 & (2) \end{cases}$$

Luego, la solución a la desigualdad es el intervalo $[2, \infty\rangle \cup \langle -\infty, -1]$ y como el rango de la variable aleatoria es $[0,5]$. Por lo tanto, los valores de y que hacen que la solución de la ecuación sean reales son las $y \in$

$[0,5] \cap ([2, \infty\rangle \cup \langle -\infty, -1])$, es decir $y \in [2,5]$. Luego, la probabilidad que las raíces de la ecuación sean reales es equivalente a calcular

$$P[2 \leq Y \leq 5] = \int_2^5 \frac{1}{5}\, dy = \frac{1}{5}[5-2] = \frac{3}{5}.$$

EJEMPLO 10 La duración (en horas) de cierto transistor de televisor es una variable aleatoria continua **X**, cuya función de densidad es

$$f(x) = \begin{cases} 150/x^2 & , \quad x > 150 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Si un transistor determinado todavía funciona después de 200 horas. - ¿Cuál es la probabilidad de que dicho transistor dure a lo más 300 horas?

(b) Se instalan 3 de tales transistores en un televisor. ¿Cuál es la probabilidad de que ninguno tenga que ser sustituído en las primeras 200 horas de uso? ¿Cuál es la probabilidad de que los tres transistores tengan que ser sustituído durante las 200 primeras horas? ¿Cuál es la probabilidad de que exactamente 1 tenga que ser sustituída en las 200 primeras horas de uso?

SOLUCION (a) Por la definición de probabilidad condicional

$$P[X \leq 300 \mid X > 200] = \frac{P[200 < X < 300]}{P[X > 200]}$$

$$= \frac{\int_{200}^{300} 150/x^2 dx}{\int_{200}^{\infty} 150/x^2\, dx} = \frac{-\frac{150}{x}\Big|_{200}^{300}}{\lim\limits_{t\to\infty} -\frac{150}{x}\Big|_{200}^{t}} = \frac{1}{3}$$

(b_1) Sea **A**, el evento: "ninguna tenga que ser sustituída en las primeras 200 horas de uso".

Luego, $P[A] = [P[X > 200]]^3$

Pero $\quad P[X > 200] = \int_{200}^{\infty} \frac{150}{x^2}\,dx = \frac{3}{4}$.

por lo tanto,

$$P[A] = (\tfrac{3}{4})^3 = \frac{27}{64} .$$

(b_2) Sea B, el evento: " los tres transistores tengan que ser sustituída en las primeras 200 horas de uso".

Entonces, $\quad P[B] = [P[X \le 200]]^3$

Pero $\quad P[X \le 200] = \int_{150}^{200} \frac{150}{x^2}\,dx = \frac{1}{4}$.

Luego, $\quad P[B] = (\frac{1}{4})^3 = \frac{1}{64}$.

(b_3) Sea C, el evento: "exactamente 1 tenga que ser sustituído en las primeras 200 horas de uso".

Entonces,

$$P[C] = 3[P[X > 200]]^2 \; P[X \le 200] = 3(\tfrac{3}{4})^2 (\tfrac{1}{4}) = \frac{27}{64} .$$

2.3.2 FUNCION DE DISTRIBUCION DE UNA VARIABLE ALEATORIA CONTINUA

DEFINICION 2.3.3 Sea **X** una variable aleatoria contínua con función de densidad $f(x)$. ***La función de distribución*** (*o* ***función de distribución acumulada***) de la variable aleatoria **X**, denotado por "F(x)", se define por

$$\boxed{F(x) = P[X \le x] = \int_{-\infty}^{x} f(t)dx, \quad \forall \; x \in \mathbb{R}}$$

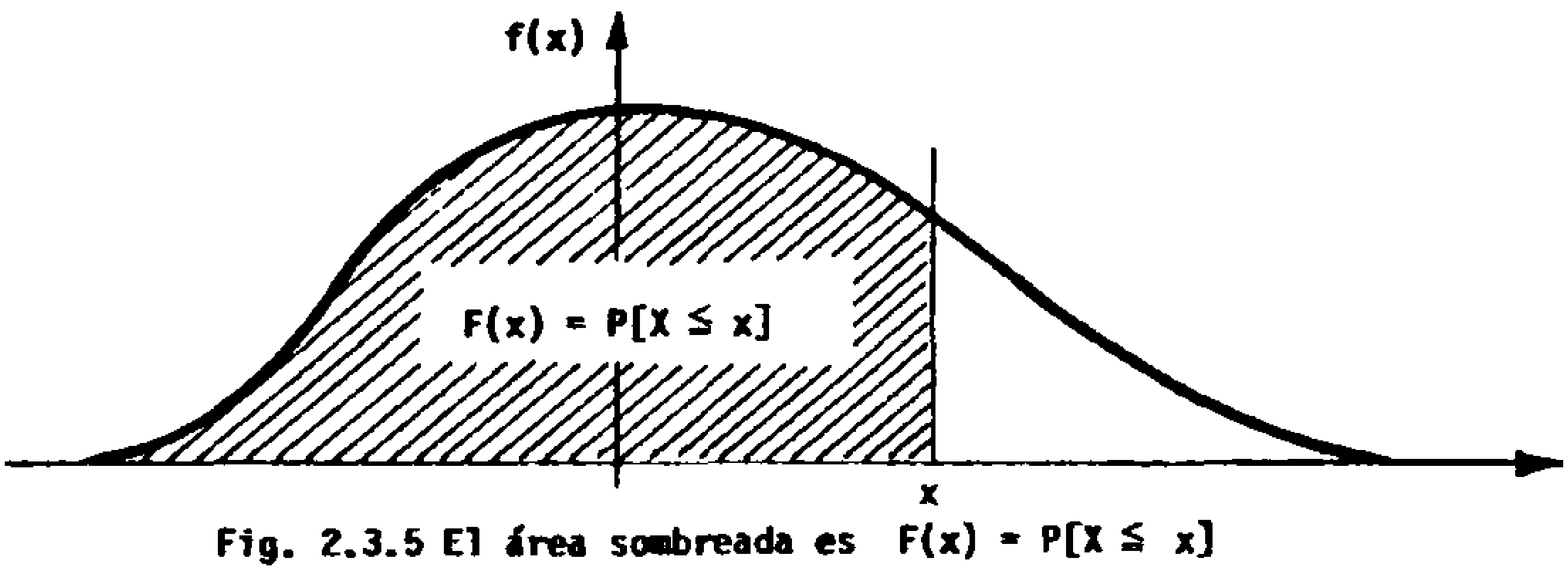


Fig. 2.3.5 El área sombreada es $F(x) = P[X \le x]$

EJEMPLO 11 Sea **X** una variable aleatoria con función de densidad definida por

$$f(x) = \begin{cases} \frac{3x}{4}(2 - x) & , \quad 0 \leq x \leq 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar la función de distribución de **X** y su gráfica.

(b) $F(1)$.

SOLUCION La función de distribución $F(x)$ está definida para todo $x \in \mathbb{R}$, - entonces consideremos los siguientes casos:

1. Si $x < 0$, $\quad F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{x} 0dt = 0$.

2. Si $0 \leq x \leq 2$, $\quad F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{x} f(t)\,dx = \int_{-\infty}^{0} 0\,dt +$

$$+ \int_{0}^{x} \frac{3}{4}\,t(2 - t)\,dt$$

$$= 0 + \frac{3}{4}\left(t^2 - \frac{t^3}{3}\right)\Big|_{0}^{x} = \frac{3}{4}\,x^2\left(1 - \frac{x}{3}\right) .$$

3. Si $x > 2$, $\quad F(x) = \mathbf{P}[X \leq x] = \int_{\infty}^{x} f(t)\,dt = 0 +$

$$+ \int_{0}^{2} \frac{3}{4}\,t\,(2 - t)dt + \int_{2}^{x} 0\,dt = 1 .$$

4. De (1), (2) y (3) la función de distribución de la variable aleatoria **X** es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{3}{4}\,x^2(1 - \frac{x}{3}) & , \quad 0 \leq x < 2 \\ 1 & , \quad x \geq 2 \end{cases}$$

La gráfica de $F(x)$ se muestra en la fig. 2.3.6

(b) $F(1) = \mathbf{P}[X \leq 1] = \frac{3}{4}\,(1)^2(1 - \frac{1}{3}) = \frac{1}{2}$.

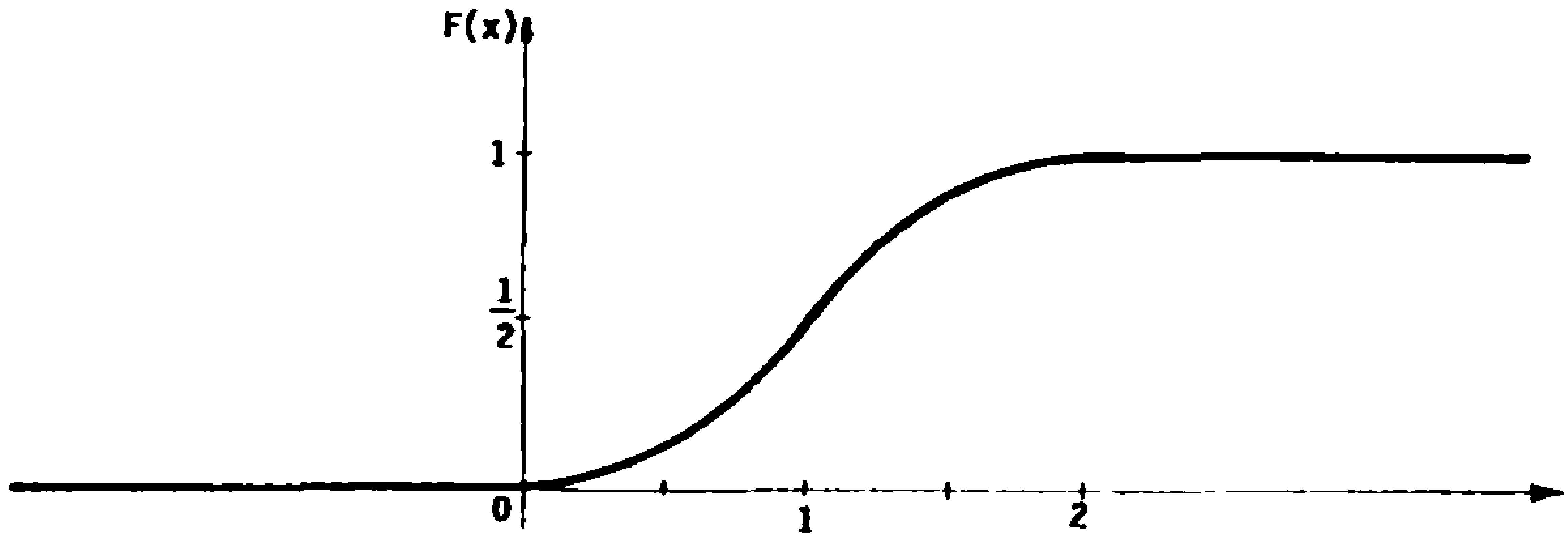


Fig. 2.3.6

EJEMPLO 12 Sea X una variable aleatoria cuya función de densidad está dada por la figura 2.3.7

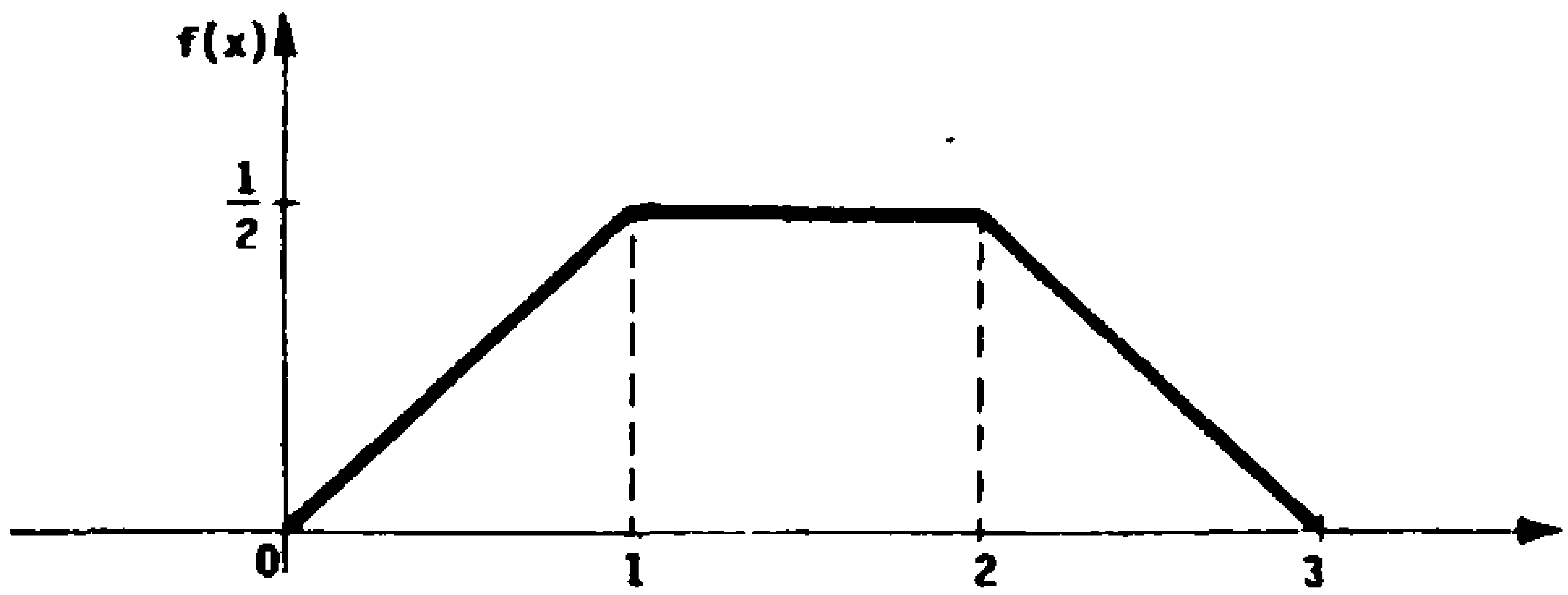


Fig. 2.3.7

(a) Hallar la función de distribución y su gráfica;

(b) Calcular $P[X \le \frac{3}{4}]$, $P[X < \frac{3}{2}]$, $P[X \le 2.4]$.

SOLUCION

1. De la figura, para $0 \le x \le 1$, la función de densidad de la variable aleatoria es una recta que pasa por los puntos $(0,0)$ y $(1, \frac{1}{2})$. Es decir, la ecuación de dicha recta es $\frac{1}{2}x$ para todo $0 \le x \le 1$.
2. Cuando $1 < x \le 2$, la función es constante e igual a $\frac{1}{2}$.
3. Para $2 \le x \le 3$, la función de densidad es una recta de pendiente negativa, pasa por los puntos $(2, \frac{1}{2})$ y $(3,0)$; por lo tanto, la ecuación es
$$-\frac{1}{2}x + \frac{3}{2} .$$
4. De (1), (2) y (3) la función de densidad de la variable aleatoria está dado por

$$f(x) = \begin{cases} \frac{1}{2}x & , \quad 0 \leq x \leq 1 \\ \frac{1}{2} & , \quad 1 < x < 2 \\ -\frac{1}{2}x + \frac{3}{2} & , \quad 2 \leq x \leq 3 \end{cases}$$

La función de distribución F(x) está definido para todo $x \in \mathbb{R}$, por lo tanto, consideremos los siguientes casos :

(*i*) Si $x < 0$, $F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{x} 0\,dt = 0$

(*ii*) Si $0 \leq x \leq 1$, $F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} \frac{1}{2}\,t\,dt = \frac{x^2}{4}$.

(*iii*) Si $1 < x < 2$, $F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{0} 0\,dt + \int_{0}^{1} \frac{1}{2}\,t\,dt + \int_{1}^{x} \frac{1}{2}\,dt$

$$= \frac{1}{4} + \frac{1}{2}x - \frac{1}{2} = \frac{1}{2}x - \frac{1}{4}.$$

(*iv*) Si $2 \leq x \leq 3$, $F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{0} 0\,dt + \int_{0}^{1} \frac{1}{2}\,t\,dt + \int_{1}^{2} \frac{1}{2}\,dt + \int_{2}^{x} (-\frac{1}{2}t + \frac{3}{2})\,dt$

$$= 1 - \frac{1}{4} + \frac{3}{2}x - \frac{1}{4}x^2 - 3 + 1$$

$$= \frac{3}{2}x - \frac{x^2}{4} - \frac{5}{4}.$$

(*v*) Si $x > 3$, $F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^{2} f(t)\,dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{3} f(t)\,dt + \int_{0}^{x} 0\,dt = 1$

Luego, la función de distribución está dado por

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1}{4}x^2 & , \quad 0 \leq x < 1 \\ \frac{1}{2}x - \frac{1}{4} & , \quad 1 \leq x < 2 \\ \frac{3}{2}x - \frac{x^2}{4} - \frac{5}{4} & , \quad 2 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

La gráfica de $F(x)$ se indica en la fig. 2.3.8

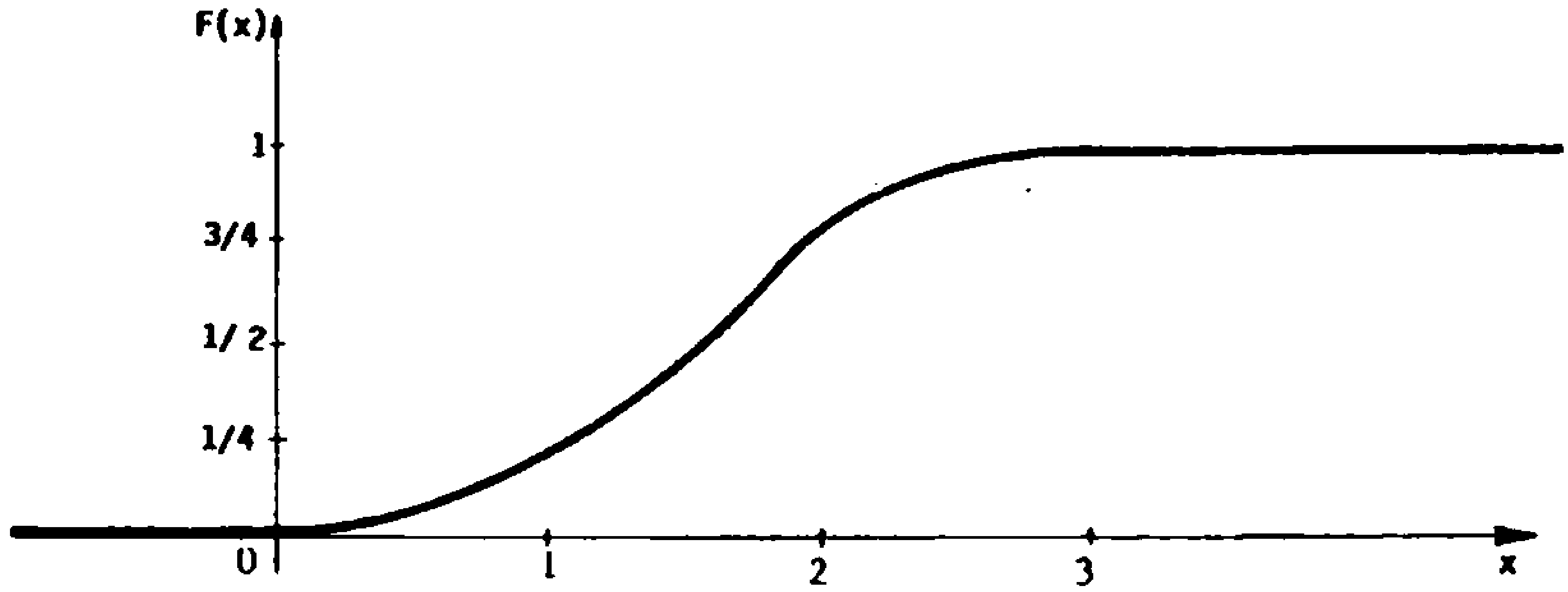


Fig. 2.3.8

(b_1) $\quad F(\frac{3}{4}) = P[X \leq \frac{3}{4}] = \frac{1}{4}(\frac{3}{4})^2 = \frac{9}{64}$;

(b_2) $\quad F(\frac{3}{2}) = P[X < \frac{3}{2}] = \frac{1}{2}(\frac{3}{2}) - \frac{1}{4} = \frac{1}{2}$;

(b_3) $\quad F(2.4) = F(\frac{12}{5}) = P[X \leq \frac{12}{5}] = \frac{3}{2}(\frac{12}{5}) - \frac{\frac{144}{25}}{4} - \frac{5}{4}$

$$= \frac{36}{10} - \frac{36}{25} - \frac{5}{4} = \frac{91}{100} = 0.91 .$$

Por definición, la función de distribución $F(x)$, da la probabilidad de eventos de la forma : $(X \leq a)$, $(X < a)$, $(X > a)$, etc. Es decir,

$$P[X \leq a] = P[X < a] = F(x) = \int_{-\infty}^{a} f(t)\,dt \qquad y$$

$$P[X > a] = 1 - P[X \leq a] = 1 - F(a)$$

La función de distribución también se puede utilizar, para determinar probabilidades de eventos de la forma :

$$(a \leq X \leq b), \quad (a < X \leq b), \quad (a < X < b) \text{ etc.}$$

En efecto: Sean $a,b \in \mathbb{R}$, tales que $a < b$, entonces

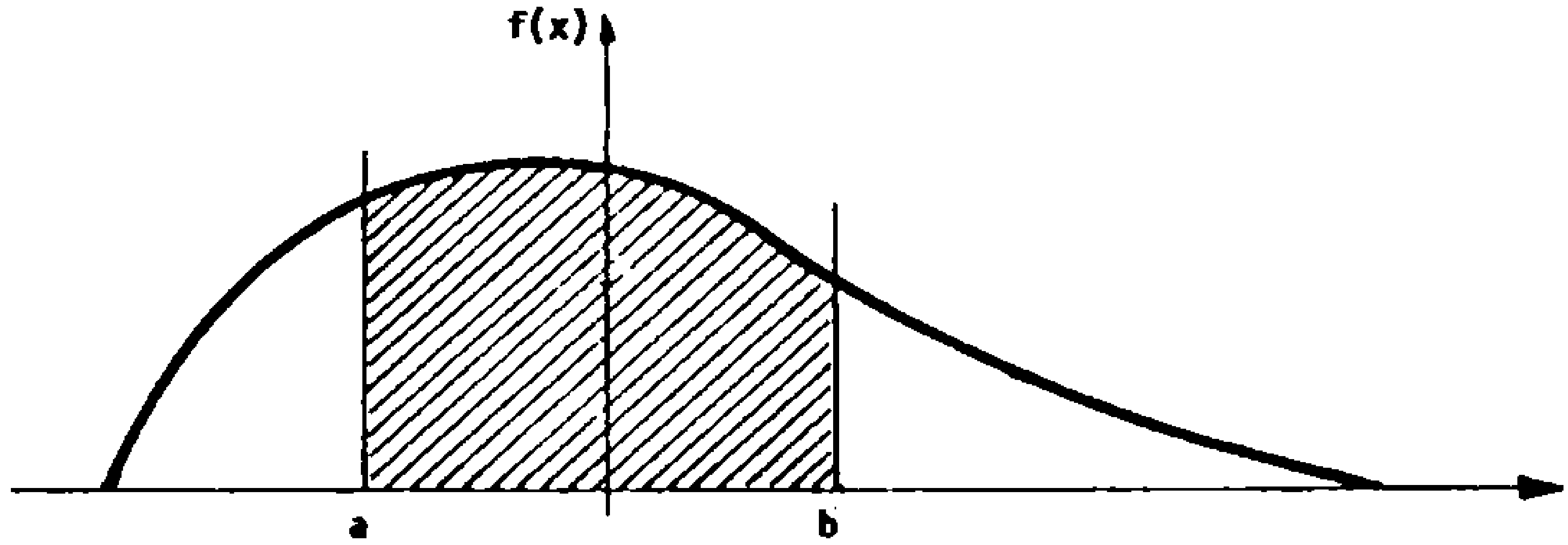


Fig. 2.3.9

$$P[a \leq X \leq b] = P[a < X \leq b] = P[a \leq X \leq b] = P[a < X < b]$$

$$= P[X \leq b] - P[X \leq a] = F(b) - F(a).$$

Así, si queremos calcular en el ejemplo anterior

$$P[\tfrac{3}{4} \leq X \leq \tfrac{5}{2}], \quad \text{tendremos}$$

$$P[\tfrac{1}{4} \leq X \leq \tfrac{5}{2}] = F(\tfrac{5}{2}) - F(\tfrac{1}{4})$$

$$= \frac{3}{2}(\frac{5}{2}) - \frac{1}{4}(\frac{5}{2})^2 - \frac{5}{4} - \frac{1}{4}(\frac{3}{4})^2 = \frac{51}{64}.$$

2.3.3 PROPIEDADES DE LA FUNCION DE DISTRIBUCION

Las tres primeras propiedades son los mismos que el caso discreto.

1. $0 \leq F(x) \leq 1, \quad \forall x \in \mathbb{R}$.

2. $$F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} \int_{-\infty}^{x} f(t)\,dt = 0 ;$$

$$F(\infty) = \lim_{x \to \infty} F(x) = \lim_{x \to \infty} \int_{-\infty}^{x} f(t)\,dt = 1.$$

3. La función de distribución es no decreciente, esto es si $a \leq b$, entonces

$$F(a) \leq F(b)$$

4. $\lim_{h \to 0} F(x + h) = F(x)$, $\forall\, x \in \mathbb{R}$, con $h > 0$; osea F es continua por la derecha, en todos los puntos.

5. Del segundo teorema fundamental del cálculo se tiene que si F(x) es una función derivable, entonces

$$\boxed{f(x) = \frac{d}{dx} F(x)}$$

EJEMPLO 13 Definida una función F(x) como sigue

$$F(x) = \begin{cases} 1 - ke^{-\frac{x}{2}} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

(a) para qué valor de k, F(x) es una función de distribución de una variable aleatoria continua **X** ;

(b) Determinar, $P[X \geq 2]$, $P[2 \leq X \leq 4]$, $P[X \geq -1]$.

SOLUCION (a) De la definición de la función, F(x), el rango de la variable aleatoria es, $0 \leq x < \infty$, entonces, se debe tener que

$$F(\infty) - F(0) = 1$$

$$1 - 0 - (1 - ke^{0}) = 1$$

de donde $k = 1$

por lo tanto, la función F(x) definida por

$$F(x) = \begin{cases} 1 - e^{-x/2} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

es una función de distribución de una variable aleatoria X con rango $R_X = [0, \infty\rangle$ y función de densidad

$$f(x) = F'(x) = \begin{cases} \frac{1}{2} e^{-x/2} & , \quad x \geq 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(b_1) $\quad P[X \geq 2] = 1 - P[X < 2] = 1 - F(2)$

$$= 1 - [1 - e^{-1}] = e^{-1}.$$

(b_2) $\quad P[2 \leq X \leq 4] = F(4) - F(2)$

$$= 1 - e^{-2} - (1 - e^{-1}) = \frac{1}{e} - \frac{1}{e^2}.$$

(b_3) $\quad P[X \geq -1] = 1 - P[X < -1] = 1 - F(1) = 1 - 0 = 1.$

EJEMPLO 14 Se selecciona al azar un punto del interior de un triángulo equilátero cuyo lado mide 6 cm. Sea **X** la distancia del punto a la base, determinar la función de distribución y la función de densidad de **X**.

SOLUCION La variable aleatoria **X** está definida por

$X(\omega)$ = la distancia del punto elegido a la base del triángulo es claro que X es una variable aleatoria contínua con rango

$$R_X = \{x \in \mathbb{R} \;/\; 0 < x \leq 3\sqrt{3}\}$$

donde $3\sqrt{3}$ es la altura del triángulo equilátero. Es decir,

$$h = \sqrt{6^2 - 3^2} = 3\sqrt{3}.$$

(a) para $x \leq 0$, la función de distribución $F(x) = 0$.

(b) para $0 < x < 3\sqrt{3}$, se tiene que

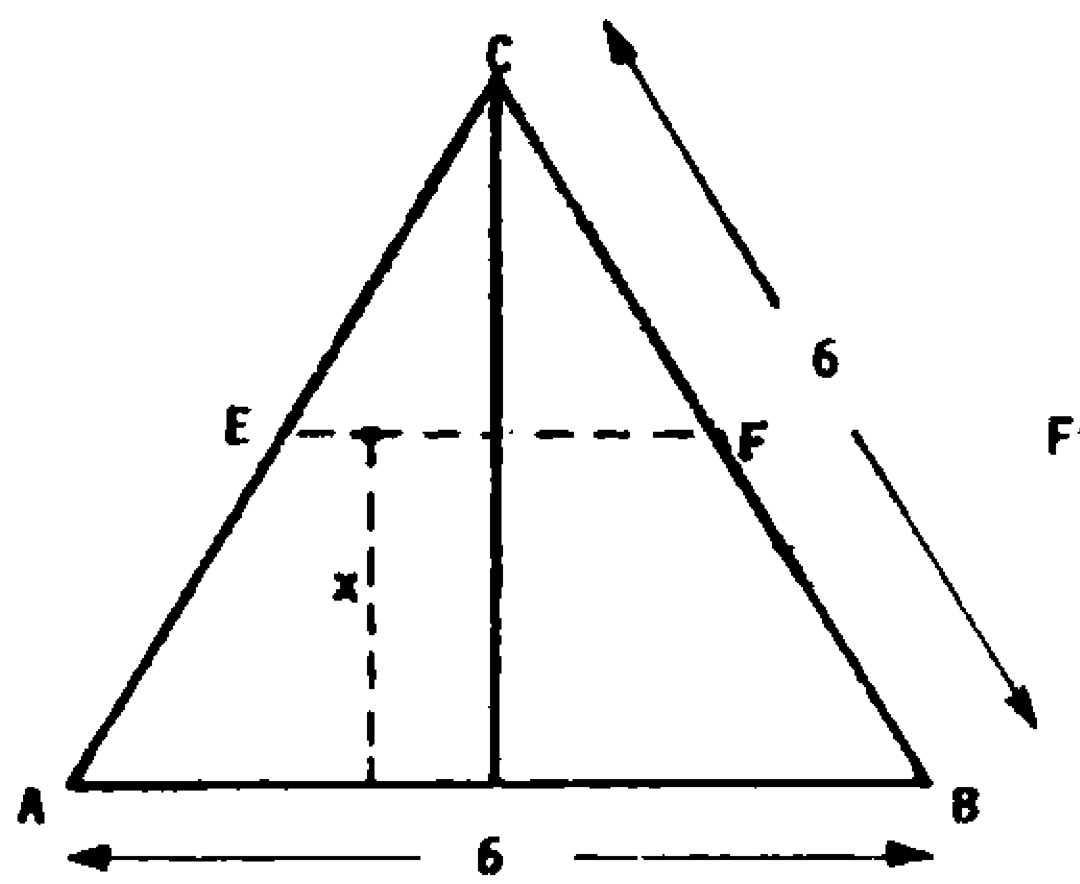


Fig. 3.3.10

$$F(x) = P[X \leq x] = \frac{\text{Area del Trapecio ABEF}}{\text{Area del Triángulo ABC}} = \frac{A_{\square}}{A_{\Delta}} \qquad (1)$$

$$A_{\Delta} = \frac{(6)(3\sqrt{3})}{2} = 9\sqrt{3} \qquad (2)$$

$$A_{\square} = \frac{1}{2}(\overline{AB} + \overline{EF})x$$

donde X es la altura del Trapecio = distancia del punto elegido a la base.

y $\overline{AB} = 6$

Cálculo de $\overline{EF}$. En los triángulos semejantes de la figura 3.3.10 se tiene

$$\frac{\frac{\overline{EF}}{2}}{3} = \frac{3\sqrt{3} - x}{3\sqrt{3}} \text{ , de donde } \overline{EF} = \frac{2(3\sqrt{3} - x)}{\sqrt{3}} .$$

Luego,

$$A_{\triangle} = \frac{1}{2}\left[6 + \frac{2(3\sqrt{3} - x)}{\sqrt{3}}\right] x = \frac{(6\sqrt{3} - x)x}{\sqrt{3}} \qquad (3)$$

reemplazando (2) y (3) en (1) obtenemos,

$$F(x) = P[X \leq x] = \frac{\frac{(6\sqrt{3} - x)x}{\sqrt{3}}}{9\sqrt{3}} = \frac{(6\sqrt{3} - x)x}{27} .$$

(c) para los $x > 3\sqrt{3}$, la función $F(x) = 1$.

De (*a*), (*b*) y (*c*) se tiene la función de distribución,

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{(6\sqrt{3} - x)x}{27} & , \quad 0 < x < 3\sqrt{3} \\ 1 & , \quad x \geq 3\sqrt{3} \end{cases}$$

y la función de densidad de la variable aleatoria X es,

$$f(x) = F'(x) = \begin{cases} \frac{6\sqrt{3} - 2x}{27} & , \quad 0 < x \leq 3\sqrt{3} \\ 0 & , \quad \text{en otros casos} \end{cases}$$

EJEMPLO 15 Una estación de servicio es provisionado de gasolina una vez semanal. El volumen X de la posible venta semanal en miles de galones tiene - la siguiente función de distribución acumulada.

$$F(x) = 1 - (1 - x)^4 , \quad 0 < x < 1$$

(a) ¿Cuál debe ser la capacidad de su tanque para que la probabilidad que - su provisión se agota en una semana sea sólo de 0.01?

(b) ¿Cuál es la probabilidad de que la venta semanal este entre 800 y 900 - galones?

SOLUCION (a) Sea C la capacidad del tanque. Por lo tanto, se quiere determinar el valor de C, tal que,

$$\begin{aligned} 0.01 &= P[X \geq C] \\ &= 1 - P[X \leq C] \\ &= 1 - [1 - (1 - C)^4] \\ &= (1 - C)^4 \end{aligned}$$

de donde, $C = 1 - \sqrt{0.1}$.

(b) $P[0.8 < X < 0.9] = F(0.9) - F(0.8) = 1 - (1 - 0.9)^4 - [1-(1-0.8)^4]$

$= (0.2)^4 - (0.1)^4 = 0.0015$.

EJEMPLO 16 La función de distribución de la variable aleatoria contínua X, es

$$F(x) = a + b \text{ arctg } (\tfrac{x}{2}) \quad , \quad x \in \langle -\infty, \infty\rangle$$

Determinar :

(a) las constantes a y b ;

(b) la función de densidad de la variable aleatoria X ;

(c) Calcular $P[2 < X < 2\sqrt{3}]$.

SOLUCION (a) Cálculo de las constantes a y b .

1. $F(-\infty) = a + b \text{ arctg } (-\infty) = 0$, propiedad de F(x) ó $a + b(-\frac{\pi}{2}) = 0$

de donde, $a = \frac{b\pi}{2}$

2. $F(-\infty) = a + b \text{ arctg}(\infty) = 1$, propiedad de F(x) ó $a + b(\frac{\pi}{2}) = 1$

de donde, $a = 1 - \frac{b\pi}{2}$

3. De las ecuaciones (1) y (2) obtenemos, $b = \frac{1}{\pi}$ y $a = \frac{1}{2}$.

por lo tanto, la función de distribución de X es,

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \text{ arctg}(\tfrac{x}{2}) \quad , \quad -\infty < x < \infty$$

(b) Cálculo de la función de densidad de X

$$f(x) = F'(x) = \frac{1}{\pi} \cdot \frac{1/2}{1 + (\frac{x}{2})^2}$$

$$f(x) = \frac{2}{\pi(4 + x^2)} \quad , \quad -\infty < x < \infty$$

(c) $P[2 < x < 2\sqrt{3}] = F(2\sqrt{3}) - F(2)$

$$= \frac{1}{2} + \frac{1}{\pi}\text{arctg}(\sqrt{3}) - [\frac{1}{2} + \frac{1}{\pi}\text{arctg}(1)]$$

$$= \frac{1}{12} \ .$$

PROBLEMAS 2.3

1. ¿Cuales de las siguientes funciones representan, función de densidad de una variable aleatoria contínua X?. Grafique las que sean función de densidad de probabilidad. Determine su rango en cada caso y la función de - distribución acumulada.

(a) $f(x) = \begin{cases} e^x & , \quad -\infty < x < 0 \\ 0 & , \quad x > 0 \end{cases}$

(b) $f(x) = \begin{cases} 1 & , \quad -\frac{1}{2} \leq x \leq \frac{1}{2} \\ 0 & , \quad \text{en otros casos} \end{cases}$

(c) $f(x) = \begin{cases} \frac{3}{4}(1 - x^2) & , \quad -1 \leq x \leq 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$

(d) $f(x) = \frac{1}{2} e^{-|x|} \quad , \quad -\infty < x < \infty$

(e) $f(x) = 1 - |1 - x| \quad , \quad 0 < x < 2$

(f) $f(x) = |x| \qquad |x| \leq 1$

2. Para qué valor de a, la función,

$$f(x) = \frac{2a}{1 + x^2} \quad , \quad -\infty < x < \infty$$

es la función densidad de probabilidad de una variable aleatoria X

3. Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} a/\sqrt{a^2 - x^2} & , \quad \text{Si} \quad |x| < a \\ 0 & , \quad \text{Si} \quad |x| \geq a \end{cases}$$

Hallar : (a) el coeficiente a ; (b) la probabilidad de que la variable aleatoria **X** se encuentre en el intervalo $\langle a/2 , a \rangle$;

(c) construir la gráfica de la función de densidad de **X**.

4. Se le da función de densidad de probabilidad de la variable aleatoria X

$$f(x) = \begin{cases} a \ \text{sen } x & , \quad \text{Si} \quad 0 \leq x \leq \pi \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar : (a) la constante a ; (b) la función de distribución.

5. Se da la función

$$f(x) = \begin{cases} a\,x^2 & , \quad \text{si} \quad 0 \leq x < 1 \\ a(2 - x)^2 & , \quad \text{si} \quad 1 \leq x < 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Para qué valor de a la función $f(x)$ es la densidad de probabilidad de una variable aleatoria X.

(b) Hallar $F(x)$.

6. El tiempo de llegada **X** de camiones (en minutos) a un depósito. Se comporta de acuerdo a la siguiente función de densidad

$$f(x) = \begin{cases} \dfrac{1}{x^2} & , \quad \text{si} \quad x > 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Se desea elegir una muestra al azar de 4 camiones. Determinar un número C tal que la probabilidad de que al menos uno de los cuatro camiones extraidas tenga un tiempo de llegada que excede a C sea 0,9375.

7. La duración en minutos de un disco de 33-rpm grabados por la compañía - Disquera Andina S.A. es una variable aleatoria **X** con una función de densidad

$$f(x) = \begin{cases} \dfrac{1}{3}x - \dfrac{x^2}{36} - \dfrac{3}{4} & , \quad \text{si} \quad 3 \leq x \leq 9 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) ¿Cuál es la probabilidad que la duración de un disco exceda a 6 minutos?

(b) Si la compañía graba 1000 discos, ¿Cuántos de ellos tiene una duración superior a 6 minutos?

8. Un agricultor encuentra que el peso en kilogramos de un melón es una variable aleatoria X con función de densidad

$$f(x) = \begin{cases} \frac{2}{3} - \frac{1}{12}x^3 & , \quad 0 \leq x \leq 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) ¿Cuál es la probabilidad que un melón pesa menos de un kilogramo?

(b) Si el agricultor cosecha 24,000 melones, ¿cuántos de ellos pesará menos de un kilogramo?

9. Para cierto negocio por correspondencia la proporción de los pedidos procesados en 24 horas tiene la función de densidad de probabilidad

$$f(x) = 20x^3(1 - x) \ , \quad 0 \leq x \leq 1$$

¿Cuál es la probabilidad que sobre un periodo relativamente largo

(a) al menos 80 % de los pedidos sean procesados dentro de las 24 horas?

(b) menos del 30 % de todos los pedidos sean procesados dentro de las 24 horas ?

(c) entre el 50, 60 % de todos los pedidos sean procesados dentro de - las 24 horas ?

10. En cierto país, el ingreso familiar tiene la función de densidad de probabilidad

$$f(x) = 2x/(1 + x^2)^2 \quad , \quad x \geq 0$$

donde x esta en miles de dolares. ¿que proporción de las familias

(a) tiene ingresos menores que \$ 6000 ?

(b) tiene ingresos entre \$ 2000 y 8000 ?

(c) tienen ingresos sobre \$ 1000 ?

11. En una tienda grande que vende al por mayor, el volumen de ventas de juguetes en el mes de Diciembre tiene la función de densidad

$$f(x) = \frac{1}{4} x \ e^{-x/2} \quad , \quad x > 0$$

donde x está en miles de docenas. Calcular la probabilidad de que se vendan

(a) menos de 10,000 docenas ;
(b) entre 600 y 12000 docenas ;
(c) al menos 15000 docenas .

12. Para un médico, el tiempo en horas que dedica a un paciente en su visita al consultorio tiene una función de densidad

$$f(x) = \begin{cases} \dfrac{4}{3(1+x)^2} & , \quad 0 \leq x \leq 3 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Para un paciente elegido al azar, ¿cuál es la probabilidad de que el tiempo dedicado por el médico es por lo menos una hora?

13. Un agricultor encuentra que el peso en kilogramos de una piña es una variable aleatoria X con función de densidad

$$f(x) = \begin{cases} \dfrac{1}{39}(x^2 - 10x + 25) & , \quad 0 \leq x \leq 3 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) para una piña elegida al azar, ¿cuál es la probabilidad que pese menos de 2 kilogramos?
(b) Si el agricultor cosecha 23,400 piñas, ¿cuántas de estas pesará menos de dos kilogramos?
(c) Se escoge al azar 3 piñas, ¿cuál es la probabilidad que al menos dos pesen menos de 2 kilogramos?

14. Considere la siguiente función

$$f(x) = \begin{cases} kx & , \quad 0 \leq x < 2 \\ k(4-x) & , \quad 2 \leq x \leq 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

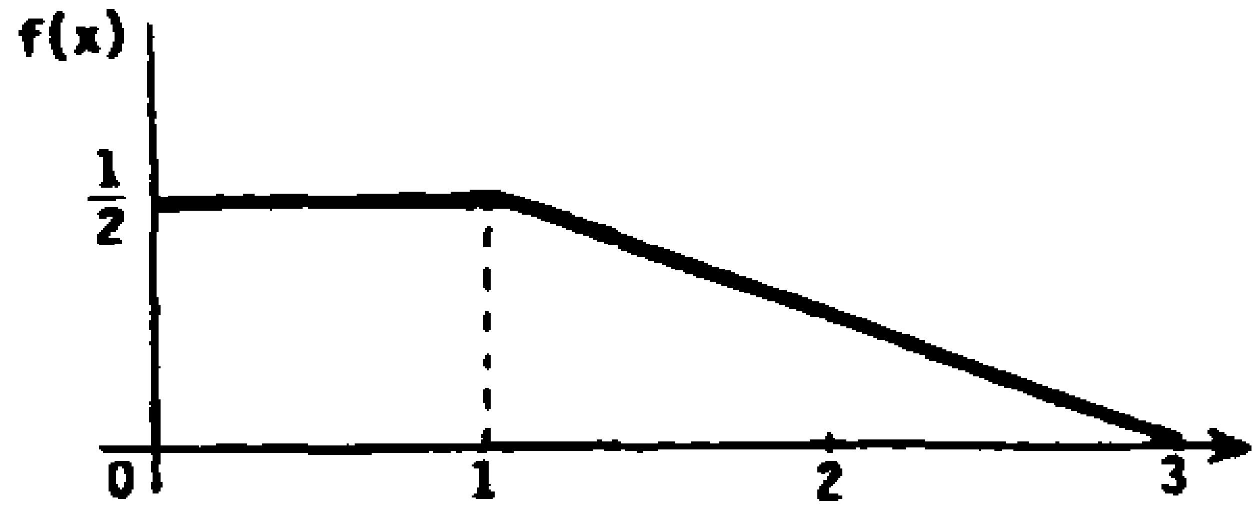
(a) Hallar el valor de k para que f sea una función de densidad.
(b) Hallar la función de distribución.

15. Sea X una variable aleatoria cuya función de densidad está dada por,

$$f(x) = \begin{cases} \dfrac{|x|}{a} & , \quad -5 \leq x \leq 5 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Encontrar el valor de a .
(b) Hallar $F(x)$ y su gráfica.

16. Suponga que X es una variable aleatoria cuya función de densidad está representada por la figura



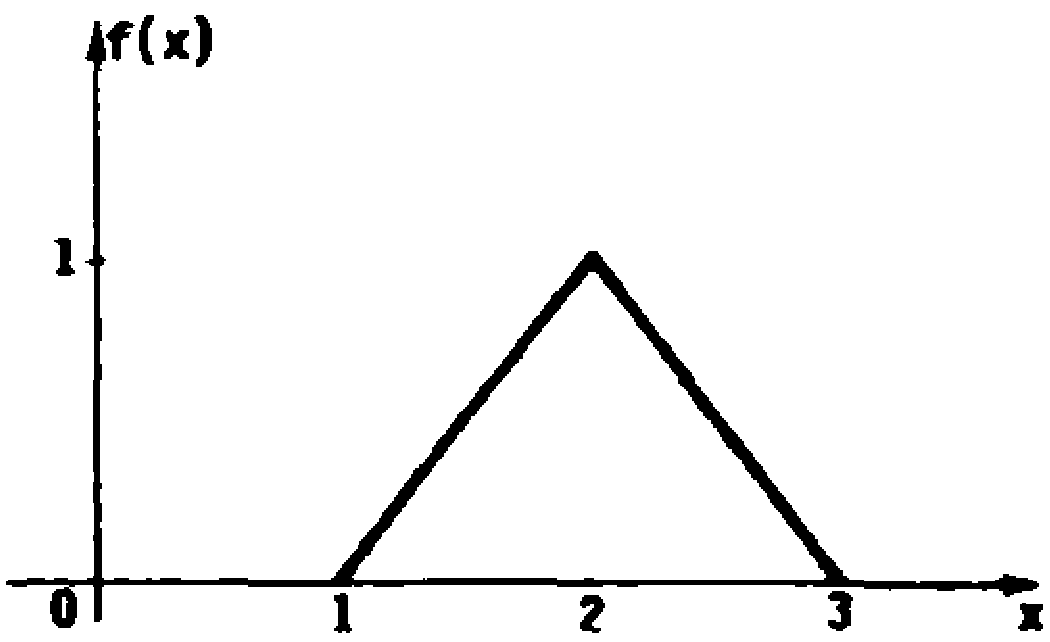
(a) Si $P[\frac{1}{3} \le X \le a] = \frac{1}{2}$

determinar el valor de a .

(b) Calcular $P[\frac{1}{2} \le X \le 2]$.

(c) Hallar $F(x)$.

17. Una estación gasolinera recibe provisión semanalmente. Las estadísticas anteriores sugieren que la función de densidad de probabilidad de las ventas semanales X, medidas en miles de galones, se aproxima a la función cuya gráfica se muestra en la figura.



(a) Hallar la función de densidad de X.

(b) la función de distribución de X .

18. Sea X una variable aleatoria cuya función de densidad está dado por

$$f(x) = \begin{cases} 0 & x < 10 \\ e^{-(x-10)}, & x \ge 10 \end{cases}$$

(a) Hallar el número c tal que X sea igualmente probable de ser mayor o menor que c.
(b) Encuentre el número d tal que la probabilidad que X sea mayor que d es igual a 0.05.

19. La función de densidad de la variable aleatoria X, la resistencia al corte de ensayos de punto de soldadura, está dada por,

$$f(x) = \begin{cases} \dfrac{x}{250{,}000} & , \quad 0 \leq x \leq 500 \\ \dfrac{1{,}000 - x}{250{,}000} & , \quad 500 \leq x \leq 1{,}000 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar el número a tal que $P[X < a] = 0.50$ y el número b tal que $P[X < b] = 0.90$.

20. ¿Cuáles de las siguientes funciones presenta la función de distribución de una variable aleatoria contínua. Grafique cada una de las funciones de distribución. Determine el rango y la función de densidad de la variable aleatoria.

(a) $$F(x) = \begin{cases} e^{x} & , \quad x \leq 0 \\ 1 & , \quad x > 0 \end{cases}$$

(b) $$F(x) = \begin{cases} \frac{1}{2} e^{x} & , \quad x \leq 0 \\ 1 - \frac{1}{2} e^{-x} & , \quad x > 0 \end{cases}$$

(c) $$F(x) = \begin{cases} 0 & , \quad x \leq -\frac{1}{2} \\ x + \frac{1}{2} & , \quad -\frac{1}{2} < \quad < \frac{1}{2} \\ 1 & , \quad x \geq \frac{1}{2} \end{cases}$$

Determine también $P[0 \leq X \leq 2]$ en cada caso.

21 Sea X la variable aleatoria que designa la vida (en horas) de una bombilla eléctrica. La función de distribución acumulativa está dado por

$$F(x) = \begin{cases} -\dfrac{1 - k}{x} & , \quad x \geq 1{,}000 \\ 0 & , \quad x < 1{,}000 \end{cases}$$

(a) Determinar el valor de k.

(b) ¿Cuál es la función de densidad para la variable aleatoria X?

(c) ¿Cuál es la probabilidad que una bombilla dure más de 1,500 horas?

(d) ¿Cuál es la probabilidad que una bombilla dure menos de 2,000 horas, si se sabe que la bombilla todavía funcionaba después de 1,500 horas?.

22. Si la función $f(x)$ es una función de densidad de una variable aleatoria X

$$f(x) = \begin{cases} 2x & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar un número n tal que, $P[X \geq n] = P[X \leq n]$

23. La función de densidad de una variable aleatoria contínua X es,

$$f(x) = \begin{cases} \dfrac{\text{sen } x}{2} & , \quad 0 \leq x \leq \pi \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Calcular el valor de m tal que, $P[X < m] = P[X > m] = \frac{1}{2}$

24. La función de densidad de una variable aleatoria X está dado por,

$$f(x) = \begin{cases} 4x^3 & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Determinar el número c tal que la probabilidad que X exceda a c sea igual a 0.05.

25. La función de densidad de la variable aleatoria X es,

$$f(x) = \begin{cases} x/2 & , \quad 0 < x < 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Si se obtienen dos valores de X. ¿Cuál es la probabilidad que ambos sean mayores que 1?.
(b) Si se obtienen 3 valores de X. ¿Cuál es la probabilidad que precisamente dos de ellos sean mayores que 1?

26. Una variable aleatoria X tiene como función de densidad,

$$f(x) = \begin{cases} 1 & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar el número a tal que sea 0.90 la probabilidad que al menos uno de cuatro valores de X estraídos al azar exceda a a.

27. La longitud de vida de una especie de planta en cierto medio ambiente es una variable aleatoria contínua X?

$$f(x) = \begin{cases} 120/x^2 & , \quad x > a \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Determinar el valor de a.

(b) Si una planta particular todavía vive después de 120 días, ¿cuál es la probabilidad que dicha planta viva más de 150 días?

(c) Si se observan 3 plantas de esta especie, ¿cuál es la probabilidad de que las tres vivan más de 150 días? ¿Cuál es la probabilidad que ninguna vive más de 150 días?

(d) ¿Cuál es la probabilidad que exactamente 2 viven más de 150 días?

(e) Determinar la función de distribución de X y su gráfica.

(f) ¿Cuál es la probabilidad que al menos uno de las tres plantas viven más de 150 días?

28. Una estación de suministro recibe gasolina una vez por semana. Si su volumen de ventas, en miles de galones, se distribuye con una función de densidad

$$f(x) = \begin{cases} 5(1 - x)^4 & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

¿Cuál deberá ser la capacidad de su depósito a fin de que la probabilidad que se agote en una semana adeterminada sea 0.01?

29. Dada la función de densidad de una variable aleatoria continua X,

$$f(x) = \begin{cases} 0 & , \quad \text{para} \quad x \leq \frac{\pi}{6} \\ 3 \text{ sen } 3x & , \quad \text{para} \quad \frac{\pi}{6} < x \leq \frac{\pi}{3} \\ 0 & , \quad \text{para} \quad x > \frac{\pi}{3} \end{cases}$$

Determinar :

(a) la función de distribución acumulativa de probabilidad de X.

(b) la probabilidad que X tome valores entre $\left[\frac{\pi}{12}, \frac{\pi}{4}\right]$.

30. Sea X una variable aleatoria con función de distribución acumulada.

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ (\frac{x}{a})^2 & , \quad 0 \leq x \leq a \\ 1 & \quad x > a \end{cases}$$

donde a es una constante no especifica.

(a) ¿Qué valor (es) de a hace (n) de $F(x)$ una función de distribución.

(b) Calcular la $P[X > 1]$, $P[\frac{1}{2} < X < \frac{3}{2}]$.

31. Sea X una variable aleatoria que tiene una función de distribución se - muestra en la figura adjunta. Determinar de la gráfica.

(a) $F(x)$

(b) $P[X < 0.5]$

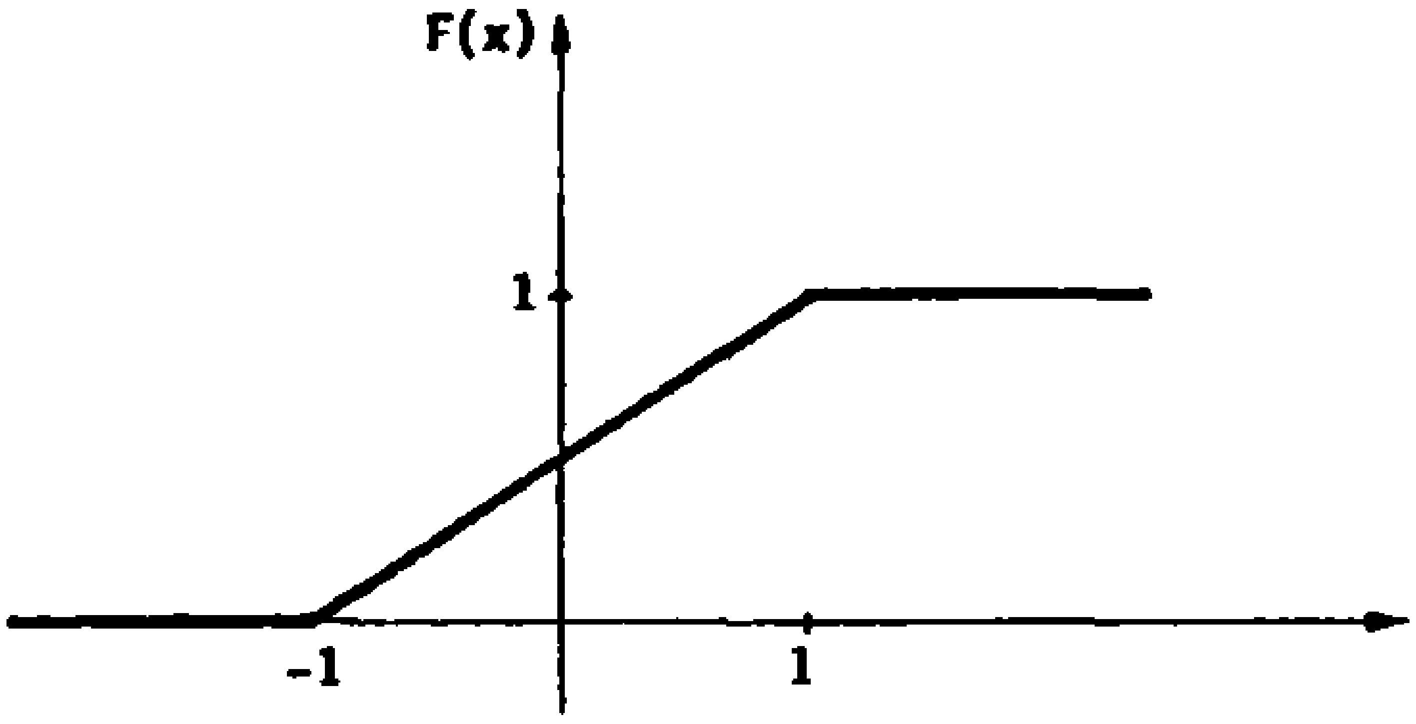
(c) $P(-0.2 < X < 0.8)$

(d) $P[X > 0.1]$

(e) $P[X < 3]$

(f) $P[X < -2]$

(g) $P[|X - 0.4| < 0.3]$



32. Sea
$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - \frac{3}{4} e^{x} & , \quad x \geq 0 \end{cases}$$

¿es su gráfica la que se muestra en la figura adjunta?
Calcular :

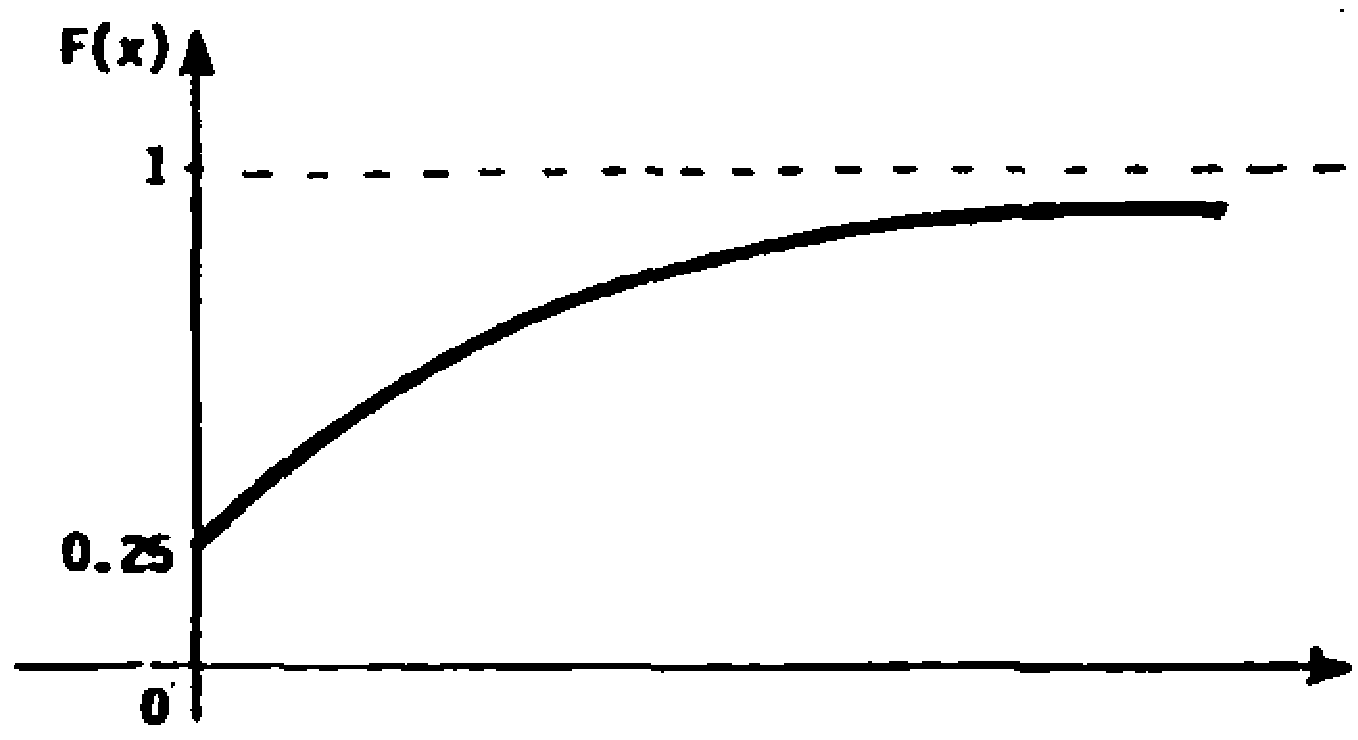
(a) $P[X > 2]$

(b) $P[X < 0]$

(c) $P[X = 0]$

(d) $P[X = 2]$

(e) $P[|X - 1| < 2]$



33. Supongamos que la vida (en horas) de cierto tipo de tubo de radio tenga una función de densidad
$$f(x) = \begin{cases} 100/x^2 & , \quad x > a \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Determinar el valor de a .

(b) Se toma una muestra aleatoria de tamaño 4 de tales tubos y se coloca en una caja. Si se selecciona aleatoriamente un tubo de la caja, ¿Cuál es la probabilidad que el tubo seleccionado tenga una duración mayor que 150 horas?

(c) ¿Cuál es el número máximo de tubos que se pueden formar en un aparato similar al de su uso, de modo que haya una probabilidad de 0.50 de que después de 150 horas de uso todavía funcionen?

2.4 DISTRIBUCIONES MIXTAS

En las secciones anteriores hemos presentado las variables aleatorias restringidas a discretas o contínuas, sin embargo hay aplicaciones en las cuales una variable aleatoria X puede tomar valores discretas, x_1, x_2, . . . con probabilidades positivas, y además tomar también todos los valores de un cierto intervalo $[a, b]$ o una colección de intervalos. Es decir, X es una combinación de los dos tipos de variables aleatorias. La distribución de probabilidad de este tipo de variables aleatorias, se llaman ***distribuciones mixtas***, se obtiene combinando la definición de distribución de probabilidad de una variable aleatoria discreta y la de función de densidad de una variable aleatoria contínua de la siguiente manera :

1. A cada uno de los valores x_i se le asigna un número $p(x_i) > 0$, y se define una función f tal que $f(x) > 0$, $\forall x \in [a, b]$ pués el rango (los valores que toma) de X es

$$R_X = \{x_1, x_2, \ldots, x_n\} \cup [a, b]$$

2. $$\sum_{x=1}^{n} p(x_i) + \int_a^b f(x)\, dx = 1$$

La probabilidad de un evento cualquiera A se define así,

$$P[A] = \sum_{x_j \in A} p(x_j) + \int_{E \subset A} f(x)\, dx$$

EJEMPLO 1 Dada la función

$$h(x) = \begin{cases} 2^{-(x+2)} & , \quad \text{para } x = 0,1,2,3, \ldots \\ \frac{1}{2}(1 - x) & , \quad \text{para } 0 < x < 1 \\ \frac{3}{4}(1 - x)^2 & , \quad \text{para } 1 < x < 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) verificar que cumple las condiciones de una función de probabilidad de una variable aleatoria X.

(b) Hallar la probabilidad de que X tome valores entre $\frac{1}{2}$ y 2 inclusive.

SOLUCION (a) Observe que el rango de la variable aleatoria es

$$R_X = \{0,1,2, \ldots\} \cup \langle 0,1\rangle \cup \langle 1,2\rangle \quad .$$

La función $h(x)$, para ser función de proababilidad debe cumplir

1. $h(x) = 2^{-(x+2)} > 0 \quad , \quad \forall\, x = 0,1,2,3, \ldots$

   y $\quad h(x) > 0 \quad , \quad \forall\, x \in \langle 0,1\rangle \cup \langle 1,2\rangle$

2. $$\sum_{x=0}^{\infty} 2^{-(x+2)} + \int_0^1 \frac{1-x}{2}\,dx + \int_1^2 \frac{3}{4}(1-x)^2\,dx = 2^{-2}\sum_{x=0}^{\infty} 2^{-x} +$$

$$+ \frac{1}{4} + \frac{1}{4} = \frac{1}{4}\left[1 + \frac{1}{2} + \frac{1}{2^2} + \ldots\right] + \frac{1}{2}$$

$$= \frac{1}{4}\left[\frac{1}{1-\frac{1}{2}}\right] + \frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1 \;.$$

de (1) y (2) obtenemos que la función $h(x)$ cumple las condiciones de una función de probabilidad de una variable aleatoria X.

(b) $$P\left[\frac{1}{2} < X \leq 2\right] = \sum_{x=1}^{2} 2^{-(x+2)} + \int_{1/2}^{1} \frac{1-x}{2}\,dx + \int_1^2 \frac{3}{4}(1-x)^2\,dx$$

$$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{16} + \frac{1}{4} = \frac{5}{8} \;.$$

PROBLEMAS 2.4

1. Dada la función

$$f(x) = \begin{cases} 3^{-(x+1)} & , \quad x = 0,1,2, \ldots \\ kx & , \quad 1 < x < 2 \\ k(4-x) & , \quad 2 < x < 3 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Determinar k de manera que f sea una función de probabilidad de una variable aleatoria X.

(b) Calcular $\quad P\left[\frac{1}{2} \leq X \leq \frac{3}{2}\right]$

2. Sea **X** una variable aleatoria con función de distribución

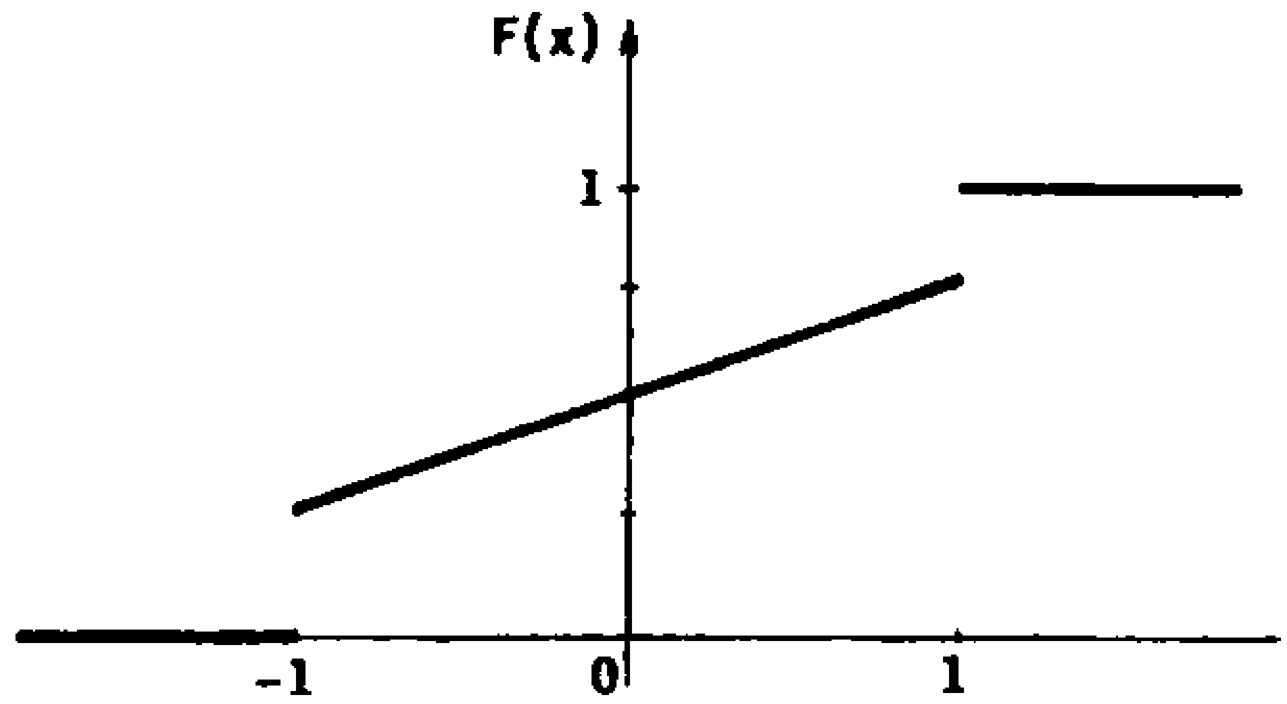
$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{4} & , \quad 0 \leq x < 1 \\ \frac{1}{2} & , \quad 1 \leq x < 2 \\ \frac{x}{3} & , \quad 2 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

Calcular :

(a) $P[0 < X \leq 1]$, (b) $P[1 \leq X \leq 2]$

(c) Hallar la función de densidad de probabilidad.

3. De la gráfica de la función de distribución



Determinar :

(a) $P[X < 1]$;

(b) $P[-1 \leq X < \frac{1}{2}]$;

(c) la función de densidad de probabilidad de **X**.

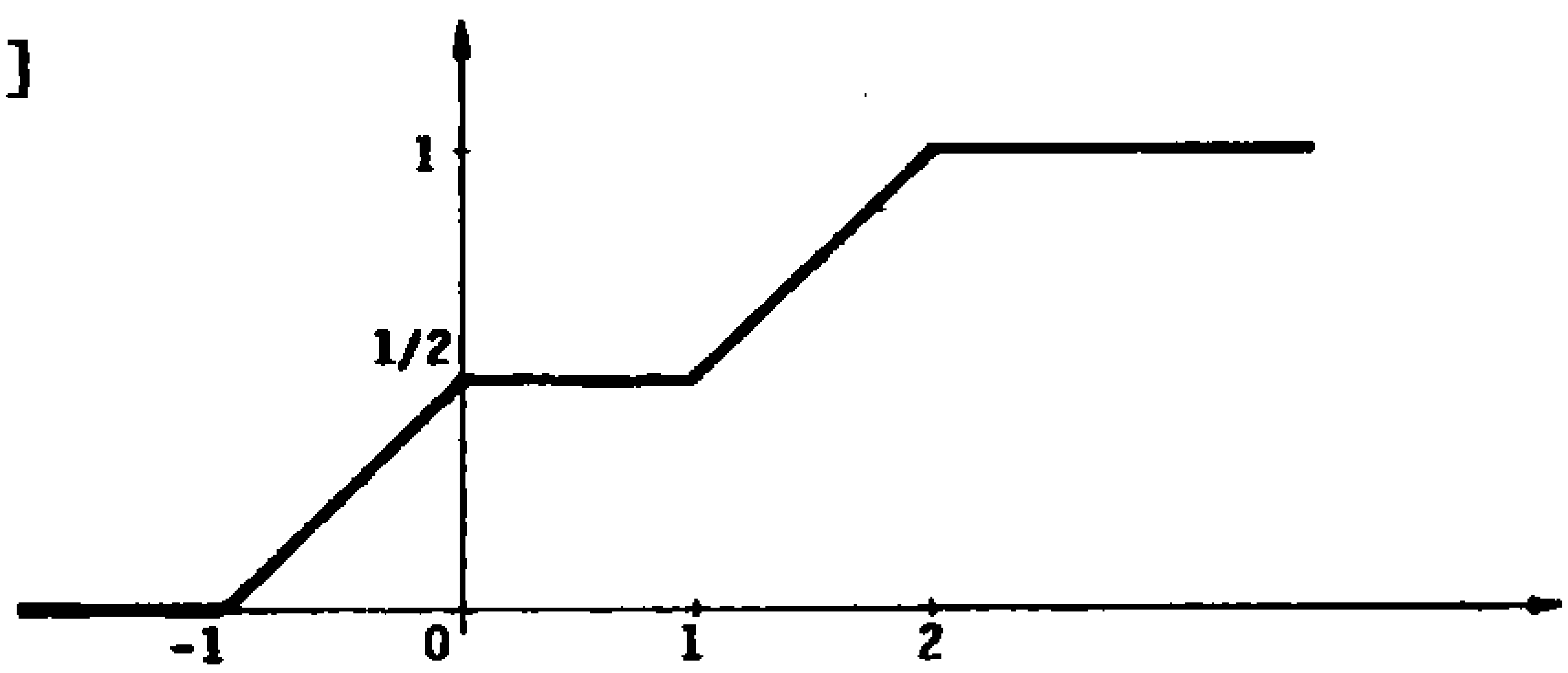
4. Una variable aleatoria **X** tiene función de distribución cuyo gráfico se muestra en la figura adjunta. Determinar :

(a) $P[-\frac{1}{2} < x < 2]$

(b) $P[|X - 1| < 1]$

(c) $P[X > 1]$

(d) $P[X = -\frac{1}{2}]$



3

ESPERANZA MATEMATICA

3.1 FUNCION DE UNA VARIABLE ALEATORIA

En muchos problemas de ingeniería, administración, etc, frecuentemente se está interesado en el comportamiento de una función, digamos H, de una variable aleatoria X, $Y = H(X)$; y cada valor de Y (esto es $y = H(x)$), se determina conociendo los valores de X. Desde que X es una variable aleatoria, Y también es una variable aleatoria, y la distribución de $Y = H(X)$ se determina conociendo la distribución de probabilidad de X. Antes, daremos un ejemplo simple con el objeto de hacer más comprensible la formalización .

EJEMPLO 1 Sea X una variable aleatoria discreta con función de probabilidad $p(x)$ como se indica en la siguiente tabla.

x	- 1	0	1	2
$p(x) = P[X = x]$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Determinar :

(a) El rango de la variable aleatoria $Y = X^2$,

(b) La función de probabilidad de la variable aleatoria $Y = X^2$.

(c) Lo mismo que en (*a*) y (*b*) para la variable aleatoria $Y = 3X - 1$.

SOLUCION

(a) Según lo expresado en el párrafo anterior, los valores de la variable aleatoria Y se determina reemplazando los valores de X en la función $H(X)$ ($y = H(x)$). Así, cuando la variable aleatoria X asume el valor $x = -1$, $Y = X^2$ toma el valor $y = H(-1) = (-1)^2 = 1$.

Si X toma el valor $x = 0$; $Y = X^2$, asume el valor $y = H(0) = (0)^2 = 0$

Si X toma el valor $x = 1$; $Y = X^2$, asume el valor
$y = H(1) = 1^2 = 1$ y

Si X asume el valor $x = 2$; $Y = X^2$, toma el valor
$y = H(2) = 2^2 = 4$

por lo tanto, el rango para la variable aleatoria Y es $R_Y = \{0,1,4\}$

(b) La función de probabilidad se calcula como sigue :

$$p_Y(0) = P[Y = 0] = P[X = 0] = p(0) = \frac{3}{8}$$

$$p_Y(1) = P[Y = 1] = P[X = -1] + P[X = 1] = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$$

$$p_Y(4) = P[Y = 4] = P[X = 2] = p(2) = \frac{1}{8}$$

Así, la distribución de probabilidad de $Y = X^2$ representando en una tabla es,

y	0	1	4
$p_Y(y)$	$\frac{3}{8}$	$\frac{4}{8}$	$\frac{1}{8}$

(c) En este caso $Y = H(X) = 3X - 1$. Los valores de Y se determinan de manera análoga al caso anterior.

Si la variable aleatoria X toma el valor $x = -1$, entonces, la variable aleatoria Y toma el valor :

$$y = H(-1) = 3(-1) - 1 = -4$$

Cuando $x = 0$, $y = H(0) = 3(0) - 1 = -1$

Si $x = 1$, $y = H(1) = 3(1) - 1 = 2$ y

Si $x = 2$, $y = H(2) = 3(2) - 1 = 5$

por lo tanto, $R_Y = \{-4, -1, 2, 5\}$.

La función de probabilidad se calcula de la siguiente manera :

$$p_Y(-4) = P[Y=-4] = P[X=-1] = p(-1) = \frac{1}{8}$$

$$p_Y(-1) = P[Y=-1] = P[X=0] = p(0) = \frac{3}{8}$$

$$p_Y(2) = P[Y=2] = P[X=1] = p(1) = \frac{3}{8}$$

$$p_Y(5) = P[Y=5] = P[X=2] = p(2) = \frac{1}{8}$$

Luego, la distribución de probabilidad está dado en la tabla

y	-4	-1	2	5
$p_Y(y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Presentaremos ahora una formulación más precisa de los conceptos anteriores.

3.1.1 EVENTOS EQUIVALENTES

Consideremos un experimento aleatorio ε con espacio muestral Ω. Sea X la variable aleatoria definida en Ω, con rango R_X. Si $Y = H(X)$ está definida de manera que los valores $y = H(x)$ en R_y (el rango de Y) son números reales, entonces Y es una variable aleatoria puesto que para cualquier suceso $\omega \in \Omega$, se determina un valor y de la función $Y = H(X)$; es decir $y = H(X(\omega))$. El diagrama de la figura 3.1.1 ilustra esto.

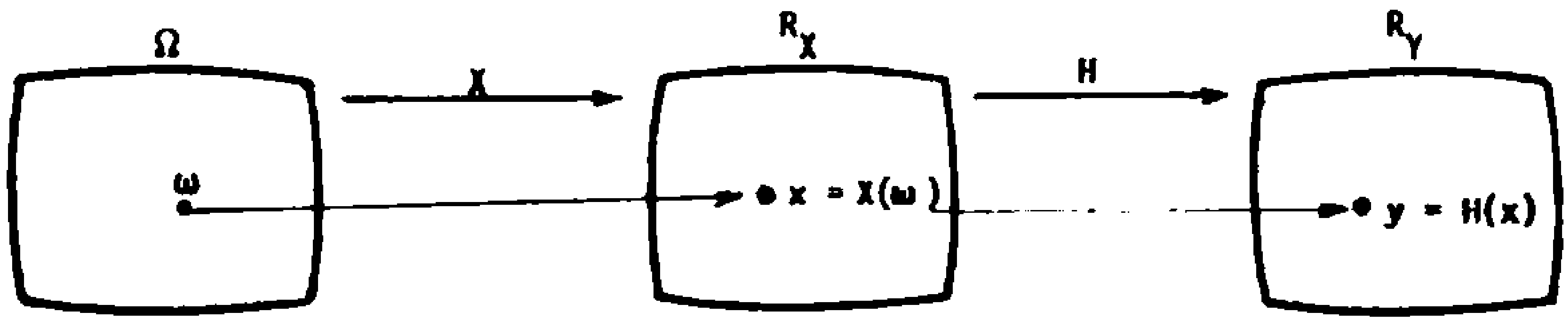


Fig. 3.1.1. Una función de una variable aleatoria $y = H(x) = H(X(\omega))$

En el capítulo anterior hemos definido eventos equivalentes en Ω y en R_X. Extenderemos ahora este concepto de la siguiente manera natural.

DEFINICION 3.1.1 Sea E_X un evento en $R_X (E_X \subset R_X)$, y F_Y un evento en R_Y $(F_Y \subset R_Y)$. E_X y F_Y son eventos equivalentes si

$$E_X = \{x \in R_X \mid H(x) \in F_Y\}$$

Informalmente hablando, los eventos E_X y F_Y son equivalentes si la ocurrencia de uno implica la ocurrencia del otro, es decir cuando ocurre E_X ocurre F_Y y recíprocamente. Supongamos ahora que A es un evento en Ω, el cual es equivalente a un evento E_X en R_X, entonces A y F_Y en R_Y son equivalentes.

NOTA Es importante notar que cuando hablamos de eventos equivalentes (en el sentido anterior) estos eventos están asociados a espacios diferentes.

EJEMPLO 2 Supongamos que $Y = H(X) = X^2$, donde X es una variable aleatoria definida como en el ejemplo 1. Entonces, los eventos $E_X = \{-1,1\}$ y $F_Y = \{1\}$ son equivalentes.

EJEMPLO 3 Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} e^{-x} & , \quad x > 0 \\ 0 & \quad\; x \leq 0 \end{cases}$$

Sea $Y = H(X) = 3X + 2$ una función de la variable aleatoria X. Por lo tanto $R_X = \{x \in \mathbb{R} \,/\, x > 0\}$; el rango de Y, se calcula así,

$y = H(x) = 3x + 2$, observe que cuando x tome valores mayores que cero, y toma valores mayores que 2; es decir

$$R_Y = \{y \in R_Y \,/\, y > 2\}$$

Supongamos ahora el evento $F_Y = (Y \geq 8)$, se tiene que $y \geq 8$ si, sólo sí $3x + 2 \geq 8$, lo que da $x \geq 2$.
Por lo tanto los eventos $F_Y = (Y \geq 8)$ y $E_X = (X \geq 2)$ son equivalentes.

DEFINICION 3.1.2 Sea X una variable aleatoria definida en el espacio muestral Ω con rango R_X. Sea H una función real, de modo que $Y = H(X)$ es una variable aleatoria con rango R_Y, entonces para cualquier evento $F_Y \subset R_Y$, definimos $P[F_Y]$ como sigue

$$P[F_Y] = P[\{x \in R_X \,/\, H(x) \in F_Y\}] = P[E_X]$$

En palabras: la probabilidad de un evento en el rango de Y está definida como la probabilidad del evento equivalente en R_X. Observe que esta probabilidad relaciona probabilidades en el espacio muestral Ω. Esto es

$$P[F_Y] = P[\{\omega \in \Omega \,/\, H[X(\omega)] \in F_Y\}]$$

La definición de probabilidad da el método a usarse en la solución de los problemas; para el cual se debe hallar el evento E_X en R_X que es equivalente el evento F_Y en R_Y, y luego se halla la probabilidad del evento E_X.

EJEMPLO 4 Sea X una variable aleatoria cuya función de probabilidad es

$$p(x) = \begin{cases} \dfrac{e^{-a}\, a^x}{a!} & , \quad x = 0,1,2,\ldots \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Si $Y = H(X) = 2X + 2$. Calcular $P[Y < 5]$, $P[Y > 4]$

SOLUCION

(a) Debemos calcular $P[Y \le 5]$; para lo cual calcularemos el evento equivalente en R_X.

$y \le 5$ se cumple sí, sólo sí $2x + 2 \le 5$ ó $x \le 3/2$ por lo tanto, el evento equivalente en R_X es, $(X \le 3/2)$

Luego,

$$P[Y \le 5] = P[2X + 2 \le 5] = P[X \le 3/2]$$

$$= p(0) + p(1) = e^{-a} + \frac{e^{-a}a}{1!} = e^{-a}[1 + a]$$

Obsérvese que el conjunto $\{x \in R_X \,/\, x \le \frac{3}{2}\}$ está en el rango de la variable aleatoria X. y la función p está definida en R_X.

(b) En este caso el evento equivalente a $y > 4$ es $2x + 2 > 4$, o sea $x > 1$. Luego,

$$P[Y > 4] = P[2X + 2 > 4] = P[X > 1] = 1 - P[X \le 1]$$

$$= 1 - e^{-a}[1 + a]$$

EJEMPLO 5 Sea X una variable aleatoria que tiene la siguiente función de densidad.

$$f(x) = \begin{cases} \dfrac{x}{8} & , \quad 0 \le x \le 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Si $Y = 2X + 10$. Calcular $P[12 \le y \le 16]$

SOLUCION Tenemos que el evento en R_Y es

$$\{y \in R_Y \,/\, 12 \le y \le 16\}$$

(el rango de Y es $R_Y = \{y \,/\, 10 \le y \le 18\}$)

El evento equivalente en R_X se calcula así;

$12 \le y \le 16$ sí sólo sí $12 \le 2X + 10 \le 16$

de donde $1 \le x \le 3$

y el conjunto $\{x \in R_X \,/\, 1 \le x \le 3\}$ está en el rango de X, y la función de densidad $f(x)$ está definida en este conjunto. Luego,

$$P[12 \le Y \le 16] = P[12 \le 2X + 10 \le 16] = P[1 \le X \le 3]$$
$$= \int_1^3 \frac{x}{8}\,dx = \left.\frac{x^2}{16}\right|_1^3 = \frac{9}{16} - \frac{1}{16} = \frac{1}{2}\,.$$

3.1.2 FUNCIONES DISCRETAS DE UNA VARIABLE ALEATORIA

PRIMER CASO Sean X e Y variables aleatorias discretas tal que

$$R_X = \{x_1, x_2, x_3, \ldots\} \qquad y \qquad p_X(x_i) = P[X = x_i]$$

la función de probabilidad de X. Supongamos que $x_{i_1}, x_{i_2}, x_{i_3}, \ldots$ representa los valores de X tal que $H(x_{i_j}) = y_i$ para algún conjunto de índices $J = \{j \,/\, j = 1,2,3,\ldots, S_i\}$. La distribución de probabilidad para Y - se denota por $p_Y(y_i)$ y está dado por

$$p_Y(y_i) = P[Y = y_i] = \sum_{y \in J} p_X(x_i)$$

por ejemplo, en la figura 3.1.2. $S_i = 4$, la probabilidad de y_i es

$$p_Y(y_i) = p_X(X_{i_1}) + p_X(X_{i_2}) + p_X(X_{i_3}) + p_X(X_{i_4})$$

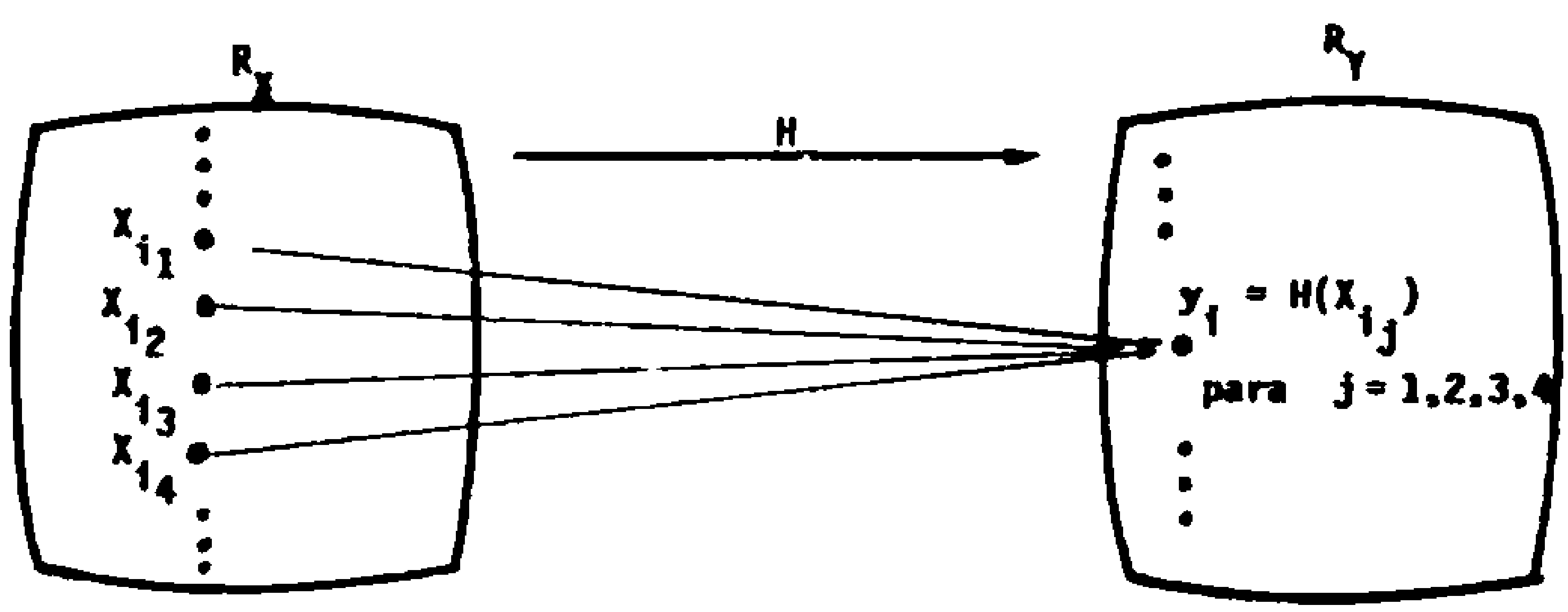


Fig. 3.1.2. Probabilidades en R_Y.

En el caso especial donde H es una función tal que para cada valor de y le hace corresponder exactamente un valor x, entonces la distribución de probabilidades de Y es

$$p_Y(Y_i) = p_X(x_i),\ \text{donde} \quad y_i = H(x_i),\ i = 1,2,\ldots$$

EJEMPLO 6 Consideremos el experimento de lanzar una moneda tres veces. Sea X el número de caras que se obtiene. Si, $Y = 2X + 1$. Calcular la distribución de probabilidad de la variable aleatoria Y.

SOLUCION X = número de caras.

$R_X = \{0,1,2,3\}$, con probabilidades $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$ y $\frac{1}{8}$ respectivamente.

Si $x = 0$, $y = 2(0) + 1 = 1$

$x = 1$, $y = 2(1) + 1 = 3$

$x = 2$, $y = 2(2) + 1 = 5$

Si $x = 3$, $y = 2(3) + 1 = 7$

por lo tanto, $R_Y = \{1,3,5,7\}$; como indica la figura 3.1.3

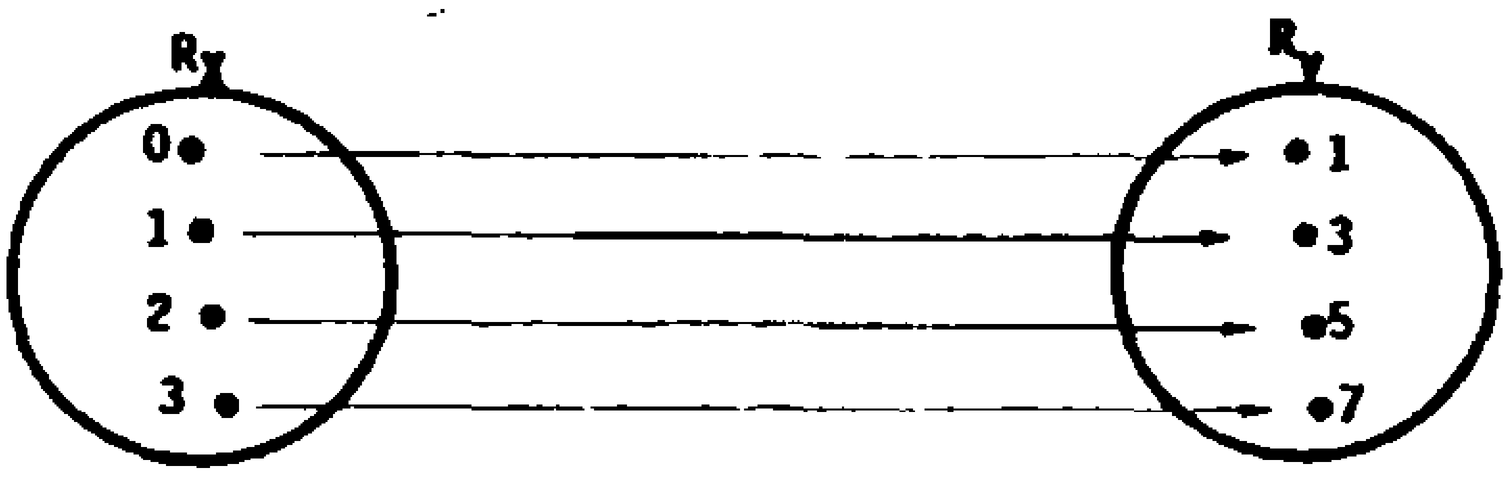


Fig. 3.1.3

En este caso la probabilidad en cada punto del rango de R_Y, es :

$$p_Y(1) = p_X(0) = \frac{1}{8} \quad ; \quad p_Y(5) = p_X(2) = \frac{3}{8}$$

$$p_Y(3) = p_X(1) = \frac{3}{8} \quad ; \quad p_Y(7) = p_X(3) = \frac{1}{8}$$

Luego, la distribución de probabilidad de Y es

y	1	3	5	7
$p_Y(y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

EJEMPLO 7 En el ejemplo anterior, supongamos que $Y = |X - 2|$. Calcular la distribución de probabilidad de Y.

SOLUCION Los valores de Y, ahora están dados por la ecuación $y = |x - 2|$. Es decir, los posibles valores de Y son 0,1,2. Como indicamos en la figura 3.1.4.

y $\quad p_Y(0) = p_X(2) = \frac{3}{8}$

$p_Y(1) = p_X(1) + p_X(3)$

$= \frac{1}{8} + \frac{3}{8} = \frac{4}{8}$

$p_Y(2) = p_X(0) = \frac{1}{8}$

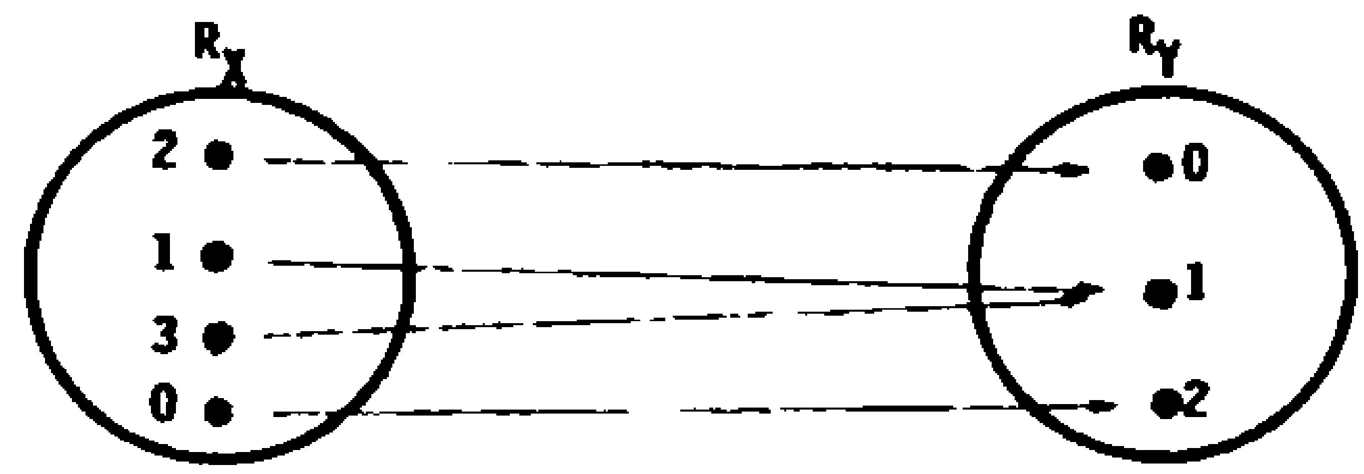


Fig. 3.1.4

entonces la distribución de probabilidad para Y es,

y	0	1	2
$p_Y(y)$	$\frac{3}{8}$	$\frac{4}{8}$	$\frac{1}{8}$

SEGUNDO CASO Cuando la variable aleatoria X es contínua, pero Y = H(X) es discreta, la formulación para la función de probabilidad de Y, $p_Y(y_i)$ está dado como sigue,

$$p_Y(y_i) = \int_{E_X} f(x)\,dx$$

donde E_X es el evento en R_X, que es equivalente al evento $(Y = y_i)$ en R_Y y $f(x)$ es la función de densidad de probabilidad asociada a la variable aleatoria X.

EJEMPLO 8 Supongamos que X es una variable aleatoria que tiene una función de densidad dada por :

$$f(x) = \begin{cases} a e^{-ax} & , \quad x \geq 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Sea Y = H(X), definida de la siguiente manera

$$Y = \begin{cases} 0 & , \quad x \leq 1/a \\ 1 & , \quad x > 1/a \end{cases}$$

Calcular la función de probabilidad de Y .

SOLUCION Se tiene que $R_Y = \{0,1\}$ Vemos que $y = 0$, sí, sólo sí - $x \leq 1/a$ por lo que el evento $E_X = \{x \,/\, X \leq 1/a\}$ es equivalente al evento $(Y = 0)$ en R_Y. Por lo tanto

$$p_Y(0) = P[X \leq 1/a\,] = \int_0^{1/a} ae^{-ax}dx = -e^{-ax}\Big|_0^{1/a} = 1 - e^{-1}.$$

También el evento $E_X = \{x \,/\, X > 1/a\}$ es equivalente al evento $(Y = 1)$ en R_Y. Luego

$$p_Y(1) = P[X > 1/a\,] = \int_{1/a}^{\infty} ae^{-ax}dx = -e^{-ax}\Big|_{1/a}^{\infty} = e^{-1}.$$

3.1.3 FUNCIONES CONTINUAS DE UNA VARIABLE ALEATORIA CONTINUA

Si X es una variable aleatoria contínua con función de densidad f, y H una función contínua; entonces $Y = H(X)$ es una variable aleatoria contínua La función de densidad de probabilidad de la variable aleatoria Y denotado por g, se obtiene siguiendo el siguiente procedimiento :

(1) Se obtiene la función de distribución G de Y, es decir

$$G(y) = P[Y \leq y\,]$$

Encontrando el evento E_X en R_X, equivalente al evento $(Y \leq y)$ en R_Y.

(2) Derivando $G(y)$ con respecto a y se obtiene $g(y)$.

(3) Se halla el rango de la nueva variable aleatoria con la condición que - $g(y) > 0$.

Ilustraremos estos tres pasos en los siguientes ejemplos.

EJEMPLO 9 Supóngase que X es una variable aleatoria con función de densidad .

$$f(x) = \begin{cases} x/10 & , \quad 0 \leq x \leq 2\sqrt{5} \\ 0 & , \quad \text{en otros casos.} \end{cases}$$

Si $Y = H(X) = 2X + 6$. Calcular la función de densidad g de Y.

SOLUCION El procedimiento dado en el párrafo anterior es,

(1) $$G(y) = P[Y \leq y\,] = P[2X + 6 \leq y] = P[X \leq \frac{y-6}{2}]$$

$$= \int_0^{\frac{y-6}{2}} \frac{x}{10}\,dx = \frac{x^2}{20}\Big|_0^{\frac{y-6}{2}} = \frac{1}{80}(y^2 - 12y + 36)$$

donde los eventos $(Y \leq y)$ en R_Y y $\left[X \leq (\frac{y-6}{2})\right]$ en R_X son equivalentes.

(2) $\quad g(y) = \frac{d}{dy}\,G(y) = \frac{y}{40} - \frac{3}{20}$.

(3) $0 \leq x \leq 2\sqrt{5}$, Entonces: $6 \leq 2x + 6 \leq 4\sqrt{5} + 6$, ó $6 \leq y \leq 4\sqrt{5} + 6$

por lo tanto la función de densidad de Y es :

$$g(y) = \begin{cases} \frac{y}{40} - \frac{3}{20} & , \quad 6 \leq y \leq 4\sqrt{5} + 6 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

EJEMPLO 10 Consideremos la variable aleatoria X definida en el ejemplo 5, y supongamos que $Y = H(X) = (X - 2)^2$. Hallar la función de densidad de Y.

SOLUCION Procediendo como en el ejemplo anterior obtenemos :

(1) $$G(y) = P[Y \leq y] = P[(X-2)^2 \leq y] = P[-\sqrt{y} \leq X - 2 \leq \sqrt{y}]$$

$$= P[2 - \sqrt{y} \leq X \leq 2 + \sqrt{y}]$$

$$= \int_{2-\sqrt{y}}^{2+\sqrt{y}} \frac{x}{8}\,dx = \left.\frac{x^2}{16}\right|_{2-\sqrt{y}}^{2+\sqrt{y}}$$

$$= \frac{1}{16}[(4 + 4\sqrt{y} + y) - (4 - 4\sqrt{y} + y)] = \frac{1}{2}\sqrt{y}$$

donde los eventos $(Y < y)$ en R_Y y $(2 - \sqrt{y} \leq X \leq 2 + \sqrt{y})$ en R_X son equivalentes.

(2) $\quad g(y) = G'(y) = \frac{1}{4\sqrt{y}}$

(3) $0 \leq x \leq 4$, entonces, Si $x = 2$, $y = 0$ y
Si $x = 0$ ó $x = 4$, $y = 4$

Como quiera que g, no está definida en $y = 0$, se tiene que el rango de, Y es, $0 < y \leq 4$, por lo tanto la función de densidad de Y es,

$$g(y) = \begin{cases} \frac{1}{4\sqrt{y}} & , \quad 0 < y \leq 4 \\ 0 & , \quad \text{en otros casos.} \end{cases}$$

TEOREMA 3.1.1 Sea X una variable aleatoria contínua con función de densidad f que satisface $f(x) > 0$ para $a < x < b$, e $y = H(x)$ es una función contínua de x, estrictamente creciente o estrictamente decreciente, entonces - la variable aleatoria $Y = H(X)$ tiene una función de densidad g dada por,

$$g(y) = f(x) \left| \frac{dx}{dy} \right|$$

Con $x = H^{-1}(y)$ expresado en función de y. Si H es creciente, $g(y) > 0$ si $H(a) < y < H(b)$; y si H es decreciente, $g(y) > 0$ si $H(b) < y < H(a)$.

DEMOSTRACION Daremos solamente para cuando H es estrictamente creciente. - similarmente se demuestra cuando H es decreciente.

(1) $$G(y) = P[Y \leq y] = P[H(X) \leq y]$$

$$= P[X < H^{-1}(y)] \,, \quad H^{-1} \text{ existe, pués H es estrictamente creciente.}$$

$$= F[H^{-1}(y)]$$

(2) $$g(y) = G'(y) = \frac{dF(x)}{dx} \cdot \frac{dx}{dy} \quad \text{por la regla de la cadena}$$

$$= f(x) \cdot \frac{dx}{dy} \,, \quad \text{donde} \quad x = H^{-1}(y)$$

Similarmente cuando H es estrictamente decreciente.

$$g(y) = G'(y) = - f(x) \frac{dx}{dy}$$

por lo tanto $g(y) = f(x) \left| \frac{dx}{dy} \right|$

EJEMPLO 11 En el ejemplo 9, donde X es una variable aleatoria con función de densidad

$$f(x) = \begin{cases} x/10 & , \quad 0 \leq x \leq 2\sqrt{5} \\ 0 & , \quad \text{en otros casos} \end{cases}$$

y $H(x) = 2x + 6$ es una función estrictamente creciente.
Luego, usando el teorema 3.1.1, tenemos

$$g(y) = f(x) \left| \frac{dx}{dy} \right| = \left(\frac{y-6}{20}\right) \left(\frac{1}{2}\right)$$

ya que $$x = \frac{y-6}{2} \,, \quad \frac{dx}{dy} = \frac{1}{2} \,, \quad f(x) = f\left(\frac{y-6}{2}\right) = \frac{y-6}{20}$$

$H(0) = 6, \quad H(2\sqrt{5}) = 4\sqrt{5} + 6$; entonces

$$f(y) = \begin{cases} \dfrac{y}{40} - \dfrac{3}{20} & , \quad 6 < y < 4\sqrt{5} + 6 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

PROBLEMAS 3.1

1. Una variable aleatoria **X** tiene la siguiente distribución de probabilidad

x	- 1	0	1
$p(x)$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{2}$

Determinar :

(a) la distribución de probabilidad de la variable aleatoria $Y = 2X + 1$.

(b) la distribución de probabilidad de la variable aleatoria $Y = X^2 - 1$.

(c) la distribución de probabilidad de la variable aleatoria $Y = (X - 1)^2$.

2. Suponga que la demanda diaria de gasolina en una estación de servicio está acotado por 1,000 galones, que se lleva a un registro diario de venta Cada galón vendido produce una utilidad de 60 soles mientras que cada galón no vendido produce una pérdida de 5 soles (debido a costo de almacenamiento). Si **X** designa la variable aleatoria que representa la demanda e **Y** la variable aleatoria utilidad, describa **Y** en función de **X**.

3. Un puerto tiene capacidad de acomodar 4 naves de cierto tipo durante la noche. Las tarifas del puerto producen una utilidad de 1,000 dólares por nave atracada. Si **X** es la variable aleatoria que representa el número de naves buscando atracadero por noche; y suponiendo que

$$g(x) = \frac{1}{6}\ , \quad \text{para} \quad x = 0,1,2,3,4,5$$

(a) Si **Y** es la variable aleatoria que representa las utilidades nocturnas, describa la variable aleatoria **Y** en términos de la variable aleatoria **X**.

(b) Determinar la distribución de probabilidad para **Y**.

4. La cantidad de magnesio en una mezcla es una variable aleatoria, la función de densidad de probabilidad está dado por

$$f(x) = \begin{cases} x/18 & , \quad 0 \le x \le C \\ 0 & , \quad \text{en otros casos} \end{cases}$$

El beneficio obtenido de esta mezcla es $P = 10 + 2x$ Hallar la función de densidad de **P**.

5. Suponga que **X** es una variable aleatoria con función de densidad,

$$f(x) = \begin{cases} 2x & , \quad 0 \leq x \leq 1 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

(a) Hallar la función de densidad de $Y = H(X) = 5x + 3$

(b) la función de densidad de $Y = H(X) = X^2$

6. Una variable **X** tiene la distribución de probabilidad siguiente

$$\begin{cases} \frac{x}{6} & , \quad x = 1,2,3 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

Si $Y = H(X) = (X - 2)^2$. Hallar la distribución de probabilidad de **Y**.

7. El número de días requeridos para la terminación de un proyecto de construcción se denota por **X** y se considera como una variable aleatoria con la distribución siguiente

x	10	11	12	13	14
$p(x)$	0.2	0.3	0.3	0.1	0.1

y $p(x) = 0$ en otros casos.

El beneficio de la contratistas es $Y = 1000(12 - X)$ intis.

Hallar la distribución de probabilidad de **Y**.

3.2 CARACTERISTICA DE UNA VARIABLE ALEATORIA

Si bien la variable aleatoria está completamente determinado por la distribución de probabilidad $[(x_i, p(x_i))\ ;\ i = 1,2,3, \ldots]$ Si es discreta, y por la función de densidad $f(x)$ si es continua, muchas veces es conveniente trabajar con algunas características descriptivas de la variable aleatoria. Introduciremos aquí las medidas descriptivas mas usadas; además una expresión general para otras medidas similares. El primero de estos es el primer momento alrededor del origen; llamado *el valor esperado o la media de la variable aleatoria* (*También Esperanza Matemática*). El segundo momento alrededor de la media, llamado *varianza* de la variable aleatoria. Finalmete la *moda* y la *mediana*.

3.2.1 VALOR ESPERADO DE UNA VARIABLE ALEATORIA

Antes de dar la definición del valor esperado de una variable aleatoria daremos un ejemplo simple para captar la idea intuitiva.

Consideremos por ejemplo el experimento de lanzar 3 monedas 20 veces y sea X el número de caras que ocurren por lanzamiento; entonces, X toma los siguientes valores, 0,1,2,3. Es decir $R_X = \{0,1,2,3\}$. Supongamos ahora que en los 20 lanzamientos de las tres monedas obtenemos 0 caras, 1 cara, 2 caras y 3 caras, un total de 4,6,7, y 3 veces respectivamente. El promedio del número de caras por lanzamiento de las 3 monedas es entonces

$$\frac{0\,(4)\ +\ 1\,(6)\ +2\,(7)\ +\ 3\,(3)}{20} = 0(\frac{4}{20}) + 1(\frac{6}{20}) + 2(\frac{7}{20}) + 3(\frac{3}{20})$$

$$= 1.45$$

este es un valor promedio y no necesariamente un posible resultado del experimento (ver figura 3.2.1). Los números $\frac{4}{20}$, $\frac{6}{20}$, $\frac{7}{20}$ y $\frac{3}{20}$ son las frecuencias relativas para los diferentes resultados posibles.

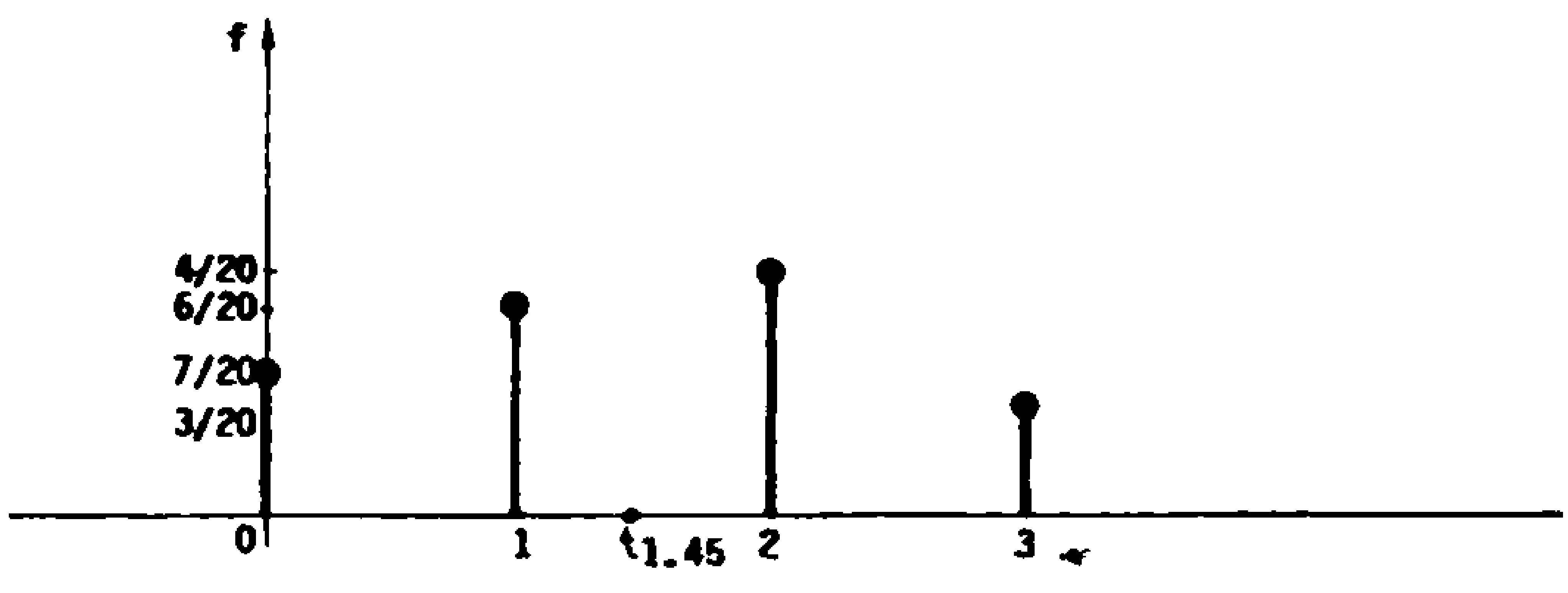


Fig. 3.2.1

Consideremos ahora el problema de calcular el número de caras por lanzamiento que podemos esperar a la larga (esto es cuando el experimento se repite indefinidamente). De la definición de probabilidad por frecuencia relativa, sabemos que a la larga aparecerá cara 1/8 de las veces ($P[X = 0] = P[\{SSS\}] = 1/8$) ; 1 cara 3/8 de veces ($P[X = 1] = P[\{CSS\}] + P[\{SCS\}] + P[\{SSC\}] = 3/8$); 2 caras 3/8, y 3 caras 1/8 de veces por lo tanto el número de caras esperado por lanzamiento a la larga *denotado por* $E(X) = \mu$

es ,

$$\mu = 0\left(\frac{1}{8}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{3}{8}\right) + 3\left(\frac{1}{8}\right) = 1.5$$

Como se indica en la figura 3.2.2 .

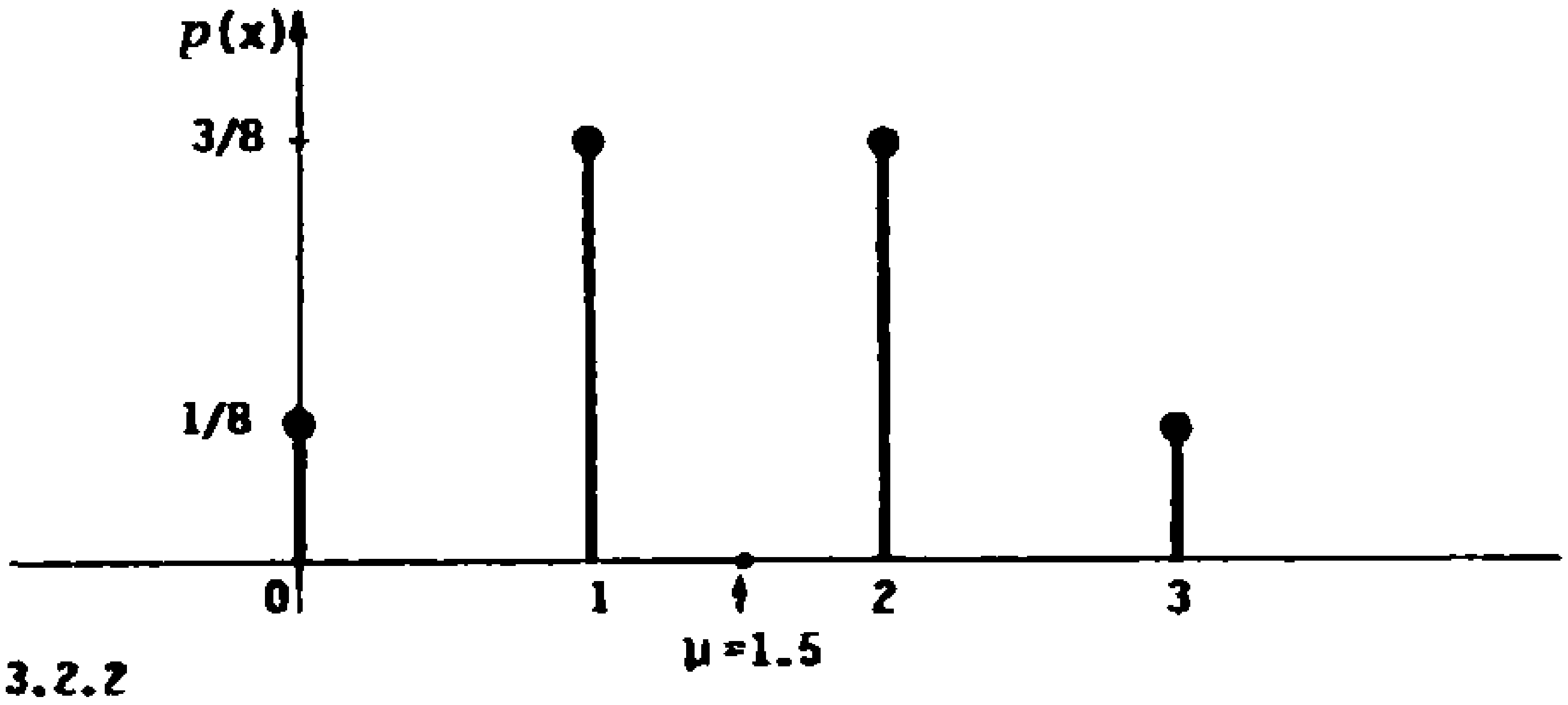


Fig. 3.2.2

La ilustración anterior sugiere que el promedio o valor esperado de cualquier variable aleatoria puede obtenerse multiplicando cada valor de la variable por sus correspondientes probabilidades y sumándoles.

DEFINICION 3.2.1 Sea X una variable aleatoria con rango R_X y función de probabilidad $p(x)$ si es discreta o función de densidad $f(x)$ si es continua. ***El valor esperado o esperanza matemática de X***, se denota por "$E(X)$" y se define

$$E(X) = \sum_{x \in R_X} x\, p(x) \text{ , Si } X \text{ es una variable aleatoria discreta.} \qquad (*)$$

$$E(X) = \int_{R_X} x\, f(x)\, dx = \int_{-\infty}^{\infty} x\, f(x) dx, \text{ si } X \text{ es una variable aleatoria continua} \qquad (**)$$

Siempre que $\sum_{x \in R_X} x\, p(x)$ sea absolutamente convergente. Es decir $\sum_{x \in R_X} |x| p(x)$- finita. y $\int_{-\infty}^{\infty} |x| f(x) dx$ finita respectivamente.

La esperanza matemática de X, se llama también, ***media de la variable aleatoria***, y se denota por μ, o sea

$$\mu = E(X) .$$

INTERPRETACION GEOMETRICA la esperanza matemática, tanto de una variable aleatoria discreta como continua, es igual a la abscisa del centro de gravedad del área acotada por la curva o el polígono de distribución de probabilidad y el eje de abscisas. Por eso cuando la curva o el polígono de distribución de probabilidad es simétrica respecto a cierta recta paralela al eje de ordenadas, la esperanza matemática coincide con la abscisa del punto de intersección de este eje de simetría con el eje de las obscisas.

EJEMPLO 1 Hallar la esperanza matemática de la variable aleatoria **X** con distribución de probabilidad dado por,

x	0	1	2	3	4
$p(x)$	0.2	0.4	0.3	0.08	0.02

SOLUCION Por la fórmula (*) encontramos la esperanza matemática

$$E(X) = 0(0.2) + 1(0.4) + 2(0.3) + 3(0.08) + 4(0.02) = 1.32 .$$

EJEMPLO 2 Sea **X** una variable aleatoria cuya función de densidad está definida por,

$$f(x) = \begin{cases} \frac{3}{4} x\,(2 - x) & , \quad 0 \le x \le 2 \\ 0 & , \quad \text{en otro caso} . \end{cases}$$

Calcular la esperanza de **X** .

SOLUCION por la fórmula (**) encontramos la esperanza matemática

$$\mu = E(X) = \int_{-\infty}^{\infty} \frac{3}{4} x^2(2 - x)\, dx = \int_0^2 \frac{3}{4} x^2(2 - x)\, dx$$

$$= \frac{3}{4}\left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_0^2 = 1 .$$

EJEMPLO 3 La función de distribución de una variable aleatoria **X** es,

$$F(x) = c\,[\frac{x^3}{3} + \frac{1}{3}] , \quad -1 \le x \le 1$$

Calcular, $E(X)$.

SOLUCION Se tiene $F(1) = c\left[\frac{1}{3} + \frac{1}{3}\right] = 1$, propiedad de $F(x)$. De donde

$c = \frac{3}{2}$; por lo tanto, $f(x) = \frac{3}{2}x^2$, $-1 \leq x \leq 1$. Luego

$$E(X) = \int_{-1}^{1} x(\tfrac{3}{2}x^2)\,dx = \frac{3}{2}\left.\frac{x^4}{4}\right|_{-1}^{1} = 0 .$$

EJEMPLO 4 Una fábrica ensambladora de televisores a colores recibe los transistores en lotes de 20. El departamento de recepción utiliza la siguiente regla de inspección : Se prueban 2 transistores de cada lote. Si ninguno es defectuoso, no se continúa probando los demás transistores. Si al menos uno de los transistores resulta defectuoso, se prueba el lote completo. ¿Cuál es el número esperado de transistores probados por lote si se sabe por experiencia que cada lote contiene exactamente el 25 % de defectuosos?

SOLUCION Definimos la variable aleatoria **X** por

$X(\omega)$ = Número de transistores probados.

$R_X = \{2,20\}$

Definimos los eventos :

D_i : "el i-ésimo transistor verificado es defectuoso, $i = 1,2$"

$\bar{D}_i$: "el i-ésimo transistor verificado es no defectuoso , $i = 1,2$"

Cada lote de 20 transistores, consta de 15 buenos y 5 defectuosos.
Cálculo de la distribución de probabilidad de **X** .

$$p(2) = P[X = 2] = P[\bar{D}_1\bar{D}_2] = P[\bar{D}_1]\ P[\bar{D}_2 \mid \bar{D}_1]$$

$$= \frac{15}{20} \times \frac{14}{19} = \frac{21}{38}$$

Puesto que en todos los otros casos se examina todo el lote se tiene que,

$$p(20) = P[X = 20] = 1 - P[X = 2] = 1 - P[\bar{D}_1\bar{D}_2]$$

$$= 1 - \frac{21}{38} = \frac{17}{38} .$$

Por lo tanto, la distribución de probabilidad de la variable aleatoria **X** es

x	2	20
$p(x)$	$\frac{21}{38}$	$\frac{17}{38}$

El número esperado de transistores probados por lote, es

$$\mu = E(X) = 2(\frac{21}{38}) + 20(\frac{17}{38}) = \frac{382}{38} = 10.05 .$$

EJEMPLO 5 Una urna contiene cuatro fichas numeradas; 1,2,3,4, y de ellas - se extraen dos fichas sin reposición. Si **X** es la variable aleatoria que representa la suma de los cuadrados de los dos números obtenidos, calcular - E(**X**).

SOLUCION La variable aleatoria **X** esta definida por

$X(\omega)$ = Suma de los cuadrados de los dos números obtenidos

R_X = {5,10,13,17,20,25}

	1	2	3	4
1	~~11~~	12	13	14
2	21	~~22~~	23	24
3	31	32	~~33~~	34
4	41	42	43	~~44~~

Cálculo de la distribución de probabilidad de **X** .

$$p(5) = P[X = 5] = \frac{2}{12} = \frac{1}{6}$$

$$p(10) = P[X = 10] = \frac{2}{12} = \frac{1}{6}$$

$$p(13) = P[X = 13] = \frac{2}{12} = \frac{1}{6}$$

por lo tanto,

$$p(x) = \frac{1}{6} \quad , \quad x = 5,10,13,17,20,25$$

es la función de probabilidad de la variable aleatoria **X**. Luego

$$E(X) = 5\ (\frac{1}{6}) + 10\ (\frac{1}{6}) + 13\ (\frac{1}{6}) + 17\ (\frac{1}{6}) + 20\ (\frac{1}{6}) + 25\ (\frac{1}{6})$$

$$= \frac{5 + 10 + 13 + 17 + 20 + 25}{6} = \frac{90}{6} = 15$$

EJEMPLO 6 Consideremos el experimento que consiste en lanzar un dado. Si - el dado muestra 1 ó 2 se selecciona un número al azar del conjunto {0,1}. - Si el dado muestra 3,4,5 ó 6 se selecciona un número al azar del conjunto - {2,3,4} .

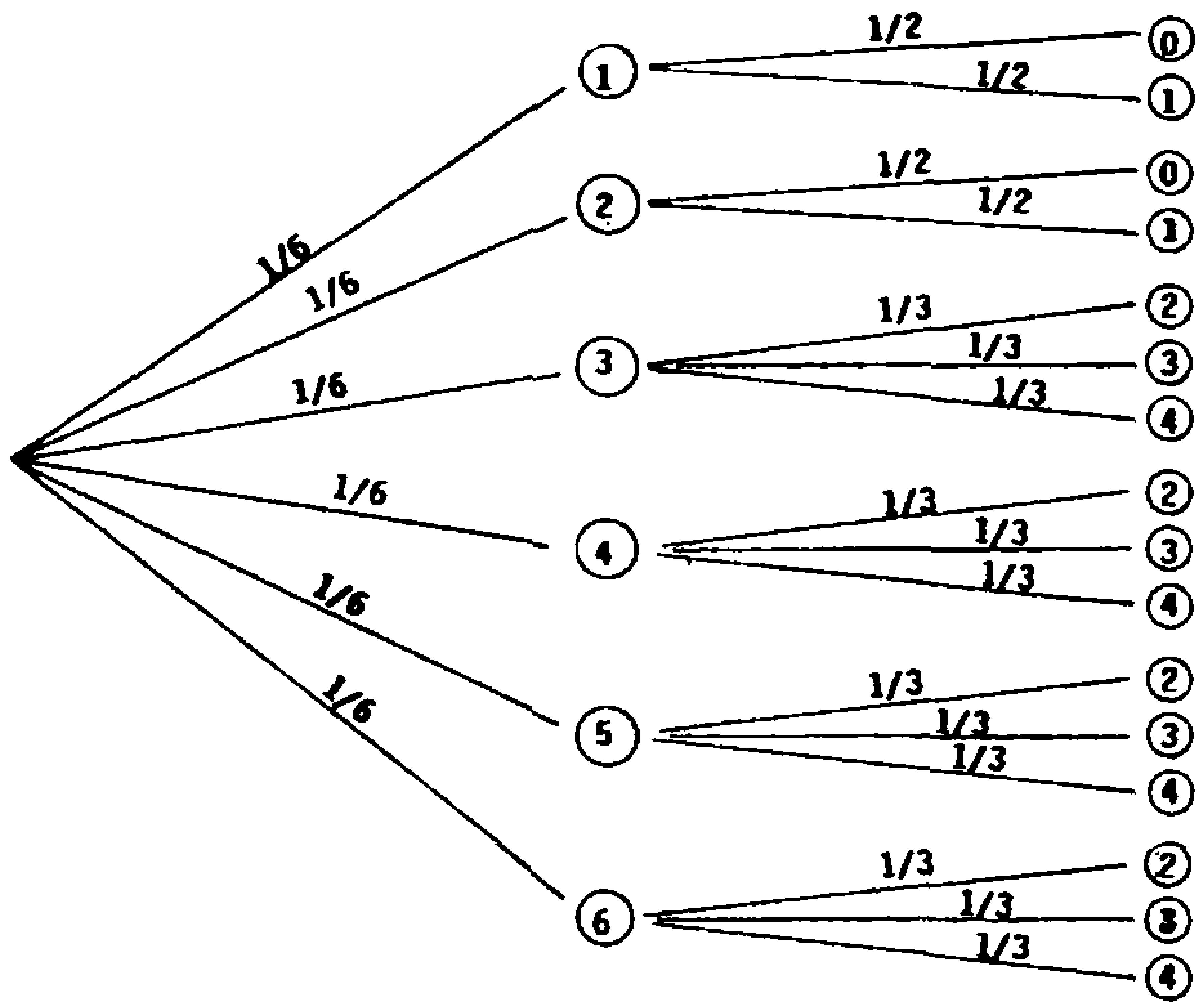
(a) Determinar la función de distribución de **X**.

(b) Hallar E(**X**). Donde **X** es la variable aleatoria que representa el número seleccionado.

SOLUCION Desde que **X** es el número seleccionado. Entonces,

$$R_X = \{0,1,2,3,4\}$$

Para determinar las probabilidades en cada punto, nos valemos del diagrama de probabilidades siguiente,



Del diagrama :

$$p(0) = P[X = 0] = \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{1}{2} = \frac{1}{6} = \frac{3}{18}$$

$$p(1) = P[X = 1] = \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{1}{2} = \frac{1}{6} = \frac{3}{18}$$

$$p(2) = P[X = 2] = \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} = \frac{4}{18} \; .$$

$$p(3) = P[X = 3] = \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} = \frac{4}{18}$$

$$p(4) = P[X = 4] = \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{3} = \frac{4}{18}$$

(a) por lo tanto, la distribución de probabilidad de **X** se da en la tabla siguiente, donde se observa también la función de distribución de **X**.

x	0	1	2	3	4
$p(x)$	$\frac{3}{18}$	$\frac{3}{18}$	$\frac{4}{18}$	$\frac{4}{18}$	$\frac{4}{18}$
$F(x)$	$\frac{3}{18}$	$\frac{6}{18}$	$\frac{10}{18}$	$\frac{14}{18}$	1

(b) La esperanza de X es,

$$E(X) = 0(\frac{3}{18}) + 1(\frac{3}{18}) + 2(\frac{4}{18}) + 3(\frac{4}{18}) + 4(\frac{4}{18}) = \frac{13}{6} = 2.17 .$$

NOTA Un juego se considera *equitativo* si la esperanza matemática o valor - esperado de la ganancia es cero.

EJEMPLO 7 Una urna contiene 4 bolas rojas, 6 negras, 8 verdes y 2 blancas. Se saca una bola de la urna. Si ésta es roja usted gana I/. 30.00 y si es negra usted gana I/. 20.00. ¿Cuánto debería pagar usted si saca una bola - verde y cuanto, si saca una bola blanca para que el juego fuera equitativo? Además, si saca una bola verde usted paga la cuarta parte de lo que paga - cuando saca una bola blanca.

SOLUCION Sea X la variable aleatoria definida por

$$X(\omega) = \text{ganancia del jugador} .$$

designemos las pérdidas asociados con la extracción de una bola verde, y - una bola blanca por L_1 y L_2 respectivamente. Entonces,

$$R_X = \{30,20, L_1, L_2\} , \quad \text{donde,} \quad 4L_1 = L_2$$

por lo tanto, la distribución de probabilidad de la variable aleatoria X es

x	30	20	L_1	L_2
$p(x)$	$\frac{4}{20}$	$\frac{6}{20}$	$\frac{8}{20}$	$\frac{2}{20}$

Luego,

$$E(X) = 30\ (\frac{4}{20}) + 20\ (\frac{6}{20}) + L_1\ (\frac{8}{20}) + L_2\ (\frac{2}{20})$$

$$= \frac{420}{20} + \frac{120}{20} + \frac{8L_1}{20} + \frac{2L_2}{20} = \frac{240}{20} + \frac{4L_2}{20} = 0$$

de donde, $L_2 = -\frac{240}{4} = -60$ y $L_1 = -15$.

Entonces, si usted paga I/. 15.00 , cuando la bola extraído es verde y I/. 60.00 cuando es blanca, el valor esperado del juego es cero y éste es equitativo.

EJEMPLO 8 En una lotería se ofrece cinco premios de la siguiente manera: un primer premio de I/. 25.00 , un segundo premio de I/. 10.00 y tres premios de I/. 5.00 cada uno. ¿Cuál debe ser el precio justo de los boletos si se venderán 100,000? ¿Cuál si se vendiera 1 millón?

SOLUCION Definimos la variable aleatoria X por

$$X(\omega) = \text{ganancia neta al comprar un boleto.}$$

Sea c el costo de un boleto en intis. Entonces

$$R_X = \{-c,\ 5-c,\ 10-c,\ 25-c\}$$

La distribución de probabilidad de X si se venden 100,000 boletos es

x	$-c$	$5-c$	$10-c$	$25-c$
$p(x)$	$\frac{99,995}{100,000}$	$\frac{3}{100,000}$	$\frac{1}{100,000}$	$\frac{1}{100,000}$

$$E(X) = \frac{-99,995c + 15 - 3c + 10 - c + 25 - c}{100,000} = \frac{50 - 100,000c}{100,000} = 0$$

$$\text{de donde } c = \frac{5}{10000} = 0.0005$$

El precio justo de cada boleto debe ser 0.0005 Intis.
La distribución de probabilidad de X si se vende 1'000,000 boletos es

x	$-c$	$5-c$	$10-c$	$25-c$
$p(x)$	$\frac{99995}{1'000,000}$	$\frac{3}{1'000,000}$	$\frac{1}{1'000,000}$	$\frac{1}{1'000,000}$

$$E(X) = \frac{-999,995c + 15 - 3c + 10 - c + 25 - c}{1'000,000} = \frac{50-1'000,000c}{1'000,000} = 0$$

$$\text{de donde } c = \frac{5}{100,000} = 0.00005$$

En este caso, el precio justo de cada boleto debe ser I/. 0.00005 .

EJEMPLO 9 (a) Calcule el peso medio de una papaya cosechado por el agricultor del ejemplo 8 de 2.3.

(b) Estime el peso total de la cosecha del agricultor.

SOLUCION

(a) por la fórmula (* *) el peso medio de una papaya es

$$\mu = \int_0^4 \frac{3x}{32}(4x - x^2)dx = \frac{3}{32}\left(\frac{4x^3}{3} - \frac{x^4}{4}\right)\Big|_0^4 = \frac{3}{32}\left(\frac{256}{3} - \frac{256}{4}\right) = 2$$

(b) Si N es el número total de papayas cosechadas, el peso total estimado de la cosecha del agricultor es

$$\mu N = 2(20000) = 40000 \text{ kg.}$$

EJEMPLO 10 Una florista estima su venta diaria de rosas en la forma siguiente,

Venta estimada diaria en docenas	Probabilidad de la venta estimada
12	0.5
13	0.4
14	0.1

La florista debe ordenar las rosas con un día de anticipación. Las rosas que no se venden en un día se pierden. Si el costo de las rosas es de \$ 100.00 por docena y su precio de venta es de \$ 300.00 por docena, ¿Cuántas docenas debe ordenar la florista para maximizar su ganancia diaria esperada?

SOLUCION La florista puede ordenar 12, 13 ó 14 docenas por día. La ganancia para cada "orden" posible depende del número de docenas que venda. Se puede considerar entonces tres variables aleatorias **X**, **Y** y **Z** donde :

X = ganancia neta (utilidad) de la florista al ordenar 12 docenas.
Y = ganancia neta (utilidad) de la florista al ordenar 13 docenas.
Z = ganancia neta (utilidad) de la florista al ordenar 14 docenas.

El rango de cada variable aleatoria (utilidad para cada orden posible), se calcula así,

utilidad = ingreso total - costo total

Cálculo del rango de **X** .

(1) Si vende 12 docenas

utilidad = 12 x 300.00 - 12 x 100.00 = 240,000

(2) Si vende 13 docenas .

El segundo valor del rango de X es también \$ 240,000, porqué la florista ordenó sólo 12 docenas aunque habría podido vender 13 docenas.

(3) Si vende 14 docenas.

De manera similar que en (2) aquí también la utilidad es de \$ 240,000.

Luego,

$$R_X = \{2,400, \ 2,400, \ 2,400\}$$

Cálculo del rango de Y

(1) si vende 12 docenas

$$\text{utilidad} = 12 \times 300.00 - 13 \times 100.00 = 2300.00$$

(2) si vende 13 docenas.

$$\text{utilidad} = 13 \times 300.00 - 13 \times 100.00 = 2,600.$$

(3) si vende 14 docenas.

Se obtiene también \$ 2,600, porque la florista ordenó sólo 13 docenas aunque podía haber vendido 14 docenas.

$$R_Y = \{2,300, \ 2,600, \ 2,600\}$$

Rango de Z .

Se obtiene de manera completamente análogo a los casos anteriores.

$$R_Z = \{2,200, \ 2,500, \ 2,800\}$$

Los resultados se resumen en la tabla siguiente :

Niveles posibles de utilidad

Probabilidad	X: ganan al ordenar 12 doc. R_X	Y: ganan al ordenar 13 doc. R_Y	Z: utilidad, al ordenar 14 doc. R_Z
0.5	\$ 2400	\$ 2,300	\$ 2,200
0.4	\$ 2400	\$ 2,600	\$ 2,500
0.1	\$ 2400	\$ 2,600	\$ 2,800

La utilidad (ganancia) esperada para cada una de las variables aleatorias - (niveles de orden) se indica en el siguiente cuadro,

X: utilidad al ordenar 12 docenas	Y: utilidad al ordenar 13 docenas	Z: utilidad al ordenar 14 docenas
2400 x 0.5 = 1200	2300 x 0.5 = 1150	2200 x 0.5 = 1100
2400 x 0.4 = 960	2600 x 0.4 = 1040	2500 x 0.4 = 1000
2400 x 0.1 = 240	2600 x 0.1 = 260	2800 x 0.1 = 280
E(X) = 2400	E(Y) = 2450	E(Z) = 2380

por lo tanto, la utilidad esperada es máxima, si la florista ordena 13 docenas por día.

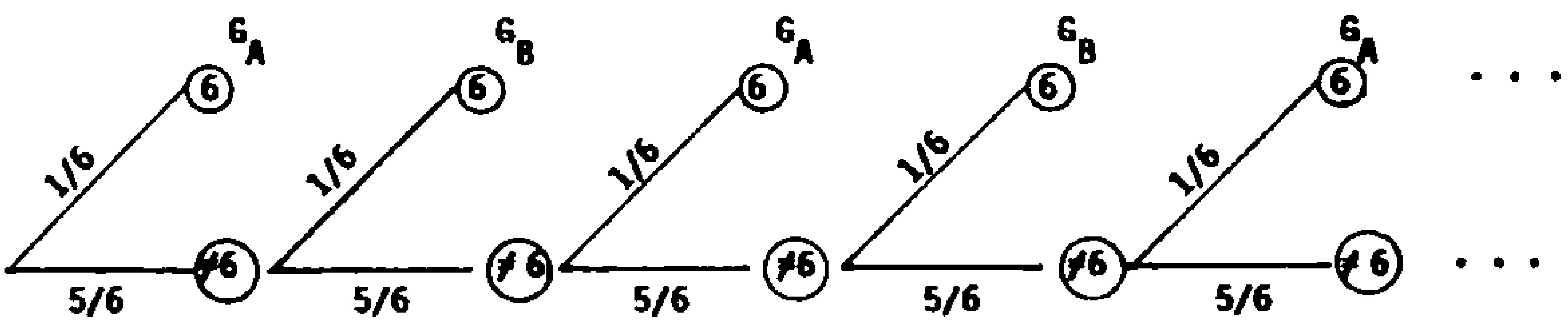
EJEMPLO 11 A y B lanzan alternativamente un dado común. El primero que saca un 6 gana. Si A lanza primero, determine usted cuál es su ganancia esperada si el premio es I/. 220 .

SOLUCION Sea X la variable aleatoria definida por

$$X(\omega) = \text{ganancia del jugador A.}$$

$$R_X = \{220, 0\}$$

Determinaremos ahora la distribución de probabilidad de X. Del diagrama de probabilidades siguiente, se obtiene



$$P[X = 220] = P[G_A] = \frac{1}{6} + \left(\frac{5}{6}\right)^2 \times \left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^4 \left(\frac{1}{6}\right) + \ldots$$

$$= \frac{1}{6}\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \ldots\right]$$

$$= \frac{1}{6}\left[\frac{1}{1 - \frac{25}{36}}\right] = \frac{6}{11} .$$

$$P[X = 0] = P[G_B] = \frac{5}{11} \quad \text{(de manera análoga que el anterior).}$$

por lo tanto, la distribución de probabilidad de X es,

x	0	220
$p(x)$	$\frac{5}{11}$	$\frac{6}{11}$

Luego, $E(X) = 0\,(\frac{5}{11}) + 220\,(\frac{6}{11}) = 120$.

La ganancia esperada del jugador A es de I/. 120.

3.2.2 PROPIEDADES DE LA ESPERANZA MATEMATICA

Daremos aquí algunas leyes útiles que simplificarán el cálculo de la esperanza matemática. Estas leyes o teoremas nos permitirán calcular esperanzas en términos de otras conocidas o esperanzas que son fáciles de calcular.

El siguiente teorema generaliza la definición de esperanza matemática a la de una función de la variable aleatoria $Y = H(X)$.

TEOREMA 3.2.1 Sea X una variable aleatoria, e $Y = H(X)$ una función de X El valor esperado de la función $H(X)$, denotado por "$E[H(X)]$", se define por

(*i*) $E[H(X)] = \sum_{x \in R_X} H(x)\,p(x)$, si X es discreta.

siempre que esta serie sea absolutamente convergente .

(*ii*) $E[H(X)] = \int_{R_X} H(x)f(x)dx$, si X es continua.

siempre que $\int_{R_X} |H(x)|\, f(x) < \infty$.

Hemos visto que si X es contínua, $H(X)$ puede ser discreta o contínua, en este caso restringiremos H a que $Y = H(X)$ sea una variable aleatoria contínua.

Como habra observado el lector en esta propiedad para evaluar $E[H(X)]$ no se necesita calcular la distribución de probabilidad de $Y = H(X)$ pués es sufi ciente el conocimiento de la distribución de probabilidades de X.

NOTA La media de la variable aleatoria X, presentado anteriormente es un caso especial de la definición anterior, pués si

$$H(X) = X \text{ vemos que, } E(H(X)) = E(X) = \mu .$$

TEOREMA 3.2.2 Si X es una variable aleatoria, a y b constantes. Entonces

(i) $E(a) = a$,

(ii) $E[aH(X)] = aE[H(X)]$

(iii) $E[aH(X) + bG(X)] = aE[H(X)] + bE[G(X)]$

DEMOSTRACION Daremos la demostración para el caso discreto. La demostración para el caso continuo es similar sólo se cambia la sumatoria por la integral

Asumimos que X es discreta con función de probabilidad $p(x)$.

(i) Tomamos $H(X) = a$, entonces

$$E[H(X)] = E(a) = \sum_{x \in R_X} ap(x) = a\sum_{x \in R_X} p(x) = a;$$ pués la suma de la derecha es 1.

(ii) $$E[aH(X)] = \sum_{x \in R_X} aH(x)p(x) = a\sum_{x \in R_X} H(x)p(x) = aE[H(X)].$$

(iii) $$E[aH(X) + bG(X)] = \sum_{x \in R_X} [aH(x) + bG(x)]p(x)$$

$$= a\sum_{x \in R_X} H(x)p(x) + b\sum_{x \in R_X} G(x)p(x)$$

$$= aE[H(X)] + bE[G(X)]$$

El teorema siguiente es un caso especial del teorema 3.2.2 (iii). Es una función lineal de X .

TEOREMA 3.2.3 Sea X una variable aleatoria y si a y b son constantes, entonces

$$E[aX \pm b] = aE(X) \pm b$$

DEMOSTRACION Demostraremos para el caso discreto, para el continuo la demostración es similar.

Asumimos que X es una variable aleatoria discreta con función de probabilidad $p(x)$, entonces

$$E[aX \pm b] = \sum_{x \in R_X} (ax \pm b)p(x)$$

$$= a \sum_{x \in R_X} x\,p(x) \pm b \sum_{x \in R_X} p(x)$$

$$= a\,E(X) \pm b.$$

CONSECUENCIA DEL TEOREMA 3.2.3 Si $b = 0$, entonces $E(aX) = a\,E(X)$. Es decir, la esperanza de una constante por una variable aleatoria, es la constante por la esperanza de la variable aleatoria.

TEOREMA 3.3.4 Sean $X_1, X_2, \ldots, X_n$, n variables aleatorias y a_i $(i = 0,1, \ldots, n)$ constantes, entonces

$$E[a_0 + \sum_{i=1}^{n} a_i X_i] = a_0 + \sum_{i=1}^{n} a_i\,E(X_i)$$

CONSECUENCIA DEL TEOREMA 3.2.4 Si X e Y son variables aleatorias y a,b constantes, entonces

$$E[aX + bY] = a\,E(X) + b\,E(Y)$$

En efecto: asumimos que X e Y son discretas con función de probabilidad - $p(x)$, entonces

$$E[aX + bY] = \sum_{x \in R_X} (ax + by)\,p(x)$$

$$= a \sum_{x \in R_X} x\,p(x) + b \sum_{x \in R_X} y\,p(x)$$

$$= a\,E(X) + b\,E(Y)$$

EJEMPLO 12 Sea X una variable aleatoria cuya distribución de probabilidad está dada por .

Calcular:

(a) $E(2X + 1)$; (b) $E(4X - 2)$

(c) $E(2X^2 + 3)$; (d) $E(4X^2 + X + 3)$

(e) $E(3X^2 + 4X - 5)$; (f) $E(X^3)$

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

SOLUCION (a) Cálculo de E(X). Por la fórmula (*) se obtiene

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{12}{8} = \frac{3}{2} \ .$$

$$E(2X + 1) = 2E(X) + 1 = 2 \times \frac{3}{2} + 1 = 4. \quad \text{por teorema 3.2.3}$$

(b) $$E(4X - 2) = 4E(X) - 2 = 4 \times \frac{3}{2} - 2 = 4 \quad \text{por teorema 3.2.3}$$

(c) Primero calculemos $E(X^2)$. Por la fórmula (i) del teorema 3.2.1. se obtiene

$$E(X^2) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = \frac{24}{8} = 3 \ .$$

$$E(2X^2 + 3) = 2E(X^2) + 3 = 2 \times 3 + 3 = 9. \quad \text{por teorema 3.2.3}$$

(d) $$E(4X^2 + X+3) = 4E(X^2) + E(X) + 3 = 4 \times 3 + \frac{3}{2} + 3 = 16.5$$

(e) $$E(3X^2 + 4X-5) = 3E(X^2) + 4E(X) - 5 = 3 \times 3 + 4 \times \frac{3}{2} - 5 = 10 \ .$$

(f) Por la fórmula (i) del teorema 3.2.1 se obtiene

$$E(X^3) = 0^3 \times \frac{1}{8} + 1^3 \times \frac{3}{8} + 2^3 \times \frac{3}{8} + 3^3 \times \frac{1}{8} = \frac{54}{8} = \frac{27}{4}$$

EJEMPLO 13 Considere la variable aleatoria X definida en el ejemplo 2. Calcular :

(a) $E(2X + 3)$; (b) $E(5X^2 - 2X + 1)$

SOLUCION (a) $E(2X + 3) = 2E(X) + 3 = 2 \times 1 + 3 = 5.$ por teorema 3.2.3

(b) Primero calculemos $E(X^2)$. Por la fórmula (ii) del teorema 3.2.1

$$E(X^2) = \int_0^2 \frac{3}{4} x^3 (2 - x)\, dx = \frac{3}{4} \left[\frac{2x^4}{4} - \frac{x^5}{5} \right]_0^2 = \frac{6}{5} \ .$$

$$E(5X^2 - 2X + 1) = 5E(X^2) - 2E(X) + 1$$

$$= 5 \times \frac{6}{5} - 2 \times 1 + 1 = 5 \ .$$

EJEMPLO 14 Se va rifar un automóvil cuyo precio es de \$ 3,000.00 dólares, se venden 10,000 boletos a 1 dólar cada uno. Si se compra 1 boleto, ¿cuál es el beneficio esperado?. Si se compran 100 boletos, ¿cuál es el beneficio esperado?

SOLUCION (a) Sea la variable aleatoria X definida por

$$X(\omega) = \text{beneficio neto al comprar 1 boleto}$$

$$R_X = \{-1,\ 2{,}999\}$$

Al comprar 1 boleto, se pierde 1 dólar o se gana el automóvil que vale - 3,000 dólares.

La distribución de probabilidad de X es,

x	-1	$2{,}999$
$p(x)$	$\frac{9{,}999}{10{,}000}$	$\frac{1}{10{,}000}$

por lo tanto, el beneficio esperado es ,

$$E(X) = -1\left(\frac{9{,}999}{10{,}000}\right) + 2{,}999\left(\frac{1}{10{,}000}\right) = -\frac{7000}{10000} = -0.70$$

(b) Supongamos ahora que $Y = H(X) = 100X$ una función de la variable aleatoria X. $Y = 100X$, es el beneficio al comprar 100 boletos. Luego,

$$E(100X) = 100\,E(X) = 100(-0.70) = -70$$

EJEMPLO 15 En una lotería, se venden 1,000 boletos a 25 ¢ (centavos de dólar) cada uno. Hay 5 premios en efectivo de \$ 25, de \$20, \$ 10, de \$ 5, y de \$ 1, primero, segundo, tercero, cuarto y quinto premio respectivamente - Calcular la ganancia esperada de un comprador de 2 boletos.

SOLUCION Definimos la variable aleatoria X tal que

$$X(\omega) = \text{es el beneficio neto al comprar un boleto.}$$

Entonces $Y = 2X$, es el beneficio neto al comprar 2 boletos y

$$E(Y) = E(2X) = 2E(X)$$

Debemos calcular entonces la distribución de probabilidad de X,

$$R_X = \{-0.25,\ 0.75,\ 4.75,\ 9.75,\ 19.75,\ 24.75\} \qquad y$$

x	-0.25	0.75	4.75	9.75	19.75	24.75
$p(x)$	$\frac{995}{1{,}000}$	$\frac{1}{1{,}000}$	$\frac{1}{1{,}000}$	$\frac{1}{1{,}000}$	$\frac{1}{1{,}000}$	$\frac{1}{1{,}000}$

por lo tanto, el beneficio esperado al comprar 1 boleto es,

$$E(X) = (-\ 0.25)\ \frac{995}{1,000} + (0.75)\ \frac{1}{1,000} + (4.75)(\frac{1}{1000}) + (9.75)(\frac{1}{1000}) + $$
$$+ (19.75)(\frac{1}{1000}) + (24.75)(\frac{1}{1,000})$$
$$= (-\ 0.25)(\frac{995}{1000}) + \frac{1}{1000}\ (0.75 + 4.75 + 9.75 + 19.75 + 24.75)$$
$$= -\ 0.24875 + 0.05975 = -\ 0.189\ .$$

Luego,

$E(2X) = 2(-\ 0.189) = -\ 0.378$, es el beneficio esperado al comprar dos boletos.

EJEMPLO 16 Si X representa los posibles valores que se obtiene al lanzar - un dado balanceado. Hallar $E(Y)$, donde

$$Y = 2X^2 - 5$$

SOLUCION $X(\omega)$ = número obtenido al lanzar un dado

$$R_X = \{1,2,3,4,5,6\}$$

La función de probabilidad de X es

$$p(x) = \frac{1}{6}\ ,\qquad x = 1,2,3,4,5,6.$$

Cálculo de $E(X^2)$. Por la fórmula (i) del teorema 3.2.1

$$E(X^2) = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = \frac{91}{6}$$

Luego,

$$E(2X^2 - 5) = 2E(X^2) - 5 = 2 \times \frac{91}{6} - 5 = \frac{76}{3}$$

EJEMPLO 17 Asumimos que un cierto tirador siempre acierta en un blanco circular de radio 1 dividido en 3 zonas A, B y C (como indica la figura); - los puntajes al acertar en las zonas son 10,7 y 3 respectivamente y la probabilidad de acertar (en dichas zonas) están dadas por la función de cuantía de la variable aleatoria Y, donde

$$Y = [\![3\,|\,X\,|]\!]$$

y X una variable aleatoria cuya función de densidad es,

$$f(x) = \begin{cases} kx^2 & , \quad -1 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

X es la variable aleatoria que da la posición horizontal del disparo del tirador. Se pide :

(a) Hallar el valor de k y la función de cuantía de Y.

(b) El puntaje medio esperado.

NOTA $[\![\ ]\!]$ Función máximo entero

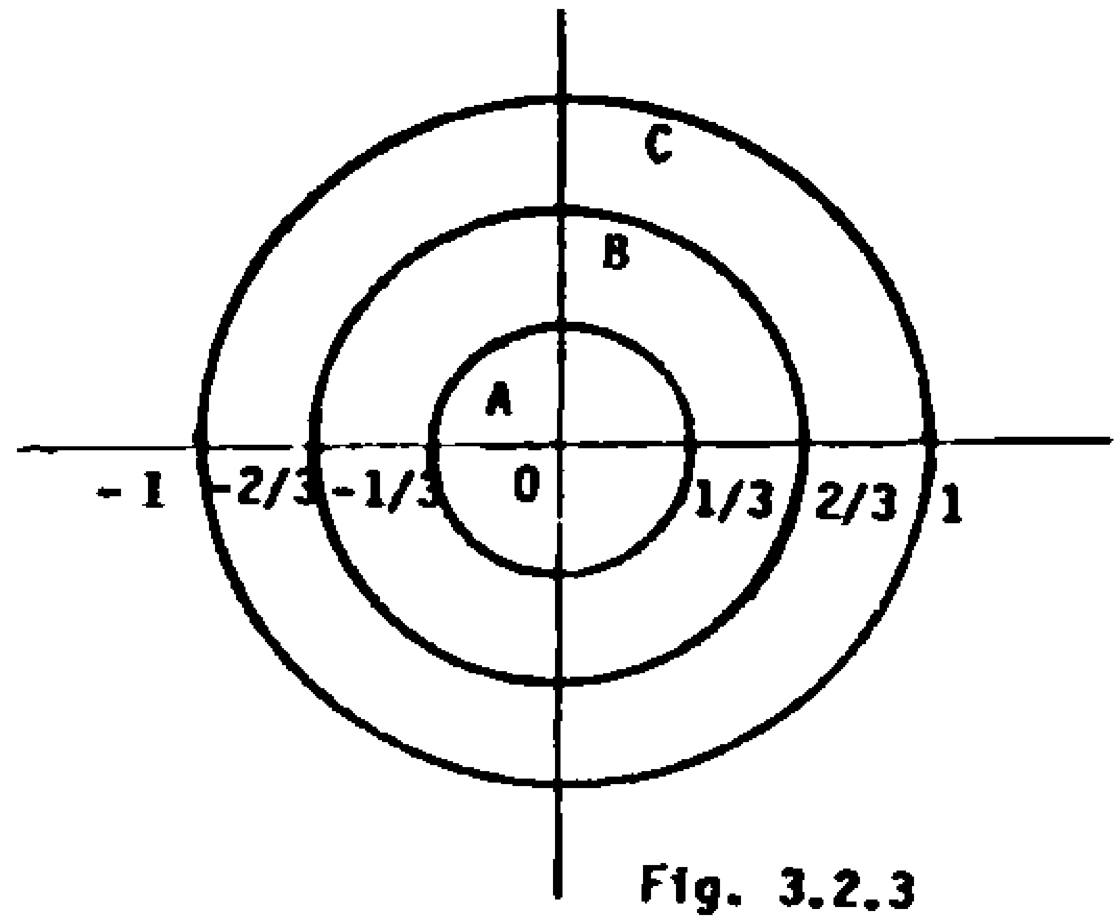


Fig. 3.2.3

SOLUCION Sea Z la variable aleatoria definida por

$Z(\omega)$ = puntaje obtenido por el tirador al hacer un disparo.

$$R_Z = \{3,7,10\}$$

La distribución de probabilidad de Z (función de cuantía) está dada por la distribución de probabilidad de

$$Y = H(X) = [\![3\,|X|]\!]$$

Obsérvese que X es una variable aleatoria contínua; pero $Y = [\![3|X|]\!]$ es una variable aleatoria discreta.

Cálculo de la distribución de probabilidad (función de cuantía) de Y (que es igual al de Z).

(a) (i) cálculo del valor de k.

$$\int_{-1}^{1} kx^2\,dx = k\,\frac{x^3}{3}\bigg|_{-1}^{1} = K\left[\frac{1}{3} + \frac{1}{3}\right] = 1;$$ por definición de función de densidad.

de donde, $k = \dfrac{3}{2}$

Luego, la funsión de densidad de X es, es,

$$f(x) = \begin{cases} \dfrac{3}{2}\,x^2 & , \quad -1 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(ii) Cálculo de la distribución de probabilidad (función de cuantía) de Y.

como, $-1 < x < 1$, entonces, $0 < 3\,|x| < 3$, luego los valores posibles que toma la variable aleatoria Y es el conjunto $\{0,1,2\}$. Es decir, el -

rango de Y, ($y = [\![3|x|]\!]$), es el conjunto $\{0,1,2\}$.

1. $y = [\![3|x|]\!] = 0$, sí, sólo sí $0 \leq 3|x| < 1$ ó $0 \leq |x| < \frac{1}{3}$ esto implica que $-\frac{1}{3} < x < \frac{1}{3}$.

   Es decir, el evento $(Y = 0)$ en R_Y es equivalente al evento $(-\frac{1}{3} < X < \frac{1}{3})$ en R_X y este es equivalente al evento $(Z = 10)$, acertar en la zona A; - ya que **X** representa la posición horizontal del disparo (ver fig. 3.2.3) Por lo tanto se tiene,

$$p_Y(0) = P[Y = 0] = P[Z = 10] = \int_{-1/3}^{1/3} \frac{3}{2}x^2\,dx = \frac{x^3}{2}\Bigg|_{-\frac{1}{3}}^{1/3} =$$

$$= \frac{1}{2 \times 27} + \frac{1}{2 \times 27} = \frac{1}{27} .$$

2. $y = [\![3|x|]\!] = 1$. sí, sólo sí $1 \leqslant 3|x| < 2$ ó $\frac{1}{3} \leqslant |x| < \frac{2}{3}$ la cual es equivalente a

$$\frac{1}{3} \leqslant |x| \;\wedge\; |x| < \frac{2}{3} \iff \begin{cases} x \leq -\frac{1}{3} \;\vee\; x \geq \frac{1}{3} \\ \wedge \\ -\frac{2}{3} < x < \frac{2}{3} \end{cases}$$

   de donde, $-\frac{2}{3} < x \leq -\frac{1}{3} \;\vee\; \frac{1}{3} \leq x < \frac{2}{3}$ es el conjunto solución

   Es decir,

   El evento $(Y = 1)$ es equivalente al evento $(-\frac{2}{3} < X \leq -\frac{1}{3}$ ó $\frac{1}{3} \leqslant X < \frac{2}{3})$ en R_X lo que a su vez es equivalente al evento $(Z = 7)$, acertar en la - zona B; por la misma razón anterior (ver fig. 3.2.3). Por lo tanto,

$$p_Y(1) = P[Y = 1] = P[Z = 7] = \int_{-2/3}^{-1/3} \frac{3x^2}{2}\,dx + \int_{1/3}^{2/3} \frac{3}{2}x^2\,dx$$

$$= \frac{x^3}{2}\Bigg|_{-2/3}^{-1/3} + \frac{x^3}{2}\Bigg|_{1/3}^{2/3} = \frac{7}{27} .$$

3. $y = [\![3|x|]\!] = 2$, sí sólo sí $2 \leq 3|x| < 3$ ó $\frac{2}{3} \leq |x| < 1$ resolviendo la desigualdad obtenemos el conjunto solución

$$\{-1 < x \leq -\frac{2}{3} \quad \text{ó} \quad \frac{2}{3} \leq x < 1\}$$

Es decir, el evento ($Y = 2$) es equivalente al evento,

$(-1 < X \leq -\frac{2}{3})$ ó $(\frac{2}{3} \leq x < 1)$ en R_X, la cual es equivalente al evento ($Z = 3$), acertar en la zona C. (ver fig. 3.2.3).
por lo tanto,

$$p_Y(2) = P[Y = 2] = P[Z = 3] = \int_{-1}^{-2/3} \frac{3}{2} x^2 dx + \int_{2/3}^{1} \frac{3}{2} x^2 dx = \frac{19}{27}.$$

4. De (1), (2) y (3), la distribución de probabilidad de Y, es

y	0	1	2
$p_Y(y)$	$\frac{1}{27}$	$\frac{7}{27}$	$\frac{19}{27}$

(b) debemos calcular el puntaje medio esperado. En otras palabras $E(Z)$.
Sabemos que las probabilidades son las mismas que las de Y. Entonces,

z	3	7	10
$p_Z(z)$	$\frac{19}{27}$	$\frac{7}{27}$	$\frac{1}{27}$

$$E(Z) = 3(\frac{19}{27}) + 7(\frac{7}{27}) + 10(\frac{1}{27}) = \frac{57 + 49 + 10}{27} = 4.29$$

Luego, el puntaje medio esperado es 4.29 .

EJEMPLO 18 Una caja contiene 4 bolas rojas y 6 azules, se sacan 3 bolas sucesivamente y con restitución. Si usted gana I/. 20.00 por cada bola roja y I/. 10.00 por cada bola azúl, ¿cuánto debería pagar por el derecho a jugar para que el juego fuese equitativo?

SOLUCION Definimos la variable aleatoria X por

$$X(\omega) = \text{ganancia total del jugador.}$$

Sea L cantidad que deposita por derecho a jugar. Entonces, para que el juego sea equitativo se debe tener,

$$E(X - L) = E(X) - L = 0$$

$$R_X = \{30, 40, 50, 60\}$$

X toma el valor 30 si las 3 bolas extraídas son azules

X toma el valor 40 si de las 3 bolas extraídas, 2 son azules y 1 roja

X toma el valor 60 si de las 3 bolas extraídas, 1 es azul y 2 rojas

X toma el valor 60 si las 3 bolas extraídas son rojas.

$$P[X = 30] = P[AAA] = P[A]\,P[A]\,P[A] = \left(\frac{6}{10}\right)^3$$

por ser la extracción con restitución.

$$P[X = 40] = P[AAR \cup ARA \cup RAA] = \binom{3}{2}\left(\frac{6}{10}\right)^2\left(\frac{4}{10}\right)$$

pués, los eventos son mutuamente excluyentes.

$$P[X = 50] = P[ARR \cup RAR \cup RRA] = \binom{3}{2}\left(\frac{6}{10}\right)\left(\frac{4}{10}\right)^2$$

$$P[X = 60] = P[RRR] = \left(\frac{4}{10}\right)^3 .$$

por lo tanto, la distribución de probabilidad de la variable aleatoria es,

x	30	40	50	60
$p(x)$	$\left(\frac{6}{10}\right)^3$	$3\left(\frac{6}{10}\right)^2\left(\frac{4}{10}\right)$	$3\left(\frac{6}{10}\right)\left(\frac{4}{10}\right)^2$	$\left(\frac{4}{10}\right)^3$

$$E(X - L) = E(X) - L = 30\left(\frac{6}{10}\right)^3 + 40\,(3)\left(\frac{6}{10}\right)^2\left(\frac{4}{10}\right) + 50\,(3)\left(\frac{6}{10}\right)\left(\frac{4}{10}\right)^2 + $$

$$+ 60\left(\frac{4}{10}\right)^3 - L = \frac{4200}{100} - L = 0$$

de donde $L = 42.00$.

Es decir, usted debe pagar I/. 42.00 por derecho a jugar para que el juego resulte equitativo.

EJEMPLO 19 En el problema anterior suponga que las bolas se sacan sin restitución, ¿cuánto debería pagarse por el derecho a jugar?

SOLUCION Sea X la variable aleatoria definida por

$$X(\omega) = \text{la ganancia total del jugador.}$$

Sea L la cantidad que deposita por el derecho a jugar. Entonces, se debe tener que

$$E(X - L) = E(X) - L = 0$$

$$R_X = \{30, 40, 50, 60\}$$

$$P[X = 30] = \frac{\binom{6}{3}}{\binom{10}{3}} = \frac{1}{6}$$

$$P[X = 40] = \frac{\binom{6}{2}\binom{4}{1}}{\binom{10}{3}} = \frac{1}{2}$$

$$P[X = 50] = \frac{\binom{6}{1}\binom{4}{2}}{\binom{10}{3}} = \frac{3}{10}$$

$$P[X = 60] = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{1}{30}$$

NOTA El lector puede calcular las probabilidades anteriores, usando probabilidad condicional.

Luego, la distribución de probabilidad de **X**, es

x	30	40	50	60
$p(x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

$$\begin{aligned} E(X - L) &= E(X) - L \\ &= 30\left(\tfrac{1}{6}\right) + 40\left(\tfrac{1}{2}\right) + 50\left(\tfrac{3}{10}\right) + 60\left(\tfrac{1}{30}\right) - L \\ &= 42 - L = 0 \end{aligned}$$

de donde $L = 42.00$.

Es decir, ud. debe pagar I/. 42.00 por derecho a jugar.

EJEMPLO 20 Un jugador **A** lanza un dado. Si en el primer lanzamiento saca un número impar de puntos, gana 5 intis, y termina el juego. Si no es así, - lanza nuevamente el dado y gana 18 intis si obtiene en el segundo lanzamiento el mismo número de puntos que obtuvo en el primer lanzamiento. Determine ud. la apuesta que debe colocar **A** para que el juego sea equitativo.

SOLUCION Sea **X** la variable aleatoria definida por

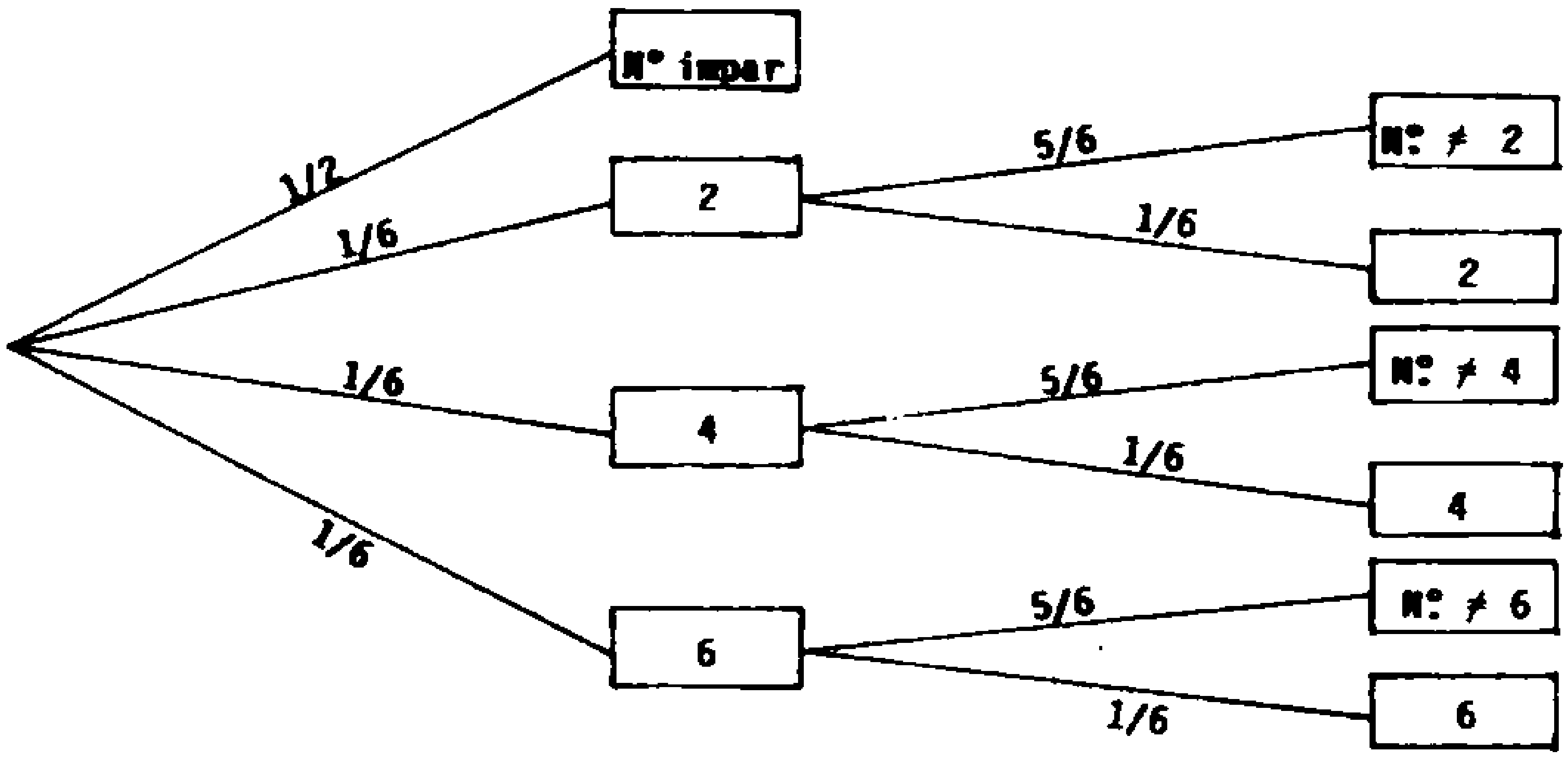
$X(\omega)$ = la ganancia del jugador A al lanzar el dado.

Sea **L** la apuesta que debe colocar **A**. Entonces, se debe tener,

$$E(X - L) = E(X) - L = 0$$

$$R_X = \{0,5,18\}$$

Para calcular las probabilidades respectivas, nos valemos del diagrama de probabilidades siguientes,



Por lo tanto,

$$p(0) = P[X = 0] = \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{5}{6} = \frac{5}{12} .$$

$$p(5) = P[X = 5] = \frac{1}{2} = \frac{6}{12} .$$

$$p(18) = P[X = 18] = \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} = \frac{1}{12} .$$

Luego, la distribución de probabilidad es,

x	0	5	18
$p(x)$	$\frac{5}{12}$	$\frac{6}{12}$	$\frac{1}{12}$

$$E(X - L) = E(X) - L$$

$$= 0\left(\frac{5}{12}\right) + 5\left(\frac{6}{12}\right) + 18\left(\frac{1}{12}\right) - L = 0$$

de donde, $L = 4$.

EJEMPLO 21 Si X es una variable aleatoria, con función de probabilidad definida por ,

$$p(x) = 2(\frac{1}{3})^x \quad , \quad x = 1,2,3, \ldots$$

Hallar $E(X)$.

SOLUCION 1 Por definición de valor esperado se tiene,

$$E(X) = \sum_{R_X} x(2\ (\frac{1}{3})^x) = \sum_{x=1}^{\infty} 2x(\frac{1}{3})^x$$

$$= 2\left[\frac{1}{3} + \frac{2}{3^2} + \frac{3}{3^3} + \frac{4}{3^4} + \ldots + \frac{n+1}{3^{n+1}} + \ldots\right]$$

2. Por otro lado $\frac{1}{3}\ E(X)$ es,

$$\frac{1}{3}\ E(X) = \frac{1}{3}\sum_{x=1}^{\infty} 2x(\frac{1}{3})^x$$

$$= \frac{2}{3}\left[\frac{1}{3} + \frac{2}{3^2} + \frac{3}{3^3} + \frac{4}{3^4} + \ldots + \frac{n}{3^n} + \ldots\right]$$

$$= 2\left[\frac{1}{3^2} + \frac{2}{3^3} + \frac{3}{3^4} + \frac{4}{3^5} + \ldots + \frac{n}{3^{n+1}} + \ldots\right]$$

3. La expresión (1) se puede escribir así,

$$E(X) = 2\left[\frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^2} + \frac{2}{3^3} + \frac{1}{3^3} + \frac{3}{3^4} + \frac{1}{3^4} + \ldots + \frac{n}{3^{n+1}} + \right.$$

$$\left. + \frac{1}{3^{n+1}} + \ldots\right]$$

4. Restando (2) de (3) obtenemos,

$$E(X) - \frac{1}{3}E(X) = 2\left[\frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \ldots + \frac{1}{3^{n+1}} + \ldots\right]$$

$$= \frac{2}{3}\left[1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \ldots + \frac{1}{3^n} + \ldots\right]$$

$$(1 - \frac{1}{3})E(X) = \frac{2}{3}\left[\frac{1}{1 - \frac{1}{3}}\right] = 1$$

de donde, $E(X) = \frac{3}{2}$.

EJEMPLO 22 Se lanzan dos dados repetidas veces hasta obtener suma 7. Determinar el número esperado de lanzamientos.

SOLUCION Sea **X** la variable aleatoria definida por

$X(\omega)$ = número de lanzamientos hasta conseguir suma 7.

$R_X = \{1,2,3,4, \ldots\}$

Determinaremos ahora la función de probabilidad de **X**.

$$P[\text{suma } 7] = \frac{6}{36} = \frac{1}{6}$$

$$P[\text{suma diferente de } 7] = \frac{30}{36} = \frac{5}{6}$$

Luego,

$$p(1) = P[X = 1] = \frac{1}{6},$$

$$p(2) = P[X = 2] = (\frac{5}{6})(\frac{1}{6}),$$

$$p(3) = P[X = 3] = (\frac{5}{6})^2 \frac{1}{6},$$

....

....

$$p(x) = P[X = x] = (\frac{5}{6})^{x-1} (\frac{1}{6})$$

por lo tanto,

$$p(x) = \frac{1}{6}(\frac{5}{6})^{x-1}, \quad x = 1,2,3,4, \ldots$$

es la función de probabilidad de **X**. El lector puede verificar fácilmente que,

(*i*) $p(x) > 0 \quad ; \quad x = 1,2,3, \ldots$ y

(*ii*) $$\sum_{x=1}^{\infty} p(x) = \sum_{x=1}^{\infty} \frac{1}{6}(\frac{5}{6})^{x-1} = 1$$

Cálculo de E(**X**). **1** por la definición de valor esperado

$$E(X) = \sum_{x=1}^{\infty} x \frac{1}{6}(\frac{5}{6})^{x-1} = \frac{1}{6}\sum_{x=1}^{\infty} x(\frac{5}{6})^{x-1}$$

$$= \frac{1}{6}\left[1 + 2(\frac{5}{6}) + 3(\frac{5}{6})^2 + \ldots + (n+1)\frac{5}{6})^n\right]$$

2. Consideremos ahora, $\frac{5}{6}$ E(**X**)

$$\frac{5}{6}\ E(X) = \frac{5}{6}\sum_{x=1}^{\infty} (\frac{1}{6})\ x\ (\frac{5}{6})^{x-1} = \frac{1}{6}\sum_{x=1}^{\infty} x\ (\frac{5}{6})^{x}$$

$$= \frac{1}{6}\left[\frac{5}{6} + 2(\frac{5}{6})^2 + \ . . . \ + n(\frac{5}{6})^n + \ . . .\right]$$

3. La expresión (1) se escribe de la siguiente manera,

$$E(X) = \frac{1}{6}[\ 1 + \frac{5}{6} + \frac{5}{6} + 2(\frac{5}{6})^2 + (\frac{5}{6})^2 + ... + n(\frac{5}{6})^n + (\frac{5}{6})^n + ...\ ]$$

4. Restando (2) de (3) obtenemos,

$$E(X) - \frac{5}{6}\ E(X) = \frac{1}{6}\ [\ 1 + \frac{5}{6} + (\frac{5}{6})^2 + \ . . . + (\frac{5}{6})^n + ...\ ]$$

$$(1 - \frac{5}{6})\ E(X) = \frac{1}{6}\ [\ \frac{1}{1-\frac{5}{6}}\ ] = 1$$

de donde, $\frac{1}{6}E(X) = 1$, Por lo tanto, $E(X) = 6$

EJEMPLO 23 Sea $H(X) = (X - a)^2$ donde a es una constante y supóngase que $E[(X - a)^2]$, existe. Hallar el valor de a para el cual $E[(X - a)^2]$, sea mínimo.

SOLUCION Escribimos,

$$G(a) = E[(X - a)^2] = E[X^2 - 2aX + a^2]$$

$$= E(X^2) - 2aE(X) + a^2$$

derivando la función $G(a)$ con respecto a a e igualando a cero

$$G'(a) = \frac{d E[(X-a)]^2}{da} = -2E(X) + 2a = 0$$

de donde, $a = E(X)$.

Es decir, la $E[(X - a)^2]$ es mínimo cuando $a = E(X)$.

3.2.3 VARIANZA DE UNA VARIABLE ALEATORIA

DEFINICION 3.2.2 La *varianza* de una variable aleatoria X, se denota por Var (X) o por la letra griega σ_X^2 (o simplemente σ^2) y se define como

$$Var(X) = \sigma^2 = E[(X - \mu)^2]$$

por lo tanto

$$\sigma^2 = E[(X - u)^2] = \sum_{x \in R_X} (x - u)^2 \, p(x) \quad , \quad \text{si } X \text{ es discreta}$$

$$= \int_{-\infty}^{\infty} (x - u)^2 f(x)\,dx \quad , \quad \text{si } X \text{ es continua.}$$

Observe que la varianza de una variable aleatoria X es un caso especial del teorema 3.2.1. Pués $H(X) = (X - u)^2$.
Observe también que si la unidad de la variable aleatoria es u, entonces la unidad de la media es la misma, en cambio la unidad de la varianza sería u^2.
Otra medida de dispersión *llamada desviación típica* de la variable aleatoria se define como la raíz cuadrada de la varianza y se denota por "σ", es decir,

$$\sigma = + \sqrt{\sigma^2}$$

Note que la unidad de σ es la misma que de la variable aleatoria; y un valor pequeño de σ indica poca dispersión, mientras que un valor grande indica gran dispersión.

NOTA En el desarrollo de la media y la varianza, hemos usado la terminología "media de la variable aleatoria" y varianza de la variable aleatoria". algunos autores usan la terminología "media de la distribución" y "varianza - de la distribución" ambas terminologías son aceptables.

EJEMPLO 24 La variable aleatoria X tiene la siguiente función de distribución

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1/8 & , \quad 0 \le x < 1 \\ 1/2 & , \quad 1 \le x < 2 \\ 5/8 & , \quad 2 \le x < 3 \\ 1 & , \quad x \ge 3 \end{cases}$$

Calcular la varianza de la variable aleatoria.

SOLUCION En el problema 23 de 2.2 hemos visto que la distribución de probabilidad, X está dado en la tabla siguiente

x	0	1	2	3	Total
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	
$xp(x)$	0	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{9}{8}$	$\frac{14}{8}$

de la tabla, $E(X) = \mu = \frac{14}{8} = \frac{7}{4}$.

La varianza de la variable aleatoria X es,

$$\sigma^2 = \sum_{x \in R_X} (x - \mu)^2 p(x)$$

$$= (0 - \tfrac{7}{4})^2 (\tfrac{1}{8}) + (1 - \tfrac{7}{4})^2(\tfrac{3}{8}) + (2 - \tfrac{7}{4})^2(\tfrac{1}{8}) + (3 - \tfrac{7}{4})^2 (\tfrac{3}{8})$$

$$= \frac{49}{16 \times 8} + \frac{27}{16 \times 8} + \frac{1}{16 \times 8} + \frac{75}{16 \times 8} = \frac{19}{16} .$$

EJEMPLO 25 Calcular la varianza de la variable aleatoria X definida en el ejemplo 2.

SOLUCION $\mu = E(X) = 1$. Entonces

$$\sigma^2 = E[(X - 1)^2] = \int_0^2 (x - 1)^2 \frac{3}{4} x(2 - x)\,dx = \frac{1}{5} .$$

3.2.4 PROPIEDADES DE LA VARIANZA Y DESVIACION TIPICA

TEOREMA 3.2.5 Si X es una variable aleatoria con media μ, la varianza de X está dado por ,

$$Var(X) = \sigma^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

DEMOSTRACION

$$Var(X) = \sigma^2 = E[(X - \mu)^2]$$

$$= E(X^2 - 2X\mu + \mu^2)$$

$$= E(X^2) - 2\mu E(X) + E(\mu^2) \qquad \text{teorema 3.2.3}$$

$$= E(X^2) - \mu^2$$

pués $E(X) = \mu$, por definición y $E(\mu^2) = \mu^2$ por ser μ^2 constante.

TEOREMA 3.2.6 Si X es una variable aleatoria, a y b constantes, entonces

$$Var(aX \pm b) = a^2\ Var(X)$$

DEMOSTRACION

$$\begin{aligned} Var(aX + b) &= E[(aX + b)^2] - [E(aX + b)]^2 && \text{teorema 3.2.5} \\ &= E[a^2X^2 + 2abX + b^2] - [aE(X) + b]^2 && \text{teorema 3.2.3} \\ &= a^2E(X^2) + 2abE(X) + b^2 - [a^2[E(X)]^2 + 2abE(X) + b^2] \\ &= a^2[E(X^2) - [E(X)]^2] \\ &= a^2\ Var(X). \end{aligned}$$

CONSECUENCIA DEL TEOREMA 3.2.6 .

1. Si $a = 0$, $Var(b) = 0$, la varianza de una constante es cero
2. Si $b = 0$, $Var(aX) = a^2\ Var(X)$.

TEOREMA 3.2.7 Si X es una variable aleatoria y c una constante, entonces

(i) $\sigma_{cX} = |c|\sigma_X$ (ii) $\sigma_{X+c} = \sigma_X$

EJEMPLO 26 Para la variable aleatoria definida en el ejemplo 1. Hallar
(a) $Var(X)$, (b) $Var(-2X + 3)$

SOLUCION (a) Por el teorema 3.2.5

$$Var(X) = E(X^2) - [E(X)]^2$$
$$E(X) = 1.32 \text{ , ejemplo 1.}$$

Cálculo de $E(X^2)$. Por la fórmula (i) del teorema 3.2.1

$$\begin{aligned} E(X^2) &= 0^2(0.2)1^2\ (0.4) + 2^2\ (0.3) + 3^2\ (0.08) + 4^2\ (0.02) \\ &= 2.64. \end{aligned}$$

Ahora hallamos

$$Var(X) = 2.64 - (1.32)^2 = 2.64 - 1.7424 = 0.8976.$$

(b) $Var(-2X) = 4\ Var(X)$ teorema 3.2.6

$$= 4(0.8976) = 3.5904.$$

EJEMPLO 27 Hallar (a) $Var(X)$, (b) $Var(-5X - 3)$ de la variable aleatoria definida en el ejemplo 3.

SOLUCION (a) Por el teorema 3.2.5 es

$$\mathrm{Var}(X) = E(X^2) - E(X)^2$$

$$E(X) = 0 \quad , \quad \text{ejemplo 3.}$$

Cálculo de $E(X^2)$. Por la fórmula (*ii*) del teorema 3.2.1

$$E(X^2) = \int_{-1}^{1} \frac{3}{2} x^4\, dx = \frac{3}{2}\, \frac{x^5}{5} \Big|_{-1}^{1} = \frac{3}{5}$$

Entonces,

$$\mathrm{Var}(X) = \frac{3}{5} - 0 = \frac{3}{5} .$$

(b) $\mathrm{Var}(-5X - 3) = 25\, \mathrm{Var}(X)$ teorema 3.2.6

$$= 25(\tfrac{3}{5}) = 15 .$$

EJEMPLO 28 Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} kx^3 & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

(a) Calcular la función de distribución y el valor de la constante k.

(b) Calcular $P[\frac{X}{X+1} \geq \frac{1}{3}]$

(c) Calcular la esperanza matemática y la varianza.

SOLUCION (a) Queda como ejercicio para el lector; verificar que $k = 4$,

y
$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ x^4 & , \quad 0 < x < 1 \\ 1 & , \quad x > 1 \end{cases}$$

La función de densidad es,

$$f(x) = \begin{cases} 4x^3 & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

(b) $P[\frac{X}{X+1} > \frac{1}{3}] = P[3X \geq X + 1] = P[X \geq \frac{1}{2}]$ pues $P(X > 0) = 1$

$$= 1 - P[X < \tfrac{1}{2}] = 1 - P[X \leq \tfrac{1}{2}]$$

$$= 1 - F(1/2) = 1 - \frac{1}{16} = \frac{15}{16} .$$

(c) $$E(X) = \int_0^1 x(4x^3)dx = \int_0^1 4x^4\,dx = \left.\frac{4x^5}{5}\right|_0^1 = \frac{4}{5}.$$

Cálculo de Var(X). Por el teorema 3.2.5

$$E(X^2) = \int_0^1 x^2(4x^3)\,dx = \left.\frac{4x^6}{6}\right|_0^1 = \frac{2}{3}.$$

Luego, $$\sigma^2 = E(X^2) - \mu^2$$

$$= \frac{2}{3} - \frac{16}{25} = \frac{2}{75}.$$

EJEMPLO 29 Sea **X** una variable aleatoria contínua con función de distribución,

$$F(x) = P[X < x] = \begin{cases} 0 & , \quad x \leq 0 \\ \dfrac{x}{2\pi} & , \quad 0 < x \leq 2\pi \\ 1 & , \quad x > 2\pi \end{cases}$$

Si $E(X) = \mu$ y $var(X) = \sigma^2$. Hallar $P[\mu - \sigma < X < \mu + \frac{\sigma}{2}]$

SOLUCION La función de densidad de la variable aleatoria es,

$$f(x) = F'(x) = \begin{cases} \dfrac{1}{2\pi} & , \quad 0 < x < 2\pi \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Cálculo de E(X).

$$\mu = E(X) = \int_0^{2\pi} \frac{x}{2\pi}\,dx = \left.\frac{x^2}{4\pi}\right|_0^{2\pi} = \pi.$$

Cálculo de Var(X).

$$E(X^2) = \int_0^{2\pi} \frac{x^2}{2\pi}\,dx = \left.\frac{x^3}{6\pi}\right|_0^{2\pi} = \frac{4}{3}\pi^2.$$

Luego, $\sigma^2 = \frac{4}{3}\pi^2 - \pi^2 = \frac{\pi^2}{3}$, de donde $\sigma = \frac{\pi}{\sqrt{3}}$.

Entonces,

$$P[\mu - \sigma < X < \mu + \frac{\sigma}{2}] = F(\mu + \frac{\sigma}{2}) - F(\mu - \sigma)$$

Puesto que $\mu + \frac{\sigma}{2}$ y $\mu - \sigma$ están entre 0 y 2π, tenemos

$$F(\mu + \frac{\sigma}{2}) - F(\mu - \sigma) = \frac{\mu + \frac{\sigma}{2}}{2\pi} - \frac{\mu - \sigma}{2\pi} = \frac{3\sigma}{4\pi} = \frac{\sqrt{3}}{4} = 0.433$$

Es decir,

$$P[\mu - \sigma < X < \mu + \frac{\sigma}{2}] = 0.433.$$

3.2.5 MODA, MEDIANA Y PERCENTILES DE UNA VARIABLE ALEATORIA

DEFINICION 3.2.3 MODA Se llama moda de una variable aleatoria discreta **X** a su valor más probable.

Se llama moda de una variable continua **X** a su valor con función de densidad máxima. En otras palabras, el valor x_o de la variable aleatoria X es una moda de **X**, si

$$p(x_o) \geq p(x), \quad \forall\, x \in R_X \quad \text{Si } X \text{ es discreta}$$

$$f(x_o) \geq f(x), \quad \forall\, x \in R_X \quad \text{si } X \text{ es continua.}$$

NOTACION La moda de una variable aleatoria denotaremos por x_{md}.

INTERPRETACION GEOMETRICA La moda es la abscisa de aquel punto de la curva o polígono de distribución de probabilidad con ordenada máxima.

DEFINICION 3.2.4 MEDIANA La *mediana* de una variable aleatoria **X** es un número x_o, tal que

$$F(x_o) = P[X \leq x_o] \geq \frac{1}{2} \quad \wedge \quad P[X \geq x_o] \geq \frac{1}{2}$$

o equivalentemente, $P[X \leq x_o] = P[X \geq x_o] = \frac{1}{2}$, para **X** continua.

Es decir, la mediana tiene la propiedad de que la variable aleatoria tiene la misma posibilidad de estar a cualquiera de los dos lados de este.

NOTACION Denotaremos por x_{me} a la mediana de una variable aleatoria.

INTERPRETACION GEOMETRICA La ordenada trazada por el punto con abscisa $x = x_{me}$ divide por la mitad el area acotada por la curva o polígono de distribución de probabilidad

NOTA 1 Una variable aleatoria **X** es simétrica si su función de probabilidad (o función de densidad de probabilidad) es simétrica*

* **Una función f es simétrica alrededor de $x = a$ si $f(x + a) = f(x - a)$. En particular, f es simétrica alrededor del origen ($a = 0$) si $f(x) = f(-x)$**

NOTA 2 Si la recta $x = a$ es el eje de simetría de la curva de distribución de probabilidad $f(x)$, entonces $x_{md} = x_{me} = E(X) = a$.

EJEMPLO 30 Hallar la moda y la mediana de la variable aleatoria **X**, cuya - distribución de probabilidad está definida por,

x	- 2	- 1	1	2
$p(x)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$

SOLUCION (a) Desde que $p(x)$ es mayor, cuando $x = -2$, y $x = 2$, - entonces $x_o = 2$ y -2 son dos modas de la variable aleatoria **X** .

(b) La mediana calculamos usando la definición, $F(x_o) = P[X < x_o] = \frac{1}{2}$ se cumple para $x_o = -1$, también para cualquier $x \in [-1, 1\rangle$. En - particular podemos tomar $x_{me} = 0$.

EJEMPLO 31 Se da la función de densidad de probabilidad de la variable aleatoria **X**,

$$f(x) = \begin{cases} x - x^3/4 & , \quad 0 \leq x \leq 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Hallar: (a) la moda (b) la mediana de esta variable aleatoria.

SOLUCION (a) Hallamos el máximo de la función $f(x)$. Para esto encontramos la primera y segunda derivada.

$$f'(x) = 1 - \frac{3}{4}x^2 \quad ; \quad f''(x) = -\frac{3}{2}x$$

De la ecuación $f'(x) = 1 - \frac{3}{4}x^2 = 0$, obtenemos $x = \pm \frac{2}{\sqrt{3}}$

Entre las dos raices de esta ecuación escogemos sólo la que está comprendida en $[0,2]$. Por lo tanto, $x = \frac{2}{\sqrt{3}}$.

Puesto que $f''\left(\frac{2}{\sqrt{3}}\right) = -\frac{3}{2}\left(\frac{2}{\sqrt{3}}\right) = -\sqrt{3} < 0$, entonces en $x = \frac{2}{\sqrt{3}}$ la función $f(x)$ presenta su máximo, o sea $M_d = \frac{2}{\sqrt{3}}$.

(b) Por la definición de mediana tenemos

$$P[X < x_o] = \int_0^{x_o} (x - \frac{x^3}{4})dx = (\frac{x^2}{2} - \frac{x^4}{16})\Big|_0^{x_o} = \frac{x_o^2}{2} - \frac{x_o^4}{16} = \frac{1}{2}$$

Así, de la ecuación $\frac{x_o^2}{2} - \frac{x_o^4}{16} = \frac{1}{2}$ ó $x_o^4 - 8x_o^2 + 8 = 0$,

obtenemos $x_o = \pm\sqrt{4 \pm \sqrt{8}}$.

Entre las cuatro raices de esta ecuación escogemos la que está comprendido en el intervalo $[0,2]$. Por lo tanto, $x_{me} = \sqrt{4 - \sqrt{8}} \simeq 1.09$.

EJEMPLO 32 Sea **X** una variable aleatoria, cuya función de densidad es,

$$f(x) = \begin{cases} \frac{2}{9}x & , \quad 0 \le x \le 3 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

Calcular : (a) la mediana ; (b) la moda de **X**

SOLUCION (a) por la definición de mediana se tiene

$$P[X \le x_o] = \int_0^{x_o} \frac{2}{9}x\,dx = \frac{x^2}{9}\Big|_0^{x_o} = \frac{x_o^2}{9} = \frac{1}{2}$$

de donde obtenemos $x_o = \pm\frac{3}{\sqrt{2}}$. De estas dos soluciones escogemos la que está comprendida en el intervalo $[0,3]$. Por lo tanto, $x_o = \frac{3}{\sqrt{2}}$

(b) Los únicos puntos críticos de la función de densidad en el intervalo $[0,3]$ son 0 y 3. Los valores de f en estos puntos críticos son

$$f(0) = \frac{2}{9}(0) = 0 \quad ; \quad f(3) = \frac{2}{9}(3) = \frac{2}{3}$$

Por lo tanto, la función de densidad toma su máximo en $x_o = 3$. Es decir la moda es 3. El gráfico de la función de densidad de **X** está dado en la fig. 4.2.4.
En esta figura se observa que la función toma su máximo en el punto $x = 3$

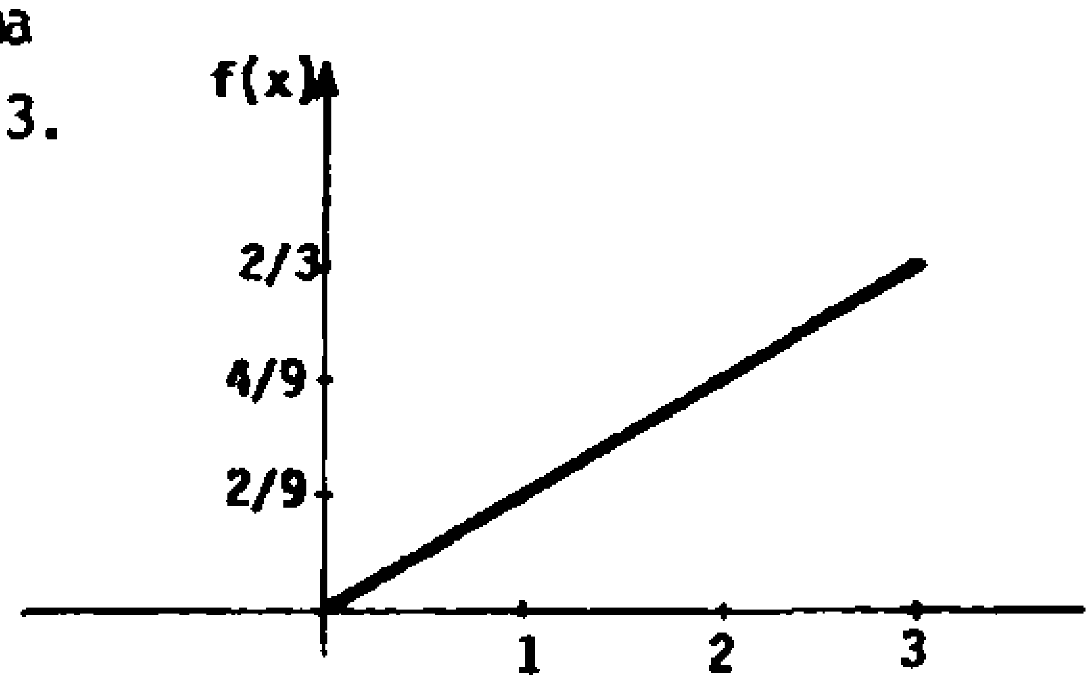


Fig. 3.2.4

DEFINICION 3.2.5 PERCENTILES Se llama el 100 k-ésimo percentil de la variable aleatoria X (denotado por x_k) al número más pequeño posible tal que la probabilidad de no excederlo es cuando menos k, con $0 < k < 1$. Es decir x_k es el valor más pequeño tal que

$$F(x_k) = P[X \leq x_k] \geq k$$

$$\wedge \qquad P[X \geq x_k] \geq 1 - k$$

Como se muestra en la fig. 3.2.5, el 30-ésimo percentil, $x_{0.30}$, es el número tal que $F(x_{0.30}) = 0.30$; el 70-ésimo percentil, $x_{0.70}$ satisface $F(x_{0.70}) = 0.70$. Por lo tanto, x_k es el número más pequeño tal que $P[X \leq x_k] = k$. Algunos percentiles reciben nombres especiales se llaman *cuartiles* de la distribución $x_{0.25}$, $x_{0.50}$ y $x_{0.75}$ debido a que cortan la función de probabilidad o función de densidad en cuatro probabilidades iguales. El 50-ésimo percentil, $x_{0.50}$ se llama también *mediana* de la variable aleatoria. A la diferencia $x_{0.75} - x_{0.25}$, se le conoce como el *rango intercuartil*, y a veces se usa como medida de la variabilidad. Se nota que el rango intercuartil esla longitud de un intervalo que encierra el 50% de la distribución de probabilidad de X.

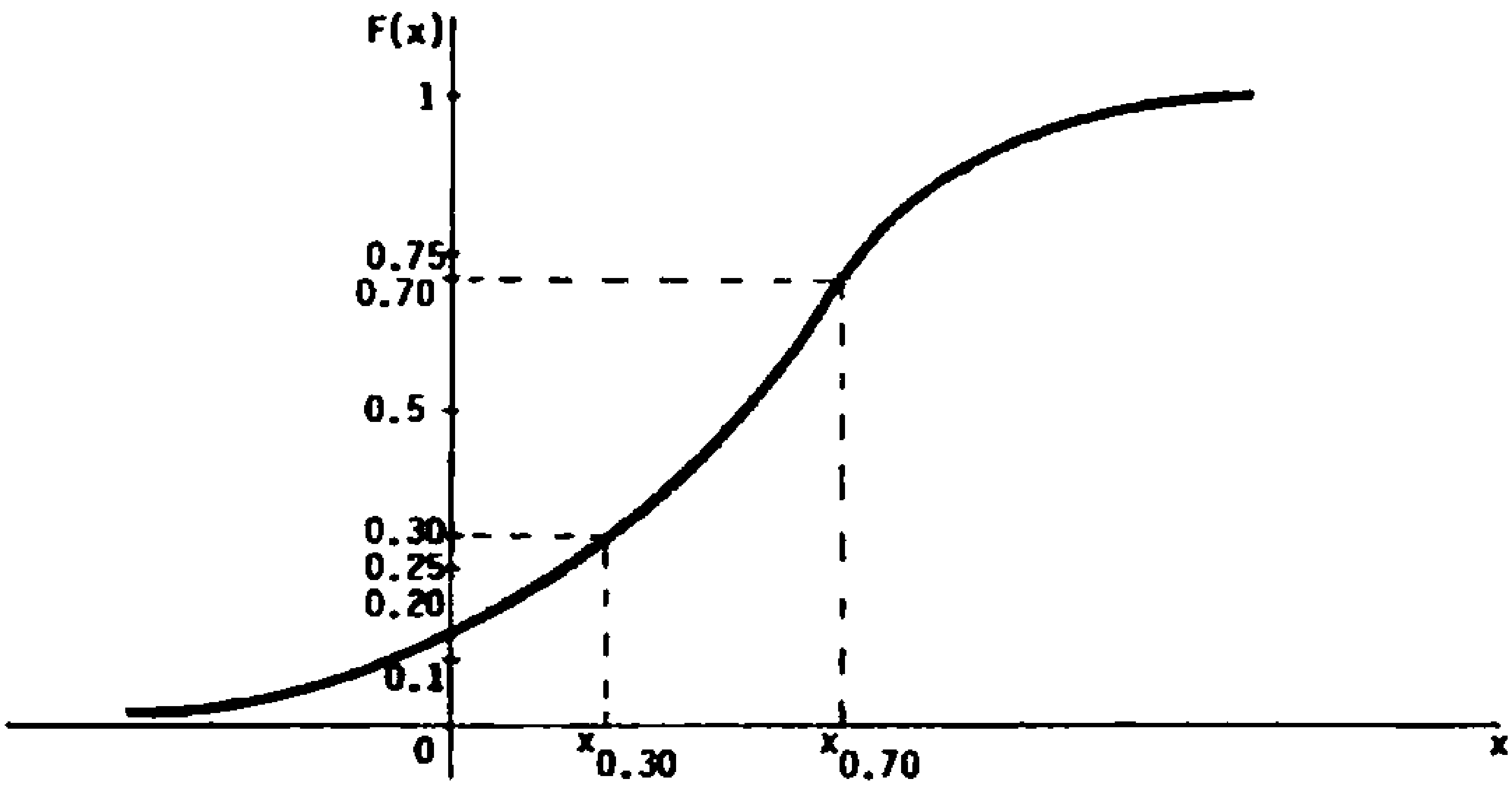


Fig. 3.2.5 Percentiles de X.

NOTA 3 En la definición del 100k-ésimo percentil se usa la desigualdad de esta manera estan seguros que se definan los percentiles para cualquier tipo de variable aleatoria. En particular, si X es una variable aleatoria discreta, su función de distribución F(x) tiene saltos y puede no existir un -

número tal que $F(x_k) = k$. Al usar la definición dada, se define el 100k-ésimo percentil como el número x_k en el cuál saltó la función de distribución más allá de k. (Ver ejemplo 34. Fig. 3.2.6)

EJEMPLO 33 Si **X** es una variable aleatoria tal que

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \dfrac{x}{100} & , \quad 0 \leq x \leq 100 \\ 1 & , \quad x > 100 \end{cases}$$

Halla : (a) $x_{0.10}$, (b) $x_{0.20}$, (c) $x_{0.55}$, (d) $x_{0.80}$, (e) la mediana y (f) el rango intercuartil

SOLUCION (a) $F(x_{0.10}) = \dfrac{x_{0.10}}{100} = 0.10$; es decir $x_{0.10} = 10.$

(b) $F(x_{0.20}) = \dfrac{x_{0.20}}{100} = 0.2$, por lo tanto $x_{0.20} = 20$

(c) $F(x_{0.55}) = \dfrac{x_{0.55}}{100} = 0.55$, se obtiene que $x_{0.55} = 55$

(d) $F(x_{0.80}) = \dfrac{x_{0.80}}{100} = 0.80$, se tiene que $x_{0.80} = 80.$

(e) La mediana está definida por

$F(x_{0.50}) = \dfrac{x_{0.50}}{100} = 0.50$, de donde $x_{0.50} = 50.$

(f) En forma análoga obtenemos

$$x_{0.25} = 100(0.25) = 25$$
$$x_{0.75} = 100(0.75) = 75$$

El rango intercuartil es $x_{0.75} - x_{0.25} = 50$.

EJEMPLO 34 Si **X** es una variable aleatoria tal que

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1/6 & , \quad 0 \leq x < 1 \\ 1/3 & , \quad 1 \leq x < 2 \\ 1/2 & , \quad 2 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

Hallar : (a) $x_{0.34}$, (b) $x_{0.51}$, (c) $x_{0.16}$, (d) $x_{0.17}$, (e) la mediana, (f) el recorrido intercuartil

SOLUCION Construimos la gráfica de F(x).

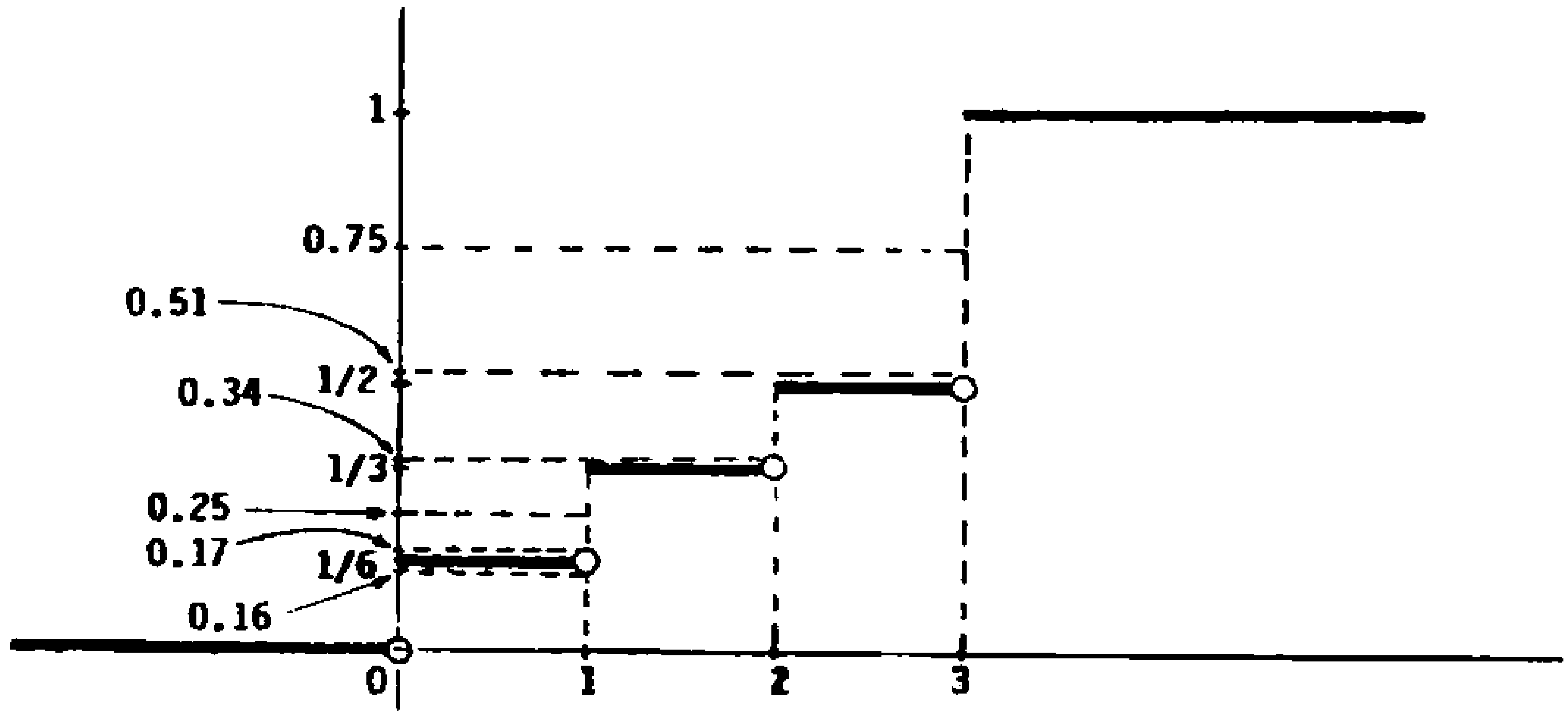


Fig. 3.2.6

(a) $x_{0.34} = 2$, (b) $x_{0.51} = 2$, (c) $x_{0.16} = 0$, (d) $x_{0.17} = 1$

(e) La mediana es $x_{0.50} = 2$, (f) $x_{0.25} = 1$, $x_{0.75} = 3$, luego, el rango intercuartil es $x_{0.75} - x_{0.25} = 2$.

3.2.6 MOMENTOS DE ORDEN SUPERIOR Y ASIMETRIA DE UNA VARIABLE ALEATORIA

DEFINICION 3.2.6 MOMENTOS Se llama *momento de orden k* alrededor del punto a de la variable aleatoria **X** a la esperanza matemática de $(X - a)^k$, y se denota por $\mu_{k,a}$. Entonces,

$$\mu_{k,a} = E[(X - a)^k] = \sum_{x \in R_X} (x - a)^k\, p(x)\ , \quad \text{si } X \text{ es discreta.}$$

$$= \int_{R_X} (x - a)^k f(x)\, dx\ , \quad \text{si } X \text{ es continua}$$

Si $a = 0$, se llaman *momentos* iniciales o *momentos al rededor del origen* de orden k de una variable aleatoria a $E(X^k)$. Es decir,

$$\mu_{k,0} = \mu'_k = E(X^k) = \sum_{x \in R_X} x^k\, p(x)\ , \quad \text{si } X \text{ es discreta.}$$

$$= \int_{R_X} x^k f(x)\,dx \quad , \text{ si } X \text{ es continua .}$$

al primer *momento inicial* , $\mu_1' = \mu = E(X)$ hemos llamado la ***media de X***.

Si $a = E(X) = \mu$, se llaman ***momentos al rededor de la*** media o momento central de orden k de una variable aleatoria a

$E[(X - \mu)^k]$. Es decir

$$\mu_{k,\mu} = \mu_k = E[(X - \mu)^k] = \sum_{x \in R_X} (x - \mu)^k p(x) \text{ , si } X \text{ es discreta .}$$

$$= \int_{R_X} (x - \mu)^k f(x)\,dx \text{ , si } X \text{ es continua}$$

al segundo momento alrededor de la media de una variable aleatoria X hemos llamado varianza de la variable aleatoria, o sea $\sigma^2 = \mu_{2,\mu}$.
El primer momento central (o alrededor de la media) para cualquier variable aleatoria es igual a cero, es decir $\mu_1 = 0$.
Los momentos iniciales y centrales de primer, segundo, tercer y cuarto orden están vinculados por las relaciones

$$\mu_1 = 0 \quad , \quad \mu_1 = \mu_2' - \mu_1'^2$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3 \quad , \quad \mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4$$

Si la distribución de probabilidad es simétrica respecto a la media, entonces todos los momentos centrales de orden impar son iguales a cero, o sea $\mu_1 = \mu_3 = \mu_5 = \ldots = 0$.

DEFINICION 3.2.7 . ASIMETRIA Se llama ***asimetría o sesgo*** a la relación

$$S_k = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{\mu_3}{\sigma^3}$$

La medida S_k es positiva , negativa o cero. Si $S_k > 0$ la distribución está sesgada a la derecha (asimétrica positiva). Si $S_k < 0$ la distribución está sesgada a la izquierda (asimétrica negativa). Si $S_k = 0$ la distribución es simétrica con respecto a la media.
En las fig. 3.2.7 y 3.2.8 se muestran las gráficas de la distribución para $S_k > 0$ y $S_k < 0$ respectivamente.

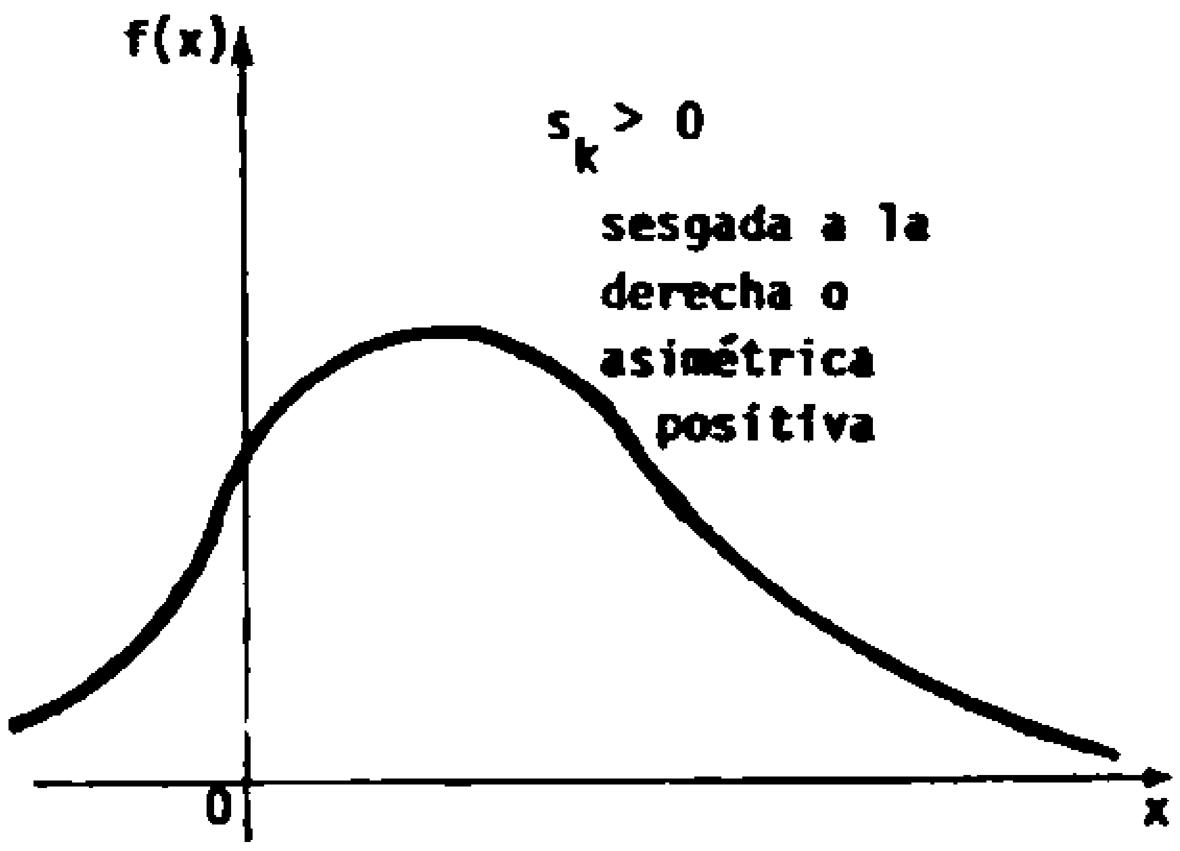


Fig. 3.2.7

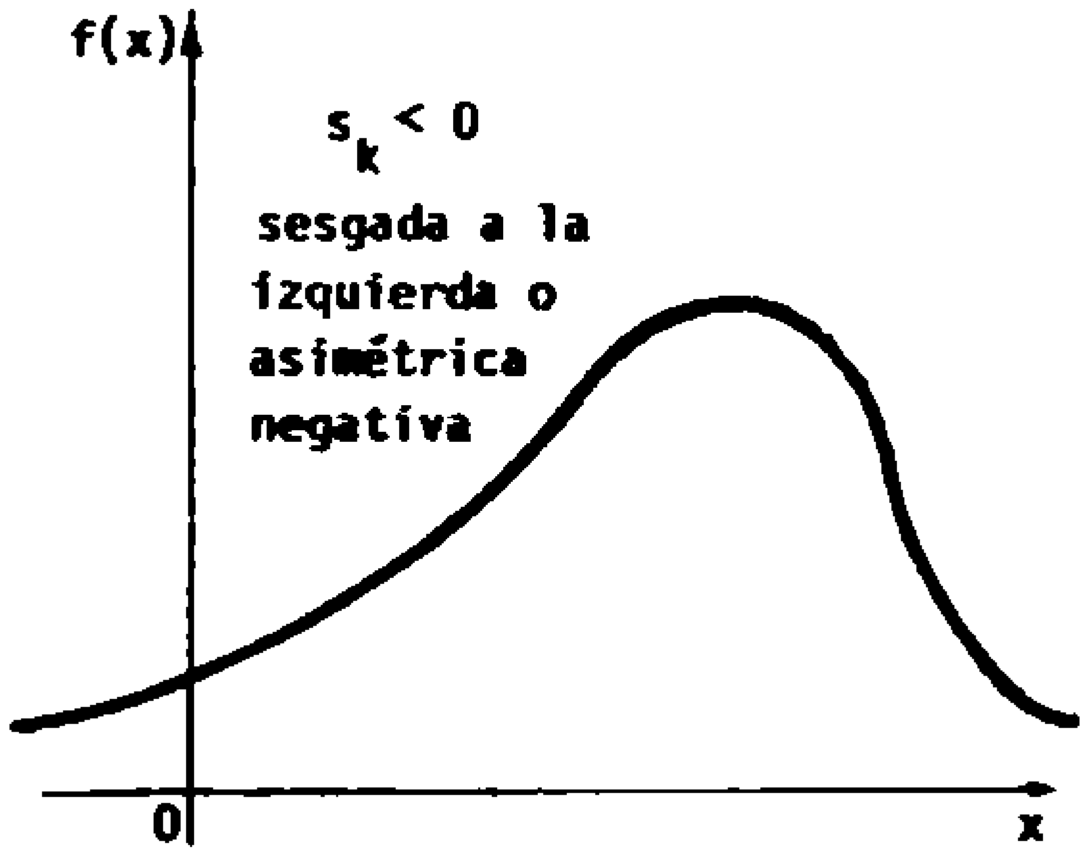


Fig. 3.2.8

DEFINICION3.2.8. CURTIOSIS Se llama curtiosis o coeficiente de curtiosis - de una variable aleatoria X a la relación

$$C_X = \frac{E[(X - u)^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4}$$

Frecuentemente se compara con la curva normal o de Gauss (Ver cap. 6) que - tiene un coeficiente de curtiosis igual a 3 y se denomina el *exceso* o sea

$$E_X = \frac{\mu_4}{\sigma^4} - 3$$

para la curva normal o de Gauss, $E_X = 0$

EJEMPLO 35 Sea **X** una variable aleatoria con distribución de probabilidad

x	2	4	6	8
$p(x)$	0.4	0.3	0.2	0.1

Hallar : (a) Los cuatro primeros momentos iniciales .
(b) Los cuatro primeros momentos centrales.
(c) La asimétrica .

SOLUCION (a_1) El momento inicial de primer orden

$$\mu_1' = 2(0.4) + 4(0.3) + 6(0.2) + 8(0.1) = 4.$$

(a_2) El momento inicial de segundo orden

$$\mu_2' = 4(0.4) + 16(0.3) + 36(0.2) + 64(0.1) = 20$$

(a_3) El momento inicial de tercer orden

$$\mu_3' = 8(0.4) + 64(0.3) + 216(0.2) + 512(0.1) = 116.8.$$

(a_4) El momento inicial de cuarto orden

$$\mu'_4 = 16(0.4) + 256(0.3) + 1296(0.2) + 4096(0.1) = 752.$$

(b_1) Determinar de los momentos centrales. El primer momento central

$$\mu_1 = 0 .$$

(b_2) El segundo momento central se encuentra por la relación

$$\mu_2 = \mu_2 - \mu_1{}^2 = 20 - 4^2 = 20 - 16 = 4 .$$

(b_3) El tercer momento central se determina por la fórmula

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2\mu_1{}^3 = 116.8 - 3 \times 4 \times 20 + 2 \times 4^3 = 4.8$$

(b_4) El cuarto momento central se utiliza la fórmula

$$\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6\mu'^2_1\mu'_2 - 3\mu'^4_1$$

$$= 752 - 4 \times 4(116.8) + 6 \times 4^2 \times 20 - 3 \times 4^4 = 35.2$$

(c) $\sigma^2 = \mu_2 = 4$, luego $\sigma = 2$. La asimetría es

$$S_k = \frac{\mu_3}{\sigma^3} = \frac{4.8}{2^3} = \frac{4.8}{8} = 0.60$$

3.2.7 DESIGUALDAD DE CHEBYSHEV

Si conocemos la distribución de probabilidad de una variable aleatoria **X** (la función de densidad en el caso contínuo, la función de probabilidad - en el caso discreto) hemos visto que podemos determinar μ y σ^2, si existen. Pero lo recíproco no es cierto. Es decir, conociendo μ y σ^2 no podemos determinar la distribución de probabilidad de **X**. Sin embargo se puede - dar una cuota superior (o inferior) para probabilidad del tipo - $P[|X - \mu| \geq k\sigma]$, este resultado, se conoce como la *desigualdad* de - CHEBYSHEV.

TEOREMA 3.2.8 (DE CHEBYSHEV) Si la variable aleatoria **X** tiene μ y σ^2 finita, entonces, para cualquier $k > 1$, se cumple

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

la cual indica que la probabilidad que **X** tome algún valor fuera del intervalo $\langle \mu - k\sigma , \mu + k\sigma \rangle$ es a lo más $1/k^2$

DEMOSTRACION El teorema es válido tanto para variables aleatorias discretas como para contínuas. Daremos aquí la demostración para el caso discreto. La demostración para el caso contínuo queda como ejercicio para el lector.
Sea $p(x)$ la función de probabilidad de X. y sea el evento,

$$A = \{x \;/\; |x - \mu| \geq k\sigma\}$$

Entonces, $$\sigma^2 = E[(X - \mu)^2] = \sum_{x \in R_X} (x - \mu)^2\, p(x)$$

$$= \sum_{x \in A} (x - \mu)^2\, p(x) + \sum_{x \notin A} (x - \mu)^2\, p(x) \qquad (1)$$

Note que $\qquad |x - \mu|^2 = (x - \mu)^2$

El segundo sumando de (1) es un número no negativo,entonces es mayor - o igual a cero. Luego,

$$\sigma^2 \geq \sum_{x \in A} (x - \mu)^2\, p(x) \qquad (2)$$

Y como el evento A, es $|x - \mu| \geq k\sigma$, se tiene $(x - \mu)^2 \geq k^2\sigma^2$ - entonces, reemplazando esta última expresión en (2),

$$\sigma^2 \geq \sum_{x \in A} k^2\sigma^2\, p(x) = k^2\sigma^2 \sum_{x \in A} p(x)$$

pero, $$\sum_{x \in A} p(x) = P[X \in A] = P[|X - \mu| \geq k\sigma]$$

Luego, $$\sigma^2 \geq k^2\sigma^2\, P[|X - \mu| \geq k\sigma] \qquad \text{ó}$$

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

CONSECUENCIAS

(a) Si $\varepsilon = k\sigma$, se tiene $\quad P[|X - \mu| > \varepsilon] \leq \dfrac{\sigma^2}{\varepsilon^2}$

(b) Puesto que $\{|X - \mu| \geq k\sigma\}$ y $\{|X - \mu| < k\sigma\}$ son eventos complementarios, entonces

$$P[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}$$

indica que la probabilidad de que X tome valores dentro del intervalo - $\langle \mu - k\sigma\, ,\, \mu + k\sigma \rangle$ es por lo menos $1 - 1/k^2$

EJEMPLO 36 Sea **X** una variable aleatoria con media 33 y varianza 16. Hallar una cota inferior para $P[23 < X < 43]$.

SOLUCION

$$P[23 < X < 43] = P[23 - 33 < X - 33 < 43 - 33]$$
$$= P[-10 < X - 33 < 10] = P[|X - \mu| < 10]$$

Observe que $10 = K\sigma$, y $\sigma = 4$, entonces $k = \frac{5}{2}$. Luego,

$$P[23 < X < 43] = P[|X - \mu| < \frac{5}{2} \times 4] \geq 1 - \frac{1}{(\frac{5}{2})^2}$$

por lo tanto, $P[23 < X < 43] \geq \frac{21}{25}$.

PROBLEMAS 3.2

1. Para participar en un juego ud. debe pagar I/. 2.00. El juego consiste - en lo siguiente: lanzar dos dados y si la suma es al menos 8, se le permitirá lanzar un dado y recibirá I/. 1.00, por cada punto que obtenga en el dado (por ejemplo, si saca 4 recibe I/.4.). ¿Qué cantidad de dinero espera ud. ganar o perder en este juego?.

2. Considere a una persona que compra un billete de una lotería que vende - 1000 billetes y que dá cuatro premios de I/. 200. , 10 premios de - I/. 100. y 20 premios de I/. 10. ¿Cuanto debería estar dispuesto a pagar la persona por un billete de esta loteria?

3. Ud. le hace una apuesta a un jugador por I/. 100. Si él pienza que su ganancia esperada es I/. 50. ¿cuál es la probabilidad de que gane la apuesta?

4. Un tirador hace tres disparos a un blanco. En cada uno de estos disparos la probabilidad de acertar es igual a 3/4. Si acierta una vez recibe 12.8 Intis, si acierta dos veces recibe 32 intis, si acierta tres veces recibe 64 intis y si ninguno de los disparos da en el blanco, tiene que pagar - 320 intis. Calcular su ganancia esperada.

5. Lotes de 40 artículos de cierto producto son aceptados si ellos contienen no más de tres defectuosos. El plan de aceptación consiste en extraer una muestra al azar de 5 artículos y si se encuentra un artículo defectuoso - se rechaza el lote.

(a) Hallar la probabilidad de que se encuentre exactamente un defectuoso en la muestra, si el lote se considera en su calidad mínima. (un lote es de calidad máxima si no tiene defectuoso).

(b) ¿Cuántos defectuosos espera encontrar en la muestra?

6. Un radio técnico debe reemplazar una válvula defectuosa. En un maletín tiene cuatro válvulas de las cuales sólo sirve una de ellas. Si las selecciona al azar, una después de otra sin reposición, ¿cuál es el número esperado de válvulas que ha de probar para que pueda arreglar el receptor?

7. Ud. lanza una moneda tres veces. Si obtiene al menos dos caras se le permitirá lanzar un dado y recibirá tantos intis como puntos obtenga en el dado. ¿Qué cantidad de dinero espera ud. ganar en este juego?

8. En una feria se deben pagar 25 ¢. para participar en un juego que consiste en tirar anillos. Se dan tres anillos a una persona, la cual trata de lanzarlos uno por uno hacia una clavija. Se da un premio de 50 ¢, si se logra ensartar un anillo en la clavija; si se logra ensartar dos anillos, el premio es de $ 1; si se ensartan los tres, entonces se otorga un premio de $ 5. Suponiendo que la probabilidad de ensartar en la clavija sea de 0.10 en cada tirada, ¿cuál es la ganancia esperada si se juega una vez? ¿Diez veces?.

9. Un jugador A paga I/. 1.00 a otro jugador B y lanza 3 dados. El jugador A recibe I/. 2.00 si aparece 1 as; I/. 4.00 si aparece 2 ases y I/. 8.00 si aparecen 3 ases en los otros casos no recibe nada. Se pregunta:

(a) ¿Es equitativo el juego? Justifique su respuesta.

(b) Si no lo fuese, ¿Cuánto debería recibir A por sacar 3 ases?.

10. Un amigo A, le hace a otro B, la siguiente apuesta: "Te doy 3 bolitas que debes lanzarla sobre cuatro casilleros, de tal suerte que, al lanzar las tres bolitas ganas, si dos de ellas (solamente dos) caen en el mismo casillero". Se sabe que:

(1°) Al ejecutar el lanzamiento las bolitas caerán, siempre, en cualquiera de los 4 casilleros, en consecuencia,la probabilidad de que una bolita caiga en un casillero, es la misma para cualquiera de los casilleros.

(2°) A, le da a B, tres oportunidades; ganando B en la primera oportunidad que logre colocar 2 bolitas en algún casillero.

Si B gana, recibira de A, $ 219.70, determinar la cantidad que debe darle B, si pierde en el juego, debiendo ser este equitativo.

11. Un borracho llega a su casa y quiere abrir la puerta de entrada. En el llavero lleva cinco llaves y prueba una tras otra al azar. Suponga que se encuentra suficientemente despierto para eliminar de tentavias posteriores las llaves probadas sin exito. Se representa por X el número de llaves que prueba hasta que abre la puerta. Hallar el número esperado de llaves que se prueba.

12. En el juego de carnaval llamado Chuck-a-Luck, un jugador paga una cantidad a como derecho de entrada al juego. Elige después uno de los números 1,2,. . ., 6 y tira tres dados. Si en los tres dados sale el número elegido por el jugador éste cobra cuatro veces su entrada; si el número sale en dos de los dados, cobra tres veces la entrada, y si sale sólo en uno de los dados cobra el doble de la entrada. Si no sale el número elegido no cobra nada. Sea X el beneficio neto del jugador en una tirada de este juego. Suponiendo que los dados son buenos. Determinar E(X).

13. Se lanza una moneda hasta que salga cara. Hallar el número esperado de tiradas.

14. Se lanza un dado hasta que aparezca el 4 ó 5. Calcular el número de lanzamientos.

15. La compañía "ELECTRON-PERU" fabrica radios y televisores. Dicha compañía recibe los transistores en cajas de 100 transistores cada una. El departamento de recepción utiliza la siguiente regla de inspección. Se prueban cuatro transistores de cada caja. Si ninguno resulta defectuoso no se continúan examinando transistores de la caja. En caso contrario se prueban todos los transistores restantes. Determine el número esperado de transistores examinados por caja, si cada caja contiene exactamente el 10% de defectuosos.

16. Una empresa que lanzará al mercado un nuevo producto ha considerado la contratación de una póliza de seguro para cubrir posibles pérdidas en la operación. Consideran que si el lanzamiento es un fracaso total, las pérdidas serán de I/. 180,000.00; y si el lanzamiento del producto es sólo modestamente satisfactorio las pérdidas serán de sólo I/. 50,000.00. Los actuarios de la empresa aseguradora, basados en encuestas del mercado han determinado que las probabilidades de un fracaso total y de un lanzamiento modestamente satisfactorio son respectivamente, 0.01 y 0.05. Si se

ignoran otras pérdidas asociadas, ¿qué monto de primas debe cobrarse para salir sin ganar ni perder?

17. Un actuario, que es un estadístico empleado por una compañía de seguros, determina las primas de seguro que la compañía debe cobrar por determinada protección. Considere el problema de determinar la prima anual para un seguro de daños de automóvil de 200,000.00 intis. La poliza cubre un tipo de eventos (siniestros) que por experiencia pasada se sabe que ocurren a 3 de cada 5000 automovilistas cada año .

18. Los dos finalistas en un torneo de tenis juegan una serie de 3 juegos, donde el ganador recibe I/. 100,000 y el segundo recibe I/. 60,000. ¿Cuáles son las esperanzas matemáticas de los dos jugadores si,
    (a) tienen las mismas posibilidades;
    (b) el mejor jugador es favorito 3 a 1?

19. Como parte de un programa promocional, un fabricante de detergente ofrece un primer premio de I/. 90,000 y un segundo de I/. 30,000 para aquellas que aceptan en usar el nuevo producto (distribuido gratuito) y enviar su nombre en la etiqueta. Los ganadores serán seleccionados al azar en un programa de T.V.
    (a) ¿Cuál sería la esperanza matemática de cada concursante si enviaran sus nombres 1'500,00 personas?
    (b) ¿Vale la pena entonces gastar un inti en estampillas para enviar una etiqueta?

20. Una compañía de seguros acepta pagar al promotor de una fiesta campestre I/. 50,000 en caso que el evento tenga que ser cancelado por lluvia. Si el actuario de la compañía cree que una prima justa a pagar por este seguro sería I/. 2,000, ¿qué probabilidad asigna a la eventualidad de que la fiesta campestre tenga que ser cancelada por lluvia?

21. Un fabricante de televisores utiliza un cierto tipo de componente electrónico en el montaje de televisores a color. Cada televisor requiere 6 de estos componentes. Un componente defectuoso no puede ser detectado hasta que el televisor ha sido totalmente montado. El costo de detección, reparación y reposición de un componente defectuoso es $ 15. El fabricante ha estado comprando estos componentes en lotes de 100 a dos diferentes proveedores. El costo de compra por lote al proveedor **A** es de $ 100, en tanto que el costo de compra por lote al proveedor **B** es $ 120. Basadas en expe-

riencias anteriores, las calidades comparadas de los lotes comprados a los dos proveedores son las siguientes:

PROVEEDOR A

Número estimado de componentes defectuosos por lote	Probabilidad
1	0.30
2	0.25
3	0.20
4	0.15
5	0.10

PROVEEDOR B

Número estimado de componentes defectuosos por lote	Probabidad
1	0.60
2	0.30
3	0.10

¿A qué proveedor debe comprar los componentes electrónicos?

22. Un comerciante estima las ventas diarias de un cierto tipo de pan especial en la forma siguiente:

Venta diaria estimada unidades	Probabilidad
4	0.50
5	0.40
6	0.10

El costo por unidad de hogaza de pan es de 25 ¢ y el precio de venta es de 50 ¢. El pan debe ser ordenado con un día de anticipación y cada unidad no vendida en el día se entrega a una institución de beneficiencia al precio de 10 ¢ por unidad. ¿Cuántas unidades debe ordenar el comerciante para maximizar su utilidad esperada diaria?

23. Un fabricante está planeando la producción de una novedad de temporada. El fabricante estima que la demanda de este artículo está dada en la forma siguiente:

Número Unidades (X)	Probabilidades
1,000	1/4
2,000	1/2
3,000	1/4

El costo de producción y comercialización del artículo consiste en un costo base fijo de \$ 5,000 y un costo variable de \$ 1 por unidad. Si el precio de venta es de \$ 5 por unidad, ¿Cuál es la ganancia esperada para el fabricante?

24. La demanda para cierto artículo particular está caracterizado por la distribución de probabilidad siguiente :

$$p(d) = \begin{cases} k\,\dfrac{d^2}{16} & , \quad d = 1,2,3,4,5 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar el valor apropiado de k; (b) Calcular la media de la demanda; (c) la varianza de la demanda.

25. A, B y C cortan una baraja de 52 naipes sucesivamente en ese orden. El primero que saque corazón gana I/. 74. Las extracciones se hacen con reposición. Determinar la esperanza de cada jugador.

26. Un jugador lanza una moneda al aire, frente a otro jugador, gana 1 sol si sale sello y pierde 1 sol si sale cara. Supongamos que lanza una vez y si gana abandona el juego; en caso contrario tira otra vez. ¿Cuál es la ganancia esperada?

27. En el problema 7 de 2.2.Hallar la utilidad esperada diaria y la varianza.
28. En el problema 3 de 3.1. Determinar la utilidad esperada por noche y la varianza.

29. Determinar la media y la varianza de la variable aleatoria Y definida en en el problema 6 de 3.1.

30. Determinar la media y la varianza de la variable aleatoria definida en el problema 7 de 3.1.

31. Una casa de suministros eléctricos está rematando cierto número de artículos, entre ellos un lote de cuatro artículos de cierto tipo al precio de I/. 40.00 por todo el lote. Un comerciante puede vender los artículos en buen estado a I/. 20.00 cada uno, pero todo artículo defectuoso representa una pérdida completa de I/. 10.00. Basada en su amplia experiencia, el comerciante asigna probabilidades de 0.1, 0.5, 0.2, 0.1 y 0.1 a los eventos que haya 0.1,2,3. y 4 artículos defectuosos en el lote, respectivamente. Si no es posible ninguna inspección. ¿Deberá comprar el lote?

32. Tres jugadores A, B y C de igual habilidad juegan de la siguiente manera: A y B juegan la primera partida mientras que C descansa, el ganador sigue

jugando y el que pierde es reemplazado por el que no jugó, haciendo lo mismo después de cada partida. El juego continúa hasta que un jugador gane dos veces seguidas. Si el premio es de I/. 210. ¿Cuál es la ganancia esperada de cada jugador.

(**a**) después de la primera partida? (suponga que la primera partida gana **A**)

(**b**) al principio del juego?

33. En un sector de Lima metropolitana, hay una playa de estacionamiento que tiene una capacidad de acomodar 8 automóviles. Las tarifas de la playa de estacionamiento producen una utilidad de I/. 4. (4 intis) cada hora por automóvil estacionado. Sea **X** la variable aleatoria que representa el número de automóviles buscando estacionamiento por hora. Suponga que la función de probabilidad de **X** está dada por

$$p(x) = \frac{1}{2^{x+1}}, \qquad x = 0, 1, 2, \ldots$$

Hallar la utilidad esperada de la playa por hora.

34. Sea **X** una variable aleatoria cuya distribución de probabilidad es la siguiente:

x	0	1	2	3	4
$p(x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$

Calcular: (**a**) $E(2X + 1)$, $V(X)$, $V(2X + 1)$, $E(X^2 + 2X + 1)$

(**b**) La moda y la mediana de **X**;

(**c**) El tercer momento alrededor de la media.

35. La siguiente tabla representa el número de televisores vendidos en cierta semana de una tienda.

x	0	1	2	3	4	5
$P[X = x]$	0.05	0.1	0.35	0.25	0.2	0.05

Hallar: (**a**) $E(3X - 2)$; (**b**) $E(-6X + 10)$; (**c**) $E(2X^2 + 3X - 5)$

(**d**) La media, la varianza y la desviación estandar de **X**;

(**e**) La moda y el tercer momento alrededor de la media de **X**;

(**f**) $\text{Var}(X^2 + 2)$; $\text{Var}(-2X^2 + 5X - 1)$; (**g**) el 60-ésimo percentil ;

(**h**) el rango intercuartil.

36. Sea X una variable aleatoria cuya distribución de probabilidad está dado por

x	1	3	5	7	9
p(x)	0.1	0.4	0.2	0.2	0.1

Hallar: (a) 20-ésimo percentil; (b) 50-ésimo percentil ;
(c) el recorrido intercuartil;
(d) los primeros cuatro momentos iniciales (o al rededor del origen)
(e) los primeros cuatro momentos centrales (o al rededor de la media)
(f) la asimetría.

37. La variable aleatoria X con media μ y varianza σ^2 tiene función de densidad,

$$f(x) = kx^2 , \quad 0 \le x \le 5$$

Calcular (a) el valor de k; (b) $F(2)$; (c) $P[|X - \mu| < \sigma^2]$.

38. La función de distribución de una variable aleatoria continua X tiene la forma;

$$F(x) = \begin{cases} 0 & , \quad x \le -1 \\ a + b \ \text{arcsen}x & , \quad -1 < x \le 1 \\ 1 & , \quad x > 1 \end{cases}$$

(a) Hallar las constantes a y b ; (b) Calcular μ y σ^2.

39. Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} x & , \quad 0 \le x < 1/2 \\ 5x - 2 & , \quad 1/2 \le x \le 1 \end{cases}$$

Hallar $E(X)$, $E[(X - 4)^2]$

40. La media y la varianza de una variable aleatoria X son 50 y 4, respectivamente. Calcular:
(a) La media de X^2 ; (b) La varianza de $2X + 3$;
(c) la desviación estandar de $2X + 3$;
(d) La varianza de $-X$.

41. Sea X una variable aleatoria con función de densidad definida por,

$$f(x) = \begin{cases} \dfrac{x}{18} & , \quad 0 \le x \le 6 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

Si $Y = 10 + 2X$. Hallar la esperanza y la varianza de Y.

42. Una variable aleatoria X, que representa el peso (en onzas) de un artículo, tiene la función de densidad dada por,

$$f(x) = \begin{cases} x - 8 & , \quad 8 \leq x \leq 9 \\ 10 - x & , \quad 9 < x < 10 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

(a) Determine la media y la varianza de la variable aleatoria X.
(b) El fabricante vende un artículo por un precio fijo de \$ 2,000.00. El garantiza el reintegro del precio de venta a cualquier cliente que encuentra que el peso de su artículo es inferior a 8.25 onzas. El costo de producción está relacionado al peso de artículos de acuerdo a la expresión $(0.05)X + 0.30$. Exprese la variable aleatoria utilidad P, en términos de la variable aleatoria X, es decir, encuentre la función $H(X)$ tal que $P = H(X)$.
(c) Determine la utilidad esperada por artículo.

43. Una máquina produce un artículo que es revisado (inspección de 100 %) antes de ser despachado. El instrumento de medición es tal que es difícil leer entre 1 y $1\frac{1}{3}$ (datos codificados). Después que se realiza el proceso de revisión, la división medida tiene la siguiente densidad.

$$f(x) = \begin{cases} kx^2 & , \quad 0 \leq x \leq 1 \\ 1 & , \quad 1 < x \leq 1\frac{1}{3} \\ 0 & , \quad \text{en otro caso} \end{cases}$$

(a) Determine el valor de K;
(b) Qué fracción de los artículos estará fuera de la zona confusa? (estará entre 0 y 1?)
(c) Calcular la media y la varianza de esta variable aleatoria.

44. Sea X una variable aleatoria con función de densidad

$$f(x) = \frac{3x^2}{2} \quad , \quad -1 \leq x \leq 1$$

Calcular :
(a) La mediana (b) la moda
(c) el momento de orden 3 alrededor de la media.

45. Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} c(a - x) & , \quad 0 \leq x \leq a \\ 0 & , \quad \text{en otro caso} \end{cases}$$

Hallar: (a) La constante c ; (b) la moda
(c) la mediana ; (d) la desviación típica.

46. La estatura de adulto de un niño de 3 años tiene la misma posibilidad de estar comprendida en el intervalo de 5 pies 6 pulgadas y 5 pies 11 pulgadas. ¿Cuál es su estatura esperada de adulto? ¿Cuál es su varianza?

47. Una estación de gasolina recibe semanalmente gasolina. Las estadísticas anteriores sugieren que la distribución de probabilidad de las ventas semanales **X**, medidas en miles de galones está dada por:

$$f(x) = \begin{cases} x - 1 & , \quad 1 < x < 2 \\ 3 - x & , \quad 2 \leq x < 3 \\ 0 & , \quad \text{en otro caso.} \end{cases}$$

(a) Calcular la probabilidad de que las ventas semanales superen a 1,500 galones.
(b) Si la provisión semanal es siempre de 500 galones más que la venta de la semana anterior. Calcular la probabilidad de una provisión menor a 3 mil galones.
(c) Calcular los límites del intervalo, $\langle \mu - k\sigma, \; \mu + k\sigma \rangle$ que ocurre con una probabilidad 0.95.

48. Cierta aleación se forma al combinar la mezcla fundida de dos metales. La aleación que resulta contiene cierto porcentaje de plomo **X**, que puede considerarse como una variable aleatoria. Supongamos que **X** tiene la siguiente función de densidad

$$f(x) = \frac{3}{5} 10^{-5} x \; (100 - x) \; , \quad 0 \leq x \leq 100$$

Suponga que **P**, la utilidad neta obtenida al vender esta aleación (por libra), es la siguiente función del porcentaje del contenido de plomo.

$P = C_1 + C_2X$. Calcular la utilidad esperada (por libra)

49. Suponga que **X** es una variable aleatoria con función de densidad de probabilidad

$$f(x) = \begin{cases} a(x - 2)(4 - x) & , \quad 2 \leq x \leq 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar: (a) el valor de a ; (b) la moda ; (c) la mediana de X; (d) el rango intercuartil

50. Suponga que X es una variable aleatoria con función de distribución

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \sqrt{x} & , \quad 0 \leq x \leq 1 \\ 1 & , \quad x > 1 \end{cases}$$

Determinar: (a) la mediana ; (b) el rango intercuartil ; (c) el 40-ésimo percentil ; (d) el 60-ésimo percentil ; (e) el 90-ésimo percentil.

51. Suponga que X es una variable aleatoria con función de densidad

$$f(x) = \begin{cases} a x^2 & , \quad 0 \leq x < 1 \\ a(2 - x)^2 & , \quad 1 \leq x < 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar: (a) el valor de a ; (b) los primeros cuatro momentos iniciales ; (c) los primeros cuatro momentos centrales ; (d) la asimetría.

52. Si $E(X) = 17$ y $E(X^2) = 298$, use la desigualdad de CHEBYSHEV para determinar la cota inferior para: $P[10 < X < 24]$

53. Sea X una variable aleatoria con media $\mu = 2$ y varianza $\sigma^2 = 1$. Use la desigualdad de CHEBYSHEV para hallar una cota inferior para:

(a) $P[|X - 2| < 4]$ (b) $P[-3 < X < 7]$

54. Sea X una variable aleatoria con media $\mu = 10$ y varianza $\sigma^2 = 4$. Hallar :

(a) $P[|X - 10| \geq 3]$; (b) $P[|X - 10| < 3]$; (c) $P[5 < X < 15]$

(d) Hallar el valor de c, tal que

$$P[|X - 10| \geq c] \leq 0.04$$

55. (a) Calcule el peso medio de un melón cosechado por el agricultor del problema 8 de 2.3.
(b) Estime el peso total de la cosecha del agricultor.

56. (a) Calcule el peso medio de una piña cosechado por el agricultor del problema 13 de 2.3.
(b) Estime el peso total de la cosecha del agricultor.

57. Una empresa que lanzará al mercado un nuevo producto ha considerado la contratación de una póliza de seguro para cubrir posibles pérdidas en la operación. Consideran que si el lanzamiento es un fracaso total, las pérdidas serán de 80,000.00 dólares; y si el lanzamiento del producto es sólo modestamente satisfactorio las pérdidas serán de sólo 25,000.00 dólares. Los actuarios de la compañía aseguradora, basados en encuestas del mercado han determinado que las probabilidades de un fracaso total y de un lanzamiento modestamente satisfactorio son, 0.1 y 0.5 respectivamente. Si se ignoran otras pérdidas asociadas, ¿qué monto debe cobrarse como prima para salir sin ganar ni perder?.

58. Calcular $P[\mu - 2\sigma < X < \mu + 2\sigma]$

donde X es una variable aleatoria, con función de densidad

$$f(x) = \begin{cases} 6x(1-x) & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

59. Sea X una variable aleatoria del tipo mixto que tiene la siguiente función de distribución

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ x^2/4 & , \quad 0 \leq x < 1 \\ \dfrac{x+1}{4} & , \quad 1 \leq x < 2 \\ 1 & , \quad x \geq 2 \end{cases}$$

Hallar : (a) $E(X)$; (b) $Var(X)$

4

VARIABLES ALEATORIAS BIDIMENCIONALES

4.1 INTRODUCCION

Hasta aquí sólo hemos considerado variables aleatorias unidimensionales. Es decir, el resultado de un experimento hemos representado por un sólo número real x. Sin embargo, un experimento aleatorio puede generar dos o más variables aleatorias de importancia. Un investigador puede estar interesado en el comportamiento simultáneo de dos o más características numéricas de un experimento aleatorio. Por ejemplo, se extrae aleatoriamente un estudiante de cierta universidad, y se anota su estatura x y su peso y; consideremos el par (x,y) como un sólo resultado del experimento. Por lo tanto, es importante estudiar la *distribución de probabilidad conjunta* de las variables aleatorias bidimensionales denotado (X, Y) que se genera al cuantificar a cada estudiante por el par (x,y); así, por ejemplo $X(\text{Juan}) = x$, $Y(\text{Juan}) = y$; donde X representa la estatura e Y el peso. Estudiaremos también las *distribuciones marginales*, es decir, la distribución de cada una de las variables aleatorias X e Y. Daremos luego las *distribución condicional*. Previamente vamos a definir el concepto de variable aleatoria *bidimensional o vector aleatorio bidimensional*

4.2 VARIABLE ALEATORIA BIDIMENSIONAL

DEFINICION 4.2.1 Si Ω es el espacio muestral asociado a un experimento alea torio ε, y X e Y dos funciones, que asigna un número real $x = X(\omega)$, $y = Y(\omega)$ a cada uno de los elementos $\omega \in \Omega$; el par (X,Y) se llama *Variable aleatoria bidimensional o vector bidimensional* (Ver fig. 4.2.1).

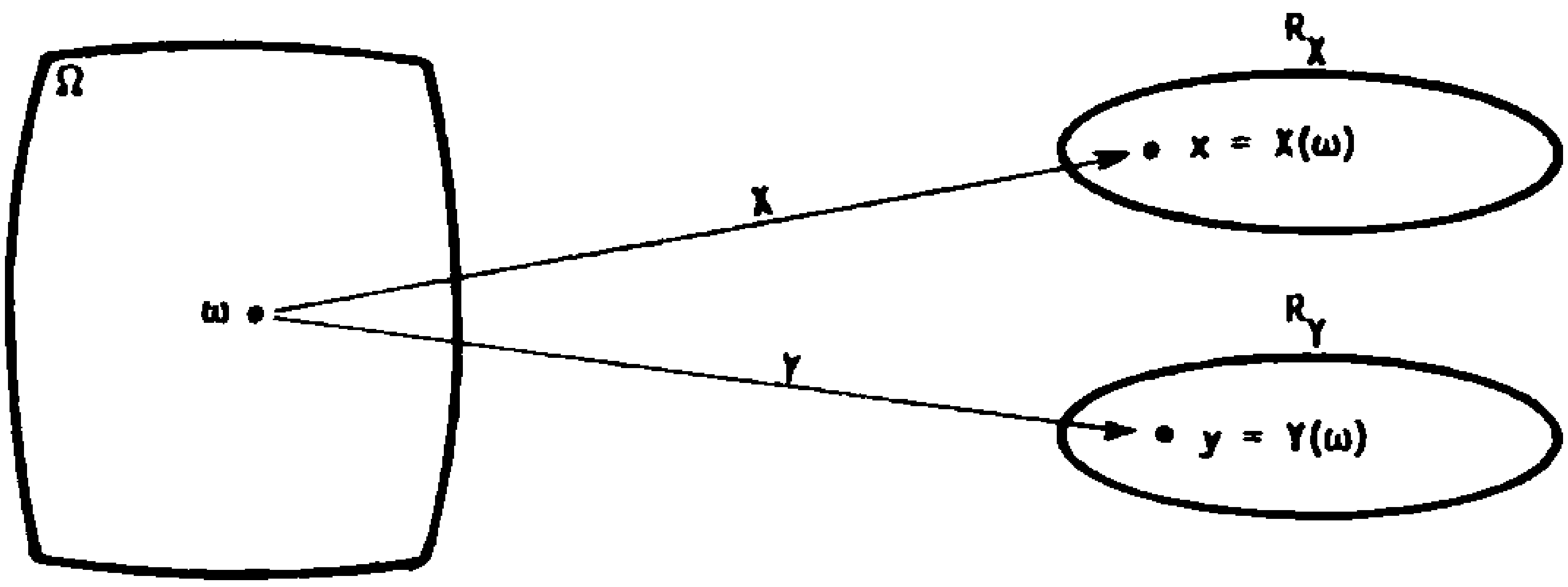


Fig. 4.2.1. Una variable aleatoria bidimencional.

La extensión de una variable aleatoria bidimensional a *n-dimensional* es directa, siguiendo los usos habituales del cálculo.

El rango de la variable aleatoria bidimensional (X,Y) es el conjunto de todos los valores posibles del vector aleatorio (X, Y), denota por $R_{X \times Y}$, - donde

$$R_{X \times Y} = \{(x,y) \;/\; x \in R_X,\;\; y \in R_Y\}$$

Es decir, es el producto cartesiano de los rangos de las variables aleatorias X e Y. La cual es un subconjunto del plano Euclidiano.

Extendiendo de manera natural las notaciones de eventos usados en el capítulo 2, tenemos, por ejemplo el evento $(X = a,\; Y = b)$, que se lee: "la variable aleatoria X toma el valor a y la variable aleatoria Y toma el valor b" similarmente $(X \leq a,\; Y \leq b)$, $(X \geq a,\; Y = b)$, etc. Las probabilidades de estos eventos se escribirá

$$P[X = a,\; Y = b]\;, \qquad P[X \geq a,\; Y = b] \;\; \text{etc.}$$

EJEMPLO 1 Una tienda comercial tiene dos vendedores A y B. Sea X el número de televisores vendidos en un día por A, e Y el número de televisores vendidos en un día por B.

Entonces, el par (X, Y) es una variable aleatoria bidimensional, donde el - rango de X, es, por ejemplo, $R_X = \{0,1,2, \ldots, n\}$, y el rango de Y es - $R_Y = \{0,1,2, \ldots, m\}$. Por lo tanto, el rango de la variable aleatoria bidimensional (X, Y) es $R_{X \times Y} = \{(x,y) \;/\; x \in R_X,\; y \in R_Y\}$

Se tiene también, por ejemplo, que $(X = 2,\ Y = 4)$ es el evento: "A vende 2 y B vende 4 televisores". $(X \geq 2,\ Y = 3)$ es el evento "A vende por lo menos 2 televisores y B vende 3" .

EJEMPLO 2 Una urna contiene 3 bolas numeradas 1,2,3 respectivamente. De la urna se extraen 2 bolas al azar una a una con reposición. Sea X el número de la primera bola que se extrae, e Y el número de la segunda bola que se extrae Se tiene que el par (X,Y) es una variable aleatoria bidimensional con rango

$$R_{X \times Y} = \{(x,y) \ / \ x,y \in \{1,2,3\}\} \quad \text{pués,}$$

$$R_X = R_Y = \{1,2,3\}$$

Como en el caso unidimensional, distinguiremos, dos tipos de variables aleatorias bidimensionales; discretas y contínuas. Puesto que en el par (X,Y) cada uno de ellas es una variable aleatoria, entonces es posible que : ambas sean discretas, ambas contínuas o que una sea discreta y la otra contínua; en este trabajo sólo se consideran los dos primeros. Si ambas son discretas, se dice que el par (X,Y) es una variable aleatoria bidimensional discreta; y si ambas son contínuas, se dice que (X,Y) es una variable aleatoria bidimensional contínua, formalizaremos esto en las definiciones siguientes.

DEFINICION 4.2.2 Si los posibles valores de (X,Y) son finitas o infinito numerable, entonces (X,Y) se llama *variable aleatoria bidimensional discreta* . Es decir, los posibles valores de (X,Y) se puede representar por,

$$(x_i, y_j), \quad i = 1,2,\ldots,n \qquad j = 1,2,\ldots,m.$$

Los dos ejemplos anteriores son variables aleatorias bidimensionales discretas.

DEFINICION 4.2.3 Si los posibles valores de (X,Y) son todos los valores de un conjunto no numerable del plano Euclideano, entonces (X,Y) se llama *variable aleatoria bidimensional contínua*.

Por ejemplo, si $a \leq x \leq b$ y $c \leq y \leq d$, se tiene que, $R_{X \times Y} = \{(x,y) \ / \ a \leq x \leq b,\ c \leq y \leq d\}$, son todos los valores de un rectángulo

EJEMPLO 3 Considere la extracción de un tornillo aleatoriamente de la producción de un día. Anotemos su diámetro y su peso. Si X representa el diámetro en pulgadas e Y representa el peso en libras, y si sabemos que $2.0 \leq x \leq 2.08$ pulgadas y $0.75 \leq y \leq 0.8$ libras; entonces el rango de (X,Y) es el conjunto ,

$$R_{X \times Y} = \{(x,y) \ / \ 2.0 \leq x \leq 2.08,\ 0.76 \leq y \leq 0.8\}$$

Esto se muestra en la figura 4.2.2

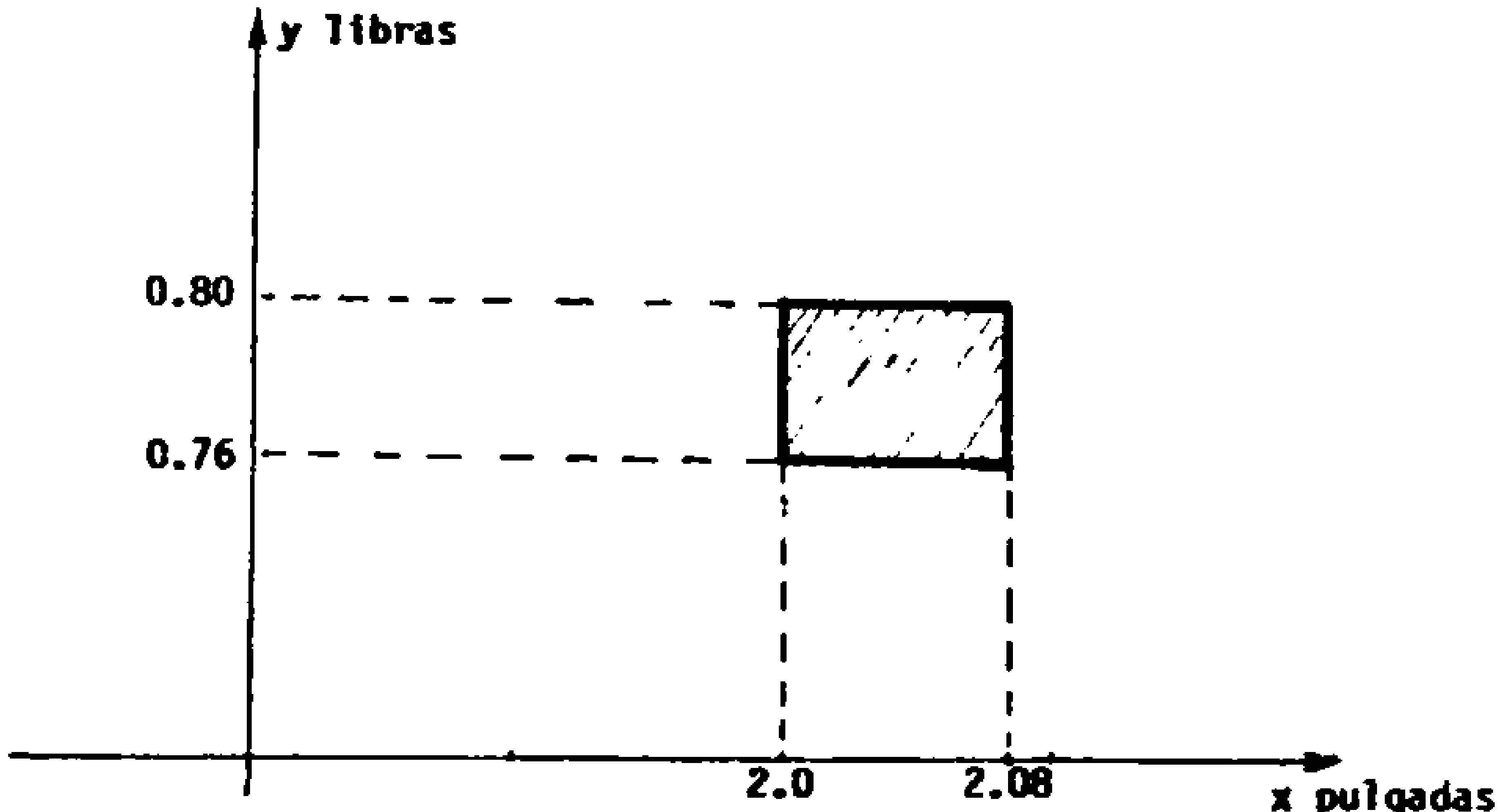


Fig. 4.2.2.

4.3 DISTRIBUCION BIDIMENSIONAL DISCRETA

DEFINICION 4.3.1 Sea (X,Y) una variable aleatoria bidimensional discreta con rango $R_{X\times Y}$. A cada posible resultado (x,y) de (X,Y), asociamos un número,

$$p(x,y) = P[X = x, \ Y = y]$$

llamado la *función de probabilidad conjunta*, que cumple las siguientes condiciones,

(1) $p(x,y) \geq 0 \ \forall \ (x,y)$;

(2) $\sum\sum_{(x,y)\in R_{X\times Y}} p(x,y) = 1$

Los valores $[(x,y), \ p(x,y)]$ para todo $(x,y) \in R_{X\times Y}$, se llama la *distribución de probabilidad conjunta.*

La probabilidad de un evento cualquiera A en $R_{X\times Y}(A \subset R_{X\times Y})$, está definido por,

$$\boxed{P[(X,Y) \in A] = \sum_{(x,y)\in A} p(x,y)}$$

Como en el caso unidimensional, la distribución de probabilidad conjunta de (X,Y) es inducida por la probabilidad de los sucesos del espacio muestral original Ω. Si la variable aleatoria bidimensional es finita, la distribución de probabilidad conjunta se representa en una tabla con dos entradas como se muestra en tabla 4.3.1.

Tabla 4.3.1. Representación Tabular

$y \backslash x$	x_1	x_2	.	.	.	x_{n-1}	x_n
y_1	$p(x_1, y_1)$	$p(x_2, y_1)$	.	.	.	.	$p(x_n, y_1)$
y_2	$p(x_1, y_2)$	$p(x_2, y_2)$	.	.	.	.	$p(x_n, y_2)$
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
y_m	$p(x_1, y_m)$	$p(x_2, y_m)$	.	.	.	.	$p(x_n, y_m)$

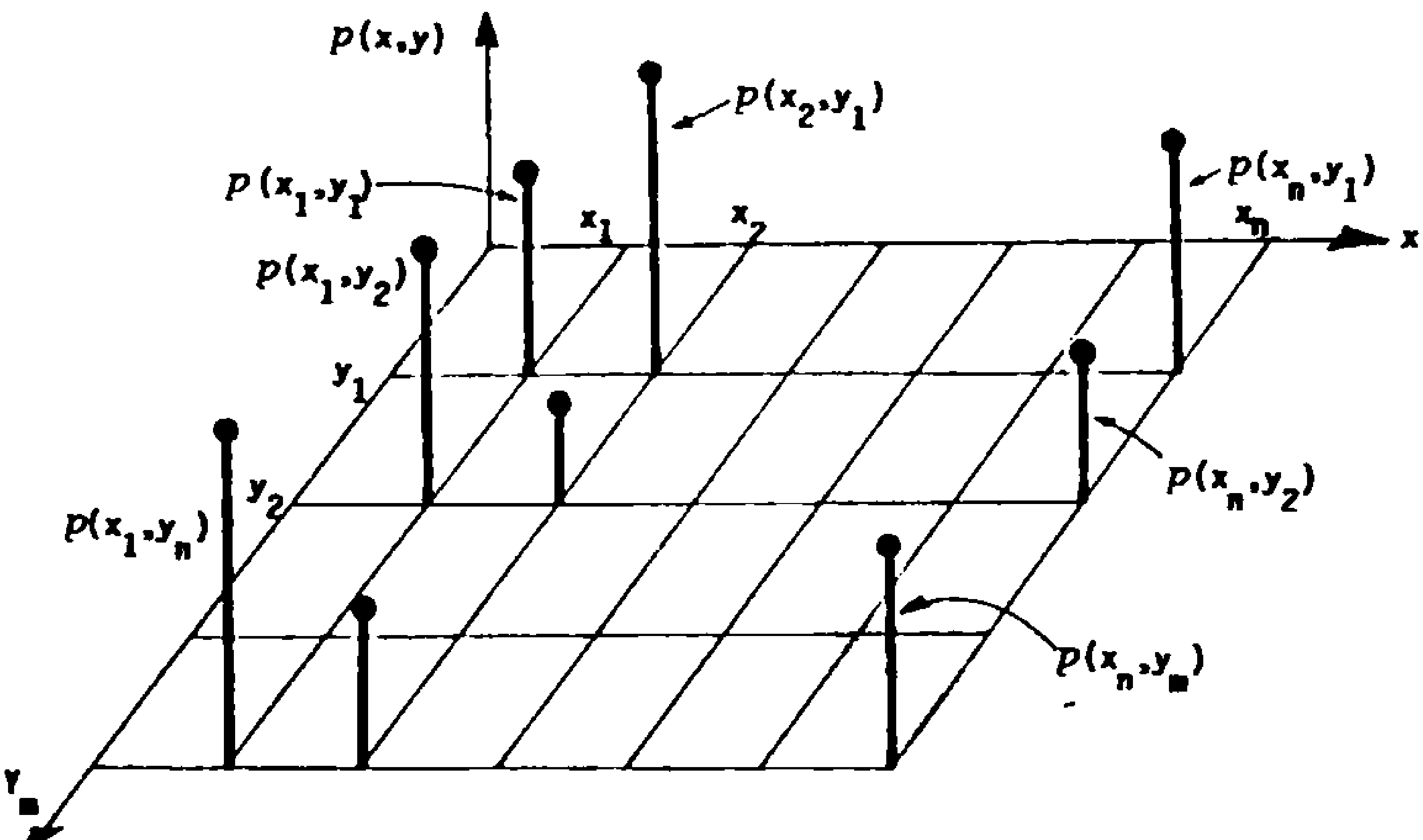


Fig. 4.3.1. Representación gráfica de una variable aleatoria bidimensional

DEFINICION 4.3.2 *La función de distribución acumulada* de la variable aleatoria bidimensional está definida por,

$$\boxed{F(x,y) = P[X \leq x,\ Y \leq y] = \sum_{u=-\infty}^{x} \sum_{v=-\infty}^{y} p(u,v)}$$

EJEMPLO 1 Hallar la distribución de probabilidad conjunta y su gráfica de la variable aleatoria bidimensional definida en el ejemplo 2 de 4.2.

SOLUCION Hemos visto que el rango de la variable aleatoria (X,Y) es,

$$R_{X \times Y} = \{(x,y) \;/\; x,y \in \{1,2,3\}\}$$

$$= \{(1,1),\ (1,2),\ (1,3),\ (2,1),\ (2,2),\ (2,3),\ (3,1),(3,2),\ (3,3)\}.$$

Luego,

$$p(1,1) = P[X = 1,\ \ Y = 1] = \frac{1}{9}$$

$$p(1,2) = P[X = 1,\ \ Y = 2] = \frac{1}{9}$$

$$p(1,3) = P[X = 1,\ \ Y = 3] = \frac{1}{9}\ ,\ \text{etc.}$$

Luego,

$$p(x,y) = \frac{1}{9}\ ,\quad x = 1,2,3;\ \ y = 1,2,3$$

es la función de probabilidad conjunta de X e Y.

Representación Tabular

y \ x	1	2	3
1	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
2	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
3	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

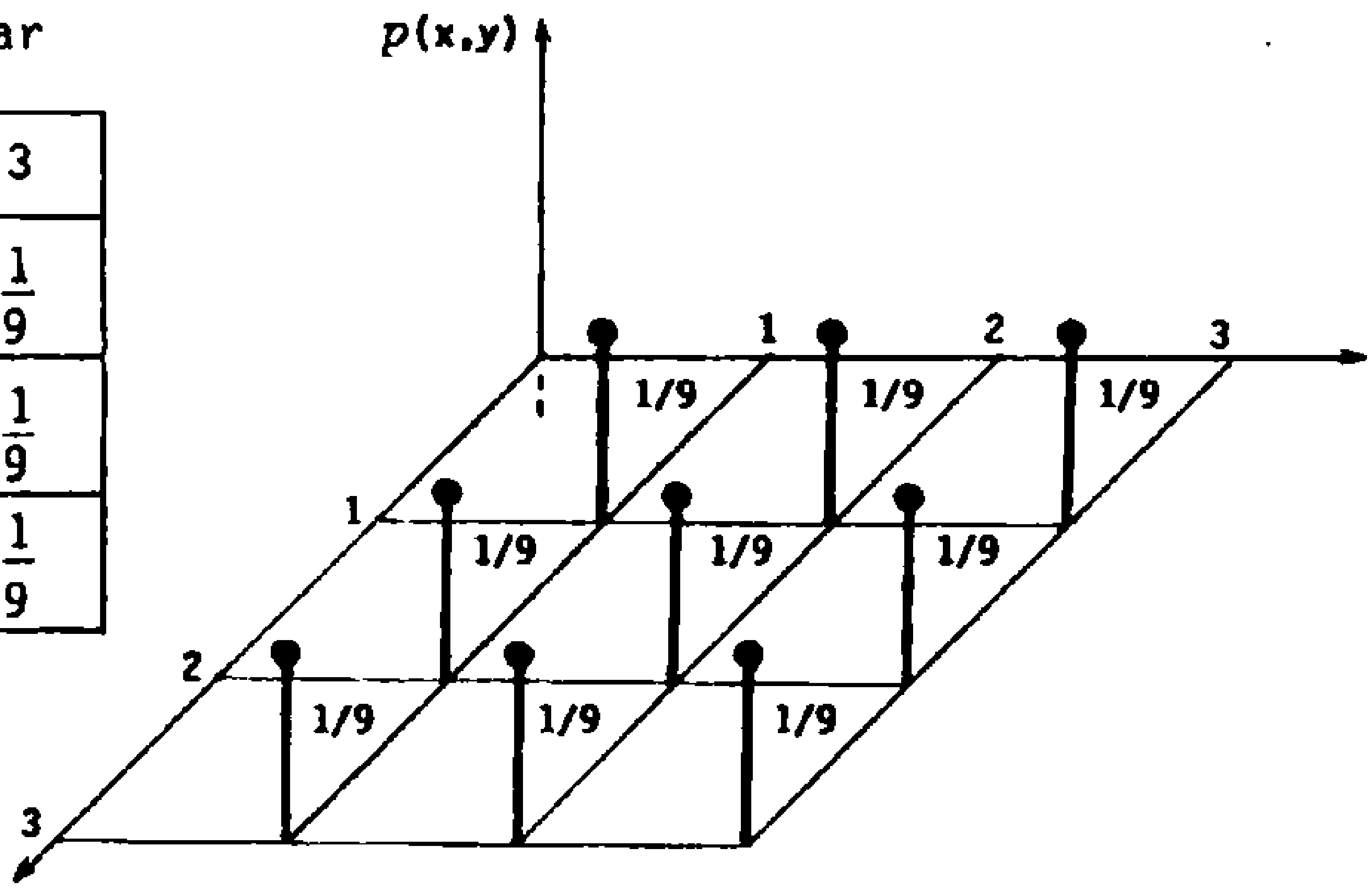


Fig. 4.3.2. Representación Gráfica del ejemplo 1 .

EJEMPLO 2 Sea (X,Y), la variable aleatoria bidimensional definida en el ejemplo 1 de 4.2. Suponga que de experiencias pasadas se sabe que la distribución de probabilidad conjunta de X e Y está dado por la siguiente tabla

y \ x	0	1	2
0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$

(a) ¿Cuál es la probabilidad que cada vendedor vende a lo más 1 televisor?

(b) ¿Cuál es la probabilidad que B venda más televisores que A?

SOLUCION (a) Debemos calcular la probabilidad del evento. $(X \leq 1,\ Y \leq 1)$

$$P[X \leq 1,\ Y \leq 1] = F(1,1) = P[X = 0,\ Y = 0] + P[X = 0,\ Y = 1] + P[X = 1,\ Y = 0] + P[X = 1,\ Y = 1]$$

$$= \frac{1}{16} + \frac{1}{8} + \frac{1}{16} + \frac{1}{8} = \frac{3}{8}\ .$$

(b) Se debe calcular la probabilidad del evento $(X < Y)$

$$P[X < Y] = P[X = 0,\ Y = 1,2] + P[X = 1,\ Y = 2]$$

$$= P[X = 0,\ Y = 1] + P[X = 0,\ Y = 2] + P[X = 1,\ Y = 2]$$

$$= \frac{1}{8} + \frac{1}{8} + \frac{1}{16} = \frac{5}{16}\ .$$

4.3.1 DISTRIBUCIONES MARGINALES

En algunos casos se puede estar interesado sólo en las distribuciones de X o de Y. Antes de dar la definición consideremos el siguiente ejemplo.

EJEMPLO 3 En el ejemplo anterior supongamos que estamos interesados, sólo en las ventas que hace el vendedor A. La distribución de probabilidad de X se obtiene de la tabla de distribución conjunta así,

$$p_X(0) = P[X = 0,\ Y = 0] + P[X = 0,\ Y = 1] + P[X = 0,\ Y = 2]$$

$$= \frac{1}{16} + \frac{1}{8} + \frac{1}{8} = \frac{5}{16}\ ;$$

$$p_X(1) = P[X = 1,\ Y = 0] + P[X = 1,\ Y = 1] + P[X = 1,\ Y = 2]$$

$$= \frac{1}{16} + \frac{1}{8} + \frac{1}{16} = \frac{4}{16}\ ;$$

$$p_X(2) = P[X = 2,\ Y = 0] + P[X = 2,\ Y = 1] + P[X = 2,\ Y = 2]$$

$$= \frac{1}{8} + \frac{1}{4} + \frac{1}{16} = \frac{7}{16}\ ;$$

En otras palabras, la probabilidad que la variable aleatoria X toma el valor 0 es 5/16, que toma el valor 1 es 4/16, y que toma el valor 2 es 7/16. Y se obtiene sumando las columnas de la tabla de distribución de probabilidad del ejemplo anterior. Luego la distribución de probabilidad de X ($p_X(x) = P[X = x]$) es,

x	0	1	2
$p_X(x)$	$\frac{5}{16}$	$\frac{4}{16}$	$\frac{7}{16}$

Similarmente sumando las filas de la tabla de distribución conjunta de X e Y, se obtiene la distribución de Y,

$$(p_Y(y) = P[Y = y])$$

y	0	1	2
$p_Y(y)$	$\frac{4}{16}$	$\frac{8}{16}$	$\frac{4}{16}$

Las distribuciones de X e Y así calculadas, se llaman *distribución de probabilidad marginal.*

Es, muchas veces, conveniente escribir estas sumas en los márgenes de la tabla de distribución conjunta de (X,Y), como se observa en la tabla siguiente

y \ x	0	1	2	Suma $P[Y = y]$
0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{4}{16}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{8}{16}$
2	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{4}{16}$
Suma $P[X = x]$	$\frac{5}{16}$	$\frac{4}{16}$	$\frac{7}{16}$	1

Distribución marginal de Y.

Distribución marginal de X.

DEFINICION 4.3.3 La función de probabilidad individual para X e Y se llaman *distribuciones de probabilidad marginal o funciones de probabilidad marginal*. *La Distribucion Marginal para* X, está dado por,

$$\boxed{p_X(x) = P[X = x] = \sum_{y \in R_Y|X=x} P[X = x,\ Y = y] = \sum_{y \in R_Y|X=x} p(x,y)}$$

.... (1)

donde $y \in R_Y|X = x$ es el rango de valores para Y, dado que X = x. Note que este rango puede ser diferente para valores diferentes de x.

La Distribución Marginal para Y, está dado por,

$$p_Y(y) \;=\; P[Y = y] \;=\; \sum_{x \in R_X | Y=y} p(x,y) \qquad (2)$$

El nombre de marginal se debe a que los valores de estas distribuciones están dadas por la suma total en los márgenes de las filas y columnas de la representación tabular de $p(x,y)$.

EJEMPLO 4 La función de probabilidad conjunta de la variable aleatoria (X,Y) está definido por,

$$p(x,y) \;=\; \frac{x + y}{32}, \qquad x = 1,2; \qquad y = 1,2,3,4.$$

Hallar la distribución de probabilidad marginal de X y de Y

SOLUCION

(a) $$p_X(x) \;=\; P[X = x] \;=\; \sum_{y \in R_Y | X=x} p(x,y) \;=\; \sum_{y=1}^{4} \frac{x + y}{32}$$

$$= \frac{x+1}{32} + \frac{x+2}{32} + \frac{x+3}{32} + \frac{x+4}{32}$$

Luego, $p_X(x) \;=\; P[X = x] \;=\; \dfrac{2x + 5}{16}, \qquad x = 1,2$

es la función de probabilidad marginal para X.

(b) $$p_Y(y) \;=\; P[Y = y] \;=\; \sum_{x \in R_X | Y=y} p(x,y) \;=\; \sum_{x=1}^{2} \frac{x + y}{32}$$

$$= \frac{1 + y}{32} + \frac{2 + y}{32}$$

Luego, $p_Y(y) \;=\; P[Y = y] \;=\; \dfrac{2y + 3}{32}, \qquad y = 1,2,3,4$

es la probabilidad marginal de Y.

4.3.2 VARIABLES ALEATORIAS INDEPENDIENTES

El concepto de independencia de variables aleatorias es muy importante en estadística. Como en el caso de independencia de eventos, dado en el capítulo 1; intuitivamente diremos que las variables aleatorias X e Y son *independientes* si el resultado de una de las variables, digamos X no influye en el resultado de Y, y viceversa.

DEFINICION 4.3.4 Sea (X,Y) una variable aleatoria bidimensional discreta con función de probabilidad marginal $p_X(x)$ y $p_Y(y)$. Se dice que X e Y son independientes si, sólo, sí

$$p(x,y) = p_X(x)\ p_Y(y) \quad , \quad \forall\ (x,y)$$

o que $\quad P[X = x,\ Y = y] = P[X = x]P[Y = y]\ ,\ \forall\ (x,y)$

En otras palabras, X e Y son independientes si, $P[X = x]$ no afecta a la $P[Y = y]$.

EJEMPLO 5 Las variables aleatorias definidas en el ejemplo 2 . no son independientes, puesto que por un lado se tiene,

$$p(0,0) = \frac{1}{16}$$

por otro lado se obtiene,

$$p_X(0)\ p_Y(0) = \frac{5}{16} \times \frac{4}{16} = \frac{5}{64}$$

de donde, $\quad p(0,0) \neq p_X(0)p_Y(0)$

EJEMPLO 6 Suponga que X e Y tienen la siguiente distribución de probabilidad conjunta,

y \ x	2	4
1	0.10	0.15
3	0.20	0.30
5	0.10	0.15

Determine si las variables aleatorias X e Y son independientes.

SOLUCION La distribución marginal de cada variable está dada en la tabla siguiente,

y \ x	2	4	$P[Y = y]$
1	0.10	0.15	0.25
3	0.20	0.30	0.50
5	0.10	0.15	0.25
$P[X = x]$	0.40	0.60	1

$p_X(2)p_Y(1) = (0.40)(0.25) = 0.10$, Luego, $p(2,1) = p_X(2)p_Y(1)$

$p_X(2)p_Y(3) = (0.40)(0.50) = 0.20$, Luego, $p(2,3) = p_X(2)p_Y(3)$

$p_X(2)p_Y(5) = (0.40)(0.25) = 0.10$, Luego, $p(2,5) = p_X(2)p_Y(5)$

$p_X(4)p_Y(1) = (0.60)(0.25) = 0.15$, Luego, $p(4,1) = p_X(4)p_Y(1)$

$p_X(4)p_Y(3) = (0.60)(0.50) = 0.30$, Luego, $p(4,3) = p_X(4)p_Y(3)$

$p_X(4)p_Y(5) = (0.60)(0.25) = 0.15$; Luego, $p(4,5) = p_X(4)p_Y(5)$

Por lo tanto, concluímos que X e Y son independientes.

4.3.3 DISTRIBUCION DE PROBABILIDAD CONDICIONAL

Dado dos eventos A y B, en el capítulo I hemos presentado, la probabilidad - condicional de que ocurra el evento A dado que ha ocurrido el evento B. Extenderemos este concepto a la distribución condicional de una variable aleatoria X, dado que la variable aleatoria Y haya tomado un valor determinado , por, ejemplo y_o.

Sea (X,Y) una variable aleatoria bidimencional con función de probabilidad - $p(x,y)$ de rango $R_{X\times Y}$, y probabilidades marginales $p_X(x)$, $p_Y(y)$. R_X y R_Y - los rangos de las variables aleatorias X e Y respectivamente.

Sean los eventos $A = (X = x)$, $B = (Y = y)$, $(x,y) \in R_{X\times Y}$, entonces , por la definición de probabilidad condicional se tiene.

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

Pero, $A \cap B = (X = x,\ Y = y)$, por lo tanto

$$P[A \cap B] = P[X = x\ ,\ Y = y] = p(x,y) \quad y \quad .$$

$$P[B] = P[Y = y] = p_Y(y) > 0, \text{ pues } y \in R_Y .$$

Luego, la probabilidad condicional del evento A, dado B es

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[X = x,\ Y = y]}{P[Y = y]} = \frac{p(x,y)}{p_Y(y)}$$

Es decir,

$$\boxed{P[X = x|Y = y] = \frac{P[X = x,\ Y = y]}{P[Y = y]} = \frac{p(x,y)}{p_Y(y)}}$$

La razón $\dfrac{p(x,y)}{p_Y(y)}$, se llama la función de *probabilidad condicional de* X, *dado* $Y = y$. La siguiente definición formaliza lo expuesto en el párrafo anterior.

DEFINICION 4.3.5 Si (X,Y) es una variable aleatoria bidimensional discreta, la *función de probabilidad condicional de* X *dado* $Y = y$, se denota por "$p_{X|Y}(x|y)$" y se define por,

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)} \quad , \quad p_Y(y) > 0$$

Similarmente, la *función de probabilidad condicional de* Y *dado* $X = x$, se define por,

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)} \quad , \quad p_X(x) > 0$$

EJEMPLO 7 Sean X e Y dos variables aleatorias cuya función de probabilidad conjunta esta definida por,

$$p(x,y) = \frac{x + y}{32} \quad , \quad x = 1,2; \quad y = 1,2,3,4.$$

Hallar :

(a) la función de probabilidad condicional de X, dado $Y = y$.

(b) la función de probabilidad condicional de Y, dado $X = x$.

SOLUCION En el ejemplo 4 hemos calculado las respectivas funciones de probabilidad marginal,

$$p_X(x) = \frac{2x+5}{16}, \qquad x = 1,2$$

$$p_Y(y) = \frac{2y+3}{32}, \qquad y = 1,2,3,4.$$

Entonces:

(a) La función de probabilidad condicional de X, dado $Y = y$ es,

$$p_{X|Y}(x|y) = \frac{\frac{x+y}{32}}{\frac{2y+3}{32}} = \frac{x+y}{2y+3}, \qquad x = 1,2;$$

cuando $y = 1$ ó 2 ó 3 ó 4.

por ejemplo,

$$P[X = 2|Y = 2] = p_{X|2}(2|2) = \frac{4}{7}.$$

(b) La función de probabilidad condicional de Y, dado $X = x$, es

$$p_{Y|X}(y|x) = \frac{\frac{x+y}{32}}{\frac{2x+5}{16}} = \frac{x+y}{4x+10}, \quad \text{para, } x = 1,2,3,4,$$

cuando $x = 1$ ó 2.

por ejemplo, $P[Y = 3|X = 1] = p_{Y|1}(3|1) = \frac{4}{14} = \frac{2}{7}$

EJEMPLO 8 La distribución de probabilidad conjunta de la variable aleatoria (X, Y) está dado por.

Hallar:

(a) $P[X = 1|Y = 0]$

(b) $P[Y = 2|X = -1]$

y \ x	-1	1
0	$\frac{1}{6}$	$\frac{3}{12}$
1	0	$\frac{4}{12}$
2	$\frac{1}{12}$	$\frac{2}{12}$

SOLUCION La distribución de probabilidad marginal, $p_X(x) = P[X = x]$ y $p_Y(y) = P[Y = y]$, están dadas por la suma de las columnas y filas de la tabla de distribución de probabilidad conjunta.

$y \backslash x$	-1	1	$P[Y = y]$
0	$\frac{1}{6}$	$\frac{3}{12}$	$\frac{5}{12}$
1	0	$\frac{4}{12}$	$\frac{4}{12}$
2	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$
$P[X = x]$	$\frac{3}{12}$	$\frac{9}{12}$	1

(a) $P[X = 1|Y = 0] = \dfrac{P[X = 1,\ Y = 0]}{P[Y = 0]} = \dfrac{3/12}{5/12} = \dfrac{3}{5}$.

(b) $P[Y = 2|X = -1] = \dfrac{P[X = -1,\ Y = 2]}{P[X = -1]} = \dfrac{1/12}{3/12} = \dfrac{1}{3}$.

El lector habrá notado que, $p_{X|Y}(x|y) > 0$, y

$$\sum_{x \in R_X|Y=y} p_{X|Y}(x|y) = \sum_{x \in R_X|Y=y} \frac{p(x,y)}{p_Y(y)} = \frac{\sum_{x \in R_X|Y=y} p(x,y)}{p_Y(y)}$$

$$= \frac{p_Y(y)}{p_Y(y)} = 1 .$$

Es decir, $p_{X|Y}(x|y)$ cumple las condiciones de una función de probabilidad , entonces se puede calcular probabilidades tales como,

$$\boxed{P[a \leq X \leq b|Y = y] = \sum_{x \in R_X|a \leq x \leq b} p_{X|Y}(x|y)}$$

$$\boxed{P[X < a|Y = y] = \sum_{x \in R_X|x < a} p_{X|Y}(x|y)} \quad , \text{etc.}$$

por ejemplo, en el problema anterior, la probabilidad condicional del evento $(0 < x < 3/2)$, dado que $Y = 2$, es

$$P[0 < X < \tfrac{3}{2}|Y = 2] = P[X = 1|Y = 2]$$

$$= \frac{P[X = 1, \; Y = 2]}{P[Y = 2]} = \frac{2/12}{3/12} = \frac{2}{3} .$$

TEOREMA 4.3.1 Sea (X,Y) una variable aleatoria bidimensional, entonces,

$$p_{X|Y}(x|y) = p_X(x) \quad , \quad \forall\ (x,y)$$
$$p_{Y|X}(y|x) \quad p_Y(y) \quad , \quad \forall\ (x,y)$$

sí, sólo si X e Y son variables aleatorias independientes .

4.3.4 ESPERANZA Y VARIANZA

En el capítulo 3 , hemos calculado, la media o valor esperado de una variable aleatoria X, la media de funciones de X, como por, ejemplo: X^2; $aX + b$; etc. En algunos problemas, se requiere calcular, el valor esperado de una función de dos o más variables aleatorias X e Y, por ejemplo, las mas simples: X + Y, XY, $X^2Y + X$, etc. La técnica para obtener el valor esperado de una función de dos variables aleatorias es la misma que en el caso de una variable, como indica la definición siguiente.

DEFINICION 4.3.6 Sea (X,Y) una variable aleatoria bidimensional discreta, y H(X,Y) una función de dos variables aleatorias X e Y. El ***valor esperado de*** H(X,Y), se define por,

$$E\ H(X,Y) = \sum_{R_{X \times Y}} \sum H(x,y)\ p(x,y)$$

Siempre que la suma sea absolutamente convergente.
En el caso particular de que H(X,Y) = X, se tiene

$$\mu_X = E(X) = \sum_{R_{X \times Y}} \sum x p(x,y) = \sum_{x \in R_X} x \sum_{y \in R_Y} p(x,y)$$
$$= \sum_{x \in R_X} x p_X(x)$$

Similarmente, si H(X,Y) = Y, $\mu_Y = E(Y) = \sum_{y \in R_Y} y p_Y(y)$. Es decir, la media de X(Y), se determina, de la función de probabilidad marginal de X(Y).
El lector, siguiendo un proceso similar, puede encontrar que la varianza de

X y de Y estan dados por

$$\sigma_X^2 = E(X - \mu_X)^2 = \sum_{x \in R_X} (x - \mu_X)^2 p_X(x) = E(X^2) - (\mu_X)^2$$

$$\sigma_Y^2 = E(Y - \mu_Y) = \sum_{y \in R_Y} (y - \mu_Y)^2 p_Y(y) = E(Y^2) - (\mu_Y)^2$$

observe que, aquí es,

$$H(X,Y) = (X - \mu_X)^2 \; ; \quad H(X,Y) = (Y - \mu_Y)^2$$

TEOREMA 4.3.2 Si (X,Y) es una variable aleatoria bidimensional, entonces,

$$E[aH(X) \pm bG(Y)] = aE[H(X)] \pm bE[G(Y)]$$

una consecuencia inmediata de este teorema, cuando,

$$H(X) = X, \text{ y } G(Y) = Y, \text{ es,}$$

$$E(X \pm Y) = E(X) \pm E(Y)$$

TEOREMA 4.3.3 Si X e Y son dos variables aleatorias independientes, entonces

(a) $E(XY) = E(X)\,E(Y)$

(b) $Var(X + Y) = Var(X) + Var(Y)$

Una generalización de la parte (a) de este teorema es la siguiente

Si X e Y son variables aleatorias independientes y H(X), G(Y) son funciones de X e Y respectivamente, entonces,

$$E[H(X)\,G(Y)] = E[H(X)]\,E[G(Y)]$$

También la parte (b) se puede generalizar a n variables aleatorias.

EJEMPLO 9 La variable aleatoria bidimensional (X,Y), tiene la siguiente distribución de probabilidad conjunta.

y \ x	0	1
1	$\frac{1}{6}$	$\frac{1}{12}$
2	$\frac{2}{6}$	$\frac{5}{12}$

Calcular:

(a) $E(X + Y)$;

(b) $E(XY)$;

(c) $E(XY^2)$, $E(2XY + XY^2)$.

SOLUCION Cálculo de la distribución de probabilidad marginal.

y \ x	0	1	$P[Y = y]$
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{3}{12}$
2	$\frac{2}{6}$	$\frac{5}{12}$	$\frac{9}{12}$
$P[X = x]$	$\frac{3}{6}$	$\frac{6}{12}$	1

(a) $$\mu_X = E(X) = \sum_{x \in R_X} x p_X(x) = 0(\tfrac{3}{6}) + 1(\tfrac{6}{12}) = \frac{6}{12} \quad ;$$

$$\mu_Y = E(Y) = \sum_{x \in R_Y} y p_Y(y) = 1(\tfrac{3}{12}) + 2(\tfrac{9}{12}) = \frac{21}{12} \quad ;$$

y por el teorema 1, se tiene,

$$E(X + Y) = E(X) + E(Y) = \frac{6}{12} + \frac{21}{12} = \frac{9}{4} \ .$$

(b) Puesto que X e Y no son independientes y por lo tanto no podemos aplicar el teorema 3, aplicaremos la definición. Entonces,

$$E(XY) = \sum_{R_{X \times Y}}\sum xyp(xy) = 0.1.(\tfrac{1}{6}) + 0.2.(\tfrac{2}{6}) + 1.1.(\tfrac{1}{12}) + 1.2.(\tfrac{5}{12}) = \frac{11}{12} \ .$$

(c) $$E(XY^2) = \sum_{R_{X \times Y}}\sum xy^2p(x,y) = 0.1.(\tfrac{1}{6}) + 0.4.(\tfrac{2}{6}) + 1.1.(\tfrac{1}{12}) + 1.4.(\tfrac{5}{12}) = \frac{21}{12} \ .$$

(d) aplicando el teorena 2 y (c) se tiene,

$$E(2XY + XY^2) = 2E(XY) + E(XY^2) = 2 \cdot \frac{11}{12} + \frac{21}{12} = \frac{43}{12} \ .$$

EJEMPLO 10 Sean X e Y dos variables aleatorias, con función de probabilidad conjunta,

$$p(x,y) = \frac{x + 2y}{18} \quad , \quad x = 1,2, \quad y = 1,2$$

Calcular:

(a) μ_X ; (b) σ_X^2 ; (c) μ_Y ; (d) σ_Y^2

SOLUCION Las funciones de probabilidad marginal son, respectivamente

$$p_X(x) = \sum_{y=1}^{2} \frac{x + 2y}{18} = \frac{2x + 6}{18} \quad , \quad x = 1,2$$

$$p_Y(y) = \sum_{x=1}^{2} \frac{x + 2y}{18} = \frac{3 + 4y}{18} \quad , \qquad y = 1,2$$

Observe que $p(x,y) \neq p_x(x)\, p_y(y)$, por lo tanto **X** e **Y** son dependientes.

(a) $\mu_X = E(X) = \sum_{x=1}^{2} x(\frac{2x + 6}{18}) = 1(\frac{8}{18}) + 2(\frac{10}{18}) = \frac{14}{9}$.

(b) También $E(X^2) = \sum_{x=1}^{2} x^2(\frac{2x + 6}{18}) = 1.(\frac{8}{18}) + 4(\frac{10}{18}) = \frac{24}{9}$.

Luego, $\sigma_X^2 = \frac{24}{9} - (\frac{14}{9})^2 = \frac{20}{81}$.

(c) $\mu_Y = E(Y) = \sum_{y=1}^{2} y\ (\frac{3 + 4y}{18}) = 1(\frac{7}{18}) + 2(\frac{11}{18}) = \frac{29}{18}$.

(d) $E(Y^2) = \sum_{y=1}^{2} y^2(\frac{3 + 4y}{18}) = 1\ (\frac{7}{18}) + 4(\frac{11}{18}) = \frac{51}{18}$.

Luego, $\sigma_Y^2 = \frac{51}{18} - (\frac{29}{18})^2 = \frac{77}{324}$.

4.3.5 ESPERANZA CONDICIONAL

La esperanza condicional (o media condicional), se define de la misma forma que una esperanza ordinaria, excepto que se usa la probabilidad condicional en vez de la probabilidad ordinaria. Esto se justifica debido a que la probabilidad condicional, cumple las condiciones de una función de probabilidad, como hemos visto anteriormente en el capítulo I.

DEFINICION 4.3.7 Sea (X,Y) una variable aleatoria bidimensional y sea y un número real, tal que $P[Y = y] > 0$. El ***valor esperado condicional de* X *dado*** Y = y, se denota por,

$\mu_{X|y}$ y se define por,

$$\boxed{\mu_{X|y} = E(X|Y = y) = \sum_{x \in R_X} xP[X = x|Y = y] = \sum_{x \in R_X} x p_{X|y}(x|y)}$$

Similarmente,

$$\mu_{Y|x} = E(Y|X = x) = \sum_{y \in R_y} yP[Y = y|X = x] = \sum_{y \in R_y} yp_{Y|x}(y|x)$$

En particular, si $P[X = x_o] > 0$, se tiene,

$$\mu_{X|x_o} = E(X|X = x_o) \sum_{x \in R_X} xP[X = x|X = x_o]$$

$$\sum_{x \in R_X} x \frac{P[X = x, X = x_o]}{P[X = x_o]}$$

$$= x_o \frac{P[X = x_o]}{P[X = x_o]} = x_o$$

Observe que habrá una $E(X|y)$ para cada valor de y en el rango de R_Y, y el valor de cada $E(X|y)$ dependerá del valor de y, que se reemplace en la - función de probabilidad condicional. Similarmente, habrá muchos valores de - $E(Y|x)$ como valores tiene x (rango de X), y el valor de $E(Y|x)$ dependerá de los valores de x, determinado por la función de probabilidad.

DEFINICION 4.3.8 La *varianza condicional de* X *dado* $Y = y$, se define por,

$$\sigma^2_{X|y} = E[(X - E(X|y))^2|y] = \sum_{x \in R_X} (x - E(X|y))^2 p(x|y)$$

Como en el caso de varianza ordinaria, puede usarse la propiedad,

$$\sigma^2_{X|y} = E(X^2|y) - (E(X|y))^2$$

La varianza $\sigma^2_{Y|x}$, se define de manera similar.

Las propiedades de la esperanza condicional son completamente similares a las de esperanza ordinaria.

TEOREMA 4.3.4 Sea (X,Y) una variable aleatoria bidimensional, y sean a, b e y números reales. H(X,Y) una función de dos variables. Entonces:

(a) $E(aX + b|Y = y) = aE(X|Y = y) + b \quad \forall$ y tal que $p_Y(y) > 0$

(b) $E[H(X,Y)|Y = y] = \sum_{x \in R_X} H(x,y)P[X = x|Y = y] = E[H(X,Y)|Y = y]$

En (a) Si $b = 0$, se tiene que $E(aX|Y = y) = aE(X|Y = y)$

TEOREMA 4.3.5 Si X e Y son variables aleatorias independientes, entonces,

$$E(X|Y = y) = E(X), \quad \forall\ y \text{ , que cumple } p_Y(y) > 0$$

EJEMPLO 11 La distribución de probabilidad conjunta de (X,Y) está definida por la siguiente tabla. Calcular :

(a) $E(X|Y = 1)$;

(b) $E(2X|Y = 1)$;

(c) $E(3X + 4|Y = 1)$;

(d) $E(XY|Y = 1)$;

(e) $E(2X + 3Y|Y = 1)$;

y \ x	-1	1
1	$\frac{1}{8}$	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{8}$

SOLUOION

(a) $$E(X|Y = 1) = \sum_{x \in R_X} xP[X = x|Y = 1]$$

$$= (-1)P[X = -1|Y = 1] + 1.\ P[X = 1|Y = 1]$$

$$= (-1)(\tfrac{1}{3}) + 1\ .\ (\tfrac{2}{3}) = \frac{1}{3}\ .$$

Ya que, $$P[X = -1|Y = 1] = \frac{P[X = -1,\ Y = 1]}{P[Y = 1]} = \frac{1/8}{3/8} = \frac{1}{3}$$

y $$P[X = 1|Y = 1] = \frac{P[X = 1,\ Y = 1]}{P[Y = 1]} = \frac{1/4}{3/8} = \frac{2}{3}$$

(b) $E(2X|Y = 1) = 2E(X|Y = 1)$ por teor 4.3.4*a*,

$$= 2(\tfrac{1}{3}) = \frac{2}{3}\ .$$

(c) $E(3X + 4|Y = 1) = 3E(X|Y = 1) + 4$ por teor. 4.3.4*a*,

$$= 3(\tfrac{1}{3}) + 4 = 5.$$

(d) $E(XY|Y = 1) = E(X.\ 1|Y = 1)$ por teor. 4.3.4*b*,

$$= E(X|Y = 1) = \frac{1}{3}\ .$$

(e) $E(2X + 3Y\ Y = 1) = E(2X + 3.1|Y = 1)$ por teor. 4*b*,

$$= E(2X + 3|Y = 1)$$

$$= 2E(X|Y = 1) + 3$$

$$= 2(\frac{1}{3}) + 3 = \frac{11}{3} \quad .$$

4.3.6 COVARIANZA Y COEFICIENTE DE CORRELACION

Daremos aquí la definición de unos números que miden como, estan relacionados los valores posibles de X con los valores posibles de Y; estos números dependen de la función de probabilidad conjunto de X e Y.

DEFINICION 4.3.9 Sean X e Y dos variables aleatorias con media μ_X y μ_Y, respectivamente. La *covarianza de* X e Y, denotado por "Cov(X,Y)", "σ_{XY}", se define por

$$\text{Cov}(X,Y) = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] =$$

$$= \sum_{(x,y)\in R_{X\times Y}} (x - \mu_X)(y - \mu_Y)p(x,y)$$

A diferencia de la varianza que siempre es no negativa, la covarianza - puede ser negativa, cero o positiva. Si los valores más pequeños de X son - asociados con valores grandes Y, y viceversa, entonces la covarianza es negativa. Es decir, si $X - \mu_X$ e $Y - \mu_y$ tienen signos opuestos. En cambio la covarianza será positiva cuando los valores grandes de X, se asocian con valores grandes de Y, y valores pequeños de X son asociados con valores pequeños de Y. Si la covarianza es cero, se dice que X e Y no están correlacionadas, aunque pueden ser dependientes o independientes. Una fórmula alternativa para calcular covarianzas, da el siguiente teorema.

TEOREMA 4.3.6 La *Covarianza* de dos variables aleatorias X e Y, con media - μ_X y μ_Y respectivamente, está dado por,

$$\text{Cov}(X,Y) = E(XY) - \mu_X\mu_Y$$

TEOREMA 4.3.7 Si X e Y son variables aleatorias independientes, entonces,

$$\text{Cov}(X,Y) = 0$$

TEOREMA 4.3.8 Si X e Y son variables aleatorias con función de probabilidad conjunta y varianzas finitas, entonces,

$$\text{Var}(aX \pm bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) \pm 2ab\text{Cov}(X,Y)$$

En particular, si X e Y son independientes, entonces Cov(X,Y) = 0, por lo tanto se cumple el teorema 3b.

DEFINICION 4.3.10 Sean X e Y variables aleatorias con desviación estandar σ_X y σ_Y respectivamente. *El coeficiente de correlación de* X e Y, denotado $\rho(X,Y)$, se define por,

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \ \sigma_Y} .$$

TEOREMA 4.3.9 PROPIEDADES DEL COEFICIENTE DE CORRELACION ρ

(a) $-1 \leq \rho \leq 1$

(b) $\rho(X,Y) = \rho(Y,X)$

(c) Si X e Y son independientes, entonces $\rho = 0$

(d) Si $\rho = \pm 1$, entonces entre X e Y existe una dependencia funcional lineal.

(e) Suponga que $Y = aX + b$, donde a y b son constantes.

Si $a > 0$, entonces $\rho(X,Y) = 1$; Si $a < 0$, entonces

$$\rho(X,Y) = -1$$

(f) Si a, b, c, d son constantes, $a > 0$ y $b > 0$, entonces;

$$\rho(aX + c, \ bY + d) = \rho(X,Y).$$

EJEMPLO 12 La función de probabilidad conjunta de (X,Y) está dado en la tabla. Calcular :

(a) Cov(X,Y) ;

(b) σ_X^2 ;

(c) σ_Y^2 ;

(d) $\rho(X,Y)$;

(e) σ_{X+Y}^2 ;

(f) $\rho(2X, 3Y + 4)$.

y \ x	0	1	$P[Y = y]$
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$
1	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
2	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$P[X = x]$	$\frac{5}{8}$	$\frac{3}{8}$	1

SOLUCION

(a) $E(X) = 0(\frac{5}{8}) + 1(\frac{3}{8}) = \frac{3}{8}$;

$E(Y) = 0(\frac{2}{8}) + 1(\frac{3}{8}) + 2(\frac{3}{8}) = \frac{9}{8}$;

$$E(XY) = 0(\frac{1}{8} + \frac{2}{8} + \frac{2}{8} + \frac{1}{8}) + 1(\frac{1}{8}) + 2(\frac{1}{8}) = \frac{3}{8} ;$$

Por lo tanto, $Cov(X,Y) = E(XY) - \mu_X \mu_Y = \frac{3}{8} - \frac{9}{8} \times \frac{3}{8} = - \frac{3}{64}$.

(b) $E(X^2) = 0(\frac{5}{8}) + 1(\frac{3}{8}) = \frac{3}{8}$;

Luego, $\sigma_X^2 = \frac{3}{8} - \frac{9}{64} = \frac{15}{64}$.

(c) $E(Y^2) = 0(\frac{2}{8}) + 1(\frac{3}{8}) + 4(\frac{3}{8}) = \frac{15}{8}$;

Luego, $\sigma_Y^2 = E(Y^2) - (\mu_Y)^2 = \frac{15}{8} - (\frac{9}{8})^2 = \frac{39}{64}$.

(d) Se tiene, $\sigma_X = \sqrt{\sigma_X^2} = \sqrt{15/64} = \frac{\sqrt{15}}{8}$

$$\sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{39/64} = \frac{\sqrt{39}}{8}$$

Por lo tanto, $\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{-3/64}{\frac{\sqrt{15 \times 39}}{64}} = - \frac{1}{\sqrt{65}}$.

(e) $\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2Cov(X,Y)$

$$= \frac{15}{64} + \frac{39}{64} + 2(- \frac{3}{64}) = \frac{3}{4} .$$

(f) por el teorema 9f, $\rho(2X, 3Y + 4) = \rho(X,Y) = - \frac{1}{\sqrt{65}}$.

PROBLEMAS 4.3

1. Una urna contiene 3 bolas numeradas 1,2,3 respectivamente. De la urna se extraen 2 bolas una a una sin reemplazo, y sea **X** el número de la primera bola extraída, **Y** el número de la segunda bola. Hallar la distribución de probabilidad conjunta de la variable aleatoria (**X**,**Y**) y su gráfica.

2. Suponga que tres objetos no diferenciables se distribuyen al azar en tres celdas numeradas. Sea **X** el número de celdas vacías e **Y** el número de objetos colocados en la primera celda. Construya la tabla de distribución de probabilidad conjunta de **X** e **Y**. ¿Son independientes dichas variables aleatorias?

3. Se elige uno de los números enteros: 1,2,3,4,5. Después de eliminar to-

dos los enteros (si los hay) menores que el elegido, se elige uno de los restantes (por ejemplo, si el primer número es el 3, la segunda elección se hace de entre los números 3,4,5). Sean X e Y los números obtenidos en la primera y segunda elección respectivamente.

(a) Hallar la distribución de probabilidad conjunta de X e Y.
(b) Determine las distribuciones de probabilidad marginal de X e Y.
(c) Determine la función de probabilidad condicional de Y, dado $X = 3$.
(d) Determine la función de probabilidad condicional de X, dado $Y = 3$.
(e) Calcular $P[X + Y > 7]$ y $P[Y - X > 0]$

4. La función de probabilidad conjunta de X e Y está dado por:

$$p(x,y) = \frac{x^2 + y^2}{32}, \qquad x = 0,1,2,3; \quad y = 0,1$$

(a) Hallar las funciones de probabilidad marginal de X e Y.
(b) Hallar la función de probabilidad condicional de X, dado $Y = y$.
(c) Hallar la función de probabilidad condicional de Y, dado $X = x$.

5. Se extraen al azar 2 naipes sin reemplazo de una baraja de 52 naipes. Sea X el número de ases que aparece e Y el número de espadas que aparecen. Obtener $p(x,y)$ y Calcular $P[X > Y]$.

6. En una urna hay 3 bolas negras y 7 blancas. Se seleccionan al azar 2 bolas sin reemplazo. Sea X el número de bolas negras sacadas, e Y el número de blancas. Obtener $p(x,y)$ y calcular $P[X \leq Y]$.

7. El espacio muestral Ω consiste de tres puntos ω_1, ω_2, ω_3. Las probabilidades asignadas a estos puntos son $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{2}$, respectivamente. Las variables aleatorias X e Y se define como :

$$X(\omega_1) = 0, \quad X(\omega_2) = 0, \quad X(\omega_3) = 1$$

$$Y(\omega_1) = 1, \quad Y(\omega_2) = 2, \quad Y(\omega_3) = 3$$

(a) Hallar la distribución de probabilidad conjunta de X e Y.
(b) Encuentre, la distribución de probabilidad marginal de Y.
(c) ¿Son independientes X e Y?
(d) Calcular la distribución de probabilidad condicional de X, dado $Y = 1$.
(e) Determine $E(Y)$.

8. Sean X e Y dos variables aleatorias que representan el número de bicicletas producidas por las líneas ensambladoras A, B respectivamente en un

día. La distribución de probabilidad conjunta $p(x,y)$ de las variables aleatorias X e Y, está dado en la siguiente tabla.

$x \backslash y$	0	1	2	3	4	5
0	0.00	0.01	0.03	0.06	0.07	0.08
1	0.01	0.02	0.04	0.07	0.07	0.09
2	0.01	0.02	0.04	0.06	0.06	0.06
3	0.00	0.01	0.05	0.05	0.05	0.04

(a) ¿Cuál es la probabilidad de que la línea B produce tres artículos, si se sabe que la línea A ha producido 1?

(b) ¿Cuál es la probabilidad de que la línea B produce más bicicletas que la línea A?

(c) ¿Son las variables aleatorias X e Y independientes?

(d) Hallar Cov(X,Y) y $\rho(X,Y)$.

9. Una moneda se lanza cuatro veces. Sea X, el número de caras que aparece y sea Y, tal que toma el valor i, si la primera cara aparece en el i-ésimo lanzamiento y es 0, si no aparece cara. Calcular:

(a) La distribución de probabilidad conjunta de X e Y.

(b) $P[X = 2,\ Y = 3]$; (c) $F(2,4)$; (d) $F(2,1)$; (e) $P[X = 3]$

(f) $P[X = 3|Y = 0]$; (g) $P[X = 3|Y = 1]$

(h) $E(2X + Y|Y = 1)$

10. La distribución de probabilidad conjunta de X e Y se define como :

$$p(x,y) = \frac{3x + 2y - 4}{54} \quad ; \quad x,y = 1,2,3.$$

(a) Hallar las distribuciones marginales de X e Y respectivamente.

(b) Calcular : $p(x|1)$; $p(x|2)$; $p(x|3)$

(c) Calcular : $p(y|1)$; $p(y|2)$; $p(y|3)$

(d) Verifique que X e Y son dependientes.

11. Dos firmas comerciales principales, A y B, controlan 50 y 30 por ciento del mercado respectivamente. Si se escoge al azar una muestra de dos compradores para una observación. ¿Cuál es la distribución de probabilidad conjunta del número de compradores que favorecen a cada firma de la mues-

tra?, Calcular, $E(X)$, $E(Y)$, $E(X + Y)$ y $E(XY)$.

12. Un puerto tiene capacidad de acomodar 4 naves de cierto tipo durante la noche. Las tarifas del puerto producen una utilidad de \$ 1,000 por nave atracada. Si X es la variable aleatoria que representa el número de naves buscando atracadero por noche y suponiendo que $P[X = k] = \frac{1}{6}$ para $k = 1,2,3,4,5$. Un segundo puerto está disponible para manejar el exceso de naves, si existen. Sea Y el número de barcos solicitando atracadero en el segundo puerto (lo cual sólo ocurrirá, si el primer puerto, está lleno) Calcular:
    (a) La distribución de probabilidad conjunta de X e Y.
    (b) La distribución de probabilidad marginal de Y.
    (c) $E(Y)$
    (d) La distribución de probabilidad de Y, dado $X = 4$.
    (e) ¿Son las variables aleatorias X e Y independientes?
    (f) σ_X^2 , σ_Y^2 ; (g) $Cov(X,Y)$; (h) $\rho(X,Y)$

13. Considere las variables aleatorias independientes X e Y, las cuales sólo pueden tomar los valores $- 1,0,1$. Suponga que $P[X = - 1] = P[X = 1] = \frac{1}{4}$ y $P[Y = - 1] = P[Y = 0] = \frac{1}{3}$.
    (a) Calcular $E(X)$, $E(Y)$;
    (b) Si $Z = 4X + 3Y$, calcular $E(Z)$;
    (c) Determine la distribución de probabilidad conjunta de X e Y .

14. La distribución de probabilidad conjunta de la variable aleatoria (X,Y) está dado por la siguiente tabla.
    Calcular :
    (a) $E(X)$;
    (b) $E(Y)$;
    (c) $E(3X + 4Y)$;
    (d) $E(Y^2)$; (e) σ_Y^2 ;
    (f) $E(XY)$;
    (g) $P[X = 1|Y = 1]$;
    (h) $E(X|Y = 1)$;
    (i) $E(2X + 1|Y = 1)$; (j) $E(2X + Y|Y = 1)$; (k) $E(XY|Y = 1)$

y \ x	0	1
0	0.1	0.05
1	0.2	0.1
2	0.3	0.25

15. La distribución de probabilidad conjunta de (X,Y) está dado por la tabla adjunta. Calcular :

(a) $E(X + Y)$;

(b) $E(XY)$;

(c) $E(2X + 3|Y = 2)$;

(d) $E(XY|Y = 2)$;

(e) $E(X + Y|Y = 2)$;

$y \backslash x$	0	1
1	0.1	0.4
2	0.2	0.3

(f) $Cov(X,Y)$; (g) $Cov(X, - Y)$; (h) $\rho(X,Y)$; (i) $\rho(X, - Y)$

16. Un producto se clasifica de acuerdo al número de defectos que contiene y a la fábrica que lo produce. Sean X e Y las variables aleatorias que representan el número de defectos por unidad (con posible valores 0,1,2 ó 3) y el número de fábricas (con posibles valores 1 ó 2), respectivamente. Los valores de la tabla representa la distribución de probabilidad conjunta - de X e Y. Determinar:

(a) La distribución de probabilidad marginal de X e Y ;

(b) $E(X + Y)$

(c) σ_X^2 ; σ_Y^2 ;

(d) $E(2Y + 3|X = 2)$

$y \backslash x$	0	1	2	3
1	1/8	1/16	3/16	1/8
2	1/16	1/16	1/8	1/4

17. Sean X e Y dos variables aleatorias que representa las producciones diarias de dos líneas ensambladoras A y B de automoviles respectivamente. La distribución conjunto $p(x,y)$ de las variables aleatorias esta dado en el cuadro adjunto. Calcular :

(a) Las distribuciones marginales $p_X(x)$ y $p_Y(x)$

(b) $P[X = 3|Y = 2]$; $P[Y = 3|X = 2]$

(c) $E(X)$ y $E(Y)$

$y \backslash x$	1	2	3
1	0	1/6	1/12
2	1/5	1/9	0
3	2/15	1/4	1/18

(d) $E(X|Y = 2)$; $E(Y|X = 3)$; $E(2X + 3|Y = 1)$; $E(2X + Y|Y = 2)$

(e) $P[X < Y]$; $Var(X)$; $Var(Y)$; $Cov(X,Y)$; $Cov(3X + 2, 2Y + 4)$

4.4 DISTRIBUCIONES BIDIMENSIONALES CONTINUAS

Extenderemos aquí la idea de la función de densidad de probabilidad de una variable aleatoria contínua a la de una variable aleatoria bidimensional contínua. Como en el caso de una variable aleatoria, las definiciones son las mismas que en el caso discreto, excepto que las integrales reemplazan a la sumatoria. Para la mayor parte, aceptaremos esta sustitución, y por lo tanto esta sección consiste principalmente de un resumen, ejemplos y problemas.

DEFINICION 4.4.1 Si (X,Y) es una variable bidimensional contínua. *La función de densidad de probabilidad conjunta*, es una función integrable $f(x,y)$ que cumple las siguientes propiedades:

(a) $f(x,y) > 0 \quad , \quad \forall \ (x,y)$

(b) $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1 \quad , \quad \left(\text{o} \iint_{R_{X\times Y}} f(x,y)dxdy = 1\right)$

La probabilidad de un evento A definido en un plano euclideano, está dado por,

$$P[(X,Y) \in A] = \iint_A f(x,y)dxdy$$

Y por lo tanto $P[(X,Y) \in A]$ es el volumen del sólido sobre la región **A** en el plano **XY** y acotado por la superficie $f(x,y)$.

Las Funciones de Densidad Marginal son :

(a) $f_X(x) = \displaystyle\int_{-\infty}^{\infty} f(x,y)dy$, función densidad marginal para X.

(b) $f_Y(y) = \displaystyle\int_{-\infty}^{\infty} f(x,y)dx$, función de densidad marginal para **Y**.

El valor esperado de **X** (media de **X**), el valor esperado de **Y** (o media **Y**) están dados por :

$$E(X) = \mu_X = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)dydx$$

$$E(Y) = \mu_Y = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)dxdy$$

En general, el valor esperado de una función $H(X,Y)$ se define por,

$$E[H(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y) f(x,y)dxdy$$

Las propiedades son las mismas que en el caso discreto. Así,

$$E[H(X,Y) \pm G(X,Y)] = E[H(X,Y)] \pm E[G(X,Y)]$$

La varianza de X e Y están definidos como sigue :

$$Var(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)^2 f(x,y)dydx$$

$$Var(Y) = \sigma_Y^2 = \int_{-\infty}^{\infty} (y - \mu_Y)^2 f_Y(y)dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y - \mu_Y)^2 f(x,y)dxdy$$

Las propiedades también son las mismas que el caso discreto. Así,

$$\sigma_X^2 = E(X^2) - (\mu_X)^2$$

EJEMPLO 1 Sea (X,Y) una variable bidimensional, con función de densidad conjunta,

$$f(x) = \begin{cases} \frac{1}{500} ; & 0 \le x \le 2.5 \;,\; 0 \le y \le 200 \\ 0 ; & \text{en otros casos} \end{cases}$$

(a) Calcular $P[1.0 \le X \le 2.0,\; 100 \le Y \le 200]$

(b) Hallar la función de densidad marginal de X e Y ;

(c) Calcular $E(X)$;

(d) σ_X^2 .

SOLUCION

(a) Esta probabilidad se calcula integrando la función $f(x,y)$ sobre la región $1.0 \le x \le 2.0$ y $100 \le y \le 200$. Es decir,

$$P[1.0 \le X \le 2.0,\; 100 \le Y \le 200] = \int_{100}^{200}\int_{1.0}^{2.0} \frac{1}{500}\, dxdy = \frac{1}{5} .$$

(b) $$f_X(x) = \int_0^{200} \frac{1}{500}\,dy = \frac{2}{5} \quad , \quad 0 \le x < 2.5$$

Por lo tanto,

$$f_X(x) = \begin{cases} \frac{2}{5} & , \quad 0 \le x \le 2.5 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

$$f_Y(y) = \int_0^{2.5} \frac{1}{500}\,dx = \frac{1}{200} \quad , \quad 0 \le y \le 200$$

Esto es,

$$f_Y(y) = \begin{cases} \frac{1}{200} & , \quad 0 \le y \le 200 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

(c) $$E(X) = \mu_X = \int_0^{2.5} x \cdot \frac{2}{5}\,dx = \frac{1}{5}(2.5)^2 = 1.25 \,.$$

(d) $$\sigma_X^2 = \int_0^{2.5} x^2 \cdot \frac{2}{5}\,dx - (1.25)^2 = \frac{2}{3 \times 5}\,(0.5)^3 - (1.25)^2 = 1.56 \cdot$$

Si (X,Y) es una variable bidimensional contínua, ***la densidad condicional de*** X, *dado* $Y = y$, es,

$$f_{X|y}(x|y) = \frac{f(x,y)}{f_Y(y)} \quad , \quad \text{siempre que} \quad f_X(x) > 0$$

Similarmente, $$f_{Y|x}(y|x) = \frac{f(x,y)}{f_X(x)} \quad ; \quad f_Y(y) > 0$$

La ***Esperanza Condicional*** para la variable aleatoria (X,Y) contínua está dado por:

$$\mu_{X|y} = E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|y}(x|y)dx \quad , \quad y$$

$$\mu_{Y|x} = E(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|x}(y|x)dy$$

EJEMPLO 2 Sea (X,Y) una variable bidimensional cuya función de densidad está dado por,

$$f(x,y) = \begin{cases} x^2 + \frac{2y}{3} & , \quad 0 \leq x \leq 1 \,, \quad 0 \leq y \leq 2 \\ 0 & , \quad \text{en otro caso} \end{cases}$$

Hallar :

(a) $f_{X|y}(x|y)$; (b) $f_{Y|x}(y|x)$; (c) $E(Y|X = x)$

SOLUCION Las respectivas densidades marginales son,

$$f_X(x) = \int_0^2 (x^2 + \frac{xy}{3})dy = 2x^2 + \frac{2x}{3} \quad , \quad 0 \leq x \leq 1$$

Es decir, $f_X(x) = \begin{cases} 2x^2 + \frac{2x}{3} & , \quad 0 \leq x \leq 1 \\ 0 & , \quad \text{de otra forma} \end{cases}$

$$f_y(y) = \int_0^1 (x^2 + \frac{xy}{3})dx = \frac{1}{3} + \frac{y}{6} \quad , \quad 0 \leq y \leq 2$$

Luego , $f_Y(y) = \begin{cases} \frac{y}{6} + \frac{1}{3} & , \quad 0 \leq y \leq 2 \\ 0 & , \quad \text{en otro caso} \end{cases}$

Por lo tanto las respectivas densidades condicionales son:

(a) $$f_{Y|x}(y|x) = \frac{x^2 + \frac{xy}{3}}{2x^2 + \frac{2x}{3}} = \frac{1}{2}\left[\frac{x + \frac{y}{3}}{x + \frac{1}{3}}\right] \quad , \quad 0 \leq y \leq 2 \,, \quad 0 < x \leq 1$$

$$= 0 \quad , \quad \text{en otro caso.}$$

(b) $$f_{X|y}(x|y) = \frac{x^2 + \frac{xy}{3}}{\frac{1}{3} + \frac{y}{6}} = \frac{x(3x + y)}{1 + \frac{y}{2}} \quad , \quad 0 < x \leq 1 \,, \quad 0 \leq y \leq$$

$$= 0 \quad , \quad \text{en otro caso}$$

(c) $$\mu_{Y|x} = E(Y|X = x) = \int_0^2 \frac{1}{2}\,\frac{x + y/3}{x + 1/3}\,dy$$

$$= \frac{9x + 4}{9x + 3} \,.$$

PROBLEMAS 4.4

1. Hallar el valor de k, de tal manera que la función :

$$f(x,y) = \begin{cases} k\,\dfrac{x}{y} & , \quad 0 < x < 1, \quad 1 < y < 2 \\ 0 & , \quad \text{en otros casos.} \end{cases}$$

Sea una función de densidad de una variable aleatoria (X,Y).

2. Sea (X,Y) una variable aleatoria contínua bidimensional con densidad.

$$f(x,y) = \begin{cases} \dfrac{1}{x} & , \quad 0 < y < x, \quad 0 < x < 1 \\ 0 & , \quad \text{en otros casos.} \end{cases}$$

Hallar las densidades marginales para X, y para Y.

3. Sea X e Y los puntajes sobre un test de inteligencia y un test de preferencias ocupacionales, respectivamente. La función de densidad de probabilidad, de la variable aleatoria (X,Y) está dado por :

$$f(x,y) = \begin{cases} \dfrac{k}{1000} & , \quad 0 \leq x \leq 100, \quad 0 \leq y \leq 10 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar el valor de K.
(b) Hallar la función de densidad marginal de X e Y.
(c) Hallar una expresión de la función de distribución acumulativa $F(x,y)$.

4. Consideremos, una situación en la cual se miden la tensión superficial y la acidez de un producto químico. Estas variables son codificadas tal que la tensión superficial se mide en una escala de $0 \leq x \leq 2$ y la acidez se mide en una escala de $2 \leq y \leq 4$. La función de densidad de (X,Y) es,

$$f(x,y) = \begin{cases} k(6 - x - y), & 0 \leq x \leq 2, \quad 2 \leq y \leq 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar el valor de k.
(b) Calcular la probabilidad de que $X < 1$, $X < 3$
(c) Calcular la probabilidad de que $X + Y \leq 4$
(d) Hallar la probabilidad de que $X < 1.5$
(e) Hallar la densidades marginales de X e Y.

5. En el examen de admisión a cierta universidad se califica en base a 100 puntos. Sea X la calificación que obtiene uno de los estudiantes (que continua sus estudios hasta graduarse) y sea Y su razón del punto de calidad al graduarse (4 puntos = A). La función de densidad conjunta de X e Y es,

$$f(x,y) = \begin{cases} \dfrac{1}{50} & , \quad 2 < y < 4 \ , \quad 25(y-1) < x < 25y \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar μ_X, μ_Y, $E(XY)$

(b) σ_X^2 , σ_Y^2 , σ_{XY} y ρ_{XY}

6. Sean X e Y variables aleatorias contínuas, cuya función de densidad conjunta es,

$$f(x,y) = \begin{cases} 4 & , \quad 0 < x < 1 \ , \quad 0 < y < \dfrac{1}{4} \\ 0 & , \quad \text{en otros casos} \end{cases}$$

¿Son independientes las variables aleatorias X e Y ?

7. La función de densidad de conjunta de la variable aleatoria bidimensional (X,Y) es,

$$f(x,y) = \begin{cases} \dfrac{1}{8}(6 - x - y) \ , & 0 \leq x \leq 2 \ , \quad 2 \leq y \leq 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Hallar las densidades condicionales $f_{X|y}(x|y)$, y $f_{Y|x}(y|x)$.

8. Se ensayan dos tubos de vacío hasta que fallan. Sea X el tiempo de falla menor y sea Y el tiempo de falla máxima. La función de densidad conjunta está dada por,

$$f(x,y) = \begin{cases} \dfrac{1}{(100)^2}\, e^{-(2/100 + \mu/100)} & , \quad 0 < x < y \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(a) Hallar las densidades marginales de X e Y
(b) ¿Son X e Y independientes?
(c) Hallar la densidad condicional de Y, dado X = 60

9. Un elemento combustible de diámetro D va a ser colocado en un tubo de diámetro T. La densidad conjunta está dada por,

$$f(d,t) = \begin{cases} 100 \quad , \quad \text{para} \quad 1.95 \le d \le 2.05 \\ \qquad\qquad\qquad\qquad\;\; 2.00 \le t \le 2.10 \\ 0 \quad , \quad \text{en otros casos} \end{cases}$$

(a) Hallar las densidades marginales de D y T.

(b) Determine E(D), E(T),

(c) Calcular σ_D^2 , σ_T^2 .

10. Un equipo que consta de dos componentes electrónicos independientes permanecerá en uso tanto tiempo como uno de sus componentes esté aún en operación. El equipo tiene una garantía del fabricante que cubre la sustitución si el aparato se vuelve inutilizable en menos de un año desde la fecha de compra. La vida X (en años) del primer componente e Y la vida (en años) del segundo, son variables aleatorias con función de densidad de probabilidad conjunta

$$f(x,y) = \frac{1}{4} e^{-x/2} e^{-y/2} \quad , \quad x \ge 0 \quad , \quad y \ge 0$$

Usted compra uno de esos equipos, seleccionado aleatoriamente entre las existencias del fabricante. ¿Cuál es la probabilidad de que la garantía expire antes de que su equipo se vuelva inutilizable?

11. Suponga que X, el tiempo (en minutos) que una persona pasa con un agente eligiendo una póliza de seguro de vida e Y el tiempo (en minutos) que el agente emplea en hacer el papeleo una vez que el cliente se ha decidido son variables aleatorias con función de densidad de probabilidad conjunta

$$f(x,y) = \frac{1}{300} e^{-x/30} e^{-y/10} \quad ; \quad x \ge 0 \; , \quad y \ge 0$$

Usted se dispone a encontrarse con un agente de seguros para suscribir una póliza de seguro de vida. ¿Cuál es la probabilidad de que toda la transacción dure más de media hora?

12. Suponga que X, el tiempo (en minutos) que una persona pasa en la sala de espera de cierto médico e Y la duración (en minutos) de un examen físico completo, son variables aleatorias con función de densidad de probabilidad conjunta

$$f(x,y) = \frac{1}{500} e^{-x/10} e^{-y/50} \quad ; \quad x \ge 0 \; , \quad y \ge 0$$

ted llega a la consulta del médico para un examen físico 50 minutos antes de tener que partir a una reunión. ¿Cuál es la probabilidad de salir tarde para la reunión?

5

DISTRIBUCIONES DISCRETAS IMPORTANTES

En este capítulo presentaremos algunas distribuciones de probabilidad discreta, desarrollando en forma analítica ciertas suposiciones básicas de un fenómeno real. Estas distribuciones tienen aplicaciones en ingeniería administración etc. La distribución *geométrica, la binomial y la binimial negativa* se basan en una sucesión de *ensayos de Bernoulli*. Presentaremos también la *distribución Hiper geométrica y la distribución de Poisson*.

5.1 ENSAYOS Y DISTRIBUCION DE BERNOULLI

Hay muchos experimentos que tienen sólo dos resultados posibles, llamados *exito* "E" y *fracaso* "F". Luego, el espacio muestral para este tipo de ensayo es $\Omega = \{E,F\}$. Por ejemplo, al lanzar una moneda se obtendrá sólo dos resultados posibles C ó S. Llamaremos éxito al evento de nuestro interés. En el ejemplo, si estamos interesados en "cara", diremos que hemos obtenido un éxito cuando en el ensayo ocurre cara, en caso contrario, diremos que hemos obtenido un fracaso.

Un experimento con esta característica se ll[illegible] *ensayo de Bernoulli*.

Definimos ahora la variable aleatoria X de tal manera que $X(\omega)$ = número de exitos en un ensayo de Bernoulli.

$$R_X = \{0,1\}.$$

Es decir: $X(F) = 0$, "si el resultado del ensayo es una F" y $X(E) = 1$, "si el resultado del ensayo es una E".
La variable aleatoria X, así definida se llama, *variable aleatoria de Bernoulli*

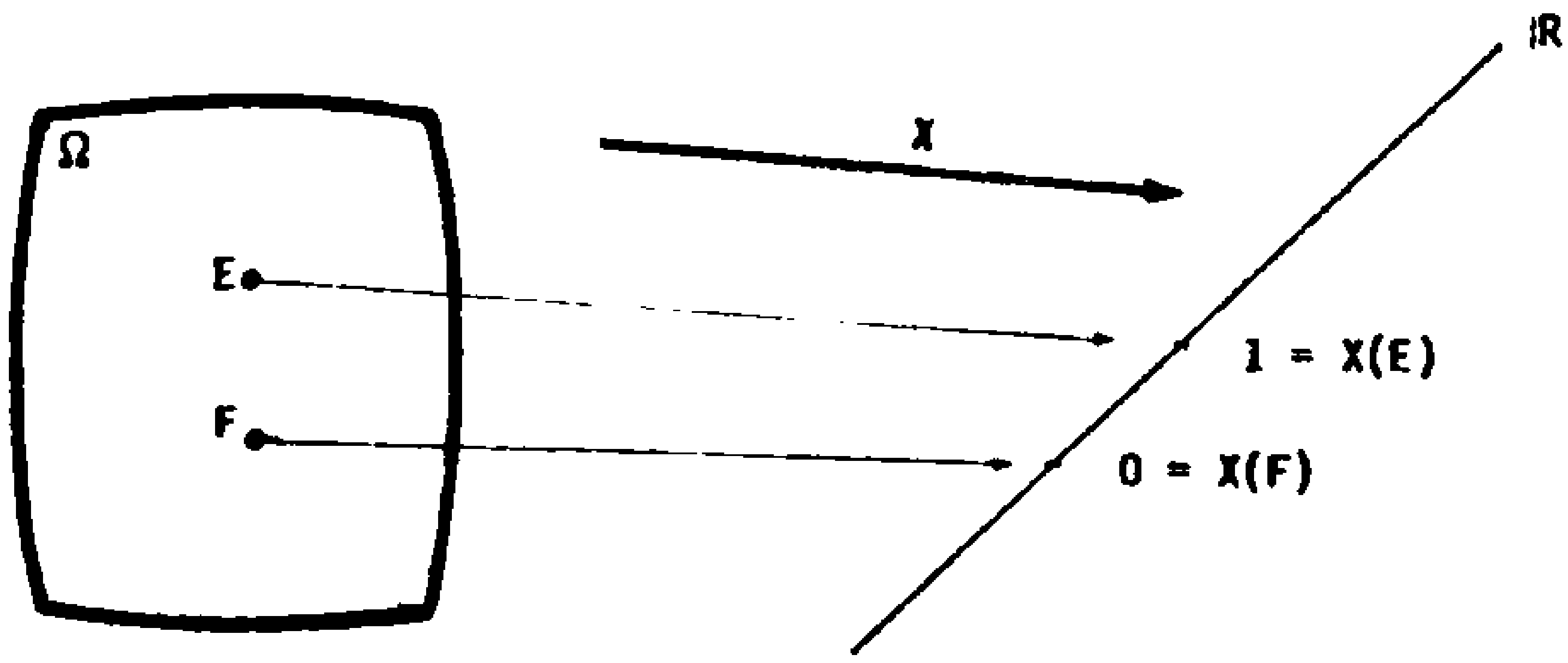


Fig. 5.1.1 Ensayo de Bernoulli.

denotemos por $p = P[E]$ y $q = 1 - p = P[F]$.

La distribución de probabilidad de la variable aleatoria X de Bernoulli, es llamada *distribución de Bernoulli* y su gráfica se muestra en la tabla 5.1.1 y la fig. 5.1.2 respectivamente.

x	0	1
$p(x)$	$1 - p = q$	p

Tabla 5.1.1

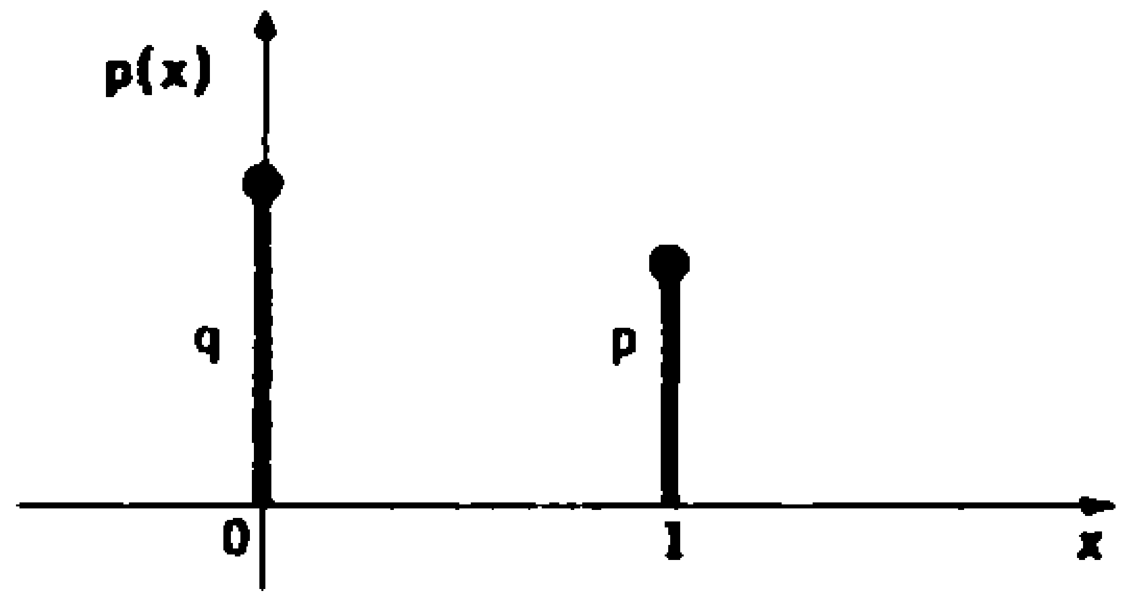


Fig. 5.1.2 Distribución de Bernoulli.

La distribución de Bernoulli se escribe también como una función así,

$$p(x) = P[X = x] = p^x(1 - p)^{1 - x} \quad , \quad x = 0 \text{ ó } 1 .$$

La media y la varianza de la variable aleatoria X, se calcula como sigue :

$$\mu = E(X) = 0.q + 1.p = p \quad .$$

$$\sigma_X^2 = \text{Var}(X) = E(X^2) - (E(X))^2$$

$$= 0^2.q + 1^2.p - p^2 = p - p^2 = p(1 - p)$$

$$= pq.$$

Hay muchos problemas en la cual el experimento consiste de n ensayos (o sub experimentos) de Bernoulli $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. Se dice que una secuencia de n ensayos iguales de Bernoulli, forman un *proceso de Bernoulli o experimento binomial* si se cumple:

(a) Cada ensayo tiene sólo dos resultados posibles E ó F.

(b) Los ensayos son independientes. Es decir, el resultado (exito o fracaso) de cualquier ensayo es independiente del resultado de cualquier otro ensayo.

(c) La probabilidad de éxito "p" permanece constante de ensayo a ensayo. Luego, la probabilidad de fracaso $q = 1 - p$ también es constante.

EJEMPLO 1 Suponga un experimento que consta de tres ensayos de Bernoulli y p la probabilidad de éxito en cada ensayo. X la variable aleatoria que representa el número de éxitos en los tres ensayos. Hallar la distribución de probabilidad de X.

SOLUCION Sea $X_i (i = 1,2,3)$ la variable aleatoria de Bernoulli que representa el número de éxitos en el ensayo i.

La variable aleatoria X se escribe, $X = \sum_{i=1}^{3} X_i$. El dominio es

Ω = {FFF, FFE, FEF, EFF, FEE, EFE, EEF, EEE} y

$R_X = \{0,1,2,3\}$

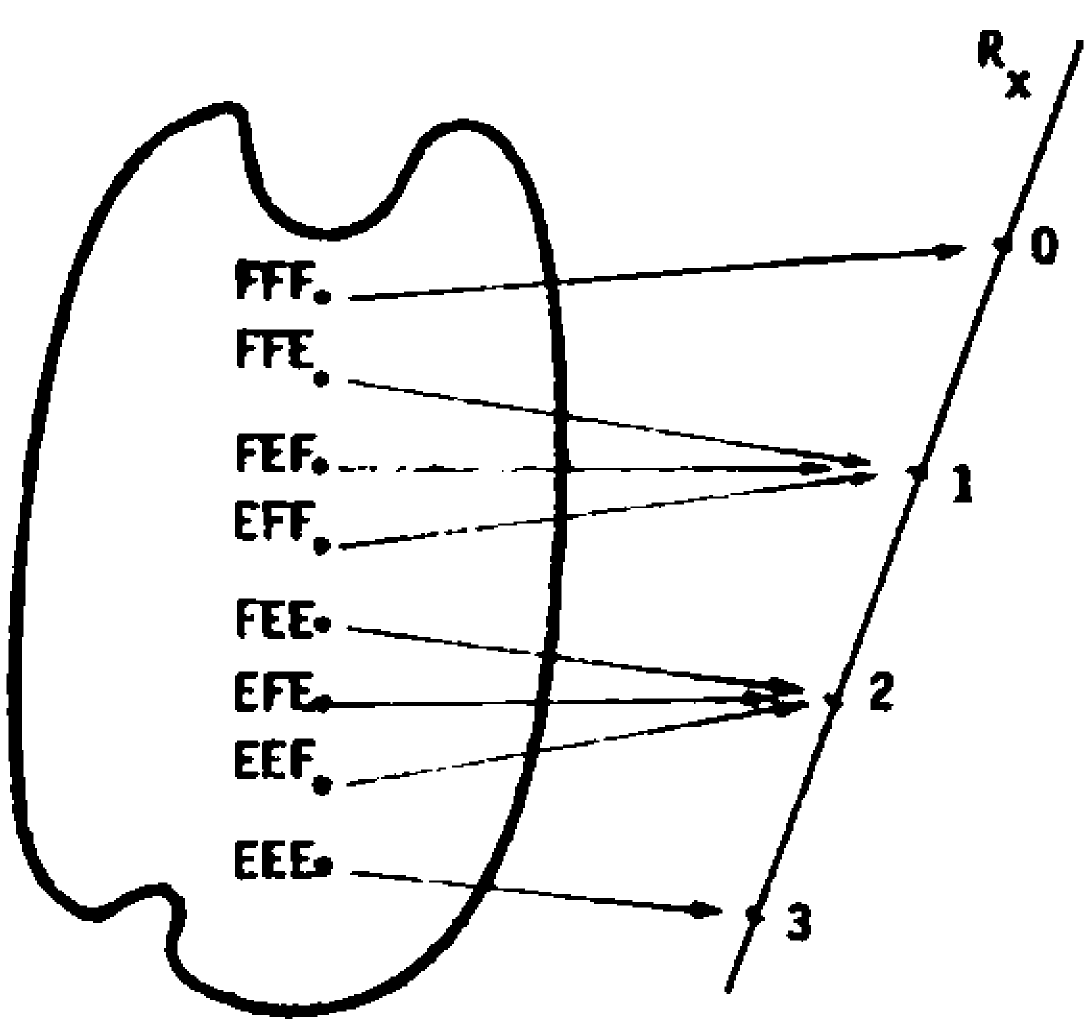
$$P[X = 0] = P[\{FFF\}] = (1 - p)(1 - p)(1 - p) = (1 - p)^3$$

$$P[X = 1] = P[\{FFE\}] + P[\{FEF\}] + P[\{EFF\}] = 3p(1 - p)^2$$

$$P[X = 2] = P[\{FEE\}] + P[\{EFE\}] + P[\{EEF\}] = 3p^2(1 - p)$$

$$P[X = 3] = P[\{EEE\}] = p^3$$

x	0	1	2	3
$p\ x$	$(1 - p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3



5.2 DISTRIBUCION BINOMIAL

A menudo estaremos interesados solamente en el ***número total de exitos*** "E" obtenidos en un proceso de n ensayos de Bernoulli, al margen del orden en que se presentan. El espacio muestral de los n ensayos, tiene 2^n elementos o sucesos los cuales son, una sucesión de n símbolos E y F.

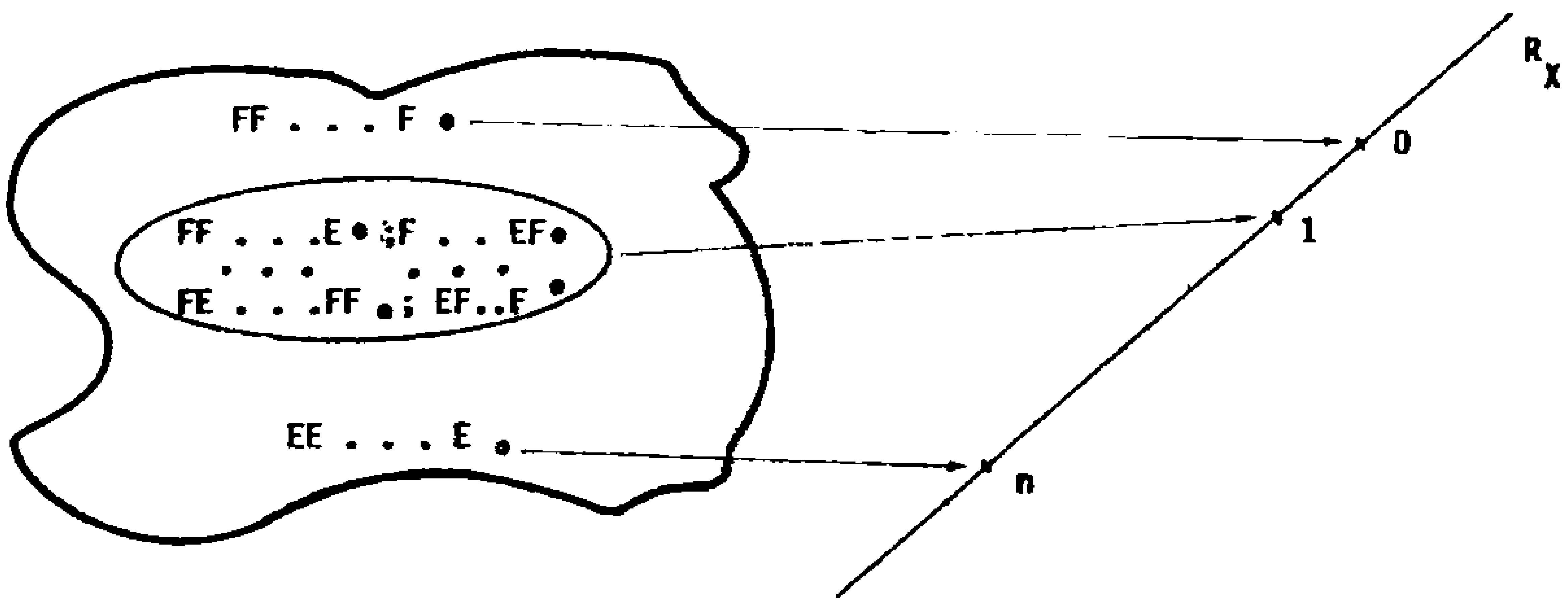
$$\Omega = \{\underbrace{EEFFEF \ldots EFE}_{n \text{ letras}}, \ldots\}$$

Definimos ahora una variable aleatoria **X** de la siguiente forma,

$X(\omega)$ = números de éxitos obtenidos en los n ensayos de Bernoulli.

$$R_X = \{0,1,2,3, \ldots, n\}$$

La variable aleatoria **X** así definida se llama una ***variable aleatoria binomial***.



La distribución de probabilidad de la variable aleatoria binomial **X**, se llama ***distribución binomial*** y se denota por $P[X = x|B; n,p]$ ó $b(x;n,p)$ que se lee : "la probabilidad de obtener exactamente **x** éxitos y $n - x$ fracasos".

Cálculo de la distribución de probabilidad de la variable aleatoria binomial **X** :

(a) La probabilidad que en los n ensayos no ocurre ningún exito es q^n. Es decir

$$P[X = 0] = P[\{\underbrace{F\ F \ldots F}_{n\text{-veces } F}\}] = q^n, \text{ por ser independientes}$$

(b) La probabilidad que en los n ensayos ocurre un éxito, es

$$P[X = 1] = P[\underbrace{\{\overbrace{FFF.\ .\ .F}^{n-1}\overbrace{E}^{1},\ldots\}}_{P_n^{n-1,1}}] = P_n^{n-1,1}\, q^{n-1}p = \binom{n}{1}q^{n-1}p$$

Ya que una E y $n - 1$ efes pueden ocurrir en los n ensayos de $P_n^{n-1} = \binom{n}{1}$ - formas diferentes.

(c) La probabilidad que en los n ensayos ocurre 2 éxitos es,

$$P[X = 2] = P[\underbrace{\{\overbrace{FF....F}^{n-2}\overbrace{EE}^{2},\ldots\}}_{P_n^{n-2,2}}] = \binom{n}{2}q^{n-2}p^2$$

ya que en los n ensayos los dos ees y $n - 2$ efes ocurren de $P_n^{n-2,2} = \binom{n}{2}$ formas diferentes.

(d) En general la probabilidad que en los n ensayos ocurre x éxitos es,

$$P[X = x] = P[\underbrace{\{\overbrace{FFF.\ .\ .F}^{n-x}\ \overbrace{EE...E}^{x},\ldots\}}_{P_n^{x,n-x}}] = \binom{n}{x}q^{n-x}p^x$$

ya que las x ees ocurren en los n ensayos de $P_n^{n-x,x} = \binom{n}{x}$ formas diferentes.

Por lo tanto, la *distribución binomial* de la variable aleatoria X es

$$\boxed{P[X = x|B;n,p] = P[X = x] = \binom{n}{x}p^x q^{n-x}, \quad x = 0,1,2,\ldots,n.}$$

Algunas veces, escribiremos simplemente

$$b(x;n,p) = \binom{n}{x}p^x q^{n-x}, \quad x = 0,1,2,\ldots,n.$$

Observe que :

(1) $\quad p(x) = P[X = x] > 0, \quad x = 0,1,2,\ldots,n$

(2) $$\sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} P[X = x] = \sum_{x=0}^{n}\binom{n}{x}p^x q^{n-x}$$

$$= (p + q)^n = 1^n = 1, \quad \text{ya que } p + q = 1.$$

La función de distribución acumulada está dada por,

$$F(x) = P[X \leq x] = P[X \leq x|B;n,p] = \begin{cases} 0 & , x < 0 \\ \sum_{k=0}^{[[x]]} \binom{n}{k} p^k q^{n-k} & , 0 \leq x < n \\ 1 & , x \geq n \end{cases}$$

CARACTERISTICAS DE LA DISTRIBUCION BINOMIAL

La media de la variable aleatoria binomial se puede calcular utilizando directamente la definición,

$$\mu = E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n(n-1)!}{(x-1)!(n-x)!} p^x q^{n-x}$$

$$= np \sum_{x=0}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

Cambiando el índice de la suma, $\ell = x - 1$, se tiene,

$$\mu = E(X) = np \sum_{\ell=0}^{n-1} \frac{(n-1)!}{\ell!(n-1-\ell)!} p^{\ell} q^{n-1-\ell}$$

$$= np(p + q)^{n-1}$$

y como $p + q = 1$, se tiene,

$$\boxed{\mu = E(X) = np}$$

Una demostración alternativa de la media usando el Teorema 3.2.4 de esperanza matemática es la siguiente:

Puesto que la variable aleatoria X, está definida como el número de exitos obtenidos en los n ensayos de Bernoulli, entonces,

$$X = X_1 + X_2 + X_3 + \ldots + X_n$$

donde X_i , $i = 1,2,\ldots,n$, es una variable aleatoria con una distribución de de probabilidad de Bernoulli. Es decir, X_i está definido por, 1, si el resultado del ensayo ε_i es un éxito y 0 si el resultado del experimento ε_i es un fracaso. Hemos visto que,

$$E(X_i) = p \quad , \quad i = 1,2,\ldots,n$$

Por lo tanto,

$$\mu = E(X) = E(X_1) + E(X_2) + \ldots + E(X_n)$$
$$= p + p + \ldots + p = np.$$

La varianza, se puede determinar aplicando, la propiedad

$$\sigma_X^2 = E(X^2) - (E(X))^2 .$$

$$E(X^2) = \sum_{x=0}^{n} x^2\binom{n}{x}p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x^2\binom{n}{x}p^x q^{n-x} - \sum_{x=0}^{n} x\binom{n}{x}p^x q^{n-x} + \sum_{x=0}^{n} x\binom{n}{x}p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\binom{n}{x}p^x q^{n-x} + \sum_{x=0}^{n} x\binom{n}{x}p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\binom{n}{x}p^x q^{n-x} + np$$

pero $x(x-1)\binom{n}{x} = x(x-1)\dfrac{n!}{x!(n-x)!} = \dfrac{n!}{(x-2)!(n-x)!}$

$$= n(n-1)\frac{(n-2)!}{(x-2)!(n-x)!} = n(n-1)\binom{n-2}{x-2}.$$

Por lo tanto,

$$E(X^2) = \sum_{x=2}^{n} n(n-1)\binom{n-2}{x-2}p^{x-2}q^{n-x}p^2 + np$$

$$= n(n-1)p^2\sum_{n=2}^{n}\binom{n-2}{x-2}p^{x-2}q^{n-x} + np$$

$$= n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^k q^{n-2-k} + np$$

$$= n(n-1)p^2(p+q)^{n-2} + np$$
$$= n(n-1)p^2 + np .$$

Luego, $\sigma_X^2 = \mathrm{Var}(X) = n(n-1)p^2 + np - n^2p^2$
$$= np^2(n-1-n) + np$$
$$= np(1-p) = npq$$

$$\boxed{\sigma_X^2 = \mathrm{Var}(X) = npq}$$

de donde, $\boxed{\sigma = \sqrt{npq}}$ la desviación estándar de **X**.

Para la varianza también una demostración alternativa para el lector que conoce el capítulo 4 es la siguiente:
Hemos visto que $\sigma_{X_i}^2 = pq$. Además sabemos que las variablea aleatorias X_i son independientes. Entonces,

$$\sigma_X^2 = \mathrm{Var}(X) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n)$$
$$= pq + pq + \ldots + pq = npq.$$

EJEMPLO 1 Se lanza un dado 10 veces. Calcular la probabilidad de obtener 4 veces seis.

SOLUCION 1 La variable aleatoria **X** está definida así,
$X(\omega)$ = número de veces que aparece el número 6 en los 10 lanzamientos del dado .

$$R_X = \{0,1,2,\ldots,10\}$$

2. El evento que nos interesa es "obtener un seis". Entonces definimos

E = "obtener un seis al lanzar un dado".
F = "obtener un número diferente de 6".

$$P[E] = \frac{1}{6} \quad y \quad P[F] = \frac{5}{6}$$

3. Se lanza el dado 10 veces $n = 10$. Cada lanzamiento puede considerarse con sólo dos resultados posibles E y F. Las probabilidades de E y F permanecen constante en cada ensayo. El resultado de cualquier ensayo es independiente de los otros. Por lo tanto, **X** es una variable aleatoria binomial y su distribución de probabilidad es,

$$P[X = x|B;\ 10,\ \frac{1}{6}\ ] = \binom{10}{x}(\frac{1}{6})^{x}(\frac{5}{6})^{10-x} \quad , \quad x = 0,1,2,\ldots,10\ .$$

4. Debemos calcular

$$P[X = 4\ ] = \binom{10}{4}(\frac{1}{6})^{4}(\frac{5}{6})^{6} = \frac{7.5^{7}}{6^{9}}\ .$$

EJEMPLO 2 En Huancayo en el mes de Octubre, la lluvia cae con un promedio - de uno cada cuatro días durante los que el cielo está nublado. Determine la distribución de probabilidad del número de días con lluvia entre los cuatro proximos días nublados, suponiendo se cumple independencia. Encuentre la media y la varianza del número de días lluviosos. Presente también la gráfica de la distribución de probabilidad.

SOLUCION 1 La variable aleatoria X está definida por

$X(\omega)$ = número de días con lluvia durante los 4 días nublados

$$R_X = \{0,1,2,3,4\}$$

2. El experimento es observar, si llueve o no cada uno de los cuatro días nublados. Obviamente cada experimento tiene sólo dos resultados posibles,

E = "día con lluvia" y F = "día sin lluvia".

$$p = P[E] = \frac{1}{4} \quad , \quad q = P[F] = \frac{3}{4}\ .$$

3. Suponiendo, que llueva o no un día determinado es independiente de que - llueva o no los otros días, se tiene que X es una variable aleatoria binomial cuya distribución de probabilidad, con parámetros $n = 4$ y $p = \frac{1}{4}$ es

$$P[X = x|4,\ \frac{1}{4}\ ] = \binom{4}{x}(\frac{1}{4})^{x}(\frac{3}{4})^{4-x} \quad , \quad x = 0,1,2,3,4.$$

4. La media y la varianza de X están dados por,

$$\mu = E(X) = np = 4(\frac{1}{4}) = 1.$$

$$\sigma^2 = npq = 4\cdot\frac{1}{4}\cdot\frac{3}{4} = \frac{3}{4}\ .$$

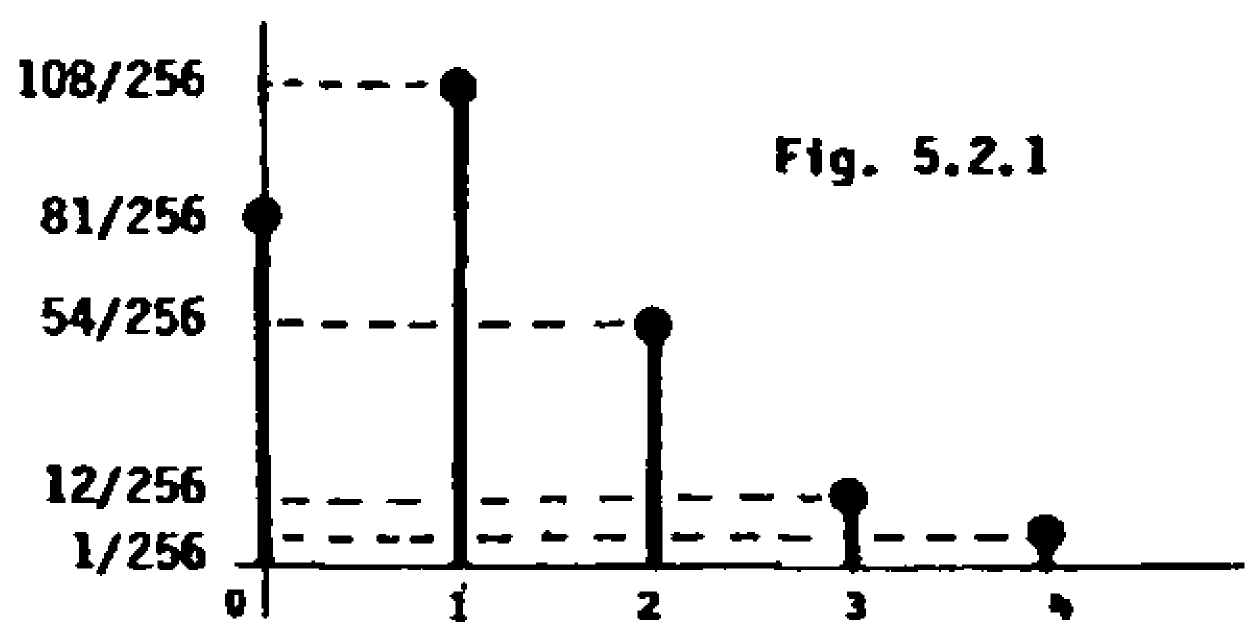


Fig. 5.2.1

EJEMPLO 3 El departamento de contabilidad de una firma comercial, tiene dos empleados a tiempo parcial: Juan y Juana, Juan trabaja los lunes, miercoles y viernes en tanto que Juana lo hace los martes y jueves. Juan archiva erróneamente uno de cada cuatro documentos, mientras que Juana lo hace uno de ca da cinco. Se selecciona un día de la semana al azar y en ese día se toma una muestra aleatoria de cuatro documentos de entre los documento archivados ese día.

(a) ¿Cuál es la probabilidad que la muestra contenga exactamente dos documentos mal archivados?

(b) Suponiendo que la muestra contenga dos documentos mal archivados. ¿Cuál es la probabilidad que ambos hayan sido archivados por Juan ?

SOLUCION 1 Elegir al azar un día de la semana. La semana se divide, en los días trabajados por Juan y en los trabajados por Juana. Se tiene así,

$$P[\text{Juan}] = 3/5 \quad , \quad P[\text{Juana}] = 2/5 \quad \text{(ver diagrama).}$$

2. De todos los documentos archivados el día elegido, se extrae una muestra de 4 documentos.

3. Teniendo en cuenta que los documentos pueden ser archivados por Juan ó - Juana (o, em el sentido de exclusión),

   Definimos :

   $X(\omega)$ = número de documentos mal archivados por Juan, en la muestra de 4 .

   $Y(\omega)$ = número de documentos mal archivados por Juana en la muestra de 4

$$R_X = \{0,1,2,3,4\} \quad , \quad R_Y = \{0,1,2,3,4\}$$

4. (*i*) La verificación de cada documento, son ensayos con sólo dos resultados posibles: mal archivados (E) o bien archivados (F).

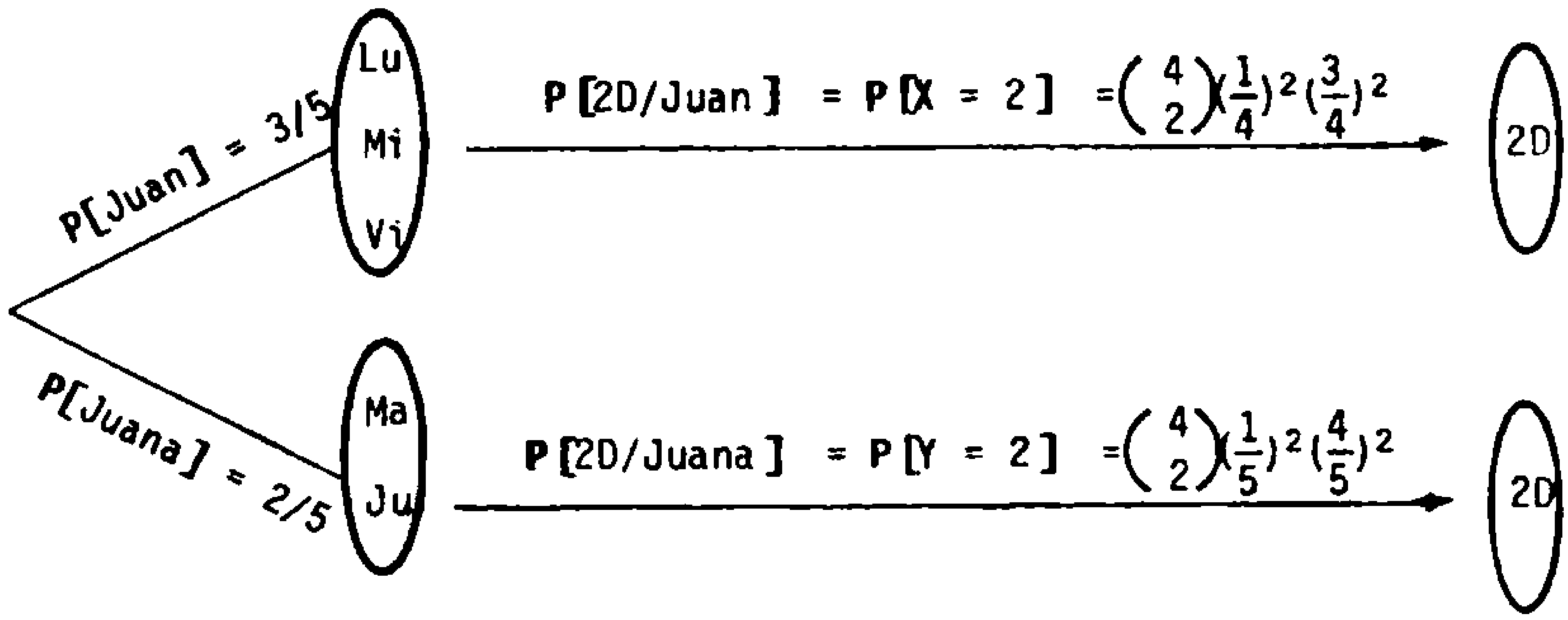
   (*ii*) Los resultados de cada ensayo son independientes debido a que la población es muy grande con respecto a la muestra (ver 5.2.1). Es decir

$$p = P[E] \quad \text{es constante en cada ensayo.}$$

5. Por lo tanto, las variables aleatorias **X** e **Y** tienen una distribución binomial con parámetros $n = 4$, $p = \frac{1}{4}$ y $n = 4$, $p = \frac{1}{5}$ respectivamente.

$$P[X = x] = \binom{4}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{4-x} \quad , \quad x = 0,1,2,3,4 \;;$$

$$P[Y = y] = \binom{4}{y}\left(\frac{1}{5}\right)^y\left(\frac{4}{5}\right)^{5-y} \quad , \quad y = 0,1,2,3,4 \;.$$



Del diagrama se obtiene

(a) $$P[2D] = \frac{3}{5} P[X = 2] + \frac{2}{5} P[Y = 2]$$

$$= \frac{3}{5}\ \frac{4!}{2!2!}\ (\tfrac{1}{4})^2(\tfrac{3}{4})^2 + \frac{2}{5}\ \frac{4!}{2!2!}\ (\tfrac{1}{5})^2(\tfrac{4}{5})^2$$

$$= \frac{81}{5 \times 128} + \frac{182}{5 \times 625} = \frac{75201}{400000} \simeq 0.188\ .$$

(b) Según el teorema de Bayes, se tiene ,

$$P[Juan|2D] = \frac{81/649}{75201/400000} = \frac{810000}{1203216} = 0.6732\ .$$

EJEMPLO 4 La figura 5.2.2 representa la gráfica de la función de distribución acumulada de una variable aleatoria **X**, que tiene una distribución binomial.

(a) Determinar n y p, μ y σ^2_X

(b) Calcular $P[X \geq 1]$

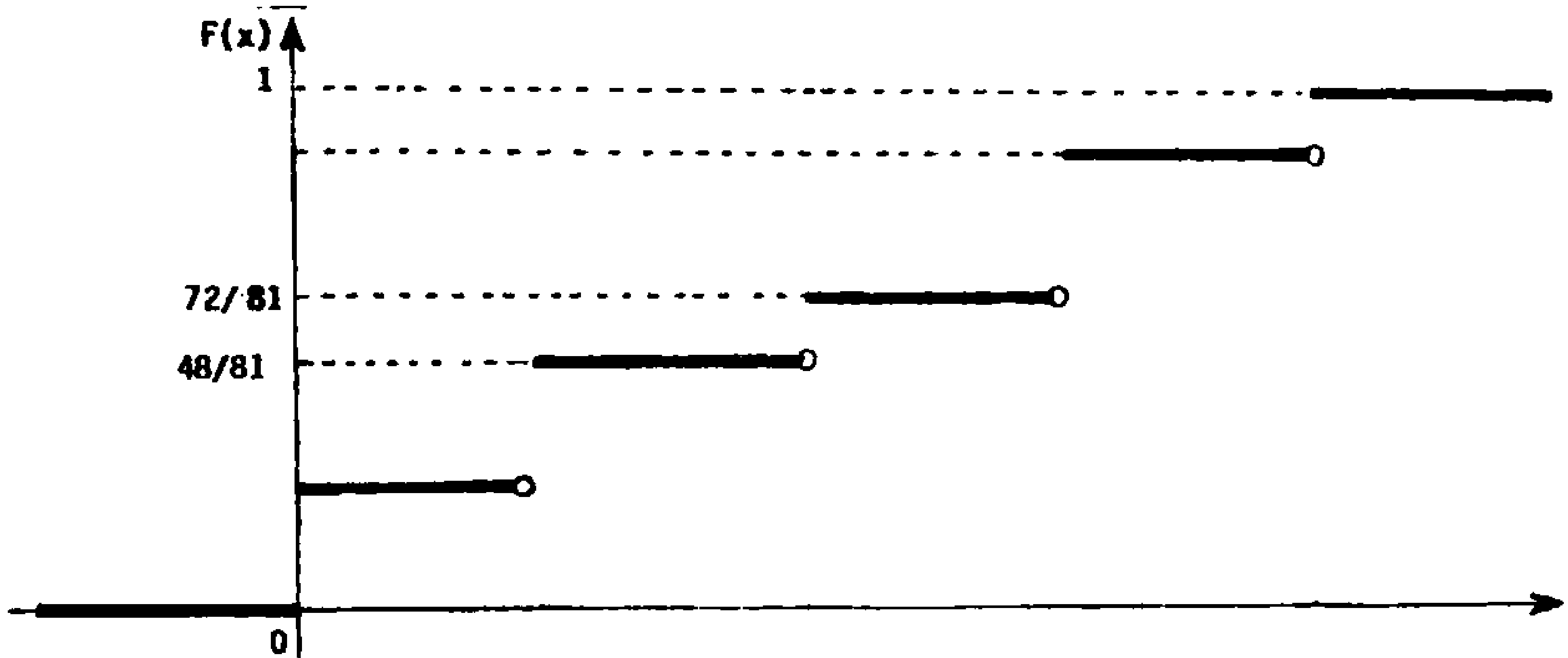


Fig. 5.2.2

SOLUCION

(a) De la gráfica obtenemos,

$$R_X = \{0,1,2,3,4\} \quad , \quad n = 4$$

Luego; $p(x) = P[X = x] = \binom{4}{x} p^x q^{4-x} \quad , \quad x = 0,1,2,3,4$

Además $F(1) = \frac{48}{81}$ y $F(2) = \frac{72}{81}$, por lo tanto, el salto en el punto $x = 2$, vale $\frac{24}{81}$, entonces, se tiene que

$$P[X = 2] = \binom{4}{2} p^2 q^2 = \frac{24}{81}$$

$$6p^2q^2 = \frac{24}{81} \quad , \quad \text{luego} \quad pq = \frac{2}{9}$$

ó $p(1 - p) = \frac{2}{9}$, ó $p^2 - p + \frac{2}{9} = 0$, de donde

$$p = \frac{1}{3} \quad , \quad y \quad q = \frac{2}{3} \quad ó \quad p = \frac{2}{3} \quad y \quad q = \frac{1}{3}$$

Entonces, la función de probabilidad de X en el primer caso es,

$$p(x) = P[X = x] = \binom{4}{x} (\tfrac{1}{3})^x (\tfrac{2}{3})^{4-x} \quad , \quad x = 0,1,2,3,4 \ .$$

La media μ es igual a,

$$\mu = np = 4(\tfrac{1}{3}) = \frac{4}{3} \ .$$

La varianza, $\sigma^2 = npq = 4 \times \frac{1}{3} \times \frac{2}{3} = \frac{8}{9}$.

(b) $P[X > 1] = 1 - P[X = 0] = 1 - \binom{4}{0} (\tfrac{1}{3})^0 (\tfrac{2}{3})^4 = \frac{65}{81}$.

EJEMPLO 5 Un estudiante se presenta a un examen de seleccion múltiple que contiene 8 preguntas cada una con tres respuestas opcionales. Si el estudiante está adivinando al responder cada pregunta y además se sabe que para aprobar el examen debe responder correctamente 6 ó más preguntas. ¿Cuál es la probabilidad de aprobar el examen?

SOLUCION 1 Definimos la variable aleatoria X tal que

$X(\omega)$ = número de respuestas correctas en las 8 preguntas.

$$R_X = \{0,1,2,3,4,5,6,7,8\}$$

2. Puesto que cada pregunta consta de una respuesta correcta y 2 respuestas no correctas, $P[E] = \frac{1}{3} = p$ y $P[F] = q = \frac{2}{3}$ (por estar adivinando)
3. Luego, la distribución de probabilidad de X es,

$$p(x) = P[X = x] = \binom{8}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{8-x} \quad , \quad x = 0,1,\ldots,8$$

4. Sea A, el evento: "aprobar el examen", entonces

$$P[A] = P[X \geq 6] = \sum_{x=6}^{8}\binom{8}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{8-x}$$

$$= \binom{8}{6}\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^2 + \binom{8}{7}\left(\frac{1}{3}\right)^7\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)^8 = \frac{43}{2187} = 0.02.$$

5.2.1 APLICACION DE LA DISTRIBUCION BINOMIAL EN UNA MUESTRA.

La extracción de una muestra de n elementos de una población puede considerarse como un experimento que consiste de n ensayos repetidos. Los n ensayos o selecciones serán independientes en los siguientes casos :

(a) Cuando los elementos de la muestra se extraen con o sin reemplazo de una población infinita. Obviamente el resultado de una extracción cualquiera es independiente de otra extracción y la proporción p de exitos ($p = P[\text{Exito}]$) permanece constante en cada extracción. Entonces, es aplicable la distribución binomial.

(b) Cuando los elementos de la muestra se extraen con reemplazamiento de una población finita. Suponga que la población tiene N elementos, k de los cuales son de cierta clase en las que estamos interesados. Definimos la variable aleatoria X tal que

$X(\omega)$ = número de elementos de la clase de nuestro interés en la muestra de tamaño n.

Las extracciones individuales son ensayos de Bernoulli, donde "elemento de clase de nuestro interés" corresponde a "exito" y el experimento de tomar una muestra de tamaño n con reeplazamiento consiste de n ensayos independientes de Bernoulli donde $p = P[\text{Exito}] = \frac{k}{N}$; es decir X tiene una distribución binomial

$$p(x) = \mathbf{P}[X = x|B;n, \frac{k}{N}] = \binom{n}{x}\left[\frac{k}{N}\right]^x \left[1 - \frac{k}{N}\right]^{n-x}, \quad x = 0,1,2,\ldots,n$$

Frecuentemente, en la práctica se extrae una muestra sin reemplazamiento de una población finita, situación en la cual no se puede aplicar la distribución binomial, ya que los ensayos claramente no son independientes. En este caso la distribución es una hipergeométrica (ver 5.5)

EJEMPLO 6 En una población grande de drosophila, el 25 % de las moscas tienen una mutación de alas. Se seleccionan aleatoriamente 300 moscas de la población para un examen de mutación de alas. Defina **X** como el número de moscas que tienen mutación en la muestra. Determinar el valor esperado, la varianza y la desviación estandar de **X**.

SOLUCION 1 $X(\omega)$ = número de moscas que tienen mutación de alas en la muestra de 300 moscas.

2. Puesto que la población es grande (infinita), no interesa como se ha extraído la muestra (con o sin reemplazamiento), se aplica la distribución binomial, con $n = 300$, $p = 0.25$, $q = 0.75$.

3. La función de probabilidad de **X** es

$$p(x) = \mathbf{P}[X = x] = \binom{300}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{300-x}, \quad x = 0,1,2,\ldots,300.$$

4. La media, varianza y desviación estándar de **X** son :

$$\mu = np = 300\left(\frac{1}{4}\right) = 75 .$$

$$\sigma^2 = npq = 75 \times \frac{3}{4} = \frac{225}{4} .$$

de donde, $$\sigma_X = \sqrt{\frac{225}{4}} = \frac{15}{2} .$$

EJEMPLO 7 Las máquinas A y B producen, en promedio, 5 % y 10 % de piezas defectuosas, respectivamente. Se extrae una muestra aleatoria de 4 piezas de la producción de cada una. ¿Cuál es la probabilidad que la muestra obtenida de la producción A tenga exactamente una pieza defectuosa y la muestra correspondiente a B contenga exactamente dos piezas defectuosas?

SOLUCION 1 Sean **X** e **Y** variables aleatorias definidas por

$X(\omega)$ = números de piezas defectuosas extraídas de la producción de **A** en la muestra de 4.

$Y(\omega)$ = número de piezas defectuosas extraídas de la producción de B en la muestra de 4

$R_X = \{0,1,2,3,4\}$, R_Y $\{0,1,2,3,4\}$

$n = 4$, $p_X = 0{,}05$, $p_Y = 0.10$.

2. X e Y así definidas son variables aleatorias binomiales y sus respectivas funciones de probabilidad son :

$$P[X = x] = \binom{4}{x}(0.05)^x(0.95)^{4-x} \quad , \quad x = 0,1,\ldots,4 .$$

$$P[Y = y] = \binom{4}{y}(0.10)^y(0.90)^{4-y} \quad , \quad y = 0,1,\ldots,4 .$$

3. Se pide calcular ,

$$P[X = 1,\ Y = 2] = P[X = 1]\,P[Y = 2] , \quad X \text{ e } Y \text{ independientes}$$

$$= \left[\frac{4!}{1!\,3!}(0.05)(0.95)^3\right]\left[\frac{4!}{2!\,2!}(0.10)^2(0.90)^2\right]$$

$$= (0.20)\left(\frac{95}{100}\right)^3 .6.\frac{1}{100}\cdot\frac{81}{100} = (0.171475)(0.0486)$$

$$= 0.0083 .$$

EJEMPLO 8 El daltonismo afecta al 1% de una población grande. Suponga que se escogen n personas aleatoriamente de ésta población. ¿Cuál es la probabilidad de que ninguna de las n personas sean daltonianos? ¿Qué tamaño debe tener n para que ésta probabilidad sea menor que el 10% ?

SOLUCION 1 Sea X la variable aleatoria definida por

$X(\omega)$ = número de personas daltonianos en las n personas seleccionadas.

$R_X = \{0,1,2,\ldots,n\}$, $p = 0.01$ y $n = ?$

2. Luego, X tiene una distribución binomial con $p = 0.01$, $q = 0.99$. O sea

$$P[X = x] = \binom{n}{x}(0.01)^x(0.99)^{n-x} \quad , \quad x = 0,1,2,\ldots, n$$

(a) $$P[X = 0] = \binom{n}{0}(0.01)^0(0.99)^n = (0.99)^n$$

(b) Se quiere determinar n tal que,

$$P[X = 0] = (0.99)^n < 0.10$$

Resolveremos la ecuación $(0.99)^n = 0.1$

Tomando logaritmos a ambos miembros, se tiene,

$$n \log(0.99) = \log(0.1)$$

$$n \log(\frac{99}{100}) = \log(\frac{1}{10})$$

de donde, $n = \dfrac{1}{2 - \log 99} = \dfrac{1}{2 - 1.9956}$

Luego, $n \simeq 227.$

EJEMPLO 9 Un dado cargado es lanzado 10 veces. Se sabe que la probabilidad de que aparezca 5 veces un número par es el doble de la probabilidad que aparezca 4 veces un número par. Si se lanza el dado una vez. ¿Cuál es la probabilidad que aparezca un número par?

SOLUCION 1 Sea X la variable aleatoria tal que

$X(\omega)$ = número de veces que aparece un número par en los 10 lanzamientos del dado cargado

$$R_X = \{0,1,2,3,4,5,6,7,8,9,10\}$$

$$n = 10 \quad , \quad p = ?$$

2. La función de probabilidad de **X** es,

$$p(x) = P[X = x] = \binom{10}{x} p^x (1 - p)^{10-x} \quad , \quad x = 0,1,\ldots,10$$

3. Además, del enunciado se tiene

$$P[X = 5] = 2P[X = 4]$$

$$\binom{10}{5} p^5 (1 - p)^5 = 2\binom{10}{4} p^4 (1 - p)^6$$

$$\frac{10!}{5!5!} p = \frac{2 \times 10!}{4!6!} (1 - p)$$

$$\frac{6}{5} p = 2(1 - p)$$

de donde $p = \dfrac{5}{8} \quad , \quad q = \dfrac{3}{8}$.

4. Por lo tanto, la probabilidad de obtener un número par en un lanzamiento de éste dado es $\frac{5}{8}$.

EJEMPLO 10 Una máquina produce artículos en los que hay una proporción p de defectuosos. El ingeniero a cargo de la producción acostumbra inspeccionar la máquina cada hora, mediante una muestra. Si la muestra no contiene artículos defectuosos, permite que la máquina siga trabajando. Admitiendo que $p = 0.10$, determinar el tamaño máximo de la muestra, de modo que la probabilidad que la máquina no sea detenida en una inspección determinada sea menor o igual que 0.01.

SOLUCION 1 Sea n el tamaño de la muestra, que queremos determinar.

2. $X(\omega)$ = número de artículos defectuosos en la muestra de tamaño n .

$$P[E] = p = 0.10, \qquad q = 0.90$$

3. La distribución de probabilidad de X es,

$$P[X = x] = \binom{n}{x}(0.10)^x(0.90)^{n-x} \qquad , \qquad x = 0,1,2,\ldots,n$$

4. La máquina no se detiene, si la variable aleatoria X toma el valor, 0, entonces,

$$P[X = 0] = \binom{n}{0}(0.10)^0(0.90)^n \leq 0.01$$

Resolveremos la ecuación,

$$(0.90)^n = 0.01$$

$$n \log(0.90) = \log(0.01)$$

$$n = \frac{-2}{-0.0458} = 43.6$$

Luego, $\quad n = 44.$

EJEMPLO 11 Fabricante de piezas de carro, envía en lotes de 20 a sus clientes. Suponer que cada pieza está defectuosa o no lo está, y que la probabilidad de que cualquiera de ellas está defectuosa es de 0.05.

(a) ¿Cuál es el número esperado de piezas defectuosas por lote?

(b) Si un cliente determinado del proveedor mencionado recibe 10 lotes. ¿Cuál es el número esperado de lotes que no tienen piezas defectuosas?.

SOLUCION 1 Definimos la variable aleatoria X, tal que
$X(\omega)$ = número de piezas defectuosas en un lote de 20.

$$R_X = \{0,1,2,\ldots,20\}$$

2. La variable aleatoria, X, tiene una distribución binomial con $n = 20$ -

$y \quad p = 0.05.$ Entonces,

(a) $\mu = E(X) = 20(0.05) = 1.$

(b) Sea Y la variable aleatoria tal que

$Y(\omega)$ = número de lotes que no tienen piezas defectuosas en 10 lotes.

$$R_Y = \{0,1,2,\ldots,10\}$$

La probabilidad que un lote no tiene piezas defectuosas se obtiene así,

$$P[X = 0] = \binom{20}{0}(0.05)^0(0.95)^{20} = 0.3585$$

La variable aleatoria Y tiene una distribución binomial con $n = 10$ y $p = 0.3585$. Entonces,

$$\mu = E(Y) = 10(0.3585) = 3.585.$$

aproximadamente hay 4 lotes que no contienen piezas defectuosas.

5.2.2 USO DE LA TABLA DE PROBABILIDAD BINOMIAL

Los valores de la función de distribución acumulada de la variable aleatoria X que tiene una distribución binomial, están dadas en tablas y el uso de éstas simplifica enormemente el cálculo de todo tipo de probabilidades binomiales.

Sea $X(\omega)$ = número de éxitos en los n ensayos de Bernoulli.

La probabilidad acumulada binomial dado en la tabla I. es $P[X \geq r]$, que designaremos por

$$P[X \geq r \mid B;\, n,p\,] \qquad \text{o} \qquad P[X \geq r \mid n,p\,]$$

En otras palabras, la tabla da, la probabilidad que variable aleatoria binomial toma valores mayores o iguales a r. Estas probabilidades se muestra en la figura 5.2.3. El diagrama muestra los posibles valores que toma la variable aleatoria, asociado con sus respectivas probabilidades y se ve también que cada probabilidad en la tabla de probabilidad acumulada es la suma de las probabilidades individuales.

$$P[X \geq r] = P[X = r] + P[X = r + 1] + \ldots + P[X = n]$$

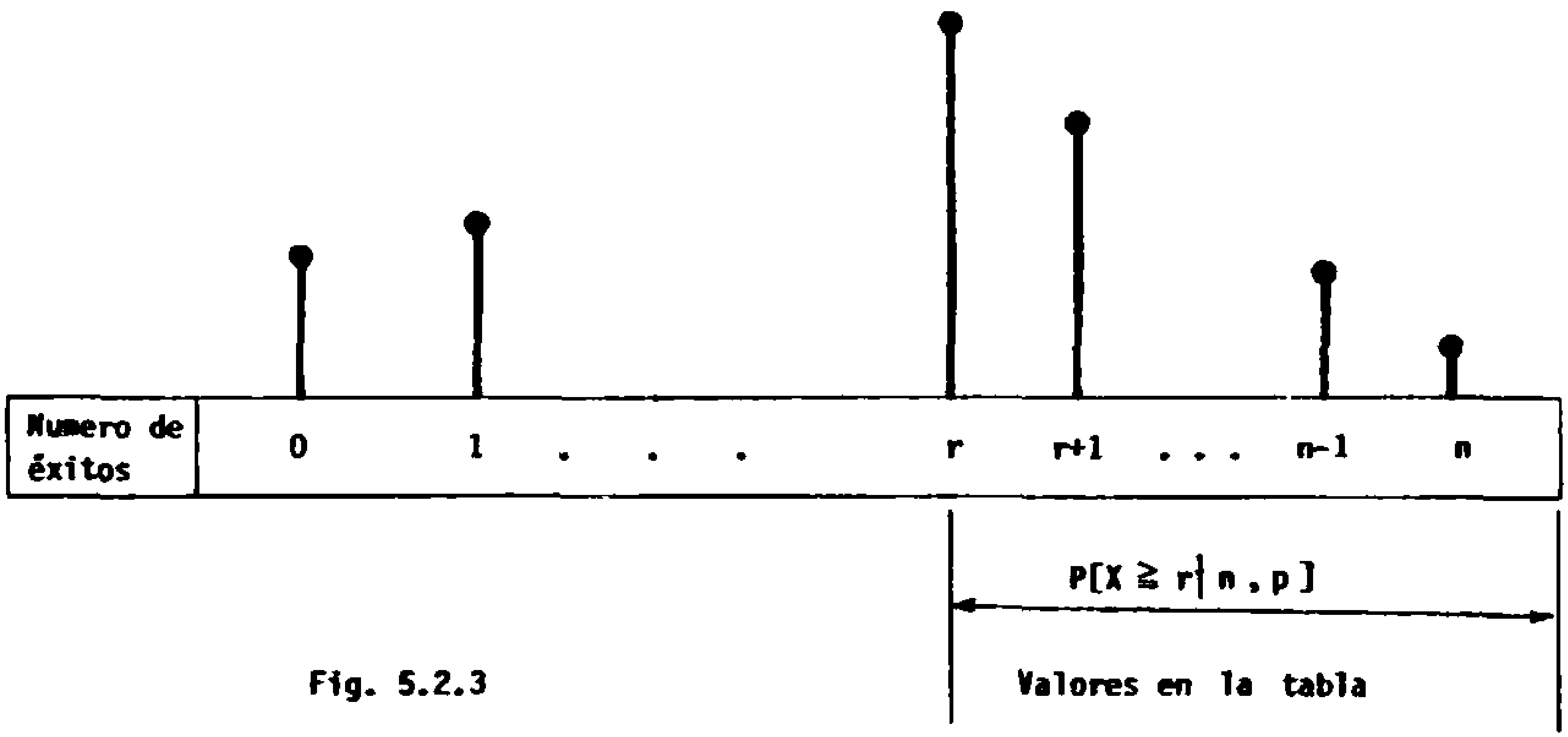


Fig. 5.2.3

Algunas de la probabilidades más simples que se presentan, se da en el siguiente cuadro

Probabilidades que queremos calcular		Probabilidades dadas en la tabla
$P[X \geq r \mid n,p]$	=	$P[X \geq r \mid n,p]$
$P[X > r \mid n,p]$	=	$P[X \geq r + 1 \mid n,p]$
$P[X = r \mid n,p]$	=	$P[X \geq r \mid n,p] - P[X \geq r + 1 \mid n,p]$
$P[X < r \mid n,p]$	=	$1 - P[X \geq r \mid n,p]$
$P[X \leq r \mid n,p]$	=	$1 - P[X \geq r + 1 \mid n,p]$

Utilizando un diagrama similar a la figura 5.2.3, se tendrá :

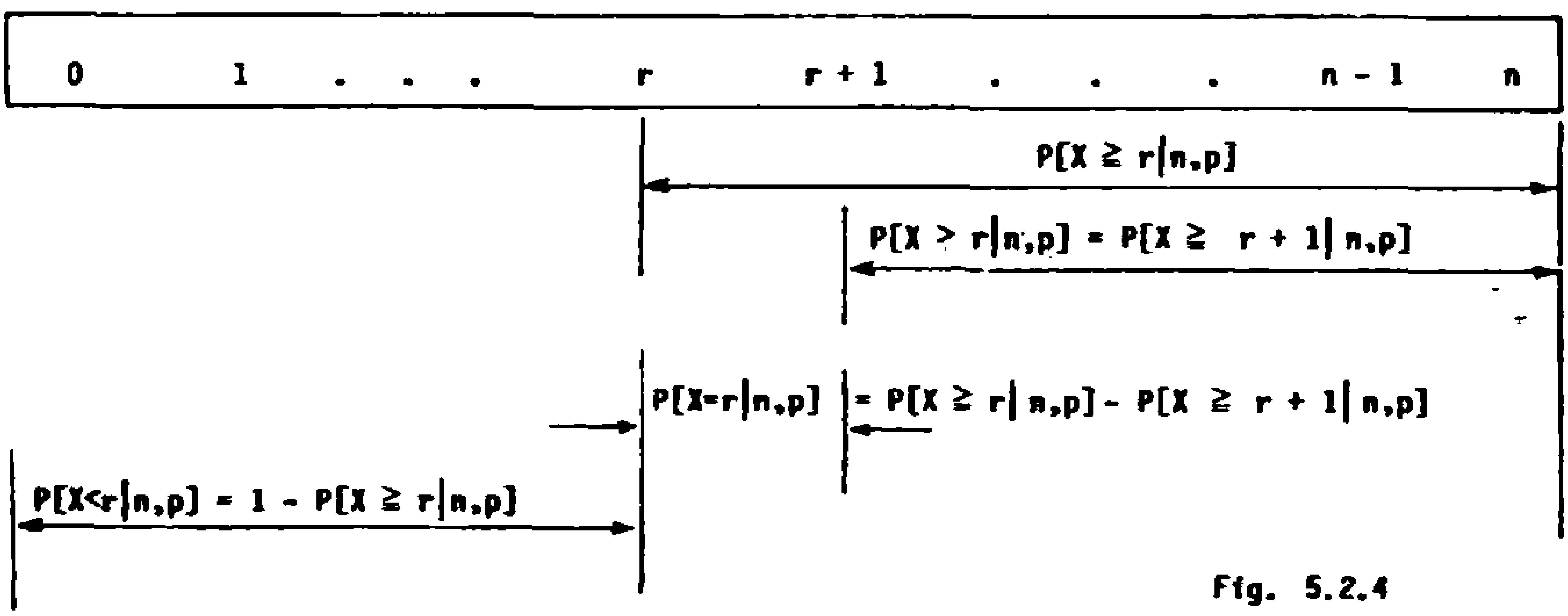


Fig. 5.2.4

EJEMPLO 12 Calcular :

(a) $P[X = 2|4,0.23]$, (b) $P[X < 10|20,0.3]$

SOLUCION

(a) $P[X = 2|4,0.23] = P[X \geq 2|4,0.23] - P[X \geq 3|4,0.23]$

$= 0.2285 - 0.00403 = 0.1882.$ (ver fig. 5.2.5)

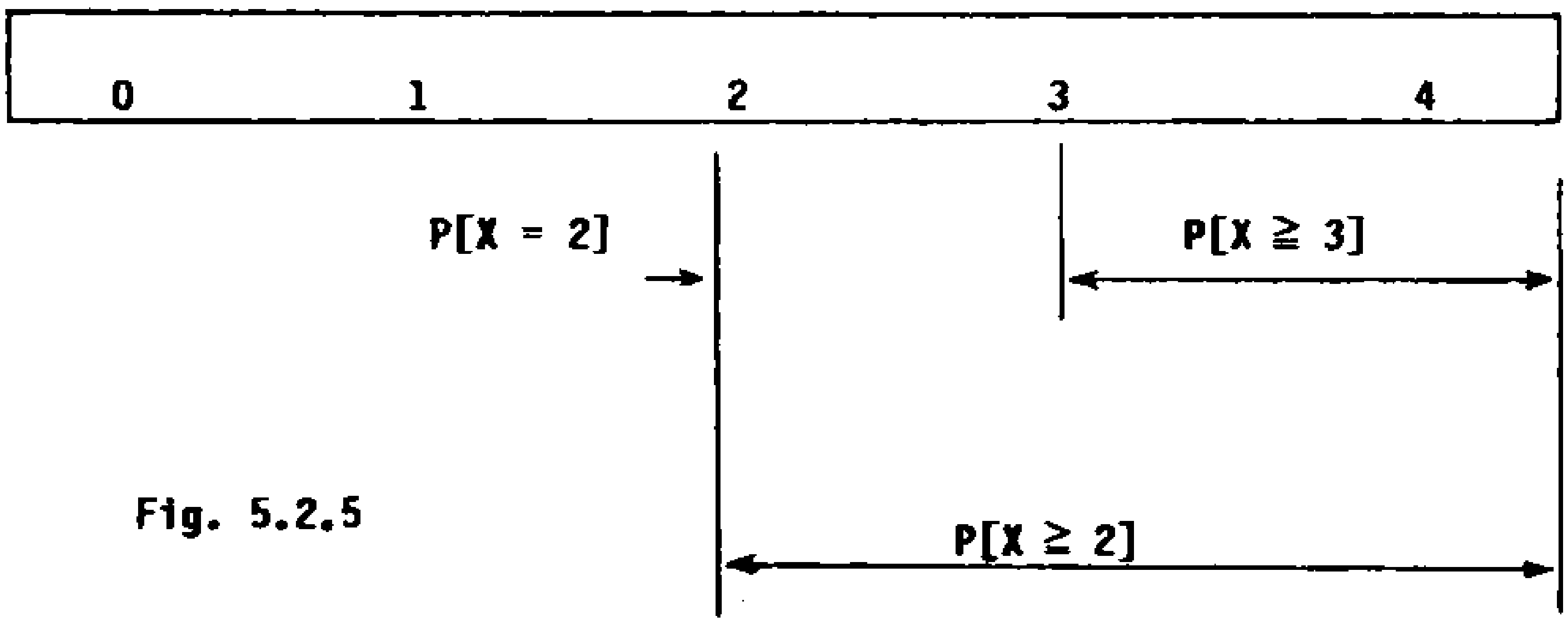


Fig. 5.2.5

De la gráfica se observa, $P[X = 2] = P[X \geq 2] - P[X \geq 3]$

(b) $P[X < 10|20,0.3] = 1 - P[X \geq 10|20,0.3] = 1 - 0.0480 = 0.9520$

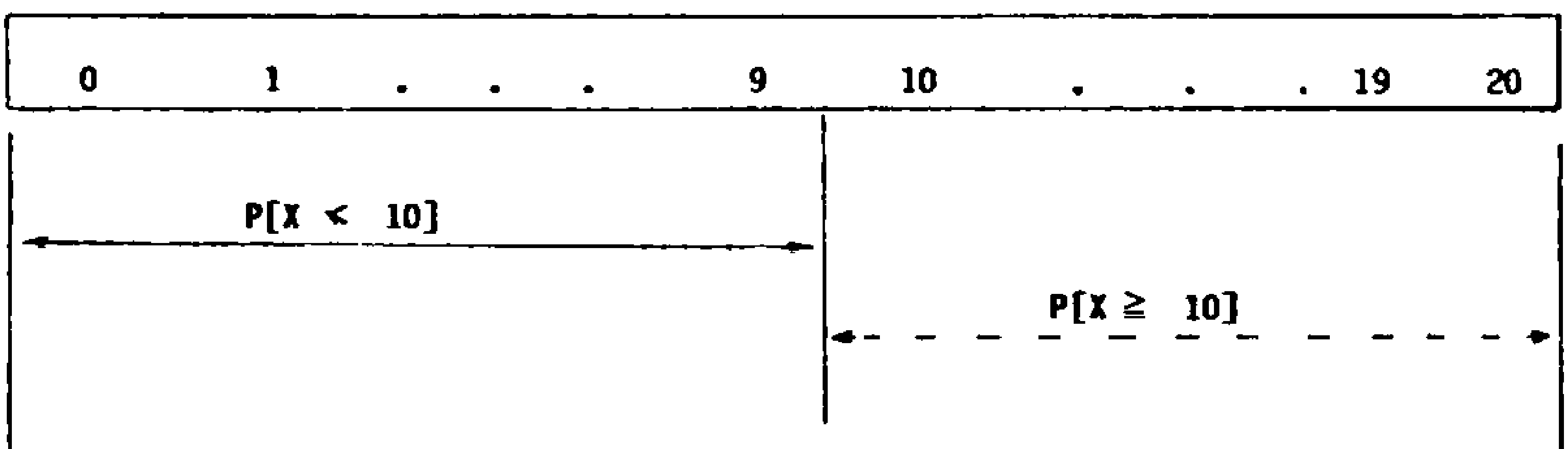


Fig. 5.2.6

de la figura: $P[X < 10] = 1 - P[X \geq 10]$

Ahora usando la simetría de la distribución binomial se ha eliminado en la tabla los valores de p que exceden a 0.5. Si se tiene un problema donde la probabilidad de éxito p es mayor que 0.5, debemos usar el siguiente procedimiento.

La figura 5.2.7 muestra como se utiliza la simetría de la distribución binomial. Consideremos las variables aleatoriamente X y X'. Primero consideremos la ocurrencia de la variable aleatoria X como el número de éxitos.

Luego sonsideremos la ocurrencia de X' como el número de fracasos.

Número de éxitos, X	$n \quad n-1 \; \ldots . \; r \; \ldots .. \; 1 \quad 0$	éxitos con probabilidad $p > 0.5$

Números de fracasos X'	$0 \quad 1 \; \ldots . \; n-r \; \ldots . \; n-1 \quad n$	Fracaso con probabilidad $1 - p < 0.5$

Fig. 5.2.7

Luego, si $p > 0.5$, se hace las siguientes sustituciones

(1) Se sustituye r por $n - r$

(2) Se sustituye p por $1 - p$

(3) Se invierte las desigualdades ($\geq$ se cambia por $\leq$; $<$ por $>$, así sucesivamente)

(4) Use el procedimiento descrito para calcular probabilidades con $p < 0.5$.

EJEMPLO 13 Calcular las siguientes probabilidades :

(a) $P[X < 4|10,0.8]$ (b) $P[X > 4|10,0.8]$

(c) $P[X \geq 4|10,0.8]$ (d) $P[X = 4|10,0.8]$

(e) $P[X \leq 4|10,0.8]$

SOLUCION

(a) Debemos determinar $P[X < 4|10, 0.8]$, entonces

$$n = 10, \qquad r = 4, \qquad p = 0.8 .$$

Luego, $n - r = 6, \qquad 1 - p = 0.2$

Por lo tanto, se tiene,

$$P[X < 4|10, 0.8] = P[X' > 6|10,0.2]$$
$$= P[X' \geq 7|10,0.2] = 0.0009$$

esta relación puede verse en la figura 5.2.8 .

Número de éxitos X	10 9 8 7 ...4 3 2 1 0	$P[E] = 0.80$

$P[X < 4|10,0.8]$

$P[X' \geq 7|10,0.2]$

Números de fracasos X'	0	1	2	3	4 ...	7	8	9	10	P[F] = 0.20

Fig. 5.2.8

Siguiendo un procedimiento similar obtenemos,

(b) $P[X > 4|10,0.8]$, aquí $n = 10$, $r = 4$, $p = 0.8$

Luego, $n - r = 6$, $1 - p = 0.2$, por lo tanto,

$$P[X > 4|10,0.8] = P[X' < 6|10,0.2]$$
$$= 1 - P[X' \geq 6|10,0.2]$$
$$= 1 - 0.0064 = 0.9936 .$$

(c) $$P[X \geq 4|10,0.8] = P[X' \leq 6|10,0.2]$$
$$= 1 - P[X' \geq 7|10,0.2]$$
$$= 1 - 0.0009 = 0.9991 .$$

(d) $$P[X = 4|10,0.8] = P[X' = 6|10,0.2]$$
$$= P[X' \geq 6|10,0.2] - P[X' \geq 7|10,0.2]$$
$$= 0.0064 - 0.0009 = 0.0055 .$$

(e) $$P[X \leq 4|10,0.8] = P[X' \geq 6|10,0.2] = 0.0064 .$$

EJEMPLO 14 Se ha elaborado un examen de selección múltiple consistente en 10 preguntas. Hay cuatro respuestas posibles para cada pregunta. Suponga que ninguno de los estudiantes que van a rendir el test concurrió a clase o que no estudió para el examen (cosa muy frecuente). El profesor que toma la prueba ha establecido que para aprobar debe contestar correctamente al menos 6 preguntas. Si hubiese 100 alumnos en la clase, ¿cuántos alumnos teóricamente aprobarián?

SOLUCION 1 Puesto que ninguno de los alumnos asistió a clase o no estudió para el examen, la elección de la respuesta en cada una de las 10 preguntas se hará al azar; Por lo tanto la elección de la respuesta en cada pregunta se considera como un ensayo de Bernoulli, con

p = Probabilidad de acertar la respuesta correcta $= \frac{1}{4}$ y $q = \frac{3}{4}$

2. El experimento se repite 10 veces. Es decir $n = 10$

3. Definimos la variable aleatoria X por

$X(\omega)$ = número de preguntas correctas en las 10 preguntas

$$R_X = \{0, 1, 2, 3, \ldots, 10\}$$

4. La variable aleatoria X, así definida es una variable aleatoria binomial, por lo tanto su distribución es,

$$P[X = x] = \binom{10}{x}\left(\frac{3}{4}\right)^{10-x}\left(\frac{1}{4}\right)^{x} \quad , \qquad x = 0,1, \ldots, 10 .$$

5. Para aprobar el examen debe contestar al menos 6 preguntas correctas. Es decir, la probabilidad de aprobar el examen es,

$$P[X \geq 6] = \sum_{x=6}^{10}\binom{10}{x}\left(\frac{3}{4}\right)^{10-x}\left(\frac{1}{4}\right)^{x}$$

buscando en la tabla I obtenemos

$$P[X \geq 6] = 0.0197$$

Por lo tanto, aprobarían teóricamente el examen,

$$100\ (0.0197) = 1.97 \approx 2, \quad \text{alumnos}$$

EJEMPLO 15 El tiempo de llegada en minutos X de camiones, a un depósito, se comporta de acuerdo a la siguiente función de densidad,

$$f(x) = \begin{cases} 1/x^2 & , \quad 1 < x < \infty \\ 0 & , \quad \text{en otro lugar} \end{cases}$$

Se desea alegir una muestra aleatoria de 8 camiones. Determinar la probabilidad que al menos dos de los camiones por elegir tengan un tiempo de llegada menor de 3 minutos.

SOLUCION 1 Calcularemos primero la probabilidad que un camión tenga un tiempo de llegada menor de 3 minutos. Para esto definimos la variable aleatoria X de la siguiente manera

$X(\omega)$ = tiempo de llegada de cada camión en minutos.

Entonces, $$P[X < 3] = \int_1^3 \frac{1}{x^2}\, dx = \frac{2}{3}$$

2. Definimos ahora la variable aleatoria Y de la siguiente manera,

   $Y(\omega)$ = número de camiones que tienen un tiempo de llegada menor de 3 minutos, en la muestra de 8.

$R_Y = \{0, 1, 2, 3, \ldots, 8\}$

3. Y es una variable aleatoria binomial, con $n = 8$, $p = \frac{2}{3}$ y $q = \frac{1}{3}$

$$P[Y = y] = \binom{8}{y}\left(\frac{1}{3}\right)^{8-y}\left(\frac{2}{3}\right)^{y}, \qquad y = 0, 1, 2, \ldots, 8$$

4. Se pide calcular $P[Y \geqslant 2]$

$$P[Y \geq 2|8, 2/3] = P[Y' \leq 6|8, 1/3] = 1 - P[Y' \geq 7|8, 1/3]$$

$$= 1 - 0.0026 = 0.9976 . \qquad \text{(Tabla I)}$$

EJEMPLO 16 Dos personas juegan a cara o sello y han convenido en continuar la partida hasta que tanto cara como sello hayan aparecido por lo menos tres veces. Hallar la probabilidad que el juego no se acabe cuando se han realizado 10 tiradas.

SOLUCION 1 Definimos la variable aleatoria X como sigue

$X(\omega)$ = número de caras obtenidas en diez tiradas.

$$R_X = \{0,1,2,3, \ldots, 10\}$$

2. X tiene una distribución binomial con parámetros $n = 10$ y $p = \frac{1}{2}$

$$P[X = x|10, 0.5] = \binom{10}{x}\left(\frac{1}{2}\right)^{10}, \quad x = 0,1, \ldots, 10.$$

3. Para que el juego no termine en 10 tiradas, debe ocurrir que :

$X \leq 2$, número de caras menores que 3 .

ó $X \geq 8$, número de sellos menores que 3 .

Es decir, la probabilidad pedida es

$$P[\{X \leq 2\} \cup \{X \geq 8\}] = P[X \leq 2] + P[X \geq 8] \quad \text{(Eventos excluyentes)}$$

$$= 1 - P[X \geq 3|10,0.5] + P[X \geq 8|10,0.5]$$

$$= 1 - 0.9453 + 0.0547 \quad \text{(tabla I)}$$

$$= 0.1094 .$$

EJEMPLO 17 Un cuerpo se encuentra en reposo, en el punto (0,0). Se lanza un dado y por cada número primo que aparece el cuerpo se desplaza 1 unidad de longitud hacia la derecha, en caso contrario se desplaza una unidad a la izquierda. Calcular la probabilidad que después de 10 lanzamientos el cuerpo se encuentre :

(a) a 8 unidades de longitud a la derecha del origen;
(b) a 3 unidades de longitud a la derecha del origen;

(c) a 2 unidades de longitud a la izquierda del origen;

(d) a más de una unidad a la derecha del origen.

SOLUCION 1 El experimento es lanzar un dado 10 veces.

2. Definimos la variable aleatoria X de siguiente manera;

$X(\omega)$ = número de veces que aparece un número primo en los 10 lanzamientos del dado.

$$R_X = \{0,1,2,3,4,5,6,7,8,9,10\}$$

$$p = P[E] = P[\{N^\circ \text{ primo}\}] = P[\{1,2,3,5\}] = \frac{4}{6} = \frac{2}{3}$$

$$q = P[F] = P[\{N^\circ \text{ no primo}\}] = P[\{4,6\}] = \frac{1}{3}\ .$$

3. La variable aleatoria X tiene una distribución binomial con parámetros

$n = 10$ y $p = \frac{2}{3}$. O sea

$$P[X = x|\,10,\ \frac{2}{3}\,] = \binom{10}{x}\ (\frac{2}{3})^x(\frac{1}{3})^{10-x}\quad ,\quad x = 0,1,\ldots,10\ .$$

4. Definimos ahora una nueva variable aleatoria Y, así

$Y(\omega)$ = posición del cuerpo con respecto al origen (0,0) .

$$R_Y = \{-10,\ -8,\ -6,\ -2,0,2,4,6,8,10\}$$

5. Observe que la variable aleatoria Y puede escribirse como una función de X,

$$y = 2\,X - 10$$

Por lo tanto, la distribución de probabilidad de Y se obtiene de la distribución de probabilidad de X.

(a) $P[Y = 8] = P[X = 9] = \binom{10}{9}(\frac{2}{9})^9(\frac{1}{3}) = 0.0898$

(b) $P[Y = 3] = 0$, pués $3 \notin R_y$.

(c) $P[Y = -2] = P[X = 4] = \binom{10}{4}(\frac{2}{3})^4(\frac{1}{-})^6 = 0.2252$.

(d) $P[Y \geq 2] = P[X \geq 6] = \sum_{x=6}^{10}\binom{10}{x}(\frac{2}{3})^x(\frac{1}{3})^{10-x} = 0.7936$.

5.2.3 NUMERO MAS PROBABLE DE REPETICIONES DE SUCESOS

Expondremos aquí dos problemas importantes con respecto a la distribución binomial. La primera es que cuando los valores de x crecen, las probabilidades

$p(x) = P[X = x|n,p]$ aumentan y, después de alcanzar un valor máximo, disminuye. Este hecho, se aprecia mejor representando la distribución de probabilidad en un diagrama (por ejemplo ver Fig. 5.2.1). Cuando n es grande, estos diagramas proporcionan un cuadro aún más convincente de la variación de $p(x)$ para los x crecientes.

La segunda es encontrar el número más probable de repeticiones de un suceso; o sea para qué valor de x la probabilidad $p(x) = P[X = x|n,p]$ alcanza su valor máximo (n y p conocidos).

A continuación proporcionaremos una solución a los problemas planteados. Calculemos primero el valor del cociente $p(x + 1) / p(x)$. Por la fórmula de la distribución binomial es

$$\frac{p(x+1)}{p(x)} = \frac{\dfrac{n!}{(x+1)!(n-x-1)!} \cdot p^{x+1}(1-p)^{n-x-1}}{\dfrac{n!}{x!(n-x)!} \cdot p^{x}(1-p)^{n-x}} = \frac{n-x}{x+1} \cdot \frac{p}{1-p}$$

$p(x + 1)$ es mayor, igual o menor que $p(x)$, según que el cociente $p(x+1)/p(x)$ sea mayor, igual o menor que uno. O en forma equivalente según que :

$$\frac{n-x}{x+1} \cdot \frac{p}{1-p} > 1 \; ; \quad \frac{n-x}{x+1} \cdot \frac{p}{1-p} = 1 \; ; \quad \frac{n-x}{x+1} \cdot \frac{p}{1-p} < 1 \qquad (1)$$

Para encontrar, para qué valores de x se cumple la desigualdad $p(x+1) > p(x)$, es suficiente saber para qué valores de x se cumple la desigualdad

$$\frac{n-x}{x+1} \cdot \frac{p}{1-p} > 1 \qquad \text{ó} \qquad (n-x)p > (x+1)(1-p)$$

de donde se obtiene

$$np - (1-p) > x \; .$$

Esto significa que para valores de x menores que $np - (1 - p)$ valdrá la desigualdad $p(x + 1) > p(x)$, es decir, cuando x crece $p(x)$ también crece.

Análogamente, partiendo de las otras dos relaciones de (1) se tendrá que :

$$p(x+1) = p(x) \quad , \quad \text{si} \quad x = np - (1-p) \; ,$$

y

$$p(x+1) < p(x) \quad , \quad \text{si} \quad x > np - (1-p)$$

Este último indica que para valores de x mayores a $np - (1 - p)$, $p(x)$ disminuye hasta alcanzar el valor $p(n)$.

En conclusión el comportamiento de la probabilidad $p(x)$ para x creciente (primero aumenta y luego disminuye) es una ley general que rige en todos los casos. Tal conclusión permite resolver inmediatamente el segundo problema -

planteado, es decir determinar el valor más probable de x .

Designaremos con x_0 el valor más probable de x. Luego debemos tener :

(i) $p(x_0 + 1) \leq p(x_0)$, la cual se cumple, si $x_0 \geq np - (1 - p)$

(ii) $p(x_0 - 1) \leq p(x_0)$, la cual se cumple, si $x_0 - 1 \leq np - (1 - p)$

ò $x_0 \leq np - (1 - p) + 1 = np + p$.

de (i) y (ii) el valor más probable x_0 de x tiene que satisfacer una doble desigualdad

$$\boxed{np - (1 - p) \leq x_0 \leq np + p}$$

El intervalo $[np - (1 - p) , np + p]$ donde está x_0, es de longitud 1.

OBSERVACIONES

(a) Cuando un extremo del intervalo, por ejemplo $np - (1 - p)$ no es un entero, el otro tampoco es un entero, existirá entonces entre estos dos extremos sólo un número entero. En tal caso existirá sólo un número x_0 más probable.

(b) Si el número $np - (1 - p)$ es un entero, tambien lo es $np + p$. En tal caso existirán dos números más probables

$$x_0 = np - (1 - p) \quad , \quad x_0 + 1 = np + p$$

sus probabilidades son iguales entre si y mayores que las probabilidades para los x restantes.

(c) Si np es un entero, el número más probables es $x_0 = np$.

EJEMPLO 18 Sea X una variable aleatoria binomial cuya distribución es $p(x) = P[X = x|20, \frac{2}{3}]$. Hallar los valores de x más probable.

SOLUCION Aqui $n = 20$ y $p = \frac{2}{3}$, reemplazando estos valores en la fórmula, $np - (1 - p) \leq x_0 \leq np + p$ obtenemos

$$20 \cdot \frac{2}{3} - (1 - \frac{2}{3}) \leq x_0 \leq 20 \cdot \frac{2}{3} + \frac{2}{3}$$

$$13 \leq x_0 \leq 14 .$$

De acuerdo con la observación (b) los números más probables son 13 y 14 .

EJEMPLO 19 Según observaciones realizadas durante muchos años en Lima, se determinó que la probabilidad de que lloviera el 1° de Agosto era igual a

4/17. ¿Cuál será, el valor más probable de este suceso en los proximos 50 - años?

SOLUCION Aquí, $n = 50$ y $p = 4/17$. Por lo tanto

$$50 \cdot \frac{4}{17} - (1 - \frac{4}{17}) \leq x_o \leq 50 \cdot \frac{4}{17} + \frac{4}{17}$$

$$11 \leq x_o \leq 12 \ .$$

Los números de días lluviosos más probables son 11 y 12

PROBLEMAS 5.2

1. Si **X** es una variable aleatoria discreta con distribución binomial de parámetros n y p tales que $E(X) = 5$ y $Var(X) = 4$. Hallar n y p y el valor más probable.

2. La probabilidad que en cierto establecimiento industrial el consumo de energía eléctrica sea normal (es decir no sobrepase un número determinado de kw) en 24 horas es igual a 3/4. Determinar la distribución de probabilidad del número de días de consumo normal de energía eléctrica en un lapso de seis días; Represente en un diagrama esta distribución. ¿Cual es la probabilidad que haya 4 días de consumo normal?.

3. Se lanza un par de dados cuatro veces. ¿Cuál es la probabilidad que la suma 9 aparezca exactamente dos veces?

4. Una máquina produce cierto tipo de piezas, de las cuales un promedio de 5% son defectuosas. En una muestra aleatoria de cinco piezas. ¿Cuál es la - probabilidad de obtener,
   (a) exactamente una pieza defectuosa?
   (b) por lo menos una pieza defectuosa?

5. En una población de drosophila, el 20% tienen mutación de alas. Si se escogen 6 moscas aleatoriamente de la población .
   (a) ¿Cuál es la probabilidad que dos tienen mutación ?
   (b) ¿Al menos uno tiene mutación?
   (c) ¿Qué menos de 5 tienen mutación?
   (d) ¿Cuál es el número esperado de moscas con mutación de alas?

6. Un tratamiento para cierta enfermedad produce una cura en 75% de los casos. Se seleccionan 6 pacientes aleatoriamente. ¿Cuál es la probabilidad de que :

(a) Todos estan curados?
(b) ninguno está curado?
(c) cuatro están curados?
(d) al menos cuatro están curados?

7. Sea X una variable aleatoria con distribución binomial, cuya media es 12 y varianza 4.8. Calcular :

(a) $P[X > 5]$, (b) $P[5 < X < 10]$; (c) $P[X \leq 10]$

8. El 90% de los tubos de ensayo soportan temperatura mayor que 80 °C , suponga que 10 de estos tubos se someten a una prueba a temperaturas mayores de 80°C. Determine la probabilidad que 3 de estos queden inutilizables.

9. La probabilidad de fallar durante el vuelo para cada uno de los seis motores de un avión es 0.0005. Suponiendo que los seis motores trabajan independientes, determine la probabilidad que en un vuelo determinado
(a) no ocurra ninguna falla de motor
(b) no ocurra más de una falla
(c) ocurra exactamente dos fallas.

10. Suponga que los motores de un avión de cierta marca, que operan independientemente, tienen una probabilidad de falla de 0.1. Suponga que un avión efectúa un vuelo exitoso si almenos la mitad de sus motores operan normalmente, determine cuál avión, uno con cuatro y otro con seis motores, tiene mayor probabilidad de efectuar un vuelo exitoso.

11. Cierto tubo de televisión tiene una probabilidad de 0.3 de funcionar más de 400 horas. Se prueban 15 tubos
(a) Hallar la probabilidad que exactamente 0, 4, 9 de ellos funcionan más de 400 horas.
(b) Cuántos tubos espera encontrar que funcionen por lo menos 400 horas.
(c) Cuál es el número de tubos más probable que funcionan por lo menos 400 horas.

12. La probabilidad de hacer una venta de cierto vendedor en un intento es $\frac{1}{2}$.
¿Cuál es la probabilidad de obtener.
(a) Exactamente dos ventas en tres intentos de ventas consecutivas?
(b) por lo menos una venta en tres intentos de ventas consecutivos?
(c) ¿Cuántos intentos de ventas consecutivas deben hacerse para obtener una seguridad de 0.9375 de obtener por lo menos una venta?

13. Suponga que la máquina A produce el doble de artículos que la máquina B. Se sabe que el 6% de los artículos que produce la máquina A son defectuosas, mientras que el 3% de los artículos producidos por la máquina B son defectuosas. Suponga que se junta la producción diaria de estas máquinas y se toma una muestra aleatoria de 10 artículos. Calcular la probabilidad de obtener 3 artículos defectuosos.

14. El departamento de contabilidad de una firma comercial tiene dos empleados a tiempo parcial: Manuel y Manuela. Manuel trabaja los lunes, miercoles y viernes en tanto que Manuela lo hace los martes, jueves y sábado. - Manuel archivo erróneamente uno de cada cinco documentos, mientras que Manuela lo hace uno de cada seis. Se elige al azar un día de la semana y en ese día se toma una muestra de seis documentos de entre los documentos archivados ese día.
    (a) ¿Cuál es la probabilidad de que la muestra contenga exactamente 3 documentos mal archivados?
    (b) Suponga que la muestra contenga exactamente 3 documentos mal archivados, ¿cuál es la probabilidad de que hayan sido archivados por Manuel?

15. Se sabe que la probabilidad de que germine una sola semilla de cierta clase es 0.9. Un agricultor quiere vender cultivos de esta planta, para lo cual asevera que cada uno contiene 100 plantas. Si siembra 110 semillas - en cada cultivo (que se supone germinarán en forma independientes). ¿Cuántas plantas se puede esperar que contenga un cultivo "promedio"?

16. Se sabe que el 10% de los vasos fabricados por determinada máquina tienen algún defecto. Si se seleccionan al azar 10 de los vasos fabricados - esta máquina, ¿cuál es la probabilidad que ninguno este defectuoso? - ¿Cuántos defectuosos esperaría encontrar?

17. Un examen consta de 20 preguntas, cada una tiene 5 respuestas de las cuales solamente una es correcta. Un estudiante que desconoce el curso contesta el examen al azar.
    (a) ¿Cuál es la probabilidad de que acierte más de 10 respuestas correctas
    (b) ¿Cuál es el número esperado de respuestas correctas?

18. Se lanza un dado 1620 veces. ¿Cuál es la media, la varianza y la desviación estandar, del número de seis obtenidos?

19. De las piezas metálicas producidas por una máquina el 5% son defectuosas y los 95% restantes son buenos. ¿Cuántas piezas debe producir la máquina

para que la probabilidad que haya al menos una pieza defectuosa sea igual o mayor que 1/2?

20. Las máquinas A y B producen un promedio, 5% y 10% de piezas defectuosos respectivamente. Se extrae una muestra de 10 piezas de la producción de cada una de las máquinas. ¿Cuál es la probabilidad que la muestra obtenida de la producción de A contenga exactamente una pieza defectuosa y la muestra correspondiente a B contenga exactamente 2 piezas defectuosas?

21. Si X es una $b(100, 1/2)$, dar una cota inferior para

$$P\left[\left|\frac{X}{100} - \frac{1}{2}\right| < 0.1\right]$$

22. En una escuela profesional de cuatro años, el 50% de los alumnos están en el primer año, el 25% en el segundo, el 15% en tercero y el 10% en cuarto. Se selecciona 5 alumnos al azar. ¿Cuál es la probabilidad que:
(a) Exactamente 2 sean del primer año?
(b) Ninguno sea del tercero o cuarto año?

23. El 60% de los televidentes de una población grande dada sintonizan un programa específico, ¿Cuál es la probabilidad que más de la mitad de las personas que forman una muestra de cinco personas, extraída aleatoriamente de la población, vean el programa de televisión?

24. De un lote de 12 tubos de televisión, 3 de ellos son defectuosos. Si se extrae una muestra aleatoria de 3 de este lote con reemplazo. ¿Cuál es la probabilidad que
(a) Exactamente uno sea defectuoso?
(b) ninguno o uno sea defectuoso? .

25. Un examen consta de 10 preguntas, con 5 respuestas cada una de las cuales sólo una es correcta. Un estudiante que desconoce el curso contesta la prueba al azar. Para aprobar el examen debe contestar correctamente al menos 6 preguntas
(a) ¿Cuál es la probabilidad que el estudiante apruebe el examen
(b) Si al examen se presentan 200 estudiantes que desconocen el curso, ¿cuántos alumnos, se espera que aprueben el examen?

26. Un laberinto para ratas tiene un corredor recto, y al final una bifurcación; en la bifurcación, la rata debe ir a la derecha o la izquierda. Suponga que se colocan 10 ratas en el laberinto, de una en una. Si cada una de las ratas toma al azar una de las dos alternativas del camino, ¿Cuál -

es la distribución del número de las que van a la derecha? ¿De las que van a la izquierda? ¿Cuál es la probabilidad que cuando menos 9 vayan al mismo lado?

27. Una variable aleatoria X tiene por función de densidad,

$$f(x) = \begin{cases} x/2 & , \quad 0 < x < 2 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

si se extraen 10 valores de X.

(a) ¿Cuál es la probabilidad que exactamente cuatro sean mayores que 1?

(b) ¿Cuál es la probabilidad que por lo menos cuatro sean mayores que 1?

28. En una feria (hace mucho tiempo), comprando un boleto de 10 soles se podía participar en un juego que consistía en lanzar 6 argollas para embocar en una botella de madera. Los premios del juego eran:
Una bolsa de caramelos (valor 1 sol) al embocar de 1 a 3 argollas,
Un tarro de duraznos (valor 4 soles) al embocar 4 argollas.
Una botella de vermouth (valor 17 soles) al embocar 5 argollas,
Una caja de cigarrillos (valor 31 soles) al embocar las 6 argollas.
Sabiendo que el jugador promedio tiene una probabilidad 1/3 de embocar cada una de las argollas, y que se vende un promedio de 729 boletos por día, determinar cuáles son los ingresos netos diarios del dueño del juego.

29. Dos personas juegan a cara o sello con una moneda normal y convienen en continuar la partida hasta que tanto cara como sello hayan aparecido por lo menos cuatro veces. Hallar la probabilidad que el juego no se acabe cuando se han realizado 12 tiradas.

30. Un cuerpo se encuentra en reposo en el punto (0,0). Se lanza un dado 10 veces y por cada número mayor que 2 que aparece, el cuerpo se desplaza una unidad de longitud hacia la derecha, en caso contrario se desplaza una unidad hacia la izquierda. Calcular la probabilidad que después de los 10 lanzamientos el cuerpo se encuentre a :

(a) 8 unidades de longitud a la derecha del origen

(b) 3 unidades a la derecha del origen.

(c) 2 unidades a la izquierda del origen .

31. Una empresa produce artículos de clase A y B que son vendidos en paquetes conteniendo 10 artículos de clase A y 20 de clase B. Se fija un precio ba

se de 250 intis por cada paquete. Pero para fijar el precio final de cada paquete, se extrae de el 10 artículos. Por cada artículo **A** que aparece se aumenta el precio en 5 intis, y por cada artículo **B** se disminuye 5 intis. Calcular:

(a) la probabilidad que un paquete sea vendido a 270 intis

(b) la probabilidad que un paquete sea vendido a un precio menor que 240 intis.

(c) ¿Cuál es el precio esperado por paquete?

Resolver el problema considerando, extracción con reposición.

32. Un vendedor de radios y televisores otorga créditos a sus clientes. Suponga que anteriormente 10 % de todos los deudores no pagaron y que el vendedor tuvo que absorver la pérdida de cada venta; el 90 % restante pagó completamente sus créditos, y el vendedor obtuvo una utilidad en esas ventas. Suponga que ese vendedor tiene 10 televisores idénticos que va a vender individual e independientemente a crédito a 10 personas. Si el comprador no paga, la pérdida es de $ 200; si el comprador paga, entonces su utilidad es de $ 100.

(a) ¿Cuál es la distribución de probabilidad del monto de la utilidad obtenida en estas 10 ventas?

(b) ¿Cuál es su utilidad esperada en esas 10 ventas ?

33. Un cuerpo **A** se encuentra en reposo en el punto (3,0). Un segundo cuerpo B se encuentra en reposo en el punto (0,0). Ud. lanza un dado 12 veces y por cada número que aparece el cuerpo B lo desplaza verticalmente una unidad hacia arriba y en caso contrario lo desplaza una unidad hacia abajo. Calcular la probabilidad que después de los 12 lanzamientos los cuerpos se encuentren a 5 unidades de distancia.

34. Una obra de 10 volúmenes se coloca al azar en un estante de 5 compartimientos. Suponga que cada volumen tiene igual probabilidad de ser ubicado en cada uno de los 5 comportamientos. Calcule la probabilidad que más de la mitad de la obra quede colocada en el tercer comportamiento del estante.

35. En los días hábiles uno de cada 6 teléfonos a que se llame en cierta ciudad estará ocupado. Si se llama 10 veces a teléfonos seleccionados al azar, ¿Cuál es la probabilidad que no más de dos estén ocupados?

36. Una oficina emplea a 10 mecanógrafas. Cada una requiere una cinta para máquina de escribir más o menos cada siete semanas. Si el empleado del alma

cén ve que al principio de esa semana sólo tiene cinco cintas, ¿Cuál es - la probabilidad que éstas se agoten esa semana?

37. Exactamente el 60 % de los trabajadores de una planta, pertenecen a un - sindicato. Si el administrador extrae una muestra aleatoria de 15 trabaja jadores. ¿Cuál es la probabilidad que
(a) exactamente 8 pertenezcan al sindicato? (b) 8 o más pertenezcan a él?

38. Una fábrica emplea un patrón de aceptación de los artículos producidos antes de embarcarlos. El plan consiste en lo siguiente: cajas de 25 artículos son preparados para su embarque; un inspector toma un artículo al azar, lo - inspecciona y lo devuelve a la caja; un segundo y un tercero efectúan el mismo procedimiento. Si cualquiera de los tres inspectores encuentra un - artículo defectuoso, la caja no se embarca. ¿Cuál es la probabilidad de:
(a) embarcar una caja que contenga tres artículos defectuosos?
(b) no embarcar una caja que contiene sólo un artículo defectuoso?

39. Una máquina es accionada por tres baterias, todas del mismo tipo, y funcionará siempre que dos de ellos trabajen adecuadamente. La probabilidad de que una batería de este tipo falle durante las primeras 8 horas de operación de la máquina es 0.2; si supera las primeras 8 horas de operación, la probabilidad de falla durante las siguientes 8 horas es 0.4. Determine la probabilidad de que la máquina funcione continuamente durante:
(a) 8 horas , (b) 16 horas.

40. En una planta química, cada lote de un producto se purifica, haciendolo - pasar a trevés de una serie de 20 rejillas reactivas, que se renuevan en cada uno de los lotes. Cuando 2 o más de las rejillas de la secuencia de 20 están inactivas, la pureza del producto es inferior a la estándar. Las rejillas se montan en conjunto y contienen un 2 % de elementos inactivos. ¿Qué porcentaje de los lotes tendrá una impureza inferior después del proceso de purificación? Dé una expresión que indique la probabilidad de que x rejillas en una secuencia de 20 sean inactivas.

41. Un plan de muestreo doble para el manejo de grandes lotes de artículos - procede como sigue: tomar una muestra de 10 artículos; si no se encuentran defectuosos aprobar el lote; si se encuentran dos o más defectuosos, rechazar el lote; en caso contrario tomar una segunda muestra, esta vez de 15 piezas. Si no se encuentra más de una defectuosa en esta segunda muestra, aprobar el lote; en caso contrario rechazarlo. Si la proporción de -

defectuosos es 0.01, calcular la probabilidad

(a) que se acepte el lote, (b) se rechace el lote.

5.3 DISTRIBUCION GEOMETRICA Y BINOMIAL NEGATIVA

5.3.1 DISTRIBUCION GEOMETRICA

La distribución geométrica está también relacionado con un proceso de Bernoulli, excepto que el número de ensayos no es fijo. Consideremos entonces una sucesión de ensayos de Bernoulli. Definimos la variable aleatoria **X** de la siguiente manera,

$X(\omega)$ = número de ensayos requeridos hasta obtener el primer éxito.

$$R_X = \{1,2,3, \ldots\}$$

El espacio muestral tiene la siguiente forma,

$$\Omega = \{E, FE, FFE, FFFE, FFFFE, \ldots\}$$

La variable aleatoria **X**, así definida se llama una *variable aleatoria geométrica*. Y la distribución de probabilidad de esta variable aleatoria se llama *distribución geométrica*. Deduciremos ahora esta distribución de probabilidad

$$p(1) = P[X = 1] = P[E] = p$$

$$p(2) = P[X = 2] = P[FE] = pq$$

$$p(3) = P[X = 3] = P[FFE] = pq^2$$

En general se obtiene que;

$$p(x) = P[X = x] = P[\{\underbrace{FF \ldots F}_{(x-1)F}E\}] = pq^{x-1}$$

Por lo tanto, la distribución de probabilidad geométrica es

$$p(x) = P[X = x] = pq^{x-1}, \qquad x = 1,2,3, \ldots$$

Es obvio que, $p(x) \geq 0$ para todo x, y

$$\sum_{x=1}^{\infty} pq^{x-1} = p\sum_{k=0}^{\infty} q^k = p\left[\frac{1}{1-q}\right] = 1$$

La función de distribución acumulada está dado por,

$$F(x) = P[X \leq x] = \sum_{k=1}^{[\![x]\!]} pq^{k-1} = p\sum_{k=1}^{[\![x]\!]} q^{k-1} = p\left[\frac{1-q^{[\![x]\!]}}{1-q}\right]$$

Es decir,

$$F(x) = \begin{cases} 0 & , \quad x < 1 \\ 1 - q^{[\![x]\!]} & \quad x \geq 1 \end{cases}$$

La media y la varianza de la distribución geométrica se obtienen como sigue:

$$\mu = E(X) = \sum_{x=1}^{\infty} x\, p\, q^{x-1} = p \sum_{x=1}^{\infty} xq^{x-1} \tag{1}$$

Se tiene que $\displaystyle\sum_{x=0}^{\infty} q^x = \frac{1}{1-q}$, ya que $q < 1$

Luego, $\displaystyle\frac{d}{dq}\left[\sum_{x=0}^{\infty} q^x\right] = \sum_{x=1}^{\infty} xq^{x-1} = \frac{1}{(1-q)^2}$

Reemplazando en (1) obtenemos ,

$$\mu = E(X) = p\left[\frac{1}{(1-q)^2}\right] = \frac{1}{p}$$

$$\boxed{\mu = \frac{1}{p}}$$

Calculemos ahora $E(X^2)$. Se tiene que ,

$$E(X^2) = \sum_{x=1}^{\infty} x^2 pq^{x-1} \tag{2}$$

y

$$q\,E(X^2) = q\sum_{x=1}^{\infty} x^2 pq^{x-1} \tag{3}$$

Restando miembro a miembro (3) de (2)

$$(1-q)E(X^2) = p + \sum_{x=1}^{\infty}\left[(x+1)^2 - x^2\right] q^x p = p + \sum_{x=1}^{\infty}(2x+1)q^x p$$

$$p\,E(X^2) = p + \sum_{x=1}^{\infty} 2xq^x p + \sum_{x=1}^{\infty} q^x p$$

$$= p + 2q\sum_{x=1}^{\infty} xq^{x-1} p + q\sum_{x=1}^{\infty} q^{x-1} p$$

$$p\,E(X^2) = p + 2q\left(\frac{1}{p}\right) + q(1)$$

Luego, $E(X^2) = \dfrac{1+q}{p^2}$

Por lo tanto, $\sigma_X^2 = \text{Var}(X) = \dfrac{1+q}{p^2} - \dfrac{1}{p^2} = \dfrac{q}{p^2}$

$$\boxed{\sigma^2 = \frac{q}{p^2}} \quad \text{y} \quad \boxed{\sigma = \frac{\sqrt{q}}{p}}$$

PROPIEDADES DE LA DISTRIBUCION GEOMETRICA

1. La distribución geométrica es decreciente, es decir,

$$p(x) < p(x-1)\,, \quad \text{para} \quad x = 2,3,\ldots$$

2. Una interesante y útil propiedad de la distribución geométrica es que no tiene memoria, esto es,

$$P[X > x + h \mid X > h] = P[X > x]$$

La distribución geométrica es la única distribución discreta con esta propiedad.

EJEMPLO 1 La probabilidad de éxito al lanzar un cohete es 0.8 suponga que el ensayo del lanzamiento ha ocurrido.
¿Cuál es la probabilidad que exactamente sean necesarios 6 ensayos?

SOLUCION 1 Sea **X** la variable aleatoria definida por,
$X(\omega)$ = número de ensayos hasta que el cohete sea lanzado con éxito.

$$R_X = \{1,2,3,4,\ldots\} \;;\quad p = P[E] = 0.8, \quad q = 0.2$$

2. Luego, la variable aleatoria **X** tiene una distribución geométrica,

$$p(x) = P[X = x] = (0.8)(0.2)^{x-1}\,, \quad x = 1,2,\ldots$$

$$P[X = 6] = (0.8)(0.2)^5 = 0.000256\,.$$

EJEMPLO 2 Se dispone de un aparato que fabrica objetos de plástico. Este aparato se utiliza hasta que aparece el primer objeto defectuoso. Se sabe que: La probabilidad que el objeto sea no defectuoso es 8/9, la probabilidad de que el objeto sea defectuoso es 1/9. Sea **X** la variable aleatoria que da el número de objetos que produce el aparato hasta antes de darle de baja. Determine la distribución de probabilidad de **X**.

SOLUCION 1 Sea **X** la variable aleatoria definida por
$X(\omega)$ = el número de objetos que produce el aparato hasta antes de darle -

de baja.

$R_X = \{1,2,3, . . .\}$, $P[F] = q = \frac{8}{9}$ y $p = \frac{1}{9}$

2. Luego, la variable aleatoria **X** tiene una distribución geométrica,

$$p(x) = P[X = x] = (\tfrac{8}{9})^{x-1}(\tfrac{1}{9}) \quad , \quad x = 1,2,. . .$$

EJEMPLO 3 Los dos tercios de los niños de un colegio están ausentes por causa de una epidemia. En una clase de 25 estudiantes, el profesor pasa lista. Defina **X** como el número de estudiantes llamados hasta que uno responda.

(a) ¿Cuál es la probabilidad que el décimo niño llamado sea el primero que responda presente?

(b) Calcular $P[X \leq 2]$, $P[X \geq 2]$.

(c) Determine $E(X)$, y la desviación estandar.

SOLUCION 1 Definimos la variable aleatoria **X** como sigue

$X(\omega)$ = número de estudiantes llamados hasta que uno responda presente.

$$R_X = \{1,2,3,. . . , 24,25\}$$

$$p = P[E] = \frac{1}{3} \quad , \quad q = P[F] = \frac{2}{3}$$

2. Luego, **X** tiene una distribución geométrica. Osea

$$p(x) = P[X = x] = (\tfrac{2}{3})^{x-1}(\tfrac{1}{3}) \quad , \quad x = 1,2, . . . , 25.$$

(a) $$P[X = 10] = (\tfrac{2}{3})^9(\tfrac{1}{3}) = \frac{2^9}{3^{10}} .$$

(b) (*i*) $$P[X \leq 2] = P[X = 1] + P[X = 2] = \frac{1}{3} + (\tfrac{2}{3})(\tfrac{1}{3}) = \frac{5}{9} .$$

(*ii*) $$P[X \geq 2] = 1 - P[X < 2] = 1 - P[X = 1] = 1 - \frac{1}{3} = \frac{2}{3} .$$

(c) $$E(X) = \frac{1}{1/3} = 3 \ ; \quad \sigma^2 = \frac{2/3}{(1/3)^2} = 6, \text{ y } \sigma = \sqrt{6} .$$

EJEMPLO 4 Cierta experiencia se repite hasta obtener un resultado exitoso. Los ensayos son independientes y el costo de ejecutar un experimento es $ 25,000. Si el experimento falla el costo del ensayo siguiente es aumentado en $ 5,000 debido a ciertos reajustes en el equipo. Si la probabilidad de exito en cualquier ensayo es 0.25, ¿cuál es el costo esperado del proyecto? Suponga que el experimentador a dispuesto un máximo de $ 500,000. Hallar la probabilidad que los trabajos experimentales costaría más de este monto.

SOLUCION Sea **X** la variable aleatoria definida como sigue

$X(\omega)$ = número de ensayos requeridos para obtener un éxito

$$R_X = \{1,2,3, \ldots\}$$

Los ensayos son independientes y la probabilidad de éxito en cada caso es 0.25. Por lo tanto **X** tiene una distribución geométrica con parámetro $p = 0.25$ es decir ,

$$p(x) = (0.25)(0.75)^{x-1} \quad , \quad x = 1,2, \ldots$$

El costo **C** del proyecto, es función de **X** y escribe así ,

$$C(X) = 25,000\,X + 5,000(X-1) = 30,000X - 5,000$$

Entonces $E[C(X)] = 30,000\,E(X) - 5,000 = 30,000\left(\frac{1}{0.25}\right) - 5,000 = \$\,115,000.$

Por otro lado,

$$P[C(X) > 500,000] = P[30,000X - 5,000 > 500,000] = P\left[X > \frac{505,000}{30,000}\right]$$

$$= P[X > 16.83] = 1 - P[X \leq 16] = 1 - F(16)$$

$$= 1 - [1 - (0.75)^{16}] = (0.75)^{16}.$$

5.3.2 DISTRIBUCION BINOMIAL NEGATIVA

La distribución binomial negativa también se fundamenta en un proceso de Bernoulli; es una extensión de la distribución geométrica. En este caso la variable aleatoria **X** está definida así,

$X(\omega)$ = Número de veces que se repite el ensayo hasta obtener r éxitos.

$$R_X = \{r, r+1, r+2, \ldots\}$$

La variable aleatoria **X** definida en la forma anterior se llama ***variable aleatoria binomial negativa*** . La distribución de probabilidad de la variable aleatoria binomial negativa **X**, se llama ***distribución binomial negativa*** o de Pascal. El espacio muestral tiene la forma siguiente,

$$\Omega = \{\overbrace{EEE\ldots E}^{r\text{-veces}},\ \underbrace{\overbrace{EE\ldots EF}^{r-1\text{ veces }1\,1}}_{P_r^{r-1,1}}\underset{\uparrow\,\text{fijo}}{E},\ \underbrace{\overbrace{EE\ldots EFF}^{r-1\text{veces }2\,1}}_{P_{r+1}^{r-1,2}}\underset{\uparrow\,\text{fijo}}{E}, \ldots\}$$

$$p(r) = P[X = r] = P[\overbrace{EE\ldots E}^{r\text{-veces}}] = p^r$$

$$p(r+1) = P[X = r+1] = P[\underbrace{\overbrace{EE\ldots EF}^{r-1}}_{P_r^{r-1,1}}\underset{\uparrow\,\text{fijo}}{E}] = P_r^{r-1,1}p^r q = \binom{r}{r-1}p^r q$$

$$p(r+2) = \mathbf{P}[X = r+2] = \mathbf{P}[\underbrace{\overbrace{EE.\,..E}^{r-1\text{veces}}\overbrace{FF}^{2}}_{P_{r+1}^{r-1,2}}\underset{\substack{\uparrow\\ \text{fijo}}}{E}] = \binom{r+1}{r-1}p^r q^2 = \binom{r+2-1}{r-1}p^r q^2$$

En general ,

$$p(x) = \mathbf{P}[X = x] = \mathbf{P}[\underbrace{\overbrace{EE\,.\,.\,.\,E}^{r-1\text{ veces}}\overbrace{FF\,.\,.\,.\,F}^{x-r}}_{P_{x-1}^{r-1,x-r}}\underset{\substack{\uparrow\\ \text{fijo}}}{\overset{1}{E}}] = \binom{x-1}{r-1}p^r q^{x-r}$$

Luego, la distribución de probabilidad de la variable aleatoria, **X** es

$$p(x) = \mathbf{P}[X = x] = \binom{x-1}{r-1}p^r q^{x-r} \qquad , \qquad x = r, r+1, r+2, \ldots$$

La función de distribución acumulativa de **X** esta dado por

$$F(x) = \mathbf{P}[X \leq x] = \begin{cases} 0 & , \quad x < r \\ \displaystyle\sum_{k=0}^{[\![x]\!]}\binom{k-1}{r-1}p^r q^{k-r} & , \quad x \geq r \end{cases}$$

La media y la varianza de la variable aleatoria binomial negativa **X**, son

$$\boxed{\mu = \frac{r}{p}} \qquad y \qquad \boxed{\sigma^2 = \frac{rq}{p^2}}$$

EJEMPLO 5 Suponga que cada vez que una persona maneja un automóvil tiene 0.001 de probabilidad de recibir una infracción de tránsito por manejar con exceso de velocidad, y suponga también que se pierde la licencia de manejo al sumar tres infracciones. Sea **X** el número de veces que se maneja el automóvil hasta recibir la tercera infracción. Obtener la función de probabilidad para **X** (suponer que cada vez que se maneja el auto se tiene la misma probabilidad de 0.001 de recibir una infracción y que las veces que ocurren esas boletas son independientes)

SOLUCION 1 La variable aleatoria **X** esta definida por

X(ω) = número de veces que maneja hasta perder la licencia (tercera infracción).

$$R_{\mathbf{X}} = \{3,4,5,6,\ .\ .\ .\ ,\}$$

$$p = 0.001 , \qquad q = 0.999 , \qquad r = 3$$

2. **X** es una variable aleatoria binomial negativa, y su distribución de proba

bilidad es,

$$p(x) = P[X = x] = \binom{x-1}{2}(0.001)^3(0.999)^{x-3} \quad , \quad x = 3,4,5,\ldots$$

EJEMPLO 6 Cierta dieta con poco yodo produce un ensanchamiento de la glándula tiroides en un 60% de los animales de una población grande. Se necesitan 4 animales con glándulas tiroides ensanchada para un experimento. Defina la variable aleatoria X como el número de animales seleccionados hasta que 4 de ellos sean con glándula tiroides ensanchada.

(a) Hallar la función de probabilidad de la variable aleatoria X.

(b) Hallar la función de distribución de X.

(c) Determinar, $P[X = 7]$, $P[X > 5]$.

(d) Calcular: $E(X)$, σ_X^2 .

SOLUCION 1 La variable aleatoria X está definida por

$X(\omega)$ = número de animales seleccionados hasta que 4 de ellos sean con glándula tiroides ensanchada.

$$R_X = \{4,5,6,\ldots\}$$

$$p = 0.6 , \quad q = 0.4 , \quad y \quad r = 4$$

2. X es una variable aleatoria binomial negativa, por lo tanto su función de probabilidad es

(a) $$p(x) = P[X = x] = \binom{x-1}{3}(0.6)^4(0.4)^{x-4} \quad , \quad x = 4,5,\ldots$$

(b) $$F(x) = \sum_{x=4}^{[x]} \binom{r-1}{3}(0.6)^4(0.4)^{r-4} = (0.6)^4 \sum_{r=4}^{[x]} (0.4)^{r-4} ,$$ es función de distribución acumulativa de X.

(c) $$P[X = 7] = \binom{6}{3}(0.6)^4(0.4)^3$$

$$P[X > 5] = 1 - P[X \le 5] = 1 - \sum_{r=4}^{5} \binom{r-1}{3}(0.6)^4(0.4)^{r-4}$$

(d) $$\mu = E(X) = \frac{4}{0.6} = \frac{40}{6} \quad y \quad \sigma^2 = \frac{4(0.4)}{(0.6)^2} = \frac{40}{9} .$$

PROBLEMAS 5.3

1. Se lanza un dado hasta que aparezca el 4. ¿Cuál es la probabilidad que haya que lanzarlo más de 6 veces?

2. En una población grande, el 25% de las personas tienen ojos azules. Escogemos aleatoriamente voluntarios de esta población, una cada vez, hasta - escoger un voluntario con ojos azules. ¿Cuál es la probabilidad que - la quinta persona es la primera que tiene los ojos azules?. ¿Cuál es el - número esperado de personas escogidas?

3. Una máquina produce artículos de tipo A y B simultaneamente; el productor afirma que la máquina produce 2 artículos tipo B por cada 3 de tipo A. - Una persona somete a prueba y de un lote grande extrae estos artículos - hasta lograr un artículo tipo A.
   (a) Suponga que la persona decide comprar estos artículos si logra extraer un artículo A antes de la cuarta extracción,¿cuál es la probabilidad que la persona compre estos artículos?
   (b) Suponga que la persona compra si logra extraer un artículo A antes de la tercera extracción, ¿cuál es la probabilidad que no compre?

4. Dos jugadores juegan a lanzar un dado, gana el primero que obtiene el primer número primo. Si el juego es empezado por A. Calcular la probabilidad de ganar de cada jugador .

5. Una fábrica de helados fabrica paletas cubiertas de chocolate que se vende a 10 intis Suponga que se pone una estrella en cada 50 paletas; cualquiera que compra una paleta con una estrella obtiene otra paleta gratuita. Si se decide comprar paletas hasta obtener una paleta gratuita. ¿Cuántas se espera comprar antes de lograr una gratis?

6. Calcule el número esperado de lanzamientos con dos dados, hasta conseguir suma 7.

7. El señor dedos largos es un ladrón aficionado. La probabilidad que - abra una caja con joyas en un intento es 0.20. ¿Cuál es la probabilidad - de que :
   (a) necesita exactamente dos intentos para abrir la caja?
   (b) abra la caja en no más de tres intentos?

8. La probabilidad de un lanzamiento exitoso es igual a 0.8. Suponga que se hacen ensayos de lanzamientos hasta que han ocurrido 3 lanzamientos exito

sos. ¿Cuál es la probabilidad que sean necesarios menos de 6 intentos?

9. Las máquinas A y B producen un promedio, 5% y 10% de piezas defectuosas respectivamente. Supon que se extraen piezas de la producción de cada una hasta que se obtenga una pieza defectuosa. ¿Cuál es la probabilidad - que de la producción A tenga que extraerse exactamente 4 piezas y de la producción B exactamente 6 piezas?

10. En una población grande se sabe que el 10% padece de cierta enfermedad - rara. Con la finalidad de hacer un diagnóstico se necesitan 10 personas - afectadas por dicha enfermedad para el análisis correspondiente. Defina X como el número de personas seleccionadas hasta que se tengan las 10 personas afectadas por la enfermedad.
    (a) Hallar la distribución de probabilidad de la variable aleatoria X.
    (b) Calcular la probabilidad de que se necesiten seleccionar exactamente 15 personas.

11. En un laboratorio de Física trabajan 10 Físicos, de los cuales 6 son doctores y 4 Licenciados en física. Cada mes se elige, al azar, a uno de los Físicos, como encargado del almacen. Calcular la probabilidad que en el quinto mes, el almacen por tercera vez esté a cargo de un doctor.

12. En el problema 8. Supóngase que cada uno de los ensayos de lanzamientos - cuesta $ 5,000. Además, un lanzamiento que falla produce un costo adicional de $ 500. Calcular el costo esperado del número de lanzamientos hasta encontrar 3 lanzamientos exitosos.

5.4 DISTRIBUCION MULTINOMIAL

Consideremos un experimento ε con espacio muestral Ω con la siguiente caracteristica :

(a) Tiene k posibles resultados $E_1, E_2, \ldots, E_k$ mutuamente excluyentes y colectivamente exhaustivos, es decir $E_i \cap E_j = \phi$, $i \neq j$, $i,j = 1,2,\ldots, k$; $\bigcup_{i=1}^{k} E_i = \Omega$

(b) $P[E_i] = p_i$, $i = 1,2, \ldots , k$, tal que $\sum_{i=1}^{k} p_i = 1$.

Observe, si $k = 2$, el experimento ε es un ensayo de Bernoulli.

Consideremos ahora una secuencia de n ensayos independientes con la característica (a) y (b). Es decir cada ensayo tiene k resultados y p_i , $i = 1,2,\ldots,k$ es constante en cada ensayo. Definimos una variable aleatoria X_i como sigue,

$X_i(\omega)$ = número de veces que el evento E_i ocurre en las n repeticiones de ε, $i = 1,2,\ldots, k$

$$R_{X_i} = \{0,1,2,\ldots, n\} \qquad i = 1,2,\ldots,k\,.$$

La distribución de probabilidad conjunta de las variables aleatorias $X_1, X_2, \ldots, X_k$, [es decir la distribución de probabilidad del vector aleatorio (X_1, X_2, . . . , X_k)] se llama distribución multinomial y se define por,

$$p(x_1, x_2, \ldots, x_k) = P[X_1 = x_1, \ldots, X_k = x_k] = \frac{n!}{x_1!x_2!\ldots x_k!}\, p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

donde $\displaystyle\sum_{i=1}^{k} x_i = n$

Observe que $X_1, X_2, \ldots, X_k$ no son variables aleatorias independientes, puesto que $\displaystyle\sum_{i=1}^{k} x_i = n$. Es decir, conocido los valores de $k - 1$ variables aleatorias cualesquiera, se conoce la que falta.

La media y la varianza de las variables aleatorias X_i son

$$E(X_i) = np_i \quad ; \quad Var(X_i) = np_i(1 - p_i) \quad , \quad i = 1,2,\ldots,k$$

EJEMPLO 1 En cierta población grande el 70 % de las personas son derechos; el 20 % izquierdos y 10 % ambidiestros. Se escogen 10 personas aleatoriamente de la población. ¿Cuál es la probabilidad que :

(a) todos sean derechos?

(b) 7 sean derechos, 2 surdos y 1 ambidiestro?

SOLUCION 1 Definimos los siguientes eventos

E_1 : una persona derecha

E_2 : una persona surda,

E_3 : una persona ambidiestra,

2. Las probabilidades correspondientes para cada evento son :

$$p_1 = 0.7 \quad , \quad p_2 = 0.2 \quad y \quad p_3 = 0.1$$

3. El experimento ε es seleccionar al azar una persona de la población. Los posibles resultados de ε son los eventos mutuamente excluyentes E_1, E_2 ó E_3. Se repite el experimento 10 veces y las probabilidades $p_i (i = 1,2,3)$ permanecen constantes en los 10 ensayos. Luego utilizando la distribución multinomial con $n = 10$ tenemos que

$X_i(\omega)$ = número de veces que ocurre el evento $E_i (i = 1,2,3)$ en los 10 ensayos.

(a) $x_1 = 10$, $x_2 = 0$, $x_3 = 0$; $p(10,0,0) = \dfrac{10!}{10!0!0!}(0.7)^{10} = (0.7)^{10} \simeq 0.028.$

(b) $x_1 = 7$, $x_2 = 2$, $x_3 = 1$; $p(7,2,1) = \dfrac{10!}{7!2!1!}(0.7)^7(0.2)^2(0.1) \simeq 0.119$

EJEMPLO 2 Las probabilidades que una declaración de impuestos sea llenada correctamente es 0.60, que contenga un error que favorezca al declarante 0.20, que lleve un error que favorezca al fisco 0.10 y que contenga ambos tipos de errores 0.10. Se escoge al azar 10 de tales declaraciones de impuestos para una auditoría,¿cuál es la probabilidad que 5 estén correctas, 3 tengan un error que favorezca al declarante, una lleve un error que favorezca al fisco y una contenga ambos tipos de errores?

SOLUCION 1 Definimos los siguientes eventos :

E_1 : una declaración de impuestos llenada correctamente.

E_2 : una declaración que tenga un error que favorezca al declarante .

E_3 : una declaración que tenga un error que favorezca al fisco .

E_4 : una declaración que contenga ambos tipos de errores.

2. Las probabilidades de la ocurrencia de cada evento son :

$p_1 = P[E_1] = 0.60$, $p_2 = P[E_2] = 0.20$, $p_3 = P[E_3] = 0.10$ y $p_4 = P[E_4] = 0.10$

3. El experimento ε es seleccionar al azar una declaración de impuestos. Los posibles resultados de ε son los eventos mutuamente excluyentes E_1, E_2, E_3 o E_4. El experimento se repite 10 veces. Las probabilidades p_i (i = 1,2,3,4) permanecen aproximadamente constantes en los 10 ensayos

Aplicando la distribución multinomial con n = 10, tenemos que

$X_i(\omega)$ = número de veces que ocurre E_i (i = 1,2,3,4) en los 10 repeticiones de ε.

Es decir $x_1 = 5$, $x_2 = 3$, $x_3 = 1$, $x_4 = 1$

$$p(5,3,1,1) = \frac{10!}{5!\ 3!\ 1!\ 1!}(0.6)^5(0.2)^3(0.1)(0.1) \approx 0.0314 .$$

PROBLEMAS 5.4

1. Las probabilidades que al conducir cierto modelo específico de automóvil ensamblado en el país se obtenga en promedio menos de 14 km. por galón, entre 14 y 16 km. por galón o más de 16 km. por galón son respectivamente 0.40, 0.40 y 0.20. Se prueban 12 de tales automóviles, ¿cuál es la probabilidad que 4 en promedio menos de 14 km. por galón, 6 entre 14 y 16 km. por galón y 2 más de 16 km. por galón?

2. Cierta enfermedad de la tiroide se trata con una terapéutica de yodo. Se halla que el 50% de los pacientes tiene una mejora rápida, el 40% no sufre ningún efecto y el 10% empeora. Se trata a 9 pacientes con esta terapéutica. ¿Cuál es la probabilidad que :
   (a) los 9 mejoran ;
   (b) 5 mejoran, tres permanecen en el mismo estado y uno empeora;
   (c) tres mejoran, tres estan en el mismo estado y 3 empeoran.

3. El género humano puede clasificarse de acuerdo al tipo de sangre en cuatro categorías mutuamente excluyentes O , A, B y AB. En una población grande la proporción de los diferentes tipos de sangre son:
   0.45, 0.40, 0.10 y 0.05 respectivamente. Supóngase que se escogen al azar seis personas de la población. ¿Cuál es la probabilidad que:
   (a) tres son de tipo A y tres de tipo B?
   (b) ninguno son del tipo AB?.

4. En un depósito grande hay un gran número de arandelas. El 50% de esas arandelas son de 1/4 pulgadas de diámetro, el 30% son de 1/8 pulgada de diámetro y el 20% restante con 3/8 pulgadas de diámetro. Se eligen 10 -

arandelas aleatoriamente, ¿cuál es la probabilidad que:

(a) haya exactamente cinco de 1/4 pulgada, cuatro de 1/8 pulgada, y una de 3/8 pulgada?

(b) sólo haya dos clases de arandelas entre las elegidas?

(c) las tres clases de arandelas estén entre las elegidas?

(d) haya tres de una clase, tres de otra, y cuatro de la tercera clase?

5. Tres companías X, Y , Z tienen probabilidades de 0.4, 0.3 y 0.3 respectivamente de obtener una orden para un tipo particular de mercadería.- Las tres ordenes se adjudican independientemente. ¿Cuál es la probabilidad que una compañía reciba todas las ordenes?

5.5 DISTRIBUCION HIPERGEOMETRICA

Consideremos una población finita con N elementos, divididos en dos clases. Una con M elementos (M < N) y la otra con N-M elementos. Llamemos "éxito" a la primera clase y "fracaso" a la segunda. Es decir, la población se constituye por M exitos y N-M fracasos. Por ejemplo, un lote de N elementos de una producción, se divide en defectuosos con M elementos y no defectuosos con N - M elementos. Otro ejemplo es, un salón de clase con N estudiantes, de los cuales M tienen cierta enfermedad, y N - M no tienen dicha enfermedad.

Consideremos el siguiente experimento, "extraer una muestra de tamaño n *sin reemplazamiento de una población finita de* N elementos". Cada extracción tiene sólo dos resultados posibles E o F. El resultado de una observación es afectado por los resultados de las observaciones previas, es decir los resultados de los ensayos no son independientes y la probabilidad de éxito P[E] no es constante de ensayo a ensayo. El experimento, así definido se llama "*experimento hipergeométrico*" .

Definimos la variable aleatoria X de la siguiente manera,

$X(\omega)$ = número de éxitos en la muestra de tamaño n sin reemplazo

$$R_X = \{0,1,2, \ldots , \min(n,M)\}$$

La variable aleatoria definida de esta forma se llama *variable aleatoria hipergeométrica*. Y la distribución de probabilidad de la variable aleatoria hipergeométrica, se llama *distribución hipergeométrica*. Se denota por $h(x;N,n,M)$. El lector puede deducir fácilmente que la distribución de probabilidad de X está dado por

$$\boxed{p(x) = h(x;N,n,M) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \qquad x = 0,1,2,\ .\ .\ .\ ,\ \min(n,M).}$$

Note, la similitud de la definición de la variable aleatoria binomial y la hipergeométrica.

La media y la varianza de la variable aleatoria hipergeométrica X, se determinan como sigue :

$$\mu = E(X) = \sum_{x=0}^{n} x\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} = n\,\frac{M}{N}\sum_{x=1}^{n}\frac{\binom{M-1}{x-1}\binom{N-M}{n-x}}{\binom{N-1}{n-1}}\ , \quad \text{si} \quad n = \min(M,n)$$

$$= n\,\frac{M}{N}\sum_{y=0}^{n-1}\frac{\binom{M-1}{y}\binom{N-1-M+1}{N-1-y}}{\binom{N-1}{n-1}} = n\,\frac{M}{N}$$

por lo tanto

$$\boxed{E(X) = n\left[\frac{M}{N}\right]} \qquad (*)$$

usando la relación $\displaystyle\sum_{i=0}^{m}\binom{a}{i}\binom{b}{m-i} = \binom{a+b}{m}$.

$$E[X(X-1)] = \sum_{x=0}^{n} x(x-1)\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= n(n-1)\frac{M(M-1)}{N(N-1)}\sum_{x=2}^{n}\frac{\binom{M-2}{x-2}\binom{N-2-M+2}{n-x}}{\binom{N-2}{n-2}}$$

$$= n(n-1)\frac{M(M-1)}{N(N-1)}\sum_{y=0}^{n-2}\frac{\binom{M-2}{y}\binom{N-2-M+2}{n-2-y}}{\binom{N-2}{n-2}}$$

$$= n(n-1)\frac{M(M-1)}{N(N-1)}$$

La propiedad de la varianza se escribe

$$\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2 \\
&= n(n-1)\frac{M(M-1)}{N(N-1)} + n\frac{M}{N} - n^2\frac{M^2}{N^2} \\
&= n\frac{M}{N}\left[(n-1)\frac{M-1}{N-1} + 1 - \frac{nM}{N}\right] \\
&= n\frac{M}{N}\left[\frac{(N-M)(N-n)}{N(N-1)}\right] = n\left[\frac{M}{N}\right]\left[1-\frac{M}{N}\right]\left[\frac{N-n}{N-1}\right]
\end{aligned}$$

$$\boxed{Var(X) = n\left[\frac{M}{N}\right]\left[1-\frac{M}{N}\right]\left[\frac{N-n}{N-1}\right]} \qquad (**)$$

La función de distribución acumulativa de la variable aleatoria hipergeométrica es

$$F(x) = P[X \le x] = \begin{cases} 0 & , \quad x < 0 \\ \displaystyle\sum_{k=0}^{[x]} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} & , \quad 0 \le x < \min(M,n) \\ 1 & , \quad x \ge \min(M,n) \end{cases}$$

EJEMPLO 1 Una urna contiene 5 bolas blancas y 6 rojas. Se extrae 4 bolas de la urna sin reposición. Hallar la distribución de probabilidad del número de bolas rojas extraídas. ¿Cuál es la probabilidad de extraer exactamente 3 bolas rojas?¿Cuál es el número esperado de bolas rojas extraídas?

SOLUCION 1 Sea **X** la variable aleatoria definida por

$X(\omega)$ = número de bolas rojas extraídas de la urna.

$$R_X = \{0,1,2,3,4\}$$

2. $N = 6 + 5 = 11, \qquad M = 6, \qquad n = 4$

(a) Se cumple las condiciones de un experimento hipergeométrico. Luego, **X** es una variable aleatoria hipergeométrica, cuya distribución de probabilidad es,

$$p(x) = h(x;11,4,6) = \frac{\binom{6}{x}\binom{5}{4-x}}{\binom{11}{4}} \quad , \qquad x = 0,1,2,3,4.$$

(b) $$P[X = 3] = \frac{\binom{6}{3}\binom{5}{1}}{\binom{11}{4}} .$$

(c) $$\mu = 4\left(\frac{6}{11}\right) = \frac{24}{11} .$$

EJEMPLO 2 En cierta clínica hay 20 enfermos de los cuales se sabe que el 30% tienen cáncer, se extrae aleatoriamente 4 pacientes para el despistaje de cáncer.

(a) ¿Cuál es la probabilidad que al menos uno tenga cáncer?

(b) ¿Cuál es el número esperado de pacientes con cáncer?

SOLUCION 1. Definimos la variable aleatoria X tal que

$X(\omega)$ = número de enfermos con cáncer en la muestra de 4 pacientes.

$$R_X = \{0,1,2,3,4\}$$

2. $N = 20$, $M = 6$, $N - M = 14$, $n = 4$

3. por lo tanto la función de probabilidad es,

$$p(x) = \frac{\binom{6}{x}\binom{14}{4-x}}{\binom{20}{4}} , \qquad x = 0,1,2,3,4 .$$

(a) $$P[X \geq 1] = 1 - P[X < 1] = 1 - P[X = 0] = 1 - \frac{\binom{14}{4}}{\binom{20}{4}} .$$

(b) $$\mu = 4\,\frac{6}{20} = \frac{6}{5} .$$

5.5.1 APROXIMACION DE LA HIPERGEOMETRICA A LA BINOMIAL

Si el tamaño n de la muestra sin reemplazo es pequeña con relación a N, la probabilidad de cada extracción varía muy levemente. En la práctica cuando n es menor que el 10% de N, (es decir, $\frac{n}{N} < 0.1$) se aproxima la distribución hipergeométrica a la distribución binomial con $p = \frac{M}{N}$ y n . Entonces

$$h(x;N,n,M) = b\left(x;n,\frac{M}{N}\right) = \binom{n}{x}\left(\frac{M}{N}\right)^{x}\left(1 - \frac{M}{N}\right)^{n-x} \qquad x = 0,1,\ldots,\min(n,M).$$

La media y la varianza se aproximan por,

$$\mu = np = n\left[\frac{M}{N}\right] \qquad (***)$$

$$\sigma^2 = npq = n\left[\frac{M}{N}\right]\left[1 - \frac{M}{N}\right]$$

Comparando las fórmulas de (***) con las de (*) y (**) vemos que la media es la misma, mientras que la varianza difiere por un factor de corrección $\frac{N-n}{N-1}$. Este es despreciable cuando n es pequeño con relación a N.

NOTA. Si se requiere mejor exactitud en la aproximación, se usa $n \leq 0.05N$.

EJEMPLO 3 Un auditor del Departamento de impuesto sobre la renta está seleccionando una muestra de seis declaraciones de impuestos de personas de una profesión particular, para una posible auditoría. Si dos o más de ellas indican deducciones "no autorizadas", se auditará todo el grupo (población) de 100 declaraciones. Si el 25% de las declaraciones es incorrecta, determinar:

(a) La verdadera distribución de probabilidad del número de declaraciones incorrectas en la muestra. ¿Cuales son los parámetros?. Halle la probabilidad de una auditoría más detallada.

(b) Utilice una aproximación a la verdadera distribución de probabilidad para hallar la probabilidad de una auditoría más detallada.

SOLUCION 1 Definimos la variable aleatoria **X** como sigue:

X(ω) = número de declaraciones incorrectas en la muestra de 6.

$$R_X = \{0,1,2,3,4,5,6\}$$

2. N = 100, M = 25, n = 6, N - M = 75

3. El muestreo es sin reemplazamiento.

(a) Se cumplen las condiciones de un experimento hipergeométrico. Luego **X** es una variable aleatoria hipergeométrica. Por lo tanto la verdadera distribución de probabilidad es la hipergeométrica,

$$h(x;100,6,25) = \frac{\binom{25}{x}\binom{75}{6-x}}{\binom{100}{6}}, \qquad x = 0,1,2,3,4,5,6.$$

los parámetros son :

$$E(X) = 6 \cdot \frac{25}{100} = \frac{3}{2} \quad y \quad Var(X) = 6 \cdot \frac{25}{100}\left[1 - \frac{25}{100}\right]\left[\frac{94}{99}\right] = \frac{47}{44}.$$

Se hará una auditoría más detallada, si X toma valores mayores o iguales que 2 - Es decir

$$P[X \geq 2] = 1 - P[X \leq 1] = 1 - [P[X = 0] + P[X = 1]]$$

$$= 1 - \frac{\binom{75}{6} + \binom{25}{1}\binom{75}{5}}{\binom{100}{6}} = 0.4691$$

(b) $n/N = \frac{6}{100} = 0.06$ es pequeño ($\frac{6}{100} < 0.1$), aproximamos la distribución hipergeométrica a la binomial con $p = \frac{25}{100} = \frac{1}{4}$. Es decir,

$$h(x;100,6,25) \approx b(x;6,\tfrac{1}{4}) = \binom{6}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{6-x}, \quad x = 0,1,\ldots,6.$$

Luego,

$$P[X \geq 2 \mid 6,0.25] = 0.4661 . \quad \text{(Tabla I)} .$$

5.5.2 EXTENSION DE LA DISTRIBUCION HIPERGEOMETRICA

Sea una población finita de N objetos que contiene M_1 objetos de primer tipo, M_2 objetos de segundo tipo, . . ., y M_k objetos del k-ésimo tipo, de tal manera que $M_1 + M_2 + \ldots + M_k = N$. Consideremos, el experimento aleatorio de "extraer una muestra sin reposición de tamaño n de la población". Entonces

(a) Cada extracción tiene k posibles resultados.

(b) El resultado de cada extracción es afectado por los resultados de las extracciones previas. Es decir, los resultados de los ensayos no son independientes.

Definimos la variable aleatoria X_i ($i = 1,2, \ldots , k$) de la siguiente manera

$$X_i = \text{número de objetos del } i\text{-ésimo tipo}$$

$$R_{X_i} = \{0,1,2, \ldots , \min (M_i, n)\}$$

Ahora nos interesa la probabilidad que en la muestra se obtenga x_1 objetos del primer tipo, x_2 objetos del segundo tipo, . . ., y x_k objetos del k-ésimo tipo. Es decir, la probabilidad conjunta de las variables aleatorias - $X_1, X_2, \ldots , X_k$, representaremos esta probabilidad por

$$p(x_1, x_2, \ldots , n_k;N, n) = P[X_1 = x_1, X_2 = x_2,\ldots, X_k = x_k;N,n]$$

1. El número de muestras posibles de tamaño n que pueden formarse con los N - objetos es $\binom{N}{n}$.

2. El número de formas de selección x_1 objetos de los M_1 del primer tipo es $\binom{M_1}{x_1}$, para cada una de éstas se puede tomar x_2 objetos de los M_2 del segundo tipo de $\binom{M_2}{x_2}$ formas. Por lo tanto, el número de formas de escoger x_1 objetos del primer tipo, x_2 objetos del segundo tipo es $\binom{M_1}{x_1}\binom{M_2}{x_2}$. Continuando de ésta forma, podemos escoger x_1 objetos del primer tipo, x_2 objetos del segundo tipo, ... , y x_k objetos del k-ésimo tipo de $\binom{M_1}{x_1}\binom{M_2}{x_2}\cdots\binom{M_k}{x_k}$ formas.

3. De (1) y (2) la probabilidad requerida esta definida por

$$p(x_1,x_2,\ldots,x_k;N,n) = \frac{\binom{M_1}{x_1}\binom{M_2}{x_2}\cdots\binom{M_k}{x_k}}{\binom{N}{n}}$$

EJEMPLO 4 En un depósito hay 20 arandelas de las cuales 10 son de 1/4 pulgada de diámetro, 6 de 1/8 pulgada de diámetro y los 4 restantes con 3/8 pulgadas de diámetro. Se eligen al azar 10 arandelas, ¿cuál es la probabilidad de que haya 5 de 1/4 pulgada de diámetro, 3 de 1/8 pulgada y 2 de 3/8 pulgada de diámetro?

SOLUCION En este caso $N = 20$, $M_1 = 10$, $M_2 = 6$ y $M_3 = 4$, $x_1 = 5$, $x_2 = 3$ y $x_3 = 2$. Entonces

$$p(5,3,2;20,10) = \frac{\binom{10}{5}\binom{6}{3}\binom{4}{2}}{\binom{20}{10}} = \frac{7560}{46189}.$$

PROBLEMAS 5.5

1. Se extraen al azar trece cartas sin reemplazo de una baraja de 52. ¿Cuál es la función de probabilidad para el número de cartas rojas en la muestra? ¿Cuál es la media y la varianza del número de naipes rojos?.

2. Una compañía quiere evaluar sus procedimientos de inspección en embarques de 50 artículos idénticos. El procedimiento consiste en tomar una muestra de 5 y aceptar el embarque si no se encuentran más de dos defectuosos. ¿Qué proporción de embarques con un 20% de artículos defectuosos será aceptada?

3. Una caja contiene 10 focos, de los cuales 8 están en buen estado, si se escogen al azar 5 focos. ¿Cuál es la función de probabilidad para los focos buenos?¿Para el número de focos que no sirven?
4. Un jurado de 7 jueces va a decidir entre dos finalistas quién es la ganadora del concurso de belleza, para lo cual bastará una mayoría simple de los jueces. Suponga que 4 jueces botan por María y que los otros 3 boten por Susana. Si se seleccionan al azar 3 jueces y se les pregunta por ouién van a votar. ¿Cuál es la probabilidad que la mayoría de los jueces de la muestra están a favor de María?
5. Una empresa manufacturera recibe un lote que contiene 100 artículos de los cuales cinco son defectuosos. La compañía revisa constantemente los lotes que recibe para establecer la calidad del material. Si la calidad de un lote recibido es baja, regresa al proveedor el lote completo. Suponga que la compañía recibe el lote y lo acepta si hay sólo 1 ó menos piezas defectuosas en una muestra de tamaño 6 sin reposición. ¿Cuál es la probabilidad de que se acepte un lote de 100 artículos que contenga 5 defectuosos?
6. En un salón universitario de 50 alumnos existen 3 corrientes políticas. Se sabe que el 50% pertenece al grupo A; el 30% al grupo B y el resto al C. Si se elige, al azar, un comité estudiantil formado por 10 alumnos, determine ud. la probabilidad de que 6 de ellos sean del grupo B.
7. Una fábrica emplea un patrón de aceptación en los artículos producidos antes de embarcárlos. El plan consta de dos etapas. Cajas con 25 artículos son preparadas para su embarque, tomándose 3 muestras para su revisión: Si se encuentra alguno defectuoso, la caja se regresa para inspeccionar el 100% de su contenido. En caso contrario, se le embarca. ¿Cuál es la probabilidad de:
   (a) embarcar una caja que contenga tres artículos defectuosos?
   (b) regresar una caja que contiene sólo un artículo defectuoso?
8. En un edificio hay 30 departamentos. De ellos, 10 están habitados por matrimonios con sólo dos hijos hombres. Los otros 20 están habitados por matrimonios con sólo dos hijas. Se va a demoler el edificio para construir uno de 60 departamentos, para lo cual se procede a desalojar parcialmente el edificio sorteando mensualmente una familia, la que debe retirarse. ¿Cuál es la probabilidad que al cabo de año y medio quede en el edificio el mismo número de mujeres y de hombres?

9. Un lote de 100 tubos de televisión a color está sujeto a un procedimiento de prueba de aceptación. El procedimineto consiste en extraer cinco tubos aleatoriamente, sin reemplazamiento, y probarlos. Si dos o menos tubos fallan, se acepta el lote. En caso contrario se rechaza el lote. Asumiendo que el lote contiene cuatro tubos defectuosos. Determinar:
   (a) la distribución de probabilidad del número de tubos defectuosos en la muestra. ¿Cuales son los parámetros?. Halle la probabilidad exacta de aceptar el lote.
   (b) Utilice una aproximación a la verdadera distribución de probabilidad para hallar la probabilidad de aceptar el lote. Compare las respuestas

10. Resolver el problema 31 de 5.2, considerando la extracción de la muestra sin reposición.

11. Una compañía recibe productos en cajas de 100. El control de calidad de los productos se efectúa en la forma siguiente: Se toma una muestra de 3 productos de una caja sin reposición. Si se encuentra a lo más un defectuoso, se acepta la caja; si se encuentra 2 ó 3 defectuosos, se elige una de 3 productos adicionales. Si en total (6) de productos elegidos hay 4 ó más defectuosos se rechaza la caja elegida, en caso contrario se acepta. Calcular la probabilidad de:
   (a) aceptar una caja con 5 productos refectuosos;
   (b) rechazar una caja con 4 defectuosos.

12. El cuerpo secretarial de un importante bufete de abogados cuenta con 25 secretarias, 10 de las cuales han estado con la compañía más de 5 años. Un ejecutivo desea seleccionar al azar cuatro secretarias para asignarlas un nuevo asunto.
   (a) ¿Cuál es la distribución de probabilidad del número de secretarias con más de 5 años en la compañía?
   (b) ¿Qué modelo discreto representa esto?
   ¿Cuál es la probabilidad que
   (c) ninguna de las secretarias tendrá más de 5 años en la compañía?
   (d) las cuatro secretarias tendrá más de 5 años en la compañía?
   (e) Calcule la media y la desviación estándar del número de secretarias con más de 5 años en la compañía, usando las fórmulas peculiares de este modelo discreto?

13. Considere un sistema eléctrico que consiste de 6 focos, conectados en serie de manera que ninguno prenderá si uno de ellos es defectuoso. Si los focos

de la instalación se selecciona al azar de un lote de 100 focos, 20 de los cuales son defectuosos. ¿Cuál es la probabilidad que funcionen todos los focos del sistema eléctrico?

14. En un estudio biológico se emplea un grupo de 15 personas. El grupo contiene 5 personas con sangre tipo O, 6 con tipo A y 4 con tipo B. ¿Cuál es la probabilidad que una muestra al azar de 6 contenga 2 personas con sangre tipo O , 3 con sangre tipo A y 1 con tipo B?

5.6 DISTRIBUCION DE POISSON

La distribución de Poisson es una de las distribuciones discretas más importantes por que se aplica en muchos problemas prácticos. La distribución de Poisson puede deducirse de dos formas. La primera se deduce a partir de un *proceso de Poisson*. La segunda se muestra la distribución de Poisson como un *límite de la distribución binomial*.

Daremos antes la idea intuitiva de un proceso de Poisson. Muchos problemas consisten en observar la ocurrencia de eventos discretos en un intervalo contínuo (unidad de medida). Por ejemplo el número de manchas (fallas) por unidad de medida (digamos 1 metro cuadrado) en el esmaltado de un refrigerador, se puede encontrar: 0 manchas, 1 mancha, 2 manchas, o tal vez más, en un metro cuadrado. Es decir podemos contar el número de fallas por unidad de medida, pero es imposible contar el número de puntos sin manchas (es infinito no numerable). Además las fallas son eventos discretos, debido a que su ocurrencia son puntos aislados en el área de 1 m^2. Si definimos la variable aleatoria X tal que $X(\omega)$ = número de manchas en un metro cuadrado. Su rango es $R_X = \{0,1,2, \ldots\}$. Otro ejemplo, es contar el número de vehículos que llegan a una estación de servicio durante un intervalo de tiempo (digamos una hora) de un día determinado; pueden llegar: 0 vehículos, 1 vehículo, 2 vehículos o más, es decir podemos contar el número de vehículos que llegan. Pero no tiene sentido contar los que no llegan. Es discreto debido a que la hora de llegada es un punto en el período contínuo de 1 hora. Aquí también, si la variable aleatoria, X se define por $X(\omega)$ = número de vehículos que llegan a la estación en una hora; los valores que puede tomar X es infinito numerable. Es decir, $R_X = \{0,1,2, \ldots\}$. Otro ejemplo, puede ser contar el número de llamadas que llegan a un tablero conmutador telefónico de una compañía grande en un in tervalo de tiempo (digamos de 8 a.m. a 10 a.m.) aquí también las llamadas que llegan al tablero es un evento discreto, ya que el tiempo de llegada de cualquiera de ellas es un punto aislado en el periodo de 2 horas. Otro caso, puede ser contar el número de bacterias en un cm^3 de agua. En este caso el intervalo contínuo es el cm^3 de agua y los eventos discretos, el número de bacterias, suponiendo que se puede considerar cada bacteria como un punto en el espacio.

Diremos que los eventos discretos que se generan en un intervalo contínuo (unidad de medida: longitud, área, volumen etc.) forman un proceso *de Poisson con parámetro* λ, si tiene las siguientes propiedades:

(1) El número promedio de ocurrencias de eventos en una unidad de medida (intervalo de tiempo, una región especificada, volumen, etc.) es conocido e igual a λ.

(2) La ocurrencia de un evento en una unidad de medida h no afecta a la ocurrencia, o no ocurrencia de otra unidad de medida h contigua. (Es decir, la ocurrencia de los eventos en unidades de medida contiguas son independientes).

(3) Sea una unidad de medida suficientemente pequeño de longitud h, luego:

(i) la probabilidad de un éxito en esta unidad pequeña es proporcional a la longitud del intervalo, esto es λh.

(ii) La probabilidad de la ocurrencia de 2 ó más éxitos en esta unidad pequeña es aproximadamente 0.

En un proceso de Poisson de parámetro λ se observa t unidades de medida, definimos X de la siguiente manera;

$X(\omega)$ = número de ocurrencia de eventos en t unidades de medida.

$$R_X = \{0,1,2,3,\ .\ .\ .\ \}$$

La variable aleatoria así definida se llama *variable aleatoria de Poisson* y la distribución de probabilidad de X se llama *distribución de Poisson* la cual está definida por,

$$\boxed{p(x) = P[X = x|P:\lambda t\,] = \frac{e^{-\lambda t}\,(\lambda t)^x}{x!}\ ,\qquad x = 0,1,2,\ .\ .\ .}$$

λt es el número promedio de ocurrencias de los eventos en las t unidades de medida.

para t fijo, ponemos $\alpha = \lambda t$ y obtenemos una expresión más simple para la distribución de Poisson

$$p(x) = P[X = x|P:\alpha] = \frac{\alpha^x e^{-\alpha}}{x!} \quad , \qquad x = 0,1,2\ldots$$

La función de distribución de X está dado por,

$$F(x) = P[X \le x|P:\alpha] = \begin{cases} 0 & , \quad x < 0 \\ \displaystyle\sum_{k=0}^{[\![x]\!]} \frac{\alpha^k}{k!} e^{-\alpha} & , \quad x \ge 0 \end{cases}$$

La *media y la varianza* de la distribución de Poisson estan dados por,

$$E(X) = \sum_{x=0}^{\infty} xP[X = x] = \sum_{x=0}^{\infty} x \frac{\alpha^x}{x!} e^{-\alpha} = e^{-\alpha} \sum_{x=1}^{\infty} \frac{\alpha^x}{(x-1)!}$$

$$= \alpha e^{-\alpha} \sum_{x-1=0}^{\infty} \frac{\alpha^{x-1}}{(x-1)!}$$

Haciendo $k = x - 1$, entonces $x = k + 1$, luego

$$E(X) = \alpha e^{-\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} = \alpha e^{-\alpha} e^{\alpha} = \alpha$$

$$\boxed{\mu = \alpha}$$

Similarmente,

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{\alpha^x e^{-\alpha}}{x!} = \sum_{x=0}^{\infty} \frac{x(x-1)\alpha^x e^{-\alpha}}{x!} + \sum_{x=0}^{\infty} \frac{x\alpha^x e^{-\alpha}}{x!}$$

$$= e^{-\alpha} \alpha^2 \sum_{x-2=0}^{\infty} \frac{\alpha^{x-}}{(x-2)!} + \alpha = \alpha^2 e^{-\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} + \alpha$$

$$\alpha^2 e^{-\alpha} e^{+\alpha} + \alpha = \alpha^2 + \alpha$$

Luego,

$$\sigma^2 = E(X^2) - (E(X))^2$$

$$= \alpha^2 + \alpha - \alpha^2$$

$$\boxed{\sigma^2 = \alpha}$$

5.6.1 TABLA DE LA DISTRIBUCION DE POISSON

Desde que, el uso de la ecuación $P[X = x|P;\alpha] = \alpha^x e^{-\alpha}/x!$ es muy tediosa, usaremos la función de distribución acumulativa $P[X \leq x|P;\alpha]$ para determinar probabilidades de cualquier tipo. La tabla II, da probabilidades acumulativas del tipo $P[X \geq x|P;\alpha]$. Mostraremos con un ejemplo el uso de esta tabla.

EJEMPLO 1 Las personas llegan aleatoriamente a la ventanilla de un banco en promedio a una razón de 24 por hora durante el periodo de tiempo entre 11.30 AM y 12.00AM de cierto día. Cuál es la probabilidad que exactamente 5 personas lleguen durante un periodo de tiempo de 12 minutos? ¿Cuál que lleguen por lo menos 10 personas? ¿ que lleguen a lo mas 12?

SOLUCION 1 $X(\omega)$ = número de personas que llegan durante un periodo de 12 minutos

$$R_X = \{0,1,2,\ldots\}$$

$$\alpha = (24)(12)/60 = 4.8$$

2. X es una variable aleatoria de Poisson y su distribución es

$$P[X = x|P:4.8] = \frac{(4.8)^x e^{-4.8}}{x!}, \quad x = 0, 1, 2, \ldots$$

3. La figura 6.6.1 muestra las probabilidades acumulativas requeridas para determinar la probabilidad pedida $P[X = 5]$

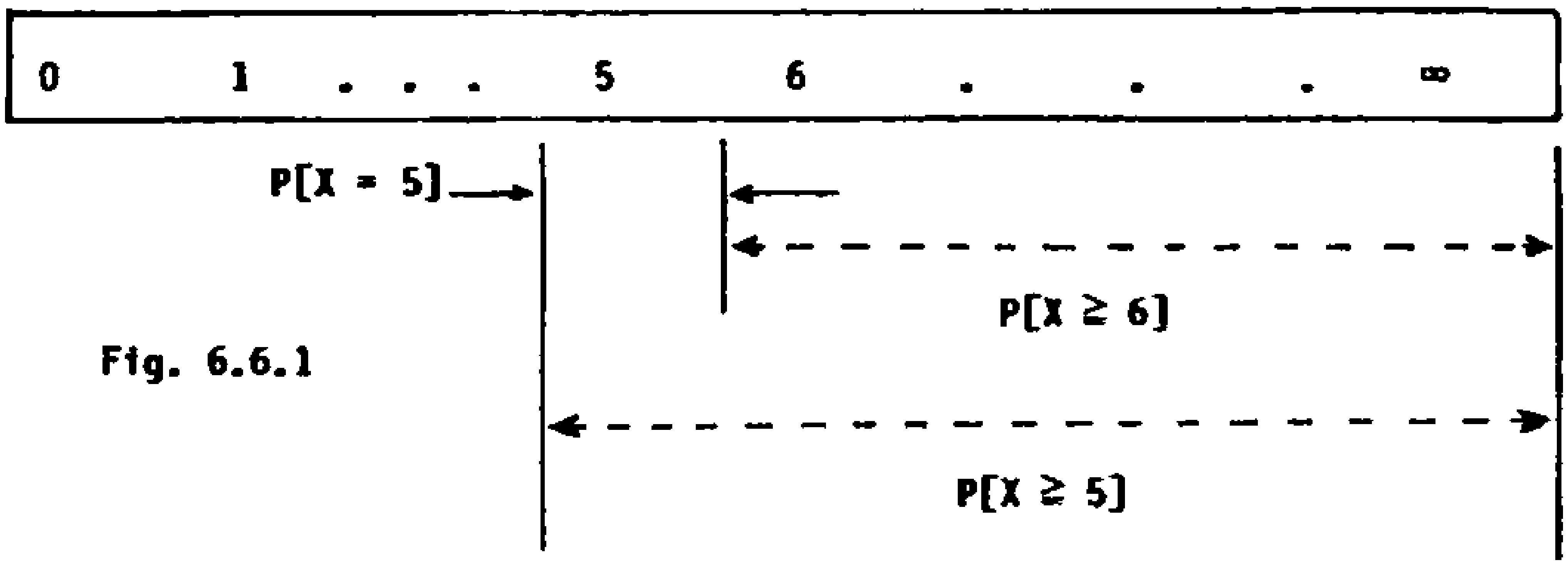


Fig. 6.6.1

$$P[X = 5|P:4.8] = P[X \geq 5|P:4.8] - P[X \geq 6|P:4.8]$$

$$= 0.5237 - 0.3490 = 0.1747\,.$$

Hay una probabilidad de 0.1747 de que exactamente 5 personas lleguen durante cualquier periodo de tiempo de 12 minutos.

EJEMPLO 2 El tablero conmutador de cierta universidad, indica un promedio de 2 llamadas cada 3 minutos. Asumiendo un proceso de poisson. ¿Cuál es la probabilidad que ocurran 5 ó más llamadas en un período de 9 minutos? ¿Cuál de que no haya llamadas en dicho período?

SOLUCION 1 Sea X la variable aleatoria definida por

$X(\omega)$ = número de llamadas en un período de 9 minutos

2. Entonces, $\alpha = E(X) = 6$ (2 llamadas por cada 3 minutos). Es decir $\lambda = 2$, $t = 3$ unidades de medida.

3. Luego, la distribución de probabilidad de X es ,

$$P[X = x] = \frac{6^x e^{-6}}{x!} , \qquad x = 0,1,2, \dots$$

(a) La figura 5.6.2 muestra la probabilidad acumulativa pedida, directamente leida en la tabla II.

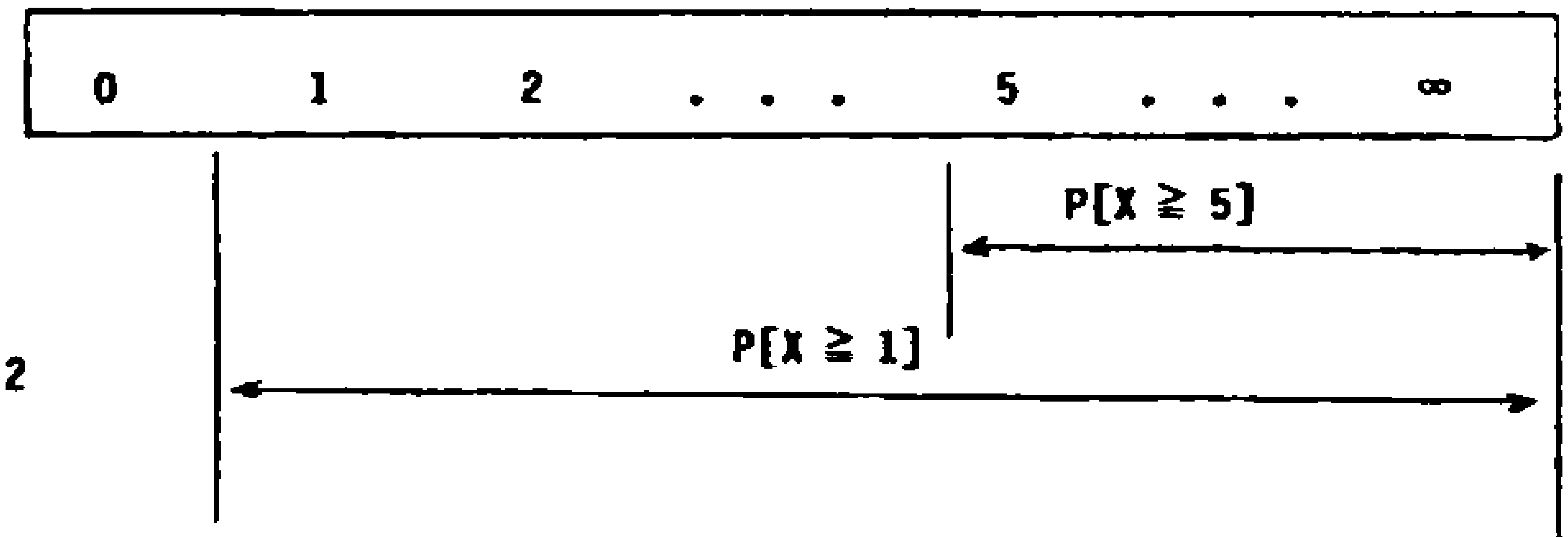


Fig. 5.6.2

$P[X \geq 5 | P;6] = 0.7149.$

(b) $P[X = 0 | P:6] = e^{-6} = 1 - P[X \geq 1 | P:6] = 1.09975 = 0.0025$.

EJEMPLO 3 Se ha observado que las cajas de cerveza pilsen se toman de los estantes de cierto supermercado a razón de 10 cajas por hora durante el período de mayor venta. ¿Cuál es la probabilidad que se saque al menos una caja durante los primeros 6 minutos de un período de mayor venta? ¿Cuál es la probabilidad que se tome del estante al menos una caja durante cada uno de 3 intervalos consecutivos de 6 minutos?

SOLUCION 1 Puesto que en 1 hora = 60 minutos, se toman un promedio de 10 cajas. Entonces en un período de 6 minutos se tomará un promedio de 1 caja - (1 cada 6 minutos, en 60 minutos 10 cajas).

2. Definimos la variable aleatoria X de la siguiente manera

$X(\omega)$ = número de caja tomadas en períodos de 6 minutos, $\alpha = \lambda t = 1$

3. Luego, la distribución de probabilidad de X es

$$P[X = x] = \frac{e^{-1}}{x!} \quad , \qquad x = 0,1,2,\ldots$$

(a) $P[X \geq 1] \approx 0.632$.

(b) Podemos considerar 3 variables aleatorias X_1, X_2 y X_3 cada una de ellas definidas como la variable aleatoria X e independientes. Luego, debemos calcular

$$P[X_1 \geq 1,\ X_2 \geq 1,\ X_3 \geq 1] = P[X_1 \geq 1]\,P[X_2 \geq 1]\,P[X_3 \geq 1]$$
$$= (0.632)^3 = 0.25244 \ .$$

EJEMPLO 4 Se sabe que un líquido particular contiene ciertas bacterias a razón de 4 bacterias por cm^3. Se toma una muestra de 1 cm^3, ¿cuál es la probabilidad que la muestra no contenga bacteria alguna? ¿cuál es la probabilidad que en 1/2 cm^3 haya por lo menos una bacteria?

SOLUCION (a) $X(\omega)$ = número de bacterias en 1 cm^3

$$R_X = \{0,1,2,\ldots\}$$

Aquí, $\lambda = 4$, $t = 1$. Entonces la distribución de X es

$$P[X = x] = \frac{4^x e^{-4}}{x!} \quad , \qquad x = 0,1,2,\ldots$$

$$P[X = 0] = e^{-4} = 0.0183 \ .$$

(b) $X(\omega)$ = número de bacterias en $\frac{1}{2}$ cm^3

$$R_X = \{0,1,2,\ldots\}$$

En este caso $\lambda = 4$, $t = \frac{1}{2}$, $\alpha = 2$. La distribución de X es

$$P[X = x] = \frac{2^x e^{-2}}{x!} \quad , \qquad x = 0,1,2,\ldots$$

$$P[X \geq 1] = 1 - P[X = 0] = 1 - e^{-2} = 0.864.$$

EJEMPLO 5 La siguiente distribución estadística da la distribución del número de días (n_i) sin accidente, con un accidente, ..., con 4 accidentes, para un periodo $n = 50$ días en una ciudad

Nº de accidentes x_i	0	1	2	3	4
Nº de días n_i	21	18	7	3	1

(a) ¿La distribución de accidentes puede considerarse como una distribución de Poisson?

(b) Si la respuesta en (a) es si, compare los valores observados (n_i) con el número teórico de días (número calculado mediante la distribución de Poisson).

SOLUCION Si para la distribución dada, $E(X) \approx Var(X) = \lambda$, entonces para resolver la cuestión del caracter de la distribución se sustituye el valor de λ en la fórmula de la distribución de Poisson y se calculan las probabilidades para $x = 0,1,2,\ldots, n$. En el caso en que los valores de estas probabilidades resultan próximas a los respectivos valores $\frac{n_i}{n}$, se puede considerar que la variable aleatoria se distribuye por una ley de Poisson.

1. Calculemos la distribución de probabilidad empírica de la variable

x	0	1	2	3	4
$p(x) = \frac{n_i}{n}$	$\frac{21}{50} = 0.42$	$\frac{18}{50} = 0.36$	$\frac{7}{50} = 0.14$	$\frac{3}{50} = 0.06$	$\frac{1}{50} = 0.02$

2. Cálculo de la esperanza matemática de la variable

$$E(X) = 0(\frac{21}{50}) + 1(\frac{18}{50}) + 2(\frac{7}{50}) + 3(\frac{3}{50}) + 4(\frac{1}{50}) = 0.9$$

3. Cálculo de la varianza de la variable

$$E(X^2) = 0^2(\frac{21}{50}) + 1^2(\frac{18}{50}) + 2^2(\frac{7}{50}) + 3^2(\frac{3}{50}) + 4^2(\frac{1}{50}) = \frac{89}{50}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{89}{50} - (\frac{45}{50})^2 = 0.97$$

4. De (2) y (3) puesto que $E(X) \approx Var(X)$, hacemos $\lambda = 0.9$, entonces

$$p(x) = P[X = x] = \frac{(0.9)^x e^{-0.9}}{x!} \quad , \quad x = 0,1,\ldots$$

5. Determinando por la fórmula (4) las $P[X = x]$ para, $x = 0,1,\ldots, 4$, obtenemos la tabla

x	0	1	2	3	4
P[X = x]	0.41	0.37	0.16	0.05	0.01

(a) De (2), (3) y (4). La variable tiene una distribución de Poisson

(b) El número teórico de días se calcula por la fórmula

$$n\,P[X = x] = 50\,P[X = x]$$

y se obtiene la siguiente tabla

x	0	1	2	3	4
frecuencia teórica f_i	20	18	8	2	1
frecuencia observada n_i	21	18	7	3	1

5.6.2 DISTRIBUCION DE POISSON COMO APROXIMACION DE LA BINOMIAL

Mostraremos ahora la distribución de Poisson como un límite de la distribución binomial, con $\lambda = np$. Aquí p debe ser suficientemente pequeño y n grande, de tal manera que np permanece casi constante. La distribución binomial para x éxitos en n ensayos de Bernoulli es

$$P[X = x|B;n,p] = \binom{n}{x}p^x q^{n-x} \quad , \qquad x = 0,1,2,\ldots,n$$

Hacemos $\lambda = np$, luego: $p = \frac{\lambda}{n}$ y $q = 1 - p = 1 - \frac{\lambda}{n}$

Reemplazando en la fórmula se tiene :

$$P[X = x] = \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n!}{(n-x)!n^x}\cdot\frac{\lambda^x}{x!}\cdot\frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-(x-1))(n-x)!}{n^x\quad(n-x)!}\cdot\frac{\lambda^x}{x!}\cdot\frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^x}$$

$$= \frac{n}{n}\,\frac{n-1}{n}\,\frac{n-2}{n}\,\cdots\,\frac{n-x+1}{n}\cdot\frac{\lambda^x}{x!}\cdot\frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^x}$$

$$= (1-\frac{1}{n})(1-\frac{2}{n})\ \ldots\ (1-\frac{x+1}{n})\,\frac{\lambda^x}{x!}\,\frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^x}$$

(1) Si n es suficientemente grande con respecto a x, entonces,

$$\frac{x}{n} \to 0, \text{ o lo que es lo mismo } \frac{n-x}{n} \to 1$$

(2) $p = \frac{\lambda}{n}$ es pequeño, entonces $(1-\frac{\lambda}{n})^x \to 1$

(3) Sabemos que, $(1-\frac{\lambda}{n})^n = e^{-\lambda}$, cuando $n \to \infty$

de (1), (2) y (3), aproximamos la $P[X = x|B:n,p]$ por

$\frac{\lambda^n}{x!}\,e^{-\lambda}$. Es decir,

$$P[X = x|B:n,p] \simeq \frac{\lambda^x}{x!}\,e^{-\lambda}, \qquad x = 0,1,2,\ \ldots$$

NOTA Teniendo en cuenta cómo se ha obtenido, resulta que las probabilidades de Poisson $P[X = x|P:np]$ da un valor aproximado de las probabilidades binomiales $P[X = x|B:n,p]$, para valores pequeños de p y n grande. Esta aproximación mejora a medida que p se acerca a cero y el valor de n se hace más grande. En la práctica se considera que la aproximación es aceptable si $p < 0.1$ y $np \leq 5$. Algunos autores consideran la aproximación aceptable cuando $p \leq 0.05$ y $n \geq 20$.

EJEMPLO 5 Una nave espacial científica tiene 200 componentes electrónicas de un tipo particular. Durante la misión de la nave espacial, la probabilidad de falla individual para cada uno de estas componentes se estima que es 0.001. - Asumiendo que las fallas de estas componentes son independientes.

(a) ¿Qué tipo de distribución de probabilidad tendría, el número de fallas - durante la misión? ¿Cuales son los valores de los parámetros de esta distribución?

Utilice una aproximación a la verdadera distribución de probabilidad para hallar la probabilidad que durante la misión:

(b) no haya fallas, (c) haya menos de 3 fallas, (d) ocurra uno o dos fallas .

SOLUCION 1 Sea **X** la variable aleatoria definida por

X(ω) = número de fallas en las 200 componentes durante la misión.

$$R_X = \{0,1,2, \ldots, 200\}$$

2. Las fallas de las componentes son independientes y la probabilidad de falla **P**[E] = 0.001 permanece constante. Por lo tanto, la variable aleatoria **X** tiene una distribución binomial con perímetros p = 0.001 y n = 200

(a) Es decir,

$$P[X = x|B;200,\ 0.001\ ] = \binom{200}{x}\ (0.001)^x(0.99)^{200-x} \quad ,\ x = 0,1,\ldots 200$$

Puesto que n es grande y p pequeño (np = 200(0.001) = 0.2 < 5) se utiliza la distribución de poisson como aproximación de la distribución binomial con $\lambda = np = 0.2$. Es decir

$$P[X = x|B;200,0.001] \simeq \frac{(0.2)^x\ e^{-0.2}}{x!} \quad , \qquad x = 0,1, \ldots, 200.$$

(b) $P[X = 0\,] = e^{-0.2} = 0.8187$

(c) $P[X < 3\,] = 1 - P[X \geq 3\,] = 1 - 0.0011 = 0.9989$.

(d) $P[X = 1 \text{ ó } 2\,] = P[X = 1\,] + P[X = 2\,]$

$= P[X \geq 1] - P[X \geq 2\,] + P[X \geq 2\,] - P[X \geq 3]$

$= P[X \geq 1\,] - P[X \geq 3] = 0.1813 - 0.0011 = 0.1802.$

EJEMPLO 6 Una fábrica textil produce ciertas piezas de dimensiones específicas. Se sabe que la probabilidad que una pieza sea defectuosa es 0.02. En un lote de 100 piezas, ¿cuál es la probabilidad que no hayan piezas defectuosas? ¿cuál es la probabilidad que haya, a lo sumo, 3 piezas defectuosas?

SOLUCION 1 Sea **X** una variable definida como sigue

X(ω) = número de piezas defectuosas en 100 piezas.

n = 100, p = 0.02, luego $\lambda = np = 2$

2. La distribución de probabilidad de la variable aleatoria **X** es la binomial con parámetros 100 y 0.02.

3. Puesto que n es grande y p pequeña (np = 100(0.02) = 2 < 5) se aproxima por la distribución de Poisson con parámetro $\lambda = np = 2$. Es decir,

$$P[X = x|B:100,0.02\,] \approx \frac{2^x\, e^{-2}}{x!} \quad , \qquad x = 0,1,2, \ldots, 100.$$

(a) $P[X = 0\,] = e^{-2} = 0.1353$.

(b) $P[X \leq 3\,] = 1 - P[X \geq 4] = 1 - 0.1429 = 0.8571$.

EJEMPLO 7 Una vacuna produce inmunidad contra la polio en un 99.99% . Suponga que la vacuna ha sido administrado a 10,000 personas.

(a) ¿Cuál es el número esperado de personas que no han sido inmunizados?

(b) ¿Cuál es la probabilidad que exactamente k personas no son inmunes?

(c) ¿Cuál es la probabilidad que menos de 2 personas no son inmunes?

SOLUCION $X(\omega)$ = número de personas que no están inmunizados.

$$R_X = \{0,1,2, \ldots, 10{,}000\}$$

la probabilidad que una persona no sea inmunizado es 0.0001, o sea $n = 10{,}000$ y $p = 0.0001$

(a) $\mu = np = (10{,}000)(0.001) = 1$

(b) $P[X = k\,] = \dfrac{e^{-1}}{k!}$

(c) $P[X \leq 1\,] = e^{-1} + e^{-1} = 2e^{-1} = 2(0.3679) = 0.7358$.

EJEMPLO 8 Un libro tiene 400 páginas y se estima que hay 400 herrores de imprenta distribuídos aleatoriamente en todo el libro. Asumiendo una distribución de Poisson. ¿Cuál es el número de páginas que contienen:

(a) ningún error?

(b) exactamente un error?

(c) más de 2 errores?

(d) Se selecciona 10 páginas de dicho libro aleatoriamente. ¿Cuál es la probabilidad que ninguna tenga errores? ¿8 páginas no tengan errores?

SOLUCION $X(\omega)$ = número de errores por página.

$$R_X = \{0,1,2, \ldots, 400\}$$

$n = 400$ número de errores

El proceso se puede considerar así: las 400 páginas del libro se tienden en una mesa, luego se lanzan los 400 errores aleatoriamente, la cual se puede considerar como 400 ensayos de Bernoulli. La probabilidad de éxito es,

$$p = P[\text{un error caiga en una página cualquiera}] = \frac{1}{400}$$

Por lo tanto, $\lambda = np = 400(\frac{1}{400}) = 1$

Luego, la distribución de probabilidad de X es ,

$$P[X = x] = \frac{1^x e^{-1}}{x!} = \frac{e^{-1}}{x!} , \qquad x = 0,1,2, ... , 400$$

Se tiene que la proporción de páginas que contienen k errores está dado por,

$$P[X = k] .$$

Y el número de páginas que contienen k errores es,

$$(\text{número páginas})P[X = k] .$$

En nuestro caso será : $400\ P[X = k]$

(a) La proporción de las 400 páginas que no contienen errores es,

$$P[X = 0] = \frac{e^{-1}}{0!} = e^{-1} = 0,3679$$

Luego, el número de páginas que no contienen errores es

$$400\ (0.3679) = 147.$$

(b) La proporción de páginas que contienen 1 error es,

$$P[X = 1] = \frac{e^{-1}}{1!} = e^{-1} = 0.3679$$

por lo tanto, el número de páginas que tienen 1 error es

$$400\ (0.3679) \simeq 147 .$$

(c) La proporción de páginas que tienen más de 2 errores

$$P[X > 2] = 1 - P[X \leq 2] = 1 - \frac{5}{2} e^{-1} = 0.08025$$

por lo tanto, el número de páginas con más de 2 errores es

$$400\ (0.080) = 32 .$$

(d) Extraemos ahora 10 páginas aleatoriamente de las 400 que tiene el libro. (es decir tomamos una muestra de tamaño 10).

Definimos ahora la variable aleatoria Y de la siguiente manera

Y = número de páginas que no tienen errores en la muestra de 10 páginas

$R_Y = \{0,1,2,3,. . . , 10\}$

$n = 10$, y la probabilidad de obtener una página que no tiene errores se ha - calculado en la parte (a). Puesto que el tamaño de la muestra 10, es pequeña con respecto a la población 400 páginas. Tendremos que $p = 0.3679 = e^{-1}$ es - constante cada vez que repite el experimento de extraer una página. Luego, **Y** es una variable aleatoria binomial y su distribución de probabilidad es,

$$P[Y = y] = \binom{10}{y}(0.3679)^{y}(0.6321)^{10-y} \quad , \quad y = 0,1,2,\ldots,10$$

(i) $P[Y = 0] = (0.6321)^{10}$

(ii) $P[Y = 8] = \binom{10}{8}(e^{-1})^{8}(1 - e^{-1})^{2} = 45e^{-8}[1 - 2^{-1} + e^{-2}]$

EJEMPLO 9 Una panadería hace galletas con pedacitos de chocolate; un lote - tiene 1,000 galletas. Se agregan 3,000 pedacitos de chocolate a la masa para un lote y se mezcla bien toda la masa. Si se elige al azar 1 galleta de un lote, ¿Cuál es la probabilidad que no contenga ningún pedacito de chocolate? ¿De que contenga exactamente 3 pedacitos de chocolate?. Calcular, ¿cuántas galletas con solamente un pedacito de chocolate podría haber en el lote?

SOLUCION El proceso de poner los pedacitos de chocolate en la masa se puede considerar como sigue:

Se tiende toda la masa para un lote en una mesa, luego se corta cuadraditos de masa de igual tamaño (cada cuadradito tienen igual área) para hacer galletas Habrá 1,000 réctangulos de igual área, Finalmente se lanza los 3,000 pedacitos de chocolate aleatoriamente uno a uno sobre la masa, este experimento se puede considerar como 3,000 ensayos de Bernoulli. La probabilidad que caiga - un pedacito de chocolate en una galleta es, $p = \dfrac{1}{1,000}$

Sea **X** = número de pedacitos de chocolate que contiene una galleta.

$$R_X = \{0,1,2\ .\ .\ .\ ,\ 3,000\}$$

Podemos considerar a **X** como una variable aleatoria de Poisson con,

$$\lambda = np = 3,000\left(\frac{1}{1,000}\right) = 3$$

Entonces, $P[X = x] = \dfrac{3^{x}e^{-3}}{x!} \quad , \quad x = 0,1,2,\ldots, 3,000$

(a) $P[X = 0] = e^{-2} = 0.0498 \simeq 0.05$.

(b) $P[X = 3] = \dfrac{3^3 e^{-3}}{3!} = \dfrac{9}{2} e^{-3} = \dfrac{9}{2}(0.0498) = 0.224$.

(c) El número de galletas que contienen exactamente 1 pedacito de chocolate - es $1,000\ P[X = 1]$

pero, $P[X = 1] = 3e^{-3} = 3(0.0498) = 0.1494$

por lo tanto hay, $1,000(0.1494) \simeq 149$ galletas que tienen exactamente - 1 pedacito de chocolate.

EJEMPLO 10 Para estudiar la distribución de nidos de hormigas en un campo - abierto, se divide el campo en 1,600 cuadrados de igual área. Se observa que exactamente 400 de los cuadrados no contienen nidos de hormigas. Asumiendo - que los nidos están distribuídos aleatoriamente sobre el campo, estimar el número de nidos en el campo.

SOLUCION Sea X = número de nidos de hormigas que contiene cada cuadrado.

$$R_X = \{0,1,2, \ldots , n\}$$

Donde n es el número total de nidos de hormigas que hay. Además estos n - nidos están distribuidos aleatoriamente sobre el campo que está dividido en - 1,600 cuadrados de igual área. Consideremos ahora uno cualquiera de estos - 1,600 cuadrados. Supongamos también que lanzamos un nido sobre el campo aleatoriamente, entonces la probabilidad que el nido caigan en este cuadrado - particular ,

$$p = \frac{1}{1,600}$$

El supuesto proceso de lanzar los n nidos en el campo abierto se puede - considerar como n ensayos de Bernoulli.

Si n es grande, se puede usar la distribución de Poisson para calcular - la probabilidad (proporción) que un cuadrado particular contenga k nidos, con

$$\lambda = np = \frac{n}{1,600}$$

Luego,

$$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!} = \frac{\left[\frac{n}{1,600}\right]^k e^{-\frac{n}{1,600}}}{k!} , \quad k = 0,1,2,\ldots,n$$

La proporción de cuadrados que no contienen nidos de hormiga es,

$$P[X = 0] = \frac{(\frac{n}{1,600})^0}{0!} e^{-\frac{n}{1,600}} = e^{-\frac{n}{1,600}}$$

Entonces, el número de cuadrados que no contienen nidos de hormiga es,

$$1,600\ P[X = 0] = 1,600\ e^{-\frac{n}{1,600}}$$

la cual debe ser igual a 400 cuadrados según el enunciado. Es decir,

$$1,600\ e^{-\frac{n}{1600}} = 400$$

$$e^{-\frac{n}{1600}} = \frac{1}{4}$$

tomando logaritmo tenemos,

$$n = \frac{1,600 \log 4}{\log e} = \frac{1,600 \times 0.602}{0.4343} \simeq 2,218$$

$$n \simeq 2,218$$

5.6.3 PROPIEDAD REPRODUCTIVA DE LA DISTRIBUCION DE POISSON

La propiedad reproductiva de algunas distribuciónes de probabilidad consiste en que, si dos o más variables aleatorias con distribucción de probabilidad del mismo tipo se suman, la variable aleatoria resultante tiene una distribución del mismo tipo que los sumandos.Esta propiedad se llama propiedad reproductiva. Estableceremos aqui para la distribución de probabilidad de Poisson.

TEOREMA Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes de Poisson con parámetros $\alpha_1, \alpha_2, \ldots, \alpha_n$ respectivamente, entonces

$Y = \sum_{i=1}^{n} X_i$ es una variable aleatoria de Poisson con parámetro $\alpha = \sum_{i=1}^{n} \alpha_i$.

EJEMPLO 11 En una fábrica el número de accidentes por semana sigue un proceso de Poisson con parámetro $\alpha = 0.2$. Determinar :

(a) La probabilidad de que haya 4 accidentes en el transcurso de tres semanas

(b) La probabilidad que haya 2 accidentes en una semana, y otros 2 accidentes en la semana siguiente

(c) Es lunes, y ya ha habido un accidente. La probabilidad que en aquella semana no haya más de 3 accidentes.

SOLUCION (a) Definimos las variables aleatorias de Poisson con parámetro $\lambda_i = 2$ $(i = 1,2,3)$ respectivamente.

X_1 = número de accidentes en la primera semana

X_2 = número de accidentes en la segunda semana.

X_3 = número de accidentes en la tercera semana.

Las tres variables aleatorias son independientes. La variable aleatoria $X_1 + X_2 + X_3 = X$, número de accidentes en las tres semanas, también sigue una distribución de poisson de parámetro 2 + 2 + 2 = 6

$$P[X = 4] = \frac{6^4 e^{-6}}{4!} = 0.1339 .$$

(b) Si X_1 y X_2 son las variables aleatorias definidas en (a). Debemos calcular,

$$P[(X_1 = 2) \cap (X_2 = 2)] = P[X_1 = 2]P[X_2 = 2] = \frac{2^2 e^{-2}}{2!} \cdot \frac{2^2 e^{-2}}{2!} = 0.0733.$$

(c) Sea X = número de accidentes en una semana. Es una variable aleatoria de Poisson de parámetro $\lambda = 2$. Debemos calcular,

$$P[X \le 3 | X \ge 1] = \frac{P[[X \le 3] \cap [X \ge 1]]}{P[X \ge 1]} = \frac{P[1 \le X \le 3]}{P[X \ge 1]}$$

$$= \frac{P[X \le 3] - P[X = 0]}{P[X \ge 1]} = \frac{P[X \ge 1] - P[X \ge 4]}{P[X \ge 1]}$$

$$= \frac{0.8647 - 0.1429}{0.8647} = 0.8348 .$$

PROBLEMAS 5.6

1. Si X es una variable aleatoria con distribución de Poisson tal que $P[X = 0] = P[X = 1]$. Hallar $E(X)$.
2. Si X tiene una distribución de Poisson y $P[X = 0] = \frac{1}{2}$. Hallar $E(X)$.
3. Suponga que X es una variable aleatoria con distribución de Poisson. Si

$$P[X = 2] = (2/3)\,P[X = 1]$$

Calcular:

(a) $P[X = 0]$ (b) $P[X = 3]$

4. Los accidentes de trabajo, que se producen por semana en una fábrica, siguen una ley de Poisson, tal que, la probabilidad que haya 5 accidentes es 16/15 de que haya 2.

   Calcular:

   (a) el parámetro de la distribución de Poisson,

   (b) la probabilidad que no haya accidentes en 3 semanas.

5. Si X tiene una distribución de Poisson con $P[X = 1] = P[X = 2]$. Hallar $P[X = 1 \text{ ó } 2]$.

6. Los defectos de cierta clase de tejido de lana ocurren al azar con un promedio de 1 por 100 pies cuadrados. Calcular la probabilidad que una pieza que mide 50 por 10 pies no contenga defectos.

7. En determinada planta manufacturera han ocurrido accidentes a razón de 1 cada 2 meses. Suponiendo que ocurren en forma independiente. ¿Cuál es el número esperado de accidentes al año? ¿Cuál es la desviación estandar del número de accidentes al año? ¿Cuál es la probabilidad que no haya accidentes en determinado mes?

8. El número de casos admitidos de emergencia en cierto hospital en 1 hora es una variable aleatoria con distribución de Poisson con $\lambda = 3$. Determinar la probabilidad que en cierta hora

   (a) ningún caso de emergencia es admitido.

   (b) más de 3 casos de emergencia son admitidos.

9. Cierto alimento produce una reacción alérgica en un 0.01% de una población grande. Si 100,000 personas comen este alimento diario en promedio.

   (a) ¿Cuál es el número esperado de personas con reacción alérgica?

   (b) ¿Cuál es la función de probabilidad del número de personas en este grupo de 100,000 son alérgicos a este alimento?

10. Suponga que cierta enfermedad rara afecta al 0.1% de la población grande. 5,000 persona se escogen aleatoriamente de esta población y son sometidos a un examen para detectar la enfermedad.

    (a) ¿Cuál es el número esperado de personas con dicha enfermedad?

    (b) ¿Cuál es la probabilidad que exactamente 10 personas son afectadas por la enfermedad?

11. Suponga que un libro de 585 páginas contiene 43 errores tipográficos. Si estos errores se distribuyen aleatoriamente a trevés del libro. ¿Cuál es la probabilidad que 10 páginas, seleccionadas al azar, no tengan errores?

12. Suponga que un libro de 1,000 páginas contiene 500 errores tipográficos. Si estos errores se distribuyen aleatoriamente a través del libro. ¿Cuál es la probabilidad que 2 páginas de 10 seleccionadas al azar, no contengan errores? ¿Cuál es el número de páginas que no contienen ningún error? ¿Cuál es el número de páginas que contienen exactamente 1 error?

13. Suponga que 1,800 células de cierto tipo se distribuyen aleatoriamente en un microscopio, el cuál mediante rejillas se ha dividido en 900 áreas iguales.
    (a) ¿Cuál es el número de áreas que no contienen células?
    (b) ¿Cuál es el número de áreas que contienen exactamente una célula?

14. Una universidad procesa 100,000 calificaciones en determinado semestre, en ocasiones anteriores, se ha descubierto que 0.1% de todas las calificaciones estaban equivocadas. Suponer que una persona estudia cinco materias en esta universidad en un semestre. ¿Cuál es la probabilidad que todas las calificaciones estén correctas?

15. Un artillero dispara a un blanco. Se sabe que la probabilidad de acertar es 0.01. ¿Cuántos disparos tendrá que hacer para tener una probabilidad mayor que 0.9 de dar en el blanco por lo menos una vez?

16. Tres máquinas contribuyen a la producción de una fábrica de cierto artículo. La máquina A produce 3% de artículos defectuosos y contribuye a la producción total con un 40%; la máquina B produce 4% de artículos defectuosos y contribuye a la producción total con un 35%; mientras que C produce 5% de artículos defectuosos y 25% de la producción total. Los artículos son almacenados en cajas, cada una contiene un número grande y fijo de éstos, producidos por una máquina; por lo tanto, el 40% de las cajas contienen artículos producidos por la máquina A y así sucesivamente. Las cajas se almacenan juntas, sin que se pueda distinguir a qué máquina pertenecen los artículos que contienen. Se escoge una caja al azar, y se toma una muestra aleatoria de 100 artículos, de los cuales 4 son defectuosos. ¿Cuál es la probabilidad que provenga de la máquina A? ¿Cuál de que provenga de la máquina B? ¿Cuál que sea de la máquina C?

17. Ciertos automóviles llegan a una garita de peaje aleatoriamente con un pro

medio de 300 autos por hora. ¿Cuál es la probabilidad que

(a) Llegue exactamente 1 automóvil durante un periodo de 1 minuto?

(b) Lleguen por lo menos 2 automóviles en un período de 1 minuto?

18. La computadora marca VELOZ se descompone a razón de 0.05 veces por hora de operación, siendo necesario darle servicio especializado de reparación. - ¿Cuál es la probabilidad que no ocurran descomposturas en un período de trabajo de 8 horas? ¿Cuál es la probabilidad que no ocurran en una semana de 40 horas?. Suponga que las descomposturas ocurren según la distribución de Poisson.

19. La probabilidad que una persona sufra una reacción alérgica a un determinado medicamento es 0.001. Determine la probabilidad que de un total de 2,000 personas que han tomado el medicamento,

(a) Exactamente 3 tengan reacción alérgica ;

(b) más de dos personas tengan reacción alérgica ;

(c) Hallar el número esperado de personas con reacciones alérgicas y su varianza.

20. Se ha observado que el tránsito promedio de automóviles en determinado punto de un camino rural es de 3 por hora. Suponer que los instantes que pasan los mismos son independientes, haciendo que **X** representa el número de los que pasan por este punto en un intervalo de 20 minutos. Calcular

$$P[X = 0] \quad , \quad P[X \geq 2]$$

21. Suponga que la probabilidad de que un motor falle en un vuelo de rutina entre dos ciudades es 0.005. Usése la aproximación de Poisson a la distribución Binomial para encontrar aproximadamente la probabilidad de,

(a) por lo menos una falla en 1000 vuelos ;

(b) por lo menos dos fallas en 1000 vuelos .

22. La probabilidad que se haga una soldadura defectuosa en una conexión dada es 10^{-4} . Considere un sistema de 5×10^4 conexiones soldadas independientemente.

(a) ¿Cuál es la distribución de probabilidad del número de uniones defectuosas en el sistema? ¿cuales son los parámetros?

(b) utilice, una aproximación, a la verdadera distribución, para calcular la probabilidad de que no se presenten defectos en el sistema.

23. Se toma una muestra de 100, sin reposición, de un embarque recién llegado de 10,000 bombillas aléctricas. Se encuentra en la muestra tres bombillas defectuosas.

(a) ¿Cuál es la distribución de probabilidad del número de bombillas defectuosas en la muestra si se sabe que hay 500 bombillas defectuosas en el embarque?

(b) Utilice una aproximación a la verdadera distribución, para calcular la probabilidad aproximada del resultado de esta muestra.

24. Una panadería hace panetones con frutillas. Un lote tiene 200 panetones. Se agregan 1,000 frutillas a la masa para un lote. Se mezcla bien toda la masa para hacer los panetones. Se elige un panetón al azar de un lote, ¿Cuál es la probabilidad que no contenga ninguna frutilla? ¿Cuántos panetones con solamente una frutilla podría haber en el lote?

25. Una panadería hace panetones con frutillas. Un lote tiene 200 panetones. Al hacer el control de calidad se encontro que un panetón no tenía frutillas. Asumiendo que las frutillas de distribuyen aleatoriamente en la masa de un lote de panetones, estimar el número de frutillas que se puso al lote .

26. Suponga que un libro de 500 páginas contiene 50 errores tipo gráficos. Si estos errores se distribuyen aleatoriamente a través del libro, ¿cuál es la probabilidad que 10 páginas de 11 seleccionadas al azar, no contengan errores?

27. Una firma comercial tiene dos vendedores a tiempo parcial, Juan y Pedro. Juan trabaja los Lunes, Miercoles y Viernes, en tanto que Pedro lo hace los Martes y Jueves. Cada intento de venta se anota en una tarjeta y se archiva. Se sabe que Juan hace sus ventas en un 4% de los casos y Pedro en un 5%. Se elige un día cualquiera de la semana y se encuentra que se hicieron 100 tarjetas de intentos de ventas. ¿Cuál es la probabilidad que se haya hecho exactamente 4 ventas ese día?

28. Suponga que las moléculas de un gas raro se encuentran a razón promedio de 3 por cm^3 de aire. Suponga que las moléculas de este gas estan distribuidas independientemente y al azar en el aire. Determinar el tamaño de la muestra (número de cm^3) de aire que se debe tomar para que la probabilidad de encontrar al menos una molécula de este gas en la muestra sea 0.99 como mínimo.

29. Los pacientes con miembros fracturados llegan a una clínica particular en forma de eventos de un proceso de Poisson a razon de uno por día.

(a) ¿Cuál es la probabilidad que lleguen a la clínica siete en un perio do de siete días?

(b) Dado que llegaran precisamente siete en un periodo de siete días,¿cuál es la probabiliaddd que llegue uno en cada uno de los siete días?

30. Cierta industria envasa en cajas sus productos. La proporción de defectuosos en cada caja es de 0.02. El control de calidad de los productos se efectúa en la forma siguiente: Se toma una muestra de 100 productos de una caja elegida al azar. Si se encuentra a lo sumo 2 defectuosos, se acepta la caja; si se encuentran 3 ó 4 productos defectuosos, se elige una muestra - de 80 productos. Si en el total (180) de productos elegidos hay 5 ó más defectuosos, se rechaza la caja, y en caso contrario se acepta. Calcular la probabilidad que la caja elegida sea aceptada.

31. Sólo 10% de todos los insectos expuestos a un insecticida, en condiciones de laboratorio pudieron sobrevivir. Si se expone una muestra de 30 insectos al insecticida, ¿cuál es la probabilidad aproximada que:

(a) sobreviva exactamente un insecto?

(b) no sobreviva ninguno de los 30 insectos?

(c) no sobreviva más de dos insectos?

32. Una de cada 100 lámparas incandescentes fabricadas por una compañía se funde antes del final de su periodo de una semana si se dejan encendidas todo el tiempo. Se instala una lámpara en cada uno de los 50 pisos de un edificio grande.

(a) ¿Cuál es la verdadera distribución de probabilidad del número de lámparas que se fundan antes del final de la semana? ¿cuales son sus parámetros? Utilice una aproximación a la verdadera distribución para - calcular la probabilidad que

(b) tres lámparas se funden antes del final de la semana.

(c) más de tres lámparas se funden antes del final de la semana.

(d) menos de tres lámparas se funden antes del final de la semana.

33. Suponga que la página impresa de un libro contiene 40 líneas, y que cada - línea contiene 75 espacios (que pueden estar en blanco u ocupadas con algún símbolo). Por lo tanto, en cada página se deben formar 3000 espacios . Suponga que el linotipista comete un error en cada 6000 espacios que forma como promedio .

(a) ¿Cuál es la distribución del número de errores por hoja?

(b) Calcular la probabilidad que una página no contenga errores.

(c) ¿Cuál es la probabilidad que un capítulo de 16 páginas no contenga errores?

34. Si 250 litros de agua han sido contaminados con 10^6 bacterias, ¿cuál es la probabilidad que una muestra de 1 ml. de agua no contenga bacterias?

35. Unas láminas de metal presentan defectos en el cromado, los cuales aparecen en forma aleatoria en un número promedio de 1 por m^2 . ¿Cuál es la probabilidad que una hoja de 1.5m x 2m tenga como máximo un defecto?

36. Un biólogo desea saber el número total de células presentes en un área determinada sobre un portaobjetos que examina al microscopio y que está dividido por una cuadrícula en 400 (20 x 20) sub-áreas iguales. Si las células están distribuidas al azar en la cuadrícula, y el número total de células en el área completa es n.

(a) Determine una fórmula para la probabilidad que una sub-área no contenga célula.

(b) Deduzca una fórmula para el número de sub-áreas que, en promedio, no contienen ninguna célula.

(c) Estima el valor de n, si se observa que 50 de las sub-áreas no contienen células.

37. Un fabricante vende cierto artículo en lotes de 5000 piezas. De acuerdo con uno de sus clientes se adopta la siguiente regla de inspección:
Se selecciona al azar una muestra de 100 artículos de cada lote que se vende. Si la muestra contiene 4 ó menos artículos defectuosos, el cliente acepta el lote. Si se encuentran más de 4 defectuosos se inspecciona cada uno de los artículos del lote.
Si la inspección cuesta \$ 75 por cada 100 artículos, y el fabricante produce 2% de artículos defectuosos, determine el costo promedio de inspección por lote.

38. Se va ha probar un lote grande de artículos por el siguiente método de muestreo doble. En la primera muestra se toman 20 artículos y si no hay artículos defectuosos, el lote se aprueba; si hay 2 ó más artículos defectuosos, el lote se rechaza y si se encuentra 1 artículo defectuoso, se toma una segunda muestra de 40 artículos siguiendo las siguientes disposiciones si sólo hay 1 o ningún artículo defectuoso se aprueba el lote, en caso contrario se rechaza. Si la proporción real de defectuosos en el lote es 0.05 ¿cuál es la probabilidad de aceptar un lote?

39. Con base en experiencia pasada, 2% de las facturas de una empresa Editora están incorrectas. Se selecciona al azar una muestra de 20 facturas.

(a) ¿Cuál es la verdadera distribución de probabilidad del número de facturas incorrectas? ¿cuales son sus parámetros?

(b) ¿cuál es el número esperado de facturas incorrectas en la muestra?

(c) ¿Cuál es la probabilidad que cuando menos una factura esté incorrecta?. Utilice, una aproximación a la verdadera distribución de probabidad, del número de facturas incorrectas, para determinar;

(d) la probabilidad que, cuando menos, una factura esté incorrecta

(e) Compare y explique sus resultados de (c) y (d).

40. El número de accidentes automovilísticos registrados diareamente en cierta ciudad durante 100 días consecutivos es la siguiente

Número de accidentes	0	1	2	3	4	5	6
Número de días	19	26	26	15	9	4	1

(a) Verifique que la distribución de accidentes puede considerarse como - una distribución de Poisson.

(b) Compare el número teórico de días con el número observado de días con 0 accidentes, 1 accidentes, . . . , 6 accidentes.

6

DISTRIBUCIONES CONTINUAS IMPORTANTES

6.1 DISTRIBUCION UNIFORME

DEFINICION 6.1.1 Sea **X** una variable aleatoria contínua, que toma todos los valores de un intevalo $[a,b]$, donde a y b son números reales y $a < b$. Si la función de densidad de probabilidad de **X** está dada por

$$f(x) = \begin{cases} \dfrac{1}{b-a} & , \quad a \leq x \leq b \\ 0 & , \quad \text{en otros casos} \end{cases}$$

se dice que **X** se *distribuye uniformemente* en el intervalo $[a,b]$. La gráfica de la función de densidad se muestra en la figura 6.1.1.

Una variable aleatoria contínua uniformemente distribuída, representa la analogía a los resultados igualmente posibles en el sentido siguiente: Para cualquier su-intervalo $[c,d]$, donde $a \leq c \leq d \leq b$, la $P[c \leq X \leq d]$ es la misma para todo los subintervalos que tienen la misma longitud.

$$P[c \leq X \leq d] = \int_c^d \frac{dx}{b-a} = \frac{d-c}{b-a}$$

sólo depende de la longitud del intervalo y no de la ubicación del intervalo. La función de distribución acumulativa está dado por

$$F(x) = P[X \leq x] = \int_{\infty}^{x} \frac{1}{b - a}\, dx = \int_{a}^{x} \frac{1}{b - a}\, dx = \frac{x - a}{b - a}$$

que se puede escribir así

$$F(x) = \begin{cases} 0 & , \quad x \leq a \\ \dfrac{x - a}{b - a} & , \quad a \leq x < b \\ 1 & , \quad x \geq b \end{cases}$$

Su gráfica está representado en la fig. 6.1.2 .

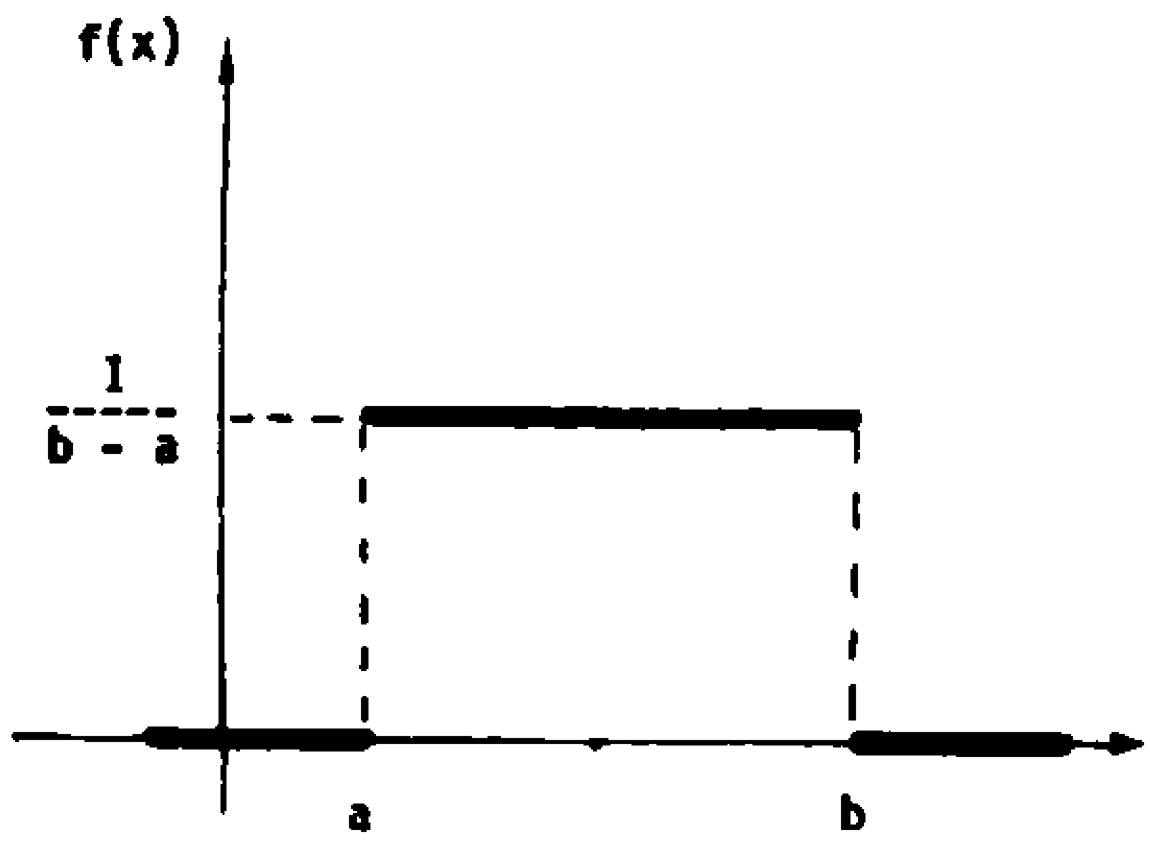


Fig. 6.1.1. Densidad uniforme

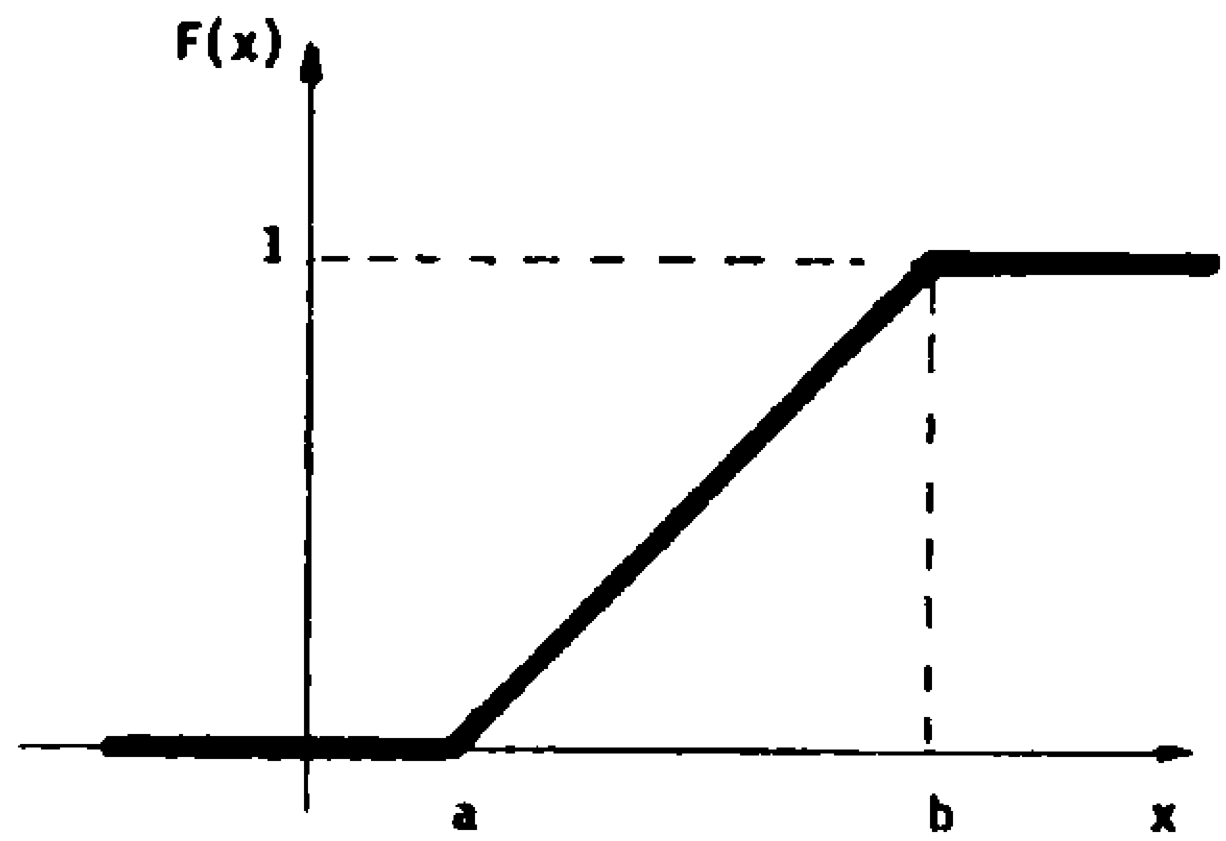


Fig. 6.1.2. Función de distribución uniforme

La *media y varianza* de la distribución uniforme son :

$$\mu = E(X) = \int_{a}^{b} \frac{x\ dx}{b - a} = \frac{x^2}{2(b - a)}\Bigg|_{a}^{b} = \frac{b^2}{2(b - a)} - \frac{a^2}{2(b - a)}$$

$$\boxed{\mu = \frac{a + b}{2}}$$

$$c^2 = Var(X) = \int_{a}^{b} \frac{x^2\ dx}{b - a} - \left(\frac{b + a}{2}\right)^2 = \frac{x^3}{3(b - a)}\Bigg|_{a}^{b} - \left(\frac{a + b}{2}\right)^2$$

$$= \frac{b^2 + ab + a^2}{3} - \frac{a^2 + ab + b^2}{4}$$

$$\boxed{\sigma^2 = \frac{(b - a)^2}{12}}$$

EJEMPLO 1 Sea **X** una variable aleatoria distribuida uniformemente en el intervalo $[0, 6]$. Calcular $P[|X - \mu| > 2]$.

SOLUCION Desde que la variable aleatoria **X** sigue una distribución uniforme en el intervalo $[0,6]$, la función de distribución acumulativa es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x}{6} & , \quad 0 \leq x < 6 \\ 1 & , \quad x \geq 6 \end{cases}$$

La esperanza de **X** es, $\mu = \frac{0 + 6}{2} = 3$.

$$\begin{aligned} P[|X - \mu| > 2] &= P[|X - 3| > 2] = P[X > 5] + P[X < 1] \\ &= 1 - P[X \leq 5] + P[X < 1] \\ &= 1 - F(5) + F(1) \\ &= 1 - \frac{5}{6} + \frac{1}{6} = \frac{1}{3} . \end{aligned}$$

EJEMPLO 2 Sea **X** una variable aleatoria continua distribuida uniformemente, con media $\mu = 1$ y varianza $\sigma^2 = 4/3$. Hallar $P[X < 0]$.

SOLUCION Suponga que **X** sigue una distribución uniforme en el intervalo $[a,b]$ Entonces,

$$\mu = \frac{a + b}{2} = 1 \quad (1) \qquad y \qquad \sigma^2 = \frac{(b - a)^2}{12} = \frac{4}{3} \quad (2)$$

Resolviendo el sistema de ecuaciones (1) y (2) se obtiene $a = -1$ y $b=3$. La función de distribución acumulativa de **X** es

$$F(x) = \begin{cases} 0 & , \quad x < -1 \\ \frac{x + 1}{4} & , \quad -1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

$$P[X < 0] = F(0) = \frac{1}{4} .$$

EJEMPLO 3 Los microbuses de cierta línea van a un horario estricto con intervalo de 8 minutos. Un pasajero llega de imprevisto a un determinado paradero. Si X es la variable aleatoria que representa el tiempo de espera del pasajero

Hallar

(a) la función de densidad de la variable aleatoria **X** .

(b) la función de distribución de **X** .

(c) la probabilidad que espere al microbus menos de 5 minutos.

(d) la probabilidad que espere más de 2 minutos.

(e) la probabilidad que espere exáctamente 7 minutos.

SOLUCION La variable aleatoria **X** está definida por

$X(\omega)$ = tiempo (en minutos) que debe esperar el pasajero.

Puesto que, el pasajero llega en cualquier instante al paradero, y todos los instantes son igualmente posibles (equiprobables), la variable aleatoria **X** sigue una distribución uniforme en un intervalo $[a,b]$ de longitud $b-a = 8$ Considerando $a = 0$ y $b = 8$, la función de densidad de **X** es

$$f(x) = \begin{cases} 1/8 & , \quad 0 < x < 8 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

(b) La función de distribución acumulativa es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ x/8 & , \quad 0 \leq x < 8 \\ 1 & , \quad x \geq 8 \end{cases}$$

(c) Sea el evento A : "el pasajero espera menos de 5 minutos"

A es equivalente a que la variable aleatoria **X** está comprendido entre 0 y 5 minutos. Es decir

$$P[A] = P[X < 5] = F(5) = \frac{5}{8}.$$

(d) B : "el pasajero espera más de 2 minutos"

$$P[B] = P[X > 2] = 1 - P[X \leq 2] = 1 - \frac{2}{8} = \frac{6}{8} = \frac{3}{4} \quad .$$

(e) $P[X = 7] = 0$, por que **X** es una variable aleatoria continua.

EJEMPLO 4 Suponer que se obtienen barras de mantequilla de un cuarto de libra con una máquina a partir de pedazos más grandes. Se supone que los pedazos más grandes son de densidad uniforme ; si la longitud de cada barra es exactamente $3\frac{3}{8}$ pulgadas, entonces la barra pesa 1/4 libra. Suponga que la longitud real **X** de una barra cortada por esta máquina tiene la misma posibilidad de estar comprendido en el intervalo entre 3.35 y 3.45 pulgadas. Suponiendo que las longitudes de las barras cortadas por esta máquina son independientes . ¿Cuál es la probabilidad que las 4 barras de un paquete determinado de mantequilla pesan cuando menos 1/4 de libra? ¿Que exactamente 3 pesen cuando menos 1/4 de libra?

SOLUCION La variable aleatoria X, definida como la longitud de una barra cortada por la máquina, tiene una distribución uniforme en el intervalo [3.35 , 3.45] . La función de distribución es

$$F(x) = \begin{cases} 0 & , \quad x < 3.35 \\ 10(x - 3.35) & , \quad 3.35 \le x < 3.45 \\ 1 & , \quad x \ge 3.45 \end{cases}$$

La probabilidad que una barra pese cuando menos 1/4 libra, es igual a la probabilidad que la longitud de la barra sea cuando menos $3\frac{3}{8}$ pulgadas. Entonces

$$P[\text{barra pesa cuando menos } \tfrac{1}{4} \text{ libra}] = P[X \ge 3\tfrac{3}{8}] \quad \text{ó}$$

$$P[X \ge 3.375] = 1 - P[X < 3.375] = 1 - 10(3.375 - 3.35) = \frac{3}{4}.$$

(a) Se extrae un paquete de 4 barras al azar. Teniendo en cuenta que las longitudes de las barras cortadas por la máquina son independientes, tenemos que

$$P[\text{las 4 pesan al menos 1/4 libra}] = P[X \ge 3.375]^4$$

$$= \left(\frac{3}{4}\right)^4 \simeq 0.316.$$

(b) $$P[\text{De las 4, exactamente 3 pesan al menos 1/4 libra}] = \binom{4}{3}\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)$$

$$= 0.422.$$

EJEMPLO 5 La variable aleatoria X está uniformemente distribuido en el intervalo [0,4]. ¿Cuál es la probabilidad que las raíces de la ecuación $y^2 + 4Xy + X + 1 = 0$ son reales?

SOLUCION Desde que la variable aleatoria X, está uniformemente distribuído en el intervalo [0,4] , su función de densidad es,

$$f(x) = \begin{cases} 1/4 & , \quad 0 \le x \le 4 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Resolviendo la ecuación cuadrática tenemos

$$y = \frac{-4X \pm \sqrt{16X^2 - 4(X+1)}}{2}$$

$y \in \mathbb{R}$, si, sólo si $16X^2 - 4X - 4 \geq 0$ ó

$$4X^2 - X - 1 \geq 0$$

Completando cuadrado obtenemos $(X - \frac{1}{8})^2 \geq \frac{17}{64}$, de donde

$$X \geq \frac{\sqrt{17}+1}{8} \quad \text{ó} \quad X \leq \frac{1-\sqrt{17}}{8}$$

la variable aleatoria X no puede tomar valores negativos, y

$$\frac{1-\sqrt{17}}{8} < 0$$

Por lo tanto, la ecuación tendrá raíces reales solamente en el caso

$$X \geq \frac{\sqrt{17}+1}{8}$$

y la probabilidad que esto ocurra es,

$$P[\frac{\sqrt{17}+1}{8} \leq X] = \int_{\frac{\sqrt{17}+1}{8}}^{4} \frac{1}{4}\,dx = 1 - \frac{\sqrt{17}+1}{32} = 0.8399$$

PROBLEMAS 6.1

1. Si X es una variable aleatoria distribuida uniformemente en el intervalo $[1,2]$. Hallar x tal que

$$P[X > x + \mu] = \frac{1}{4}.$$

2. Sea X una variable aleatoria contínua distribuída uniformemente en el intervalo $[-2,2]$. Calcular:

$$P[|X| > \frac{3}{2}], \quad P[|X-\mu| > 1], \quad P[\mu - 2\sigma \leq X \leq \mu + 2\sigma]$$

3. Sea X una variable aleatoria contínua uniformemente en $[-a,a]$, donde $a > 0$. Cada vez que sea posible, determinar a, de manera que se cumpla lo siguiente:

(a) $P[|X| > 1] = P[|X| < 1]$, (b) $P[X > 1] = \frac{1}{3}$

(c) $P[X < \frac{1}{2}] = 0.7$

4. Sea X una variable aleatoria distribuída uniformemente en [a,b] y simétrica al rededor del origen con varianza 1. Hallar los valores para a y b.

5. El tiempo medio en minutos, que cierta persona invierte en ir de su casa a la estación del tren es un fenómeno aleatorio que obedece una ley de distribución uniforme en el intervalo de 20 a 25 minutos. ¿Cuál es la probabilidad que alcance el tren que sale de la estación a las 7.28 a.m. en punto, si sale de su casa exactamente a las 7.05 a.m.

6. Los omnibuses de una línea interprovincial parten de la estación cada 30 minutos. Un viajero llega de imprevisto. Hallar
   (a) la función de distribución del "tiempo de espera del viajero".
   (b) la probabilidad que espere al omnibus menos de 15 minutos.
   (c) la probabilidad que espere por lo menos 10 minutos.

7. El minutero de un reloj eléctrico se mueve a saltos al final de cada minuto. Hallar la probabilidad que en un instante dado el reloj indique un tiempo que se diferencie del tiempo real a lo más 20 segundos. (Sug: defina X = diferencia del tiempo real. X está distribuida uniformemente, en [- 1,1])

8. El valor de una división de la escala de un equipo de medida es igual a 0.2. La indicación del equipo se redondea hasta la división entera próxima Determinar la probabilidad que al leer se cometa un error:
   (a) menor que 0.04
   (b) mayor que 0.05
   (Sug: defina X = error del redondeo de la lectura. X está distribuida uniformemente en el intervalo entre dos divisiones enteras contiguas a y b tal que $b - a = 0.2$. Considerar $a = 0$ y $b = 0.2$)

9. El petroleo es separado por destilación en las fracciones, listado en el cuadro siguiente

Fracción	Temperatura de destilación (°C)	Precio de venta por galón ($)
Gas	menos de 20	C_1
Petroleo ether	20 - 60	C_2
Ligroin	60 - 100	C_3

Suponga que C dólares es el costo de producir un galón de petroleo y la temperatura de destilación T está distribuida uniformemente en [0,100] .

Hallar el beneficio neto esperado (por galón) por las fracciones.

10. Cierto médico ordena a una persona seguir una dieta específica durante 3 - semanas. Suponiendo que el peso perdido tiene la misma posibilidad de estar comprendido entre 5 y 10 kg. Calcular:
    (a) la probabilidad que pierda más de 8kg.
    (b) la probabilidad que pierda a lo más 8kg.
    (c) la cantidad promedio que se espera perder .

11. Cierto semáforo permanece rojo 45 segundos cada vez. Usted llega (al azar) al semáforo y encuentra en rojo. Use una función de densidad uniforme apropiada para hallar la probabilidad que el semáforo se ponga verde en menos de 15 segundos.

12. Una película de dos horas se proyecta en sesión continua en un cine de la localidad. Usted sale para el cine sin comprobar primero las horas de proyección . Use una función de densidad uniforme apropiada para calcular la probabilidad que llegue al cine con menos de 10 minutos de adelanto ó de retraso respecto del comienzo de la película.

13. Una panadería termina una nueva hornada de pan de yema cada 45 minutos. - Usted llega (al azar) a la panadería esperando comprar un pan recien hecho. Use una función de densidad uniforme apropiada para hallar la probabilidad de que llegue dentro de los 5 minutos (antes o depués) del momento en que los panes salen del horno.

6.2 DISTRIBUCION EXPONENCIAL

DEFINICION 6.2.1 Sea **X** una variable aleatoria contínua. Se dice que **X** tiene una *distribución exponencial* con parámetro real λ, si su función de densidad está dado por

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

donde el parámetro real λ, es una constante positiva .

La gráfica de la función de densidad se muestra en la figura 6.2.1

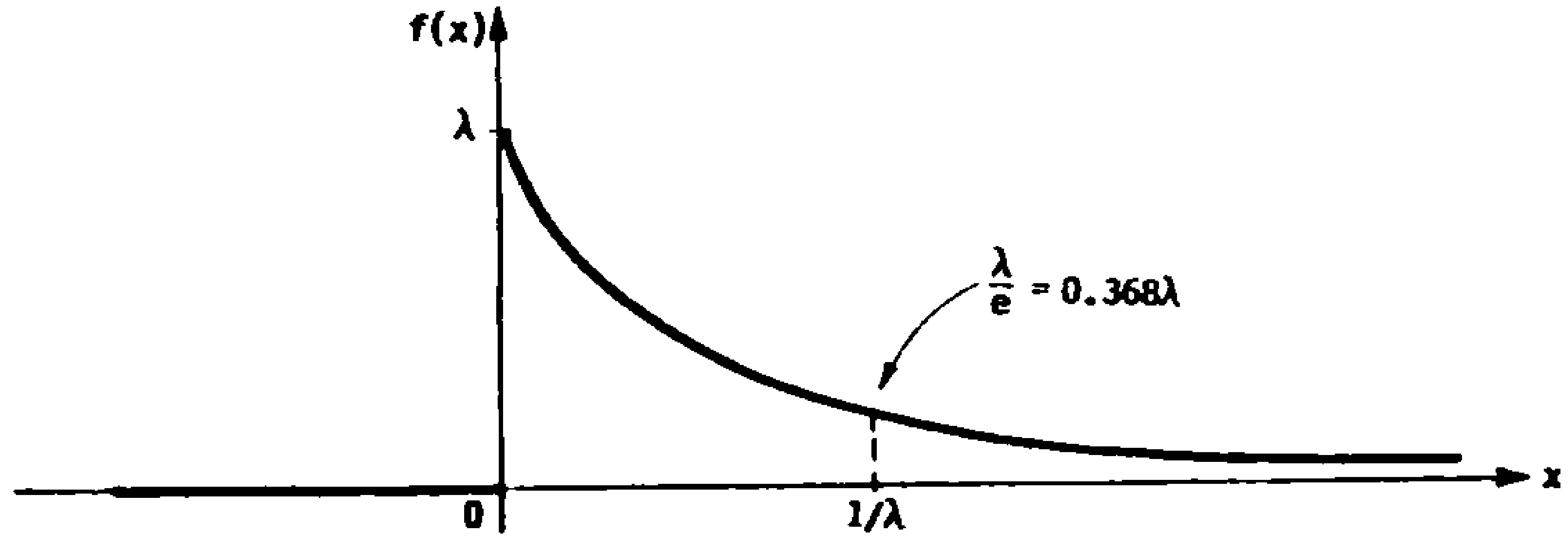


Fig. 6.2.1

La función de distribución de la variable aleatoria **X** con distribución exponencial es

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$$

Luego,
$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-\lambda x} & , \quad x \geq 0 \end{cases}$$

La figura 6.2.2 muestra la gráfica de $F(x)$

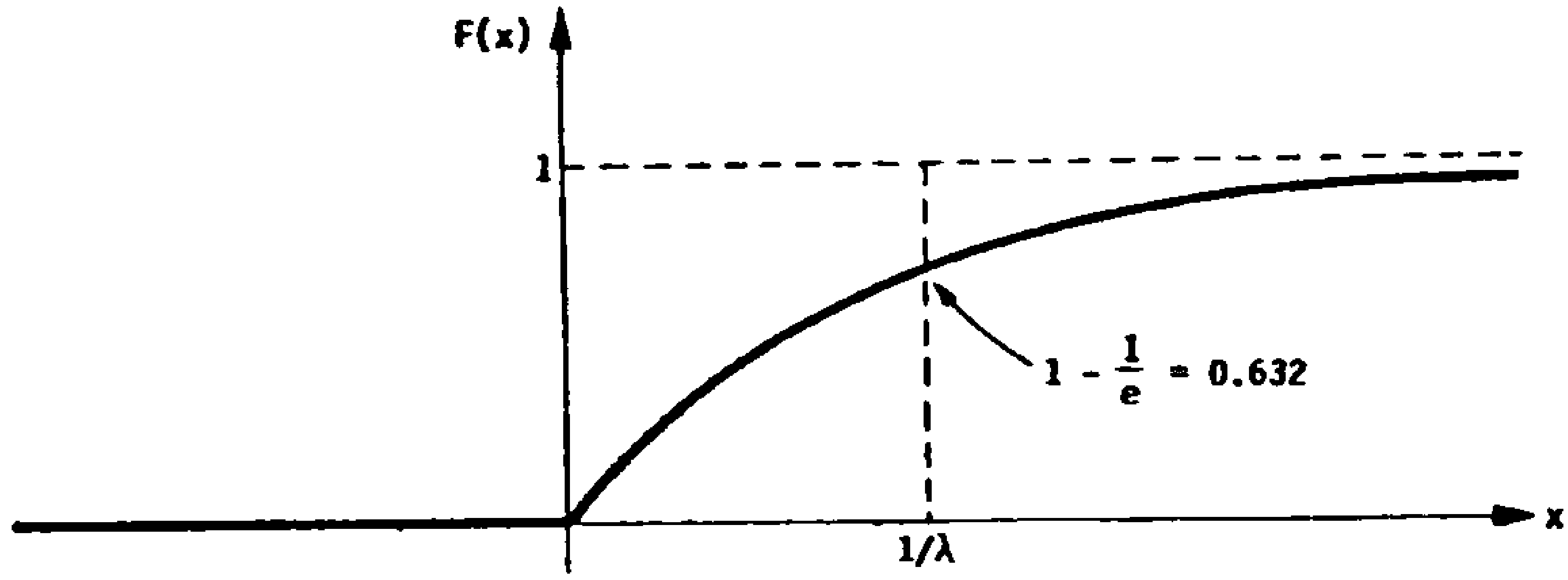


Fig. 6.2.2

Cálculo de esperanza (o media) de **X**.

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)dx = \int_{0}^{\infty} x\lambda e^{-\lambda x} dx$$

Aplicando integración por partes:

$$\mu = x \quad ; \qquad du = dx$$

$$dV = \lambda e^{-\lambda x}, \quad \text{de donde } V = -e^{-\lambda x}$$

Luego,

$$\mu = E(X) = \lim_{t\to\infty}\left\{ -xe^{-\lambda x}\Big|_0^t + \int_0^t e^{-\lambda x}\,dx \right\} = 0 + \lim_{t\to\infty}\left[\left(-\frac{1}{\lambda}\right) e^{-\lambda x}\Big|_0^t \right]$$

$$= \frac{1}{\lambda}$$

$$\boxed{\mu = \frac{1}{\lambda}}$$

Cálculo de la varianza de la variable aleatoria X.

$$E(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x}\,dx$$

Aplicando el método de integración por partes 2 veces obtenemos,

$$E(X^2) = \frac{2}{\lambda^2}$$

Luego, $\sigma^2 = E(X^2) - (E(X))^2 = \dfrac{2}{\lambda^2} - \dfrac{1}{\lambda^2} = \dfrac{1}{\lambda^2}$

$$\boxed{\sigma^2 = \frac{1}{\lambda^2}}$$

EJEMPLO 1 Suponga que la vida de cierto tipo de tubos electrónicos tiene una distribución exponencial con vida media de 500 horas. Si **X** representa la vida de un tubo (el tiempo que dura el tubo).

(a) Hallar la probabilidad que se queme antes de las 300 horas?

(b) ¿Cuál es la probabilidad que dure por lo menos 300 horas?

(c) Si un tubo particular ha durado 300 horas. ¿Cuál es la probabilidad de - que dure otras 400 horas?

SOLUCION Sabemos que $\mu = \dfrac{1}{\lambda}$, entonces $\dfrac{1}{\lambda} = 500$, de donde $\lambda = \dfrac{1}{500}$.

La función de densidad de la variable aleatoria **X** es,

$$f(x) = \begin{cases} \dfrac{1}{500} e^{-\frac{x}{500}} & , \quad x \geq 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

La función de distribución es

$$F(x) = \begin{cases} 0 & , \quad x < \\ 1 - e^{-\frac{x}{500}} & \quad x \geq \end{cases}$$

(a) $P[X < 300] = P[X \leq 300] = F(300) = 1 - e^{-\frac{300}{500}} = 1 - e^{-3/5}$

(b) $P[X > 300] = 1 - P[X \leq 300] = 1 - [1 - e^{-3/5}] = e^{-3/5}$.

(c) $$P[X \geq 700 | X > 300] = \frac{P[X \geq 700 \cap X > 300]}{P[X > 300]}$$

$$= \frac{P[X \geq 700]}{P[X > 300]} = \frac{1 - [1 - e^{-7/5}]}{1 - [1 - e^{-3/5}]}$$

$$= \frac{e^{-7/5}}{e^{-3/5}} = e^{-4/5} .$$

Observe que, $P[X \geq 700 \ X > 300] = P[X \geq 400]$

este, es una propiedad de la distribución exponencial que se conoce como la - de ***no tener memoria*** . En general, dado cualquier real a, $b > 0$, se tiene

$$P[X \geq (a + b) | X > a] = \frac{P[X \geq a + b]}{P[X > a]} = \frac{e^{-\lambda(a + b)}}{e^{-\lambda a}} = e^{-\lambda b}$$

Es decir, se cumple

$$P[X \geq (a + b) | X > a] = P[X \geq b]$$

EJEMPLO 2 El 5% de los tubos producidos por cierta factoría son defectuosos El tiempo de vida T de un tubo defectuoso es una variable aleatoria exponencial con media 0.5 (años), mientras que el tiempo de vida T_1 de un tubo no defectuoso es una variable aleatoria exponencial con media 2(años). ¿Cuál es la probabilidad que un tubo escogido al azar dure (a) a lo más 2 años?
(b) a lo más 4 años? , (c) menos de 2 años?.

SOLUCION Definimos la variable aleatoria **X** y el evento **D** como sigue :

$X(\omega)$ = tiempo de vida del tubo escogido al azar.

D = el tubo escogido es defectuoso.

$\bar{D}$ = el tubo escogido es no defectuoso

Debemos calcular $P[X \leq 2]$, $P[X \leq 4]$ y $P[X < 2]$

Usando el teorema de probabilidad total para los eventos $[X \leq k]$, D y $\bar{D}$ se tiene

$$P[X < k] = P[D]\,P[T < k|D] + P[\bar{D}]\,P[T_1 < k|\bar{D}]$$

$$= (0.05)\,P[T < k] + (0.95)\,P[T_1 < k] \qquad (1)$$

T y T_1 son variables aleatorias exponencial con parámetros $\lambda = \frac{1}{\mu} = \frac{1}{0.5} = 2$ y $\lambda_1 = \frac{1}{\mu_1} = \frac{1}{2} = 0.5$ respectivamente. La función de distribución de T y T_1 son

$$F(x) = \begin{cases} 0 & , \quad t < 0 \\ 1 - e^{-2t} & \quad t \geq 0 \end{cases} \quad \text{y} \quad F_1(x) = \begin{cases} 0 & t < 0 \\ 1 - e^{-0.5t} & , \; t \geq 0 \end{cases}$$

respectivamente. La expresión (1) se escribe

$$P[X \leq k] = (0.05)(1 - e^{-2k}) + (0.95)(1 - e^{-0.5k})$$

Las probabilidades pedidas son :

$$P[X \leq 2] = (0.05)(1 - e^{-4}) + (0.95)(1 - e^{-1})$$

$$= 1 - (0.05)e^{-4} - (0.95)e^{-1} = 0.6487$$

$$P[X \leq 4] = (0.05)(1 - e^{-8}) + (0.95)(1 - e^{-2})$$

$$= 1 - (0.05)e^{-8} - (0.95)e^{-2} = 0.8715$$

$$P[X < 2] = P[X \leq 2] = 0.6487$$

EJEMPLO 3 Suponga que un fabricante tenga que decidir entre dos procesos de fabricación de cierta componente electrónica. El costo del proceso A es de c dólares y del proceso B es kc dólares por unidad de componente, donde $k > 1$ Las componentes electrónicas tienen un tiempo de falla exponencial con una razón de 200^{-1} fallas por hora para las fabricada por el proceso A y 300^{-1} fallas por hora para B. El fabricante debe garantizar, de manera que, si una componente dure menos de 400 horas, pagará una multa de K dólares. ¿Qué proceso deberá usar?

SOLUCION Definimos las variable aleatorias siguientes :

T = tiempo de falla de cada componente electrónica.

C_A = costo (por unidad) de la componente electrónica fabricado por el proceso A.

C_B = costo (por unidad) de la componente electrónica fabricado por el proceso B .

para el proceso A, tenemos

$$C_A = \begin{cases} c & T \geq 400 \\ c + K \quad , & T < 400 \end{cases}$$

para el proceso B, tenemos

$$C_B = \begin{cases} kc \quad , & T \geq 400 \\ kc + K \quad , & T < 400 \end{cases}$$

La función de distribución de la variable T es

$$F(t) = \begin{cases} 0 \quad , & T < 0 \\ 1 - e^{\lambda t} \quad , & T \geq 0 \end{cases}$$

donde λ es 200^{-1} y 300^{-1} para las componentes fabricados por los procesos A y B respectivamente.

Calculemos el costo esperado para cada uno de los procesos

$$E(C_A) = (c + k)\,P[T < 400\,] + cP[T > 400\,]$$

$$= (c + K)(1 - e^{-\frac{400}{200}}) + ce^{-\frac{400}{200}} = c + K(1 - e^{-2})$$

$$E(C_B) = (kc + K)\,P[T < 400\,] + kc\,P[T > 400\,]$$

$$= (kc + K)(1 - e^{-\frac{400}{300}}) + kc\,e^{-\frac{400}{300}} = kc + K(1 - e^{-4/3})$$

$E(C_A)$ será mayor, igual o menor que $E(C_B)$, según que $E(C_A)/\,E(C_B)$ sea mayor, igual o menor que uno. Es decir

$$\frac{E(C_A)}{E(C_B)} > 1 \quad ; \quad \frac{E(C_A)}{E(C_B)} = 1 \quad ; \quad \frac{E(C_A)}{E(C_B)} < 1 \quad .$$

cuando se cumple la primera relación escogemos el proceso B; en cambio si ocurre la segunda, escogemos cualquiera, y cuando se cumple la tercera relación escogemos el proceso A. En nuestro caso

$$\frac{E(C_A)}{E(C_B)} = \frac{c + K(1 - e^{-2})}{kc + K(1 - e^{-4/3})} > 1 \text{ ó } 1 - \frac{K}{c}[\,e^{-2} - e^{-4/3}\,] > k \qquad (1)$$

si k cumple la relación (1) entonces la razón $E(C_A)\,/\,E(C_B) > 1$, escogemos el proceso B .

6.2.1 RELACION ENTRE LA DISTRIBUCION EXPONENCIAL Y POISSON

La distribución exponencial tiene una relación especial con la distribución de Poisson. Hemos visto que la distribución de Poisson describe el número de ocurrencias de eventos (éxitos) por unidad de medida (intervalo de tiempo, una área determinada etc.). Por ejemplo, número de llamadas telefónicas por minuto. Otro ejemplo es: número de carros que llegan a la caseta de peaje cada hora. En cambio la distribución exponencial describe el valor de la medida por ocurrencia de eventos (éxito). Por ejemplo, el tiempo transcurrido entre llamadas telefónicas sucesivas. El tiempo transcurrido entre llegadas sucesivas de automóviles a la caseta de peaje. La longitud de tiempo entre llamadas es continuo, lo mismo que la longitud entre llegadas de los automóviles a la caseta. El tiempo entre llegadas sucesivas (o longitud entre ocurrencias sucesivas), se denomina: *Tiempo entre llegadas*.

Así pués, las dos distribuciones se puede utilizar para describir el mismo fenómeno, la distribución de Poisson describe el número de ocurrencias de eventos por unidad de media, y la exponencial, describe valor de la medida entre ocurrencias sucesivas de eventos.

Consideremos un proceso de Poisson como sigue:

$X(\omega)$ = número de ocurrencias de un evento en un período t . definimos el parámetro λ del proceso de Poisson como el número esperado de ocurrencia por unidad de tiempo. Entonces, el número esperado de ocurrencias en el intervalo t será λt, luego la distribución de probabilidad de X es

$$P[X = x] = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad , \qquad x = 0,1,2,\ldots \qquad (1)$$

Supongamos ahora que estamos interesados en la distribución de probabilidad del intervalo de tiempo entre ocurrencias de los eventos. Es decir, la variable aleatoria está definida ahora por :

$T(t)$ = intervalo de tiempo entre ocurrencias sucesivas de eventos.

Luego, T es contínua obviamente y su rango es

$$R_T = \{t \in \mathbb{R} / t > 0\}$$

La distribución de probabilidad de T, se deduce de la distribución de Poisson dado en (1) .

Observe, que la probabilidad de la no ocurrencia de eventos en el período t, es

$$P[X = 0] = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

y la probabilidad de no ocurrencia de eventos en el intervalo t es igual a la probabilidad de que T excede a t así, pués

$$P[T > t] = P[X = 0] = e^{-\lambda t}$$

de donde obtenemos que,

$$P[T \leq t] = 1 - P[T > t] = 1 - e^{-\lambda t}$$

por lo tanto, la función de distribución es

$$F(t) = \begin{cases} 0 & , \quad t < 0 \\ 1 - e^{-\lambda t} & , \quad t \geq 0 \end{cases}$$

de donde la función de densidad de probabilidad, la que hemos llamado *distribución exponencial* , es

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & , \quad t \geq 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

EJEMPLO 4 Un vendedor vende períodicos en una esquina. Los períodicos que vende son eventos de un proceso de Poisson con parámetro $\lambda = 50$ por hora. Si alguien acaba de comprarle un períodico. ¿Cuál es la probabilidad que transcurran al menos 2 minutos antes que venda otro? ¿De que no pasen más de 5 minutos?. Si ya han transcurridos 5 minutos desde la última venta. ¿Cuál es la probabilidad que transcurran al menos 2 minutos más para su siguiente venta?

SOLUCION Sea **X** la variable aleatoria definida por

$X(\omega)$ = longitud de tiempo entre ventas.

En el proceso de Poisson $\lambda = 50$ por hora. Entonces el promedio de ventas por minuto será, $\frac{50}{60} = \frac{5}{6}$. La función de distribución de **X** es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-\frac{5}{6}x} & , \quad x \geq 0 \end{cases}$$

(a) $P[X > 2] = 1 - P[X \leq 2] = 1 - [1 - e^{-5/3}] = e^{-5/3}$.

(b) $P[X \leq 5] = F(5) = 1 - e^{-25/6}$.

(c) $P[X > 7 | X > 5] = \dfrac{P[X > 7 \cap X > 5]}{P[X > 5]} = \dfrac{P[X > 7]}{P[X > 5]} = e^{-5/3}$

EJEMPLO 5 Suponga que el tiempo entre la llegada sucesiva de clientes a la ventanilla de una cajera de un banco, se sabe que es una exponencial con media de 0.20 minutos. ¿Cuál es la probabilidad de un intervalo de menos de 10 segundos entre una llegada y la siguiente?

SOLUCION Definimos la variable aleatoria **X** por

$X(\omega)$ = intervalo de tiempo entre llegadas sucesivas de los clientes.

$$\mu = \frac{1}{\lambda} = 0.20 \text{ mint.} , \quad \text{entonces } \lambda = 5 .$$

La función de distribución acumulada es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-5x} & , \quad x \geq 0 \end{cases}$$

10 seg = $\frac{1}{6}$ mint. Por lo tanto

$$P[X < \tfrac{1}{6}] = 1 - e^{-\frac{5}{6}} = 1 - e^{-0.633} .$$

EJEMPLO 6 Suponga que el tiempo que necesita un cajero de un banco para atender a un cliente tiene una distribución exponencial con una media de 40 segundos.

(a) ¿Cuál es la probabilidad que el tiempo necesario para atender un cliente dado sea mayor que 20 minutos?

(b) ¿Cuál es la probabilidad que el tiempo necesario para atender a un cliente esté comprendido entre 1 y 2 minutos.

SOLUCION Sea **X** la variable aleatoria definida por

$X(\omega)$ = intervalo de tiempo necesario para atender a un cliente .

$$\mu = \frac{1}{\lambda} = 40 \text{ segundos} = \frac{40}{60} \text{ minutos} = \frac{2}{3} \text{ minutos}$$

de donde $\lambda = \frac{3}{2}$ minutos. Por lo tanto, la función de distribución es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-\frac{3}{2}x} & , \quad x \geq 0 \end{cases}$$

(a) $P[X > 2] = 1 - P[X \leq 2] = 1 - [1 - e^{-\frac{3}{2}(2)}] = e^{-3}$

(b) $P[1 < X < 2] = F(2) - F(1) = 1 - e^{-3/2(2)} - [1 - e^{-3/2(1)}]$

$$= e^{-3/2} - e^{-3}$$

EJEMPLO 7 En promedio, los barcos de carga llegan aleatoriamente a cierto puerto, a razón de 1 cada dos días. ¿Cuál es la probabilidad que el tiempo entre dos llegadas consecutivas de buques sea mayor de dos días?

SOLUCION Sea X la variable aleatoria definida por

$X(\omega)$ = longitud de tiempo entre llegadas consecutivos de los barcos

λ = promedio de llegadas por cada unidad de medida = 1 cada dos días

La definición de distribución es

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-x} & , \quad x \geq 0 \end{cases}$$

$$P[X > 2] = 1 - F(2) = e^{-2} .$$

6.2.2 APLICACION DE LA EXPONENCIAL EN LA TEORIA DE LA CONFIABILIDAD

La distribución exponencial se usa en muchas aplicaciones de la teoría de la confiabilidad, que consiste en la elaboración de modelos de cuán confiables o no son las componentes o sistemas. Supongamos que la componente o el sistema comienza a trabajar en el instante $t_0 = 0$ y observamos hasta que falla. Supongamos que ocurre en el instante t. Designaremos por T la variable aleatoria continua "el tiempo para fallar" o "la duración del tiempo de trabajo sin falla". Entonces, la probabilidad de falla de la componente o sistema en el tiempo t (o período t) es

$$P[T \leq t] = F(t)$$

donde $F(t)$ es la función de distribución de la variable aleatoria T. Por lo tanto, la probabilidad de trabajo sin falla de la componente o sistema en el tiempo t (o período t) es

$$P[T > t] = 1 - F(t) .$$

DEFINICION 6.2.2 La confiabilidad $R(t)$ de una componente (o un sistema) en determinado medio durante un período t se define como la probabilidad de que su tiempo para fallar excede a t (o sea que trabaja satisfactoriamente en el período t); es decir

$$R(t) = P[T > t] = 1 - F(t) .$$

Si el tiempo T para fallar es una variable aleatoria exponencial, entonces

$$R(t) = 1 - [1 - e^{-\lambda t}] = e^{-\lambda t} .$$

EJEMPLO 8 Suponga que la compañía Eléctrica RMC encuentra que el tiempo de fallo en meses **X** de una bombilla para iluminación tiene la siguiente función de densidad

$$f(x) = \begin{cases} \frac{1}{10} e^{-x/10} & , \quad x \geq 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar la confiabilidad de la bombilla para un período de un año

SOLUCION En este caso la variable aleatoria **X** está definida por

$X(\omega)$ = tiempo de falla (en meses) de una bombilla

$x = 12$, periodo de un año

Luego,

$$\begin{aligned} R(12) &= P[X > 12] \\ &= e^{-12/10} \quad , \qquad \text{por ser } X \text{ una exponencial} \\ &= 0.3012 . \end{aligned}$$

6.2.3 LA DISTRIBUCION GAMMA

Una función usado en la definición de la distribución gamma es la *función gamma* denotada por Γ y definida como sigue:

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx \quad , \qquad \text{para} \quad n > 0 \tag{1}$$

Se puede demostrar que cuando $n > 0$ la integral impropia anterior converge. Es decir, existe el siguiente límite

$$\lim_{k \to \infty} \int_0^k x^{n-1} e^{-x} dx$$

Una importante relación recursiva se obtiene integrando por partes la ecuación (1) con $u = x^{n-1}$ y $dv = e^{-x} dx$

$$\Gamma(n) = \lim_{t \to \infty} - e^{-x} x^{n-1} \Big|_0^t - \int_0^\infty e^{-x} (n-1)x^{n-2}\, dx$$

$$= 0 + (n-1)\int_0^\infty e^{-x}x^{n-2}\, dx = (n-1)\quad \Gamma(n-1)$$

Si n es un *entero positivo* , entonces

$$\Gamma(n) = (n-1)\quad \Gamma(n-1) = (n-1)(n-2)\quad \Gamma(n-2) = \ldots$$

$$= (n-1)(n-2)\ldots \qquad \Gamma(1)$$

Desde que $\Gamma(1) = \int_0^\infty e^{-x}\, dx = 1$ obtenemos

$$\Gamma(n) = (n-1)!$$

Es decir, la función gamma es una generalización de la función factorial. - también se verifica que

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty x^{-1/2}\, e^{-x}\, dx = \sqrt{\pi}$$

DEFINICION 6.2.3 Una variable aleatoria continua X se dice que tiene una - distribución gamma, con parámetros $r > 0$ y $\lambda > 0$ si su función de densidad está dada por

$$f(x) = \begin{cases} \dfrac{\lambda}{\Gamma(r)}\,(\lambda x)^{r-1}e^{-\lambda x} & , \quad x > 0 \\ 0 & , \quad \text{en otros casos.} \end{cases}$$

PROBLEMAS 6.2

1. Hallar, si existe la función de densidad exponencial que cumple la siguiente condición,

$$P[X \le 2\,] = \frac{2}{3}\,P[X \le 3\,]$$

2. Si $f(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x > 0\ , \quad \lambda > 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$

   Hallar el valor de a tal que, $P[X \le a\,] = 2P[X > a\,]$

3. La longitud de vida de una cierta clase de bacterias en un cierto medio - ambiente es una variable aleatoria continua X, cuya distribución de probabilidad es aproximadamente una distribución exponencial. Si el promedio - de duración de vida es 12 horas. Calcular la probabilidad
   (a) de que una bacteria particular muera antes de las 12 horas.
   (b) de que una bacteria, la cual ha vivido 12 horas, muera antes de las - 12 horas más?

4. En un estudio de dispersión de una población de insectos, un gran número - de hormigas son soltadas en un punto dado. Después de 1 minuto, se observa que la proporción de hormigas que estan a una distancia mayor que r metros del punto donde fueron soltados es aproximadamente e^{-2r}.
   (a) ¿Qué proporción de hormigas han recorrido más de 1 metro del punto de partida?
   (b) ¿Cuál es el promedio de la distancia recorrida por las hormigas del - punto de partida?

5. El tiempo (en años) que un satélite permanece en el espacio es una variable aleatoria exponencial T, cuya función de distribución está dado por

$$F(x) = 1 - e^{-0.5t} , \qquad x > 0$$

   (a) Hallar la probabilidad que un satélite permanece en el espacio entre - uno y tres años.
   (b) ¿Cuál es la probabilidad que un satélite permanece en el espacio más - de 4 años?
   (c) Si son lanzados 3 satélites simultáneamente, ¿cuál es la probabilidad de que por lo menos uno permanece en el espacio más de 4 años?.

6. Suponga el tubo de imagen de televisión tiene una longitud de vida X (en - años) la cual es una variable aleatoria exponencial con vida media de 5 años. El costo de fabricación de un tubo es \$ 40. El fabricante vende el tubo a \$ 75, pero garantiza un reintegro total, si el tubo no dura 4 años ¿Cuál es el beneficio esperado por tubo del fabricante?

7. La longitud de vida de una especie de planta en cierto medio ambiente es - una variable aleatoria continua X, que tiene una distribución exponencial con una longitud media de vida de 1,000 días.
   (a) ¿Qué proporción de plantas de esta especie mueren antes de los 1,000 - días?

(b) ¿Si una planta individual vive durante 800 días, ¿Cuál es la probabilidad que viva otros 400 días?.

8. Considere unos focos producidos por una máquina de los que sabemos la duración X, en horas, de un foco producido por la máquina es una variable aleatoria con distribución exponencial y media 1,000 horas.
   (a) ¿Cuál es la probabilidad de que una muestra aleatoria de 5 focos no contenga focos con duración menor que 1020 horas?
   (b) Supongamos ahora que la muestra de 5 focos se coloca en una caja. Si se selecciona aleatoriamente un foco de la caja, ¿Cuál es la probabilidad que el foco seleccionado tenga una duración mayor que 1020 horas?

9. El tiempo que tarda una persona en ser atendida en una cafeteria es una variable aleatoria que tiene una distribución exponencial con media de 4 minutos. ¿Cuál es la probabilidad que una persona sea atendida en menos de 3 minutos, al menos en 4 de los 6 días siguientes?.

10. Supongamos que una componente electrónica tiene una duración, en años dada por la variable aleatoria T, distribuida exponencialmente con parámetro $\lambda = 5$. En un circuito se instalan cinco de estas componentes, ¿cuál es la probabilidad que al menos dos de ellos continua funcionando al termino de 8 años?
    (Sug. defina X = número de componentes que funcionan después de 8 años, X es una binomial con $p = P[T > 8]$)

11. En un conmutador telefónico se reciben llamadas de acuerdo a un proceso de Poisson con parámetro $\lambda = 5$ por hora. Si hay una persona en el conmutador. ¿Cuál es la probabilidad que transcurran al menos 15 minutos antes de la siguiente llamada? ¿De que no pasan más de 10 minutos? Si ya han transcurrido 10 minutos desde la última llamada, ¿cuál es la probabilidad de que transcurran a lo más 5 minutos más para la siguiente llamada?

12. Los barcos llegan a cierto puerto según la ley de Poisson con un promedio de dos horas entre dos llegadas. Determinar la probabilidad que :
    (a) Transcurran cinco horas sin ninguna llegada;
    (b) como máximo llegan 3 barcos durante un intervalo de 5 horas.

13. El tiempo en minutos entre las llegadas de dos vehículos sucesivos a un puesto de peaje es una variable aleatoria X exponencialmente distribuído con función de densidad

$$f(x) = \begin{cases} 0.5\, e^{-0.5x} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

(a) Determine la probabilidad que un par de coches sucesivos aleatoriamente seleccionados lleguen al puesto de peaje con al menos 6 minutos de diferencia.

(b) Halla el intervalo medio de tiempo entre las llegadas de coches sucesivos al puesto de peaje.

14. La vida en años de cierto aparato eléctrico es una variable aleatoria X exponencialmente distribuida con función de densidad

$$f(x) = \begin{cases} 0.08\, e^{-0.08x} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

El aparato lleva una garantía de fábrica por dos años. Suponga que adquiere uno de esos aparatos, seleccionando al azar entre las existencias del fabricante. ¿Cuál es la probabilidad que la garantía expire antes que su aparato se vuelva inútil ?

15. Un fabricante de mecanismos electrónicos para mini computadoras encuentra que el tiempo en años al cabo del cual el mecanismo requiere reparaciones es una variable aleatoria X con función de densidad exponencial con parámetro 1/25.

(a) ¿Cuál es la probabilidad que el mecanismo no requiera reparación durante 6 años?

(b) El fabricante desea garantizar el mecanismo para que el 90% de las computadoras no requiere reparación dentro del período de garantía. ¿Que tan largo debe escogerse el periodo de garantía?

16. La probabilidad de buen funcionamiento de un elemento de cierto equipo se distribuye según una ley exponencial

$$f(t) = \begin{cases} 0.02\, e^{-0.02t} & , \quad t > 0 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar la confiabilidad del elemento en un período de 50 horas.

6.3 DISTRIBUCION NORMAL

La distribución de probabilidad contínua más importante en todo el campo de la estadística, es con toda seguridad, la *distribución normal*, debido a que en la práctica muchos fenómenos, industriales, cientificos, o de la vida diaria pueden describirse por esta distribución. También mediante esta distribución se obtienen aproximaciones a otras leyes de probabilidad.

DEFINICION 6.3.1 Una variable aleatoria contínua X, se dice que está *distribuída normalmente*, con media μ $(-\infty < \mu < \infty)$ y varianza $\sigma^2 > 0$, si su función de densidad de probabilidad esta dado por,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(x-\mu)/\sigma]^2} \quad , \quad -\infty < x < \infty$$

donde $\pi = 3.14159 \ldots$ y $e = 2.71828 \ldots$

NOTACION Una notación muy usada para la distribución normal es,

$$X \rightarrow N(\mu , \sigma^2)$$

que se lee: "la variable aleatoria X se distribuye normalmente con media μ y varianza σ^2" (se usa también la notación $X \rightarrow N(\mu,\sigma)$, que se lee: "X tiene - una distribución normal con media μ y desviación estándar σ").

La gráfica de la distribución normal está dado en la fig. 6.3.1 . Para - una mejor comprensión daremos algunas pautas generales de cómo se obtiene esta gráfica.

1. $$f'(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(x-\mu)/\sigma]^2} \left[-\frac{(x-\mu)}{\sigma^2}\right]$$

$$f'(x) = -\frac{x-\mu}{\sigma^3\sqrt{2\pi}} e^{-(1/2)[(x-\mu)/\sigma]^2}$$

Observe que $f'(x) = 0$ si, sólo si $x = \mu$ y .

$$f(\mu) = \frac{1}{\sigma\sqrt{2\pi}} \quad , \quad \text{entonces} \quad \left(\mu, \frac{1}{\sigma\sqrt{2\pi}}\right) \text{ es un punto crítico}$$

2. Si $x < \mu$, $f'(x) > 0$, luego f es creciente en $\langle -\infty, \mu \rangle$
3. Si $x > \mu$, $f'(x) < 0$, luego f es decreciente en $\langle \mu, \infty \rangle$

4. $\forall x$, $f(x) < f(\mu)$, por lo tanto la función f tiene un máximo absoluto en μ

5. $$f''(x) = \frac{(x-\mu)^2}{\sigma^5\sqrt{2\pi}}\, e^{-(1/2)[(x-\mu)/\sigma]^2} - \frac{1}{\sigma^3\sqrt{2\pi}}\, e^{-(1/2)[(x-\mu)/\sigma]^2}$$

$f''(x) = 0$ si sólo si $x = \mu \pm \sigma$

Por lo tanto, los puntos $\left(\mu - \sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right)$ y $\left(\mu + \sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right)$

son de inflexión. Entonces, la gráfica de f es cóncava hacia arriba en $\langle -\infty, \mu - \sigma\rangle$, cóncava hacia abajo en $\langle \mu - \sigma, \mu\rangle$ y $\langle \mu, \mu+\sigma\rangle$, cóncava hacia arriba en $\langle \mu + \sigma, \infty\rangle$

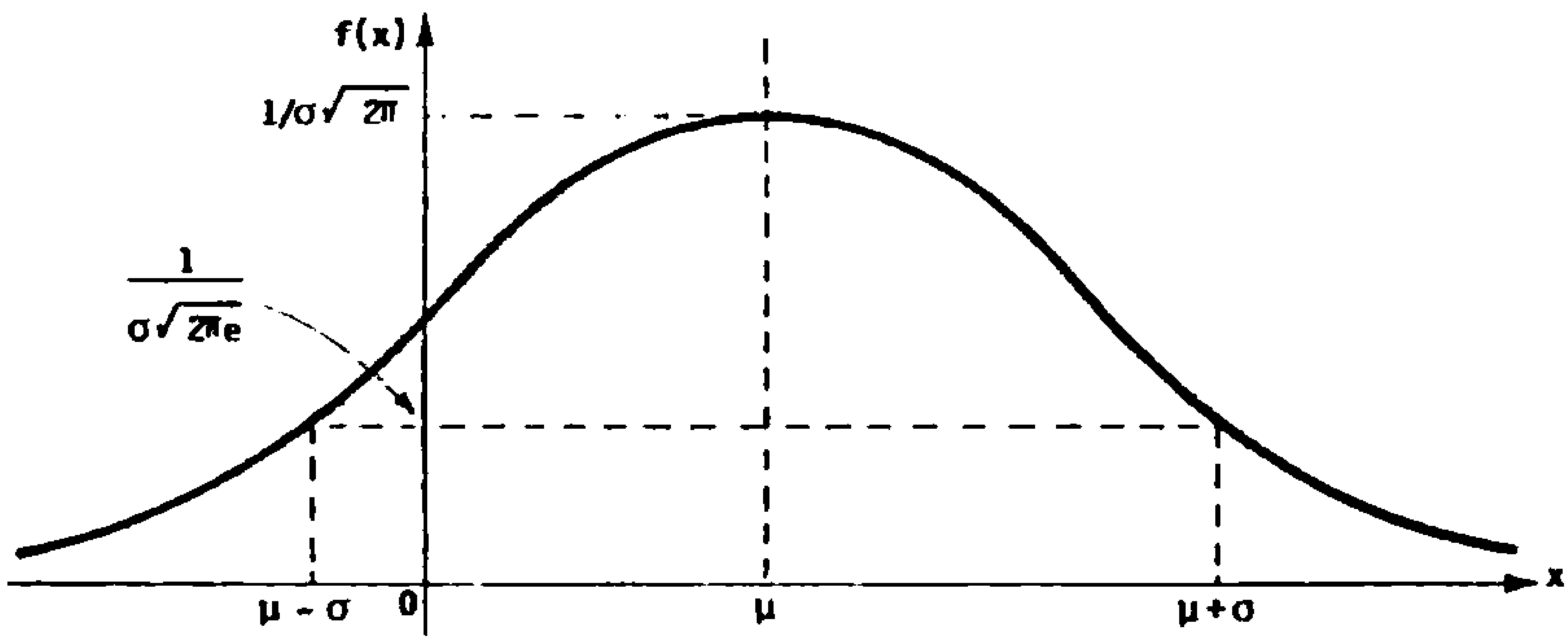


Fig. 6.3.1 La distribución normal.

Una propiedad importante de la distribución normal es su simetría respecto a la recta $x = \mu$. En efecto

$$f(\mu + y) = f(\mu - y) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(1/2)(y/\sigma)^2}$$

La función de distribución acumulada de la variable aleatoria **X** distribuída normalmente está dada por,

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(1/2)[(r-\mu)/\sigma]^2}\, dr$$

Esta integral no puede calcularse directamente, sin embargo, se puede representar por el área sombreada en la fig. 6.3.2 y su valor se determina con la ayuda de tablas como se verá en la sección 6.3.2. La gráfica de la función de distribución acumulativa se representa en la figura 6.3.3.

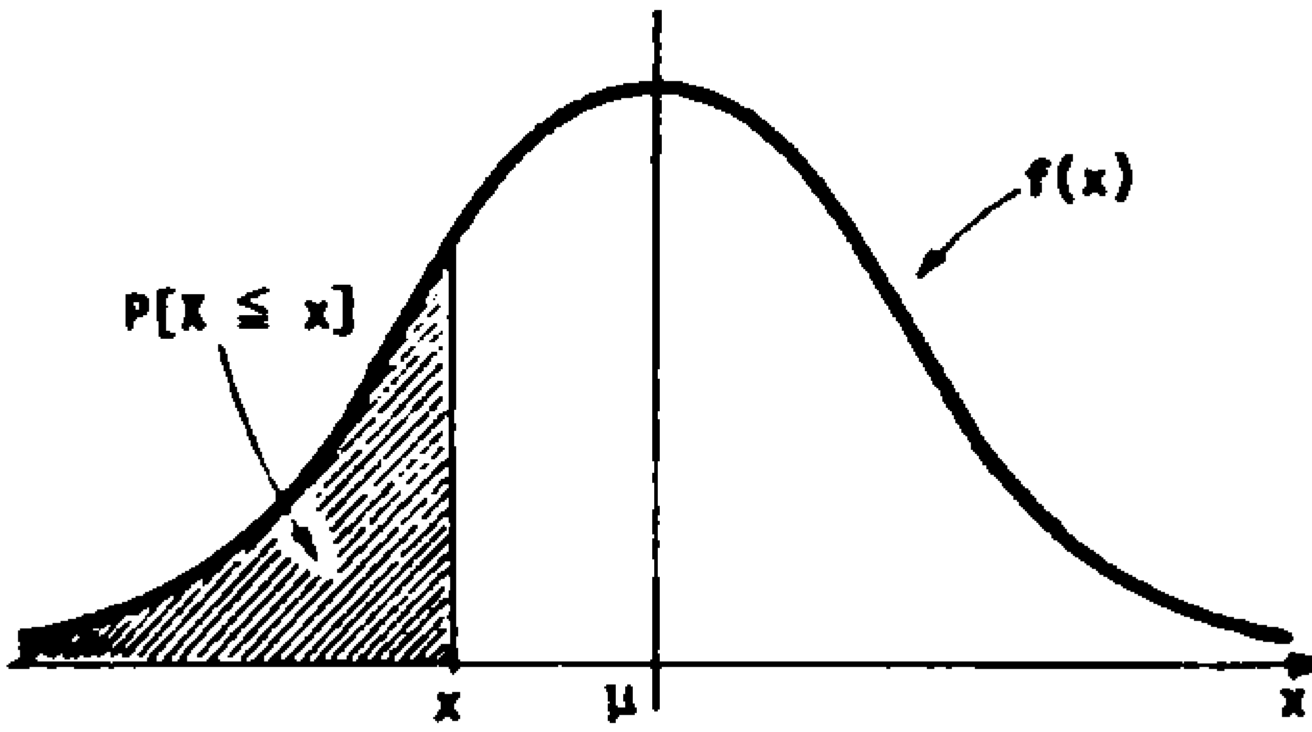


Fig. 6.3.2

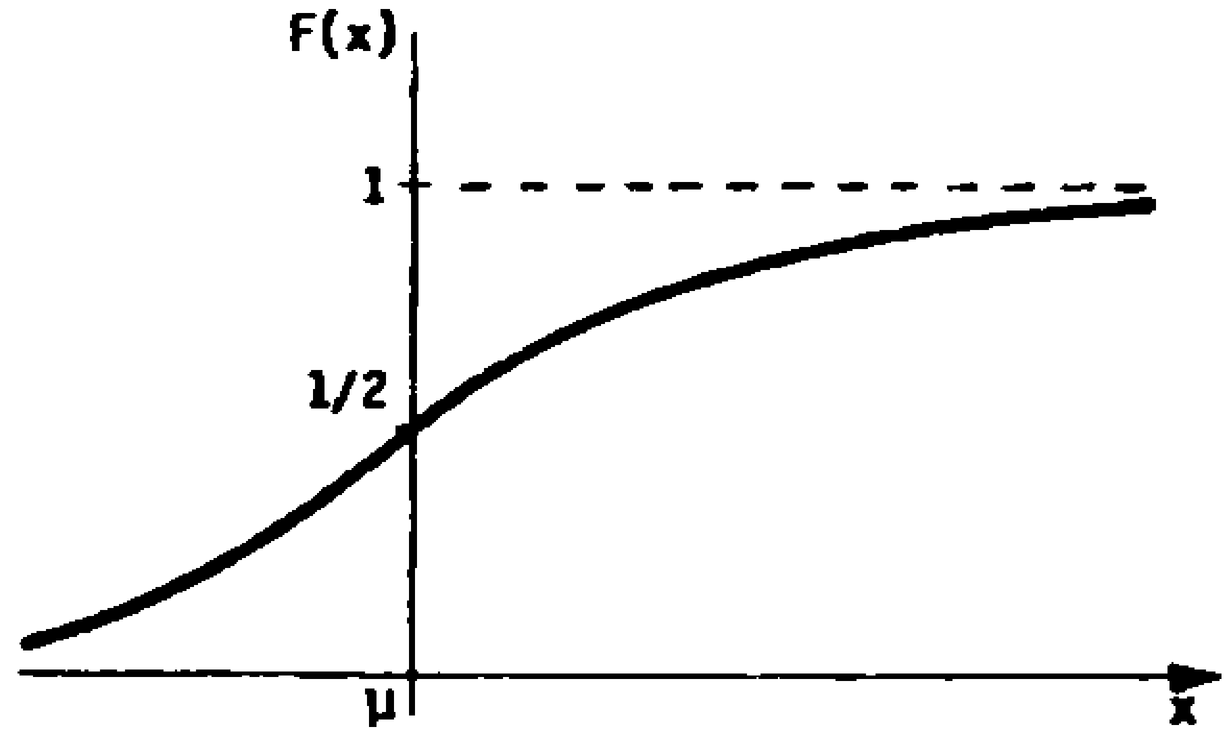


Fig. 6.3.3 Función de distribución

La *media* y la *varianza* de la distribución normal están dadas por,

$$E(X) = \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} \; e^{-(1/2)[(x-\mu)/\sigma]^2} \, dx$$

Sea $r = \dfrac{x-\mu}{\sigma}$, entonces $dx = \sigma dr$. Luego,

$$E(X) = \int_{-\infty}^{\infty} \frac{\sigma r + \mu}{\sqrt{2\pi}} \; e^{-r^2/2} \, dr = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} r e^{-r^2/2} \, dr + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-r^2/2} \, dr$$

pero $\dfrac{\sigma}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} r e^{-r^2/2} \, dr = \lim_{t \to \infty} \frac{-\sigma}{\sqrt{2\pi}} \; e^{-r^2/2} \Big|_{-t}^{t} = 0$ y

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \; e^{-r^2/2} \, dr = 1$$

Pues esta integral es la densidad de probabilidad normal con media $\mu = 0$ y varianza $\sigma^2 = 1$. (El valor de la integral es 1). Por lo tanto,

$$E(X) = 0 + \mu = \mu .$$

La varianza

$$\mathrm{Var}(X) = E(X - \mu)^2 = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma\sqrt{2\pi}} \; e^{-(1/2)[(x-\mu)/\sigma]^2} \, dx$$

haciendo $r = \dfrac{x-\mu}{\sigma}$, obtenemos $dx = \sigma dr$. Entonces,

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} \frac{\sigma^2 r^2}{\sqrt{2\pi}} \; e^{-r^2/2} \, dr = \sigma^2 \int_{-\infty}^{\infty} \frac{r^2}{\sqrt{2\pi}} \; e^{-r^2/2} \, dr$$

integrando por partes, se tiene

$$\mathrm{Var}(X) = \sigma^2 \left[\lim_{t \to \infty} \frac{-r e^{-r^2/2}}{\sqrt{2\pi}} \Bigg|_{-t}^{t} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-r^2/2} \, dr \right]$$

$$= \sigma^2 [\, 0 + 1 \,] = \sigma^2$$

Es decir, los dos parámetros que determinan completamente la distribución normal son, su media y su varianza.

Los valores medios pueden representarse en las gráficas de las funciones de densidad como se muestra en la fig. 6.3.4. Hemos visto que la función de densidad normal es simétrico respecto a su media.

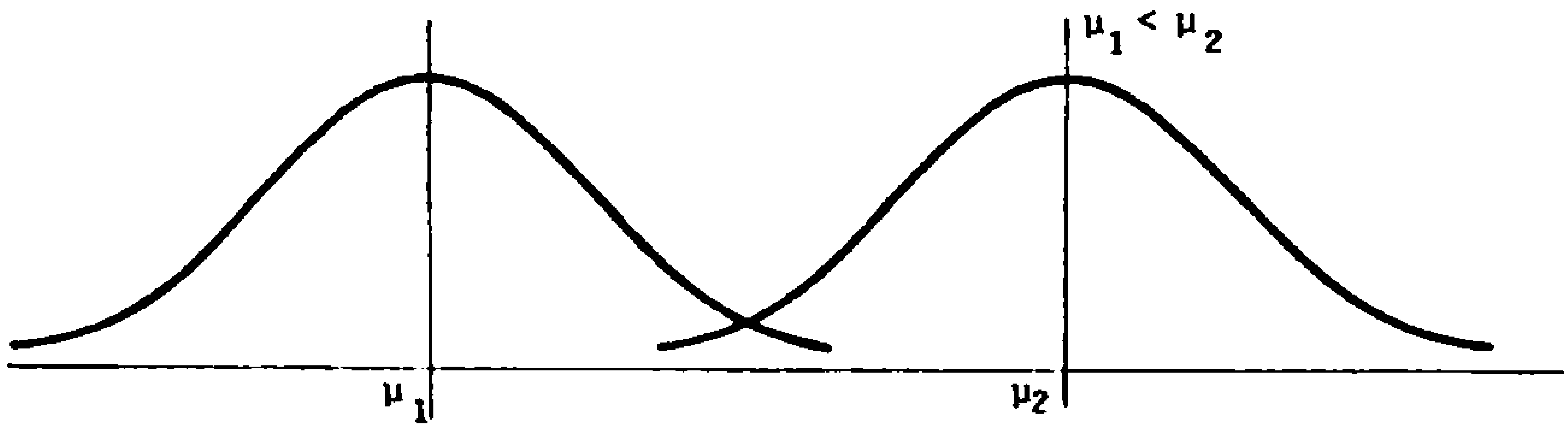


Fig. 6.3.4. Comparación de dos distribuciones normales con medias diferentes y varianzas iguales, $\mu_1 < \mu_2$ y $\sigma_1^2 = \sigma_2^2$

La varianza es una medida de la variabilidad o dispersión de la variable aleatoria. A mayor varianza, mayor variabilidad. Esto puede mostrarse gráficamente en la fig. 6.3.5. La desviación estándar, puede usarse para localizar los puntos de inflexión de la función de densidad. Hemos visto que los puntos de inflexión corresponden a $\mu - \sigma$ y $\mu + \sigma$.

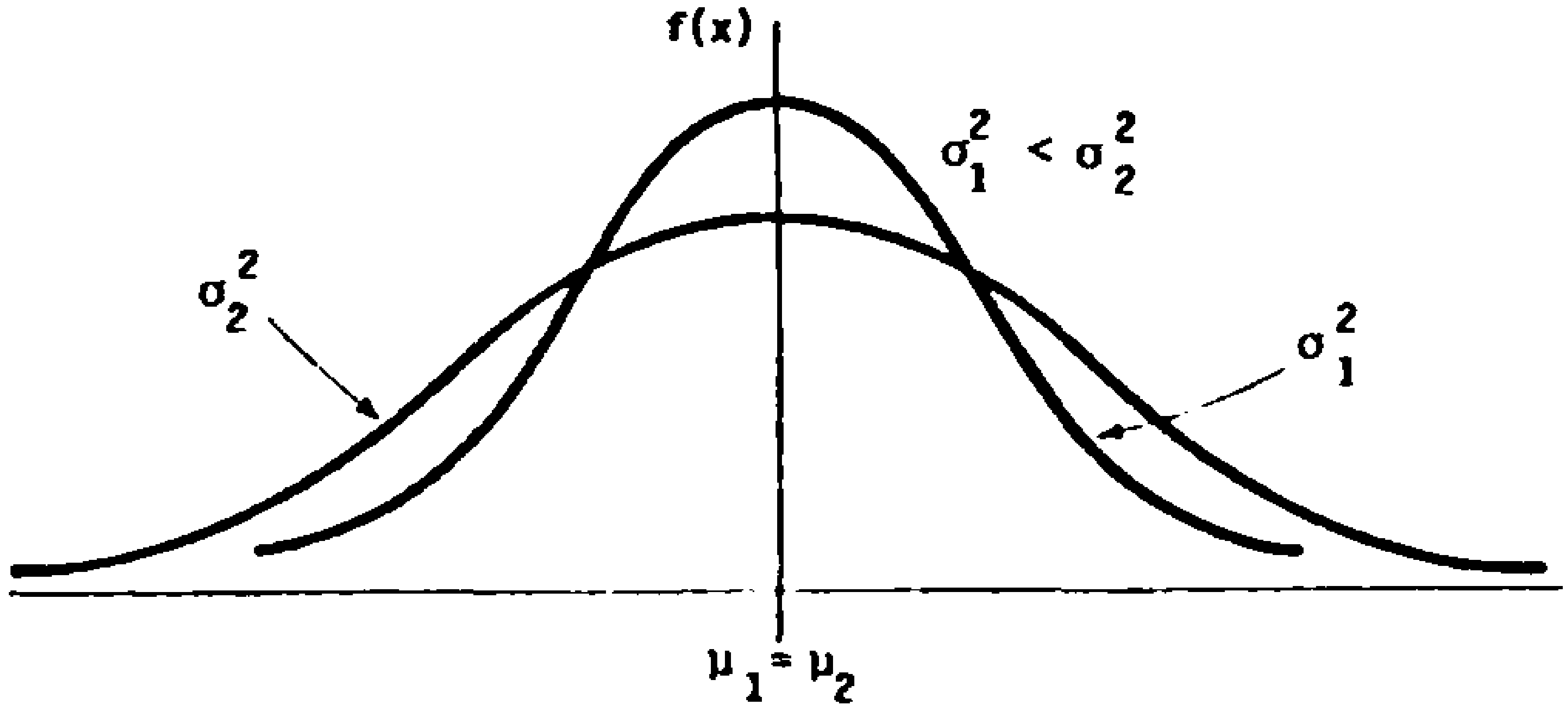


Fig. 6.3.5. Comparación de dos distribuciones normales con varianzas diferentes $\sigma_1^2 < \sigma_2^2$ y $\mu_1 = \mu_2$

6.3.1 DISTRIBUCION NORMAL ESTANDAR

DEFINICION 6.3.2 Si Z es una variable aleatoria que tiene una distribución normal, con media $\mu = 0$ y varianza $\sigma^2 = 1$, entonces Z se llama variable aleatoria normal estándar, su función de densidad es,

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

La función de distribución de Z, se denota por $\Phi(z)$ ó $N_Z(z)$ y está dado por,

$$\Phi(z) = P[Z \leq z] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

Los valores de esta función están tabulados y se da en la tabla III. En esta tabla, los valores de $\Phi(z)$ están dados para valores de z desde -3 a 3. Cualquier variable aleatoria X normal con media μ y varianza σ^2 puede ser transformado a una variable aleatoria estandarizada Z, por la siguiente transformación

$$Z = \frac{X - \mu}{\sigma}$$

Obviamente $\quad E\left(\frac{X - \mu}{\sigma}\right) = 0 \quad$ y $\quad Var\left(\frac{X - \mu}{\sigma}\right) = 1$

La relación intuitiva de la función de densidad de X y Z se muestra en la fig. 6.3.6.

Note que el 99.74% del área de una distribución normal está entre $\mu-3\sigma$ y $\mu + 3\sigma$. El área en ambas colas es despreciable. Por esta razón es que las tablas dan los valores de $\Phi(z)$ para valores de -3 a 3.

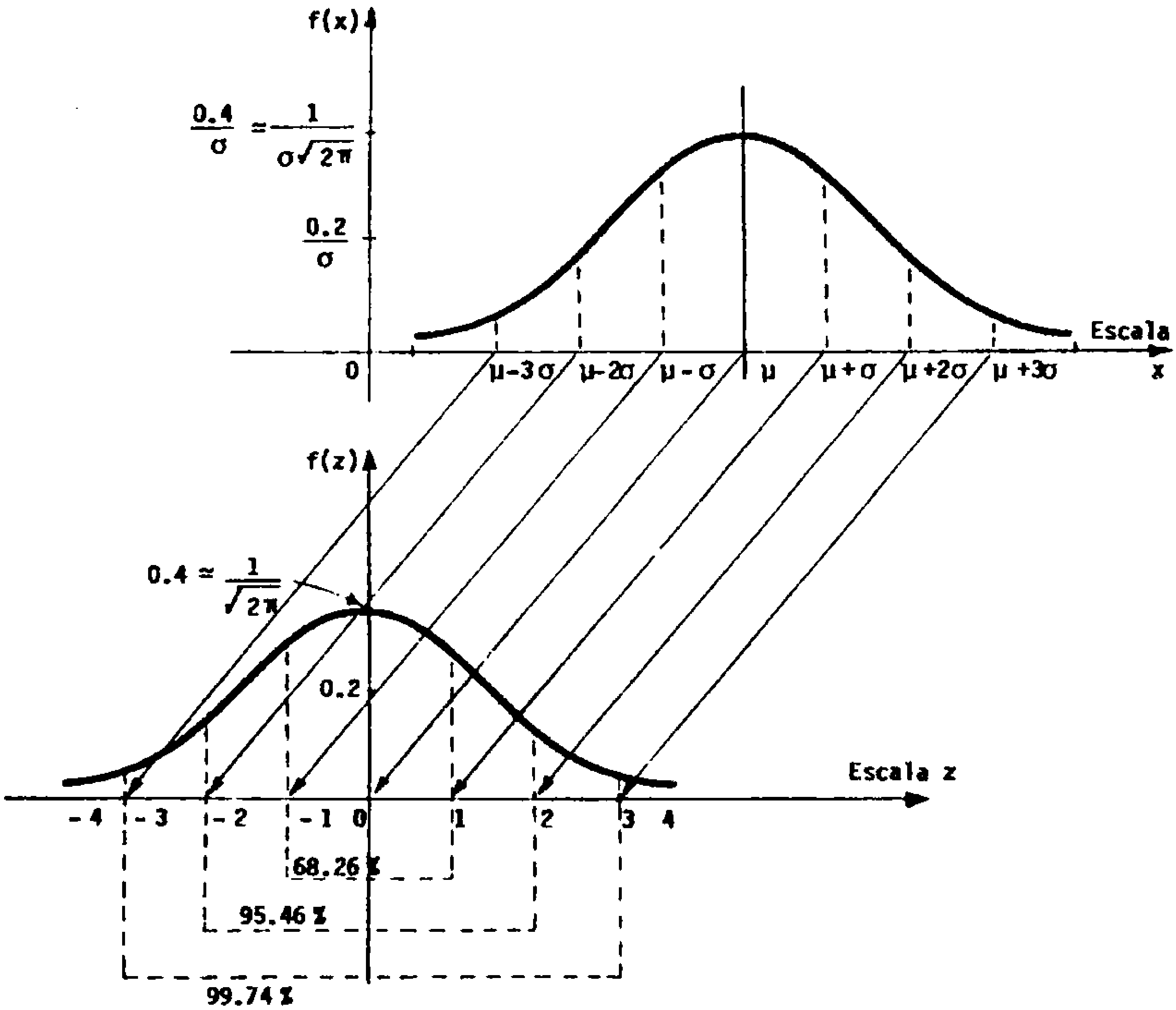


Fig. 6.3.6. Relación entre la función de densidad de X y Z

Observe que la escala vertical son diferentes para $f(x)$ y $f(z)$, pues los coeficientes de estas funciones, de densidad difieren por un factor σ, como se ve de sus respectivas definiciones. $f(z)$ toma su valor máximo en $z = 0$ y es igual a $\frac{1}{\sqrt{2\pi}}$. La función $f(x)$ toma su valor máximo en $x = \mu$ y es igual a

$\frac{1}{\sigma\sqrt{2\pi}}$. Recordar que estos son valores de la función de densidad, no valores de probabilidad.

El teorema siguiente justifica la transformación de cualquier variable aleatoria X con distribución normal a una normal estándar Z.

TEOREMA 6.3.1 Si X es una variable aleatoria con distribución normal con media μ y varianza σ^2, entonces $Z = \frac{X - \mu}{\sigma}$ es una variable aleatoria con distribución normal estándar.

DEMOSTRACION $X \to N(\mu,\sigma^2)$, entonces la función de distribución es,

$$F(x) = P[X \leq x] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(1/2)[(t - \mu)/\sigma]^2} dt$$

Si $Z = \frac{X - \mu}{\sigma}$, la función de distribución de Z está dado por

$$\Phi(z) = P[Z \leq z] = P\left[\frac{X - \mu}{\sigma} \leq z\right] = P[X \leq \mu + \sigma z] = F(\mu + \sigma z)$$

Luego,

$$\Phi(z) = P[Z \leq z] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu + \sigma z} e^{-(1/2)[(t - \mu)/\sigma]^2} dt$$

haciendo $r = \frac{t - \mu}{\sigma}$, obtenemos

$$\Phi(z) = P[Z \leq z] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{z} e^{-r^2/2} \sigma dr = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-r^2/2} dr$$

ya que $dt = \sigma dr$, y $r = z$, cuando $t = \mu + \sigma z$

6.3.2 USO DE TABLAS

Conocido los valores de la media y la varianza de una variable aleatoria X con distribución normal, es sólo cuestión de cálculo para encontrar probabilidades tales como:

(a) $P[X \leq b] = P\left[\frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right] = P\left[Z \leq \frac{b - \mu}{\sigma}\right] = \Phi\left[\frac{b - \mu}{\sigma}\right]$

(b) $P[X > b] = 1 - P[X \leq b] = 1 - \Phi\left[\frac{b - \mu}{\sigma}\right]$

(c) $P[a \leq X \leq b] = P\left[\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right]$

$$= P\left[\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right]$$

$$= \Phi\left[\frac{b - \mu}{\sigma}\right] - \Phi\left[\frac{a - \mu}{\sigma}\right]$$

Es decir, tal como hemos visto en 6.3.1 es suficiente una tabulación de la - función de distribución de la variable aleatoria normal estandar. Los valores dados en la tabla III, representa al área bajo la curva de la función de densidad de una variable aleatoria normal estandar Z desde menos infinito hasta z_α, o sea, el área sombreada que se muestra en la fig. 6.3.7. Y está definida por ,

$$\Phi(z_\alpha) = P[Z \leq z_\alpha] = \int_{-\infty}^{z_\alpha} \frac{1}{\sqrt{2\pi}} e^{-r^2/2} dr = \alpha = \text{área sombreada}$$

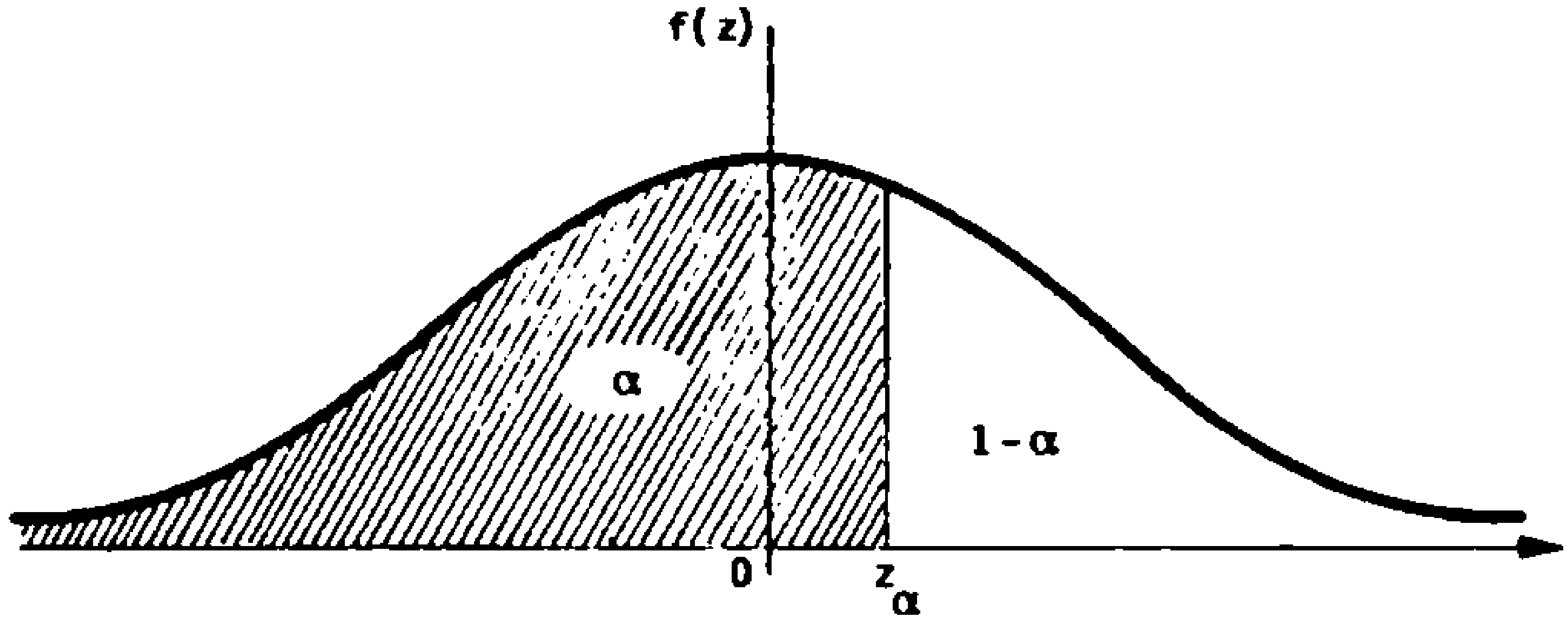


Fig. 6.3.7

Existe otra tabla, también muy usada que da $\Phi(z_\alpha) - 0.5$ cuya área bajo la curva está dada en la fig. 6.3.8, para valores positivos de z_α, y está definida por

$$P[0 \leq Z \leq z_\alpha] = \int_0^{z_\alpha} \frac{1}{\sqrt{2\pi}} e^{-r^2/2} dr = \alpha = \text{área sombreada}$$

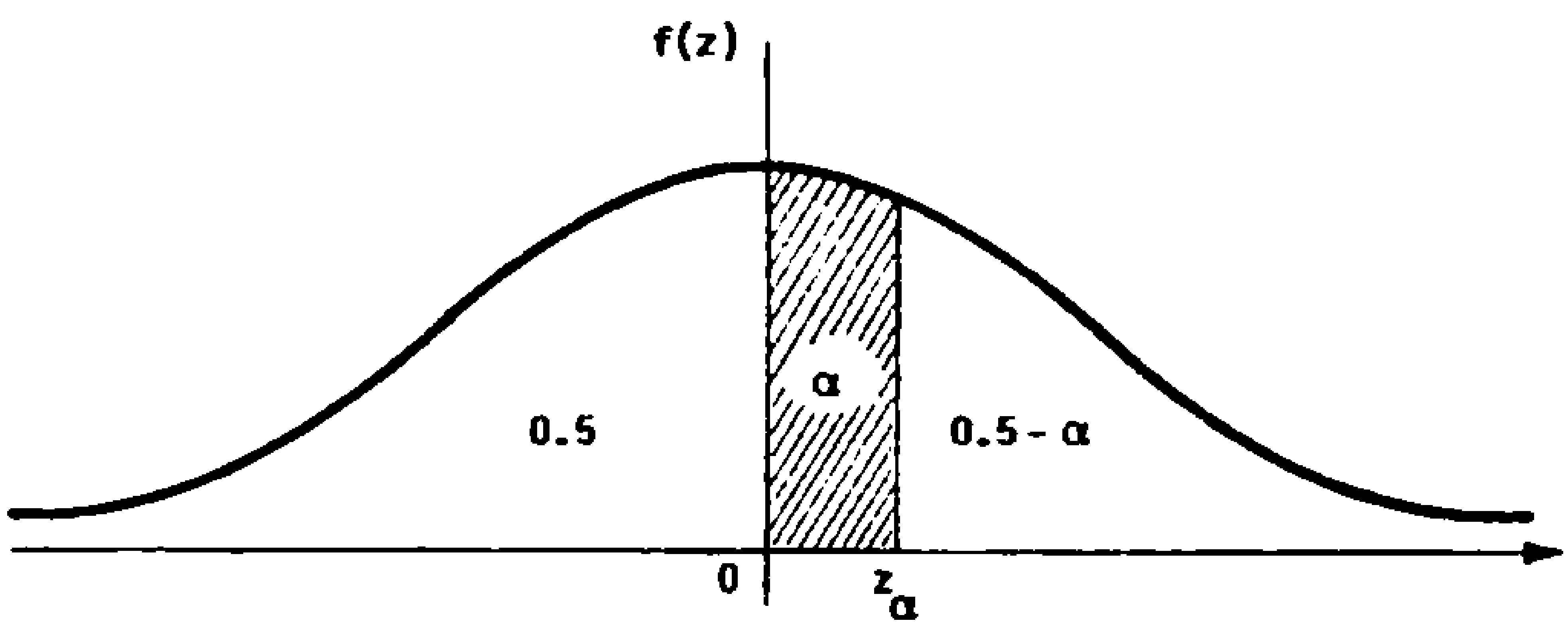


fig. 6.3.8

Otra tabla, es la que da el área bajo la curva de $f(z)$ desde z_α hasta mas infinito para valores positivos de z_α . Es decir, el área sombreada en la fig. 6.3.9 . En símbolos

$$P[Z \geq z_\alpha] = \int_{z_\alpha}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx = \alpha = \text{área sombreada}$$

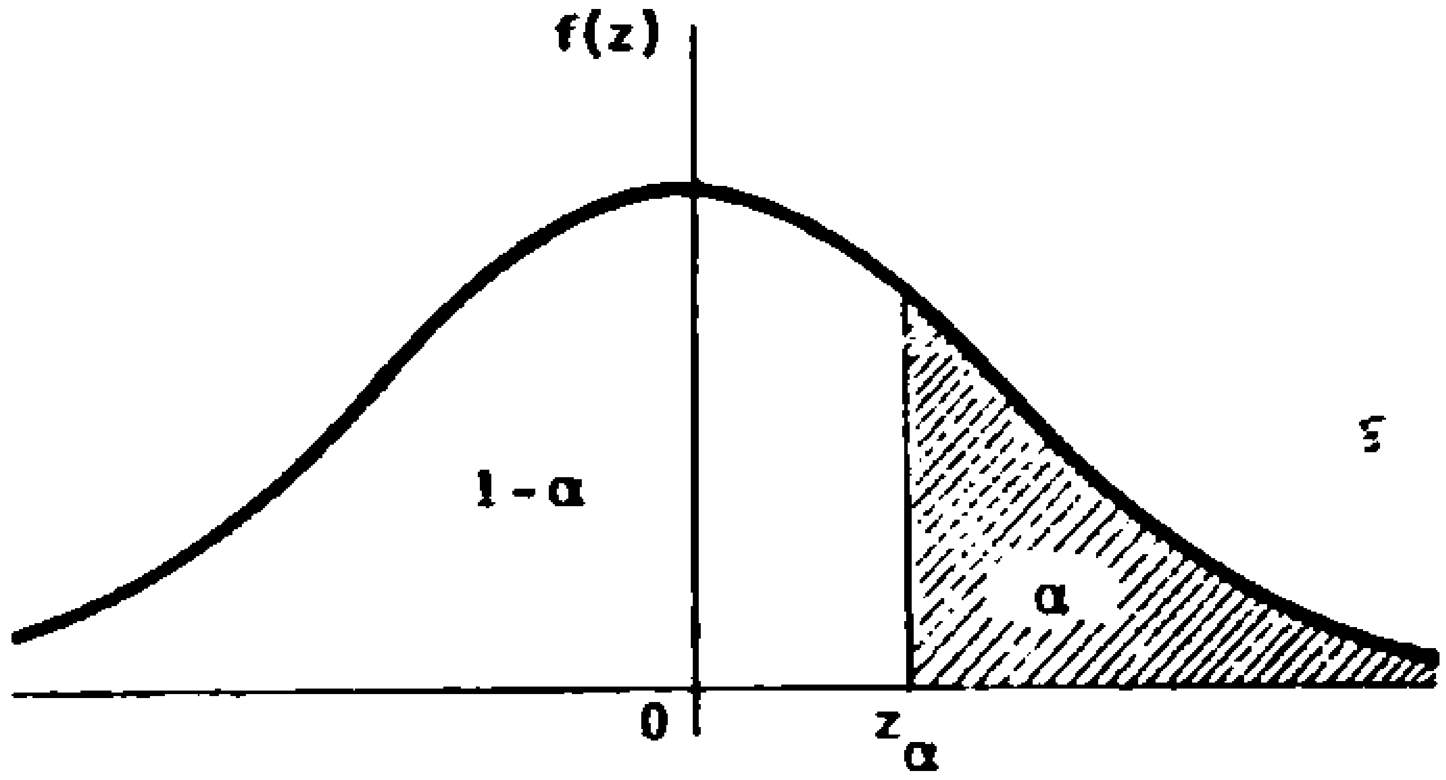


Fig. 6.3.9

El lector puede usar cualquier tipo de tabla, para calcular probabilidades, utilizando la simetría de la distribución normal y las probabilidades complementarias según sea el caso para las que no se encuentran directamente en la tabla. Por ejemplo, si disponemos de la tabla de la fig. 6.3.9, y estamos interesados en calcular $P[Z \leq -z_\alpha]$. Por simetría de la distribución normal es,

$$P[Z \leq -z_\alpha] = P[Z \geq z_\alpha] \qquad (1)$$

Como indica las áreas sombreadas en la figura 6.3.10 .

Por definición el primer miembro de (1) es

$$P[Z \leq -z_\alpha] = \Phi(-z_\alpha)$$

El segundo miembro de (1) se escribe

$$P[Z \geq z_\alpha] = 1 - P[Z \leq z_\alpha] = 1 - \Phi(z_\alpha)$$

Luego,

$$\Phi(-z_\alpha) = 1 - \Phi(z_\alpha)$$

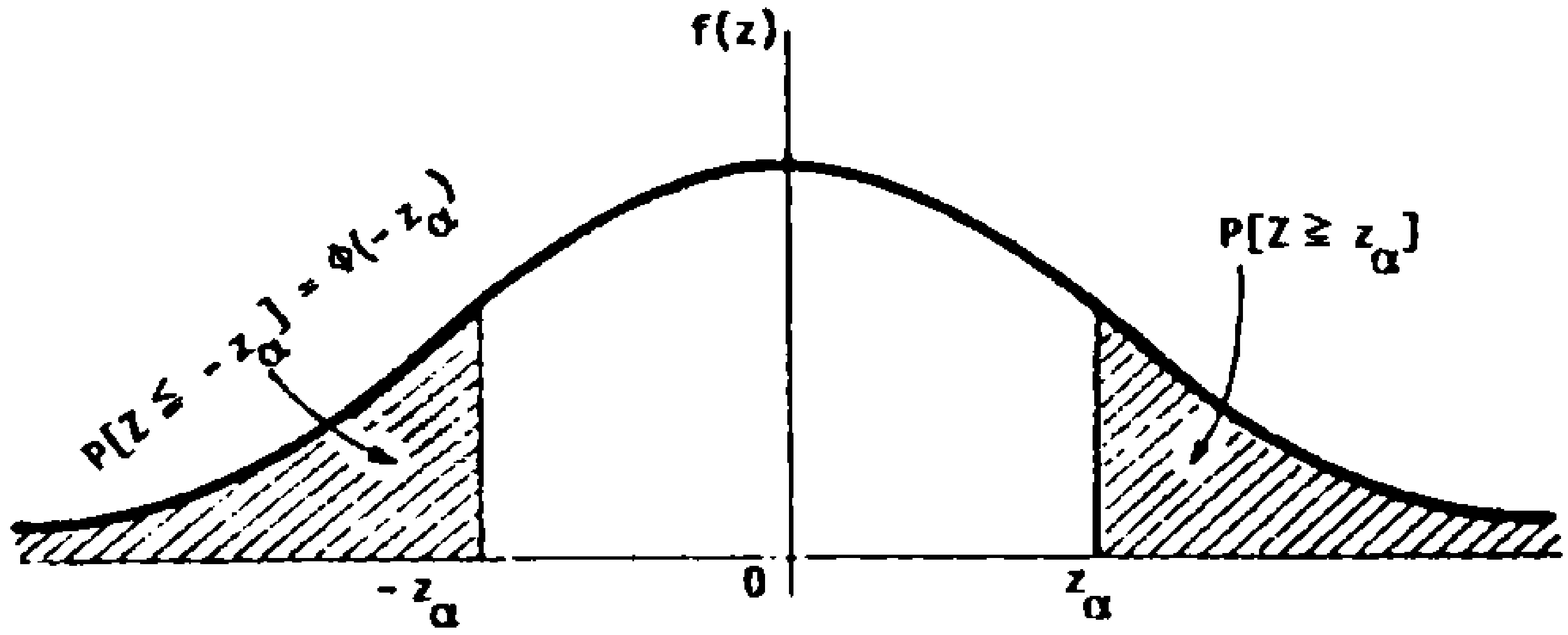


Fig. 6.3.10

LECTURA DE LA TABLA III Se debe tener en cuenta lo siguiente:

1. La tabla proporciona área bajo la curva normal estandar desde $-\infty$ hasta z es decir área correspondiente a $P[Z \leq z] = \Phi(z)$ (ver fig. 6.3.7)

2. Los valores de z están dados en centésimos (o sea con dos decimales) desde -3.49 hasta 3.49. Por ejemplo $z = 1.86$

3. Es una tabla con dos entradas encabezado por la letra z. En la primera columna (primera entrada) se lee el valor de z. En décimos (es decir con un decimal), por ejemplo 1.8. En la primera fila se lee al centecimal (segundo decimal de z) 0.06, luego $z = 1.86$.

4. La tabla ayuda resolver dos tipos de problemas:

   (a) conocido z hallar el área (b) conocido el área hallar z.

   (a) z conocido, digamos $z = 1.86$, es decir queremos calcular $P[Z \leq 1.86]$. En la primera columna se ubica el valor de z con un decimal 1.8 y el segundo decimal se ubica en la primera fila 0.06 por ambos puntos se traza una recta horizontal y una vertical respectivamente, el número - que corresponde a la intersección de ambas rectas 0.9686 (ver cuadro 1) es el área deseado, es decir $P[Z \leq 1.86] = \Phi(1.86) = 0.9686$.

   (b) Es un procedimiento inverso al anterior. Se ubica el área en el cuerpo de la tabla, por este punto se traza una horizontal y una vertical. z se obtiene sumando el punto de la intercección de la horizontal con la primera columna con el punto de la intersección de la vertical con la primera fila. Por ejemplo se pide calcular z tal que

   $P[Z \leq z] = 0.9382$. En efecto:

   Se ubica el número 0.9382 (ver cuadro 1) en el cuerpo de la tabla. Por este punto se traza un horizontal y una vertical, las que intersectan a la primera columna y primera fila en 1.5 y 0.04 respectivamente, luego $z = 1.5 + 0.04 = 1.54$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
⋮										
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
⋮										

Cuadro 1. Estracto de la tabla III.

EJEMPLO 1 Sea X una variable aleatoria N(5,4). ¿Cuál es la probabilidad de que X tome valores entre 4 y 7? ¿Cuál es la probabilidad que tome valores mayores que *10*?

SOLUCION

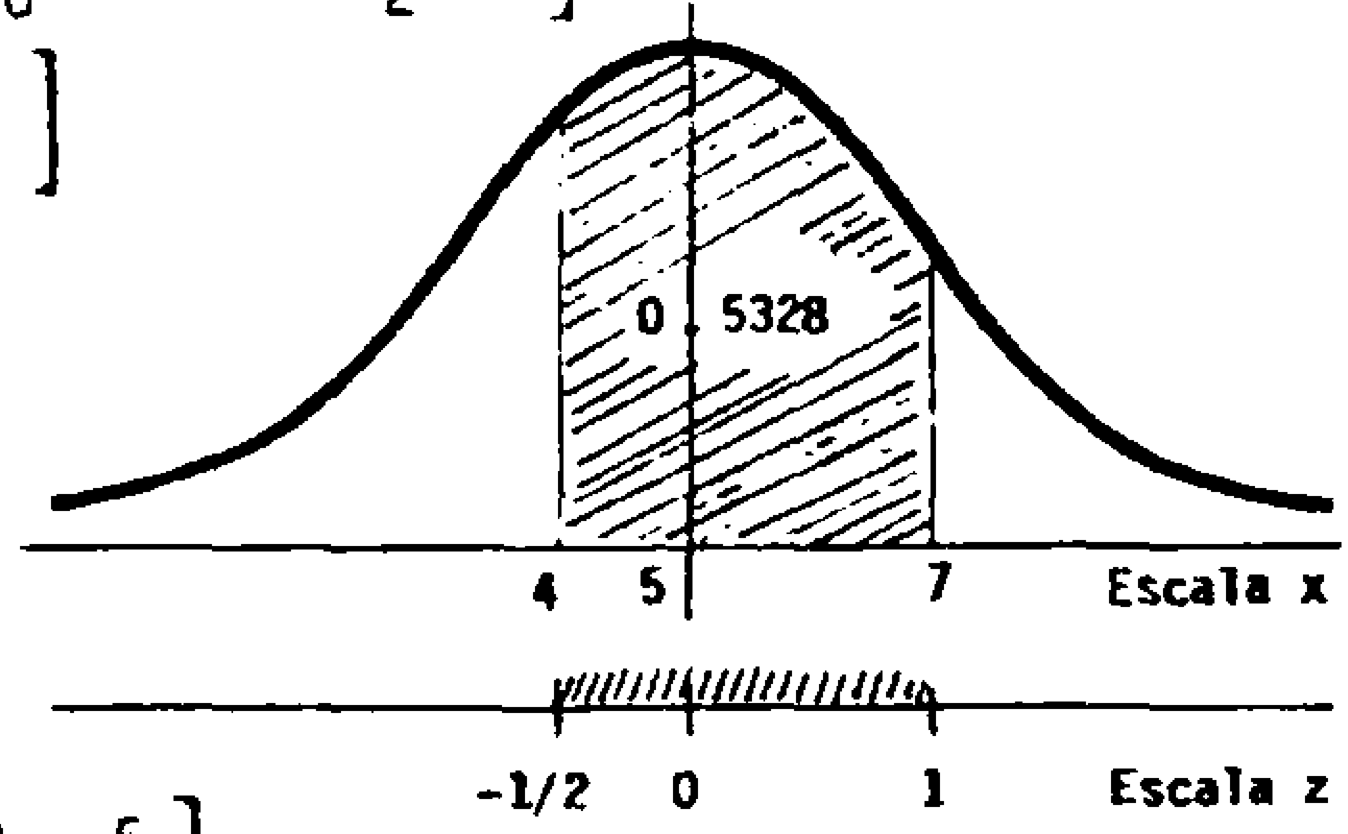
(a) $P[4 < X < 7] = P\left[\frac{4-5}{2} < \frac{X-\mu}{\sigma} < \frac{7-5}{2}\right]$ (teorema 6.3.1)

$= P\left[-\frac{1}{2} < Z < 1\right]$

$= \Phi(1) - \Phi(-1/2)$

$= 0.8413 - 0.3085$

$= 0.5328$.



(b) $P[X > 10] = 1 - P[X \leq 10]$

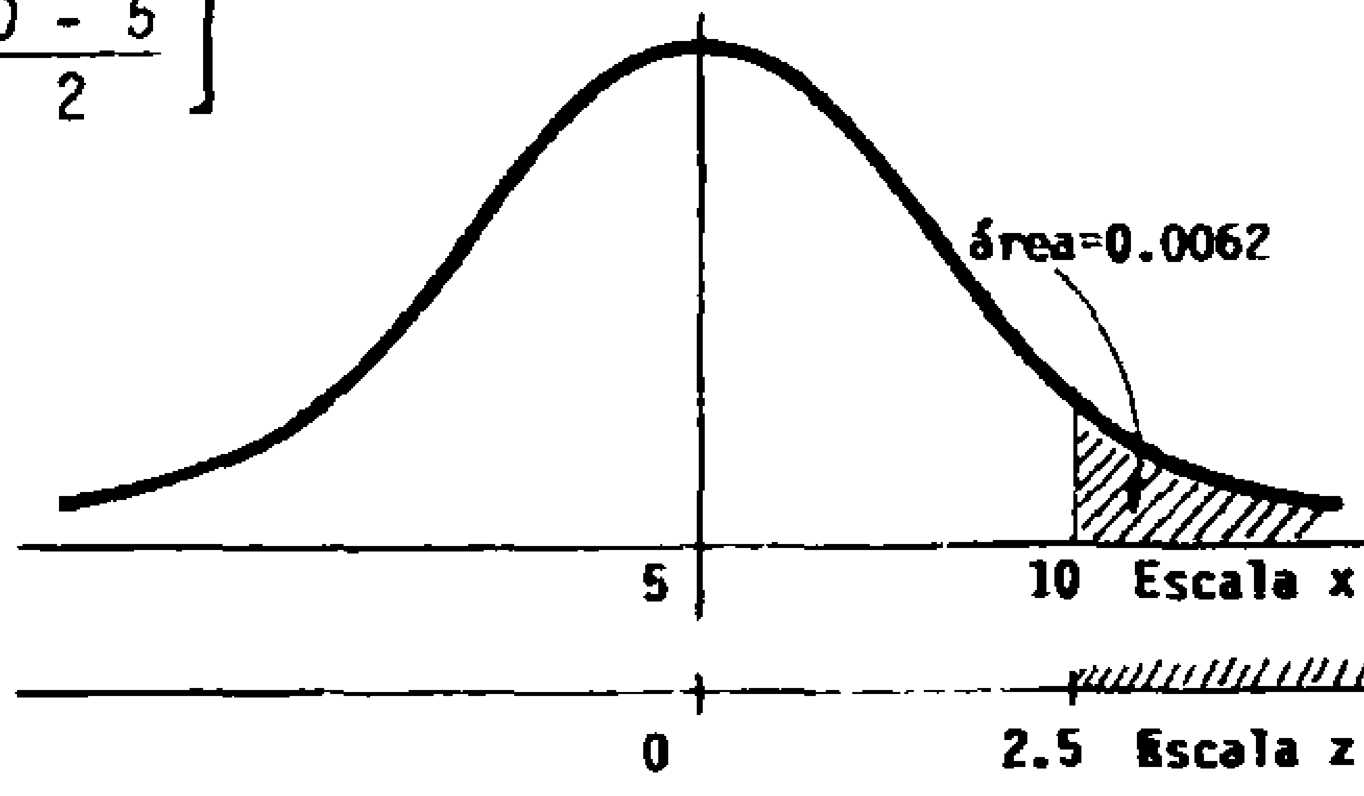
$= 1 - P\left[\frac{X-\mu}{\sigma} \leq \frac{10-5}{2}\right]$

$= 1 - P\left[Z \leq \frac{5}{2}\right]$

$= 1 - \Phi(2.5)$

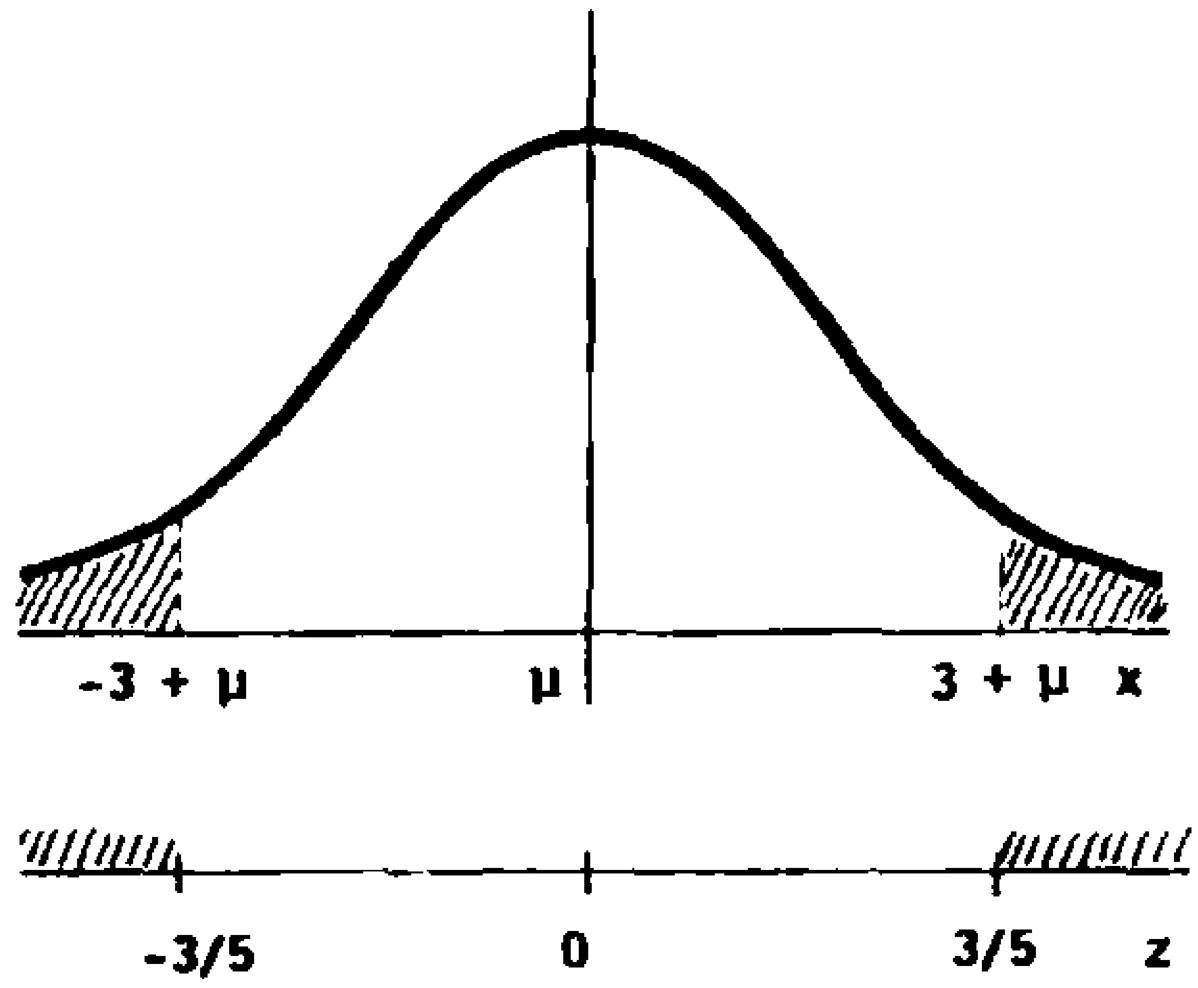
$= 1 - 0.9938$

$= 0.0062$.



EJEMPLO 2 Sea X una variable aleatoria $N(\mu, 25)$. Calcular

$$P[|X - \mu| > 3]$$



SOLUCION

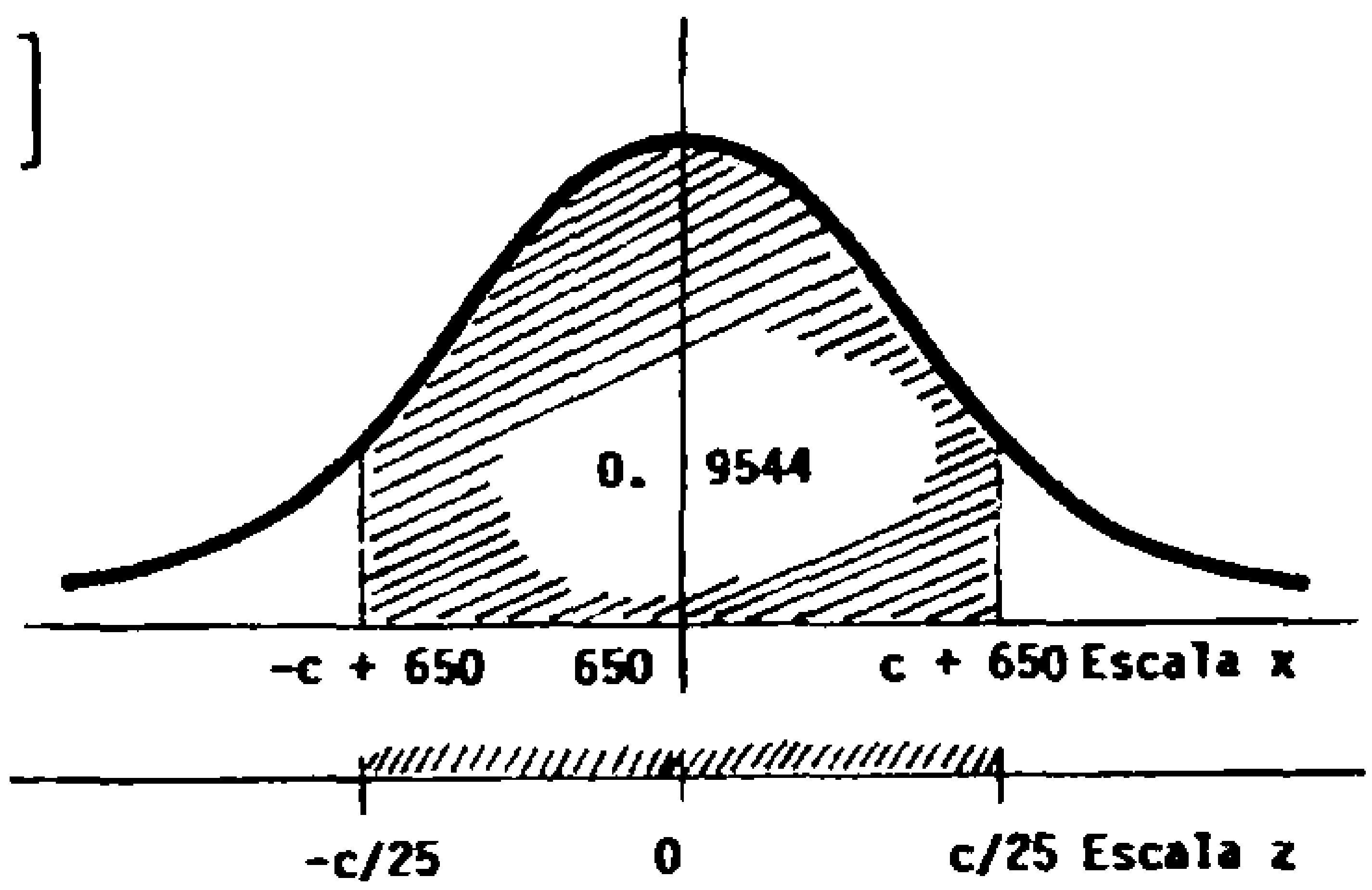
$$\begin{aligned} P[|X - \mu| > 3] &= 1 - P[|X - \mu| \leqslant 3] \\ &= 1 - P[-3 \leqslant X - \mu \leqslant 3] \\ &= 1 - P\left[-\frac{3}{5} \leqslant \frac{X - \mu}{\sigma} \leqslant \frac{3}{5}\right] \\ &= 1 - P[-0.6 \leqslant Z \leqslant 0.6] \\ &= 1 - [\Phi(0.6) - \Phi(-0.6)] \\ &= 1 - 0.7557 + 0.2743 \\ &= 0.5186. \end{aligned}$$

EJEMPLO 3 Si X es una variable aleatoria $N(650, 625)$. Hallar la constante $c > 0$ tal que,

$$P[|X - 650| \leqslant c] = 0.9544 \cdot$$

SOLUCION

$$\begin{aligned} 0.9544 &= P[|X - 650| \leqslant c] = P[-c \leqslant X - 650 \leqslant c] \\ &= P\left[-\frac{c}{25} \leqslant \frac{X - 650}{25} \leqslant \frac{c}{25}\right] \\ &= P\left[-\frac{c}{25} \leqslant Z \leqslant \frac{c}{25}\right] \\ &= \Phi\left(\frac{c}{25}\right) - \Phi\left(-\frac{c}{25}\right) \\ &= \Phi\left(\frac{c}{25}\right) - \left[1 - \Phi\left(\frac{c}{25}\right)\right] \\ 0.9544 &= 2\Phi\left(\frac{c}{25}\right) - 1 \end{aligned}$$



de donde $\Phi\left(\frac{c}{25}\right) = 0.9772$, y de la tabla III obtenemos

$$\frac{c}{25} = 2.0, \qquad \text{luego} \qquad c = 50 .$$

EJEMPLO 4 En una distribución normal se tienen los siguientes datos

$$P[X < 45] = 0.31 ; \qquad P[X > 64] = 0.08$$

Hallar la media y la desviación típica de la distribución

SOLUCION

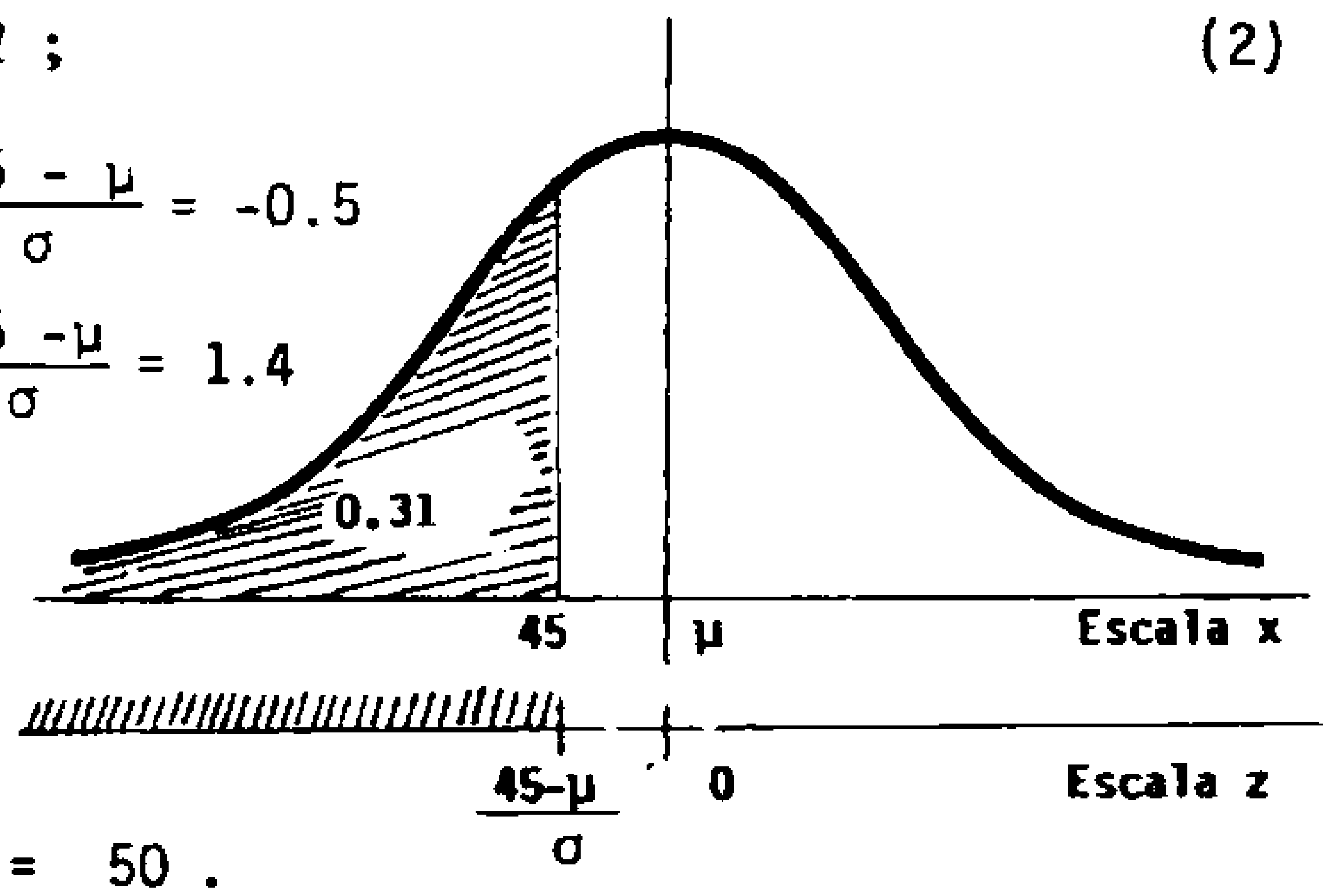
$$P[X < 45] = P\left[\frac{X-\mu}{\sigma} < \frac{45-\mu}{\sigma}\right] = P\left[Z < \frac{45-\mu}{\sigma}\right]$$

$$= \Phi\left[\frac{45-\mu}{\sigma}\right] = 0.31 \qquad (1)$$

$$P[X > 64] = 1 - P[X \leq 64] = 1 - P\left[\frac{X-\mu}{\sigma} \leq \frac{64-\mu}{\sigma}\right]$$

$$= 1 - \Phi\left[\frac{64-\mu}{\sigma}\right] = 0.08$$

ó $$\Phi\left[\frac{64-\mu}{\sigma}\right] = 0.92 \; ; \qquad (2)$$



de (1) y la tabla III obtenemos, $\frac{45-\mu}{\sigma} = -0.5$

de (2) y la tabla III obtenemos, $\frac{45-\mu}{\sigma} = 1.4$

de donde $$\begin{aligned} -45 &= -\mu + 0.05\sigma \\ 64 &= \mu + 1.4\sigma \\ \hline 19 &= 1.9\sigma \end{aligned}$$

Luego, $\sigma = 10$ y $\mu = 50$.

EJEMPLO 5 Una pequeña ciudad es abastecida de agua cada 2 días; el consumo en volumen de agua para esa pequeña ciudad tiene una distribución normal con media 20,000 litros y desviación típica 1,000 litros (se entiende el consumo cada 2 días). Se trata de hallar la capacidad de su tanque de agua para que sea de sólo 0.01, la probabilidad que en un período de 2 días el agua no sea suficiente para satisfacer toda la demanda.

SOLUCION Sea **X** la variable aleatoria definida por

$X(\omega)$ = consumo de agua en un período de dos días.

X se distribuye normalmente con μ = 20,000 litros y σ = 1,000 litros. Si c es la capacidad de su tanque, se tiene

$$0.01 = P[X > c] = P\left[\frac{X-\mu}{\sigma} > \frac{c - 20{,}000}{1{,}000}\right]$$

$$= 1 - P\left[Z \leq \frac{c-20{,}000}{1{,}000}\right] = 1 - \Phi\left[\frac{c-20{,}000}{1{,}000}\right]$$

de donde, $$\Phi\left[\frac{c-20{,}000}{1{,}000}\right] = 0.99$$

de la tabla III obtenemos

$$\frac{c - 20{,}000}{1{,}000} = 2.33$$

ó $$c = 2{,}330 + 20{,}000 = 22{,}330$$

Es decir, la capacidad del tanque de agua de la ciudad debe ser de 22,330 litros .

EJEMPLO 6 Si la distribución de los períodos de duración de los postes telefónicos de madera es tal que el 9.51% tienen período de duración que exceden los 15 años y que el 62.55% tienen períodos de duración que exceden los 9 años, determinar la desviación estándar o desviación típica de los períodos de duración si se sabe que la distribución de dichos períodos es normal.

SOLUCION Sea **X** = período de duración de los postes telefónicos de madera.

$X \to N(\mu, \sigma^2)$, debemos calcular μ y σ. Sabemos que,

$$0.0951 = P[X > 15] = P\left[\frac{X-\mu}{\sigma} > \frac{15-\mu}{\sigma}\right]$$
$$= 1 - P\left[Z \leq \frac{15-\mu}{\sigma}\right]$$

luego, $$\Phi\left[\frac{15-\mu}{\sigma}\right] = 0.9049$$

de la tabla III se obtiene $$\frac{15-\mu}{\sigma} = 1.31\ ; \qquad (1)$$

$$0.6255 = P[X > 9] = 1 - P[X \leq 9] = 1 - P\left[\frac{X-\mu}{\sigma} \leq \frac{9-\mu}{\sigma}\right]$$

o sea $$\Phi\left[\frac{9-\mu}{\sigma}\right] = 0.3745$$

luego, de la tabla III $$\frac{9-\sigma}{\sigma} = -0.32 \qquad (2)$$

Resolviendo el sistema de ecuaciones (1) y (2) obtenemos,

$$\sigma = \frac{6}{1.63} = 3.68 \quad \text{y} \quad \mu = 9 + (0.32)\left(\frac{6}{1.63}\right) \simeq 10.18$$

EJEMPLO 7 La fábrica de neumáticos "DURAMAS" produce un tipo de neumáticos que tiene una vida útil media de 80,000 km. y una desviación estándar de 8,000 km. Suponiendo que esta vida útil está distribuída normalmente:

(a) ¿Cuál es la probabilidad que un neumático dure más de 96,000 km.?

(b) El 50% de los neumáticos duran entre x_1 y x_2 kilométros. Hallar los valores de x_1 y x_2, si ellos son simétricos respecto a la media.

(c) El fabricante garantiza que reemplazará gratis cualquier neumático cuya duración sea inferior a x. Determinar el valor de x de modo que tenga que reemplazar sólo el 1% de los neumáticos.

SOLUCION X = vida útil de un neumático en km.

$X \to N(80{,}000, \quad (8{,}000)^2) = N(\mu, \sigma^2)$

(a) $P[X > 96{,}000] = 1 - P[X \leq 96{,}000]$

$$= 1 - P\left[\frac{X-\mu}{\sigma} \leq \frac{96{,}000 - 80{,}000}{8{,}000}\right]$$

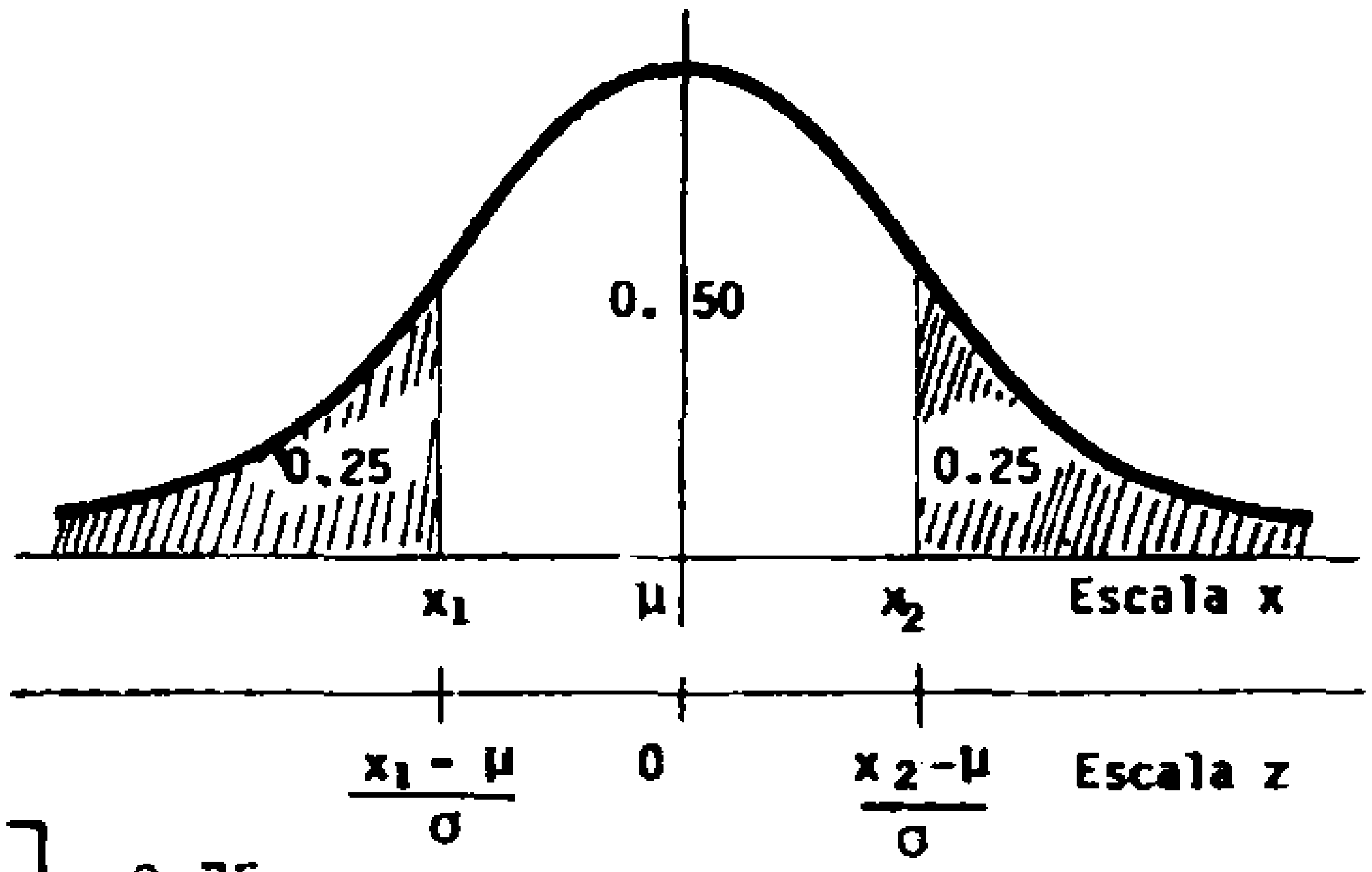
$$= 1 - P[Z \leq 2] = 1 - 0.9772 = 0.0228$$

(b) $P[x_1 \leq X \leq x_2] = 0.5$, además, x_1 y x_2, simétricos respecto a la media. Por lo tanto

$P[X < x_1] = 0.25$ y $P[X \leq x_2] = 0.75$ ó

$$P\left[\frac{X-\mu}{\sigma} \leq \frac{x_1-\mu}{\sigma}\right] = 0.25$$

$$\Phi\left[\frac{x_1-\mu}{\sigma}\right] = 0.25$$



de la tabla III obtenemos

$$\frac{x_1 - 80{,}000}{8{,}000} = -0.67$$

$$x_1 = 74{,}640 \text{ km.}$$

$$P[X \leq x_2] = P\left[\frac{X-\mu}{\sigma} \leq \frac{x_2-\mu}{\sigma}\right] = 0.75$$

ó $\Phi\left[\frac{x_2-\mu}{\sigma}\right] = 0.75$

de donde, $\dfrac{x_2 - 8{,}000}{8{,}000} = 0.67$

$$x_2 = 85{,}360.$$

(c) $P[X < x] = P\left[\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right] = 0.01$ ó

$$\Phi\left[\frac{x-\mu}{\sigma}\right] = 0.01$$

de la tabla III se obtiene

$$\frac{x - 80{,}000}{8{,}000} = -2.33$$

$$x = 61{,}360.$$

EJEMPLO 8 El tiempo T requerido para la digestión de una unidad de alimento por un protozoario es una variable aleatorio normal con media 31 minutos y desviación estándar 5 minutos.

(a) ¿Cuál es la probabilidad que una unidad de alimento es digerido en me nos de 35 minutos?

(b) Si una unidad particular de alimento se observa que no está digerido completamente en 30 minutos. ¿Cuál es la probabilidad que sea digerido antes de los 35 minutos?

SOLUCION

T = tiempo requerido para la digestión de una unidad de alimento.

$T \to N(\mu,\sigma) = N(31,5)$

(a) $$P[T < 35] = P\left[\frac{T-\mu}{\sigma} < \frac{35-31}{5}\right] = P[Z < 4/5] = \Phi(0.8)$$
$$= 0.7881.$$

(b) $$P[T < 35 \mid T > 30] = \frac{P[30 < T < 35]}{P[T > 30]}$$

El numerador

$$P[30 < T < 35] = P\left[\frac{30-31}{5} < \frac{T-\mu}{\sigma} < \frac{35-31}{5}\right]$$
$$= P[-0.2 < Z < 0.8] = \Phi(0.8) - \Phi(-0.2)$$
$$= 0.7881 - 0.4207 = 0.3674.$$

El denominador

$$P[T > 30] = 1 - P[T \leq 30] = 1 - P\left[\frac{T-\mu}{\sigma} \leq \frac{30-31}{5}\right]$$
$$= 1 - \Phi(-0.2) = 1 - 0.4207 = 0.5793.$$

Por lo tanto,

$$P[T < 35 \mid T > 30] = \frac{0.3674}{0.5793} \simeq 0.63 .$$

EJEMPLO 9 Los registros de pérdida de peso por evaporación de cierto producto empacado muestran una pérdida media de 6.45 gramos con una desviación estándar de 1.30. Asumiendo una distribución normal. ¿Cuál es la probabilidad de que si se extraen dos paquetes al azar de un lote ambas muestran una pérdida de más de 8.00 gramos?

SOLUCION Sea X = peso del primer paquete.

Y = peso del segundo paquete.

ambas variables aleatorias se distribuyen normalmente con $\mu = 6.45$ gr. y $\sigma = 1.30$ gr. Teniendo en cuenta que X e Y son independientes debemos determinar

$$
\begin{aligned}
P[(X > 8.00) \cap (Y > 8.00)] &= P[X > 8.00]\, P[Y > 8.00] \\
&= P\left[\frac{X-\mu}{\sigma} > \frac{8 - 6.45}{1.3}\right] P\left[\frac{Y-\mu}{\sigma} > \frac{8-6.45}{1.3}\right] \\
&= \left[P\left[Z > \frac{8-6.45}{1.3}\right]\right]^2 = [1 - \Phi(1.19)]^2 \\
&= (0.1170)^2 = 0.0137 .
\end{aligned}
$$

EJEMPLO 10 En el problema anterior. ¿Cuál es la probabilidad al seleccionar cinco paquetes que, por lo menos, uno de ellos muestre pérdida de más de 8.00 gramos?

SOLUCION Y = número de paquetes que muestran una pérdida de peso más de 8.00 gr. en los 5 paquetes.

$$R_Y = \{0,1,2,3,4,5\}$$

$$p = 0.1170 \qquad y \qquad n = 5$$

Y tiene una distribución binomial. Pues observe que el experimento es extraer un paquete y observar su peso

$$E = [X > 8.0], \qquad F = [X < 8.0] \qquad y \qquad P[E] = 0.1170$$

La distribución de probabilidad de Y es;

$$P[Y = y] = \binom{5}{y}(0.1170)^y(0.8830)^{5-y} \quad , \qquad y = 0,1,2,\ldots,5$$

Luego,
$$
\begin{aligned}
P[Y \geq 1] &= 1 - P[Y = 0] \\
&= 1 - \binom{5}{0}(0.1170)^0(0.8830)^5 \\
&= 1 - (0.8830)^5 = 0.4632 .
\end{aligned}
$$

EJEMPLO 11 Una característica cuantitativa X se distribuye normalmente con media $\mu = 50$ y desviación típica $\sigma = 10$. Una segunda característica Y, independiente de la primera se distribuye normalmente con $\mu = 20$ y $\sigma = 5$. En una muestra aleatoria de un elemento de cada población. ¿Cuál es la probabilidad de obtener un valor de X mayor que 60, al mismo tiempo que Y menor que 17?

SOLUCION

$X \to N(50,10^2)$; $\qquad Y \to N(20,5^2)$, e independientes. Se pide

$$
\begin{aligned}
P[(X > 60) \cap (Y < 17)] &= P[X > 60]\, P[Y < 17] \\
&= P\left[\frac{X-\mu}{\sigma} > \frac{60-50}{10}\right] P\left[\frac{Y-\mu}{\sigma} < \frac{17-20}{5}\right] \\
&= [1 - P[Z \leq 1]]\, P[Z \leq -\tfrac{3}{5}] = [1 - \Phi(1)]\,\Phi(-0.6)
\end{aligned}
$$

$$= (1 - 0.8413)(0.2743) = (0.1587)(0.2743)$$
$$= 0.0435 \; .$$

EJEMPLO 12 Los diámetros de una partida grande de rodamientos están distribuídos normalmente con una media de 2,0 pulgadas y una desviación estándar de 0.01 pulgadas. Se necesitan cuatro rodamientos de diámetro mayor que 2.02 pulgadas para un aparato especial. ¿Cuál es la probabilidad de probar exactamente 10 rodamientos?

SOLUCION Determinaremos primero la probabilidad que el diámetro de un rodamiento cualquiera sea mayor que 2.02 pulgadas.

Sea X = diámetro de un rodamiento.

$X \to N(2.0, \; (0.01)^2)$

Por lo tanto,

$$P[X > 2.02] = 1 - P[X \leq 2.02]$$
$$= 1 - P\left[\frac{X - \mu}{\sigma} \leq \frac{2.02 - 2.0}{0.01}\right]$$
$$= 1 - P[Z \leq 2] = 1 - \Phi(2)$$
$$= 1 - 0.9773 = 0.0227 \; .$$

Sea ahora la variable aleatoria Y definida de la siguiente manera

Y = número de rodamientos probados hasta conseguir, 4 con diámetro mayor que 2.02

$R_X = \{4, 5, 6, \ldots\}$

$p = 0.02227$; (probabilidad que el diámetro de un rodamiento sea mayor que 2.02 pulgadas).

Luego, la variable Y tiene una distribución binomial negativa. Y su función de probabilidad es

$$P[Y = y] = \binom{y-1}{3} (0.0227)^4 (0.9773)^{y-4} \quad , \qquad y = 4, 5, 6, \ldots$$

Entonces,

$$P[Y = 10] = \binom{9}{3} (0.0227)^4 (0.9773)^6 \; .$$

EJEMPLO 13 El tiempo de máquina necesario para fabricar una unidad del producto "TONIX" está distribuído normalmente con media 50 minutos y desviación estándar 5 minutos. Se debe fabricar una partida de 40000 unidades de dicho producto.

(a) ¿Cuántas unidades requerirán, un tiempo de máquina más de 53 minutos?

(b) ¿Cuántas unidades requerirán un tiempo de máquina comprendido entre 48 y 53 minutos inclusive?

(c) El 50% de las unidades requieren de un tiempo comprendido entre x_1 y x_2 minutos. Determinar x_1 y x_2 si son simétricos son respecto al tiempo medio.

SOLUCION X = tiempo de máquina necesario para fabricar una unidad del producto "TONIX".

$X \to N(50, 25)$.

(a) Cálculo de la probabilidad que una unidad requiera un tiempo de máquina, más de 53 minutos .

$$P[X > 53] = P\left[\frac{X-\mu}{\sigma} > \frac{53-50}{5}\right] = P\left[Z > \frac{3}{5}\right]$$

$$= 1 - \Phi(0.6) = 1 - 0.7257 = 0.2743 .$$

El número de unidades que requerirán un tiempo de máquina más de 53 minutos - está dado por

[Número de unidades producidas] $P[X > 53] = 40000(0.2743) = 1097.2$ unidades

(b) $$P[48 \leq X \leq 53] = P\left[\frac{48-50}{5} \leq \frac{X-\mu}{\sigma} \leq \frac{53-50}{5}\right]$$

$$= P\left[-\frac{2}{5} \leq Z \leq \frac{3}{5}\right] = \Phi(0.6) - \Phi(-0.4)$$

$$= 0.7257 - 0.3446 = 0.3811 .$$

El número de unidades que requerirán un tiempo de máquina comprendido entre 48 y 53 minutos inclusive, es

[Número de unidades producidas] $P[48 \leq X \leq 53] = 40000(0.3811) = 1524$ unidades.

(c) Puesto que x_1 y x_2 son simétricos respecto a la media, entonces $x_2 = 50 + x$, $x_1 = 50 - x$. Luego se tiene que

$$P[50 - x \leq X \leq 50 + x] = P\left[-\frac{x}{5} \leq \frac{X-\mu}{\sigma} \leq \frac{x}{5}\right]$$

$$= \Phi(\tfrac{x}{5}) - \Phi(-\tfrac{x}{5})$$

$$= \Phi(\tfrac{x}{5}) - [1 - \Phi(\tfrac{x}{5})]$$

$$= 2\Phi(\tfrac{x}{5}) - 1 = 0.5$$

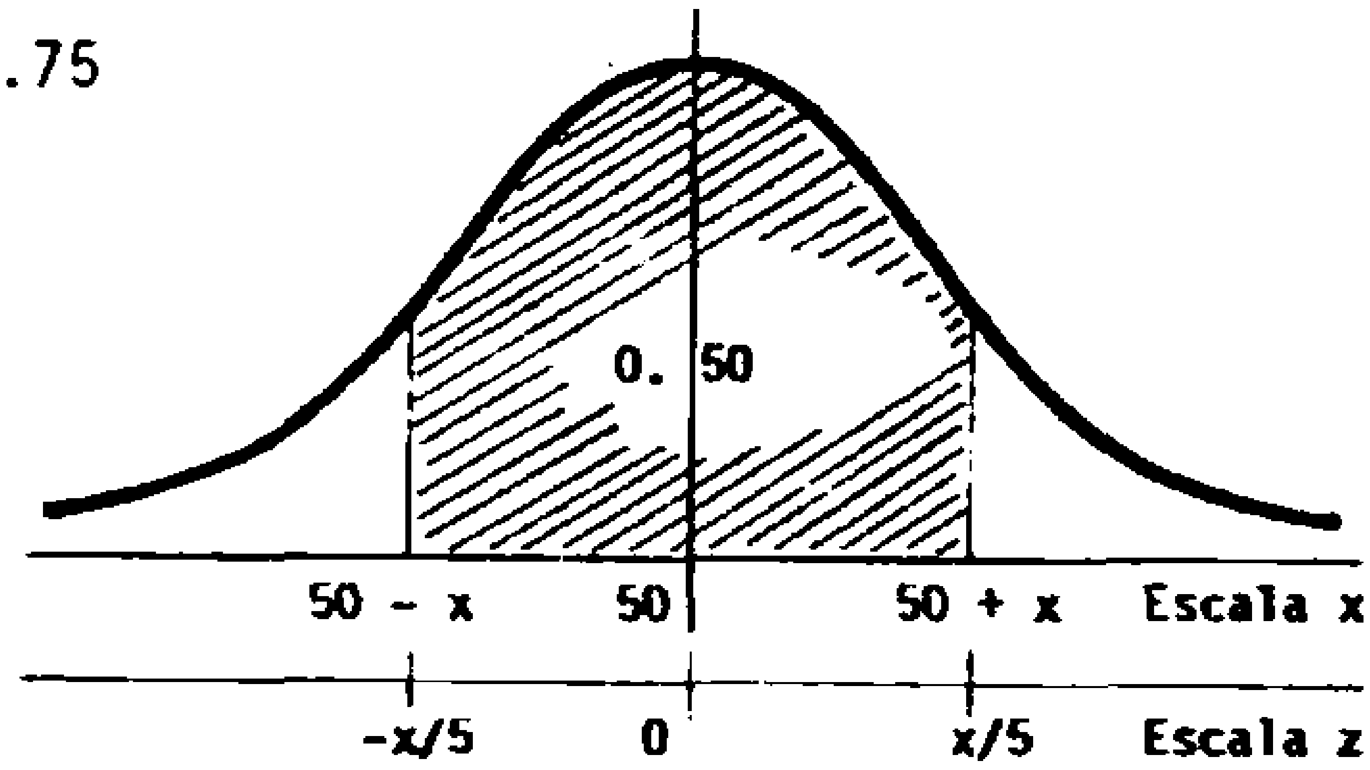
de donde $\Phi(\frac{x}{5}) = \frac{1 + 0.5}{2} = 0.75$

y de la tabla III, es $\frac{x}{5} = 0.675$

es decir $x = 3.375$

Por lo tanto $x_1 = 50 - 3.375 = 46.625$

$x_2 = 50 + 3.375 = 53.375$



EJEMPLO 14 Un combustible para cohetes va a contener cierto porcentaje (llamado **X**) de un compuesto particular. Las especificaciones exigen que **X** esté entre 30 y 35 por ciento. El fabricante tendrá una utilidad neta en el combustible (por galón) que es la siguiente función de **X**:

$$T(X) = \begin{cases} \$\ 0.10 \text{ por galón si } 30 < X < 35 \\ \$\ 0.05 \text{ por galón si } 35 \leq X < 40 \text{ ó } 25 < X \leq 30 \\ \$-0.10 \text{ por galón , en otros casos} \end{cases}$$

Si **X** tiene una distribución $N(33,9)$. Calcule $E(T)$

SOLUCION T = utilidad neta del fabricante

$R_T = \{0.10, 0.05, -0.10\}$

La distribución de probabilidad de **T** está dado en el siguiente cuadro

t	0.10	0.05	- 0.10
$P[T = t]$	0.5899	0.3964	0.0137

donde $P[T = 0.10] = P[30 < X < 35] = P\left[\frac{30-33}{3} < \frac{X-\mu}{\sigma} < \frac{35-33}{3}\right]$

$= (0.67) - \Phi(-1) = 0.7486 - 0.1587 = 0.5899.$

$P[T = 0.05] = P = [35 \leq x < 40] \cup [25 < x \leq 30]$

$= P = \left[\frac{35-33}{3} \leq \frac{x-\mu}{\sigma} < \frac{40-33}{3}\right] + P\left[\frac{25-33}{3} < \frac{x-\mu}{\sigma} \leq \frac{30-33}{3}\right]$

$= \Phi(2.33) - \Phi(0.67) + \Phi(-1) - \Phi(-2.67)$

$= 0.9901 - 0.7486 + 0.1587 - 0.0038$

$= 0.3964 .$

$P[T = -0.10] = 1 - [0.5899 + 0.3964] = 0.0137.$

Por lo tanto $E(T) = (0.10)(0.5899) + (0.05)(0.3964) - (0.10)(0.0137)$

$= 0.0774 .$

6.3.3 PROPIEDAD REPRODUCTIVA DE LA DISTRIBUCION NORMAL

Algunas distribuciones de probabilidad tienen la propiedad siguiente: Si dos o mas variables aleatorias que tienen distribución de probabilidad del mismo tipo se suman, la variable aleatoria resultante tiene una distribución del mismo tipo que los sumandos. Esta propiedad, se llama *propiedad reproductiva*. Estableceremos aquí para la distribución normal.

TEOREMA 6.3.2 Sean X_1 , $X_2, \ldots, X_n$, n variables aleatorias independientes donde $X_i \to N(\mu_i, \sigma_i^2)$, para $i = 1,2,\ldots, n$. Si

$$Y = X_1 + X_2 + \ldots + X_n$$

Entonces, la variable aleatoria Y se distribuye normalmente con media, .

$$\mu_Y = \sum_{i=1}^{n} \mu_i \quad \text{y} \quad \text{varianza} \quad \sigma_Y^2 = \sum_{i=1}^{n} \sigma_i^2$$

Es decir,

$$Y \to N\left(\sum_{i=1}^{n} \mu_i \,,\ \sum_{i=1}^{n} \sigma_i^2\right)$$

EJEMPLO 15 Un brazo mecánico consta de 3 partes. Suponga que X_1, X_2 y X_3 - son producidos por diferentes fábricas y cuya distribución de la longitud de cada una está dado por :

$$X_1 \to N(12,\ 0.02) = N(\mu,\sigma^2)$$
$$X_2 \to N(24,\ 0.03) = N(\mu,\sigma^2)$$
$$X_3 \to N(18,\ 0.04) = N(\mu,\sigma^2)$$

donde la media está dado en centímetros y la varianza en centímetros cuadrados. Calcular la probabilidad que la longitud del brazo esté comprendido - entre 53.8 y 54.2.

SOLUCION Sea Y = longitud del brazo mecánico.

Entonces, $Y = X_1 + X_2 + X_3$

la variable aleatoria Y está distribuída normalmente de acuerdo a la propiedad reproductiva de la distribución normal, con media

$$\mu_Y = \mu_1 + \mu_2 + \mu_3 = 12 + 24 + 18 = 54 \quad \text{y varianza}$$
$$\sigma_Y^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 0.02 + 0.03 + 0.04 = 0.09$$

Entonces

$$P[53.8 < Y < 54.2] = P\left[\frac{53.8 - 54}{0.3} < \frac{Y - \mu}{\sigma} < \frac{54.2 - 54}{0.3}\right]$$

$$= P\left[-\frac{2}{3} < Z < \frac{2}{3}\right] = \Phi(0.67) - \Phi(-0.67)$$

$$= 0.7486 - 0.2514 = 0.49 .$$

El resultado anterior se puede generalizar a una combinación lineal de variables aleatorias normales independientes. Sea Y una combinación lineal de las variables aleatorias normales independientes $X_1, X_2, \ldots, X_n$. Es decir,

$$Y = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n$$

entonces Y tiene una distribución normal con media

$$\mu_Y = a_0 + \sum_{i=1}^{n} a_i\mu_i$$

y varianza

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2\sigma_i^2$$

EJEMPLO 16 Sea $Y = \left[\frac{X_1 + X_2}{2}\right] - X_3$

donde $X_i (i = 1,2,3)$ son variables aleatorias normales independientes con

$$\mu_1 = 20; \quad \mu_2 = 16; \quad \mu_3 = 25; \quad \sigma_1^2 = 5; \quad \sigma_2^2 = 11 \quad y \quad \sigma_3^2 = 5$$

(a) ¿Cuál es la distribución de probabilidad de Y?

(b) Calcular $P[Y > 0]$

SOLUCION $X_i \to N(\mu_i, \sigma_i^2)$, X_i independientes

(a) Y es una combinación lineal de las X_i , por lo tanto tiene una distribución normal con $E(Y) = \mu_Y$ y $Var(Y) = \sigma_Y^2$. Donde

$$\mu_Y = E(Y) = E\left[\left(\frac{X_1 + X_2}{2}\right) - X_3\right] = \frac{1}{2}[E(X_1) + E(X_2)] - E(X_3)$$

$$= \frac{1}{2}[\mu_1 + \mu_2] - \mu_3 = \frac{1}{2}[20 + 16] - 25 = -7,$$

$$\sigma_Y^2 = Var(Y) = Var\left[\left(\frac{X_1 + X_2}{2}\right) - X_3\right] = \frac{1}{4}[Var(X_1) + Var(X_2)] + Var(X_3)$$

$$= \frac{1}{4}[\sigma_1^2 + \sigma_2^2] + \sigma_3^2 = \frac{1}{4}[5 + 11] + 5 = 9 .$$

(b) $P[Y > 0] = P\left[\frac{Y - \mu_Y}{\sigma_Y} > \frac{0 - (-7)}{3}\right] = P\left[Z > \frac{7}{3}\right] = 1 - \Phi(2.33)$

$= 1 - 0.9901 = 0.0099$.

6.3.4 TEOREMA CENTRAL DEL LIMITE

El teorema central del límite, es uno de los conceptos más importantes en estadística. Este teorema justifica la importancia de la distribución normal.

TEOREMA 4.3.3 CENTRAL DEL LIMITE Sea $X_1, X_2, \ldots, X_n, \ldots$ una sucesión de variables aleatorias independientes con

$$E(X_i) = \mu_i \quad y \quad Var(X_i) = \sigma_i^2 \quad \text{(ambos finitos)} .$$

Si $Y_n = X_1 + X_2 + \ldots + X_n$, entonces bajo ciertas condiciones generales, la variable aleatoria Z_n definida por

$$Z_n = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

tiene una distribución aproximadamente N(0,1), cuando n es grande. Es decir si F_n es la función de distribución de Z_n , se tiene

$$\lim_{n \to \infty} F_n(z) = \Phi(z)$$

Las "condiciones generales" mencionado en el teorema, son resumidos informalmente como sigue: los términos X_i tomados individualmente, contribuye con una cantidad despreciable a la variación de la suma, y no es probable que un simple término haga una gran contribución a la suma.

Observe que la variable aleatoria $Y_n = \sum_{i=1}^{n} X_i$ puede ser aproximada por una variable aleatoria distribuída normalmente, cualquiera que sea la distribución de las X_i . Esta es una buena razón para la importancia de la distribución normal en estadística.

En muchos problemas la variable aleatoria que se considera, se pueden considerar como la suma de n variables aleatorias independientes, y por lo tanto, su distribución puede aproximarse por la distribución normal.

Veremos ahora un caso especial del teorema central del límite, cuando cada X_i tiene la misma distribución.

TEOREMA 4.3.4 Sean X_1 , X_2 , . . . , X_n , n variables aleatorias independientes, idénticamente distribuídas con $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$ (media y varianza común y finitas). Si $Y_n = X_1 + X_2 + \ldots + X_n$. Entonces la variable aleatoria

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

donde $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

tiene aproximadamente una distribución N(0,1). Es decir, para todo $z \in \mathbb{R}$ se tiene que

$$\lim_{n\to\infty} F_{Z_n}(z) = \lim_{n\to\infty} P[Z_n < z] = \Phi(z)$$

EJEMPLO 17 La vida de cierto foco electrónico tiene una distribución exponencial con media $\mu = 23$ días. Tan pronto como un foco se quema, es reemplazado por otro. ¿Cuál es la probabilidad que durante un año se necesitan más de 25 focos?

SOLUCION

vida útil del i-ésimo foco ($i = 1,2,\ldots, 25$)

$Y_{25} = \sum_{i=1}^{25} X_i$, denota la suma de los tiempos en que los 25 focos se queman.

X_i es una variable aleatoria con distribución exponencial, cuya media es

$$\mu = \frac{1}{\alpha} = 23 \quad y \quad \sigma = \frac{1}{\sqrt{\alpha^2}} = 23 \quad \text{para cada } i = 1,2,\ldots,25$$

Debemos calcular,

$$P[Y_{25} < 365] = P\left[\frac{\sum_{i=1}^{25} X_i - \mu}{\sigma\sqrt{n}} \leqslant \frac{365 - 25(23)}{5(23)}\right]$$

$= P[Z_{25} \leq -1.826] = \Phi(-1.83)$ (teorema central del limite)

$= 0.0336$.

EJEMPLO 18 La longitud a que se puede estirar sin ruptura un filamento de Nylon es una variable aleatoria exponencial con media de 5,000 pies. ¿Cuál es la probabilidad (aproximada) que la longitud promedio de 100 filamentos este comprendida entre 4,750 y 5,550 pies?

SOLUCION

X_i = longitud a que se puede estirar un filamento de Nylon sin ruptura.

$$E(X_i) = \frac{1}{\alpha} = 5{,}000, \quad \forall i = 1,2,\ldots,100$$

$$Var(X_i) = \frac{1}{\alpha^2} = (5{,}000)^2, \quad \forall i = 1,2,\ldots,100 \quad \text{y}$$

$$\bar{X}_{100} = \frac{1}{100}\sum_{i=1}^{100} X_i$$

Entonces,

$$P[4750 < \bar{X}_{100} < 5550] = P\left[\frac{4750 - 5{,}000}{\frac{5{,}000}{\sqrt{100}}} < \frac{\bar{X}_{100} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{5{,}550 - 5{,}000}{\frac{5{,}000}{\sqrt{100}}}\right]$$

$$= P\left[-0.5 < \frac{\bar{X}_{100} - \mu}{\sigma/\sqrt{n}} < 1.1\right] = \Phi(1.1) - \Phi(-0.5)$$

$$= 0.8643 - 0.3085 = 0.5558 .$$

EJEMPLO 19 Las cajas entregadas por una fábrica tienen un peso medio de 300 libras y una desviación estándar de 50 libras. ¿Cuál es la probabilidad de que 25 cajas tomadas al azar y cargadas en un camión exceden de la capacidad especificada del camión, que se sabe es de 8,200 libras?

SOLUCION

Sea X_i = la variable aleatoria que representa el peso de la caja i.

$E(X_i)$ = 300 libras $\forall i$, y la desviación estándar de la caja X_i, $\sigma_{X_i} = 50$ libras, $\forall i$.

Se toman 25 cajas al azar, y sean $X_1, X_2, \ldots, X_{25}$ los pesos de las cajas tomadas al azar. Entonces, el peso total de las 25 cajas es $\sum_{i=1}^{25} X_i = Y_{25}$

Se pide determinar la $P[Y_{25} > 8{,}200]$

$$P[Y_{25} > 8{,}200] = P\left[\frac{Y_{25} - n\mu}{\sigma\sqrt{n}} > \frac{8{,}200 - 25(300)}{50\sqrt{25}}\right]$$

$$= P\left[Z_{25} > \frac{700}{250}\right] = P\left[Z_{25} > 2.8\right]$$

$$= 1 - P[Z_{25} \leqslant 2.8] = 1 - 0.9974$$

$$= 0.0026 .$$

EJEMPLO 20 Suponga que 30 instrumentos electrónicos, $D_1, D_2, \ldots, D_{30}$, se usan de la manera siguiente: tan pronto como D_1, falla empieza a actuar D_2, cuando D_2 falla empieza a actuar D_3. etc. Suponga que el tiempo para fallar D_i, es una variable aleatoria distribuída exponencialmente con parámetro - $\alpha = 0.1$ por hora. Sea T el tiempo total de operación de los 30 instrumentos. ¿Cuál es la probabilidad que T exceda a 350 horas?

SOLUCION Sean T_i = tiempo que dura el instrumento D_i, $i = 1,2,\ldots, 30$. - Además cada T_i tiene una distribución exponencial con $\alpha = 0.1$ por hora. Por lo tanto,

$$\mu = \frac{1}{\alpha} = 10 \quad , \quad \sigma^2 = \frac{1}{\alpha^2} = 100 \quad y \quad T = \sum_{i=1}^{30} T_i$$

debemos calcular,

$$P[T > 350] = P\left[\sum_{i=1}^{30} T_i > 350\right]$$

$$= P\left[\frac{\sum_{i=1}^{30} X_i - n\mu}{\sigma\sqrt{n}} > \frac{350 - 30(10)}{10\sqrt{30}}\right]$$

$$= P\left[Z_{30} > \frac{50}{54.77}\right] \qquad \text{teorema central del límite}$$

$$= 1 - \Phi\left[\frac{50}{54.77}\right] = 1 - \Phi(0.91) = 1 - 0.8186 = 0.1814 .$$

PROBLEMAS 6.3

1. Si X es una variable aleatoria distribuída normamalmente con media $\mu = 6$ y varianza $\sigma^2 = 25$. Hallar:

   (a) $P[6 \leqslant X \leqslant 12]$ (b) $P[0 \leqslant X \leqslant 8]$

(c) $P[-2 < X < 0]$ (d) $P[X > 21]$

(e) $P[|X - 6| < 5]$ (f) $P[|X - 6| < 10]$

(g) $P[|X - 6| < 15]$

2. Si X es $N(25,36)$, determinar la constante c tal que

$$P[|X - 25| \leqslant c] = 0.9544$$

3. Si X es $N(\mu, 4)$. Calcular, $P[|X - \mu| > 3]$

4. Si X es $N(50,25)$. Calcular:

(a) $P[X > 62]$ (b) $P[|X - 50| < 8]$

5. Si X es $N(5,9)$. Hallar los valores a y b tal que $P[a < X < b] = 0.80$ donde a y b son simétricos con respecto a la media.

6. Si X es $N(3,4)$. Hallar el número c tal que

$$P[X \geqslant c] = 2P[X < c]$$

7. Una variable aleatoria X se distribuye normalmente . Si $E(X^2) = 68$ y $P[X < 10] = 0.8413$; determinar μ y σ^2 .

8. Los tubos fabricados por cierta máquina tiene un diámetro medio de - $\mu = 9.8$ mm. con desviación estándar $\sigma = 0.536$ mm. ¿Qué porcentaje de tubos será rechazado, si no se aceptan diámetros inferriores a 9.0 mm?. Asuma - que los diámetros tienen una distribución normal.

9. Los límites de aceptación para los diámetros de los balones producidos por cierta máquina son $\mu \pm \sigma$. ¿Qué porcentaje de balones serán aceptados? - (Sug. $P[\mu - \sigma < X < \mu + \sigma]$)

10. Para cierto examen la calificación media es de 11 y $\sigma = 2$. Se desea desaprobar al 40% de los examinados . ¿Cuál debe ser la calificación máxima - desaprobatoria?

11. Un ictiólogo está interesado en estimar cuanto tiempo puede sobrevivir - cierto tipo de pez de mar en aguas del rio Amazonas. Luego, de una serie - de experimentos llega a estimar que la vida media de este tipo de pez alcanza a los 210 días después de haber sido colocado en el agua del rio, - con una desviación estándar de 40 días. El ictiólogo estima que la distribución de los días vividos es normal. Un pez particular ha sobrevivido 230 días, ¿cuál es la probabilidad de que viva más de 240 días?

12. Se está construyendo un grupo de 100 casas en la Urbanización San Borja. - El material empleado en las redes de desague es tal que el 9.512% de las tuberías de desague tienen períodos de duración que exceden los 15 años y que el 62.556% tienen períodos de duración que exceden los 9 años. Considerando que la distribución de probabilidad de los períodos de duración de éstas tuberías es normal, determínese la media y la varianza de esta distribución .

13. El gerente de producción de una fábrica piensa que la vida útil de una máquina M está distribuída normalmente con una media de 3000 horas. Si además el gerente piensa que hay una probabilidad 0.50 de que la máquina dure menos de 2632 ó más de 3368 horas. ¿Cũal es la desviación estándar?

14. Un rodamiento es considerado defectuoso y por lo tanto es rechazado si su diámetro es mayor que 2.02 pulgadas o menor que 1.98 pulgadas. ¿Cuál es el número esperado de rodamientos rechazados, si los diámetros de una partida de 10,000 rodamientos están distribuídos normalmente con una media de 2 - pulgadas y una desviación estándar de 0.01 pulgadas?

15. Los diámetros de una partida grande de rodamientos están distribuídos normalmente con una media de 2.0 pulgadas y una desviación estándar de 0.01 - pulgadas. Suponga que se necesita un rodamiento con diámetro mayor que - 2.02 pulgadas. ¿Cuál es la probabilidad de tener que probar 10 rodamientos?

16. Un super mercado almacena 30 kilogramos de queso fundido cada semana. Si - la demanda semanal de queso fundido está normalmente distribuida con media 24 kilogramo y desviación típica 5 kilogramo. Determinar la probabilidad - que el supermercado agote los quesos fundidos durante una semana seleccionada al azar.

17. Los errores de medida resultantes del uso de cierta balanza están normalmente distribuidos con media 0 onzas y desviación típica 0.1 onzas. Determinar la probabilidad que el peso medido de un objeto seleccionado al azar difiere del peso verdadero del objeto en más de 0.12 onzas.

18. Un análisis estadístico de 10000 llamadas telefónicas de larga distancia - hechas desde una central telefónico indica que la duración de esas llamadas tiene una distribución normal con media 129.5 segundos y desviación típica 30.0 segundos.
   (a) ¿Cuál es la probabilidad que una llamada particular haya durado entre 89.5 y 169.5 segundos?

(b) ¿Cuantas llamadas duraron menos de 60 segundos o más de 150 segundos?

(c) ¿Cuál debe ser la duración de una llamada particular, si sólo 1% de todas las llamadas son más cortas?

19. Se extraen 4 observaciones al azar de una población normal ¿Cuál es la probabilidad que por lo menos 3 de ellos difieran de la media de la población en mas de una desviación estandar?

20. Si de una población distribuida normalmente, extraemos una muestra de 5 observaciones al azar. ¿Cuál es la probabilidad que 3 de las observaciones difieren de la media de la población en más de media desviación estándar?

21. Una compañía envasadora de conservas, ofrece en el mercado latas de arvejas en cuya etiqueta establece que el peso es de 17 onzas. La compañía no desea sobrellenar las latas ni tampoco desea que no tengan el peso completo. Los pesos de las latas se encuentran normalmente distribuídos. La media del peso neto de las latas se puede controlar ajustando la máquina llenadora, mientras, que la desviación estándar no se puede alterar, puesto que refleja la capacidad de la máquina para producir llenado uniforme. Por experiencias pasadas se sabe que la desviación estándar es de 0.18 onzas. Si la compañía no desea producir más de 0.5% de latas con un peso neto menor que el que se ofrece en la etiqueta. ¿A qué peso neto promedio debe ajustarse a la máquina llenadora?

22. En el problema 15. Se selecciona 4 rodamientos al azar de la partida. ¿Cuál es la probabilidad que exactamente 2 rodamientos tengan un diámetro mayor que 2.02 pulgadas?

23. En una industria alimenticia se comercializa harina en paquetes se "PESO NETO 500 GRS". El proceso automático de llenado de los paquetes puede regularse de modo que la cantidad media de harina por paquete puede ajustarse al nivel que se desee. Suponiendo que la cantidad de harina por paquete se distribuye normalmente con una desviación estándar de 0.2 onzas.

(a) ¿A qué nivel debe ajustarse el llenado medio de modo que sólo el 0.001 de los paquetes tengan un peso neto inferior a 12 onzas?

(b) ¿A qué nivel debe ajustarse el llenado medio de modo que sólo el 0.05 de los paquetes tengan un peso superior a 12.4 onzas?

24. El bar "un par más" ha instalado una máquina automática para la venta de cerveza. La máquina puede regularse de modo que la cantidad media de cerve

za por vaso sea la que se desea; sin embargo, en cualquier caso esta cantidad tendrá una distribución normal con una desviación estándar de 5.9 mililitros.

(a) Si el nivel se ajusta a 304.6 mililitros, ¿Qué porcentaje de los vasos contendrán menos de 295.7 mililitros?

(b) ¿A qué nivel medio debe ajustarse la máquina para que sólo el 2.28% de los vasos contengan menos de 295.7 mililitros?

(c) ¿A qué nivel medio debe ajustarse la máquina para que el 84.13% de los vasos contenga menos de 313.6 mililitros?

25. En una distribución normal hay 40% de valores inferiores a 50 y 30% superiores a 70. Determinar la proporción de valores entre 55 y 70.

26. Una persona viaja diariamente de su casa a la oficina y ha encontrado que el tiempo en el viaje le corresponde una media $\mu = 35.5$ con una desviación estándar $\sigma = 3.11$ minutos. Si sale de su casa todos los días a las 8.20 y debe estar en su oficina a las 9.00 ¿Cuántos días al año espera llegar a las 9?. Suponer 240 viajes anuales. Suponga distribución normal.

27. La presión sanguínea media en hombres de 20 a 25 años de edad es 123 unidades con desviación típica de 13.7 unidades. Si se selecciona al azar uno de estos hombres, calcule la probabilidad que su presión sanguínea esté comprendida entre 120 y 128 unidades. Suponer distribución normal.

28. Una linterna grande es alimentada por cinco baterías. Suponga que la vida de una batería está normalmente distribuído con media $\mu = 120$ horas y desviación estándar $\sigma = 10$ horas. La linterna dejará de funcionar si se agota una o más de sus baterías. Suponiendo que las vidas de las baterías son independientes. ¿Cuál es la probabilidad que la linterna funcione más de 100 horas?

29. En una distribución normal con media 163 y desviación típica 12, se hace 3 observaciones diferentes. Determinar la probabilidad que 2 de estas observaciones sean mayores o iguales que 175.

30. Si el espesor de cierto tipo de tuercas tiene media igual a 1.95 mm. y una desviación estándar de 0.12 mm. ¿Cuántas tuercas de un grupo de 1,000 tendrán un espesor comprendido entre 1.80 mm. y 2.10 mm.?

31. El gerente de crédito de una tienda de departamentos estima las pérdidas por deudas impagos en el año en la forma siguiente; la pérdida tiene una

distribución normal con una media de \$ 30,000 y hay una porbabilidad de 0.50 de que sea mayor que \$ 35.000 o menor que \$ 25,000. ¿Cuál es la desviación estándar?

32. La duración de ciertos tubos electrónicos empleados en una máquina tienen distribución normal. Los tubos de tipo **A** tienen una duración media de 30 horas con una desviación estándar de 5 horas. Los tubos de tipo B tienen una vida media de 35 horas con una desviación estándar de 3 horas. Si la máquina debe funcionar durante 35 horas, ¿qué tipo de tubos será preferible? justifique plenamente su respuesta.

33. Hay dos procedimientos para poner en disposición de despegue los aviones cazas. El procedimiento **A** requiere un tiempo medio de 27 minutos con una desviación estándar de 5 minutos. Para el procedimiento **B**, $\mu = 30$ y $\sigma = 2$ minutos respectivamente. ¿Qué procedimiento debe usarse si el tiempo disponible es 30 minutos? ¿34 minutos?

34. Las estaturas de 1000 estudiantes están distribuidas normalmente con una media de 174.5 centímetros y una desviación estándar de 6.9 centímetros. Suponiendo que las estaturas se redondean al medio centimetro más cercano, ¿cuántos de estos estudiantes se espera que tengan estaturas
(a) menores de 160.0 cm.?
(b) entre 171.5 y 182.0 cm. inclusive?
(c) mayores o iguales a 188 cm?
(d) iguales a 175.0 cm. ?

35. En un examen de estadística la puntuación promedio fue de 82 y la desviación estándar de 5. Todos los estudiantes con puntuaciones desde 88 hasta 94 obtuvieron una calificación de B. Si las puntuaciones tienen aproximadamente una distribución normal y ocho estudiantes recibieron una calificación de B, ¿cuántos estudiantes se presentaron al examen?

36. El número de días entre la facturación y el pago de las cuentas corrientes de crédito en una tienda de departamentos grande tiene una distribución aproximadamente normal con una media de 18 días y una desviación estándar de 4 días. ¿Qué proporción de las facturas será pagada,
(a) entre 12 y 18 días? (b) entre 20 y 23 días? (c) en menos de 8 días?
(d) ¿Dentro de cuántos días estará pagado el 99.5 % de las facturas?
(e) ¿Entre cuales dos valores simetricamente distribuídos al rededor de la media recaerá el 98 % de las facturas?

37. La medición del diámetro de contacto de la rosca de una unión se distribuye normalmente con media 0.4008 pulgadas y una desviación estándar 0.003 pulgadas. Los límites de especificación están dados como 0.4000 ± 0.0010 pulgadas. ¿Cuál es la probabilidad que ocurra una defectuosa?

38. Los diámetros de las fichas rin, para llamar por teléfono, se distribuyen normalmente con media 1.0 pulgadas y una desviación estándar de 0.03 pulgadas. Las especificaciones requeridas por los tornos de acceso son 1.00 ± 0.05 pulgadas. ¿Qué porcentaje de nuevas fichas estarán dentro de los límites de especificación?

39. Cierta dimensión de una parte mecánica producida en gran escala tiene un valor nominal de 100 mm. con tolerancias aceptables de ± 1 mm. Si el proceso de fabricación produce partes para los cuales los valores de está dimensión tiene distribución normal con media 100.2 mm. y desviación estándar de 0.5 mm. ¿qué porcentaje de las partes tendrán que rechazarse por estar fuera de los límites de tolerancia?

40. La durabilidad de un lote de componentes para radio sigue una distribución normal con una media de 500 horas y una desviación estándar de 50 horas. Un comprador requiere que el 95% de las componentes como mínimo tengan una durabilidad mayor a 400 horas. ¿Cumplirá este lote la especificación del comprador?

41. El análisis de los datos pasados ha mostrado a un fabricante que el espesor del centro de un determinado tipo de engranaje tiene una distribución normal con media de 50 mm. y desviación estándar de 1 mm. Estime cuántos engranajes en una partida de producción de 5000, tendran un espesor de centro (a) mayor que 51.5 mm, (b) entre 49.2 y 50.8 mm, (c) entre 49.0 y 49.5 mm.

42. La especificación con la que se fabrican los tornillos de acero de sección transversal circular requiere que sus longitudes se encuentren entre 8.45 y 8.65 cm. y sus diámetros entre 1.55 y 1.60 cm. Los tornillos producidos por cierta máquina tienen longitudes que siguen una distribución normal con media de 8.65 cm y desviación estándar 0.05 cm, y diámetros que siguen también una distribución normal independiente con media 1.58 cm y desviación estándar 0.001 cm. Determine los siguientes datos para tornillos que fabrica esta máquina:

(a) El porcentaje que estará fuera de los límites de longitud especificados

(b) El porcentaje que estará fuera de los límites especificados de diámetro.

(c) El porcentaje que no cubra la especificación.

43. Ciertas varillas se fabrican con una longitud nominal de 100 mm. pero de hecho sus longitudes forman una distribución normal con media 100.4 mm. y desviación estándar 0.8 mm. La fabricación de cada varilla cuesta \$12 y su uso inmediato si su longitud se encuentra entre 99 mm y 101 mm. Las varillas más cortas no pueden emplearse, pero tienen un valor de desecho de \$ 2. Las que sobrepasan a 101 mm. pueden cortarse a una longitud aceptable con un costo adicional de \$ 4. Determine el costo promedio de una varilla útil.

44. Suponga que **X**, la resistencia a la ruptura de una cuerda (en libras) tiene distribución normal con media 100 y varianza 16. Cada 100 pies de alambre para cuerda produce una utilidad de \$ 25, si $X > 95$. Si $X \leq 95$, la cuerda puede utilizarse con un objetivo diferente y se obtiene una utilidad de - \$ 10 por alambre. Encuentre la utilidad esperada por alambre.

45. Un fabricante de juguetes empaca éstos a razón de 24 por caja para enviarlos a las jugueterías. El peso de cada uno de estos juguetes es una variable aleatoria normal con media 12 (onzas) y una desviación estándar - 0.1. Se representa por **X** el peso de una de estas cajas.

(a) ¿Cuál es la distribución de probabilidad de **X**?

(b) ¿Cuál es la probabilidad que una caja pese entre 18 y 19 libras? (**NOTA**: 16 onzas = 1 libra).

46. En una placa junto a la puerta de un ascensor se lee. "Capacidad 6 personas 990 libras". Suponer que las personas que usan este ascensor se seleccionan de una distribución normal con $\mu = 140$ libras, $\sigma = 30$ libras .

(a) Seis personas entran al elevador, ¿cuál es la probabilidad que su peso combinado exceda la capacidad máxima de 990 libras?

(b) resuelva la pregunta (a) si entran siete personas el ascensor.

47. Un eje consiste de seis secciones diferentes colocadas una al lado de la - otra. Se sabe que las longitudes de estas secciones componentes son variables aleatorias distribuídas normalmente con las siguientes medias y varianzas ;

(8.10, 0.05) , (7.25, 0.04) , (9.75, 0.06) , (3.45 , 0.04), (17.15, 0.10), (6.20, 0.07). Si las especificaciones establecen que el eje ensamblado -

tenga una longitud de 52 ± 1.5 pulgadas. ¿Cuál es la probabilidad que un eje ensamblado cumpla con las especificaciones?

48. Sea X_1 y X_2 variables aleatorias independientes distribuídas normalmente con $\mu_1 = 0$, $\sigma_1^2 = 4$, $\mu_2 = 1$, $\sigma_2^2 = 9$ como medias y varianzas, respectivamente. Calcular

(a) $P[X_2 \leqslant 0]$; (b) $P[X_1 + X_2 > 1]$

49. Sean X_1, X_2, X_3 y X_4 variables aleatorias independientes tales que:

$X_1 \to N(2,3)$; $X_3 \to N(4,4)$

$X_2 \to N(2,4)$; $X_4 \to N(4,2)$

Si $X = X_1 + 2X_2 + X_3 + X_4$

(a) Hallar la media y la varianza de X .

(b) Calcular $P[10 \leqslant X \leqslant 14]$

50. Sean X_1, X_2, X_3 y X_4 variables aleatorias normales independientes con $\mu_1 = 2$; $\mu_2 = 5$; $\mu_3 = 4$; $\mu_4 = 6$; $\sigma_1^2 = 4$; $\sigma_2^2 = 6$; $\sigma_3^2 = 6$; $\sigma_4^2 = 5$

Si $Y = [\dfrac{X_1 + 2X_2}{4}] - X_3 + \dfrac{X_4}{2}$

Calcular $P[|Y| > 3]$

51. Las pastillas metálicas cilíndricas que se utilizan en un reactor se fabrican en serie y puede suponerse que sus longitudes siguen una distribución normal con media 0.290 cm. y desviación estándar 0.016 cm. Nueve de estas pastillas deben ajustarse, extremo con extremo, en un recipiente que ocupa una longitud no mayor de 2.670 cm. Si las nueve pastillas se ensamblan al azar, ¿qué proporción de estos no se ajustará en el espacio requerido?

52. En una de las etapas de un proceso de ensamble un tapón cilíndrico tiene que ajustarse a una abertura circular seleccionando cada elemento al azar en un suministro continuo. Los diámetros del tapón y de los casquillos en mm, son $N(24.9, 0.03^2)$ y $N(25, 0.04^2)$ respectivamente. Si para que el ajuste sea satisfactorio se requiere un claro en el diámetro de cuando menos 0.02 mm, ¿en qué proporción de los casos el ajuste no será satisfactorio? (claro del diámetro = Diámetro del casquillo - diámetro del tapón).

53. Un tipo de sensor tiene una vida de 15,000 horas con una desviación estándar de 1000 horas. Tres de estos sensores se conectan en un dispositivo de

tector de incendios de manera que cuando uno falla el otro lo sustituye. Suponiendo que la durabilidad tiene una distribución normal, ¿cuál es la probabilidad que el dispositivo (a) funcione durante 40000 horas, (b) falle antes de 39,000 horas?

54. Suponga que los pesos de los pasajeros que viajan por aire en los vuelos establecidos que parten de un aeropuerto grande siguen una distribución normal con media 78 kg y desviación estándar 10 kg, encuentre los límites (simétricos alrededor de la media) de tal manera que el 95% de los pasajeros tengan un peso límite dentro de estos valores. ¿Cuál es la probabilidad de que el peso total de una muestra al azar de 100 pasajeros exceda a 8000 kg?

55. Varias pruebas han demostrado que la temperatura más alta (en °C) que pueden soportar los condensadores de un determinado tipo es N(125,9). En los sistemas en que se utilizan, la temperatura máxima (en °C) a que se sujeta un condensador individual es N(116, 16). ¿Qué proporción de los condensadores fallará por sobrecalentamiento?

56. Dos marcas de focos "Económica" y "vida eterna" tienen durabilidad (en horas) que son $N(1400, 200^2)$ y $N(2000, 250^2)$ respectivamente. Si se prueba la vida de uno de los focos correspondientes a cada una de las marcas, ¿cuál es la probabilidad que la "económica" dure más?

57. Al sumar números, un computador aproxima cada número al entero más próximo Supongamos que todos los errores de aproximación son independientes y distribuídos uniformemente entre $\langle -0.5 , 0.5 \rangle$

(a) Si se suman 1,500 números, ¿Cuál es la probabilidad que la magnitud del error total exceda a 15?

(b) ¿Cuántos números pueden sumarse juntos a fin que la magnitud del error total sea menor que 10, con probabilidad 0.90?

58. Supóngase que las variables aleatorias $X_1, X_2, \ldots, X_{50}$ representan las vidas útiles de 50 tubos electrónicos $T_1, T_2, \ldots, T_{50}$ y se usan de la siguiente manera: tan pronto como falla T_1 empieza a funcionar T_2 y cuando falla T_2 empieza a funcionar T_3, etc. Suponga que las X_i, $i = 1, 2, \ldots, 50$, tienen igual distribución de probabilidad y la probabilidad de que un tubo electrónico esté funcionando después de 1 hora es $e^{-1/500}$. ¿Cuál es la probabilidad que el tiempo de funcionamiento de los 50 tubos esté comprendido entre 26,000 y 28,000 horas?

59. 250 artículos son empacados en cajas grandes. Los pesos de los artículos son variables aleatorias independientes con media 0.5 libras y desviación estándar 0.10 libras. 20 cajas son colocadas en un camión. Calcular la probabilidad que el peso de los empaques colocados en el camión exceda a 2510 libras. (considere despreciable el peso de la caja).

60. Un camión de reparto transporta cajones cargados de artículos varios. Si el peso de cada cajón es aproximadamente distribuído normalmente con media 50 libras y desviación estándar 5 libras, ¿cuántos cajones pueden ser transportados en el camión de tal forma que la probabilidad que la carga total exceda a 1 tonelada sea sólo 0.10?

61. El 40% de las llamadas de un vendedor resultan ventas. Si el desea hacer 60 ventas en ésta semana, ¿cuál es la probabilidad que haga más de 165 llamadas? (sug. defina X_i , $i = 1,2, \ldots, 60$. Como el número de llamadas efectuadas hasta obtener la i-ésima venta. Y $Y_{60} = X_1 + X_2 + \ldots + X_{60}$)

6.4 APROXIMACION DE LAS DISTRIBUCIONES DISCRETAS A LA NORMAL

En esta sección se mostrará, cómo se usa la distribución normal para aproximar varias distribuciones de probabilidad discreta importantes, como la binómial, la hipergeométrica y la de Poisson. Se describirá en forma simple esos métodos de aproximación. Por otra parte, siempre que se utiliza una aproximación de las distribuciones discretas a una distribución continua (como la normal) se tiene una aproximación más exacta si se emplea un ajuste o *corrección por continuidad.*

6.4.1 APROXIMACION BINOMIAL A LA NORMAL

En la sección 5.2 se vió que la variable aleatoria binomial X podía considerarse como una suma de n variables aleatorias independientes de Bernoulli $X_1, X_2, \ldots, X_n$. Es decir,

$$X = \sum_{i=1}^{n} X_i$$

hemos visto también que la *media* de X es $\mu = \sum_{i=1}^{n} \mu_i = np$ *y la varianza*

$$\sigma^2 = \sum_{i=1}^{n} \sigma_i^2 = npq$$

Entonces, por el teorema central del límite, la variable aleatoria

$$Z = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} = \frac{X - np}{\sqrt{npq}} = \frac{X/n - p}{\sqrt{pq/n}}$$

donde $\quad X/n = \frac{1}{n} \sum_{i=1}^{n} X_i$

tiene aproximadamente una distribución N(0,1), cuando n es grande. Para p cercano a $\frac{1}{2}$ y $n > 10$, la aproximación de la binomial a la normal, es buena Para otros valores de p, el valor de n debe ser grande. En general la experiencia indica que la aproximación es buena cuando $np > 5$ para $p \leqslant \frac{1}{2}$ ó $nq > 5$ para $p > \frac{1}{2}$.

Cuando p es muy pequeña y n grande hemos visto que la aproximación binomial a la Poisson es buena.

Observe que la distribución binomial discreta, cuya gráfica se muestra en la fig. 6.4.1, se aproxima por el área bajo la curva normal que se indica en la fig. 6.4.2

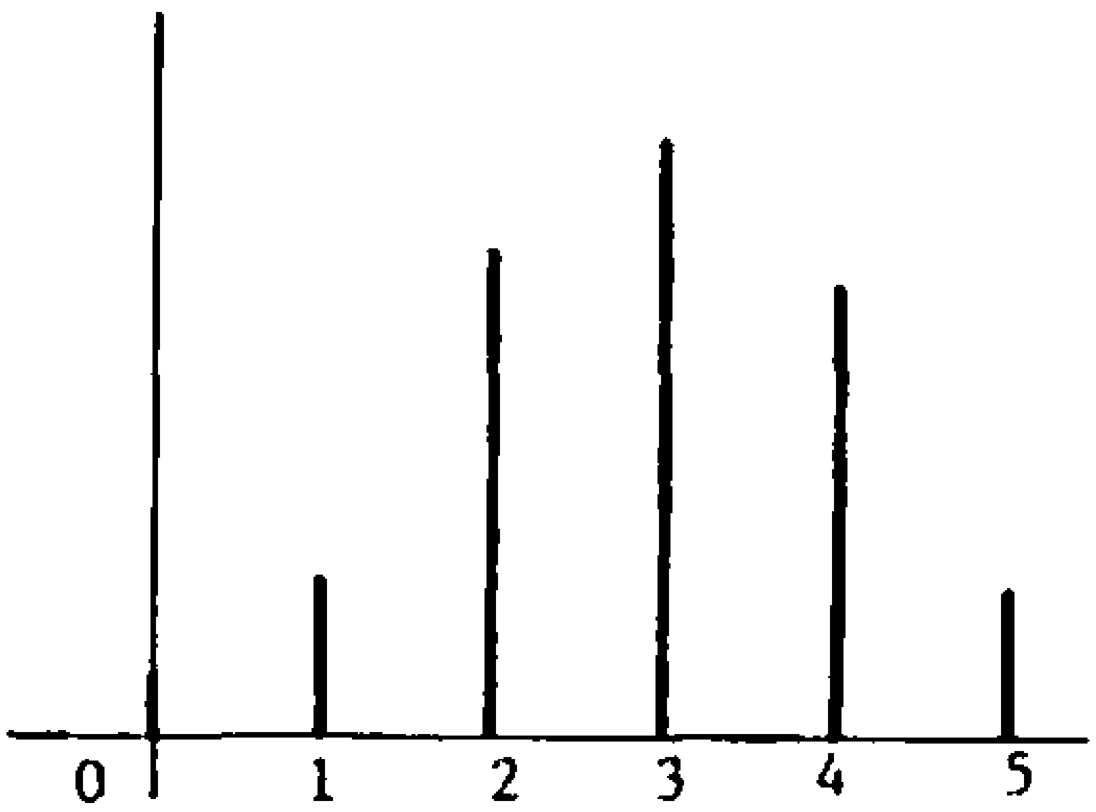


Fig. 6.4.1 Distribución binomial

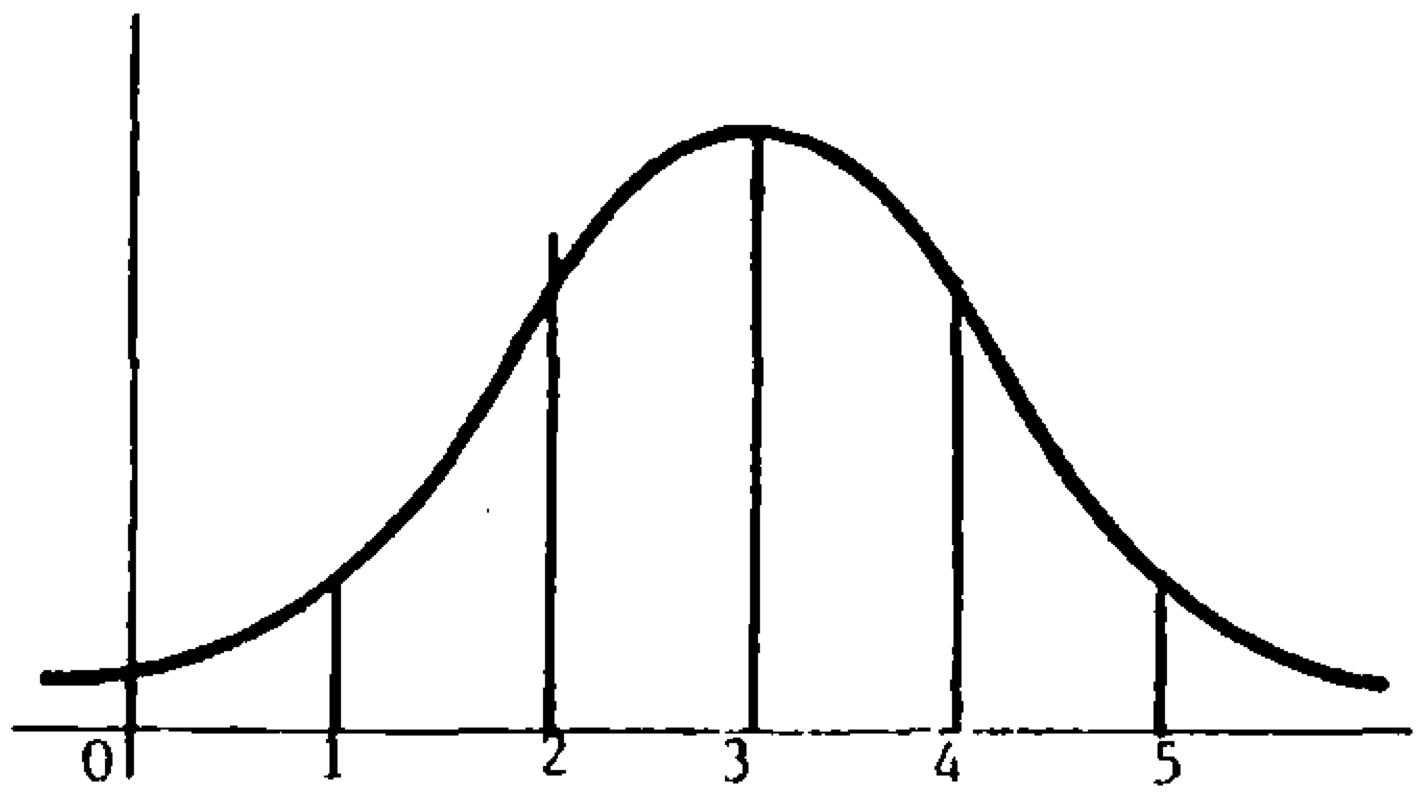


Fig. 6.4.2. Curva normal como aproximación de la binomial

Así, la probabilidad de obtener exactamente un valor x , es aproximada mediante el área bajo la curva normal entre $x - \frac{1}{2}$ y $x + \frac{1}{2}$ veáse el área sombreada en la fig. 6.4.3 .

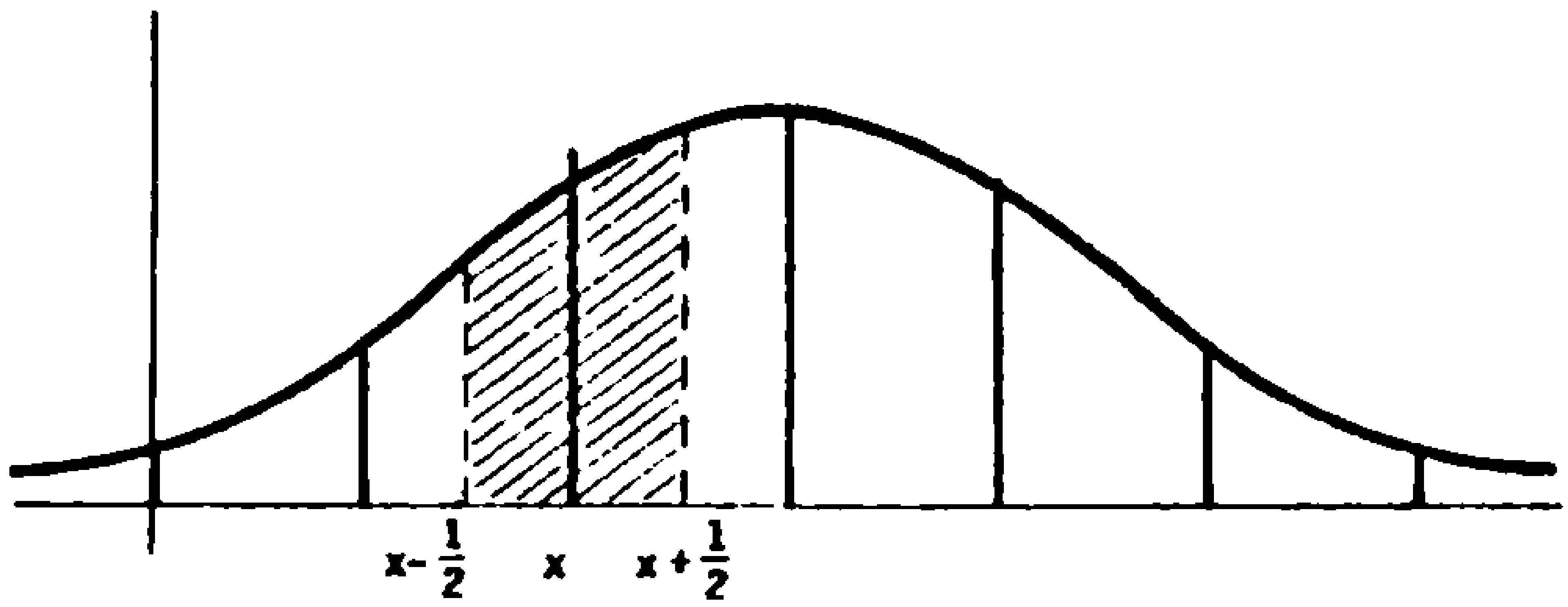


Fig. 6.4.3

Es decir, $P[X = x] \simeq P\left[x - \frac{1}{2} \leq X \leq x + \frac{1}{2}\right]$

Si se quiere aproximar la probabilidad que $a \leq X \leq b$, se tendrá que $P[a \leq X \leq b]$ está aproximada por el área bajo la curva normal entre $a - \frac{1}{2}$ y $b + \frac{1}{2}$. Es decir

$$P[a \leq X \leq b] \simeq P\left[a - \frac{1}{2} \leq X \leq b + \frac{1}{2}\right]$$

El ajuste de $\frac{1}{2}$ unidad a cada lado del intervalo discreto, se conoce conoce el *factor de corrección de continuidad* para las aproximaciones discretas a la normal.

El cuadro siguiente, presenta algunas aproximaciones de las probabilidades binomiales por la curva normal.

Probabilidad binomial del evento que se desea calcular.	Con la corrección por continuídad	en términos de la función de distribución normal estándar Φ
$P[X = x]$	$= P\left[x - \frac{1}{2} \leq X \leq x + \frac{1}{2}\right]$	$= \Phi\left[\frac{x + \frac{1}{2} - np}{\sqrt{npq}}\right] - \Phi\left[\frac{x - \frac{1}{2} - np}{\sqrt{npq}}\right]$
$P[X \leq x]$	$= P\left[X \leq x + \frac{1}{2}\right]$	$= \Phi\left[\frac{x + \frac{1}{2} - np}{\sqrt{npq}}\right]$
$P[X < x] = P[X \leq x - 1]$	$= P\left[X \leq x - 1 + \frac{1}{2}\right]$	$= \Phi\left[\frac{x - \frac{1}{2} - np}{\sqrt{npq}}\right]$
$P[X \geq x]$	$= P\left[X \geq x - \frac{1}{2}\right]$	$= 1 - \Phi\left[\frac{x - \frac{1}{2} - np}{\sqrt{npq}}\right]$
$P[X > x] = P[X \geq x + 1]$	$= P\left[X \geq x + 1 - \frac{1}{2}\right]$	$= 1 - \Phi\left[\frac{x + \frac{1}{2} - np}{\sqrt{npq}}\right]$
$P[a \leq X \leq b]$	$= P\left[a - \frac{1}{2} \leq X \leq b + \frac{1}{2}\right]$	$= \Phi\left[\frac{x + \frac{1}{2} - np}{\sqrt{npq}}\right] - \Phi\left[\frac{x - \frac{1}{2} - np}{\sqrt{npq}}\right]$

EJEMPLO 1 En un almacén se tiene 300 fusibles, y por experiencia se sabe que hay 2% de defectuosos en la producción de fusibles.

(a) ¿Cuál es la distribución de probabilidad del número de fusibles defectuosos en el almacén? ¿cuales son sus parámetros?

Utilice una aproximación a la verdadera distribución de probabilidad del número de defectuosos, para calcular la probabilidad que:

(b) haya exactamente 1 defectuoso.

(c) no hay más de 5 defectuosos.

SOLUCION (a) X = número de fusibles defectuosos en el almacén.

$$R_X = \{0, 1, 2, \dots, 300\}$$

la variable aleatoria así definida es una binomial. Su distribución de probabilidad con $n = 300$ y $p = 0.02$, es

$$P[X = x|B,300,0.02] = \binom{300}{x}(0.02)^x(0.98)^{300-x} \quad , \; x = 0,1,2,\ldots,300.$$

Los parámetros de **X** son $n = 300$ y $p = 0.02$

Desde que $np = 300(0.02) = 6 > 5$, aproximamos la distribución binomial a la distribución normal con $\mu = np = 6$ y $\sigma = \sqrt{npq} = \sqrt{5.88}$

(b) $P[X = 1] = P[0.5 \leq X \leq 1.5] = P\left[\frac{0.5 - 6}{\sqrt{5.88}} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{1.5 - 6}{\sqrt{5.88}}\right]$

$= P[-2.27 \leq Z \leq -1.86] = \Phi(-1.86) - \Phi(-2.27)$

$= 0.0314 - 0.0116 = 0.0198$.

(c) $P[X \leq 5] = P[X \leq 5.5] = P\left[\frac{X - np}{\sqrt{npq}} \leq \frac{5.5 - 6}{\sqrt{5.88}}\right]$

$= P[Z \leq -0.21] = \Phi(-0.21) = 0.4168.$

EJEMPLO 2 En una población grande de drosophila, 25% tienen mutación de alas. Una muestra de 300 insectos se escoge aleatoriamente. Calcular aproximadamente la probabilidad que más de 60 pero no más que 90 insectos de la muestra tienen mutación de alas.

SOLUCION

X = número de insectos que tienen mutación de alas en la muestra

$R_X = \{0,1,2, \ldots , 300\}$

Puesto que la población es grande, la extracción de los 300 insectos puede considerarse como 300 ensayos de Bernoulli con $p = 0.25$ constantes en cada extracción. Por lo tanto, la variable aleatoria **X** tiene una distribución binomial con $n = 300$ y $p = 0.25$.

Pero, $np = 300(0.25) = 75 > 5$, aproximamos la binomial por la normal con media $\mu = np = 75$ y desviación estándar $\sigma = \sqrt{npq} = \sqrt{300(0.25)(0.75)} = 7.5$

$$P[60 < X \leq 90] = P[60.5 \leq X \leq 90.5]$$

$$= P\left[\frac{60.5 - 75}{7.5} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{90.5 - 75}{7.5}\right]$$

$$= P\left[-\frac{14.5}{7.5} \leq Z \leq \frac{15.5}{7.5}\right] = \Phi(2.06) - \Phi(-1.93)$$

$$= 0.9803 - 0.0268 = 0.9535.$$

EJEMPLO 3 Se sabe que cierto virus ha invadido la UNI y ataca la mitad de los estudiantes. Se toma una muestra aleatoria * de 200 alumnos y se pide calcular:

(a) La probabilidad que en dicha muestra el 49% sean atacadas por el virus.

(b) La probabilidad que en la misma muestra ninguno presenta síntoma de virus.

SOLUCION X = número de alumnos atacados por el virus en la muestra.

$$R_X = \{0,1,2, \ldots, 200\}$$

X tiene una distribución binomial con $n = 200$ y $p = 1/2$.

Debido a que n es grande, y $np = 100$, daremos una aproximación de las probabilidades pedidas por la distribución normal, con

$$\mu = np = 100 \quad y \quad \sigma = \sqrt{npq} = \sqrt{50}$$

(a) $x = 49\% \,.\, 200 = 98$

$$P[X = 98] = P[97.5 \leq X \leq 98.5]$$

$$= P\left[\frac{97.5 - 100}{7.07} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{98.5 - 100}{7.07}\right]$$

$$= \Phi(-0.21) - \Phi(-0.35) = 0.4168 - 0.3632 = 0.0536.$$

(b) $$P[X = 0] = P[-0.5 \leq X \leq 0.5]$$

$$= P\left[\frac{-100.5}{7.07} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{-99.5}{7.07}\right] = 0 .$$

EJEMPLO 4 Se lanza una moneda 100 veces. El jugador recibe \$ 10 . Si el número de caras es 57 ó más, y pierde \$ 5 si el número de caras es 47 ó menos, en otros casos no recibe ni pierde nada. Determinar la esperanza del jugador en este juego.

SOLUCION X = ganancia neta del jugador en el juego.

$$R_X = \{-5, 0, 10\}$$

Para determinar la distribución de probabilidad de X, definimos la variable aleatoria

Y = número de caras obtenidas en los 100 lanzamientos de la moneda.

*** Una muestra aleatoria de tamaño n, es una muestra que se escoge de tal manera que cada subconjunto de n elementos (observaciones) de la población tienen la misma probabilidad de salir.**

Y es una variable aleatoria que tiene una distribución binomial con

$$p = \frac{1}{2} \quad y \quad n = 100$$

puesto que $np = 50$ es grande usaremos la aproximación de la binomial a la curva normal con

$$\mu = np = 100(\tfrac{1}{2}) = 50, \quad y \quad \sigma = \sqrt{npq} = \sqrt{100 \times \tfrac{1}{2} \times \tfrac{1}{2}} = 5.$$

Entonces;

$$P[X = 10] = P[Y \geq 57] = P\left[\frac{Y - np}{\sqrt{npq}} \geq \frac{57 - \frac{1}{2} - 50}{5}\right]$$

$$= P[Z \geq 1.3] = 1 - \Phi(1.3)$$

$$= 1 - 0.90324 = 0.0968 .$$

$$P[X = -5] = P[Y \leq 47] = P\left[\frac{Y - np}{\sqrt{npq}} \leq \frac{47 + \frac{1}{2} - 50}{5}\right]$$

$$= P[Z \leq -0.5] = \Phi(-0.5) = 0.3085 .$$

Finalmente,

$$P[X = 0] = P[47 < X < 57]$$

$$= P\left[\frac{47.5 - 50}{5} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{56.5 - 50}{5}\right]$$

$$= P[-0.5 \leq Z \leq 1.3] = \Phi(1.3) - \Phi(-0.5)$$

$$= 0.9032 - 0.3085 = 0.5947 .$$

Por lo tanto, la función de distribución de **X** es

x	- 5	0	10
P[X = x]	0.3085	0.5947	0.0968
x**P[X = x]**	- 1.5425	0	0.9680

$$E(X) = \sum_{x \in R_X} xP[X = x] = -1.5425 + 0 + 0.9680 = -0.5745 .$$

EJEMPLO 5 Una firma comercializa sus productos sólo por correo a una lista - de 100,000 clientes potenciales. Para decidir a cerca de la comercialización de un nuevo artículo, la firma selecciona una muestra de 100 personas de su - lista para ofrecerle el artículo. Si 30 ó más de estos clientes están dispuestos a adquirirlo, procederá a su comercialización en caso contrario no lo hará.

(a) ¿Cuál es la probabilidad que comercialice el artículo si en realidad - sólo el 20% de todos los clientes lo comprarían?

(b) ¿Cuál es la probabilidad que no comercialice el artículo si en realidad el 36% de todos los clientes lo comprarían?

SOLUCION

(a) X = número de clientes que comprarían el artículo en la muestra de 100.

$R_X = \{0,1,2, \ldots, 100\}$

Puesto que sólo el 20% de todos los clientes comprarían el artículo, entonces, $p = 0.20$, es la probabilidad de obtener un cliente que compraría el artículo. Luego, X tiene una distribución binomial con

$n = 100$ y $p = 0.2$. Por lo tanto,

$$\mu = np = 100(0.2) = 20, \qquad \sigma = \sqrt{npq} = \sqrt{100(0.2)(0.8)} = 4$$

$$P[X \geq 30] = 1 - P[X < 30] = 1 - P\left[\frac{X - np}{\sqrt{npq}} < \frac{30 - \frac{1}{2} - 20}{4}\right]$$

$$= 1 - P[Z < 2.38] = 1 - \Phi(2.38)$$

$$= 1 - 0.9913 = 0.0087 .$$

(b) El artículo no se comercializa si en la muestra, de los 100, menos de 30 comprarían el artículo.

En este caso $p = 0.36$, $n = 100$.

$$P[X < 30] = \Phi\left[\frac{29.5 - 100(0.36)}{\sqrt{100(0.36)(0.64)}}\right] = \Phi[- \frac{6.5}{4.8}]$$

$$= \Phi(- 1.35) = 0.0885.$$

EJEMPLO 6 La probabilidad que un disparo acierte en el blanco es de $\frac{9}{19}$. Calcular la probabilidad que en 100 disparos se acierte por lo menos 45.

SOLUCION

X = número de veces que se acierta en el blanco en los 100 disparos.

$R_X = \{0,1, \ldots, 100\}$

X tiene una distribución binomial con $n = 100$ y $p = P[E] = 9/19$.
Aproximaremos la distribución binomial a la distribución normal con

$$\mu = np = \frac{900}{19}, \qquad y \qquad \sigma = \sqrt{100(\tfrac{9}{19})(\tfrac{10}{19})} = \frac{30\sqrt{10}}{19} = \frac{94.8}{19}$$

$$P[X \geq 45\,] = 1 - P[X < 45\,] = 1 - \Phi\left[\frac{44.5 - \frac{900}{19}}{\frac{94.8}{19}}\right] = 1 - \Phi(-\,0.58)$$

$$= 1 - 0.2810 = 0.7190\ .$$

EJEMPLO 7 Sea $f(x) = 1/x^2$, $1 < x < \infty$ la función de densidad de la variable aleatoria **X**. Considerando una muestra de tamaño 72 de **X**, determinar aproximadamente la probabilidad que más de 50 de dichas observaciones muestrales sean menores que tres.

SOLUCION La probabilidad de obtener un valor de **X** menor que 3 en una extracción cualquiera es

$$P[X < 3\,] = \int_1^3 \frac{1}{x^2}\,dx = -\frac{1}{x}\Big|_1^3 = -\frac{1}{3} + 1 = \frac{2}{3}\ .$$

Cada extracción es un ensayo de Bernoulli con $E = [X < 3\,]$ y $F = [X \geq 3]$. Puesto que la población es continua (infinito no numerable) $P[E] = \frac{2}{3}$ es constante en cada extracción y ademas los resultados de cada ensayo son independientes . Entonces, si

Y = número de observaciones muestrales con valores menores que 3.

$R_Y = \{0,1,2, \ldots , 72\}$

Y tiene una distribución binomial con $n = 72$ y $p = \frac{2}{3}$.

$$P[Y = y\,] = \binom{72}{y}(\tfrac{2}{3})^y(\tfrac{1}{3})^{72-y} \qquad , \qquad y = 0,1,2, \ldots , 72$$

Puesto que $np = 72.\,\frac{2}{3} = 48 > 5$, aproximamos la distribución binomial a la normal con

$$\mu = np = 48 \qquad y \qquad \sigma = \sqrt{npq} = \sqrt{72 \cdot \tfrac{2}{3} \cdot \tfrac{1}{3}} = 4\ .$$

$$P[Y > 50\,] = P[Y \geq 51] = P\left[\frac{Y - np}{\sqrt{npq}} \geq \frac{50.5 - 48}{4}\right]$$

$$= \Phi 1 - \Phi(0.63) = 1 - 0.7357 = 0.2643\ .$$

6.4.2 APROXIMACION DE LA DISTRIBUCION DE POISSON A LA NORMAL

La aproximación de la *distribución de Poisson a la normal*, se hace teniendo en cuenta que si, $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes de Poisson, cada una con parámetro λ, entonces

$$X = \sum_{i=1}^{n} X_i$$

es una variable aleatoria de Poisson con parámetro $n\lambda$. (propiedad reproductiva). Por el teorema central del límite, la variable aleatoria

$$Z = \frac{X - n\lambda}{\sqrt{n\lambda}} = \frac{\bar{X}_n - \lambda}{\sqrt{\lambda/n}}$$

tiene aproximadamente una distribución N(0,1), para n suficientemente grande. La aproximación de la distribución de Poisson a la normal se mejora conforme aumenta el valor del parámetro $n\lambda$, de la suma. En la práctica se considera una aproximación buena cuando $n\lambda > 5$. Por lo tanto, si el parámetro comun λ de los sumandos es pequeña, el número de variables aleatorias n debe ser grande. Si λ es grande, n puede reducirse en forma correspondiente.

La distribución normal es contínua y de Poisson es discreta, por lo tanto, para aproximar la distribución de Poisson por el área bajo la curva normal se usa el ***factor de corrección de continuidad***. Es decir

$$Z \simeq \frac{X \pm 0.5 - n\lambda}{\sqrt{n\lambda}}$$

El cuadro siguiente presenta algunas aproximaciones de la distribución de Poisson a la normal.

Probabilidad de Poisson del evento que se desea calcular	Probabilidad con la corrección por continuidad	en términos de la función de distribución normal estándar Φ
$P[X = x]$	$= P[x - 0.5 \le X \le x + 0.5]$	$= \Phi\left[\frac{x+0.5-n\lambda}{\sqrt{n\lambda}}\right] - \Phi\left[\frac{x-0.5-n\lambda}{\sqrt{n\lambda}}\right]$
$P[X \le x]$	$= P[X \le x + 0.5]$	$= \Phi\left[\frac{x + 0.5 - n}{\sqrt{n\lambda}}\right]$
$P[X < x] = P[X \le x - 1]$	$= P[X \le x - 1 + 0.5]$	$= \Phi\left[\frac{x - 0.5 - n\lambda}{\sqrt{n\lambda}}\right]$
$P[X \ge x]$	$= P[X \ge x - 0.5]$	$= 1 - \Phi\left[\frac{x - 0.5 - n\lambda}{\sqrt{n\lambda}}\right]$
$P[X > x] = P[X \ge x + 1]$	$= P[X \ge x + 1 - 0.5]$	$= 1 - \Phi\left[\frac{x + 0.5 - n\lambda}{\sqrt{n\lambda}}\right]$
$P[a \le x \le b]$	$= P[a - 0.5 \le X \le b + 0.5]$	$= \Phi\left[\frac{b+0.5-n\lambda}{\sqrt{n\lambda}}\right] - \Phi\left[\frac{x-0.5-n\lambda}{\sqrt{n\lambda}}\right]$

EJEMPLO 8 El número de vehículos que llegan por minuto a la caseta de peaje de una determinada autopista tiene una distribución de Poisson con $\mu = 2.5$. Determinar la probabilidad que en cualquier período dado de 10 minutos

(a) lleguen no más de 20 vehículos.

(b) lleguen entre 20 y 30 vehículos inclusive.

SOLUCION Definimos la siguiente variable aleatoria

X = número de vehículos que llegan por cada 10 minutos a la caseta de peaje

$R_X = \{0,1,2, \ldots\}$

puesto que la variable aleatoria X se escribe $X = X_1 + X_2 + \ldots + X_{10}$, donde

X_i = número de vehículos que llegan por minuto. $(\lambda = 2.5 = \mu)$

X tiene una distribución de Poisson con parámetro $n\lambda = 10(2.5) = 25 > 5$. por lo tanto se aproxima la distribución de Poisson por la normal con $\mu = 25$ y $\sigma = \sqrt{25} = 5$.

(a) $P[X \le 20] = P[X \le 20.5] = P\left[\frac{X - n\lambda}{\sqrt{n\lambda}} \le \frac{20.5 - 25}{\sqrt{25}}\right]$

$= \Phi(-0.9) = 0.1841$.

(b) $P[20 \leqslant X \leqslant 30] = P[19.5 \leqslant X \leqslant 30.5] = P\left[\frac{19.5 - 25}{5} \leqslant \frac{X - n\lambda}{\sqrt{n\lambda}} \leqslant \frac{30.5 - 25}{5}\right]$

$= \Phi(1.1) - \Phi(-1.1) = 0.8643 - 0.1357 = 0.7286$.

EJEMPLO 9 Las llamadas telefónicas que se reciben en un conmutador de una industria llegan como eventos de un proceso de Poisson a razón de 120 por hora. ¿Cuál es la probabilidad que lleguen entre 110 y 125 llamadas inclusive entre las 9 y las 10 a.m. de cualquier día?

SOLUCION X = número de llamadas que recibe el conmutador en una hora

$\lambda = 120$ es grande. Por lo tanto, resolveremos el problema utilizando aproximación a la normal. Esto es

$$P[110 \leq X \leq 125] = P\left[\frac{110 - 0.5 - 120}{\sqrt{120}} \leqslant \frac{X - \lambda}{\sqrt{\lambda}} \leqslant \frac{125 + 0.5 - 120}{\sqrt{120}}\right]$$

$$= \Phi(0.50) - \Phi(-0.96) = 0.6915 - 0.1685$$

$$= 0.5230.$$

6.4.3 APROXIMACION DE LA HIPERGEOMETRICA A LA NORMAL

El lector recordará que la distribución hipergeométrica se relaciona con problemas en donde el muestreo se obtiene sin reemplazo de una población finita.

Sea N el tamaño de la población finita constituída por objetos de dos clases A y B. Suponga que hay M objetos de clase A y $N - M$ de clase B. Se extrae una muestra de tamaño n sin reemplazo de la población, y se define la variable aleatoria como sigue

X = número de objetos de clase A en la muestra de tamaño n .

$R_X = \{0,1,2, \ldots, \min(M,n)\}$

Se recordará también que la variable aleatoria así definida tiene una distribución hipergeométrica con media $E(X) = n(\frac{M}{N})$ y $Var(X) = n(\frac{M}{N})(1 - \frac{M}{N})(\frac{N - n}{N - 1})$

Definimos ahora la variable aleatoria siguiente

X_i = número de objetos de clase A obtenida en la i-ésima extracción, $i = 1,2,\ldots, n$

$R_{X_i} = \{0,1\}$

Entonces, la variable aleatoria X se escribe $X = \sum_{i=1}^{n} X_i$

Suponiendo que los X_i son aproximadamente independientes, para n suficiente -

mente grande, por el teorema central del límite la variable aleatoria

$$Z \approx \frac{X - n(\frac{M}{N})}{\sqrt{n(\frac{M}{N})(1 - \frac{M}{N})}\ \sqrt{\frac{N - n}{N - 1}}}$$

tiene aproximadamente una distribución normal estándar.

La expresión $\sqrt{\frac{N - n}{N - 1}}$ se llama el FACTOR DE CORRECCION PARA POBLACION FINITA

En la práctica, la aproximación de la distribución hipergeométrica a la distribución normal se utiliza siempre que tanto $n(\frac{M}{N}) > 5$ y $n(1 - \frac{M}{N}) > 5$.

Usando el factor de corrección por continuidad se utiliza la relación

$$Z \approx \frac{X \pm 0.5 - n(\frac{M}{N})}{\sqrt{n(\frac{M}{N})(1 - \frac{M}{N})}\ \sqrt{\frac{N - n}{N - 1}}}$$

El lector puede construír una tabla de aproximaciones similar a las de la binomial y Poisson.

NOTA 1 El cálculo de la desviación estándar por la fórmula $\sqrt{n(\frac{M}{N})(1-\frac{M}{N})}\sqrt{\frac{N-n}{N-1}}$ sólo se hace cuando el tamaño n de la muestra seleccionada sin reemplazo de una población finita, es grande en comparación con el tamaño N de la población $n \geq 0.1N$. Cuando se requiere mayor exactitud se suele tomar $n \geq 0.05N$.

2 Cuando el tamaño n de la muestra sin reemplazo de una población finita, es pequeña en comparación con el tamaño N de la población ($n < 0.1N$ ó con mayor presición $n \leq 0.05N$) se puede pasar por alto al factor de corrección para población finita, se considera el problema como una aproximación binomial a la normal.

EJEMPLO 10 El decano de la facultad de Ingeniería de cierta universidad desea formar una junta directiva con los catedráticos, con 20 miembros de entre el cuerpo docente de 100. Si la selección va a ser al azar y en la facultad 25 de los catedráticos son de matemáticas, ¿cuál es la probabilidad que la junta directiva incluya

(a) Cuando menos dos catedráticos de matemáticas?

(b) Entre dos y seis catedráticos de matemáticas inclusive?

SOLUCION X = número de catedráticos de matemáticas en la junta directiva de 20.

$$R_X = \{0,1,2,\ .\ .\ .\ ,\ 20\}$$

La población es finita $N = 100$. El muestreo es sin reemplazo $n = 20$. Por lo tanto X tiene una distribución hipergeométrica.

$$P[X = x\ ] = \frac{\binom{25}{x}\binom{75}{20-x}}{\binom{100}{20}} \quad , \qquad x = 0,1,2,\ .\ .\ .\ ,\ 20$$

Desde que $n(\frac{M}{N}) = 20(\frac{25}{100}) = 5$ y $n = 20 > 0.1N$, aproximamos la distribución hipergeométrica a la distribución normal con $\mu = n(\frac{M}{N}) = 5$ y

$$\sigma = \sqrt{n(\frac{M}{N})(1-\frac{M}{N})}\ \sqrt{\frac{N-n}{N-1}} = \sqrt{5.\ \frac{3}{4}}\ \sqrt{\frac{80}{99}} \approx 1.741\ .$$

(a) $$P[X \geq 2\ ] = P[X \geq 1.5\ ] = P\left[\frac{X - n(\frac{M}{N})}{\sqrt{n(\frac{M}{N})(1-\frac{M}{N})}\ \sqrt{\frac{N-n}{N-1}}} \geq \frac{1.5 - 5}{1.741}\right]$$

$$= P[Z \geq -2.585] = 1 - P[Z \leq -2.585\ ] = 1 - 0.00485$$

$$= 0.99515\ .$$

(b) $$P[2 \leq X \leq 6\ ] = P[1.5 \leq X \leq 6.5\ ] = P\left[\frac{1.5-5}{1.741} \leq \frac{X - n(\frac{M}{N})}{\sqrt{n(\frac{M}{N})(1-\frac{M}{N})}\ \sqrt{\frac{N-n}{N-1}}} \leq \frac{6.5-5}{1.741}\right]$$

$$= \Phi(0.86) - \Phi(-2.585) = 0.8051 - 0.00485 = 0.80025.$$

PROBLEMAS 6.4

1. Con base en experiencia pasada, el 40% de todos los clientes de cierta estación de gasolina pagan sus compras con tarjeta de crédito. Si se selecciona una muestra aleatoria de 200 clientes, ¿cuál es la probabilidad que:
   (a) cuando menos 75 paguen con tarjeta de crédito?
   (b) no más de 70 paguen con tarjeta de crédito?
   (c) entre 70 y 75 clientes, inclusive, paguen con tarjeta de crédito?

2. Si una muestra de 100 tiene 3 ó menos artículos defectuosos se acepta el lote de 100. Si se supone que la probabilidad de producir artículos buenos del proceso de producción es de 0.90, ¿cuál es la probabilidad que se acepte el lote?

3. Si el 10% de los tubos de los receptores de radio se queman antes que la garantía haya expirado: Un comerciante ha vendido 100 tubos. ¿Cuál es la probabilidad que:
   (a) se vea abligado a sustituir por lo menos 20 de ellos?
   (b) sustituya por lo menos 5 y no más de 15 tubos?

4. En la playa de estacionamiento de cierta empresa grande, el registro de los automóviles de los empleados reveló, que la razón de los automóviles de manufactura nacional a extranjera es 2 a 1, es decir dos de cada tres automóviles son de manufactura nacional. Si se elige al azar 72 propietarios de autos y asumiendo una población suficientemente grande,
   (a) ¿Qué tipo de distribución de probabilidad tendrá el número de propietarios de automóviles de manufactura nacional en la muestra? ¿cuáles son los valores de sus parámetros?
   (b) Utilice una aproximación a la verdadera distribución de probabilidad para determinar la probabilidad que en la muestra haya a lo más 48 propietarios de autos de manufactura nacional.

5. Una compañía estima que ha de recibir de vuelta alrededor del 30% de los cupones especiales de premio que planea enviar por correo para un programa de promoción de ventas. Si se envián 500 cupones. ¿Cuál es la probabilidad de que se reciban de vuelta más de 165 respuestas?

6. Para decidir a cerca de un proyecto de remodelación de un sector de Lima, el consejo municipal decide seleccionar al azar 100 unidades habitacionales de este sector. Si el 40% ó maś de ellas están en mal estado, se procederá a la remodelación, en caso contrario, no se hará la remodelación.
   (a) ¿Cuál es la probabilidad que se haga la remodelación si sólo el 36% de todas las viviendas de ese sector estan en mal estado?
   (b) ¿Cuál es la probabilidad que no se haga la remodelación si el 50% de las viviendas estan en mal estado?

7. Suponga que el 10% de los neumáticos de un fabricante tienen defectos en la superficie, y que los embarca en lotes de 100.
   (a) ¿Cuál es la probabilidad (aproximada) que un lote contenga 8 ó menos neumáticos con defectos en la superficie?
   (b) Un comprador mayorista recibe 500 lotes. ¿Cuál es la probabilidad (aproximada) que al menos 140 lotes contengan 8 ó menos neumáticos con defectos en la superficie cada uno?

8. Sean $X_1, X_2, \ldots, X_{30}$ variables aleatorias de Poisson distribuídos cada

una con $\lambda = \frac{2}{3}$. Calcular :

(a) $P[15 < \sum_{i=1}^{30} X_i < 22]$ (b) $P[21 \leq \sum_{i=1}^{30} X_i < 27]$ (c) $P[\frac{1}{2} \leq \bar{X}_{30} \leq \frac{3}{4}]$

9. El número de accidentes en un tramo de 100 km de una autopista es una variable aleatoria de Poisson con media 2 por semana. ¿Cuál es la probabilidad (aproximada) que hayan menos de 100 accidentes en este tramo durante un año?

10. Se utiliza la siguiente regla para controlar el funcionamiento de una máquina que produce cierto tipo de artículos. Se selecciona una muestra aleatoria de 400 artículos cada hora. Si el número de artículos defectuosos es de 12 ó más, se detiene la máquina; y si el número de artículos defectuosos es inferior a 12, se deja que la máquina siga funcionando. ¿Cuál es la probabilidad de:
 (a) detener la máquina si está produciendo un 2% de artículos defectuosos en promedio?
 (b) dejar que la máquina siga funcionando si está produciendo un 4% de artículos defectuosos, en promedio?

11. Se tiene un telar que teje 1,000 m^2 de tela por día. La probabilidad de que no halla falla en 250 m^2 de tela es e^{-1}. ¿Cuál es la probabilidad (aproximada) que ocurra 100 ó más pero no más de 140 fallas en la producción de un mes (30 días) de trabajo del telar?

12. Tornillos de hierro de media pulgada fabricados por cierta empresa ocacionalmente no tienen ranura. Esto ocurre al azar, y la probabilidad de este hecho y de que escape a la inspección es 0.02. En una remesa de 2,500 de tales tornillos. ¿Cuál es la probabilidad que:
 (a) 64 ó más carezcan de ranuras? (b) ¿36 ó menos? ¿entre 36 y 64 inclusive?

13. Se sabe que el 5% de los tubos de radio producidos por cierto fabricante son defectuosos. Si el fabricante envía a un mayorista 1,000 lotes, cada uno de los cuales contiene 100 tubos, ¿en cuántos de estos lotes espera tener : (a) menos de 90 tubos buenos? (b) 98 ó más tubos buenos?

14. Una caja contiene 80 tarros de conserva de las que 60% son de duraznos y 40% de damascos. De un total de 50 muestras de 20 tarros cada una, sacadas de la caja con reemplazamiento. ¿En cuántos cabe esperar
 (a) igual número de tarros de cada fruta?
 (b) 12 tarros de duraznos y 8 de damascos?
 (c) 8 tarros de duraznos y 12 de damascos?

(d) 10 ó más tarros de damascos?

15. Si X es una variable aleatoria $b(n,0.55)$. Hallar el menor valor de n de modo que $P[X/n > 1/2] \geq 0.95$ (aproximadamente). Sugerencia: observe que $P[X/n > 1/2] = P[X > n/2]$

16. Un funcionario de asitencia pública en cierta área de una ciudad grande sospecha que un 10% de los niños sufren grave desnutrición. En esa área viven 2000 niños. Se selecciona una muestra de 80 niños.
    (a) ¿Qué distribución de probabilidad tiene el número de niños desnutridos en la muestra? ¿cuales son sus parámetros?
    (b) ¿Cuál es la probabilidad (aproximada) que cuando menos 5 niños estén desnutridos?

17. Los automóviles llegan a un servicio de lavado autómatico a razón de nueve cada media hora. Determine la probabilidad que en cualquier período dado de 8 horas
    (a) lleguen cuando menos 120 automóviles.
    (b) lleguen entre 120 y 150 automóviles.

18. (a) Un sistema está formado por 100 componentes que funcionan independientemente. La probabilidad que cualquier componente falle durante el período de operación es igual a 0.10. El sistema funciona si funcionan al menos 85 componentes. Calcular la probabilidad que funcione el sistema.
    (b) Suponga que el sistema anterior esté formado por n componentes, cada una con una confiabilidad de 0.90. El sistema funcionará si al menos el 80% de las componentes funcionen correctamente. Determinar n de modo que el sistema tenga una confiabilidad de 0.95.

19. Considere un sistema eléctrico en serie que contiene 30 focos, conectados de manera que ninguno prenderá si uno de ellos es defectuoso. Si los focos de la instalación se seleccionan al azar de un lote de 600 focos, 100 de los cuales son defectuosos. Determinar la probabilidad que funcionen todos los focos del sistema eléctrico.

7

DISTRIBUCIONES MUESTRALES

7.1 POBLACION Y MUESTRA

La palabra ***población*** es muy común en el lenguaje cotidiano y su uso es muy general. Al pasar diremos simplemente que ***población es el conjunto de todas las observaciones (resultados) posibles que puede tomar una variable aleatoria* X**. Según ésta definición, la distribución de la población es la distribución de la variable aleatoria **X** y la población será discreta o contínua según sea **X**.

En muchos problemas es imposible o innecesario tener todos los datos de la población. Los datos de sólo una parte de la población pueden dar la información necesaria para generalizar acerca de los parámetros de la población que por lo general son desconocidas. Una parte (subconjunto) de la población se llama una ***muestra***.

DEFINICION (*MUESTRA ALEATORIA*) Sea **X** una variable aleatoria con función de distribución $f(x)$ (función de probabilidad o función de densidad), media μ y varianza σ^2 . Una ***muestra aleatoria*** de tamaño n , de **X**, es un conjunto de n variables aleatorias $X_1, X_2, \ldots, X_n$ que cumplen :

1. Cada X_i ($i = 1,2, \ldots, n$), tiene la misma distribución que **X**. Es decir,

$$F_{X_i}(x) = F_X(x) \;, \quad \text{para todo } x \quad , \quad i = 1,2, \ldots, n$$

o $f_{X_i}(x) = f_X(x)$ para todo x , $i = 1,2, \ldots, n$

2. Las variables aleatorias $X_i (i = 1,2, \ldots, n)$ son independientes

CONSECUENCIA (a) Cada X_i tiene la misma distribución que X. Entonces

$$\mu_{X_i} = E(X_i) = E(X) = \mu$$

$$\sigma^2_{X_i} = Var(X_i) = Var(X) = \sigma^2$$

(b) La función de probabilidad conjunta de la muestra aleatoria $X_1, X_2, \ldots, X_n$ está dado por

$$p_{X_1,\ldots,X_n}(x_1,x_2,\ldots,x_n) = p_X(x_1)\ldots p_X(x_n) \quad , \quad \text{si } X \text{ es discreta.}$$

$$f_{X_1,\ldots,X_n}(x_1,x_2,\ldots,x_n) = f_X(x_1)\ldots f_X(x_n) \quad , \quad \text{si } X \text{ es continua.}$$

NOTA 1 La definición anterior se cumple, cuando la muestra proviene de una población infinita discreta o continua y cuando la muestra se extrae con reemplazamiento de una población finita.

2. Cuando la muestra se extrae sin reemplazo de una población finita, evidentemente no satisface la definición de la muestra aleatoria, pues las variables aleatorias: $X_1, X_2, \ldots, X_n$ no son independientes. Sin embargo, si el tamaño n de la muestra es muy pequeña en comparación con el tamaño de la población, se cumple aproximadamente la definición.

EJEMPLO 1 De una población normal con media 10 y varianza 12 se extrae una muestra aleatoria, $X_1, X_2, \ldots, X_{10}$. Calcular

$$P[X_1 - X_5 + X_8 \geq 13]$$

SOLUCION Si X, es la variable aleatoria de la población normal, $X \rightarrow N(10,12)$. Entonces, por ser $X_1, X_2, \ldots, X_{10}$ una muestra aleatoria, se cumple:

1. $X_i \rightarrow N(10,12)$, $i = 1,2, \ldots, 10$
2. X_i , $i = 1,2, \ldots, 10$, son variables aleatorias independientes.
3. Sea $Y = X_1 - X_5 + X_8$, entonces $Y \rightarrow N(\mu_Y, \sigma^2_Y)$, por ser suma de variables aleatorias normales e independientes (propiedad reproductiva), donde

$$\mu_Y = E(Y) = E(X_1 - X_5 + X_8) = E(X_1) - E(X_5) + E(X_8)$$

$$= 10 - 10 + 10 = 10 .$$

$$\sigma_Y^2 = Var(Y) = Var(X_1 - X_5 + X_8)$$
$$= Var(X_1) + Var(X_5) + Var(X_8)$$
$$= 12 + 12 + 12 = 36$$

de donde $\sigma_Y = 6$.

$$P[X_1 - X_5 + X_8 \geq 13] = P[Y \geq 13]$$
$$= P[\frac{Y - \mu_Y}{\sigma_Y} \geq \frac{13 - 10}{6}]$$
$$= P[Z \geq \frac{1}{2}] = 1 - \Phi(\frac{1}{2})$$
$$= 1 - 0.6915 = 0.3085 .$$

EJEMPLO 2 Suponga que X es una variable aleatoria normal estándar y que X_1,X_2,X_3,X_4 es una muestra aleatoria de X; se define

$\bar{X} = \sum_{i=1}^{4} X_i/4$. Calcular la probabilidad que $|\bar{X}| > \frac{1}{2}$

SOLUCION $X \to N(0,1)$

Desde que X_i ,(i = 1,2,3,4) es una muestra aleatoria, se cumple:

1 $X_i \to N(0,1)$, $i = 1,2,3,4$

2 X_i,(i = 1,2,3,4) son variables aleatorias independientes. Entonces

$$\bar{X} \to N(\mu_{\bar{X}},\sigma_{\bar{X}}^2)$$

por ser $\bar{X}$ una combinación lineal de variables aleatorias normales. Donde

$$\mu_{\bar{X}} = E(\sum_{i=1}^{4}X_i/4) = \frac{1}{4}\sum_{i=1}^{4}E(X_i) = 0 .$$

$$\sigma_{\bar{X}}^2 = Var(\sum_{i=1}^{4}X_i/4) = \frac{1}{4^2}\sum_{i=1}^{4}Var(X_i) = \frac{4}{4^2} = \frac{1}{4} .$$

luego, $\sigma_{\bar{X}} = \frac{1}{2}$.

$$P[|\bar{X}| > 1/2] = P[(\bar{X} > 1/2) \cup (\bar{X} < - 1/2)]$$
$$= P[\bar{X} > 1/2] + P[\bar{X} < - 1/2]$$
$$= P\left[\frac{\bar{X} - \mu}{\sigma} > \frac{1/2 - 0}{1/2}\right] + P\left[\frac{\bar{X} - \mu}{\sigma} < \frac{- 1/2 - 0}{1/2}\right]$$
$$= P[Z > 1] + P[Z < - 1]$$

$$= 2\,P[Z > 1] = 2[1 - P[Z \leq 1]]$$

$$= 2[1 - 0.8413] = 0.3174 .$$

EJEMPLO 3 Una urna contiene 4 bolas blancas y 1 roja. Se van a extraer de la urna bolas con reemplazamiento hasta obtener la roja. Sea X el número de sacadas necesarias. Si se toma una muestra aleatoria de 5 observaciones de X. ¿Cuál es la función de probabilidad conjunta de la muestra?

SOLUCION Obtendremos primero la función de probabilidad de X. El experimento es "extraer una bola con reemplazo de la urna hasta obtener la roja".

X = número de extracciones hasta obtener la roja.

$R_X = \{1,2,3,4, \ldots\}$

$$P[X = 1] = \frac{1}{5} = \frac{4^{1-1}}{5}$$

$$P[X = 2] = P[BR] = P[B]\,P[R|B] = \frac{4}{5}\cdot\frac{1}{5} = \frac{4^{2-1}}{5^2}$$

$$P[X = 3] = P[BBR] = P[B]\,P[B|B]\,P[R|BB] = \frac{4}{5}\cdot\frac{4}{5}\cdot\frac{1}{5} = \frac{4^{3-1}}{5^3}$$

En general se cumple,

$$p_X(x) = P[X = x] = \frac{4^{x-1}}{5^x} , \qquad x = 1,2,3, \ldots \qquad (1)$$

Se toma ahora una muestra aleatoria de 5 observaciones de la variable aleatoria X con función de probabilidad definida por (1).

Sea X_1, X_2, X_3, X_4, X_5, la muestra aleatoria con valores x_1, x_2, x_3, x_4, y x_5 respectivamente. Entonces,

$$p_{X_1 \ldots X_5}(x_1, x_2, \ldots, x_5) = p_X(x_1)p_X(x_2) \ldots p_X(x_5)$$

$$= \frac{4^{x_1-1}}{5^{x_1}} \cdot \frac{4^{x_2-1}}{5^{x_2}} \cdot \frac{4^{x_3-1}}{5^{x_3}} \cdot \frac{4^{x_4-1}}{5^{x_4}} \cdot \frac{4^{x_5-1}}{5^{x_5}}$$

$$= 4^{\sum_{i=1}^{5}(x_i - 1)} \cdot 5^{-\sum_{i=1}^{5} x_i}$$

EJEMPLO 4 Suponga que determinados bulbos electrónicos tienen vidas medias que están distribuídas exponencialmente con parámetro λ. Si se toma una muestra aleatoria de 10, de estos bulbos y se representa con X_i = la duración del i-ésimo bulbo, $i = 1,2, \ldots, 10$, ¿cuál es la función de densidad conjunta de la muestra?

SOLUCION La función de densidad de la población es,

$$f_X(x) = \lambda e^{-\lambda x} \quad , \quad x > 0$$

Se toma una muestra aleatoria de tamaño 10 ,

X_i , $i = 1,2, \ldots , 10$ donde X_i = la duración del i-ésimo bulbo,

$i = 1,2, \ldots , 10$.

Entonces, la función de densidad conjunta de la muestra es

$$\begin{aligned} f_{X_1,\ldots,X_{10}}(x_1,x_2,\ldots,x_{10}) &= f_X(x_1)f_X(x_2)\ldots f_X(x_{10}) \\ &= \lambda e^{-\lambda x_1} . \lambda e^{-\lambda x_2} \ldots \lambda e^{-\lambda x_{10}} \\ &= \lambda^{10} e^{-\lambda \sum_{i=1}^{10} x_i} \quad , \quad x_i > 0 \end{aligned}$$

PROBLEMAS 7.1

1. Suponga que $X_1,X_2, \ldots , X_{10}$ es una muestra aleatoria de una variable aleatoria normal estándar. Calcular

$$P[\bar{X} < 0.271] \quad \text{donde} \quad \bar{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$

2. De una población normal con $\mu = 10$ y $\sigma = 2$ se extraerá una muestra aleatoria X_1,X_2,X_3,X_4 . Calcular:

   (a) $P[X_1 + X_2 + X_3 + X_4 \geq 48]$; (b) $P[X_3 > 13]$

3. Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una variable aleatoria X con función de probabilidad $f(x) = \frac{x}{6}$, $x = 1,2,3$. Calcular la media y la varianza de $X_1 + X_2$.

4. Sea X una variable aleatoria con función de densidad $f(x) = 6x(1 - x)$, $0 < x < 1$. Se extrae una muestra aleatoria $X_1,X_2, \ldots, X_{10}$ de esta población Calcular

   (a) la media y la varianza de $\sum X_i$

   (b) la función de densidad conjunta de la muestra.

5. Suponga que todos los ladrones tienen una probabilidad constante p de ser sorprendidos mientras cometen un robo; y sea X el número de robos cometidos hasta que el ladrón es atrapado. Si se toma una muestra aleatoria de 10 ladrones, ¿cuál es la función de probabilidad conjunta del número de robos efectuados por cada uno de ellos antes de ser atrapados por primera vez?

6. Suponga X está distribuída uniformente en el intervalo $[0,1]$. Si se toma una muestra aleatoria de 10 observaciones de X, ¿cuál es la función de densidad conjunta de la muestra?

7. Suponga que una variable aleatoria se distribuye normalmente con media μ y varianza σ^2. Se extrae una muestra aleatoria de 5 observaciones. Determinar la función de densidad conjunta de la muestra.

8. Suponga que los resultados de los exámenes de manejo están distribuídos normalmente con media 50 y varianza 25, y que cualquiera que obtenga más de 60 puntos en la prueba será un buen piloto. Se presentan al examén 10 personas. ¿Cuál es la probabilidad que:
   (a) no hayan buenos pilotos en el grupo de 10 personas?
   (b) haya uno ó más buenos pilotos en el grupo de 10?

9. Un lote consiste de N transistores, y de estos $M(M < N)$ son defectuosos. Se selecciona aleatoriamente dos transistores sin reemplazamiento de éste lote y se determina si son defectuosos o no defectuosos. Se define la variable aleatoria

$$X_i = \begin{cases} 1 & , \quad \text{si el } i\text{-ésimo transistor es no defectuoso} \\ 0 & , \quad \text{si el } i\text{-ésimo transistor es defectuoso} \end{cases}$$

$i = 1,2$.

Determinar la función de probabilidad conjunta para X_1 y X_2. ¿Cuales son las funciones de probabilidad marginal para X_1 y X_2 ? ¿son las variables aleatorias X_1 y X_2 independientes?

7.2 DISTRIBUCIONES MUESTRALES

Un problema central en estadística (como hemos mencionado en 7.1) es, estudiar una población con función de densidad $f(x,\theta)$, donde la fórmula de la función de densidad es conocida (o se supone conocida) pero contiene un parámetro desconocido θ. Si se conoce θ, la función de densidad estará completamente especificado. El procedimiento es tomar una muestra aleatoria $X_1, X_2, \ldots, X_n$ de tamaño n de la población y buscar alguna función de esta muestra que estime el parámetro desconocido θ. Este problema se formulará con más detalle en el proximo capítulo. En esta sección daremos la definición de estadística, distribuciones muestrales y los momentos muestrales.

7.2.1. ESTADISTICO Y MOMENTOS MUESTRALES

DEFINICION 7.2.1 Un *estadístico* es una variable aleatoria que depende solamente de la muestra observada.

EJEMPLO 1 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una población X, entonces

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{y} \qquad S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

son estadísticos.

DEFINICION 7.2.2 La distribución de probabilidad de un estadístico de llama *distribución muestral*

DEFINICION 7.2.3 La desviación estándar de la distribución muestral de un estadístico, se llama el *error estándar del estadístico* .

DEFINICION 7.2.4 MOMENTOS MUESTRALES Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población con función de densidad $f(x)$. Entonces, el *r-ésimo momento muestral alrededor del origen* 0, denotado por M_r' , se define así,

$$M_r' = \frac{1}{n}\sum_{i=1}^{n} X_i^r \quad .$$

En particular, si $r = 1$, tenemos la *media muestral*, la cual usualmente se denota por $\bar{X}$ o $\bar{X}_n$; esto es:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

El *r-ésimo momento muestral al rededor de* $\bar{X}_n$, denotado por M_r , se define

$$M_r = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^r$$

NOTA Los momentos muestrales son ejemplos de estadístico .

DEFINICION Varianza muestral. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población X; entonces $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, para $n > 1$, se llama varianza muestral.

En el capítulo 3 , hemos definido el r-ésimo momento al rededor del origen 0 de una variable aleatoria X así, $E[X^r] = \mu_r'$. En adelante diremos que

$E(X^r)$ es el r-ésimo momento de la población con distribución de probabilidad $f(x) = f_X(x)$.

TEOREMA 7.2.1 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población con distribución de probabilidad $f(x)$. El valor esperado del r-ésimo momento muestral (alrededor del origen 0) es igual al r-ésimo momento de la población; es decir

$$E[M'_r] = \mu'_r \qquad \text{(si } \mu'_r \text{ existe) .}$$

y $$\operatorname{Var}[M'_r] = \frac{1}{n}\{[E[X^{2r}] - (E[X^r])^2\} = \frac{1}{n}[\mu'_{2r} - (\mu'_r)^2] \quad \text{(si } \mu'_{2r} \text{ existe)}$$

En particular, si $r = 1$, tenemos el siguiente corolario

COROLARIO 7.2.1 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población X, distribuida con media μ y varianza σ^2. Sea $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ la media muestral; entonces

$$E(\bar{X}) = \mu \qquad \text{y} \qquad \operatorname{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

DEMOSTRACION

$$E(\bar{X}) = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E(X_i) \qquad \text{propiedad de la esperanza}$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu = \mu . \qquad \text{por ser } X_i \text{ muestra aleatoria}$$

$$\operatorname{Var}(\bar{X}) = \operatorname{Var}[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n^2}\operatorname{Var}[\sum_{i=1}^{n} X_i] \qquad \text{propiedad de la varianza}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{Var}(X_i) \quad , \quad X_i \text{ muestra aleatoria (independientes)}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n} \quad , \quad X_i \text{ muestra aleatoria}$$

En la investigación, en las encuestas en los negocios etc., el muestreo se -

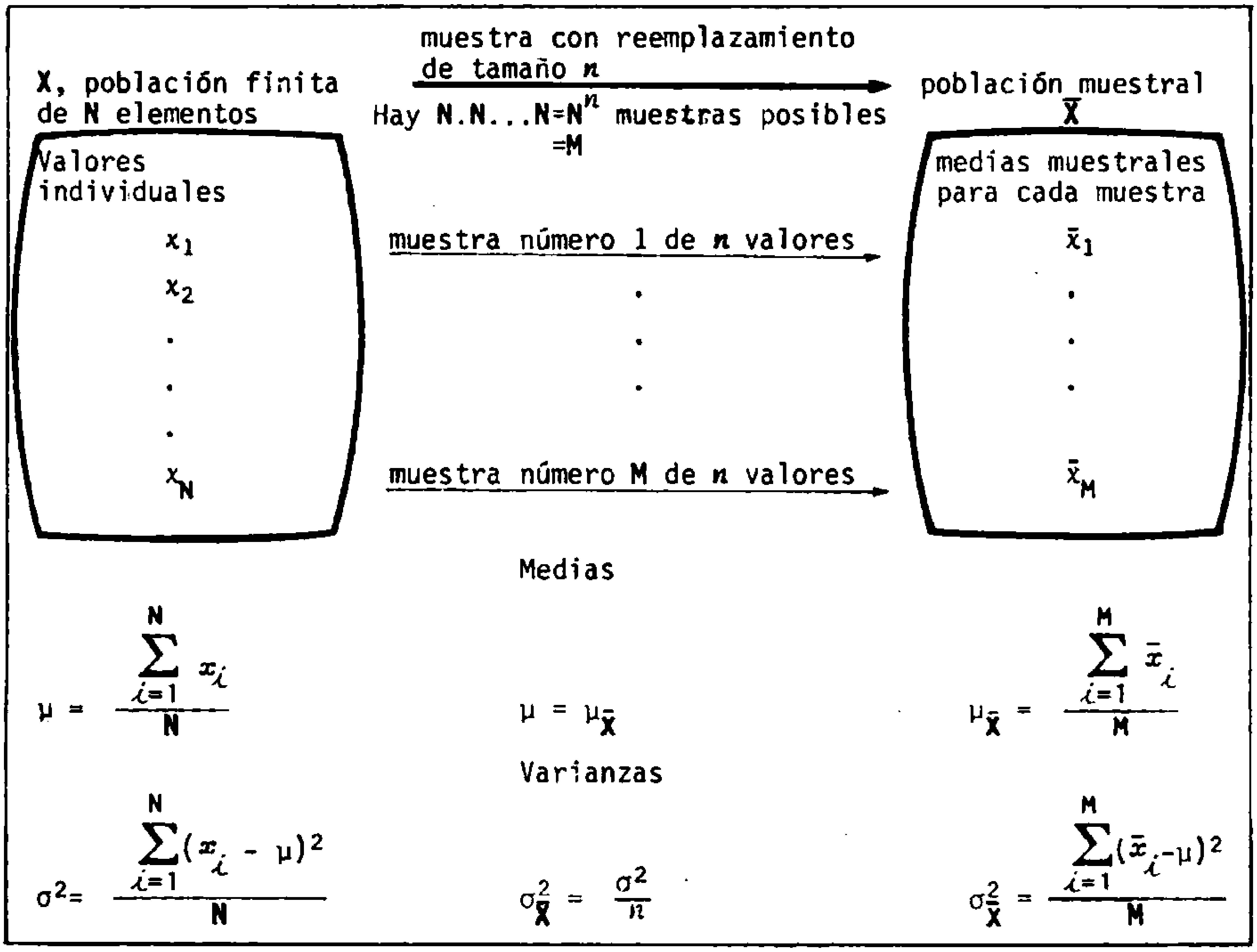
efectua sin reemplazo en poblaciones finitas. En estos casos se cumple el corolario siguiente.

COROLARIO 7.2.2 Sea $X_1, X_2, \ldots, X_n$ una muestra de tamaño n, extraida sin reemplazamiento, de una población finita de N elementos, distribuida con media μ y varianza σ^2, entonces

$$E(\bar{X}) = \mu \qquad \text{y} \qquad Var(\bar{X}) = \frac{\sigma^2}{n}\cdot\frac{N-n}{N-1}$$

El factor $\frac{N-n}{N-1}$ se llama *factor de corrección* para población finita

Los diagramas siguientes, da una visión esquemática de los corolarios 7.2.1 y 7.2.2, para una población finita.



X, población finita de N elementos

Valores individuales

x_1, x_2, ..., x_N

muestra sin reemplazo de tamaño n →

Hay $P_N^n = \frac{N!}{(N-n)!} = M$ muestras posibles considerando el orden

$N(N-1)\ldots(N-n+1)$

muestra número 1 de n valores →

⋮

muestra número M de n valores →

población muestral $\bar{X}$

medias muestrales para cada muestra

$\bar{x}_1$, ..., $\bar{x}_M$

Medias

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \qquad \mu = \mu_{\bar{X}} \qquad \mu_{\bar{X}} = \frac{\sum_{i=1}^{M} \bar{x}_i}{M}$$

Varianzas

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \qquad \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\cdot\frac{N-n}{N-1} \qquad \sigma_{\bar{X}}^2 = \frac{\sum_{i=1}^{M}(\bar{x}_i - \mu)^2}{M}$$

NOTA La desviación estándar muestral está dado por :

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$, si el muestreo es con reemplazamiento

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$, si el muestreo es sin reemplazamiento de una población finita.

EJEMPLO 2 Una población X consta de cinco elementos

$$X = \{1,2,3,4,5\}$$

(a) Calcule la media μ y la desviación estándar de la población,

(b) grafique la distribución de probabilidad de la población X.

Se extrae una muestra de tamaño 2 de la población (1) con reemplazamiento (2) sin reemplazamiento.

(c) Escriba todas las muestras posibles de tamaño 2 de X (en ambos casos)

(d) Calcular las medias $\bar{X}$ de cada una de estas muestras.

(e) Calcular la media $\mu_{\bar{X}}$ de las $\bar{X}$. ¿Que se verifica?

(f) Calcular la desviación estándar $\sigma_{\bar{X}}$ de las $\bar{X}$ determinando la desviación estándar de todas las posibles medias de las muestras obtenidas en (d) con respecto a la media de la población, μ . Luego calcule la desviación estándar de la media basado en que σ se conoce. Verifique que los dos valores de la desviación estándar son iguales.

(g) Grafique la distribución de probabilidad de $\bar{X}$. Compare con la gráfica de la distribución de X.

SOLUCION (a) Se obtiene de dos formas

PRIMERA FORMA Usando la definición de media y varianza poblacional.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3,$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{5} = \frac{10}{5} = 2.$$

luego $\sigma = \sqrt{2}$.

SEGUNDA FORMA Se determina la distribución de probabilidad de la población - X .

x	1	2	3	4	5
$p(x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

$$\mu = \sum_{x \in R_X} x\,p(x) = \frac{1}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5} + \frac{5}{5} = 3.$$

$$E(X^2) = \sum_{x \in R_X} x^2 p(x) = \frac{1}{5} + \frac{4}{5} + \frac{9}{5} + \frac{16}{5} + \frac{25}{5} = 11.$$

$$\sigma^2 = E(X^2) - (\mu)^2 = 11 - 9 = 2 .$$

luego $\sigma = \sqrt{2}$.

(b)

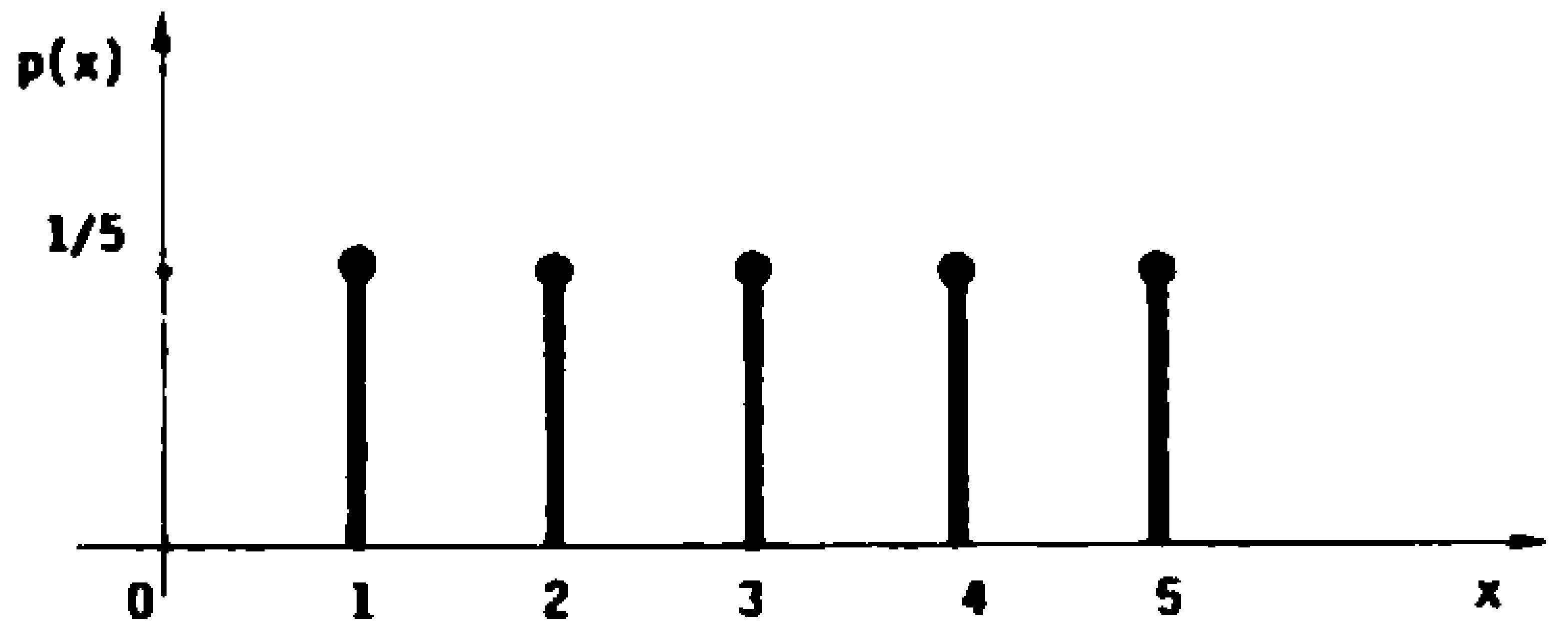


Fig. 7.2.1. Gráfica de la distribución de probabilidad de la población

1. Muestreo con reemplazamiento. $n = 2$. Hay $M = N^2 = 5^2 = 25$ posibles muestras.

(1C)

1° Extracción / 2° Extracción	1	2	3	4	5
$\bar{x}$	1.0	1.5	2.0	2.5	3.0
1	11	12	13	14	15
$\bar{x}$	1.5	2.0	2.5	3.0	3.5
2	21	22	23	24	25
$\bar{x}$	2.0	2.5	3.0	3.5	4.0
3	31	32	33	34	35
$\bar{x}$	2.5	3.0	3.5	4.0	4.5
4	41	42	43	44	45
$\bar{x}$	3.0	3.5	4.0	4.5	5.0
5	51	52	53	54	55

Tabla de muestras posibles. En la parte superior de cada muestra está su media

(1d) Las medias muestrales llevados a una tabla de frecuencias

$\bar{x}_i$	f_i	$f_i\bar{x}_i$	$\bar{x}_i - \mu$	$(\bar{x}_i - \mu)^2$	$f_i(\bar{x}_i - \mu)^2$
1.0	1	1.0	- 2.0	4.00	4.00
1.5	2	3.0	- 1.5	2.25	4.50
2.0	3	6.0	- 1.0	1.00	3.00
2.5	4	10.0	- 0.5	0.25	1.00
3.0	5	15.0	0.0	0.00	0.00
3.5	4	14.0	0.5	0.25	1.00
4.0	3	12.0	1.0	1.00	3.00
4.5	2	9.0	1.5	2.25	4.50
5.0	1	5.0	2.0	4.00	4.00
Total	25	75.0			25.00

(1e) $$\mu_{\bar{X}} = \frac{\sum f_i\bar{x}_i}{M} = \frac{75.0}{25} = 3.0$$

Se verifica el corolario 1. Es decir $\mu_{\bar{X}} = \mu = 3$.

(1f) $$\sigma_{\bar{X}} = \sqrt{\frac{\sum f_i(\bar{x}_i - \mu)^2}{M}} = \sqrt{\frac{25}{25}} = 1 .$$

Aplicando el corolario 1. $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{\sqrt{2}}{\sqrt{2}} = 1$.

Por lo tanto se verifica que $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$.

(1g) Cálculo de la distribución de probabilidad de $\bar{X}$. Se obtiene de dos formas

(i) Directamente de la tabla de distribución de frecuencias $p(x_i) = \dfrac{f_i}{M}$, por ejemplo $p(1.0) = \dfrac{1}{25}$, $p(3.5) = \dfrac{4}{25}$, etc.

(ii) El muestreo es con reemplazo. Entonces el resultado de la segunda extracción no depende del resultado de la primera extracción. Así

$p(1.0) = P[\{(1,1)\}] = \dfrac{1}{5} \cdot \dfrac{1}{5} = \dfrac{1}{25}$; $p(1.5) = P[\{(1,2),(2,1)\}]$

$= \dfrac{1}{5} \times \dfrac{1}{5} + \dfrac{1}{5} \times \dfrac{1}{5} = \dfrac{2}{25}$, etc.

$\bar{x}_i$	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
$p(\bar{x}_i)$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{5}{25}$	$\frac{4}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{1}{25}$

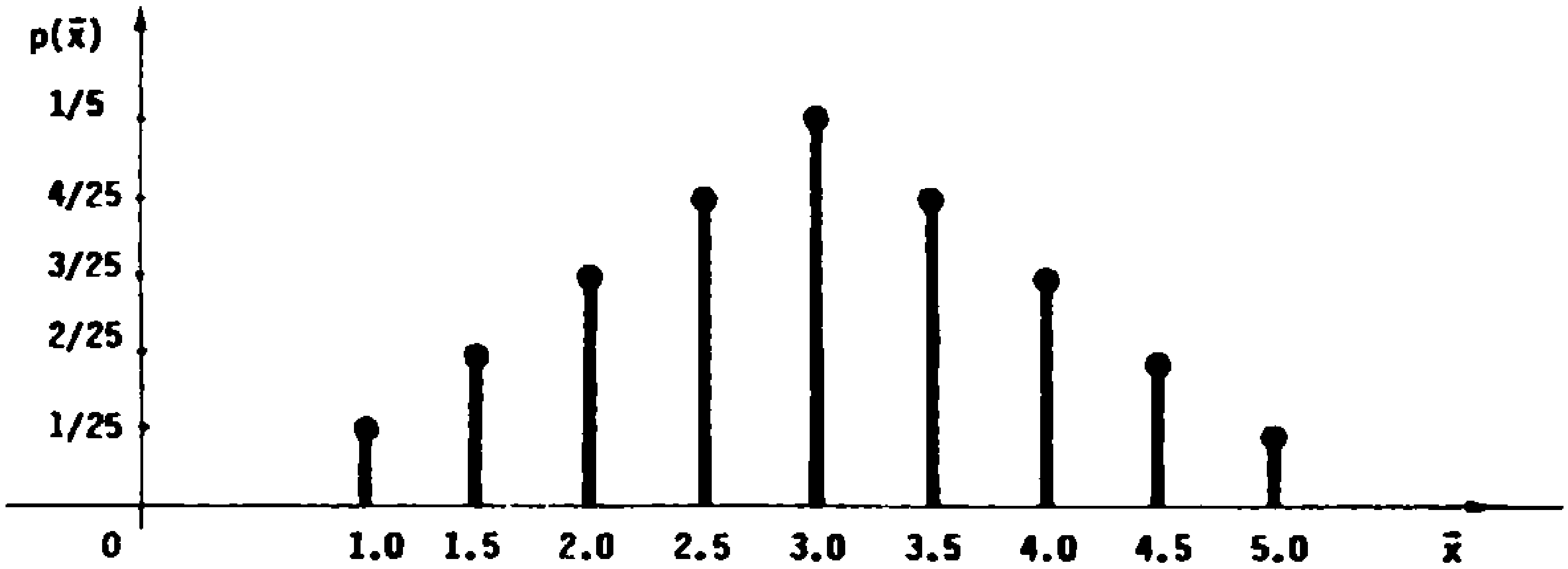


Fig. 7.2.2. Gráfica de la distribución de probabilidad de $\bar{X}$

COMPARACION La gráfica de X es uniforme, de $\bar{X}$ es un triángulo isosceles.

2. Muestreo sin reemplazamiento $n = 2$. Hay $N(N - 1) = 5 \times 4 = 20$ muestras posibles. Teniendo en cuenta que el orden no interesa, es suficiente trabajar con $M = \binom{N}{n} = \dfrac{N!}{n!(N-n)!} = \dfrac{5!}{2!.3!} = 10$ muestras.

(2c)

2° Extracción \ 1° Extracción	1	2	3	4	5
$\bar{x}$		1.5	2.0	2.5	3.0
1		1 2	1 3	1 4	1 5
$\bar{x}$			2.5	3.0	3.5
2	2 1		2 3	2 4	2 5
$\bar{x}$				3.5	4.0
3	3 1	3 2		3 4	3 5
$\bar{x}$					4.5
4	4 1	4 2	4 3		4 5
$\bar{x}$					
5	5 1	5 2	5 3	5 4	

Tabla de muestras posibles. En la parte superior de cada muestra esta su media.

(2d) Las medias muestrales tabulados en una tabla de distribución de frecuencias.

$\bar{x}_i$	f_i	$f_i\bar{x}_i$	$\bar{x}_i - \mu$	$(\bar{x}_i - \mu)^2$	$f_i(\bar{x}_i - \mu)^2$
1.5	1	1.5	- 1.5	2.25	2.25
2.0	1	2.0	- 1.0	1.00	1.00
2.5	2	5.0	- 0.5	0.25	0.50
3.0	2	6.0	0.0	0.00	0.00
3.5	2	7.0	0.5	0.25	0.50
4.0	1	4.0	1.0	1.00	1.00
4.5	1	4.5	1.5	2.25	2.25
Total	10 = M	30.0			7.50

(2e) $$\mu_{\bar{X}} = \frac{\sum f_i \bar{x}_i}{M} = \frac{30.0}{10} = 3 .$$

Se verifica el corolario 2. Es decir $\mu_{\bar{X}} = \mu = 3$.

(2f) $$\sigma_{\bar{X}} = \sqrt{\frac{\sum f_i(\bar{x}_i - \mu)^2}{M}} = \sqrt{\frac{7.50}{10}} = \frac{\sqrt{3}}{2}$$

Por otro lado aplicando el corolario 2, obtenemos

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}} = \frac{\sqrt{2}}{\sqrt{2}} \sqrt{\frac{5 - 2}{5 - 1}} = \frac{\sqrt{3}}{2}$$

Es decir, se verifica que $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} \sqrt{\dfrac{N - n}{N - 1}}$

(2g) La distribución de probabilidad de $\bar{X}$ se obtiene de dos formas

(*i*) Directamente de la tabla de distribución de frecuencia, $p(x_i) = \dfrac{f_i}{M}$

(*ii*) De la tabla de muestras posibles, teniendo en cuenta que el resultado de la segunda extracción depende de la primera, se aplica probabilidad condicional así

$$p(1,5) = P[\{(1,2),(2,1)\}] = \frac{1}{5} \times \frac{1}{4} + \frac{1}{5} \times \frac{1}{4} = \frac{2}{20} = \frac{1}{10}$$

$$p(2.0) = P[\{(1,3),(3,1)\}] = \frac{1}{5} \times \frac{1}{4} + \frac{1}{5} \times \frac{1}{4} = \frac{2}{20} = \frac{1}{10} \quad . \text{ etc.}$$

$\bar{x}_i$	1.5	2.0	2.5	3.0	3.5	4.0	4.5
$p(\bar{x}_i)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

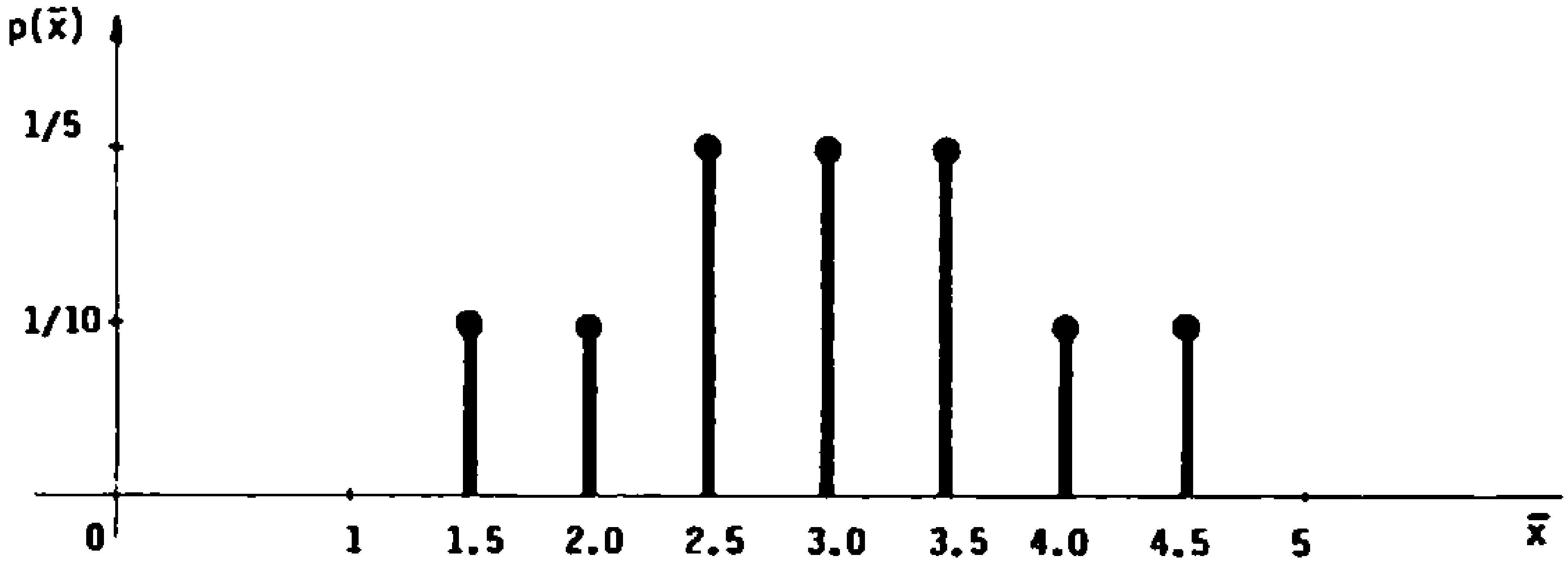


Fig. 7.2.3

7.2.2 DISTRIBUCION MUESTRAL DE LA MEDIA

Sea X una población con distribución de probabilidad $f(x)$, media μ y varianza σ^2. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de X. La media muestral es $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Entonces

1. $\mu_{\bar{X}} = \mu$ y $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$ corolario 7.2.1

2. Para n suficientemente grande, por el teorema central del límite, la variable aleatoria $\bar{X}$ se distribuye aproximadamente por una normal con media μ y varianza σ^2/n . En símbolos

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Por lo tanto, la variable aleatoria

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma}$$

tiene aproximadamente una distribución normal estándar.

3. Si la población X tiene una distribución normal con media μ varianza σ^2 , la muestra aleatoria $X_1, X_2, \ldots, X_n$ son variables aleatorias distribuidas normalmente e identicamente con media μ y varianza σ^2. Entonces $\bar{X}$ tiene una distribución normal con media μ y varianza σ^2/n para todo n . En consecuencia, la variable aleatoria

$$Z = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma}$$

tiene una distribución normal estándar.

La media de $\bar{X}$ es la misma de la población, μ y su varianza se reduce a σ^2/n . Una comparación de la distribución de la media muestral $\bar{X}$ y una de las variables originales distribuídas normalmente se muestra en la fig. 7.2.4

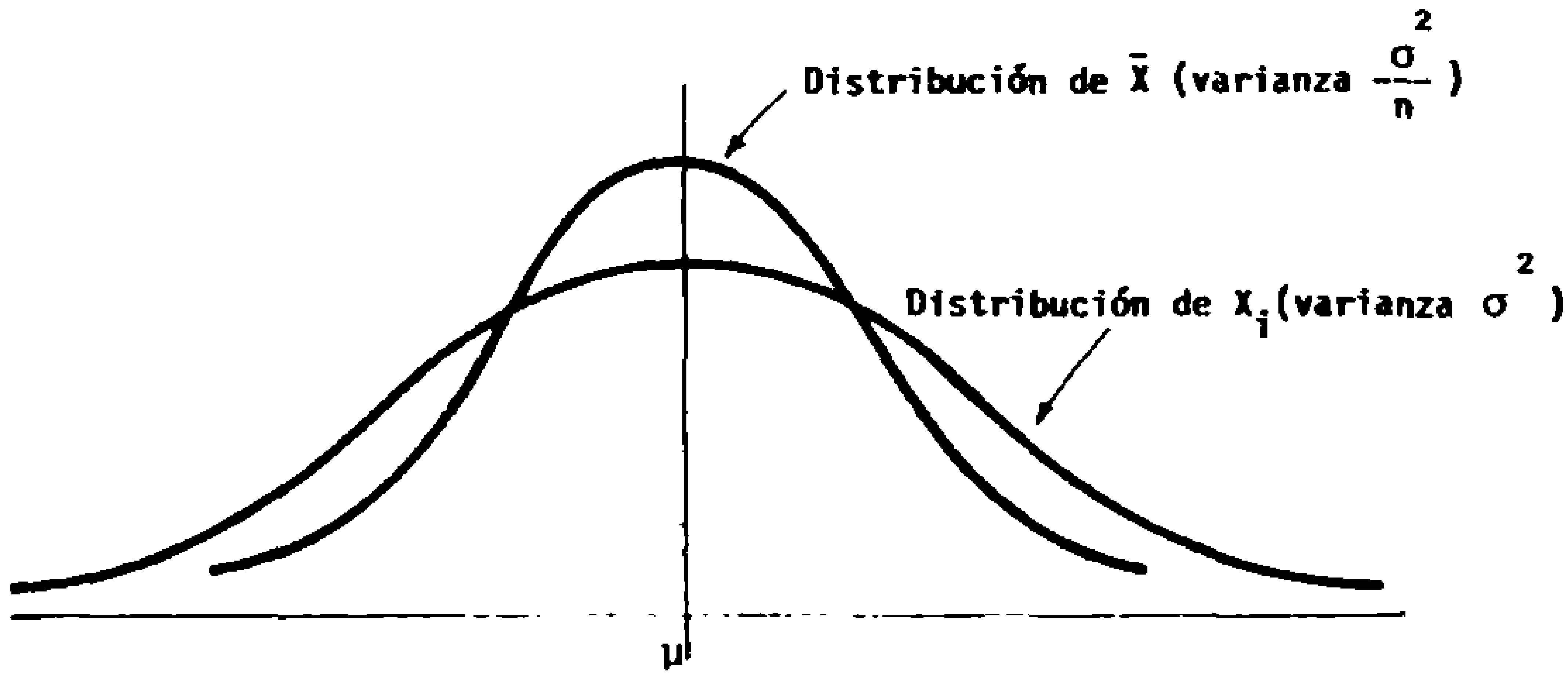


Fig. 7.2.4. Comparación de la distribución de probabilidad de X_i con la distribución de probabilidad de $\bar{X}$

El teorema 7.2.2 formaliza lo expresado en el párrafo anterior.

TEOREMA 7.2.2 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de tamaño n de una variable aleatoria X con media μ y varianza σ^2 , entonces la distribución de la media muestral $\bar{X}$ es aproximadamente una distribución normal con media μ y varianza σ^2/n. Y la variable aleatoria

$$Z = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma}$$

tiene aproximadamente una distribución $N(0,1)$.

NOTA El teorema 7.2.2 es válido para cualquier población finita o infinita discreta o contínua, cuando $n \geq 30$. Si la población es normal el teorema se cumple cualquiera que sea el tamaño de n.

Cuando la población es finita de N elementos y el muestreo es sin reemplazamiento, las variables aleatorias X_i no son independientes, entonces la distribución de $\bar{X}$ obedece a una distribución de probabilidad Hipergeométrica. Por lo tanto

$$\mu_{\bar{X}} = \mu \ , \quad y \quad \sigma^2_{\bar{X}} = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) \ . \qquad \text{Corolario 7.2.2}$$

TEOREMA 7.2.3 Si $X_1, X_2, \ldots, X_n$ es una muestra de tamaño n extraída sin reemplazamiento de una población finita de tamaño N con media μ y varianza σ^2 entonces la media muestral $\bar{X}$ tiene aproximadamente una distribución normal con media μ y varianza $\frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$. Y la variable aleatoria

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$$

tiene aproximadamente una distribución normal estándar.

EJEMPLO 3 Sea 1,1,1,3,4,5,6,6,6,7 una población. Se extrae una muestra de tamaño 36 con reemplazamiento de esta población. Calcular:

(a) la media μ y la desviación estándar σ de la población

(b) la media $\mu_{\bar{X}}$ y la desviación estándar $\sigma_{\bar{X}}$ de la media muestral $\bar{X}$

(c) $P[3.6 \leq \bar{X} \leq 4.4]$

SOLUCION Consideremos la población como una variable aleatoria X cuyos valores que podrían tomar son 1,2,3,4,5,6 y 7. Entonces la distribución de probabilidad de la población, es

x	1	3	4	5	6	7
P[X = x]	0.3	0.1	0.1	0.1	0.3	0.1

(a) Cálculo de la media de la población.

$$\mu = E(X) = 1(0.3) + 3(0.1) + 4(0.1) + 5(0.1) + 6(0.3) + 7(0.1)$$

$$= 0.3 + 0.3 + 0.4 + 0.5 + 1.8 + 0.7 = 4.$$

Cálculo de la varianza de la población.

$$E(X^2) = 1(0.3) + 9(0.1) + 16(0.1) + 25(0.1) + 36(0.3) + 49(0.1)$$

$$= 0.3 + 0.9 + 1.6 + 2.5 + 10.8 + 4.9 = 21.$$

$$\sigma^2 = E(X^2) - \mu^2 = 21 - 16 = 5 .$$

Luego, $\sigma = \sqrt{5}$.

(b) Por el corolario 7.2.1 $\mu_{\bar{X}} = 4$ y $\sigma_{\bar{X}} = \frac{\sqrt{5}}{\sqrt{36}} = \frac{\sqrt{5}}{6}$

(c) $n = 36 > 30$, por el teorema 7.2.2, $\bar{X}$ tiene aproximadamente una distribución normal con media $\mu = 4$ y varianza $\frac{\sigma^2}{n} = \frac{5}{36}$. Entonces

$$P[3.6 \leq \bar{X} \leq 4.4] = P\left[\frac{(3.6 - 4)6}{\sqrt{5}} \leq \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \leq \frac{(4.4 - 4)6}{\sqrt{5}}\right]$$

$$= P\left[-\frac{2.4}{2.24} \leq Z \leq \frac{2.4}{2.24}\right] = \Phi(1.07) - \Phi(-1.07)$$

$$= 2\Phi(1.07) - 1 = 2(0.8577) - 1 = 0.7154 .$$

EJEMPLO 4 En su camino al trabajo una persona pasa tres semáforos cada mañana. Los semáforos operan independientemente y debido a que la distancia entre ellos es grande, también operan independientemente respecto a una persona que camina de uno hacia otro. La probabilidad de una luz roja es 0.4,0.8 y 0.5, - respectivamente, para cada uno de los semáforos. Sea **X** el número de luces rojas que la persona encuentra en su camino de ida. Considere que la persona, - durante un año hace 250 viajes a su trabajo. Sea $\bar{X}$ la media del número de luces rojas que encuentra en cada uno de estos viajes. Determinar

(a) la media y la desviación estándar de $\bar{X}$.

(b) $P[\bar{X} \geq 1.5]$.

SOLUCION **X** = número de luces rojas que la persona encuentra en su camino de ida.

$R_X = \{0,1,2,3\}$

Cálculo de la distribución de probabilidad de **X**. Definimos los siguientes - eventos

R_i : "el i-ésimo semáforo está en rojo, $i = 1,2,3$" .

$P[R_1] = 0.4$, $P[R_2] = 0.8$ y $P[R_3] = 0.5$

$P[X = 0] = P[\bar{R}_1\bar{R}_2\bar{R}_3] = (0.6)(0.2)(0.5) = 0.06$. $R_i, i = 1,2,3$ independientes

$P[X = 1] = P[R_1\bar{R}_2\bar{R}_3 \cup \bar{R}_1R_2\bar{R}_3 \cup \bar{R}_1\bar{R}_2R_3] = (0.4)(0.2)(0.5) + (0.6)(0.8)(0.5) + (0.6)(0.2)(0.5)$
$= 0.34$.

$P[X = 2] = P[R_1R_2\bar{R}_3] + P[R_1\bar{R}_2R_3] + P[\bar{R}_1R_2R_3] = (0.4)(0.8)(0.5) + (0.4)(0.2)(0.5) + (0.6)(0.8)(0.5)$
$= 0.44$.

$P[X = 3] = P[R_1R_2R_3] = (0.4)(0.8)(0.5) = 0.16$.

x	0	1	2	3
$p(x)$	0.06	0.34	0.44	0.16

Cálculo de la media y varianza poblacional.

$$\mu = \sum_{x \in R_X} x p(x) = 0(0.06) + 1(0.34) + 2(0.44) + 3(0.16) = 1.7 \,.$$

$$E(X^2) = 0(0.06) + 1(0.34) + 4(0.44) + 9(0.16) = 3.54 \,.$$

$$\sigma^2 = E(X^2) - (E(X))^2 = 3.54 - (1.7)^2 = 0.65.$$

(a) Sea X_i = número de luces rojas que la persona encuentra en el i-ésimo viaje $i = 1,2,\ldots, 250$.

por ejemplo, X_5 = número de luces rojas que la persona encuentre en el quinto viaje, etc.

$$\bar{X} = \frac{1}{250}\sum_{i=1}^{250} X_i$$

por el corolario 7.2.1 $\mu_{\bar{X}} = 1.7$ y $\sigma_{\bar{X}} = \dfrac{\sqrt{0.65}}{\sqrt{250}}$

(b) Por el teorema 7.2.2, $\bar{X}$ tiene aproximadamente una distribución normal, entonces

$$P[\bar{X} \geq 1.5] = P\left[\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \geq \frac{(1.5 - 1.7)\sqrt{250}}{\sqrt{0.65}}\right] = P[Z \geq -3.91] \simeq 1$$

EJEMPLO 5 Sea $\bar{X}_{36}$ la media de una muestra aleatoria de tamaño 36 de la variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} (\frac{3}{2})x^2 & , \quad -1 < x < 1 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Calcular $P[0.03 \leq \bar{X}_{36} \leq 0.15]$

SOLUCION

1. La media y la varianza de la población están dados por

$$\mu = E(X) = \int_{-1}^{1} \frac{3}{2} x^3 dx = \frac{3}{2}\left[\frac{x^4}{4}\right]_{-1}^{1} = 0 \,.$$

y $$\sigma^2 = E(X^2) = \int_{-1}^{1} \frac{3}{2} x^4 dx = \frac{3}{5} .$$

2. $n = 36 > 30$, por el teorema 7.2.2 $\bar{X}_{36}$ tiene aproximadamente una distribución normal con

$$\mu_{\bar{X}} = \mu = 0 , \quad y \quad \sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{3/5}{36} .$$

3. $$P[0.03 \le \bar{X}_{36} \le 0.15]$$
$$= P\left[\frac{(0.03 - 0)\sqrt{36}}{\sqrt{3/5}} \le \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \le \frac{(0.15 - 0)\sqrt{36}}{\sqrt{3/5}}\right]$$
$$= P\left[\frac{0.18}{0.775} \le Z \le \frac{0.90}{0.775}\right] = \Phi(1.16) - \Phi(0.23)$$
$$= 0.8770 - 0.5910 = 0.2860 .$$

EJEMPLO 6 Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \frac{x^3}{4} , & 0 < x < 2 \\ 0 , & \text{en otro lugar} \end{cases}$$

De la cual se toma una muestra aleatoria de tamaño 32. Hallar :

(a) $P[\bar{X} < 1.6]$

(b) $P[(1.5 < \bar{X} < 1.6) | \bar{X} < 1.6]$

SOLUCION

1. La media y la varianza de la población son

$$\mu = E(X) = \int_0^2 \frac{x^4}{4} dx = \frac{x^5}{20}\Big|_0^2 = \frac{8}{5} .$$

$$E(X^2) = \int_0^2 \frac{x^5}{4} dx = \frac{x^2}{24}\Big|_0^2 = \frac{8}{3}$$

$$\sigma^2 = \frac{8}{3} - \frac{64}{25} = \frac{8}{75} .$$

2. $n = 32 > 30$, por el teorema 7.2.2, $\bar{X}$ tiene aproximadamente una distribución normal con

$$\mu_{\bar{X}} = \mu = \frac{8}{5} \quad y \quad \sigma^2 = \frac{\sigma^2}{n} = \frac{8/75}{32} = \frac{1}{300} .$$

(a) $P[\bar{X} < 1.6] = P\left[\dfrac{\bar{X} - \mu}{\sigma\sqrt{n}} < \dfrac{1.6 - 8/5}{\sqrt{1/300}}\right] = P[Z < 0.0] = 0.500$.

(b) $P[(1.5 < \bar{X} < 1.6) | \bar{X} < 1.6]$

$$= \frac{P[(1.5 < \bar{X} < 1.6) \cap (\bar{X} < 1.6)]}{P[\bar{X} < 1.6]} = \frac{P[1.5 < \bar{X} < 1.6]}{P[\bar{X} < 1.6]}$$

$$= \frac{P\left[\dfrac{1.5 - 1.6}{\sqrt{1/300}} < \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 0\right]}{P[Z < 0]}$$

$$= \frac{P[-1.73 < Z < 0]}{P[Z < 0]}$$

$$= \frac{0.5000 - 0.0418}{0.500} = 0.9164 .$$

EJEMPLO 7 El número de horas de duración de una pila para transistores, tiene una distribución normal con μ = 100 horas y σ = 20 horas. Si se seleccionan muestras aleatorias de 16 pilas,

(a) ¿Qué proporción de las medias muestrales estará entre 100 y 125 horas?

(b) ¿Por abajo de qué valor en horas caerá el 95 % de las medias muestrales?

(c) ¿Dentro de qué límites caerá el 99 % de las medias muestrales al rededor de la media de la población?

(d) ¿Es necesario que se cumpla el teorema central del límite para contentar (a), (b) y (c)? Explique.

SOLUCION 1 La población, **X** = número de horas de duración de una pila; tiene una distribución normal con μ = 100 horas y σ = 20 horas.

2. Se extraen muestras aleatorias de 16 pilas, $X_1, X_2, \ldots, X_{16}$ de esta población. Entonces cada $X_i (i = 1,2, \ldots ,16)$ tiene distribución normal con $\mu = 100$ y $\sigma = 20$.

3. $\bar{X} = \dfrac{1}{16}\sum_{i=1}^{16} X_i$, tiene una distribución normal (propiedad reproductiva) con $\mu_{\bar{X}} = 100$ y $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{20}{\sqrt{16}} = 5$.

(a) $$P[100 < \bar{X} < 125] = P\left[\frac{(100-100)\sqrt{16}}{20} < \frac{(X-\mu)\sqrt{n}}{\sigma} < \frac{(125-100)\sqrt{16}}{20}\right]$$

$$= P[0 < Z < 5] = 0.5 .$$

(b) Sea a el valor que debemos determinar, entonces

$$P[\bar{X} < a] = 0.95$$

$$P\left[\frac{(\bar{X}-\mu)\sqrt{n}}{\sigma} < \frac{(a-100)\sqrt{16}}{20}\right] = 0.95$$

$$P\left[Z < \frac{a-100}{5}\right] = 0.95$$

de la tabla III obtenemos $\frac{a-100}{5} = 1.645$

de donde se obtiene $a = 100 + 5(1.645) = 108.225.$

(c) Suponiendo que los límites x_1 y x_2 dentro de los cuales caerá el 99% de medias muestrales alrededor de la media poblacional son simétricos respecto a dicha media. Es decir $x_1 = \mu - x$ y $x_2 = \mu + x$. Entonces ,

$$P[\mu - x < \bar{X} < \mu + x] = 0.99$$

$$P\left[\frac{-x\sqrt{16}}{20} < \frac{(\bar{X}-\mu)\sqrt{n}}{\sigma} < \frac{x\sqrt{16}}{20}\right] = 0.99$$

$$P\left[-\frac{x}{5} < Z < \frac{x}{5}\right] = 0.99$$

$$2\,\Phi\left(\frac{x}{5}\right) - 1 = 0.99$$

$$\Phi\left(\frac{x}{5}\right) = 0.005$$

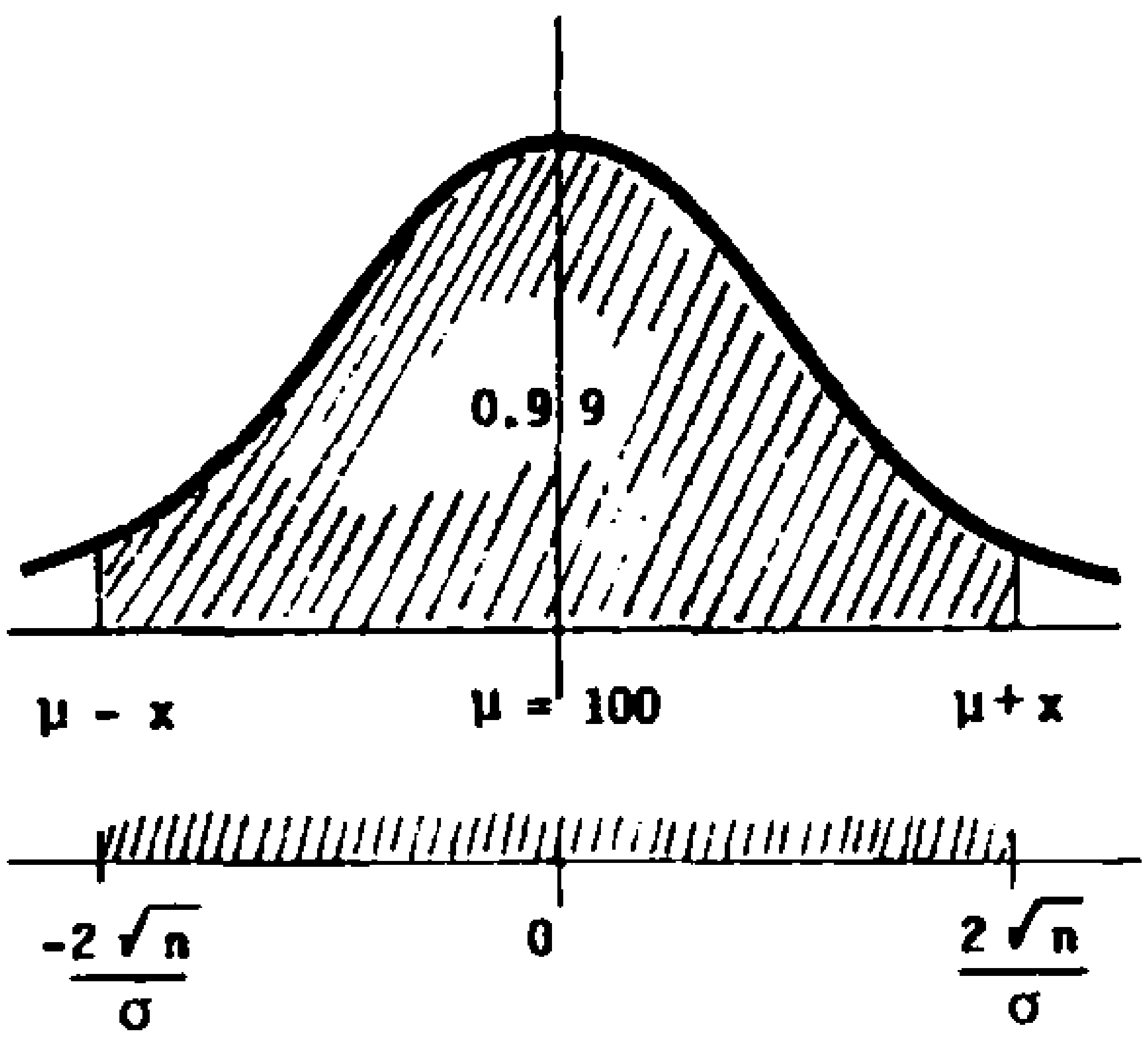
De la tabla III obtenemos $\frac{x}{5} = 2.575$

de donde $x = 12.875$

Por lo tanto, los límites pedidos son :

$$x_1 = 100 - 12.875 = 87.125 .$$

$$x_2 = 100 + 12.875 = 112.875.$$



(d) No es necesario que se cumpla el teorema central del límite, puesto que la población tiene una distribución normal.

EJEMPLO 8 Una cadena de tiendas a nivel nacional, vende una marca muy conocida de calculadora de bolsillo. Para poder lograr el máximo descuento por volumen de compra, todas las tiendas deben hacer un nuevo pedido de calculadoras al mismo tiempo. La decisión para el nuevo pedido, es hacer el pedido cuando el inventario promedio en una muestra de tiendas es menor de 25 calculadoras. Con base en datos anteriores, se supone que la desviación estándar es de 10 - calculadoras. Si se selecciona una muestra de 25 tiendas, ¿cuál es la probabilidad que se vuelva a ordenar el pedido de calculadoras ,

(a) Cuando el inventario promedio real de todas las tiendas es de 20 calculadoras?

(b) cuando el inventario promedio real de todas las tiendas es de 30 calculadoras?.

(c) ¿Qué suposición se debe hacer para contestar (a) y (b).

(d) ¿Cuál sería su respuesta a (a) y (b) si el tamaño de la muestra es 36?

SOLUCION 1 La población, X = número de calculadoras en inventario de una tienda, tiene media μ calculadoras y $\sigma = 10$ calculadoras.

2. Se extrae de la población una muestra de 25 tiendas , $X_1, X_2, \ldots, X_{25}$. Y $\bar{X}$ la media muestral de las calculadoras de las 25 tiendas.

3. Por el corolario 7.2.1 $\mu_{\bar{X}} = \mu$, $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{10}{\sqrt{25}} = 2$.

4. Suponiendo que se cumple el teorema central del límite, se tiene que

$$\bar{X} \sim N(\mu, \frac{100}{25})$$

5. Se hace un nuevo pedido, si $\bar{X} < 25$. Entonces:

(a) cuando $\mu = 20$.

$$P[\bar{X} < 25] = P\left[\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} < \frac{(25 - 20)\sqrt{25}}{10}\right] = P[Z < 2.5] = 0.9938 .$$

(b) cuando $\mu = 30$

$$P[\bar{X} < 25] = P\left[\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} < \frac{(25 - 30)\sqrt{25}}{10}\right] = P[Z < -2.5] = 0.0062.$$

(c) Se debe considerar válido el teorema central del límite (ver paso 4)

(d)1. La población tiene media μ y $\sigma = 10$ calculadoras

2. Se extrae una muestra de 36 tiendas $X_1, X_2, \ldots, X_{36}$ y $\bar{X}$ la media muestral de las calculadoras de las 36 tiendas.

3. $n = 36 > 30$, por el teorema 7.2.2 es

$$\bar{X} \sim N(\mu, \frac{100}{36})$$

4. Se hace un nuevo pedido si $\bar{X} < 25$. Entonces

(i) cuando $\mu = 20$.

$$P[\bar{X} < 25] = P\left[\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} < \frac{(25 - 20)\sqrt{36}}{10}\right] = P[Z < 3] = 0.9987 .$$

(ii) cuando $\mu = 30$.

$$P[\bar{X} < 25] = P\left[\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} < \frac{(25 - 30)\sqrt{36}}{10}\right] = P[Z < -3] = 0.0013 .$$

EJEMPLO 9 Con referencia al ejemplo 8. Suponga que la cadena tuviera un total de 250 tiendas, ¿cuales serían sus respuestas a las preguntas (a) y (b) de ese problema?

SOLUCION 1 El tamaño de la población es $N = 250$ tiendas.

2. La población tiene media μ y desviación estándar $\sigma = 10$

3. Se toma una muestra de 25 tiendas $X_1, X_2, \ldots, X_{25}$ de la población finita. Y $\bar{X}$ la media muestral de las calculadoras de las 25 tiendas .

4. Por el corolario 7.2.2, $\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}} = \frac{10}{\sqrt{25}} \sqrt{\frac{250-25}{150-1}}$

Observe que $n = 25 = 0.1N$, es grande con respecto al tamaño de la población .

5. Suponiendo que se cumple el teorema central del límite, se tiene

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n} \frac{N - n}{N - 1}\right)$$

(a) Si $\mu = 20$.

$$P[\bar{X} < 25] = P\left[\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N - n}{N - 1}}} < \frac{25 - 20}{\frac{10}{\sqrt{25}}\sqrt{\frac{250 - 25}{250 - 1}}}\right] = P[Z < 0.63]$$

$$= 0.9957$$

(b) Si $\mu = 30$.

$$P[\bar{X} < 25] = P\left[\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N - n}{N - 1}}} < \frac{25 - 30}{\frac{10}{\sqrt{25}}\sqrt{\frac{250 - 25}{250 - 1}}}\right] = P[Z < -0.63]$$

$$= 0.0043 .$$

EJEMPLO 10 Un investigador quiere estimar la media de una población utilizando una muestra suficientemente grande, para que la probabilidad que la media muestral no difiera de la media poblacional en más de 25 % de la desviación estándar, sea 0.95. ¿Qué tamaño deberá adoptar para la muestra?

SOLUCION Sea n el tamaño de la muestra y $\bar{X}$ la media de esta muestra; si μ y σ^2 son la media y varianza, respectivamente de la población, entonces $\bar{X}$ tiene aproximadamente una distribución normal con $\mu_{\bar{X}} = \mu$ y $\sigma^2_{\bar{X}} = \sigma^2/n$

Del enunciado del problema se tiene;

$$P[|\bar{X} - \mu| \leq 0.25\sigma] = 0.95$$

$$P[- 0.25\sigma \leq \bar{X} - \mu \leq 0.25\sigma] = 0.95$$

$$P\left[\frac{-0.25\sigma}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{0.25\sigma}{\sigma/\sqrt{n}}\right] = 0.95$$

$$P[- 0.25\sqrt{n} \leq Z \leq 0.25\sqrt{n}] = 0.95$$

$$\Phi(0.25\sqrt{n}) - \Phi(- 0.25\sqrt{n}) = 0.95$$

$$\Phi(0.25\sqrt{n}) - [1 - \Phi(0.25\sqrt{n})] = 0.95$$

$$2\,\Phi(0.25\sqrt{n}) = 1.95$$

$$\Phi(0.25\sqrt{n}) = 0.975$$

De la tabla III obtenemos, $\quad 0.25\sqrt{n} = 1.96$

de donde $\quad n = \left(\frac{196}{25}\right)^2 = 61.4656$

por lo tanto $\quad n = 62$

EJEMPLO 11 Los calificativos de un examen a nivel nacional se distribuyen con media $\mu = 11.80$ y $\sigma = 3.6$ puntos, ¿cuántos estudiantes se deben tomar en una muestra para garantizar, con un 95 % de probabilidad, que la media muestral sea mayor que 11.08.

SOLUCION 1 La población se distribuye con media $\mu = 11.80$ y varianza $\sigma^2 = (3.6)^2$.

2. Sea $\bar{X}$ la media muestral de una muestra de tamaño n .

3. $\bar{X} \sim N(\mu, \sigma^2/n)$ por el teorema 7.2.2

4. Según el enunciado se tiene

$$P[\bar{X} > 11.08\ ] \quad = \quad 0.95$$

$$P[\frac{(\bar{X} - n)\sqrt{n}}{\sigma} > \frac{(11.08 - 11.80)\sqrt{n}}{3.6}\ ] = \quad 0.95$$

$$P[Z > -\frac{0.72}{3.6}\sqrt{n}\ ] \quad = \quad 0.95$$

$$P[Z > -\ 0.2\sqrt{n}\ ] \quad = \quad 0.95$$

$$P[Z < 0.2\sqrt{n}\ ] \quad = \quad 0.95$$

De la tabla III, $\quad 0.2\sqrt{n} \ = \ 1.645$

$$\sqrt{n} = \frac{16.45}{2} = 8.225$$

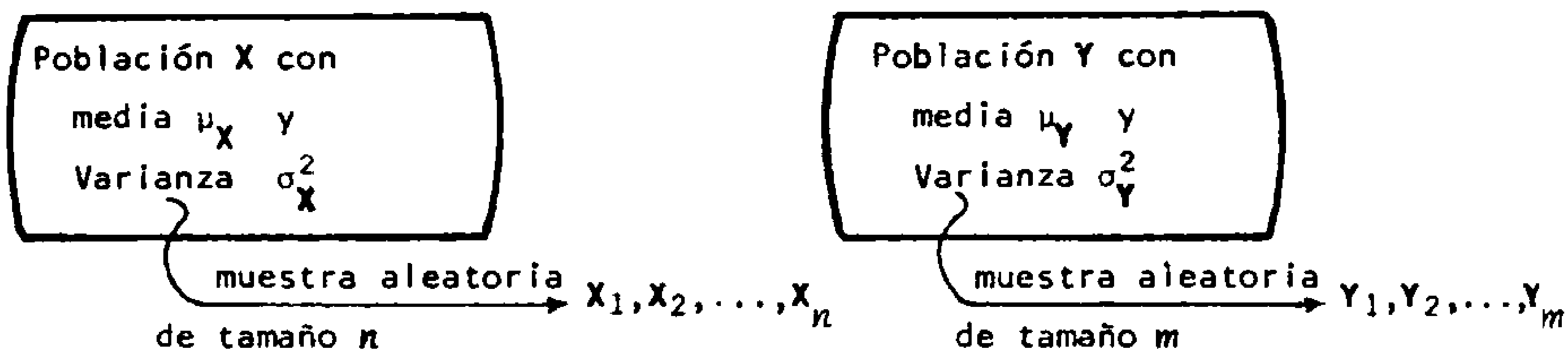
$$n = 67.65$$

luego, $\quad n = 68$.

7.2.3 DISTRIBUCION MUESTRAL DE LA DIFERENCIA DE DOS MEDIAS, MUESTRAS INDEPENDIENTES

En muchos problemas se está interesado en comparar los parámetros, (en particular las medias) de dos poblaciones (o dos variables aleatorias). La comparación puede hacerse sobre la base de dos muestras aleatorias *independientes*. Supongamos ahora que tenemos dos poblaciones **X** e **Y**, la primera con media μ_X y varianza σ_X^2 , y la segunda con media μ_Y y varianza σ_Y^2 . Sea $\bar{X}$ la media de la muestra aleatoria de tamaño n , extraída de la primera población, e $\bar{Y}$ la media de la muestra aleatoria de tamaño m , tomada de la segunda población. La distribución de la diferencia de dos medias muestrales $\bar{X} - \bar{Y}$, se llama, ***distribución muestral de la diferencia de dos medias.***

El siguiente diagrama da una visión esquemática de lo expresado en el parrafo anterior



$\bar{X} = \frac{1}{n}\sum_{i=1}^{n}\bar{X}_i$ la media muestral $\bar{Y} = \frac{1}{m}\sum_{i=1}^{m}Y_i$

1. Por el corolario 7.2.1 $\mu_{\bar{X}} = \mu_X$; $\sigma^2_{\bar{X}} = \dfrac{\sigma^2_X}{n}$; y $\mu_{\bar{Y}} = \mu_Y$, $\sigma^2_{\bar{Y}} = \dfrac{\sigma^2_Y}{m}$

2. $\bar{X} - \bar{Y}$ es una variable aleatoria con media

$$\mu_{\bar{X}-\bar{Y}} = E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_{\bar{X}} - \mu_{\bar{Y}} = \mu_X - \mu_Y .$$

y varianza

$$\mu^2_{\bar{X}-\bar{Y}} = \mathrm{Var}(\bar{X} - \bar{Y}) = \mathrm{Var}(\bar{X}) + \mathrm{Var}(\bar{Y}) = \sigma^2_{\bar{X}} + \sigma^2_{\bar{Y}}$$
$$= \frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m} .$$

de donde $\sigma_{\bar{X}-\bar{Y}} = \sqrt{\dfrac{\sigma^2_X}{n} + \dfrac{\sigma^2_Y}{m}}$

3. Para n y m suficientemente grande por el teorema 7.2.2, $\bar{X}$ se distribuye aproximadamente por una normal con media μ_X y varianza σ^2_X/n e $\bar{Y}$ se distribuye aproximadamente por una normal con media μ_Y y varianza σ^2_Y/m. En símbolos

$$\bar{X} \frown\!\!\rightarrow N(\mu_X,\ \sigma^2_X/n) \quad , \qquad \bar{Y} \frown\!\!\rightarrow N(\mu_Y,\ \sigma^2_Y/m) .$$

4. Por la propiedad reproductiva de la normal, la distribución muestral de la diferencia de medias $\bar{X} - \bar{Y}$ es aproximadamente una distribución normal con media $\mu_X - \mu_Y$ y varianza $\dfrac{\sigma^2_X}{n} + \dfrac{\sigma^2_Y}{m}$. En símbolos

$$\bar{X} - \bar{Y} \frown\!\!\rightarrow N\left(\mu_X - \mu_Y,\ \frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m}\right)$$

y la variable aleatoria

$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\dfrac{\sigma^2_X}{n} + \dfrac{\sigma^2_Y}{m}}}$$

tiene aproximadamente una distribución $N(0,1)$.

Observe que los resultados obtenidos para la distribución de $\bar{X} - \bar{Y}$ son válidos:

(a) cuando el muestreo es con reemplazamiento de dos poblaciones finitas.

(b) cuando el muestreo es con reemplazamiento o sin reemplazamiento de dos poblaciones infinitas, discretas o contínuas.

(c) cuando el muestreo es sin reemplazamiento de dos poblaciones finitas, cuyos tamaños N_1 y N_2 son grandes con respecto a los tamaños n y m de la muestra, respectivamente.

En el caso que la población es pequeña y el muestreo es sin reemplazamiento, entonces se debe calcular $\sigma^2_{\bar{X}}$ y $\sigma^2_{\bar{Y}}$ usando el teorema 7.2.3

El teorema siguiente formaliza la distribución muestral de la diferencia de medias.

TEOREMA 7.2.4 Si $\bar{X}$ y $\bar{Y}$ son las medias de dos muestras aleatorias de dos poblaciones X e Y, con media μ_X y μ_Y, varianza σ^2_X y σ^2_Y, respectivamente, entonces la distribución muestral de la diferencia de medias $\bar{X} - \bar{Y}$, es aproximadamente una distribución normal con media $\mu_{\bar{X}-\bar{Y}} = \mu_X - \mu_Y$

y varianza $\sigma^2_{\bar{X}-\bar{Y}} = \dfrac{\sigma^2_X}{n} + \dfrac{\sigma^2_Y}{m}$ y la variable aleatoria

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\dfrac{\sigma^2_X}{n} + \dfrac{\sigma^2_Y}{m}}}$$

tiene aproximadamente una distribución N(0,1)

Si n y m son mayores o iguales que 30, la aproximación normal para la distribución de $\bar{X} - \bar{Y}$ es optíma.

Si las poblaciones son normales, el teorema se cumple cualquiera que sea los tamaños de las muestras.

EJEMPLO 12 Cierta universidad en formación tiene 100 profesores, 60 de los cuales tienen el doctorado. Dos muestras, con $n_1 = n_2 = 30$, son extraídas independiente de este profesorado, con reposición, y se anotan los números de los que tienen el doctorado, ¿cuál es la probabilidad que las dos muestras difieran en 8 ó más en el número con doctorado?

SOLUCION Sean las siguientes variables aleatorias:

X_1 = número de doctores en la primera muestra de tamaño 30.

X_2 = número de doctores en la segunda muestra de tamaño 30.

$R_{X_1} = \{0,1,2,\ldots, 30\}$

$R_{X_2} = \{0,1,2,\ldots, 30\}$

Entonces, X_1 y X_2 son variables aleatorias con distribuciones binomiales y parámetros,

$$p_1 = p_2 = \frac{60}{100} = \frac{3}{5}, \quad n = 30 \quad y \quad q_1 = q_2 = \frac{40}{100} = \frac{2}{5}$$

luego, $\mu_{X_1} = \mu_{X_2} = np = 30(\frac{3}{5}) = 18$

$\sigma^2_{X_1} = \sigma^2_{X_2} = npq = 18(\frac{2}{5}) = \frac{36}{5}$

por lo tanto, $\mu_{X_1-X_2} = \mu_{X_1} - \mu_{X_2} = 0$

y $\sigma^2_{X_1-X_2} = \sigma^2_{X_1} + \sigma^2_{X_2} = 2(\frac{36}{5}) = \frac{72}{5}$

de donde $\sigma_{X_1-X_2} = \sqrt{\sigma^2_{X_1} + \sigma^2_{X_2}} = \sqrt{\frac{72}{5}} = 3.795$

Se pide calcular:

$$P[|X_1 - X_2| \geq 8] = P[X_1 - X_2 \geq 8] + P[X_1 - X_2 \leq -8]$$

$$= P\left[\frac{(X_1-X_2)-(\mu_{X_1}-\mu_{X_2})}{\sqrt{\sigma^2_{X_1} + \sigma^2_{X_2}}} \geq \frac{8-0}{3.795}\right] + P\left[\frac{(X_1-X_2)-(\mu_{X_1}-\mu_{X_2})}{\sqrt{\sigma^2_{X_1} + \sigma^2_{X_2}}} \leq \frac{-8-0}{3.795}\right]$$

$$= P[Z \geq 2.11] + P[Z \leq -2.11]$$

$$= P[Z \geq 2.11] + P[Z \geq 2.11]$$

$$= 2P[Z \geq 2.11] = 2[1 - P[Z \leq 2.11]] = 2[1 - 0.9826]$$

$$= 0.0348 .$$

NOTA En la aproximación de la diferencia de las dos variables binomiales por la normal hemos ignorado el factor de corrección de continuidad, aunque realmente debería usarse.

EJEMPLO 13 De dos máquinas que embolsan automáticamente café, se han extraído muestras de 64 bolsitas de cada uno de ellas.
La distribución de probabilidad del peso de cada bolsita para ambas poblaciones tienen idénticas medias y desviaciones típicas de 6.40 grs. y 7.20 grs. - respectivamente. Determinar la probabilidad que la diferencia entre las medias de las muestras excede a 0.60 en valor absoluto.

SOLUCION 1 Sea **X** e **Y** definidos por

X, población de bolsitas de café embolsados por la primera máquina, con media $\mu_X = \mu$ y desviación típica $\sigma = 6.40$.	Y, población de bolsitas de café embolsados por la segunda máquina, con media $\mu_Y = \mu$ y desviación típica $\sigma = 7.20$
muestra de tamaño $n = 64$ → $X_1, X_2, \ldots, X_{64}$	muestra de tamaño $m = 64$ → $Y_1, Y_2, \ldots, Y_{64}$

$\bar{X} = \frac{1}{64} \sum X_i$ media muestral $\qquad\qquad$ $\bar{Y} = \frac{1}{64} \sum Y_i$ media muestral

2. $\mu_{\bar{X}-\bar{Y}} = \mu_X - \mu_Y = 0$, $\sigma^2_{\bar{X}-\bar{Y}} = \frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m} = \frac{(6.4)^2}{64} + \frac{(7.2)^2}{64}$;

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{(6.4)^2}{64} + \frac{(7.2)^2}{64}} = 1.204$$

3. Por el teorema 7.2.4, la distribución de la diferencial de medias muestrales es aproximadamente normal con media $\mu_X - \mu_Y = 0$ y varianza $\frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m}$. Es decir

$$(\bar{X} - \bar{Y}) \sim N(0, \frac{(6.4)^2}{64} + \frac{(7.2)^2}{64}) .$$

4. Entonces, del enunciado se tiene

$$P[|\bar{X} - \bar{Y}| > 0.6] = P[\bar{X} - \bar{Y} > 0.6] + P[\bar{X} - \bar{Y} < -0.6]$$

$$= P\left[\frac{(\bar{X}-\bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m}}} > \frac{0.6-0}{1.204}\right] + P\left[\frac{(\bar{X}-\bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m}}} < -\frac{0.6-0}{1.204}\right]$$

$$= P[Z > 0.5] + P[Z < -0.5] = 2P[Z > 0.5]$$

$$= 2[1 - P[Z < 0.5]] = 2(1 - 0.6915) = 0.6170 .$$

EJEMPLO 14 Sean $\bar{X}_1$ y $\bar{X}_2$ las medias muestrales de dos muestras aleatorias independientes de tamaño n . Las observaciones que comforman la muestra estan distribuídas normalmente con media común μ y varianza común iqual a 2. Determinar n de tal forma que la probabilidad que $\bar{X}_1$ y $\bar{X}_2$ difieran en menos de 2 - sea 0.98 .

SOLUCION Las observaciones $X_1, X_2, \ldots, X_n$ que conforman las medias se distribuyan normalmente con media común μ y varianza común 2. Es decir

$$X_i \to N(\mu, 2), \quad i = 1,2,\ldots, n. \quad \text{Entonces}$$

$$\bar{X}_1 \to N(\mu, \frac{2}{n}) \quad \text{y} \quad \bar{X}_2 \to N(\mu, \frac{2}{n})$$

por lo tanto, $\mu_{\bar{X}_1 - \bar{X}_2} = \mu - \mu = 0.$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{n}} = \sqrt{2(\frac{2}{n})} = \frac{2}{\sqrt{n}}$$

Se pide calcular n tal que

$$0.98 = P[|\bar{X}_1 - \bar{X}_2| < 2] = P[-2 < \bar{X}_1 - \bar{X}_2 < 2]$$

$$= P\left[\frac{-2-0}{\frac{2}{\sqrt{n}}} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu - \mu)}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{n}}} < \frac{2-0}{\frac{2}{\sqrt{n}}}\right]$$

$$= P[-\sqrt{n} < Z < \sqrt{n}] = \Phi(\sqrt{n}) - \Phi(-\sqrt{n})$$

$$= \Phi(\sqrt{n}) - [1 - \Phi(\sqrt{n})] = 2\Phi(\sqrt{n}) - 1$$

ó $\Phi(\sqrt{n}) = 0.99$, y de la tabla III se obtiene $\sqrt{n} = 2.33$

ó $n = 5.4289$

Luego, $n = 6$.

7.2.4 DISTRIBUCION DE UNA PROPORCION

Hemos visto que una variable aleatoria binomial está definida por, X = número de éxitos ocurridos en n ensayos de Bernoulli. Luego,

$$R_X = \{0,1,2,\ldots,n\} \quad , \quad p = P[E]$$

y se dice que X tiene una distribución binomial, con parámetros n y p . Entonces, la ***proporción de éxitos***, $\frac{X}{n}$, es una variable aleatoria que se denota por

$$\bar{P} = \frac{X}{n}$$

los valores que toma la variable aleatoria, $\frac{X}{n}$ son números comprendidos entre 0 y 1. Es decir, el rango de esta variable aleatoria es

$$R_{\bar{P}} = \{0\ ,\ \frac{1}{n}\ ,\ \frac{2}{n}\ ,\ \frac{3}{n}\ ,\ .\ .\ .\ ,\ 1\}$$

la *media y la varianza* de la proporción de éxitos son:

$$\mu_{\bar{P}} = E(\frac{X}{n}) = \frac{1}{n} E(X) = \frac{1}{n}(np) = p\ .$$

$$\sigma^2_{\bar{P}} = Var(\frac{X}{n}) = \frac{1}{n^2} Var(X) = \frac{1}{n^2}(npq) = \frac{pq}{n}\ .$$

de donde $\sigma_{\bar{P}} = \sqrt{\frac{pq}{n}}\ .$

Para evaluar probabilidad, por ejemplo del tipo $P[\bar{P} < p_o]$, donde p_o es un número entre 0 y 1, observe lo siguiente

$$P[\bar{P} \leq p_o] = P\left[\frac{X}{n} \leq p_o\right] = P[X \leq np_o]$$

Desde que np_o posiblemente no siempre sea un entero, se tiene que

$$P[\bar{P} \leq p_o] = P[X \leq np_o] = \sum_{x=0}^{[\![np_o]\!]} \binom{n}{x} p^x q^{n-x}$$

(donde $[\![\ ]\!]$ es la función máximo entero). Es decir, la distribución de probabilidad de la proporción de éxitos obedece a una distribución de probabilidad binomial, y se escribe

$$p(\frac{x}{n}) = P\left[\bar{P} = \frac{x}{n}\right] = P\left[\frac{X}{n} = \frac{x}{n}\right] = P[X = x] = \binom{n}{x} p^x q^{n-x}\ , \quad x = 0,1,\ldots,n$$

EJEMPLO 15 Una compañía tiene un número grande de empleados. La probabilidad de que un empleado seleccionado aleatoriamente participe en un programa de inversión de acciones en la compañía es 0.40. Si se escoge aleatoriamente 10 empleados. ¿Cuál es la probabilidad que la proporción de participantes sea exactamente 0.60? ¿cuál es la probabilidad que la proporción de participantes sea por lo menos 0.80?

SOLUCION $n = 10$, $p = 0.40$

(a) $P[\bar{P} = 0.60] = P\left[\bar{P} = \frac{6}{10}\right] = P\left[\frac{X}{n} = \frac{6}{10}\right] = P[X = 6]$

$$= \binom{10}{6}(0.40)^6(0.60)^4 = \frac{10!}{6!4!}(0.004096)(0.1296)$$

$$\approx 0.1115\ .$$

(b) $P[\bar{P} \geq 0.8] = P[X \geq 8] = 0.0123$ (tabla I)

La importancia de la variable aleatoria proporción de éxitos es principal mente por sus aplicaciones muestrales. Suponga que se tiene una población binomial (*cualquier colección de objetos, donde cada una puede ser clasificado - como un "éxito" o un "fracaso"*) con parámetro p de la cual se extrae una muestra aleatoria de n observaciones, evidentemente cada observación se clasifica como *éxito o fracaso* y sea X el número de éxitos en la muestra. Observe el - diagrama siguiente

población binomial con parametro p p = proporción de éxitos de la población	$\xrightarrow[\text{de tamaño } n]{\text{muestra aleatoria}}$	$X_1, X_2, \ldots, X_n$

1. Note que cada X_i es una variable aleatoria de Bernoulli, entonces

$$X = \sum_{i=1}^{n} X_i \quad \text{es una variable aleatoria binomial}$$

2. La proporción de éxitos $\bar{P} = \frac{X}{n}$ es una variable aleatoria con media

$$\mu_{\bar{P}} = p \quad \text{y varianza} \quad \sigma^2_{\bar{P}} = \frac{pq}{n}$$

3. Para n suficientemente grande, por el teorema central del límite, la variable aleatoria $\bar{P} = \frac{X}{n}$ se distribuye aproximadamente por una normal con media p y varianza $\frac{pq}{n}$. En simbolos

$$\bar{P} = \frac{X}{n} \sim N\left(p, \frac{pq}{n}\right)$$

y,
$$Z = \frac{\bar{P} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{X}{n} - p}{\sqrt{\frac{pq}{n}}}$$

tiene aproximadamente una distribución $N(0,1)$. Por lo tanto

$$P[\bar{P} \le p_o] = P\left[Z \le \frac{p_o - p}{\sigma_{\bar{P}}}\right] = \Phi\left[\frac{p_o - p}{\sigma_{\bar{P}}}\right]$$

NOTA La exposición anterior se cumple :

1. Para una población infinita, cualquiera que sea el tipo de muestreo
2. Para población finita, cuando el muestreo es con reemplazamiento .

Si el muestreo se hace sin reposición, de una población binomial finita, la distribución por muestreo de $\bar{P}$ obedece a la distribución de probabilidad - hipergeométrica. Es decir

$$P\left[\bar{P} = \frac{x}{n}\right] = p\left(\frac{x}{n}\right) = \binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}$$

Entonces, la desviación estándar debe ser ajustado por el *factor de corrección de población finita*, y en este caso está dado por

$$\sigma_{\bar{P}} = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$$

Si, n es grande, por el teorema central del límite, la variable aleatoria

$$Z = \frac{\frac{X}{n} - p}{\sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}}$$

tiene aproximadamente una distribución $N(0,1)$.

Cuando n es muy pequeña, puede obtenerse aproximaciones normales introduciendo el *factor de corrección de continuidad* de $1/2n$ (se emplea $1/2n$ en lugar de $1/2$, por que en este caso la proporción de éxitos es el número de éxitos - dividido por n). Entonces

$$P[\bar{P} < p_o] = \Phi\left[\frac{p_o + \frac{1}{2n} - p}{\sigma_{\bar{P}}}\right]$$

El lector puede escribir las demás fórmulas.

NOTA En una población binomial finita de N elementos, la proporción de éxitos de la población es $p = \frac{M}{N}$, donde M es el número de éxitos en la población

EJEMPLO 16 Una firma de pedidos por correo, sabe por experiencias anteriores de las circulares que envía por correo, el 10 % tendrán respuestas. Suponga - que se envían 20 circulares como prueba de mercado en una nueva región geográfica. Suponiendo que se puede aplicar la tasa de respuestas del 10 % en la nueva región. Calcular la probabilidad que menos del 20 % de la gente responde. Calcular la probabilidad que contesten entre el 20 % y el 30 % de la - gente.

SOLUCION X = número de circulares respondidas de las 20 enviadas.

$R_X = \{0,1,2,\ldots,20\}$, $n = 20$, $p = 0.10$

debemos calcular

(a) $$P\left[\frac{X}{n} < 0.20\right] = P\left[\frac{\frac{X}{n} - \frac{1}{2n} - p}{\sqrt{\frac{pq}{n}}} \le \frac{0.20 - \frac{1}{40} - 0.10}{\sqrt{\frac{(0.1)(0.9)}{20}}}\right]$$

$$= P\left[Z \le \frac{\frac{3/40}{0.3}}{4.472}\right] = \Phi(\frac{13.416}{12}) = \Phi(1.12)$$

$$\simeq 0.8686.$$

(b) $$P\left[0.20 < \frac{X}{n} < 0.30\right] = P\left[\frac{X}{n} < 0.30\right] - P\left[\frac{X}{n} \le 0.20\right]$$

$$= P\left[\frac{X/n - 1/2n - p}{\sqrt{\frac{pq}{n}}} \le \frac{0.30 - \frac{1}{40} - 0.1}{\sqrt{\frac{(0.1)(0.9)}{20}}}\right]$$

$$- P\left[\frac{X/n + \frac{1}{2n} - p}{\sqrt{\frac{pq}{n}}} \le \frac{0.2 + \frac{1}{40} - 0.10}{\sqrt{\frac{(0.1)(0.9)}{20}}}\right]$$

$$= P[Z \le 2.61] - P[Z \le 1.86]$$

$$= \Phi(2.61) - \Phi(1.86) = 0.9955 - 0.9686 = 0.0269.$$

EJEMPLO 17 Historicamente, el 10% de un embarque grande de piezas para maquinaria es defectuosa. Suponga que el embarque consta de 5000 piezas de máquina. Si se seleccionan muestras sin reemplazo de 400 piezas, ¿qué proporción de las muestras tendrá

(a) entre 9% y 10% de piezas defectuosas?

(b) menos de 8% de piezas defectuosas?

(c) ¿por arriba de qué valor en porcentaje caerá el 95% de los porcentajes de la muestra?

SOLUCION 1 La población binomial es finita con $N = 5000$ y parámetro $p = 0.10$

2. De la población finita se extrae una muestra de 400 piezas. Sea X = número de piezas defectuosas en la muestra .

3. $\bar{P} = \frac{X}{n}$ es la variable aleatoria proporción de defectuosos en la muestra .

4. Puesto que $n = 400$ es grande, la variable aleatoria $\bar{P}$ tiene aproximadamente una distribución normal con $\mu_{\bar{P}} = p = 0.10$ y $\sigma_{\bar{P}} = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{(0.1)(0.9)}{400}}\sqrt{\frac{5000 - 400}{5000 - 1}} = 0.0144$

(a) $$P[0.09 < \bar{P} < 0.1] = P\left[\frac{0.09 - 0.1}{0.0144} < \frac{\bar{P} - p}{\sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}} < \frac{0.1 - 0.1}{0.0144}\right]$$

$$= P[-0.69 < Z < 0] = \Phi(0.0) - \Phi(-0.69)$$

$$= 0.5 - 0.2451 = 0.2549 .$$

(b) $$P[\bar{P} < 0.08] = P\left[\frac{\bar{P} - p}{\sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}} < \frac{0.08 - 0.1}{0.0144}\right] = P[Z < -1.39]$$

$$= 0.0823$$

(c) Sea a el valor que debemos encontrar, entonces

$$0.95 = P[\bar{P} > a] = P\left[\frac{\bar{P} - p}{\sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}} > \frac{a - 0.1}{0.0144}\right]$$

$$= P\left[Z > \frac{a - 0.1}{0.0144}\right] = 1 - P\left[Z \leq \frac{a - 0.1}{0.0144}\right]$$

ó $$P\left[Z \leq \frac{a - 0.1}{0.0144}\right] = 0.05$$

de la tabla III, $\frac{a - 0.1}{0.0144} = -1.645$ de donde $a \approx 0.076$.

Es decir, arriba del 7.6% .

7.2.5 DISTRIBUCION DE LA DIFERENCIA DE DOS PROPORCIONES

Supongamos que se extraen dos muestras independientes de tamaño n_1 y n_2 de dos poblaciones binomiales. Sean p_1 y p_2 las probabilidades de éxito respectivamente. Y sean: X_1 = número de éxitos en la muestra de tamaño n_1; X_2 = número de éxitos en la muestra de tamaño n_2. Hemos visto que las variables aleatorias independientes $\bar{P}_1 = \frac{X_1}{n_1}$ y $\bar{P}_2 = \frac{X_2}{n_2}$ tienen respectivamente medias $\mu_{\bar{P}_1} = p_1$, $\mu_{\bar{P}_2} = p_2$, varianzas $\sigma^2_{\bar{P}_1} = p_1(1 - p_1)/n_1$ y $\sigma^2_{\bar{P}_2} = p_2(1 - p_2)/n_2$. Entonces la diferencia de proporciones de éxitos $\frac{X_1}{n_1} - \frac{X_2}{n_2}$ tiene media $\mu_{\frac{X_1}{n_1} - \frac{X_2}{n_2}} = p_1 - p_2$ y $\sigma^2_{\frac{X_1}{n_1} - \frac{X_2}{n_2}} = \frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$ ...(1)

Por el teorema central del límite, para n_1 y n_2 suficientemente grande, las variables aleatorias $\frac{X_1}{n_1}$ y $\frac{X_2}{n_2}$ tienen aproximadamente una distribución normal, entonces, la variable aleatoria

$$\frac{X_1}{n_1} - \frac{X_2}{n_2}$$

tiene aproximadamente una distribución normal con media y varianza dada en (1) Es decir

$$Z = \frac{(\frac{X_1}{n_1} - \frac{X_2}{n_2}) - (p_1 - p_2)}{\sqrt{p_1(1 - p_1)/n_1 + p_2(1 - p_2)/n_2}}$$

tiene aproximadamente una distribución $N(0,1)$

EJEMPLO 18 Dos compañías A y B producen pilas. La compañía A cree que el 10% de su producción son defectuosas y B, el 5%. Se toma una muestra al azar de 300 unidades de la línea de producción de la compañía A y se encuentra que 24 son defectuosas. Se toma una muestra al azar de 400 unidades de la línea de producción de la compañía B y se encuentra que 20 son defectuosas. ¿Cuál es la probabilidad de obtener esta diferencia o una menor en proporción si la creencia acerca de los parámetros de la población es correcta?

SOLUCION Si la creencia acerca de los parámetros es correcta, la distribución de $\bar{P}_1 - \bar{P}_2$ por muestreo estaría definida por

$$\mu_{\bar{P}_1 - \bar{P}_2} = p_1 - p_2 = 0.1 - 0.05 = 0.05$$

y

$$\sigma_{\bar{P}_1 - \bar{P}_2} = \sqrt{\frac{(0.1)(1 - 0.1)}{300} + \frac{(0.05)(1-0.05)}{350}} = 0.0205 .$$

Es decir $\bar{P}_1 - \bar{P}_2$ se distribuye aproximada por una normal con media $\mu_{\bar{P}_1 - \bar{P}_2} = 0.05$ y $\sigma_{\bar{P}_1 - \bar{P}_2} = 0.0205$

Se tiene que $\bar{P}_1 - \bar{P}_2 = \frac{24}{300} - \frac{20}{400} = 0.08 - 0.05 = 0.03$ por lo tanto, la probabilidad pedida es

$$P[\bar{P}_1 - \bar{P}_2 \leq 0.03] = P\left[\frac{(\bar{P}_1 - \bar{P}_2) - \mu_{\bar{P}_1 - \bar{P}_2}}{\sigma_{\bar{P}_1 - \bar{P}_2}} \leq \frac{0.03 - 0.05}{0.0205}\right]$$

$$= P\left[Z \leq - \frac{0.02}{0.0205}\right] = P[Z \leq - 0.98]$$

$$= 0.1635 .$$

PROBLEMAS 7.2

1. Una población consiste de las edades de los niños de una familia de cuatro niños. Estas edades son : 2,4,6 y 8 años.
   (a) Determinar la media μ y la desviación estándar σ de la población.
   (b) Enumerar todas las muestras posibles (sin reemplazo) de 2 niños que pueden seleccionarse en esta familia y determine $\bar{x}$ para cada muestra.
   (c) Calcular la media $\mu_{\bar{X}}$ y la desviación estándar $\sigma_{\bar{X}}$ de las medias muestrales. Y verifique que se cumple $\mu_{\bar{X}} = \mu$, $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} \sqrt{\dfrac{N-n}{N-1}}$

2. Evalúe la distribución de la población X por muestreo con $n = 2$ de una población { 0,1,2,3,4,5,6,7,8,9 } . Suponga que se hace el muestreo (a) con reposición (b) sin reposición.
   NOTA La evaluación de una distribución de muestreo supone enumerar todos los valores posibles de la estadística con probabilidad asociadas, y calcular su valor esperado y su error estándar.

3. Sea X una población constituída por {2,4,6} . Calcular
   (a) la media μ y la desviación estándar σ de la población. Se extrae una muestra de tamaño 54 con reemplazamiento de la población. Calcular
   (b) la media $\mu_{\bar{X}}$ y la desviación estándar $\sigma_{\bar{X}}$ de las $\bar{X}$.
   (c) $P[4.1 < \bar{X} < 4.4]$.

4. Considere una población X que consiste en 8 billetes de \$ 5 cada uno y 2 billetes de \$ 10 cada uno. Determine E(X) y Var(X) .
   Ponga los diez billetes en una urna y seleccione al azar dos billetes con reposición. Sea $\bar{X}$ la media muestral.
   (a) halle la distribución de probabilidad de $\bar{X}$.
   (b) Calcule $E(\bar{X})$ y $Var(\bar{X})$ ¿Qué se verifica?
   Repita el experimento anterior con la excepción de que se seleccionan los dos billetes al azar sin reposición.
   (c) Halle la distribución de probabilidad de $\bar{X}$.
   (d) Calcule $E(\bar{X})$ y $Var(\bar{X})$ ¿Qué se verifica?

5. De una población con media 25 y varianza 10 se extrae una muestra aleatoria de 25 observaciones.
   (a) ¿Cuál es la probabilidad que la media muestral se encuentra entre 24 y 27 .
   (b) ¿Qué supone para responder (a)?

6. Sea $X_1, X_2, \ldots, X_{36}$ una muestra aleatoria de tamaño 36 de una población con distribución geométrica cuya función de probabilidad es

$$f(x) = \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right), \qquad x = 0, 1, 2, \ldots$$

Calcular: (a) $P[1/4 < \bar{X}_{36} < 1/2]$.

(b) $P[10 < \sum_{i=1}^{36} X_i < 13]$.

7. Sea $\bar{X}_{12}$ la media de una muestra aleatoria de 12 observaciones, de una variable aleatoria con función de distribución uniforme en el intervalo $[0,1]$. Calcular (aproximadamente) $P[1/2 \leq \bar{X}_{12} \leq 2/3]$

8. En determinada ciudad grande 1/3 de las familias no tienen automóvil, 1/3 tiene uno, 1/6 tiene dos, 1/12 tiene tres y 1/12 tiene cuatro automóvil. Cada automóvil tiene cinco llantas. Sea X la variable aleatoria que representa el número de llantas por familia. Se toma una muestra aleatoria de 100 familias. Determinar
   (a) la media $\mu_{\bar{X}}$ y la desviación estándar $\sigma_{\bar{X}}$ de la media muestral.
   (b) $P[\bar{X} < 5]$.

9. Una máquina vendedora de refrescos está regulada de modo que la cantidad despachada tenga una distribución normal con $\mu = 7$ onzas y $\sigma = 0.5$ onzas Si se toman muestras de nueve vasos
   (a) ¿de qué valor exederá el 95% de las medias de la muestra?
   (b) ¿Es necesario que se cumpla el teorema central del límite para responder (a) ? Explique .

10. Las cuentas de gastos de representación de los ejecutivos de una agencia de publicidad tiene una media de \$ 100 por persona y una desviación estándar de \$ 16 por persona. Si se seleccciona muestras aleatorias de 16 cuentas,
   (a) ¿Por abajo de qué valor en dinero caerá el 99% de las medias muestrales.?
   (b) ¿Qué proporción de las medias muestrales estará entre \$ 90 y \$ 110 ?
   (c) ¿Qué suposición se debe hacer para resolver (a) y (b) ?

11. De sus archivos , un ingeniero mecánico observa que el tiempo empleado en ensamblar cierto dispositivo a un equipo está distribuido normalmente con

media μ = 22 minutos y desviación estándar σ = 6 minutos. El ingeniero planea ensamblar 16 de estos dispositivos hoy. Suponga:

(1) que el tiempo en colocar un dispositivo es independiente del tiempo en ensamblar otro; y (2) que estos 16 ensamblajes representan una muestra aleatoria de la experiencia pasada.

(a) ¿Cuál es la probabilidad que 25 minutos o más sea el tiempo promedio por dispositivo para este ingeniero?

(b) ¿Cuál es la probabilidad de emplear 20 minutos o menos en el primer ensamble?

(c) Con el fin de poder llegar a una cita para jugar golf, el ingeniero tiene que emplear un promedio de 20 minutos o menos por dispositivo ¿Cuál es la probabilidad de llegar tarde a la cita?

(d) El ingeniero empieza a las 8.a.m. Si en el almuerzo se demora 40 minutos, ¿a qué hora es la cita para el golf?

12. El número de clientes por semana en cada tienda de una cadena de autoservicio tiene una media poblacional μ = 5000 clientes y una desviación estándar σ = 500. Si se selecciona una muestra aleatoria de 25 tiendas

(a) ¿Cuál es la probabilidad que la media muestral sea inferior a 5075 clientes por semana?

(b) ¿Dentro de qué límites se puede tener la certeza que caerá el 95% de la media muestral alrededor de la media poblacional?

13. Cierta marca de bombillas tiene una vida media de 257.1 horas y una desviación estándar de 20 horas. Un pasadiso sin ventanas de un edificio de apartamentos, tiene una instalación eléctrica, planeada para iluminar continuamente. El pasadiso consiste de cuatro bombillas, pero sólo una se enciende a la vez. Cuando ésta se quema, la proxima bombilla se enciende automáticamente. Este proceso continúa hasta que se quemen las cuatro bombillas. Cada semana al medio día, el administración viene y reemplaza las cuatro bombillas. ¿Cuál es la probabilidad que se quemen las cuatro bombillas antes que llegue el administrador para reemplazarlas?

14. Un fabricante de radios recibe semanalmente un cargamento de 100,000 pilas de 6 voltios. Para decidir si acepta o rechaza el cargamento, utiliza la siguiente regla de muestreo: mide la vida útil de 36 pilas de cada cargamento. Si la media de la muestra es de 50 ó más horas acepta el cargamento y en caso contrario, lo rechaza

(a) ¿Cuál es la probabilidad de aceptar un cargamento que tiene una vida -

útil media de 49 horas y una desviación estándar de 3 horas?

(b) ¿Cuál es la probabilidad de rechazar un cargamento que tiene una vida útil media de 50.5 y una desviación estándar de 3 horas?

(c) ¿Cuál es la probabilidad de rechazar un cargamento que tiene una vida útil media de 50 horas?¿cuál de aceptarlo?

15. Un procesador de alimentos envasa café en frascos de 400 grm. Para controlar el proceso, se utiliza la siguiente regla de muestreo: se selecciona 64 frascos cada hora. Si su peso medio es inferior a un valor crítico L, se detiene el proceso y se reajusta; en caso contrario, se continúa la operación sin detener el proceso. Determinar el valor de L de modo que haya una probabilidad de sólo 0.05 de detener el proceso cuando está envasado a un promedio de 407.5 grm. con una desviación estándar de 2.5 grm.

16. Un fabricante de café instantáneo envasa su producto en frascos de 300 grm neto. Para controlar el proceso automático de llenado, se selecciona cada hora una muestra de 36 frascos. Si el peso neto medio de la muestra, $\bar{X}$, está entre 301 y 302 gramos, el proceso se continúa, y en caso contrario, se detiene y se reajusta la máquina.

(a) ¿Cuál es la probabilidad de detener un proceso que está operando con una media de 301.5 grm. y una desviación estándar de 7.5 gr.

(b) ¿Cuál es la probabilidad de dejar que continúe un proceso que opera con una media de 302 grm. y una desviación estándar de 7.5 grm.?

17. En una partida grande de pilas eléctricas, la vida útil de ellas está distribuída normalmente con una media de 400 horas. Se sabe además que el 90% de las pilas tienen una vida útil comprendida entre 318 y 482 horas. Si se selecciona una muestra aleatoria de 100 pilas de esta partida, ¿cuál es la probabilidad que la media muestral sea mayor que 420 horas?

18. Una partida grande de rodamientos tiene un diámetro medio de 2.00 pulgadas con una desviación estándar de 0.02 pulgadas.

(a) Obtener un intervalo para el cual haya una probabilidad de 0.95 que el diámetro medio de una muestra aleatoria de 400 rodamientos esté incluído en él.

(b) ¿Cuál es la probabilidad que el diámetro medio de una muestra aleatoria de 100 rodamientos sea mayor que 2.003 pulgadas?

19. Calcular E(X) y Var(X), sabiendo que:

(a) X se distribuye normalmente ;

(b) $P[\bar{X} < 6] = 0.0228$

(c) $P[\bar{X} > 8] = 0.8413$

(d) $\bar{X}$ corresponde a una muestra de tamaño 4.

20. Con referencia al problema 10, si hay una población de 500 comprobantes de gastos de presentación, ¿cuales serían sus respuestas a las preguntas (a) y (b) de ese problema?

21. Un lote de 1000 cajas de cereal tiene un peso medio de doce onzas y una desviación estándar de 0.6 onzas. Se extrae una muestra al azar de 100 cajas sin reposición de ésta población. ¿Cuál es la probabilidad que el peso total sea:
(a) menor que 1190 onzas? (b) mayor que 1195 onzas? (c) entre 1190 y 1195 onzas?

22. Un lote de 500 cajas de galletas tienen un peso medio de 5.02 kg. y una desviación estándar de 0.3 kg.; se extrae una muestra al azar sin reemplazo de 100 cajas del lote. ¿Cuál es la probabilidad que tenga un peso medio (a) entre 4.96 y 5.00 kg.? (b) superior a 5.10 kg.?

23. En un colegio grande hay 500 niños matriculados en el primer grado. Si la desviación estándar del peso de los niños es de 2.5 kg.,¿cuál es la probabilidad que el peso medio de una muestra al azar sin reemplazo de 100 de estos niños y el peso medio de todos los niños difieren en más de medio kilogramo?

24. Una población está constituida por sólo 100 elementos. La población tiene una distribución normal con media 30 y desviación estándar 8. Calcular la probabilidad que el promedio muestral basada en una muestra al azar de tamaño 16 sin reemplazo, (a) sea menor que 32, (b) exceda a 28, (c) sea menor que 25, (d) se encuentre entre 33 y 34.

25. Periódicamente un fabricante determina el contenido de azufre en un producto químico y en cierto período de tiempo ha encontrado que el contenido promedio de azufre es 0.35% con desviación estándar 0.05%. Los lotes de este producto se envian a un cliente para quién el contenido de azufre es importante y que por lo tanto verifica la calidad llevando a cabo determinaciones de este dato en 4 muestras tomadas de cada lote. Si el contenido promedio de azufre de las cuatro excede a 0.5% se aplica una sanción al fabricante. ¿Puede el fabricante tener la suficiente confianza que esto

no sucederá, si el proceso de fabricación permanece bajo control?

26. Se encuentra que un cierto proceso de trituración los diámetros de las roscas se distribuyen en forma normal con media $\mu = 1.5$ cm. y desviación estándar $\sigma = 0.3$ cm.

    (a) Si d representa el diámetro medio calculado en muestras de tamaño 100, calcular la probabilidad que esta media muestral esté comprendida entre 1.2 cm. y 1.6 cm.

    (b) ¿Qué tamaño debe tener una muestra aleatoria de diámetros de roscas para afirmar con un 10% de probabilidad que el diámetro medio sea inferior a 1.4 cm.?

27. En un examen de caracter nacional las calificaciones produjeron media $\mu = 72$ y desviación estándar $\sigma = 10$. ¿Qué tan grande debe ser una muestra aleatoria de candidatos de una universidad para que tengan un 10% de probabilidad que la calificación media sea inferior a 70?

28. Suponga que la variable aleatoria X se distribuye exponencialmente con parámetro $\alpha = 0.1$. ¿Cuántas observaciones se debe hacer para afirmar con un 95% de probabilidad que la media muestral sea mayor que 8?

29. Se sabe que la vida de bombillas eléctricas es una variable aleatoria distribuída normalmente con media desconocida μ y desviación estándar 200 horas. El valor de un lote de 1,000 bombillas es $1000(\frac{1}{5000})\mu$ dólares. Un posible comprador propone tomar una muestra aleatoria de n bombillas y pagar al productor $1000(\frac{1}{5000})\overline{X}$ dólares por el lote de 1,000 bombillas. ¿Cuál debe ser, n (tamaño de la muestra) para que la probabilidad que el comprador no sobrepague ni subpague al productor en más 20 dólares, sea 0.95?

30. Suponga que las lámparas fabricadas mediante un proceso tienen una vida media $\mu = 2000$ horas y desviación estándar $\sigma = 250$ horas. Se considera aconsejable sustituir el proceso, si la vida media puede aumentarse al menos en un 10%. Un ingeniero desea poner a prueba un nuevo proceso admitiendo que la desviación estándar de la distribución de vidas de las lámparas es aproximadamente la misma que para el proceso considerado al principio. ¿Qué tamaño de muestra debe examinar si quiere que la probabilidad de no adoptar el nuevo proceso sea 0.01 aproximadamente cuando con él se obtiene en efecto lámparas con vida media de 2250 horas.

31. Sea $\bar{X}_1$ la media de una muestra de tamaño $n_1 = 2$, con reemplazamiento, de - la población finita 2,3 y 7. Similarmente $\bar{X}_2$ es la media de una muestra de tamaño $n_2 = 2$, con reemplazamiento, de la población finita 1,1 y 3. Hallar

$$\mu_{\bar{X}_1 - \bar{X}_2} \quad y \quad \sigma^2_{\bar{X}_1 - \bar{X}_2}$$

32. Una muestra de tamaño 25 se toma de una población normal con media de 80 - y desviación estándar de 5. Una segunda muestra de tamaño 36 se toma de una población normal con media 75 y desviación estándar de 3. Hallar la probabilidad que la media de la muestra de 25 observaciones excede a la media de la muestra de las 36 observaciones es por lo menos 3.4 pero menos que - 5.9.

33. Un industrial compró la producción total de 1 año de los tubos de imagen - de T.V. producidos por una fábrica determinada. Los datos técnicos proporcionados por los fabricantes son los siguientes:

    Duración media de vida de los tubos, 2800 hs.

    Desviación típica, 500 horas.

    Si se consideran 2 muestras de tubos, una de tamaño 120 y la otra de tamaño 200, calcular:

    (1°) La probabilidad que la duración media de vida de la primera muestra no sea superior en más de 100 horas a la duración media de vida de la 2a. muestra.

    (2°) La probabilidad que sea superior en más de 200 horas.

34. Dos marcas de focos "economico" y "vida eterna" tienen durabilidades (en - horas) que son $N(1400,200^2)$ y $N(2000,250^2)$ respectivamente. Si se prueba - la durabilidad de 4 focos de cada marca, ¿cuál es la probabilidad que - la vida de los económicos sea mayor que la vida media de los de vida eterna?

35. Suponga que se sabe que los resultados de un método que mide la dureza de los metales siguen una distribución normal alrededor del valor real (o sea no hay error sistemático), con desviación estándar σ . El método se va a - utilizar para estimar la diferencia entre dos aleaciones A y B, o sea para estimar $\mu_A - \mu_B$, donde $\mu_A - \mu_B$ son los valores de dureza (desconocidos) de las dos aleaciones. Las pruebas en las aleaciones A y B darán resultados - que son $N(\mu_A , \sigma^2)$ y $N(\mu_B , \sigma^2)$ respectivamente. Si n_A representa el nú-

mero de pruebas realizadas con la aleación A y n_B el número de pruebas realizadas con la aleación B, ¿en más de qué valor $\bar{X}_A - \bar{X}_B$ no difiere de $\mu_A - \mu_B$ con 95% de probabilidad?

36. Cierta marca de rodamiento de bolas tiene un peso medio de 0.5 onza y una desviación estándar de 0.02 onza. Se toman independientemente dos muestras al azar, con reposición de cierto día de producción, con $n_1 = 500$ y $n_2 = 800$. ¿Cuál es la probabilidad que las dos medias de las muestras difieran (a) en más de 0.002 onza? (b) en menos de 0.001 onza?

37. Con referencia al ejemplo 12, si las muestras se extraen sin reposición, ¿cuál sería su respuesta a ese problema?

38. Supongamos que $\bar{X}_1$ y $\bar{X}_2$ son medias de dos muestras de tamaño n de una población con varianza σ^2. Determínese n de modo que la probabilidad que las dos medias muestrales difieren en un valor superior a σ sea, aproximadamente 0.01.

39. Supongamos que $\bar{X}_1$ y $\bar{X}_2$ son medias de dos muestras aleatorias independientes de tamaño n. Cada una de las observaciones se supone normalmente distribuida con media y varianza común 2. Determine, n, de manera que la probabilidad de que $\bar{X}_1$ y $\bar{X}_2$ difieran en menos de 2 con 95% de probabilidad.

40. Una máquina empaqueta (envuelve y sella) porciones iguales de cereales. Si el paquete no está derecho se considera defectuoso; la máquina produce un 10% de paquetes defectuosos. Un lote grande de paquetes acaba de ser producido. Se selecciona una muestra al azar de cinco paquetes del lote de producción Determinar
   (a) la distribución de probabilidad para la proporción de paquetes defectuosos, y haga su gráfica.
   (b) $E(\bar{P})$ y $Var(\bar{P})$, ¿se cumple que $E(\bar{P}) = p$ y $Var(\bar{P}) = (1/n)p(1-p)$?
   (c) ¿Cómo está sesgada ésta distribución muestral para $\bar{P}$?
   (d) Halle la distribución de probabilidad para X, número de paquetes defectuosos y haga su gráfica.

41. En una urna hay diez bolas, cinco de las cuales son negras y cinco blancas. Se extrae una muestra al azar de seis, sin reposición, evalué la distribución por muestreo de la proporción de bolas blancas. (NOTA: véase la nota del problema 2)

42. Si de una gran población con $p = 1/3$ se extrae una muestra al azar de 180 unidades. Calcular : (a) $\mu_{\bar{p}}$ y $\sigma_{\bar{p}}$. (b) la probabilidad que

$50/180 < \bar{P} < 70/180$.

43. Con base en datos pasados, el 30% de las compras con tarjeta de crédito en una tienda muy conocida son por cantidades superiores a \$100 . Si se seleccionan muestras aleatorias de 100 compras; (a) ¿qué proporción de las muestras es posible que tengan entre 20% y 30% de compras mayores que \$100 ? (b) ¿Dentro de qué límites simétricos del porcentaje de la población caerá el 95% de los porcentajes de la muestra?

44. Del profesorado de cierta universidad, 1/6 son mujeres. Si de esta población se extrae una muestra al azar de 180. Calcular:
(a) $P[\bar{P} \geq 0.3]$ (b) $P[0.1 \leq \bar{P} \leq 0.25]$

45. Una población de 5 tiendas va a ser muestreada con el fin de estimar la proporción de las tiendas de la población que su línea comercial tienen una cierta marca de televisor. Suponga que la población es, de hecho, la siguiente:

Tienda	*Característica*
A	tiene la marca de televisor
B	no tiene la marca de televisor
C	no tiene la marca de televisor
D	tiene la marca de televisor
E	tiene la marca de televisor

Se toma una muestra de dos tiendas de esta población.
(a) ¿Qué proporción de las tiendas de la población tienen este tipo de televisor?
(b) Obtenga la distribución de muestreo de $\bar{P}$ por enumeración de todas las combinaciones posibles para las muestras.
(c) ¿Cuál es la probabilidad $\bar{P}$ que la proporción de la muestra sea (1) 1, (2) menor que 0.50.
(d) ¿Cuál es la probabilidad que la proporción de la muestra no difiera de la proporción de la población por más de 0.20?
(e) Calcule la media de la distribución de muestreo de $\bar{P}$ y la desviación estándar de la distribución de muestreo de $\bar{P}$.

46. En una población de 5 archivos, la proporción de las que tienen una parte incorrectamente llenada es $p = 1/5$. Se va a elegir una muestra aleatoria de 3 archivos. Hallar la probabilidad que la proporción de la muestra sea $\bar{P} = 1/3$.

47. La probabilidad que un nuevo empleado esté con la misma firma al cabo - de un año es de 0.45. Suponga que se aplica la distribución binomial.
    (a) Obtenga la distribución muestral de $\bar{P}$, la proporción de siete nuevos empleados que están todavía en la firma al cabo de un año.
    (b) ¿Cuál es la probabilidad que la proporción de la muestra sea (1) - 4/7, (2) menor que 3/7. ¿Cuál es la media de la distribución muestral de $\bar{P}$?. ¿Cuál es su desviación estándar?

48. Al reparar cierto tipo de máquina empacadora ocacionalmente ocurre una complicación que requiere asistencia técnica exterior. Se desea estimar la - proporción de trabajos de reparación que requieren asistencia exterior, basándose en una muestra aleatoria simple de 100 trabajos de reparación terminada recientemente. Suponga que en realidad se requiere asistencia exterior en un 15% de los trabajos de reparación, esto es, que la proporción del proceso es $p = 0.15$.
    (a) ¿Cuál es la media y la desviación estándar de la distribución muestral de $\bar{P}$?
    (b) ¿Cuál es la probabilidad que $\bar{P}$ esté en un intervalo de radio 0.05 centrado en la proporción del proceso?
    ¿Qué $\bar{P}$ esté entre 0.12 y 0.20?
    (c) ¿Dentro de qué intervalo caerá la proporción de la muestra el 90% de las veces?. Utilice límites simétricos alrededor de la proporción del proceso.
    (d) ¿Cuál es el intervalo que corresponde a la parte c para una muestra - aleatoria simple de 400 trabajos? ¿Qué efecto tiene el incremento en - el tamaño de la muestra?
    (e) Para una muestra de tamaño dado, ¿es la distribución muestral de $\bar{P}$ más variable si p está cercana a 0, si p es 0.5, o si p está cercana a 1?

49. Se observa una máquina en puntos aleatorios del tiempo para estimar la proporción del tiempo en que no se encuentra en estado productivo. Suponga - que la proporción de tiempo fuera de la producción, para la máquina es realmente $p = 0.25$. ¿Cuál debe ser el tamaño de la muestra, n si la desviación estándar de la distribución muestral de $\bar{P}$ debe ser $\sigma_{\bar{P}} = 0.02$?

50. Un embarque de 5,000 piezas va a ser muestreado para determinar qué tan - aceptable es su calidad. Suponer que la proporción defectuosa en el embarque es realmente 0.10. Ignorando el factor de corrección finita, ¿Cuál de

be ser el tamaño de la muestra, n , si la desviación estándar de la distribución muestral de $\bar{P}$ debe ser 0.01?

7.3 OTRAS DISTRIBUCIONES DE PROBABILIDAD USADAS EN PRUEBAS

En la sección anterior hemos visto las distribuciones muestrales para una muestra grande ($n \geq 30$), estas distribuciones eran aproximadamente una distribución normal o que el tamaño de la muestra, era suficientemente grande de tal manera que se cumplía el teorema central del límite. Sin embargo cuando la muestra es pequeña ($n < 30$), no podemos aplicar el teorema central del límite y si la población no es normal, tampoco podemos suponer que la distribución muestral es normal. Daremos aqui distribuciones relacionados con la distribución normal, la chi-cuadrado, t-student y la distribución F. Estas distribuciones a menudo son asumidas para muestras pequeñas.

7.3.1 DISTRIBUCION CHI-CUADRADO

DEFINICION 7.3.1 Sean $Z_1, Z_2, \ldots, Z_n$, variables aleatorias independientes distribuídas normalmente, cada una con media 0 y varianza 1. La variable aleatoria

$$X^2 = Z_1^2 + Z_2^2 + \ldots + Z_n^2$$

Se dice que es una variable aleatoria *chi-cuadrado* con n grados de libertad (o que tiene una distribución chi-cuadrado con n grados de libertad) si su función de densidad está dado por

$$f_{X^2}(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})} x^{n/2-1} e^{-x/2} & , \quad 0 < x < \infty \\ 0 & , \quad \text{en otros casos} \end{cases}$$

donde Γ representa la función gamma (ver 6.2.3)

Observe que los valores que toma la variable aleatoria chi-cuadrado, son todos los reales positivos, debido a que es una suma de cuadrados.

Grado de Libertad n, es el número de variables aleatorias independientes que -

se suman. También el grado de libertadad se puede concebir como un parámetro - asociado con la distribución de probabilidad o como al número de variables que pueden variar libremente.

NOTACION Cuando una variable aleatoria X tiene una distribución chi-cuadrado - con r grados de libertad, escribiremos abreviadamente que X es χ^2_r .

La *media y la varianza* de la variable aleatoria chi-cuadrado con r grados de libertad son :

$$\mu = E(\chi^2) = r$$

$$\sigma^2 = Var(\chi^2) = 2r$$

Es decir, la media es igual al número de grados de libertad y su varianza es - igual a dos veces el número de grados de libertad. En otras palabras estos momentos se expresan en términos de los grados de libertad.

La fig. 7.3.1 da un bosquejo de la función de densidad de la variable aleatoria chi-cuadrado, para distintos grados de libertad.

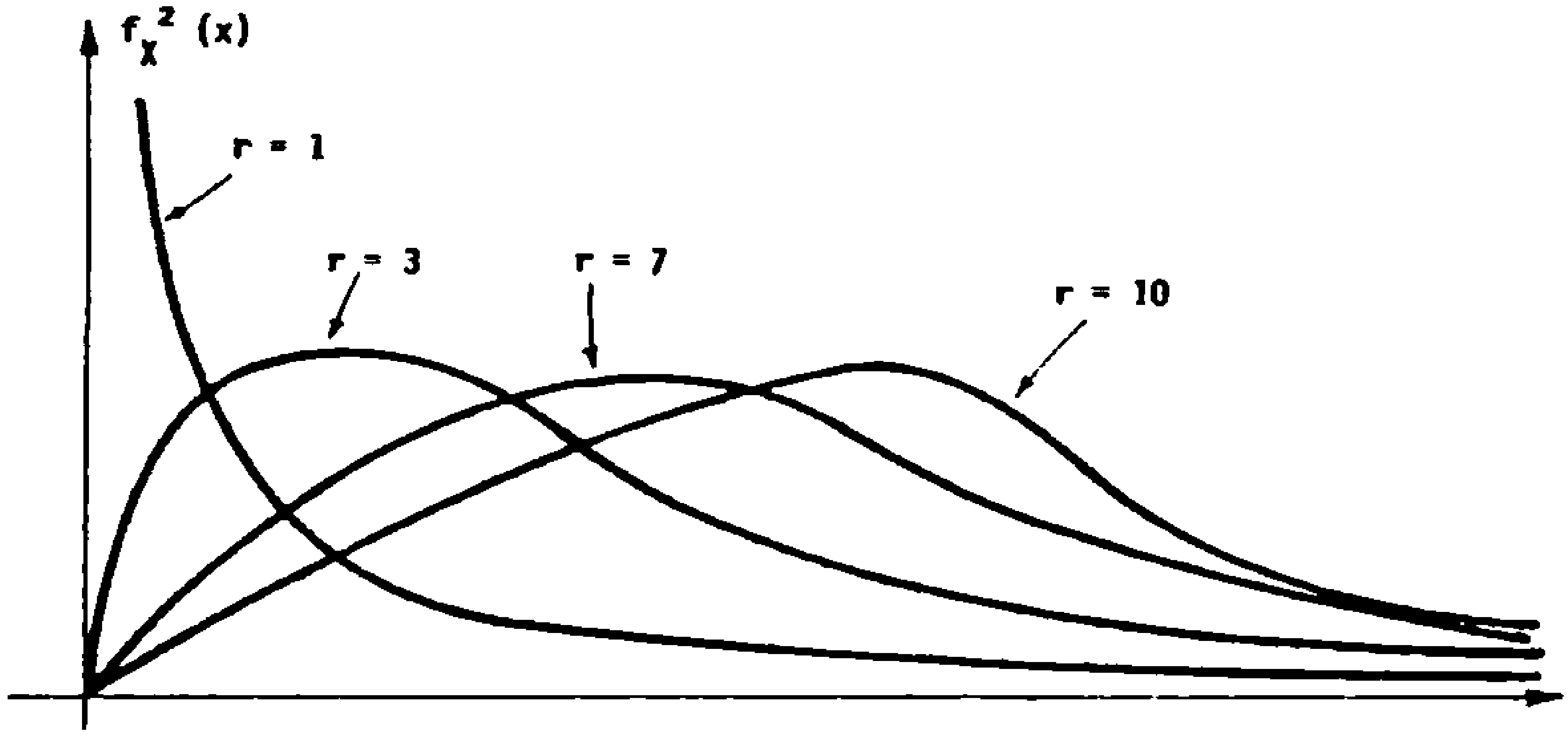


Fig. 7.3.1. Función de densidad de la variable aleatoria chi-cuadrado

Observe que las distribuciones chi-cuadrado son una familia de distribuciones contínuas positivamente asimétricas; sin embargo cuando r, (r = grado - de libertad) aumenta la chi-cuadrado se aproxima a una distribución normal. - por esta razón, es que en la práctica, cuando r es grande ($r > 30$), la probabilidad de la chi-cuadrado puede calcularse empleando aproximación normal como - veremos posteriormente.

Debido a que la distribución chi-cuadrado es importante en las aplicaciones, principalmente en inferencia estadística algunas de las cuales citaremos posteriormente; la *función de distribución* $F(x)$ están preparadas en tablas (ver tabla IV), para valores seleccionados de r y x^2. Por lo tanto, se puede encontrar en la tabla, la probabilidad que la variable aleatoria X que tiene una distribución χ^2_r $(1 \leq r \leq 30)$ sea menor o igual a un valor constante χ^2_α, representado por

$$P[X < \chi^2_\alpha] = \alpha \quad , \qquad 0 < \alpha < 1$$

Es decir,

$$P[X < \chi^2_\alpha] = \int_0^{\chi^2_\alpha} f(x)dx = \int_0^{\chi^2_\alpha} \frac{1}{2^{r/2}\,\Gamma(\frac{r}{2})} x^{\frac{r}{2}-1} e^{-\frac{r}{2}} dx = \alpha$$

Vea la figura 7.3.2.

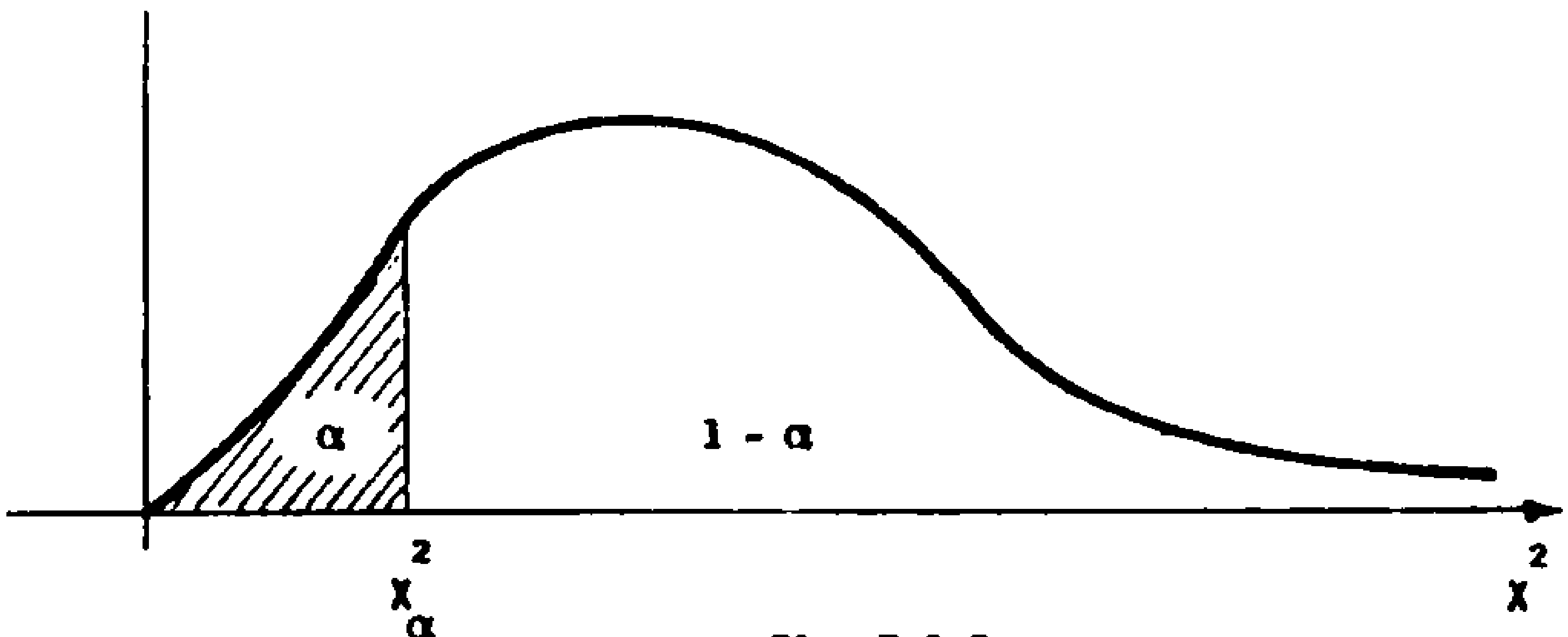


Fig. 7.3.2

Note que $P[X > \chi^2_\alpha] = 1 - \alpha$

Obsérvese, puesto que existe una distribución chi-cuadrado diferente para cada valor de r, resulta impráctico proporcionar tablas de áreas completas. En lugar de esto, la tabla IV presenta un resumen de la información más esencial acerca de la distribución.

Nótese que la columna de la izquierda de ésta tabla tiene como encabezado "grados de libertad". Cada fila en la tabla contiene información sobre la distribución chi-cuadrado correspondiente a los grados de libertad indicados, r. Es decir, cada fila de esta tabla corresponde a una distribución chi-cuadrado particular. Por, ejemplo, si $r = 5$, $\chi^2_{0.10} = 1.61$. Por lo tanto,

$$P[X \leq \chi^2_{0.10}] = P[X \leq 1.61] = 0.10$$

EJEMPLO 1 Si X es una variable aleatoria χ^2_{17}. Calcular :

(a) $P[X < 7.564]$; (b) $P[X > 27.59]$; (c) $P[6.408 < X \leqslant 27.59]$

SOLUCION

(a) $P[X < 7.564] = P[X \leqslant \chi^2_{0.025}] = 0.025$

(b) $P[X > 27.59] = 1 - P[X \leqslant 27.59] = 1 - P[X \leqslant \chi^2_{0.95}]$

$= 1 - 0.95 = 0.05$

(c) $P[6.408 < X \leqslant 27.59] = P[X \leqslant 27.59] - P[X \leqslant 6.408]$

$= P[X \leqslant \chi^2_{0.95}] - P[X \leqslant \chi^2_{0.01}]$

$= 0.95 - 0.01 = 0.94$

EJEMPLO 2 Si X es una variable aleatoria con una distribución χ^2_{23} . Hallar - a y b tal que

$$P[a < X < b] = 0.95 \quad y \quad P[X < a] = 0.025$$

SOLUCION $r = 23$, $a = \chi^2_{0.025} = 11.69$

$$P[a < X < b] = P[X \leqslant b] - P[X \leqslant a] = 0.95$$

de donde $P[X \leqslant b] = P[X \leqslant a] + 0.95 = 0.025 + 0.95$

$= 0.975$

Luego, $b = \chi^2_{0.975} = 38.08$

EJEMPLO 3 Sea $X_1, X_2, \ldots, X_{10}$ una muestra aleatoria de tamaño 10 de una variable aleatoria con distribución χ^2_{19} . Hallar la probabilidad de que exactamente dos de los diez valores muestrales exceda a 30.14

SOLUCION Sea X la variable aleatoria con χ^2_{19}. Debemos calcular antes

$$P[X > 30.14] = 1 - P[X \leqslant 30.14] = 1 - P[X \leqslant \chi^2_{0.95}]$$

$$= 1 - 0.95 = 0.05 .$$

Sea la variable aleatoria Y definida por: número de valores mayores que 30.14 en la muestra de 10. Luego

$$P[Y = 2] = \binom{10}{2}(0.05)^2(0.95)^8$$

Usando tablas obtenemos

$$P[Y = 2] = 0.0746 .$$

Si queremos calcular $P[X < \chi^2_\alpha]$ donde X tiene una distribución χ^2_r con $r > 30$, de manera que no se puede obtener directamente de la tabla de la distribución chi-cuadrado, aplicaremos la aproximación normal usando la siguiente propiedad.

TEOREMA 7.3.1 Si X es una variable aleatoria que tiene una distribución chi-cuadrado con r grados de libertad. Entonces para r suficientemente grande, la variable aleatoria $\sqrt{2X}$, tiene aproximadamente una distribución $N(\sqrt{2r-1}, 1)$ Es decir, si X es χ^2_r, con r grande, entonces

$$\sqrt{2X} \longrightarrow N(\sqrt{2r-1},\ 1)$$

Luego, la variable aleatoria, $Z = \sqrt{2X} - \sqrt{2r-1}$ tiene aproximadamente una distribución $N(0,1)$. Por lo tanto, para $r > 30$

$$\begin{aligned}
\alpha &= P[X \leq \chi^2_\alpha] = P[\sqrt{2X} \leq \sqrt{2\chi^2_\alpha}] \\
&= P[\sqrt{2X} - \sqrt{2r-1} \leq \sqrt{2\chi^2_\alpha} - \sqrt{2r-1}] \\
&= P[Z \leq \sqrt{2\chi^2_\alpha} - \sqrt{2r-1}] = \Phi(\sqrt{2\chi^2_\alpha} - \sqrt{2r-1}) \\
&= \Phi(z_\alpha)
\end{aligned}$$

En consecuencia, para $r > 30$, se tiene

$$\sqrt{2\chi^2_\alpha} - \sqrt{2r-1} \approx z_\alpha$$

de manera que $\boxed{\chi^2_\alpha = \frac{1}{2}[z_\alpha + \sqrt{2r-1}]^2}$

por, ejemplo si $r = 41$, se tiene

$$\begin{aligned}
\chi^2_{0.10} &= \frac{1}{2}[z_{0.1} + \sqrt{2(41)-1}]^2 = \frac{1}{2}[-1.28 + 9]^2 \\
&= 29.799
\end{aligned}$$

en tanto que

$$\begin{aligned}
\chi^2_{0.95} &= \frac{1}{2}[z_{0.95} + \sqrt{81}]^2 = \frac{1}{2}[1.64 + 9]^2 \\
&= 56.605
\end{aligned}$$

El siguiente teorema relaciona la distribución chi-cuadrado con la distribución $N(\mu,\sigma^2)$

TEOREMA 7.3.2 Si la variable aleatoria X es $N(\mu,\sigma^2)$, entonces la variable aleatoria, $Y = (X-\mu)^2/\sigma^2$ es una χ^2_1.

EJEMPLO 4 Si X es una variable aleatoria N(7,4). Calcular

$$P[15.36 \leq (X - 7)^2 < 20.08]$$

SOLUCION $X \to N(7,4)$, entonces la variable aleatoria

$$Y = \frac{(X-7)^2}{4}$$

tiene una distribución χ_1^2 . Luego ,

$$\begin{aligned} P[15.36 \leq (X - 7)^2 < 20.08] &= P\left[\frac{15.36}{4} \leq \frac{(X-7)^2}{4} < \frac{20.08}{4}\right] \\ &= P\left[3.84 \leq \frac{(X-7)^2}{4} < 5.02\right] \\ &= P[Y \leq 5.02] - P[Y \leq 3.84] = 0.975-0.95 \\ &= 0.025 . \end{aligned}$$

El teorema siguiente es una generalización del teorema anterior.

TEOREMA 7.3.3 Sea $X_1, X_2, \ldots, X_n$, una muestra aleatoria de una variable aleatoria normal X con media μ y varianza σ^2 . Entonces, la variable aleatoria

$$Y = \sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$$

tiene una distribución χ_n^2 .

Igualmente el lector recordará que $\bar{X} = \frac{1}{2}\sum_{i=1}^{n} X_i$ es una variable aleatoria normal con media μ y varianza $\frac{\sigma^2}{n}$, por lo tanto

$$\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \quad \text{es} \quad N(0,1)$$

entonces, $\frac{(\bar{X} - \mu)^2 n}{\sigma^2}$ es una variable aleatoria que tiene una distribución χ_1^2 .

Otra propiedad igualmente importante de la distribución chi-cuadrado es la propiedad reproductiva.

TEOREMA 7.3.4 (PROPIEDAD ADITIVA DE LA CHI-CUADRADO) Sea $\chi_1^2, \chi_2^2, \ldots, \chi_p^2$ variables aleatorias chi-cuadrados independientes con grados de libertar $r_1, r_2, \ldots, r_p$ respectivamente, entonces la variable aleatoria

$$\chi^2 = \chi_1^2 + \chi_2^2 + \ldots + \chi_p^2$$

Sigue una distribución chi-cuadrado con grado de libertad igual a

$$r = \sum_{i=1}^{p} r_i$$

7.3.2 DISTRIBUCION DE LA VARIANZA MUESTRAL

TEOREMA 7.3.5 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de una población normal con media μ y varianza σ^2. Sea $\bar{X}$ y S^2 la media muestral y varianza muestral respectivamente, entonces

(a) Las variables aleatorias $\bar{X}$ y S^2 son independientes .

(b) La función de la varianza muestral $\dfrac{(n-1)S^2}{\sigma^2}$ tiene una distribución χ^2_{n-1}

DEMOSTRACION Demostraremos sólo la parte (b)

1. $$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2}$$

2. La variable aleatoria
$$\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2}$$
tiene una distribución χ^2_n, puesto que cada término $(X_i - \mu)/\sigma$ son variables aleatorias normales estándars e independientes (teorema 7.3.3)

3. Consideremos
$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}\left[(X_i - \bar{X}) + (\bar{X} - \mu)\right]^2$$
$$= \sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(\bar{X} - \mu)^2 + 2(\bar{X} - \mu)\sum_{i=1}^{n}(X_i - \bar{X})$$
$$= \sum_{i=1}^{n}(X_i - \bar{X})^2 + n(\bar{X} - \mu)^2$$

por lo tanto n

$$\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} + \frac{(\bar{X} - \mu)^2}{\sigma^2/n}$$
$$= \frac{(n-1)\,S^2}{\sigma^2} + \frac{(\bar{X} - \mu)^2}{\sigma^2/n}$$

Desde que $(\bar{X} - \mu)^2/(\sigma^2/n)$ tiene una distribución χ_1^2 . Por la parte (a) $\bar{X}$ y -

y S^2 son independientes, y $\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$ tiene una distribución χ_n^2, por la propiedad aditiva de la chi-cuadrado (teorema 7.3.4), concluímos que la distribución de $(n - 1)S^2/\sigma^2$ es χ_{n-1}^2 .

NOTA 1 Es notable que $\bar{X}$ y S^2 resulta variables aleatorias independientes, siendo ambas funciones que se calculan de la misma muestra aleatoria.

2. Observe, que se usa S^2/σ^2 como una medida de aproximación en vez de $S^2 - \sigma^2$, debido a que la distribución de S^2/σ^2 se obtiene fácilmente como se ve en la demostración del teorema 7.3.5, en cambio la distribución de $S^2 - \sigma^2$ es difícil de obtener.

EJEMPLO 5 Calcular $P[0.618 \leq S^2/\sigma^2 \leq 1.60]$, si S^2 está basado en una muestra aleatoria de 11 observaciones de una variable aleatoria distribuída normalmente con media desconocida μ y varianza desconocida σ^2 .

SOLUCION Tamaño de la muestra $n = 11$, entonces, la variable aleatoria $\frac{10\,S^2}{\sigma^2}$ tiene una distribución chi-cuadrado con 10 grados de libertad. Por lo tanto

$$\begin{aligned} P[0.618 \leq S^2/\sigma^2 \leq 1.60] &= P[6.18 \leq 10S^2/\sigma^2 \leq 16.0] \\ &= P[10S^2/\sigma^2 \leq 16.0] - P[10S^2/\sigma^2 \leq 6.18] \\ &= P[10S^2/\sigma^2 \leq \chi^2_{0.90}] - P[10S^2/\sigma^2 \leq \chi^2_{0.20}] \\ &= 0.90 - 0.20 = 0.7 \end{aligned}$$

$\alpha = 0.2$ se ha obtenido interpolando, ya que el valor 6.18 no está en la tabla.

7.3.3 DISTRIBUCION t - DE STUDENT

DEFINICION 7.3.2 Sea **Z** una variable aleatoria normal con media 0 y varianza 1 Sea **Y** una variable aleatoria que tiene una distribución chi-cuadrado con r grados de libertad, y si **Z** e **Y** son independientes, entonces la variable aleatoria

$$T = \frac{Z}{\sqrt{Y/r}} = \frac{Z\sqrt{r}}{\sqrt{Y}}$$

se dice que tiene una distribución t - de student (o simplemente tiene una distribución t), con r grados de libertad, y su función de densidad está dado por

$$f(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left[1 + \frac{t^2}{n}\right]^{-(n+1)/2}, \qquad -\infty < t < \infty$$

Observe que la distribución de la variable aleatoria T, está completamente determinado sólo por el parámetro n. Entonces, hay una distribución t correspondiente a cada grado de libertad. En la figura 7.3.3 se presenta un bosquejo de la función de densidad de la variable aleatoria T, para diferentes grados de libertad. En la misma figura se da, la gráfica de la normal estándar. Note, la simetría de la distribución t alrededor de $t = 0$ y varía de menos infinito a más infinito.

La media y la varianza de la distribución t con n grados de libertad están dados por,

$$\mu = E(T) = 0 \quad , \quad n > 1$$

$$\sigma^2 = Var(T) = \frac{n}{n-2}, \quad n > 2$$

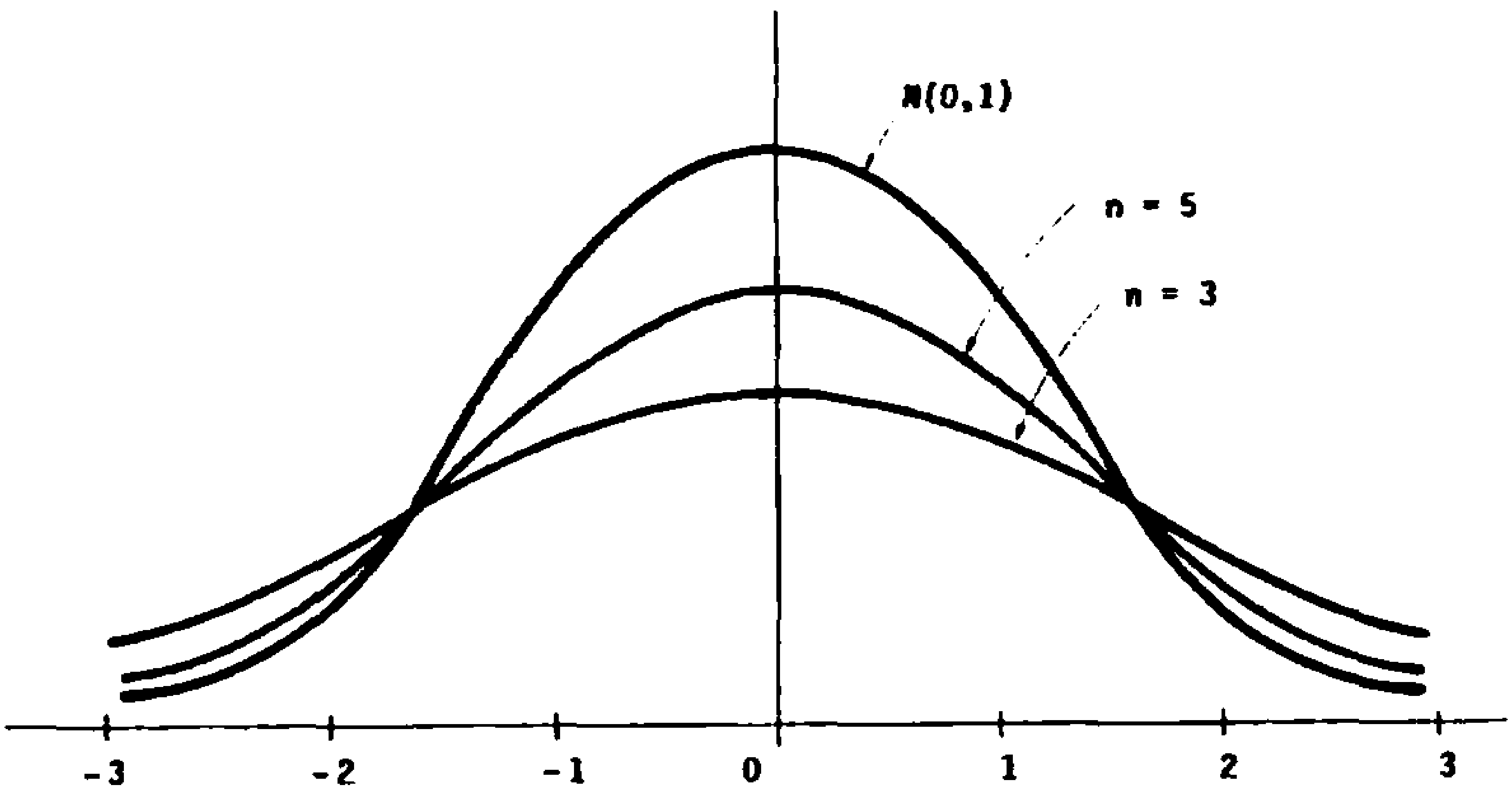


Fig. 7.3.3

Así, la distribución t no tiene media cuando $n = 1$ y la varianza no existe cuando $n = 1$ ó 2. La distribución t es muy similar a la distribución normal estándar, ya que ambas varían de $-\infty$ a ∞, son simétricas y centradas alrededor de $t = 0$, es decir su media es cero, pero la distribución t tiene mayor dispersión que la distribución normal estándar, esto se observa de la varianza $\sigma^2 = \frac{n}{n-2}$, pués: la varianza de la distribución t, es mayor que 1, la varianza de la distribución t , varía para diferentes grados de libertad (existe una distribución t , diferente para cada grado de libertad). Finalmente la varianza de la distribución t , se aproxima a 1 cuando el grado de libertad es grande. Por lo tanto, la distribución t , se aproxima a la distribución normal estándar cuando el grado de libertad es suficientemente grande. En la práctica se trata a la distribución t , como N(0,1) cuando $r > 30$.

Tabulación de la distribución t. Debido a la importancia de la distribución t, en inferencia estadística y la dificultad para evaluar la función de distribución de la variable aleatoria T, éstas se dan en una tabla (ver tabla V). Puesto que existe una distribución t , distinta para cada grado de libertad, no es práctico proporcionar una tabla de áreas completas para todas las distribuciones t , que corresponden a los diferentes grados de libertad, se presenta en la tabla V sólo un resumen de la información de cada una de estas distribuciones, para $r = 1,2,\dots,30$. En el encabezado de la columna de la izquierda, dice grados de libertad y cada fila de esta tabla corresponde a una distribución t, particular . Por lo tanto, la probabilidad que la variable aleatoria T sea menor o igual a una constante $t = t_\alpha$, es decir

$$P[T \leq t_\alpha] = \int_{-\infty}^{t_\alpha} f_T(z)dz = \alpha$$

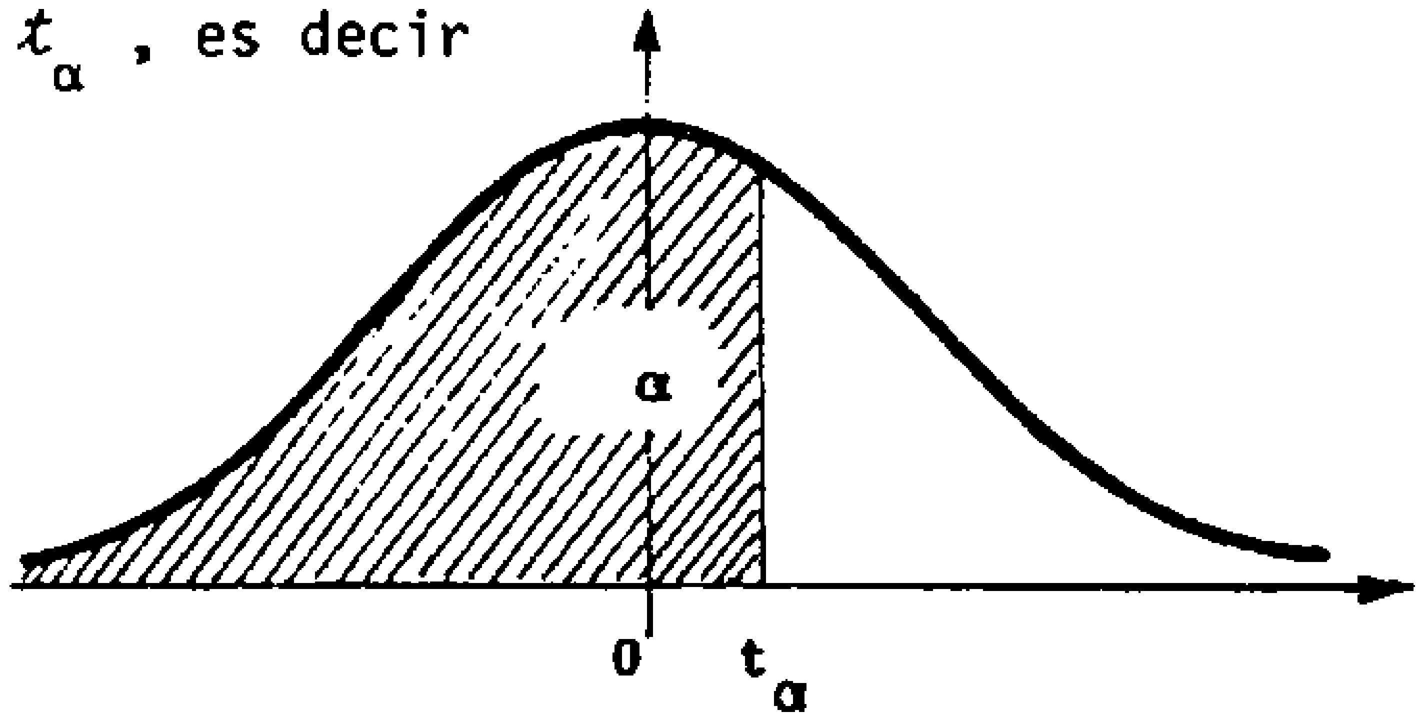


Fig. 7.3.4

Se encuentra en la tabla V. (Ver fig. 7.3.4)
por, ejemplo si $r = 5$

$$P[T \leq t_{0.90}] = P[T \leq 1.476] = 0.90$$

$$P[T \leq t_{0.10}] = P[T \leq -t_{0.9}] = P[T \leq -1.476] = 0.10$$

Ya que $t_\alpha = -t_{1-\alpha}$,por la simetría de la distribución t . (ver fig. 7.3.5).

NOTA Los valores t_α, para $\alpha < 0.5$, pueden obtenerse de la relación

$$t_\alpha = -t_{1-\alpha}$$

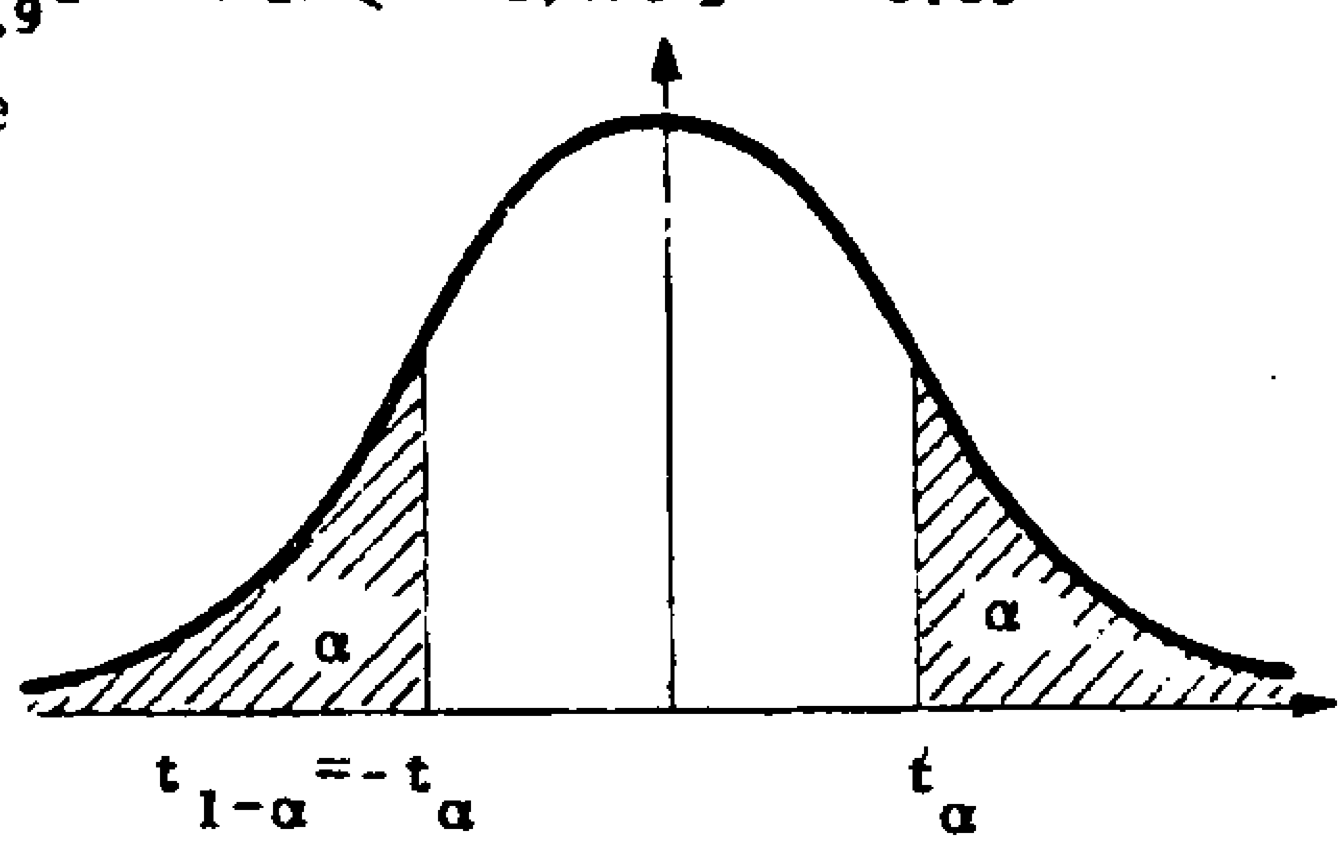


Fig. 7.3.5

EJEMPLO 6 Sea T una variable aleatoria que tiene una distribución t con varianza $\sigma^2 = \frac{5}{4}$. Calcular,

$$P[-1.812 \leq T \leq 2.228]$$

SOLUCION Desde que $\sigma^2 = \frac{r}{r-1} = \frac{5}{4}$, $r = 10$, entonces

$$\begin{aligned} P[-1.812 \leq T \leq 2.228] &= P[T \leq 2.228] - P[T \leq -1.812] \\ &= P[T \leq t_{0.975}] - [1 - P[T \leq 1.812]] \\ &= P[T \leq t_{0.975}] - 1 + P[T \leq t_{0.9}] \\ &= 0.975 + 0.95 - 1 = 0.925 . \end{aligned}$$

EJEMPLO 7 Sea T una variable aleatoria que tiene una distribución t con 14 grados de libertad. Hallar una constante a tal que

$$P[|T| < a] = 0.90$$

SOLUCION

$$\begin{aligned} P[|T| < a] &= P[-a < T < a] = P[T < a] - P[T < -a] \\ &= P[T < a] - [1 - P[T < a]] \\ &= 2P[T < a] - 1 = 0.90 \end{aligned}$$

entonces $P[T < a] = 0.95$, de donde $a = 1.761$

7.3.4 DISTRIBUCION DE $(\bar{X}-\mu)\sqrt{n}/s$

En 7.2.2 se mostro que la distribución de la variable aleatoria $(\bar{X}-\mu)\sqrt{n}/\sigma$ es $N(0,1)$. Esta variable aleatoria es útil para hacer inferencias (tomar decisiones) con respecto a μ cuando se conoce σ. Por ejemplo la expresión.

$$P\left[-1.96 \leq \frac{(\bar{X}-\mu)\sqrt{n}}{\sigma} \leq 1.96\right] = 0.95$$

puede escribirse, como

$$P\left[\frac{-1.96\sigma}{\sqrt{n}} \leq \bar{X}-\mu \leq \frac{1.96\sigma}{\sqrt{n}}\right] = 0.95$$

Así, la diferencia entre el estimado, $\bar{X}$ y la cantidad a estimarse μ, está entre $\pm 1.96\sigma/\sqrt{n}$ con probabilidad de 0.95. Pero cuando no se conoce σ, es natural reemplazar σ por un estimado S y obtener la variable aleatoria, $(\bar{X}-\mu)\sqrt{n}/S$, que desafortunadamente no tiene una distribución $N(0,1)$. En esta secccién se determina la distribución de esta variable aleatoria.

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n, de una variable aleatoria X con distribución $N(\mu, \sigma^2)$, hemos visto :

1. la variable aleatoria $(\bar{X} - \mu)\sqrt{n}/\sigma$ tiene una distribución $N(0,1)$.
2. la variable aleatoria $(n - 1)S^2/\sigma^2$ tiene una distribución chi-cuadrado con $n - 1$ grados de libertad. (teorema 7.3.5)
3. $\bar{X}$ y S^2 son variables aleatorias independientes (teorema 7.3.5).

   De acuerdo con la definición de la variable aleatoria T, de (1) y (2), la variable aleatoria

$$\frac{\dfrac{(\bar{X} - \mu)\sqrt{n}}{\sigma}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/n-1}} = \frac{(\bar{X} - \mu)\sqrt{n}}{S}$$

tiene una distribución t con $n - 1$ grados de libertad. Esta variable aleatoria se usa para estimar μ cuando no se conoce la desviación estándar .

EJEMPLO 8 Calcular $P[- 1.383 \leq (\bar{X} - \mu)\sqrt{10}/S]$, donde $\bar{X}$ está basada en 10 observaciones y S en 20 observaciones diferentes.

SOLUCION Ya que S se basa en una muestra de tamaño 20, entonces, la variable aleatoria $\dfrac{(\bar{X} - \mu)\sqrt{10}}{S}$, tiene una distribución t con 19 grados se libertad. Por lo tanto

$$P[- 1.383 \leq (\bar{X} - \mu)\sqrt{10}/S] = P[(\bar{X} - \mu)\sqrt{10}/S \leq 1.383] = 0.91$$

(interpolando)

7.3.5 DISTRIBUCION DE LA DIFERENCIA DE DOS MEDIAS MUESTRALES CON VARIANZAS DESCONOCIDAS PERO IGUALES

Otro resultado importante, que frecuentemente se está interesado en examinar, es la distribución de la variable aleatoria $\bar{X} - \bar{Y}$, cuando no se conoce la varianza.

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n, de una variable aleatoria X con distribución $N(\mu_X, \sigma^2)$; Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de tamaño m de una variable aleatoria Y, con distribución $N(\mu_Y, \sigma^2)$, sean también las dos muestras independientes. Entonces se tiene :

1. La distribución de la variable aleatoria

$$\frac{(n-1)S_X^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2}$$

es una chi-cuadrado con $n - 1$ grados de libertad y es independiente de $\bar{X}$ y de $\bar{Y}$.

2. La distribución de la variable aleatoria

$$\frac{(m-1)S_Y^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{\sigma^2}$$

es chi-cuadrado con $m - 1$ grados de libertad y es independiente de $\bar{X}$, $\bar{Y}$ y S_X^2 .

3. La distribución de la variable aleatoria

$$U = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

es normal estándar.

4. de (1) y (2) usando la propiedad reproductiva de la distribución chi-cuadrado, la variable aleatoria

$$V = \frac{(n-1)S_X^2}{\sigma^2} + \frac{(m-1)S_Y^2}{\sigma^2}$$

tiene una distribución chi-cuadrado con $n + m - 2$ grados de libertad, y además está distribuída independientemente de $\bar{X} - \bar{Y}$; por lo tanto U y V son independientes. Entonces de acuerdo con la definición de una variable aleatoria T de (3) y (4), la variable aleatoria

$$\frac{U}{\sqrt{\frac{V}{n+m-2}}} = \frac{\dfrac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}}{\sqrt{\dfrac{[(n-1)S_X^2 + (m-1)S_Y^2]/\sigma^2}{n+m-2}}} = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\dfrac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

tiene una distribución t con $n + m - 2$ grados de libertad. Observe que esta variable aleatoria no contiene σ.

7.3.6 DISTRIBUCION - F

En muchas situaciones estaremos interesados en comparar las varianzas de dos variables aleatorias independientes distribuídas normalmente, la que se - describe en 7.3.7. En estos casos recurrimos a la distribución F que se define como sigue:

Sea U y V dos variables aleatorias independientes que tienen distribuciones chi-cuadrado, con r_1 y r_2 grados de libertad, respectivamente. Entonces, la variable aleatoria

$$F = \frac{U/r_1}{V/r_2}$$

tiene una distribución F con r_1 y r_2 grados de libertad. Su función de densidad está dada por

$$f_F(z) = \begin{cases} \dfrac{\Gamma(\frac{r_1 + r_2}{2}) r_1^{r_1/2} r_2^{r_2/2}}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})} \cdot \dfrac{z^{(r_1/2) - 1}}{(r_1 + r_2)^{(r_1 + r_2)/2}} & , \quad 0 < z < \infty \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Observe que la distribución F depende de dos parámetros, r_1 y r_2 en ese orden. El primer parámetro es el número de grados de libertad en el numerador y el segundo es el número de grados de libertad en el denominador.
En la figura 7.3.6. Se da un bosquejo de la función de densidad de la variable aleatoria F para tres pares diferentes de grados de libertad.

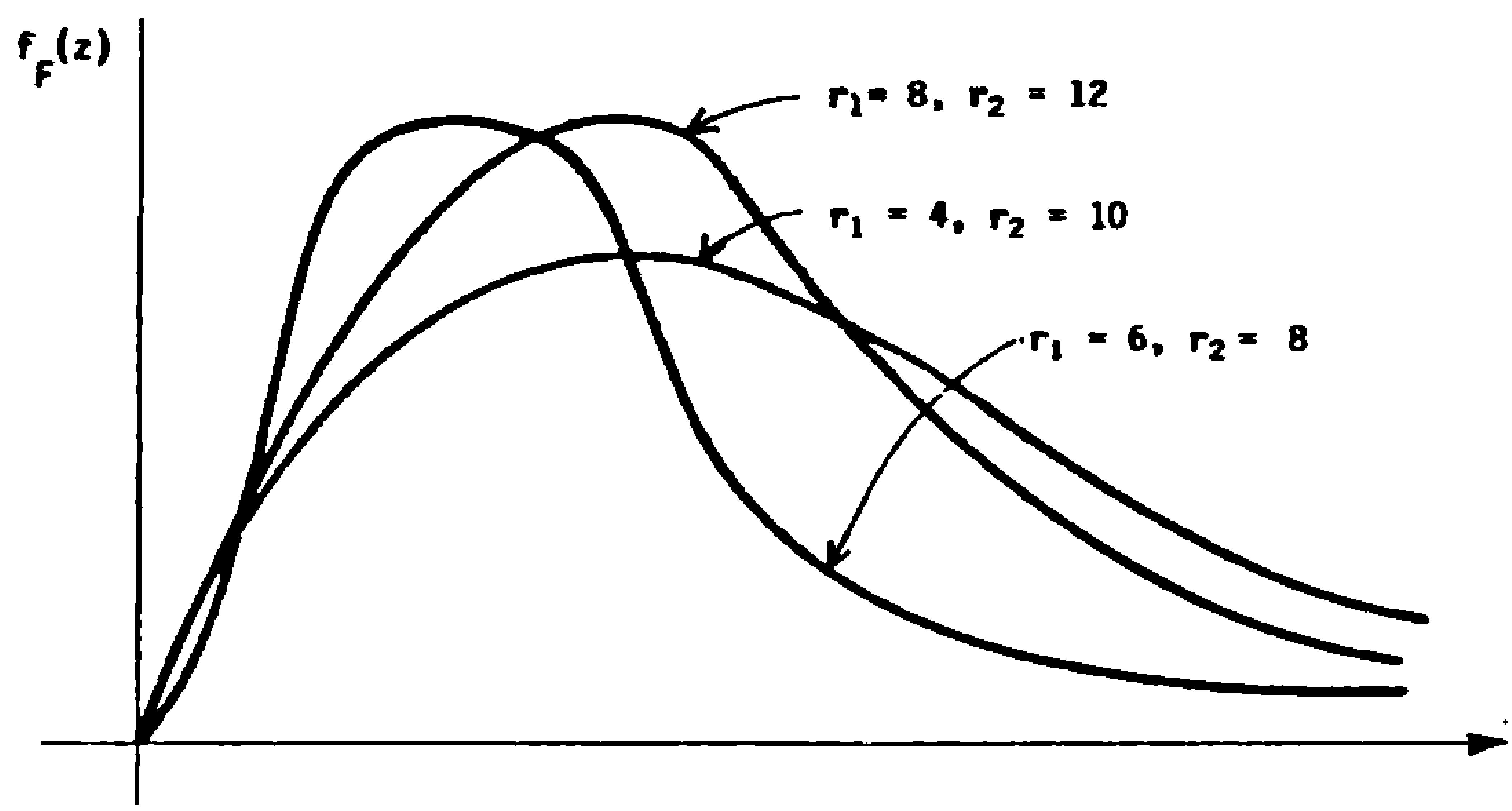


Fig. 7.3.6

Observe que las distribuciones F son una familia de distribuciones continuas, asimétricas hacia la derecha. Existe una distribución F separada para cada par de valores de sus parámetros r_1 y r_2. Por lo tanto se presenta en tablas sumarios de información esencial acerca de las funciones de distribución de la variable aleatoria F.

La media y la varianza de la variable aleatoria F están dados por

$$\mu_F = E(F) = \frac{r_2}{r_2 - 2}, \quad r_2 > 2$$

$$\sigma_F^2 = Var(F) = \frac{2r_2^2(r_1 + r_2 - 2)}{r_1(r_2 - 2)^2(r_2 - 4)}, \quad r_2 > 4$$

La probabilidad que la variable aleatoria F sea menor o igual que una constante f_α está dada por

$$P[F < f_\alpha] = \int_0^{f_\alpha} f_F(z)dz = \alpha$$

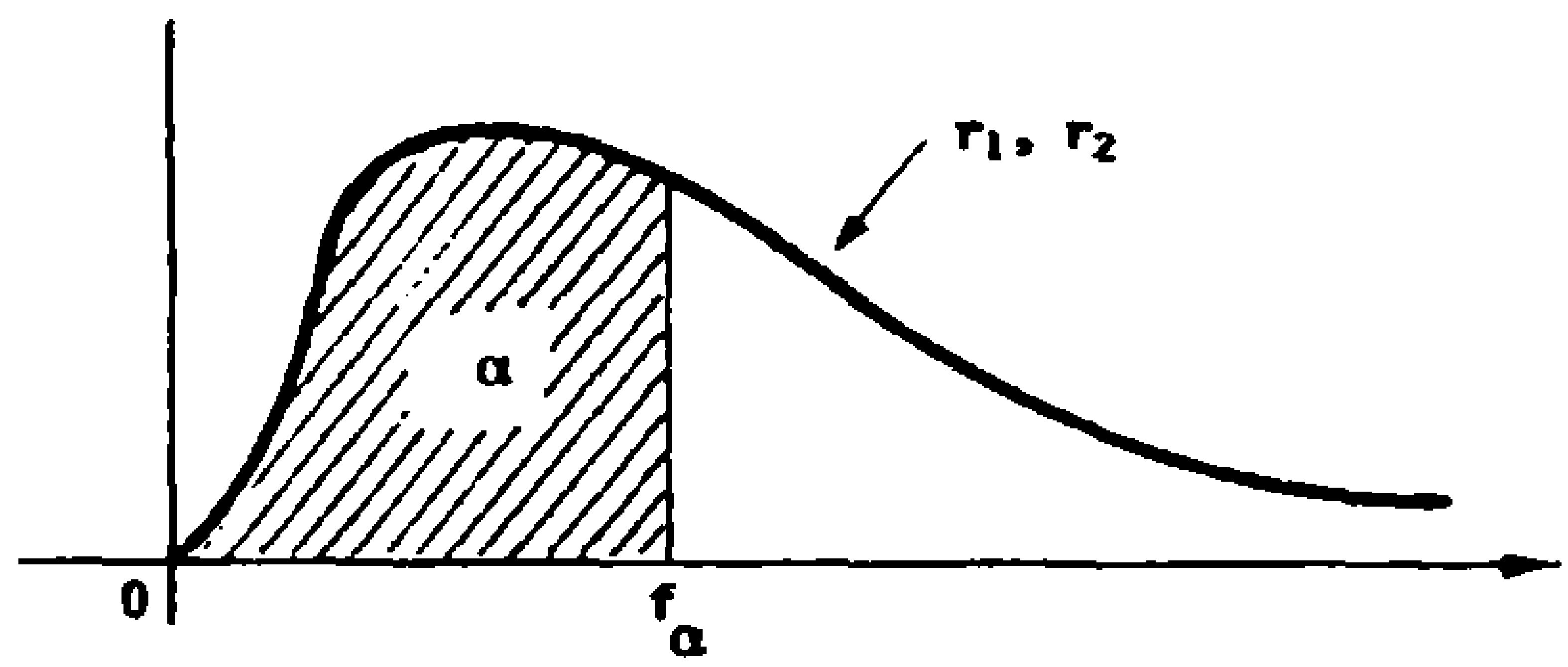


Fig. 7.3.7

Como la distribución F depende de los dos parámetros r_1, y r_2, se necesita una tabla con tres entradas para tabular el valor de F que corresponde a diferentes probabilidades y valores de r_1 y r_2. Una tabla de este tipo se da en la tabla VI, donde se dan puntos porcentuales para valores de $\alpha > 50\%$, o sea el área sombreada en la figura 7.3.7. ($\alpha = 95$, $\alpha = 0.975$, $\alpha = 0.99$) para valores de $\alpha < 50\%$, se obtiene usando la siguiente igualdad

$$\alpha = P\left[\frac{U/r_1}{V/r_2} < f_{\alpha,r_1,r_2}\right] = P\left[\frac{V/r_2}{U/r_1} > \frac{1}{f_{\alpha,r_1,r_2}}\right]$$

$$= 1 - P\left[\frac{V/r_2}{U/r_1} \leq \frac{1}{f_{\alpha,r_1 r_2}}\right]$$

o $$P\left[\frac{V/r_2}{U/r_1} \leq \frac{1}{f_{\alpha,r_1,r_2}}\right] = 1 - \alpha$$

Pero $(V/r_2)/(U/r_1)$ también tiene una distribución F con r_2 y r_1 grados de libertad, es decir

$$P\left[\frac{V/r_2}{U/r_1} \leq f_{1-\alpha,r_2,r_1}\right] = 1 - \alpha$$

Entonces, $$f_{1-\alpha;r_2,r_1} = \frac{1}{f_{\alpha,r_1,r_2}}$$

o $$f_{\alpha,r_1,r_2} = \frac{1}{f_{1-\alpha,r_2,r_1}}$$

por ejemplo, el punto porcentual para $\alpha = 0.05$, $r_1 = 6$, $r_2 = 8$, $f_{0.05;6,8}$. Se obtiene así,

$$f_{0.05;6,8} = \frac{1}{f_{0.95,8,6}} = \frac{1}{4.15} = 0.241 \text{ .}$$

EJEMPLO 9 Sea F una variable aleatoria que tiene una distribución F con r_1 y r_2 grados de libertad. Hallar

(a) $P[F \leq 14.66]$, con $r_1 = 9$, $r_2 = 4$

(b) $P[F \geq 9.01]$, con $r_1 = 5$, $r_2 = 3$

(c) $P[F \leq 0.244]$, con $r_1 = 6$, $r_2 = 9$

(d) Hallar los valores a y b tales que

$P[a \leq F \leq b] = 0.90$ con, $r_1 = 8$, $r_2 = 6$

SOLUCION Según la tabla.

(a) $P[F \leq 14.66] = P[F \leq f_{0.99,9,4}] = 0.99$

(b) $P[F \geq 9.01] = 1 - P[F \leq 9.01] = 1 - 0.95 = 0.05$

(c) $P[F \leq 0.244] = P[\frac{1}{F} \geq 1/0.244] = 1 - P[\frac{1}{F} \leq 4.10]$

$$= 1 - P[\frac{1}{F} \leq f_{0.95,9,6}] = 1 - 0.95 = 0.05$$

(d) Se observa de la tabla VI, los únicos valores para α que se pueden obtener son, 0.95, 0.05; 0.975, 0.025; 0.99 y 0.01, de los cuales 0.95 y 0.05 da la diferencia 0.90. Por lo tanto,

$$P[F \leq b\,] = 0.95 \quad , \quad \text{donde} \quad b = 4.15$$

y $$P[F \leq a\,] = P[\tfrac{1}{F} \geq \tfrac{1}{a}\,] = 1 - P[\tfrac{1}{F} \leq \tfrac{1}{a}\,] = 0{,}05$$

o $$P[\tfrac{1}{F} \leq \tfrac{1}{a}\,] = 0.95, \quad \text{con } r_1 = 8, \quad r_2 = 6$$

Luego, $$\frac{1}{a} = 3.58, \quad \text{de donde} \quad a = 0{,}279$$

7.3.7 DISTRIBUCION DE LA RAZON DE DOS VARIANZAS MUESTRALES

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de una variable aleatoria X, $N(\mu_X, \sigma_X^2)$. Sea $Y_1, Y_2, \ldots, Y_m$ una muestra aleatoria de tamaño m de una variable aleatoria Y con distribución $N(\mu_Y, \sigma_Y^2)$. Además suponga que X e Y son independientes. Entonces, la variable aleatoria,

$$\frac{(n-1)S_X^2}{\sigma_X^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma_X^2}$$

tiene una distribución chi-cuadrado con $n - 1$ grados de libertad. De modo similar, la variable aleatoria

$$\frac{(m-1)S_Y^2}{\sigma_Y^2} = \frac{\sum_{i=1}^{m}(Y_i - \bar{Y})^2}{\sigma_Y^2}$$

tiene una distribución chi-cuadrado con $m - 1$ grados de libertad. Además, las dos variables aleatorias chi-cuadrado son independientes por que X e Y son independientes. Entonces, de acuerdo con la definición de la variable aleatoria F, la variable aleatoria

$$\frac{\dfrac{(n-1)S_X^2}{\sigma_X^2}\Big/(n-1)}{\dfrac{(m-1)S_Y^2}{\sigma_Y^2}\Big/(m-1} = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$$

tiene una distribución F con $n - 1$ y $m - 1$ grados de libertad.

PROBLEMAS 7.3

1. (a) Hallar $\chi^2_{0.01}$, Si $n = 10$

   (b) Hallar $\chi^2_{0.975}$, Si $n = 29$

   (c) Hallar χ^2_α tal que $P[X < \chi^2_\alpha] = 0.99$ con $n = 4$

2. Si X es $\chi^2(12)$, hallar las constantes a y b tales que

   $P[a < X < b] = 0.90$ y $P[X < a] = 0.05$

3. Si X es una variable aleatoria que tiene distribución chi-cuadrado con 19 grados de libertad, calcular:

   (a) $P[X \leq 20]$; (b) $P[X \geq 15]$; (c) $P[16 \leq X \leq 21]$

4. Suponga que $X_1, X_2, \ldots, X_{10}$ es una muestra aleatoria de una variable aleatoria normal estándar. Calcular

$$P[2.56 < \sum_{i=1}^{10} X_i^2 < 18.3]$$

5. En el problema 4. Calcular $P[S^2 < 1.88]$, donde S^2 es la varianza de la muestra.

6. Si X_1 tiene una distribución chi-cuadrado con 7 grados de libertad y X_2 tiene una distribución chi-cuadrado con 3 grados de libertad, independiente de X_1, calcular la probabilidad que $X_1 + X_2$ exceda a 12.

7. ¿Cuántas observaciones son necesarias para asegurar que $P[0.618 \leq S^2/\sigma^2 \leq 1.60] \geq 0.95$? Donde S^2 es la varianza de la muestra, tomada de una variable aleatoria distribuída normalmente con media μ y varianza σ^2.

8. La temperatura de encendido de un interruptor controlado termostáticamente se distribuye normalmente con media y varianza desconocidas. Se va a tomar una muestra aleatoria y a determinar la varianza muestral.
   ¿Cuántas observaciones son necesarias para asegurar que
   $P[S^2/\sigma^2 \leq 1.83] \geq 0.99$?

9. (a) Hallar $t_{0.99}$, si $n = 10$

   (b) Hallar t_α tal que $P[-t_\alpha < T < t_\alpha] = 0.90$, cuando $n = 23$.

10. Sea T una variable aleatoria que tiene una distribución t con n grados de libertad. Calcular,

(a) $P[|T| > 2{,}228]$, cuando $n = 10$

(b) $P[-1.753 \leq T \leq 2.602]$ cuando $n = 15$

11. Si T tiene una distribución t con 25 grados de libertad. Hallar :

(a) $P[T \geq \frac{10}{9}]$, (b) $P[-\frac{20}{9} \leq T]$, (c) $P[-2.787 \leq T \leq 2.787]$

12. Hallar $P[-1.383 \leq (\bar{X} - \mu)\sqrt{10}/S]$ donde $\bar{X}$ y S están basadas en 10 observaciones.

13. Si $\bar{X}$ y S^2 son la media y la varianza de una muestra aleatoria de tamaño 17 de una distribución $N(\mu,\sigma^2)$. Hallar la constante c tal que

$$P[-c < \frac{4(\bar{X} - \mu)}{S} < c] = 0.95$$

14. Si la variable aleatoria F tiene una distribución F, con n_1 y n_2 grados de libertad, respectivamente. Calcular :

(a) $f_{0.05;9,7}$; (b) $f_{0.95,7,9}$

(c) $P[F \geq 4.76]$ con $n_1 = 3$, $n_2 = 6$

(d) $P[F \leq 3.50]$ con $n_1 = 7$, $n_2 = 8$

(e) Hallar los números a y b tal que :

$$P[a < F < b] = 0.98, \text{ con } n_1 = 8, \; n_2 = 6$$

15. Sea $X_1, X_2, \ldots, X_7$ e $Y_1, Y_2, \ldots, Y_9$ muestras aleatorias independientes de distribuciones normales ambas son medias cero y varianzas $\frac{1}{4}$ y $\frac{1}{3}$ respectivamente .

(a) Hallar $P[4 \sum_{i=1}^{7} X_i^2 > 6 \sum_{i=1}^{9} Y_j^2]$

(b) Determine el valor de la constante a tal que

$$P[a\bar{X} > S_Y] = 0.05$$

donde $\bar{X}$ es la media muestral de las X_i, $i = 1, \ldots, 7$ y S_Y es la desviación estándar muestral de las Y_j, $j = 1, \ldots, 9$.

16. Sean $X_1, X_2, \ldots, X_n, X_{n+1}$, $n + 1$ variables aleatorias independientes de una distribución normal $N(\mu,\sigma^2)$.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

encontrar la constante c tal que, el estadístico

$c\left[\dfrac{\bar{X} - X_{n+1}}{S}\right]$ tenga una distribución t .

17. Sea $X_1, X_2, \ldots, X_9$ una muestra aleatoria de tamaño 9 de una población - $N(54,10)$ y sea Y_1, Y_2, Y_3, Y_4 una muestra aleatoria de tamaño 4 de una población $N(54,12)$. Calcular

$$P\left[0.546 \leq \frac{\sum_{i=1}^{9}(X_i - \bar{X})^2}{\sum_{i=1}^{4}(Y_i - \bar{Y})^2} \leq 61.09\right]$$

18. La afirmación que la varianza de una población normal es $\sigma^2 = 21.3$ se rechaza si la varianza de una muestra aleatoria de tamaño 15 excede a 39.74 . ¿Cuál es la probabilidad que la afirmación sea rechazada a pesar que - $\sigma^2 = 21.3$? .

19. Si muestras aleatorias independientes de tamaños $n_1 = n_2 = 8$ provienen de poblaciones normales con la misma varianza, ¿cuál es la probabilidad que la varianza muestral de alguna de las dos muestras sea al menos siete veces - mas grande que la otra?

20. Se halla que la duración de transistores fabricados por una compañía tienen una media de 2000 horas y una desviación típica de 60 horas. Se seleccionan 10 transistores al azar, determinar la probabilidad que la desviación típica muestral
(a) no exceda a 50 horas, (b) se encuentre entre 50 y 70 horas.

21. Dos compañías A y B fabrican transistores. La duración para los fabricados por A tiene una desviación típica de 40 horas en tanto que la duración para los de B tiene una desviación de 50 horas. Se toma una muestra de 8 transistores de A y 16 de B. Hallar la probabilidad que la varianza de la primera muestra sea mayor que
(a) dos veces, (b) 1.2 veces, que la segunda.

8

ESTIMACION

8.1 INTRODUCCION

Los números que aparecen en las distribuciones de probabilidad de las variables aleatorias tales como: p en la distribución binomial, λ en la distribución de Poisson, μ y σ en la distribución normal etc. se llaman *parámetros* o características de la población.

La *estimación* estadística es el proceso mediante el cual se aproxima el valor del parámetro de la población a partir de la información de una muestra. Daremos aquí brevemente los métodos para hacer estimaciones de los parámetros desconocidos de la población a partir de una muestra.

La estimación de un parámetro puede adoptar la forma de un sólo *punto*, es decir la estimación de *un sólo valor* del parámetro de interés; o de un *intervalo* es decir la estimación de un *rango de valores* dentro del cual se espera el valor del parámetro. La primera se llama *estimación puntual y la segunda estimación por intervalo.*

8.2 ESTIMACION PUNTUAL

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una variable aleatoria X y sean $x_1, x_2, \ldots, x_n$ los valores que toma la muestra. Hemos visto que toda función $Y = G(X_1, X_2, \ldots, X_n)$ de estas variables aleatorias que no depende de parámetros desconocidos, se llama un *estadístico*, y su valor está dado por $y = G(x_1, x_2, \ldots, x_n)$.

DEFINICION 8.2.1 ESTIMADOR Cualquier estadístico $Y = G(X_1, X_2, \ldots, X_n)$ cuyo valor $y = G(x_1, x_2, \ldots, x_n)$ se utiliza para estimar el parámetro θ, se dice que es un *estimador* de θ y se escribe así,

$$\hat{\theta} = G(X_1, X_2, \ldots, X_n)$$

La *estimación puntual* de un parámetro desconocido θ de la población consiste en elegir un estimador $\hat{\theta}$ de θ, o sea una función de la muestra $\hat{\theta} = G(X_1, X_2, \ldots, X_n)$, cuyo valor numérico particular $\hat{\theta} = G(x_1, x_2, \ldots, x_n)$ se toma como una aproximación de θ; este valor se llama, *estimado puntual* de θ. Por ejemplo, el valor $\bar{x}$ del estadístico $\bar{X}$, calculado de una muestra de tamaño n, es un estimado puntual del parámetro μ de la población. Es decir, $\bar{X}$ es un estimador de μ y su valor $\bar{x}$ es un estimado de μ. Similarmente $\bar{P} = \frac{X}{n}$ es un estimador de la proporción p y $\hat{p} = \frac{x}{n}$ es un estimado puntual del valor de la proporción p de una distribución binomial.

OBSERVACION Es importante observar que un estimador (en particular el estimador puntual) es una función de n variables aleatorias independientes observables (valores de la muestra). Las estimaciones obtenidas de tal función variarán de una muestra a otra. Por lo tanto, cada estimador tiene su propia distribución,

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = f(x_1)\, f(x_2) \ldots f(x_n)$$

Observe también que puede haber varios posibles estimadores diferentes para un parámetro. Por ejemplo, si queremos estimar la media de una variable aleatoria, se puede considerar, la media muestral o la mediana muestral o la observacion más pequeña en la muestra como estimadores puntuales.

Considerando lo dicho en el párrafo anterior, para estimar θ debe escogerse una función de la muestra que dé el *mejor estimador* de θ. Intuitivamente, parece obvio que la distribución de un *buen* estimador debería concentrarse lo más cerca posible del verdadero valor del parámetro de la población. Supongamos por ejemplo, que θ es el valor verdadero de un parámetro poblacional y que $\hat{\theta}_1, \hat{\theta}_2$ y $\hat{\theta}_3$ son diferentes estimadores de θ con funciones de densidad de probabilidad que se ilustran en la fig. 8.1.1. El estimador $\hat{\theta}_1$ se consideraría como el mejor que $\hat{\theta}_2$ ó $\hat{\theta}_3$, porque su distribución se concentra cerca del verdadero valor de θ.

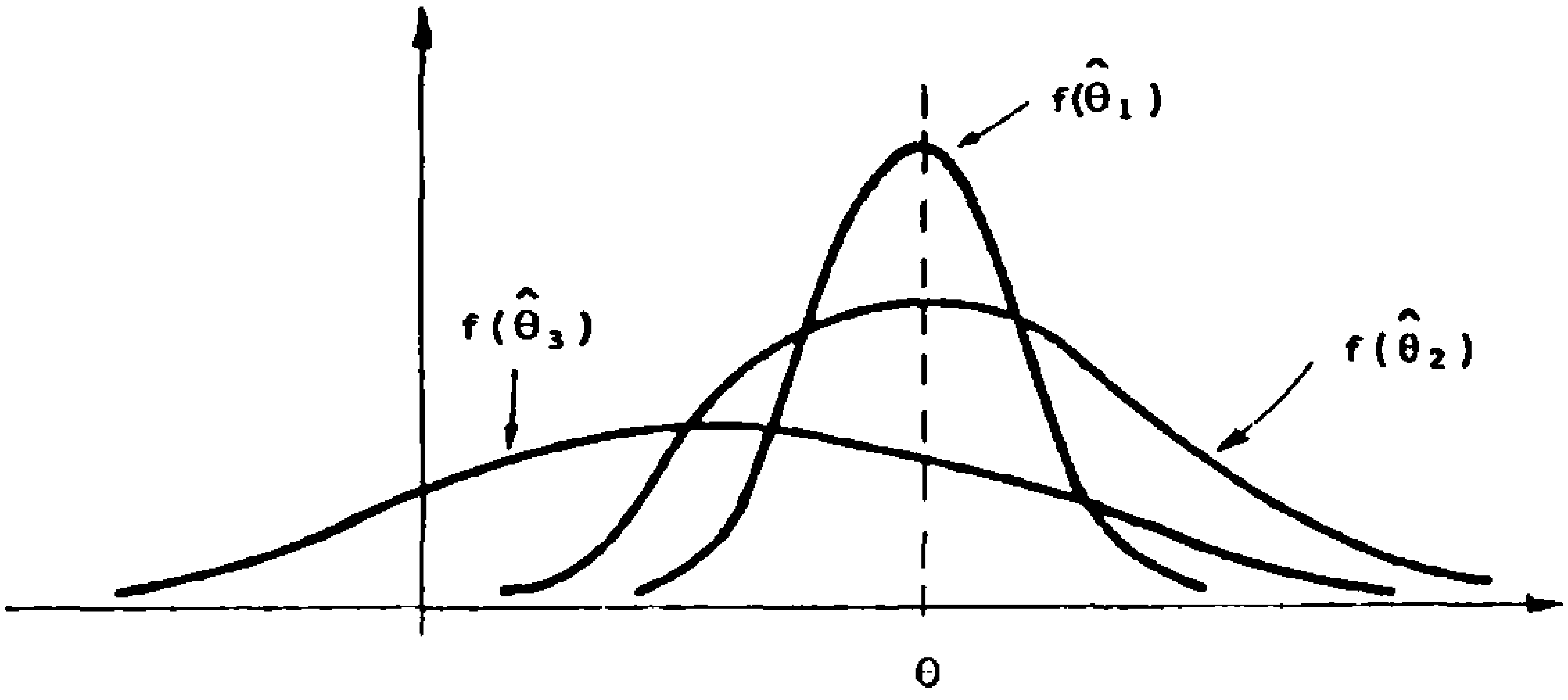


Fig. 8.1.1

Para decidir, qué estimador puntual de un parámetro particular θ es mejor, se necesita estudiar sus propiedades estadísticas y desarrollar algún criterio para comparar estimadores.

8.2.1 PROPIEDADES DE UN ESTIMADOR

INSESGADO Un estimador $\hat{\theta} = G(X_1, X_2, \ldots, X_n)$ se dice que es *insesgado* para el parámetro θ, si $E(\hat{\theta}) = \theta$.

si no se cumple la definición anterior, el estimador se dice que es *sesgado.*

EJEMPLO 1 Suponga que X es una variable aleatoria con media μ y varianza σ^2. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de X. Demostrar que la media muestral $\bar{X}$ y la varianza muestral S^2 son estimadores insesgados de μ y σ^2 respectivamente.

SOLUCION (a) Consideremos

$$E(\bar{X}) = E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=1}^{n} E(X_i)$$

y desde que $E(X_i) = \mu$, para todo $i = 1,2,\ldots,n$, se tiene

$$E(\bar{X}) = \frac{1}{n} \sum_{i=1}^{n} \mu = \mu$$

Por lo tanto, la media muestral $\bar{X}$ es un estimador insesgado de la media poblacional μ.

NOTA El teorema de CHEBYSHEV permite acotar la probabilidad del error que se comete al tomar $\bar{X}_n$ como el valor de μ

$$P[|\bar{X}_n - \mu| \geq k\,\sigma_{\bar{X}}] \leq \frac{1}{k^2}\ ,\ k > 1$$

Si $\quad k\sigma_{\bar{X}} = \varepsilon \quad$ y como $\quad \sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$, se tiene

$$P[|\bar{X}_n - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{n\varepsilon^2}$$

Cuando $n \to \infty$, $\quad P[|\bar{X}_n - \mu| \geq \varepsilon] \to 0$ *

Por lo tanto, el estimador $\bar{X}$ de μ es tanto mejor cuanto mayor sea la muestra.

(b) Consideremos ahora,

$$E(S^2) = E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}\right] = \frac{1}{n-1} E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right]$$

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n}(X_i^2 + \bar{X}^2 - 2\bar{X}X_i)\right] = \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 + n\bar{X}^2 - 2n\bar{X}^2\right]$$

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)\right]$$

Desde que $\quad E(X_i^2) = \mu^2 + \sigma^2 \quad$ y $\quad E(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$, se tiene

$$E(S^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}(\mu^2 + \sigma^2) - n(\mu^2 + \frac{\sigma^2}{n})\right] = \frac{1}{n-1}(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2)$$

$$= \sigma^2 .$$

Es decir, la varianza muestral S^2 es un estimador insesgado de la varianza pobla-

* En cursos avanzados, se dice que $\bar{X}_n$ converge en probabilidad a μ

cional σ^2.

NOTA Esta es otra razón por la cuál en muchos textos se define la varianza muestral como $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/(n - 1)$ en vez de $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/n$

CONSISTENCIA Si $\hat{\theta}_n = G(X_1, X_2, \ldots, X_n)$ es una secuencia de estimadores del parametro θ basado en una muestra aleatoria de tamaño n, se dice que $\{\hat{\theta}_n\}$ es *consistente* para θ, si

$$\lim_{n \to \infty} P[|\hat{\theta}_n - \theta| < \varepsilon] = 1$$

EJEMPLO 2 Demostrar que la media muestral $\bar{X}_n$, es un estimador consistente de μ.

SOLUCION

$$P[|\bar{X}_n - \mu| < k\, \sigma_{\bar{X}}] \geq 1 - \frac{1}{k^2}$$

Tomando $\varepsilon = k\sigma_{\bar{X}}$, entonces $k = \frac{\varepsilon}{\sigma_{\bar{X}}} = \frac{\varepsilon\sqrt{n}}{\sigma}$ se tiene

$$P[|\bar{X}_n - \mu| < \varepsilon] \geq 1 - \frac{\sigma^2}{n\varepsilon^2}$$

cuando $n \to \infty$, $P[|\bar{X}_n - \mu| < \varepsilon] \to 1$

ERROR CUADRATICO MEDIO DE UN ESTIMADOR Sea $\hat{\theta} = G(X_1, X_2, \ldots, X_n)$ un estimador de θ. *El error cuadrático medio* del estimador $\hat{\theta}$ se define por $E(\hat{\theta} - \theta)^2$.

NOTACION El error cuadrático medio del estimador $\hat{\theta}$ se denota por $MSE(\hat{\theta})$. Es decir, $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$

OBSERVACION El error cuadrático medio de $\hat{\theta}$ se escribe así

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + [\theta - E(\hat{\theta})]^2$$

De modo que, si $\hat{\theta}$ es un estimador insesgado de θ, entonces $MSE(\hat{\theta}) = Var(\hat{\theta})$

En efecto:

$$\begin{aligned} MSE(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 = E[(\hat{\theta} - E(\hat{\theta})) - (\theta - E(\hat{\theta}))]^2 \\ &= E[(\hat{\theta} - E(\hat{\theta}))^2] - 2(\theta - E(\hat{\theta}))\, E[\hat{\theta} - E(\hat{\theta})] + E[(\theta - E(\hat{\theta}))^2] \\ &= Var(\hat{\theta}) + (\theta - E(\hat{\theta}))^2 \end{aligned}$$

El término $\theta - E(\hat{\theta})$ se llama *sesgo* del estimador $\hat{\theta}$ y puede ser positivo, negativo o cero. La observación muestra que el error cuadrático medio es la suma de dos cantidades no negativas.

EFICIENCIA RELATIVA El error cuadrático medio es un criterio importante para comparar dos estimadores. Sean $\hat{\theta}_1$ y $\hat{\theta}_2$ dos estimadores del parámetro θ y sean $MSE(\hat{\theta}_1)$ y $MSE(\hat{\theta}_2)$ los errores cuadráticos medios de $\hat{\theta}_1$ y $\hat{\theta}_2$ respectivamente. La *eficiencia relativa* de $\hat{\theta}_2$ con respecto a $\hat{\theta}_1$ se define como el cociente

$$\frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)}$$

si esta eficiencia relativa es menor que 1 concluimos que $\hat{\theta}_1$ es un estimador más *eficiente* que $\hat{\theta}_2$, en el sentido que tiene menor error cuadrático medio.

EJEMPLO 3 Suponga que estamos interesados en estimar la media μ de una población con varianza σ^2.

En efecto : 1 Tomemos $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de la población.

2 Queremos comparar dos posibles estimadores para μ: la media muestral $\bar{X}$ y una simple observación de la muestra, digamos X_i. Note que $\bar{X}$ y X_i son estimadores insesgados de μ; consecuentemente, el error cuadrático medio de ambos es simplemente la varianza,

Para la media muestral $\quad MSE(\bar{X}) = Var(\bar{X}) = \sigma^2/n$

Para la observación individual $\quad MSE(X_i) = Var(X_i) = \sigma^2$

Luego, la eficiencia relativa de X_i con relación a $\bar{X}$ es

$$\frac{MSE(\bar{X})}{MSE(X_i)} = \frac{\sigma^2/n}{\sigma^2} = \frac{1}{n}$$

Desde que $1/n < 1$ para muestras de tamaño $n > 2$, concluimos que la media muestral es mejor estimador de μ que una simple observación X_i.

EFICIENCIA Considerando todos los posibles estimadores insesgados de un parámetro θ; el que tiene varianza mínima se llama estimador *eficiente* de θ o estimador de *varianza mínima* de θ.

NOTA 1 Hallar un estimador eficiente para el parámetro θ, es encontrar entre los estimadores insesgados, el estimador que tiene varianza más pequeña; tal estimador se llama tambien *ESTIMADOR INSESGADO DE MINIMA VARIANZA*.

2 Es posible obtener una cota inferior de las varianzas de todos los estimadores insesgados de θ. Sea $\hat{\theta}$ un estimador insesgado del parámetro θ , basado en una muestra aleatoria de n observaciones y sea $f(x;\theta)$ la distribución de probabilidad de la variable aleatoria X. Entonces, una cota inferior de la varianza de $\hat{\theta}$ es

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{nE[\frac{l}{d\theta}\ ln f(x;\theta)]^2}$$

esta desigualdad se llama la cota inferior de RAO-CRAMER **

3 Si la varianza de un estimador insesgado $\hat{\theta}$ satisface la desigualdad de **RAO-CRAMER** como una igualdad, este es un estimador insesgado de mínima varianza o eficiente de θ.

EJEMPLO 4 Demostrar que la media muestral $\bar{X}$ es un estimador insesgado de mínima varianza de la media μ de una variable aleatoria X con distribución normal de varianza conocida σ^2.

SOLUCION 1 $\bar{X}$ es un estimador insesgado de la media μ.

2 La función de densidad de la variable aleatoria X es

$$f(X;\mu) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}(\frac{X-\mu}{\sigma})^2} \quad , \qquad -\infty < X < \infty$$

Tomando logaritmo natural a ambos miembros de la función de densidad

$$ln f(X;\mu) = -\ ln(\sigma\sqrt{2\pi}) - \frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2$$

3 La desigualdad de RAO-CRAMER para la varianza de $\bar{X}$ es

$$\mathrm{Var}(\bar{X}) > \frac{1}{n\,E\{\frac{d}{d\mu}[(-\ ln(\sigma\sqrt{2\pi}) - \frac{1}{2}(\frac{X-\mu}{\sigma})^2]\}^2}$$

$$\geq \frac{1}{n\,E\left[0 - \frac{1}{2} \times 2\left(\frac{X-\mu}{\sigma}\right)\left(-\frac{1}{\sigma}\right)\right]^2}$$

$$\geq \frac{1}{n\,E\left[\frac{X-\mu}{\sigma^2}\right]^2}$$

** El lector interesado puede ver HOGG - CRAIG o BICKEL MATH

$$\geq \frac{1}{n\,\dfrac{E(X-\mu)^2}{\sigma^4}}$$

$$\geq \frac{1}{n\,\dfrac{\sigma^2}{\sigma^4}}$$

$$\geq \frac{\sigma^2}{n}\ .$$

Puesto que $Var(\bar{X}) = \frac{\sigma^2}{n}$, vemos que $Var(\bar{X})$ satisface la cota inferior de RAO-CRAMER como una igualdad. Entonces $\bar{X}$ es un estimador insesgado de mínima varianza o eficiente de μ.

8.2.2 METODOS DE ESTIMACION PUNTUAL

1. METODO DE MAXIMA VEROSIMILITUD

El principio de máxima verosimilitud, en su aplicación a la obtención de estimadores puntuales, consiste en seleccionar aquel estimador que maximice la probabilidad de obtener la muestra realmente observada.

DEFINICION 8.2.2 Sea X una variable aleatoria cuya función de probabilidad, $f(x;\theta)$ depende de un parámetro θ. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de X y sean $x_1, x_2, \ldots, x_n$ los valores observados de la muestra. La *función de verosimilitud* de la muestra se define así,

$$V(\theta) = f(x_1, x_2, \ldots, x_n;\ \theta) = f(x_1;\theta)\,f(x_2;\theta)\ \ldots f(x_n;\theta)$$

$$= \prod_{i=1}^{n} f(x_i;\theta)$$

El método de máxima verosimilitud, consiste en tomar como valor estimado de θ, el valor que hace máxima la función $V(\theta)$.

Observe que si $\hat{\theta}$ hace máxima a $V(\theta)$, entonces también hace máxima a su logaritmo $ln\,V(\theta)$. Para convertir el producto en suma, se suele tomar la función $L = ln V(\theta)$ y determinar θ por la ecuación,

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{n} \frac{\partial\, ln f(x_i\,;\,\theta)}{\partial\,\theta} = 0$$

Si, la distribución de probabilidad tiene varios parámetros desconocidos, $\theta_1, \theta_2, \ldots, \theta_k$, entonces en lugar de *una* ecuación, tendremos k ecuaciones,

$$\frac{\partial L}{\partial \theta_1} = 0 \quad , \quad \frac{\partial L}{\partial \theta_2} = 0, \ldots, \frac{\partial L}{\partial \theta_k} = 0$$

a partir de las cuales se obtiene las estimaciones para estos parámetros.

EJEMPLO 5 Obtener el estimador de máxima verosimilitud de la probabilidad de éxito para una variable aleatoria **X** de Bernoulli.

SOLUCION 1 Desde que **X** es una variable aleatoria de Bernoulli, entonces la función de distribución de **X** está dado por

$$p(x;p) = p^x(1-p)^{1-x}, \quad x = 0,1 \quad ; \quad 0 < p < 1$$

2 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de **X** cuyos valores son $x_1, x_2, \ldots, x_n$ ($x_1, x_2, \ldots, x_n$ es una secuencia de ceros y unos). Entonces

$$p(x_i;p) = p^{x_i}(1-p)^{1-x_i} \quad , \quad i = 1,2,\ldots,n.$$

3 La función de verosimilitud es

$$V(p) = p(x_1,x_2,\ldots,x_n;p) = p(x_1;p)p(x_2;p), \ldots, p(x_n;p)$$

$$= \prod_{i=1}^{n} p(x_i;p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$= p^{\Sigma x_i}(1-p)^{\Sigma(1-x_i)}$$

4 Tomando logaritmo natural a $V(p)$

$$L = \ln V(p) = \sum_{i=1}^{n} x_i \ln p + \left[n - \sum_{i=1}^{n} x_i\right] \ln(1-p)$$

5 Derivando con respecto a p obtenemos

$$\frac{\partial L}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p} + \frac{n - \sum_{i=1}^{n} x_i}{1-p}(-1) = 0$$

ó $$\frac{\sum_{i=1}^{n} x_i}{p} = \frac{n - \sum_{i=1}^{n} x_i}{1 - p}$$

ó $$\frac{1 - p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

de donde $$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

es, simplemente, el número promedio de "éxitos" observados en una serie de n ensayos.

Observe que $$= E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = p$$

Es decir, $\hat{p}$ es un estimador insesgado de p.

EJEMPLO 6 Suponga que el número diario de ventas de autos nuevos que efectúa determinado concesionario, es una variable aleatoria de Poisson con parámetro λ. Dado que en 20 días la venta total de autos fué de 30, ¿cuál es el estimador de máxima verosimilitud de λ?

SOLUCION 1 Comenzaremos determinando, el estimador de máxima verosimilitud del parámetro λ de la distribución de Poisson.

$$p(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad , \quad x = 0,1,2, \ldots$$

2 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de la variable aleatoria de Poisson, con valores $x_1, x_2, \ldots, x_n$. Entonces,

$$p(x_i;\lambda) = P[X = x_i] = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \quad , \quad i = 1,2,\ldots,n.$$

3 La función de verosimilitud es,

$$V(\lambda) = \prod_{i=1}^{n} p(x_i;\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!}$$

4 tomando logaritmo a $V(\lambda)$ obtenemos,

$$L = \ln V(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \ \ln \lambda - \ln(\prod_{i=1}^{n} x_i!)$$

5 Derivando con reespecto a λ, se tiene,

$$\frac{\partial L}{\partial \lambda} = -n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0$$

de donde $\hat{\lambda} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$.

Expresado en palabras, la media muestral es el estimador de máxima verosimilitud del parámetro λ de la distribución de Poisson. Por lo tanto, empleando los datos del problema, $n = 20$ días, $\sum_{i=1}^{20} x_i = 30$ autos, se tiene,

$$\hat{\lambda} = \frac{30}{20} = 1.5 \ .$$

EJEMPLO 7 Suponga que se compra una docena de naranjas de huando, se pesa cada - una, y se encuentra que

$$\sum_{i=1}^{12} x_i = 180 \quad , \qquad \sum_{i=1}^{12} x_i^2 = 2799$$

en (onzas). Calcular los estimados de máxima verosimilitud de μ y σ^2, suponiendo que las 12 naranjas compradas son una muestra aleatoria de todas las naranjas de huando (y que sus pesos estan distribuídos normalmente).

SOLUCION 1 Determinaremos primero el estimador de máxima verosimilitud de μ y de σ^2 de la población normal X, con función de densidad

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad , \quad -\infty < x < \infty$$

2 Sea $X_1, X_2, \ldots, X_n$, una muestra aleatoria de tamaño n, de una variable aleatoria X con distribución normal y $x_1, x_2, \ldots, x_n$ los valores de la muestra. Entonces

$$f(x_i;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/2\sigma^2}$$

3 La función de verosimilitud es,

$$V(\mu,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/2\sigma^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-1/2\,\Sigma(x_i-\mu)^2/\sigma^2}$$

4 Tomando logaritmo se tiene

$$L = ln\ V(\mu,\sigma^2) = -\frac{n}{2} ln\ 2\pi - \frac{n}{2} ln\ \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

5 Derivando con respecto a μ y σ^2 , obtenemos

$$\frac{\partial L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)$$

y

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

6 Resolviendo el sistema de ecuaciones para las estimadas

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

y

$$-\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

Se obtiene

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

por lo tanto, los estimadores de máxima verosimilitud, de μ y de σ^2 son $\hat{\mu} = \bar{x}$

y $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{n-1}{n}S^2$$

Para nuestro problema tenemos

$$\hat{\mu} = \frac{1}{12}\sum_{i=1}^{12} x_i = \frac{180}{12} = 15.$$

$$\sigma^2 = M(X^2) - [M(X)]^2 = \frac{2799}{12} - 225 = 8.25$$

2 METODO DE LOS MOMENTOS

1 Suponga que X es ó una variable aleatoria continua con función de densidad - $f(x;\theta_1,\theta_2,\ldots,\theta_k)$ o una variable aleatoria discreta con distribución de probabilidad $p(x;\theta_1,\theta_2,\ldots,\theta_k)$ caracterizado por k parámetros desconocidos.

2 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de X. Los primeros k ***momentos muestrales*** alrededor del origen se define como

$$M'_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \quad , \qquad r = 1,2,\ldots,k$$

3 Los primeros k momentos de la población alrededor del origen son respectivamente

$$\mu'_r = E(X^r) = \int_{R_X} x^r f(x;\theta_1,\theta_2,\ldots,\theta_k)\,dx \quad , \quad r = 1,2,\ldots,k \;;$$

si X es continua,

$$\mu'_r = E(X^r) = \sum_{R_X} x^r p(x;\theta_1,\theta_2,\ldots,\theta_k) \;, \qquad r = 1,2,\ldots,k \;;$$

si X es discreta.

4 Los momentos μ'_r ($r = 1,2,\ldots,k$) de la población son en general función de los k parámetros desconocidos $\theta_1,\theta_2,\ldots,\theta_k$. El método de los momentos consiste en - igualar los momentos muestrales y momentos poblacionales, obteniendo k ecuaciones simultaneos con k, parámetros desconocidos $\theta_1,\theta_2,\ldots,\theta_k$. Esto es

$$\mu'_r = M'_r \quad , \qquad r = 1,2,\ldots,k$$

5 Las soluciones de las ecuaciones en (4) denotados por $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ constituyen los momentos estimados de $\theta_1, \theta_2, \ldots, \theta_k$ respectivamente.

EJEMPLO 8 Suponga que X es una variable aleatoria distribuida normalmente con media μ y varianza σ^2 desconocidos. Encuentre los estimadores para μ y σ^2 por el método de los momentos.

SOLUCION 1 La distribución normal tiene dos parámetros μ y σ^2, en concecuencia habrá dos ecuaciones. Los dos primeros momentos poblacionales para una población normal son

$$\mu'_1 = E(X) = \mu$$

$$\mu'_2 = E(X^2) = \sigma^2 + \mu^2$$

2 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de X. Los dos primeros momentos muestrales son

$$M'_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$M'_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

3 De las ecuaciones de los pasos (1) y (2) se obtiene el sistema

$$\mu = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

y $$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \quad \text{ó} \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \mu^2$$

4 El sistema de ecuaciones de (3) dan las soluciones

$$\hat{\mu} = \bar{X} .$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - 2\bar{X}\sum_{i=1}^{n} X_i + \sum_{i=1}^{n}\bar{X}^2\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2\bar{X}X_i + \bar{X}^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$= \frac{n-1}{n}S^2.$$

Es decir, el estimador por el método de los momentos de μ es $M_1' = \bar{X}$ y el estimador por el método de los momentos de σ^2 es $M_2' - \bar{X}^2 = \frac{n-1}{n}S^2$.

EJEMPLO 9 Sea $X_1, X_2, \ldots, X_n$ una muestra de una población con distribución de Poisson con parámetro λ . Estime λ por el método de los momentos.

SOLUCION 1 La distribución de Poisson tiene solamente un parámetro, en consecuencia se tendrá sólo una ecuación.

2 El primer momento poblacional es $\mu_1' = E(X) = \lambda$

3 El primer momento muestral es $M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

4 De los pasos (2) y (3) se tiene la ecuación

$$\mu_1' = \lambda = M_1' = \bar{X}$$

En concecuencia $\hat{\lambda} = \bar{X}$.

EJEMPLO 10 Sea X una variable aleatoria distribuida uniformemente en el intervalo $[0, a]$. Hallar un estimador de a por el método de los momentos.

SOLUCION 1 En este caso la distribución uniforme tiene solo un parámetro, entonces se tendra una ecuación. El primer momentos poblacional alrededor del origen es

$$\mu_1' = E(X) = \int_0^a \frac{x}{a}\,dx = \left.\frac{x^2}{2a}\right|_0^a = \frac{a}{2}$$

2 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de X, el primer momento muestral es

$$M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

3 De los pasos (1) y (2) se tiene

$$\frac{a}{2} = \overline{X}$$

Es decir, el estimador de momentos de a es $\hat{a} = 2\overline{X}$

EJEMPLO 11 Sea X una variable aleatoria distribuida uniformemente en el intervalo $[a, b]$, con a y b desconocidos. Estimar a y b por el método de los momentos.

SOLUCION 1 La distribución uniforme tiene dos parámetros a y b ,luego se tendrá dos ecuaciones. Los dos primeros momentos poblacionales son

$$\mu'_1 = E(X) = \int_a^b \frac{x}{b-a}\,dx = \frac{a+b}{2}$$

$$\mu'_2 = E(X^2) = \int_a^b \frac{x^2}{b-a}\,dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}\,.$$

2 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de tamaño n de X, los dos primeros momentos muestrales son

$$M'_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$$M'_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

3 De los pasos (1) y (2) se tiene el sistema de ecuaciones

$$\frac{a+b}{2} = \overline{X}$$

$$\frac{a^2+ab+b^2}{3} = \frac{1}{n}\sum_{i=1}^{n} X^2 = M'_2$$

4 De la primera ecuación de (3) se despeja $a = 2\overline{X} - b$ y reemplazando en la otra ecuación se obtiene las soluciones

$$(2\overline{X} - b)^2 + (2\overline{X} - b)\,b + b^2 = 3M'_2$$

$$b^2 - 2\overline{X}b + 4\overline{X}^2 - 3M'_2 = 0$$

$$b = \frac{2\overline{X} \pm \sqrt{4\overline{X}^2 - 4\,[4\overline{X}^2 - 3M'_2]}}{2}$$

$$b = \bar{X} + \sqrt{3(M_2' - \bar{X}^2)}$$

Por lo tanto, los estimadores de los momentos para a y b son

$$\hat{a} = \bar{X} - \sqrt{3(M_2' - \bar{X}^2)}$$

$$\hat{b} = \bar{X} + \sqrt{3(M_2' - \bar{X}^2)}$$

PROBLEMAS 8.2

1 La primera observación de una muestra aleatoria podría utilizarse como un estimador de la media poblacional. ¿Es este un estimador:
(a) insesgado? (b) consistente? (c) eficiente? .

2 Suponga que $\hat{\theta}$ es un estimador de θ basado en una muestra aleatoria de tamaño n. Si $E(X_i) = \theta$ y $\hat{\theta} = \Sigma a_i X_i$, ¿qué restricciones tiene que cumplir a_i para que $\hat{\theta}$ sea un estimador insesgado de θ?

3 Sea X una variable aleatoria con media μ y varianza σ^2. Suponga que

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Se usa como estimador de σ^2. Demostrar que es un estimador sesgado.

4 Suponga que tiene una muestra de tamaño $2n$ de una población X con $E(X) = \mu$ y $Var(X) = \sigma^2$. Sean

$$\bar{X}_1 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \qquad y \qquad \bar{X}_2 = \frac{1}{n} \sum_{i=1}^{n} X_i$$

dos estimadores de μ , ¿cuál es el mejor estimador de μ ? justifique su respuesta.

5 Sea $X_1, X_2, \ldots, X_7$ una muestra aleatoria de una población que tiene una media μ y varianza σ^2. Considere los siguientes estimadores de μ

$$\hat{\theta}_1 = \frac{X_1 + X_2 + \ldots + X_7}{7} \quad ; \quad \hat{\theta}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

¿son estimadores insesgados? ¿cuál es el mejor estimador? ¿en que sentido es mejor?

6 Suponga que $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores de θ con
$E(\hat{\theta}_1) = \theta$, $E(\hat{\theta}_2) = \theta/2$, $Var(\hat{\theta}_1) = 6$, $Var(\hat{\theta}_2) = 2$. ¿Cúal es mejor estimador de θ ? ¿por qué?

7 Sea X una variable aleatoria con media μ y varianza σ^2. Dada dos muestras aleatorias de tamaños n_1 y n_2 con medias muestrales $\bar{X}_1$ y $\bar{X}_2$ respectivamente, demostrar que

$$\bar{X} = a\bar{X}_1 + (1-a)\bar{X}_2 \quad , \quad 0 < a < 1$$

es un estimador insesgado de μ. Asumiendo que $\bar{X}_1$ y $\bar{X}_2$ son independientes, hallar el valor de a que minimiza la varianza de $\bar{X}$.

8 Se realizan 15 pruebas independientes de Bernoulli y se observan 12 éxitos, si no se conoce p. Hallar el estimador de máxima verosimilitud de p.

9 Se instalan 30 tubos electrónicos escogidos aleatoriamente y se registra el tiempo de falla de cada uno de ellos. Si se supone que los tiempos registrados son tales que $\Sigma x_i = 32,916$ horas. ¿Cuál es el estimado de máxima verosimilidad para el parámetro de la distribución exponencial de las duraciones de los tubos?

10 Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores independientes insesgados de un parámetro desconocido θ, con varianzas conocidas σ_1^2 y σ_2^2 respectivamente:

a) Demostrar que $\hat{\theta} = (1-a)\hat{\theta}_1 + a\hat{\theta}_2$ también es un estimador insesgado de θ, para cualquier valor de a.

b) Encontrar el valor de a que minimiza la varianza de $\hat{\theta}$.

11 Basado en su experiencia, un vendedor de televisores a colores piensa que el número de ventas de televisores a colores que hace cada día laborable es una variable aleatoria de Poisson con parámetro λ. Examina sus registros del año anterior (que tuvo 310 días habiles) y encuentra que vendió en total 279 televisores. Calcular el estimado de máxima verosimilitud de la probabilidad que no venda televisores en su próximo día de trabajo.

12 Sea X una variable aleatoria uniformemente distribuída en el intervalo $[a - \frac{1}{2}, a + \frac{1}{2}]$. ¿Cuál es el estimador de máxima verosimilitud para a en base a una muestra aleatoria de tamaño n de X?

13 Sea X una variable aleatoria binomial con parámetro p y n (n conocida). Determinar el estimador de máxima verosimilitud para p en base a una muestra aleatoria de m observaciones de X.

14 Sea X una variable aleatoria geométrica con parámetro p. Hallar el estimador de máxima verosimilitud de p, basado en una muestra aleatoria de tamaño n.

15 Sea X una variable aleatoria binomial con parámetros n (desconocido) y p. Determinar el estimador de máxima verosimilitud de p, en base a una muestra aleatoria de tamaño m.

16 Hallar el estimador del parámetro α de la distribución exponencial por el método de los momentos, en base a una muestra aleatoria de tamaño n.

17 Sea X una variable aleatoria geométrica con parámetro p. Hallar el estimador de p por el método de los momentos, basado en una muestra aleatoria de tamaño n.

18 Sea X una variable aleatoria de Bernoulli con parámetro p. Determinar el estimador de p por el método de los momentos, en base a una muestra aleatoria de tamaño n.

19 Sea X una variable aleatoria binomial con parámetros n (conocida) y p. Hallar el estimador de p, por el método de los momentos, basado en una muestra aleatoria de tamaño m.

20 Sea X una variable aleatoria binomial con parámetros n y p, ambos desconocidos. Calcular el estimador de n y p por el método de los momentos, en base a una muestra aleatoria de tamaño m.

8.3 ESTIMACION DE INTERVALOS DE CONFIANZA

En la sección anterior, hemos visto la *estimación puntual* del parámetro de la población, la cual es un valor numérico que se obtiene a partir de alguna fórmula conocida mediante el uso de una muestra de la población.

En las aplicaciones, cualquier estimador puntual del parámetro desconocido de la población, aún cuando tenga todas las propiedades deseadas, tales como, ser: insesgado, eficiente, etc; tiene la limitación que no proporciona información acerca de la precisión de la estimación obtenida. Es decir, de la *magnitud del error debido al muestreo.* Por ejemplo, la estimación puntual del tiempo promedio de servicio de los 10,000 empleados de "SIDER-PERU", $\bar{x}$ = 8.0 años basados en una muestra, no da una indicación de la precisión de la estimación.

En el capítulo anterior, hemos visto que la media basada en una muestra aleatoria puede tomar muchos valores diferentes, indicados por la distribución muestral de $\bar{X}$. Además estos valores posibles de $\bar{X}$ están agrupados en la distribución muestral de $\bar{X}$ alrededor de la media μ de la población ($E(\bar{X}) = \mu$), y la mayoría de las $\bar{x}$ caen a la derecha o a la izquierda de la media de la población. Por lo tanto, es casi seguro que cualquier $\bar{x}$; en otras palabras cualquier estimación puntual de μ no será correcta en el sentido de que proporcione el valor real de la media de la población. Así pues, es bastante seguro concluir que 8.0 años no es realmente el promedio del tiempo de servicio de los 10,000 empleados. Es decir, a partir de una muestra no podemos obtener conclusiones acerca de la población correspondiente que sean 100 % verdaderas. Por esta razón es que la estimación puntual tiene poco uso. A menos que se tenga alguna indicación sobre la precisión de la estimación.

La precisión de la estimación puntual puede evaluarse en una muestra, por estimación de un intervalo junto con una medida de la seguridad que tal intervalo contenga al parámetro desconocido de la población. Dichos intervalos se llaman *intervalos de confianza o estimación por intervalo* del parámetro desconocido. Formalizaremos esto en las definiciones siguientes.

DEFINICION 8.3.2 *Intervalo aleatorio*, es un intervalo en el cual por lo menos uno de sus extremos es una variable aleatoria.

EJEMPLO 1 Sea $2 < X < 6$, un evento, donde X es una variable aleatoria. Pero $2 < X < 6$, se puede escribir así

$$X < 6 \quad y \quad X > 2$$

ó $\quad X < 6 \quad y \quad 3X > 6 \quad$ o sea $\quad X < 6 < 3X$

$X < 6 < 3X$, es un intervalo aleatorio ya que sus extremos son variables aleatorias. Y el evento $2 < X < 6$ es equivalente al evento $X < 6 < 3X$ luego;

$$P[2 < X < 6] = P[X < 6 < 3X]$$

EJEMPLO 2 Sea X una variable aleatoria cuya función de densidad es,

$$f(x) = \begin{cases} \frac{1}{6}x + k & , \quad 0 < x < 3 \\ 0 & , \quad \text{en otros casos} \end{cases}$$

Determinar la probabilidad que el intervalo aleatorio $\langle X, 3X\rangle$ incluya al punto 3

SOLUCION Queda como ejercicio para el lector verificar que $k = \frac{1}{12}$. Debemos calcular $P[X < 3 < 3X]$. La desigualdad $X < 3 < 3X$ se escribe así,

$$X < 3 \quad y \quad 3 < 3X$$

ó $\quad X < 3 \quad y \quad 1 < X$, de donde $\quad 1 < X < 3$.

Es decir, el evento $X < 3 < 3X$ es equivalente al evento $1 < X < 3$, por lo tanto

$$P[X < 3 < 3X] = P[1 < X < 3]$$

$$= \int_1^3 \left(\frac{x}{6} + \frac{1}{12}\right) dx = \frac{x^2}{12} + \frac{x}{12}\Big|_1^3 = \frac{5}{6} .$$

DEFINICION 8.3.1 INTERVALOS DE CONFIANZA Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población X, cuya distribución de probabilidad es $f(x,\theta)$.
Sea $\theta_1 = G(X_1, X_2, \ldots, X_n)$ y $\theta_2 = G(X_1, X_2, \ldots, X_n)$ dos estadísticos tales que $\theta_1 < \theta_2$, para los cuales se cumple

$$P[\theta_1 \leq \theta \leq \theta_2] = \gamma = 1 - \alpha \qquad (*)$$

donde γ no depende de θ ; entonces el intervalo aleatorio $[\theta_1, \theta_2]$ Se llama el intervalo del 100 γ% ó 100$(1 - \alpha)$% de confianza para θ .
La relación $(*)$ se interpreta como sigue: γ es igual a la probabilidad que el intervalo aleatorio $[\theta_1, \theta_2]$ contenga a θ .

θ_1 , se llama el límite inferior de confianza para θ .

θ_2 , se llama el límite superior de confianza para θ .

$\gamma = 1 - \alpha$, se llama el nivel (o coeficiente) de confianza.

La elección del nivel de confianza depende del investigador y los valores más utilizados son

$$\gamma = 0.90 \;; \quad \gamma = 0.95 \;; \quad \gamma = 0.975 \;; \quad \gamma = 0.98 \;; \quad \gamma = 0.99$$

Por ejemplo, si se escoge $\gamma = 0.95$; se establece que $[\theta_1, \theta_2]$ es un intervalo del 95 % de confianza para θ. En otras palabras estamos diciendo: de las muestras que podemos obtener cerca del 95 % producirán intervalos que incluyan al valor θ, mientras que el 5 % no. En general, la interpretación de un intervalo del 100 γ% de confianza para θ es, que si se seleccionan muchas muestras aleatorias y se calcula para cada muestra el intervalo de confianza del 100 γ % para θ, entonces el 100 γ % de estos intervalos incluirán el verdadero valor de θ. Ahora frecuentemente en la práctica se obtiene sólo una muestra, por tanto se calcula sólo un intervalo de confianza. Desde que este intervalo contiene o no al verdadero valor de θ, no es razonable aplicarle probabilidad a este resultado específico. Si θ_1 y θ_2 son los resultados específicos, carece de sentido, la expresión $P[\theta_1 \leq \theta \leq \theta_2] = \gamma$, por lo que escribiremos simplemente $[\theta_1, \theta_2]$; $\theta_1 \leq \theta \leq \theta_2$. El valor $[\theta_1, \theta_2]$ ó $\theta_1 \leq \theta \leq \theta_2$ del intervalo aleatorio $[\theta_1, \theta_2]$ se llaman también intervalo del 100γ% de confianza para θ.

El intervalo de confianza $[\theta_1, \theta_2]$ se llama también *Intervalo de confianza de dos lados o bilateral*, por que especifíca el límite inferior y superior.

INTERVALO DE CONFIANZA UNILATERAL O DE UN LADO Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población X con distribución de probabilidad $f(x;\theta)$

a) Si $\theta_1 = G(X_1, X_2, \ldots X_n)$ es un estadístico, para el cual ,

$$P[\theta_1 \leq \theta] = \gamma = 1 - \alpha$$

entonces el intervalo $[\theta_1, \infty\rangle$ se llama el intervalo inferior del 100 γ% ó 100(1 - α)% de confianza para θ.

θ_1, se llama límite unilateral inferior de confianza para θ.

b) Similarmente, si $\theta_2 = G(X_1, X_2, \ldots, X_n)$ es un estadístico tal que

$$P[\theta \leq \theta_2] = \gamma = 1 - \alpha$$

entonces el intervalo $\langle -\infty, \theta_2]$ se llama el intervalo unilateral superior del 100 γ % ó 100(1 - α)% de confianza para θ.

θ_2, se llama límite unilateral superior de confianza para θ.

OBSERVACION el lector habrá observado, que obtener un intervalo de confianza para un parámetro θ consiste en:

i) Elegir una probabilidad γ cercano a 1 ó α cercano a cero.

ii) Hallar dos estadísticos Θ_1 y Θ_2 de θ tal que $P[\Theta_1 \leq \theta \leq \Theta_2] = \gamma = 1 - \alpha$

8.3.1 INTERVALO DE CONFIANZA PARA LA MEDIA CON VARIANZA CONOCIDA, MUESTRA GRANDE

Sea X una población distribuida con media μ desconocida y varianza σ^2 conocida. Debemos hallar un intervalo de confianza para μ. Es decir, encontrar dos estadísticos Θ_1 y Θ_2 tal que $P[\Theta_1 < \mu < \Theta_2] = \gamma$, con γ conocido. Entonces,

1 Elegimos el nivel de confianza $\gamma = 1 - \alpha$

2 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de X, y $\bar{X}$ la media muestral

3 Sabemos que $\bar{X}$ es adecuada para estimar μ, por lo tanto podemos usar la distribución muestral de $\bar{X}$ para establecer un intervalo de confianza para μ

4 Para n suficientemente grande ($n \geq 30$) por el teorema central del límite se tiene que

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Si X es una población normal, entonces $\bar{X}$ es normal para todo n. Además se obtiene

$$Z = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \sim N(0,1) \qquad (*)$$

5 Observe que, si bien la variable aleatoria Z(de $*$)depende del parámetro μ,su distribución no. Por lo tanto, elegido $\gamma = 1 - \alpha$, puede determinarse dos valores z_1 y z_2 tal que

$$P\left[z_1 \leq Z \leq z_2\right] = \gamma = 1 - \alpha$$

6 Hay infinidad de formas de escoger z_1 y z_2 que cumplen tal condición. El más simple es escoger $z_2 = -z_1 = z_0$(Ver Fig. 8.3.1)

$$P[-z_0 \leq Z \leq z_0] = \gamma = 1 - \alpha$$

7 Por la simetría de la curva normal, se tiene

$$P[-z_0 \leq Z \leq z_0] = 2P[Z \leq z_0] - 1 = \gamma$$

Es decir,

$$2P[Z \leq z_0] = 1 + \gamma \qquad \text{ó}$$

$$P[Z \leq z_0] = \frac{1}{2}(1 + \gamma)$$

$$= 1 - \frac{\alpha}{2}$$

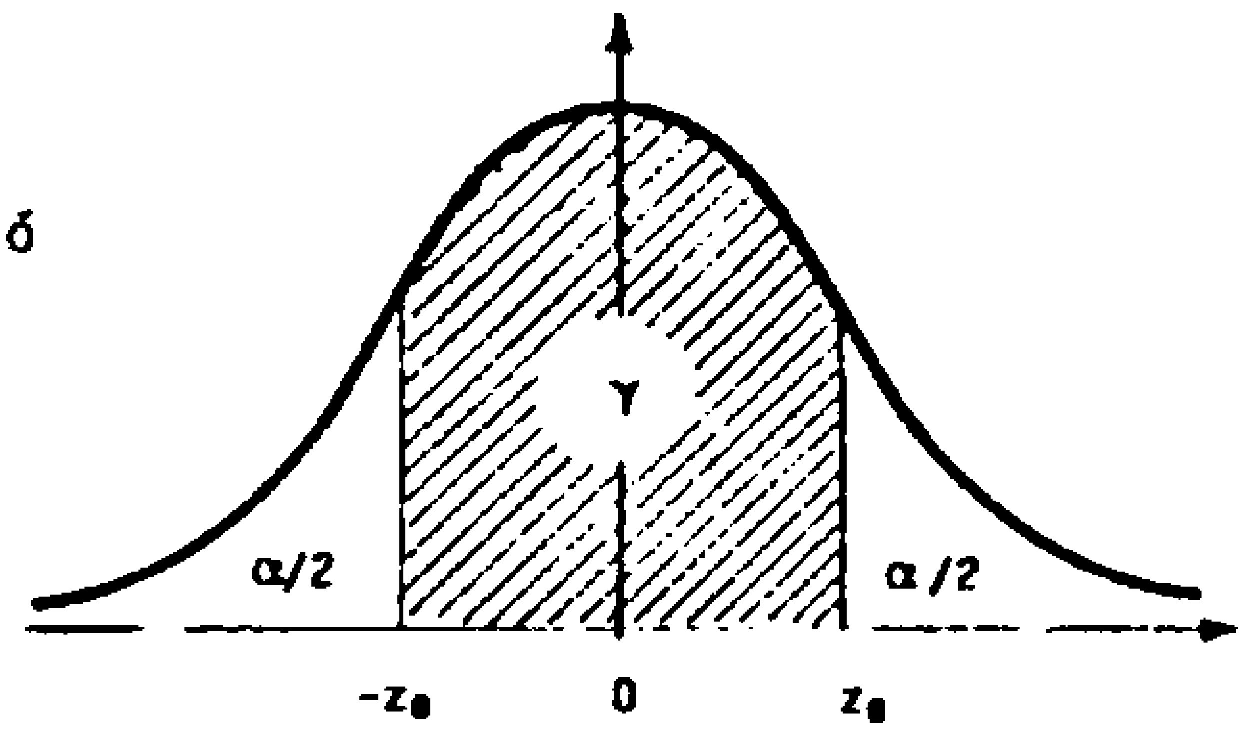


Fig. 8.3.1

8 Del paso 7 y la tabla III (de la distribución normal), se encuentra el valor de z_0.

9 Por otro lado, la expresión del paso 6 sustituyendo Z, se escribe equivalentemente

$$P[-z_0 \leq Z \leq z_0] = P[-z_0 \leq \frac{\bar{X} - \mu}{\sigma}\sqrt{n} \leq z_0] = \gamma$$

$$= P[-\frac{z_0\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq \frac{z_0\sigma}{\sqrt{n}}] = \gamma$$

$$= P[-\bar{X} - \frac{z_0\sigma}{\sqrt{n}} \leq -\mu \leq -\bar{X} + \frac{z_0\sigma}{\sqrt{n}}] = \gamma$$

$$= P[\bar{X} - \frac{z_0\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{z_0\sigma}{\sqrt{n}}] = \gamma$$

De donde se obtiene el intervalo $[\bar{X} - \frac{z_0\sigma}{\sqrt{n}}, \bar{X} + \frac{z_0\sigma}{\sqrt{n}}]$ del $100\gamma\%$ de confianza para el parámetro μ .

10 Para un valor particular de la muestra $x_1, x_2, \ldots, x_n$, sustituyendo el valor de $\bar{x}$ de $\bar{X}$ en el intervalo aleatorio, se obtiene el siguiente intervalo del $100\gamma\%$ ó $100(1 - \alpha)\%$ de confianza para θ

$$\bar{x} - \frac{z_0\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{z_0\sigma}{\sqrt{n}} \quad .$$

El teorema siguiente formaliza y resume el resultado anterior

TEOREMA 8.3.1 INTYERVALO DE CONFIANZA PARA μ CON σ^2 CONOCIDA

Si $\bar{x}$ es la media de una muestra de tamaño n ($n \geq 30$) tomada de una población distribuída con media μ (desconocida) y varianza σ^2 conocida, entonces

$$[\bar{x} - \frac{\sigma z_o}{\sqrt{n}} \quad , \quad \bar{x} + \frac{\sigma z_o}{\sqrt{n}}]$$

es aproximadamente un intervalo del 100 γ% de confianza para la media μ. Donde z_o es tal que $P[Z \leq z_o] = \frac{1-\gamma}{2}$

NOTA 1 Si se desconoce σ y $n \geq 30$, se puede usar la desviación estándar muestral S para aproximar σ.

NOTA 2 Si la variable aleatoria **X** se distribuye normalmente con media μ y varianza σ^2 (conocida), entonces, el teorema anterior se cumple también para $n < 30$.

NOTA 3 En el capítulo anterior (C.7.2.2), se vio que, cuando el muestreo es sin reemplazo en poblaciones finitas, se usa al factor de corrección para población finita $(N - n)/(N - 1)$. Al estimar parámetros de población con muestras sin reemplazo, se debe utilizar el factor de corrección para población finita. Entonces para población finita de N elementos, σ conocida, muestreo sin reemplazo y $n \geq 30$, el intervalo del 100 γ % de confianza para μ es

$$\left[\bar{X} - \frac{\sigma z_o}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \; , \; \bar{X} + \frac{\sigma z_o}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right]$$

Si σ es desconocida por la nota 1 el intervalo de confianza es

$$\left[\bar{X} - \frac{s z_o}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \; , \; \bar{X} + \frac{s z_o}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right]$$

OBSERVACIONES

1 z_o es una función creciente del nivel de confianza γ. Por lo tanto, cuanto más cercano a 1 se escoge γ, debemos esperar intervalos de confianza más grandes.

2 El tamaño de la muestra aparece en el denominador de σz_o, entonces, muestras más grandes darán intervalos de confianza más cortos, por lo tanto información más precisa.

EJEMPLO 3 Construir un intervalo del 95 % de confianza para la media de la población, a partir de una muestra de tamaño 64 extraída de una población con $\sigma = 10$. La media muestral resultó $\bar{x} = 48.5$.

SOLUCION 1 $\gamma = 0.95$

2. reemplazando en la formula del paso 7

$$P[Z \leq z_0] = \frac{1 + 0.95}{2} = 0.975$$

de la tabla III se obtiene

$$z_0 = 1.96$$

3. Cálculo de $\theta_1 = 48.5 - \dfrac{10(1.96)}{8} = 46.05$

$$\theta_2 = 48.5 + \frac{10(1.96)}{8} = 50.95$$

4. luego $[46.05 \, , \, 50.95]$ es el intervalo del 95 % de confianza para μ .

Los pasos que se siguen para determinar un intervalo de confianza se pueden resumir en el siguiente cuadro:

1 Elegir un nivel de confianza $\gamma = 1 - \alpha$

2 Determinar z_0 mediante la tabla III y la fórmula

$$P[Z \leq z_0] = \frac{1 + \gamma}{2}$$

γ	0.90	0.95	0.98	0.99
z_0	1.645	1.960	2.33	2.576

3 Calcular $\sigma z_0 / \sqrt{n}$.

4 Cálculo de $\bar{x}$.

5 Cálculo de $\bar{x} - \dfrac{\sigma z_0}{\sqrt{n}}$ y $\bar{x} + \dfrac{\sigma z_0}{\sqrt{n}}$

6 El intervalo del $100\gamma\%$ de confianza para la media de la población es,

$$\left[\bar{x} - \frac{\sigma z_0}{\sqrt{n}} \, , \, \bar{x} = \frac{\sigma z_0}{\sqrt{n}}\right]$$

NOTA 4 Una propiedad de la media muestral es que se puede utilizar para estimar la *cantidad total* de una población en casos apropiados. Por ejemplo, si la media del salario mensual para una muestra de 6 obreros calificados de una fábrica es

I/. 5000, y en la fábrica trabajan 100 obreros calificados, el gasto mensual en obreros calificados es I/. 100 x 5000 = I/. 500,000. Es decir, si

N = Tamaño de la población

$\bar{X}$ = media muestral

τ = Total = $N\bar{X}$

Para situaciones en los cuales resulta apropiado utilizar la media muestral para estimar una cantidad total, la desviación típica para el total es

$$S_{total} = N\,S$$

Entonces, el intervalo del 100γ % de confianza para el total "τ" de la población, se obtiene multiplicando, los límites del intervalo de confianza del 100γ % para μ por el número total

$$[N\bar{X} - z_o \frac{N\,S}{\sqrt{n}}\ ,\ N\bar{X} + z_o \frac{N\,S}{\sqrt{n}}\ ]$$

EJEMPLO 4 Cuarenta y nueve reses recibieron una alimentación especial por un período de cuatro meses, siendo el aumento medio de peso en este período de 60 kg. con una desviación típica de 7 kg. ¿Con qué nivel de confianza puede afirmarse que esta dieta produce un aumento medio de peso de 59.25 a 60.75 en un período de cuatro meses?

SOLUCION Datos $n = 49$, $S = 7$ kg.

1. Queremos calcular $\gamma = ?$
2. Puesto que n es grande, aproximando σ por S tenemos

$$\bar{x} - z_o \frac{S}{\sqrt{n}} = \bar{x} - z_o = 59.25$$

$$\bar{x} + z_o \frac{S}{\sqrt{n}} = \bar{x} + z_o = 60.75$$

3. Resolviendo el sistema se obtiene $z_o = 0.75$
4. Reemplazando el valor de z_o en la fórmula

$$P[Z \le z_o] = P[Z \le 0.75] = 0.7734 = \frac{1}{2}(1 + \gamma)$$

ó $1 + \gamma = 1.5468$, de donde $\gamma = 0.5468$.

EJEMPLO 5 Una auditoría del inventario de una compañía se realizó seleccionando una muestra al azar de 100 productos en existencia. El precio de venta promedio obtenido en la muestra fue de I/. 17.5 con una desviación típica de I/.6.75. Construya un intervalo de confianza del 95% para el precio promedio de todos los artículos en existencia.

SOLUCION 1 $\gamma = 0.95$.

2. $P[Z \le z_o] = \dfrac{1+0.95}{2} = 0.975$, de la tabla III, $z_o = 1.96$.

3. $\bar{x} = 17.5$, $n = 100$ grande aproximamos σ por $S = 6.75$.

4. Cálculo de $\dfrac{z_o S}{\sqrt{n}} = \dfrac{(1.96)(6.75)}{\sqrt{100}} = \dfrac{13.23}{10} = 1.323$.

5. $\bar{x} - \dfrac{z_o S}{\sqrt{n}} = 17.5 - 1.323 = 16.177$.

$\bar{x} + \dfrac{z_o S}{\sqrt{n}} = 17.5 + 1.323 = 18.823$.

6. El intervalo del 95 % de confianza para el precio promedio de todos los artículos en existencia es $[16.1777, 18.823]$.

EJEMPLO 6 De un embarque de 2,200 secadoras de mano se probó 81 secadoras al azar. La vida promedio en la muestra fué de 3.2 horas con una desviación estandar de 0.9 horas. Construya un intervalo de confianza del 95 % para la vida media de las secadoras del embarque.

SOLUCION 1 $\gamma = 0.95$

2. $P[Z \le z_o] = \dfrac{1+0.95}{2} = 0.975$ y de la tabla III, $z_o = 1.96$

3. $\bar{x} = 3.2$, $S = 0.9$

4. Puesto que $n = 81$ es grande, y $N = 2,200$ tamaño de la población, por la nota 3, tenemos que

$$\frac{z_o S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{(1.96)(0.9)}{\sqrt{81}}\sqrt{\frac{2200-81}{2200-1}} = (0.196)\sqrt{0.963} = (0.196)(0.981)$$
$$= 0.192$$

5. $\bar{x} - \dfrac{z_o S}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}} = 3.2 - 0.192 = 3.008$

$\bar{x} + \dfrac{z_o S}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}} = 3.2 + 0.192 = 3.396$

6. El intervalo del 95 % de confianza para la vida promedio del embarque de secadoras es $[3.008 , 3.396]$

EJEMPLO 7 Se va a vender un nuevo tipo de leche en polvo para niños como prueba de mercado durante un mes en las tiendas de una cadena de autoservicio. Los resultados de una muestra de 36 tiendas indicaron ventas promedio de I/. 1200 con una desviación estándar de I/. 180. Si la cadena tiene 200 tiendas, establezca una estimación del intervalo con 99 % de confianza de las ventas totales de esta leche en la cadena de tiendas de autoservicio.

SOLUCION 1 $\gamma = 0.99$

2. De la fórmula $P[Z \le z_0] = \frac{1+\gamma}{2} = 0.995$, y la tabla III, $z_0 = 2.576$.

3. $\bar{x} = 1200$, $S = 180$, $n = 36$

4. $z_0 \frac{S}{\sqrt{n}} = (2.576)\frac{180}{\sqrt{36}} = (2.576)30 = 77.28$

5. $\bar{x} - \frac{z_0 S}{\sqrt{n}} = 1200 - 77.28 = 1122.72$

$\bar{x} + z_0 \frac{S}{\sqrt{n}} = 1200 + 77.28 = 1277.28$

6. Multiplicando los límites de este intervalo por N = 200 se tiene (ver Nota - 4)

$$200(1122.72) = 224544$$
$$200(1277.28) = 255456$$

Por lo tanto, [224544, 255456] es el intervalo del 99 % de confianza de las ventas totales de la nueva leche en la cadena de tiendas de autoservicio.

8.3.2 TAMAÑO MUESTRAL PARA ESTIMAR UNA MEDIA

El intervalo del 100 γ % de confianza para μ, es un intervalo con centro $\bar{x}$ (Ver fig. 8.3.2), entonces si μ es el valor central del intervalo del 100 % de confianza para μ, $\bar{x}$ estima a μ sin error. Desde que la mayoría de las veces, $\bar{x}$ no es exactamente igual a μ, hay un error de muestreo en la estimación de μ por intervalo. El tamaño del error de estimación será la longitud de la diferencia $\bar{x} - \mu$; es decir $|\bar{x} - \mu|$. Y esta diferencia es menor o igual que $z_0\sigma/\sqrt{n}$ con 100 γ % de confianza. La figura 8.3.2 muestra el error de estimación de μ por $\bar{x}$. Note que la amplitud del intervalo es $2z_0\sigma/\sqrt{n}$.

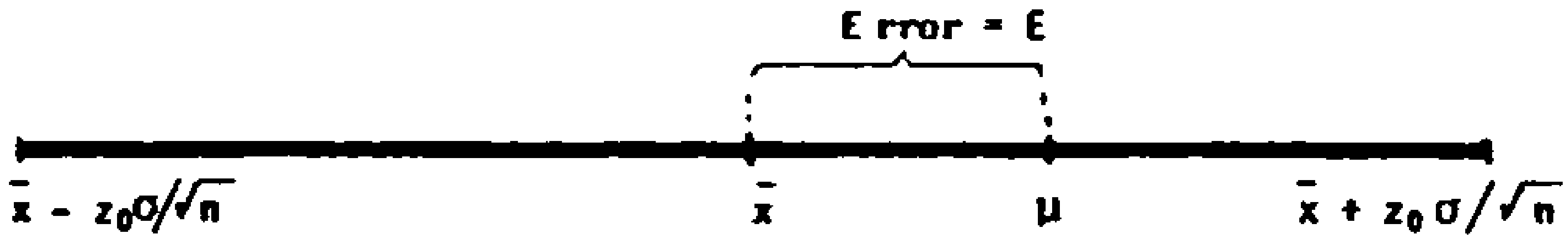


Fig. 8.3.2

El procedimiento para seleccionar el tamaño de muestra es

1. Elija E el error máximo permisible (cota para el error de estimación) y un coeficiente de confianza $\gamma = 1 - \alpha$
2. Resuelva la ecuación $E = z_o\sigma/\sqrt{n}$ para n. Es decir,

$$n = [\frac{z_o\sigma}{E}]^2$$

EJEMPLO 8 Una firma constructora desea estimar la resistencia media de las barras de acero utilizadas en la construcción de edificios de departamentos. ¿Qué tamaño muestral se requiere para garantizar que haya un riesgo de sólo 0.001 de sobrepasar un error de 5 kg. o más en la estimación?. La desviación típica de la resistencia de este tipo de barras se estima en 25 kg.

SOLUCION $E = 5$ kg , $\sigma = 25$ kg. $n = (z_o\sigma/E)^2$

Desde que el riesgo de sobrepasar este error es de sólo 0.001, entonces el grado de confianza para que este error sea el máximo permisible es $\gamma = 1 - 0.001 = 0.999$.

$$P[Z \le z_o] = \frac{1}{2}(1 + \gamma) = 0.9995, \text{ de donde } z_o = 3.09$$

Luego, $n = (3.09)^2(25)^2/5^2$, ó $n = 238.7 \simeq 239$.

EJEMPLO 9 Para su producción total de bombillas, la gerencia de una firma electrónica está segura que los límites superior e inferior de vida no difieren por más de 600 horas. Para un nivel de confianza del 90 %, ¿Qué tan grande debe tomarse la muestra para encontrar la vida promedio de una bombilla dentro de más y menos 30 horas?

SOLUCION $E = 30$ horas , $\gamma = 0.90$, $n = (z_o\sigma/E)^2$

$$P[Z \le z_o] = \frac{1 + \gamma}{2} = 0.95, \text{ de Tabla III, } z_o = 1.645$$

Cálculo de σ. En el capítulo 6 hemos visto que el intervalo $[-2\sigma , 2\sigma]$ -

incluye el 95.46 % del área bajo la curva normal; es decir este intervalo incluye la mayor parte de las observaciones. La fig. 8.3.3 muestra la relación aproximada entre el rango y la desviación típica de la población. Entonces

$4\sigma \sim 600$ (aproximadamente)

$\sigma \sim 150.$

Por lo tanto

$$n = \left(\frac{(1.645)150}{30}\right)^2 = 67.65$$

$$n = 68$$

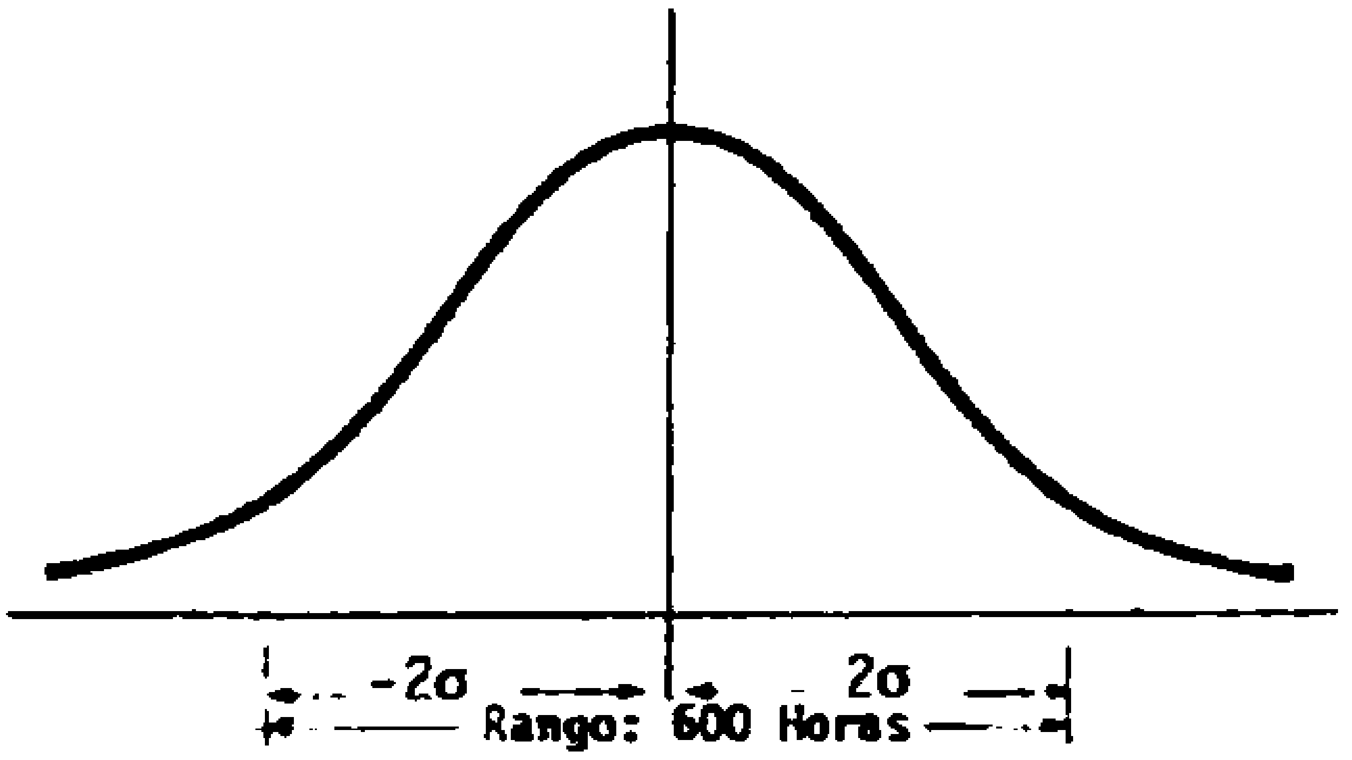


Fig. 8.3.3

8.3.3 INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE MEDIAS DE DOS DISTRIBUCIONES CON AMBAS DESVIACIONES TIPICAS CONOCIDAS, MUESTRAS GRANDES

Sea X una variable aleatoria distribuída con media μ_X y varianza σ_X^2 (conocida). Sea Y una variable aleatoria independiente de X distribuída con media μ_X y varianza σ_X^2 (conocida). Queremos hallar un intervalo de confianza del 100 γ %, para la diferencia de las medias $\mu_x - \mu_y$. Es decir debemos encontrar dos estadísticos

θ_1 y θ_2 tal que $P[\theta_1 \leq \mu_X - \mu_Y \leq \theta_2] = \gamma$. Entonces

1. Elegimos el nivel de confianza $\gamma = 1 - \alpha$.
2. Consideremos una muestra aleatoria de tamaño n ($n \geq 30$) de X y una muestra aleatoria de tamaño m ($m \geq 30$) de Y. Sean $\overline{X}$ e $\overline{Y}$ las medias muestrales respectivamente.
3. Sabemos que la estadística adecuada para estimar $\mu_X - \mu_Y$ es $\overline{X} - \overline{Y}$, entonces usaremos la distribución muestral de $\overline{X} - \overline{Y}$ para establecer un intervalo de confianza para $\mu_X - \mu_Y$
4. Para n y m suficientemente grande ($n \geq 30$, $m \geq 30$); la variable aleatoria

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

tiene aproximadamente una distribución normal estándar.

Si X e Y son poblaciones con distribución normal, $\bar{X}$ e $\bar{Y}$ tienen distribución normal para todo n y m, por lo tanto $\bar{X} - \bar{Y}$ también tiene una distribución normal.

5. Si bien la variable aleatoria Z depende de $\mu_X - \mu_Y$, sin embargo su distribución no. Por lo tanto, elegido γ, puede encontrarse z_0 tal que

$$P[-z_0 \le Z \le z_0] = \gamma \qquad (1)$$

de donde, usando la simetría de la curva normal, se obtiene

$$P[Z \le z_0] = \frac{1}{2}(1 + \gamma) \qquad (2)$$

de (2) utilizando la tabla de la distribución normal se obtiene z_0.

6. Por otro lado la expresión (1), sustituyendo Z, se escribe como sigue

$$\gamma = P[-z_0 \le Z \le z_0] = P\left[-z_0 \le \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \le z_0\right]$$

$$= P\left[-z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le \bar{X} - \bar{Y} - (\mu_X - \mu_Y) \le z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

$$= P\left[-(\bar{X} - \bar{Y}) - z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le -(\mu_X - \mu_Y) \le -(\bar{X} - \bar{Y}) + z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

$$= P\left[(\bar{X} - \bar{Y}) - z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le \mu_X - \mu_Y \le (\bar{X} - \bar{Y}) + z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

Haciendo $\theta_1 = (\bar{X} - \bar{Y}) - z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$

$$\theta_2 = (\bar{X} - \bar{Y}) + z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$$

obtenemos el intervalo aleatorio

$$\left[(\bar{X} - \bar{Y}) - z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\ ,\quad (\bar{X} - \bar{Y}) + z_0\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

que con probabilidad γ contiene a $\mu_X - \mu_Y$. En otras palabras, el 100 γ % de las veces $\mu_X - \mu_Y$ estará contenido en dicho intervalo.

Para dos muestras aleatorias independientes de tamaño n y m extraídas de dos poblaciones cuyas varianzas σ_X^2 y σ_Y^2 son conocidas, sean $\bar{x}$ e $\bar{y}$ las medias de las muestras observadas. Entonces el intervalo de confianza del 100 γ % para $\mu_X - \mu_Y$, está dado por el teorema siguiente que resume el resultado anterior.

TEOREMA 8.3.2 INTERVALO DE CONFIANZA PARA $\mu_X - \mu_Y$; CON σ_X^2 y σ_Y^2 CONOCIDOS

Si $\bar{x}$ e $\bar{y}$ son las medias de dos muestras aleatorias independientes de tamaños n y m, extraídas de dos poblaciones distribuídas con medias (desconocidas) μ_X, μ_Y y varianzas σ_X^2, σ_Y^2(conocidas), respectivamente. Entonces

$$\left[(\bar{x} - \bar{y}) - z_0 \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \quad , \quad (\bar{x} - \bar{y}) + z_0 \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \right]$$

es el intervalo de confianza 100γ% para diferencia de medias $\mu_X - \mu_Y$. Donde z_0 es tal que

$$P[Z \leq z_0] = \frac{1}{2}(1 + \gamma)$$

Como en el caso anterior, el procedimiento para encontrar el intervalo de confianza para $\mu_X - \mu_Y$ con σ_X^2 y σ_Y^2 conocidos, se resume en el cuadro siguiente.

1. Se escoge el nivel de confianza $\gamma = 1 - \alpha$
2. Se determina z_0 ,de la ecuación $P[Z \leq z_0] = \frac{1}{2}(1 + \gamma)$ con ayuda de la tabla de la distribución normal.
3. se calcula $\bar{x}$, $\bar{y}$
4. se calcula $z_0 \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$
5. Y se escribe el intervalo de confianza

$$\left[(\bar{x} - \bar{y}) - z_0 \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \quad , \quad (\bar{x} - \bar{y}) + z_0 \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \right]$$

NOTA 1 Si X e Y son variables aleatorias normales con varianzas conocidas y n , $m < 30$; el procedimiento anterior, se aplica para estimar el intervalo de confianza para la diferencia de medias.

NOTA 2 El procedimiento dado, se aplica también para estimar intervalos de confianza para diferencia de medias,si σ_X^2 y σ_Y^2 son estimados a partir de muestras grandes.

EJEMPLO 10 La media y la desviación típica de las cargas máxima soportada por 100 cables producidos por la Compañía-DURAMAS son 20 toneladas y 1.1 tonelada. La media y la desviación típica de 60 cables producidos por la COMPAÑIA CABLECO son 16 toneladas y 0.8 toneladas, respectivamente. Determinar un intervalo de confianza del 95 % para la diferencia de cargas máximas medias.

SOLUCION 1 $\gamma = 0.95$

2. $P[Z \le z_o] = \dfrac{1 + 0.95}{2} = 0.975$

de la tabla III, se obtiene $z_o = 1.96$

3. $n = 100$, $\bar{x} = 20$, $s_X = 1.1$

$m = 60$, $\bar{y} = 16$, $s_Y = 0.8$

4. $$z_0\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}} = (1.96)\sqrt{\frac{(1.1)^2}{100} + \frac{(0.8)^2}{60}} = 0.29596$$

Luego, $\theta_1 = (20 - 16) - 0.29596 = 3.70404$

$\theta_2 = (20 - 16) + 0.29596 = 4.2956$

Por lo tanto

$$[3.7 \ , \ 4.3]$$

es un intervalo de confianza del 95 %, para la diferencia de medias $\mu_X - \mu_Y$.

8.3.4 INTERVALO DE CONFIANZA PARA UNA PROPORCION, MUESTRAS GRANDES

Un estimador puntual de la proporción p en la distribución binomial está dado por $\hat{P} = \frac{X}{n}$. Por lo tanto, la proporción muestral observada $\hat{p} = \frac{x}{n}$, se usa como un estimado puntual del parámetro p .

Si p es desconocido se puede establecer un intervalo de confianza para p, $(p_1 \le p \le p_2)$ considerando la distribución muestral de $\hat{P}$, la cual es el mismo de la variable aleatoria X, tal como hemos visto en el capítulo 7. Para una muestra aleatoria de tamaño n suficientemente grande ($n \ge 30$), la variable aleatoria $\hat{P}$ tiene una distribución aproximadamente normal con media

$\mu_{\hat{P}} = E(\hat{P}) = p$, y varianza $\sigma^2_{\hat{P}} = \frac{pq}{n}$.

Y la variable aleatoria $Z = \frac{\hat{P} - \hat{p}}{\sqrt{pq/n}}$ tiene una distribución aproximadamente normal estándar para n grande.

Si bien la variable aleatoria Z depende de p, sin embargo su distribución no. Entonces, elegido γ , se puede encontrar z_0 tal que

$$P[- z_0 \leq Z \leq z_0] = \gamma \tag{1}$$

o

$$P[Z \leq z_0] = \frac{1}{2}(1 + \gamma) \tag{2}$$

de (2) y la tabla III (de la distribución normal) se obtiene z_0 .

Por otro lado sustituyendo, Z, la expresión (1) se escribe,

$$P[- z_0 \leq Z \leq z_0] = P[- z_0 \leq \frac{\hat{P} - p}{\sqrt{pq/n}} \leq z_0] = \gamma$$

$$= P[- z_0 \sqrt{pq/n} \leq \hat{P} - p \leq z_0 \sqrt{pq/n}] = \gamma$$

$$= P[- \hat{P} - z_0 \sqrt{pq/n} \leq - p \leq - \hat{P} + z_0 \sqrt{pq/n}] = \gamma$$

$$= P[\hat{P} - z_0 \sqrt{pq/n} \leq p \leq \hat{P} + z_0 \sqrt{pq/n}] = \gamma \tag{3}$$

Para obtener un intervalo que no contenga a p en ninguno de los extremos, debe recurrirse a manipulación más complicada. Cuando n es grande, se aproxima - por $\hat{p} = \frac{x}{n}$ en la expresión $\frac{pq}{n} = \frac{p(1 - p)}{n}$. Entonces (3) se escribe

$$P[\hat{p} - z_0 \sqrt{\frac{(x/n)(1 - x/n)}{n}} \leq p \leq \hat{P} + z_0 \sqrt{\frac{(x/n)(1 - x/n)}{n}}] = \gamma$$

ó

$$P[\hat{P} - z_0 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \leq p \leq \hat{P} + z_0 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}] = \gamma$$

Si el valor observado de X es x , en una muestra aleatoria de tamaño n (n grande), la proporción muestral es $\hat{p} = \frac{x}{n}$, entonces el intervalo de confianza del 100γ% para p, está dado aproximadamente por

$$[\hat{p} - z_0 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} , \quad \hat{p} + z_0 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}]$$

ó $\left[\frac{x}{n} - z_0\sqrt{\frac{(x/n)(1-x/n)}{n}} \quad , \quad \frac{x}{n} + z_0\sqrt{\frac{(x/n)(1-x/n)}{n}}\right]$

TEOREMA 8.3.3 INTERVALO DE CONFIANZA PARA p CON $n \geq 30$ Si $\hat{p} = x/n$ es la proporción de éxitos en una muestra de tamaño n, el intervalo de confianza del 100γ% para p, es aproximadamente

$$\left[\hat{p} - z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad , \quad \hat{p} + z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

donde z_0, es el valor de la curva normal estándar, cuya área está dada por

$$P[Z \leq z_0] = \frac{1}{2}(1+\gamma)$$

Se deja a cargo del lector resumir en un cuadro el procedimiento seguido para encontrar el intervalo de confianza para p.

NOTA De acuerdo con la nota 3 de 8.3.1, el intervalo del 100γ% de confianza de la proporción, para un muestreo sin reemplazo de una población finita es

$$\left[\hat{p} - z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}} \quad , \quad \hat{p} + z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}}\right]$$

EJEMPLO 11 Durante cierta semana, una tienda de departamentos observó y registró que 5,750 de las 12,500 personas que entraron en la tienda hicieron por lo menos una compra. Tratando esto como una muestra al azar de todos los clientes potenciales, hallar el intervalo de confianza del 99 % para la proporción real de personas que entran en la tienda y que harán por lo menos una compra.

SOLUCION 1 $\gamma = 0.99$

2. $P[Z \leq z_0] = \frac{1}{2}(1 + 0.99) = 0.995$

y de la tabla de la distribución normal obtenemos $z_0 = 2.576$.

3. El tamaño de la muestra es, $n = 12{,}500$, el número de éxitos en la muestra $x = 5{,}750$, luego la proporción muestral es

$$\hat{p} = \frac{5{,}750}{12{,}500} = 0.46 \quad y \quad 1 - \hat{p} = 0.54 .$$

4. $z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (2.576)(0.00446) = 0.01148896 \approx 0.01$

si $p_1 = \frac{x}{n} - z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.46 - 0.01 = 0.45$

$$p_2 = \frac{x}{n} + z_0 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.46 + 0.01 = 0.47$$

se obtiene que $[0.45\ ,\ 0.47]$ es un intervalo de confianza del 99% para la proporción real de personas que entran a la tienda.

EJEMPLO 12 El gerente de una agencia bancaria que tiene 1,000 cuenta habitantes quiere determinar la proporción de sus cuentahabitantes a los cuales les paga el sueldo por semana. Se selecciona una muestra de 100 cuentahabitantes, de los cuales 30 indican que se les pagó por semana. Determinar el intervalo del 90% de confianza de la proporción real de cuentahabitantes a quienes se les paga por semana.

SOLUCION La población es finita con $N = 1000$ y es obvio que el muestreo es sin reemplazo, por lo tanto, se usa el factor de corrección para población finita,

1. $\gamma = 0.90$

2. $P[Z \le z_0] = \frac{1}{2}(1 + 0.90) = 0.95$ y de la tabla III, $z_0 = 1.645$.

3. $n = 100\ ,\quad x = 30 \quad$ y $\quad \hat{p} = \frac{30}{100} = 0.3 \quad , \quad 1 - \hat{p} = 0.7.$

4. $$z_0 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \sqrt{\frac{N-n}{N-1}} = (1.645)\sqrt{\frac{(0.3)(0.7)}{100}} \sqrt{\frac{1000-100}{100-1}}$$

$$= (1.645)\,\frac{\sqrt{21}}{100}\,\frac{10}{\sqrt{111}} \approx 0.0716$$

5. $$\hat{p} - z_0 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \sqrt{\frac{N-n}{N-1}} = 0.3 - 0.0716 = 0.2284$$

$$\hat{p} + z_0 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \sqrt{\frac{N-n}{N-1}} = 0.3 + 0.0716 = 0.3716\ .$$

luego, $[0.2284,\ 0.3716]$ es el intervalo del 90% de confianza de la proporción de cuentahabitantes a quienes se les paga por semana.

8.3.5 TAMAÑO MUESTRAL PARA ESTIMAR UNA PROPORCION

Si p es el valor central del intervalo de confianza del $100\gamma\%$, entonces $\hat{p}$ estima a p sin error. Pero casi todas las veces, p no es exactamente igual a $\hat{p}$, por lo tanto la estimación puntual tiene un error. El tamaño de este error será el valor absoluto de la diferencia entre p y $\hat{p}$, $|\hat{p} - p|$, y $100\gamma\ \%$ es la confianza de que este error es menor que $z_o \sqrt{\hat{p}(1-\hat{p})/n}$. Esto se muestra en la figura 8.3.4.

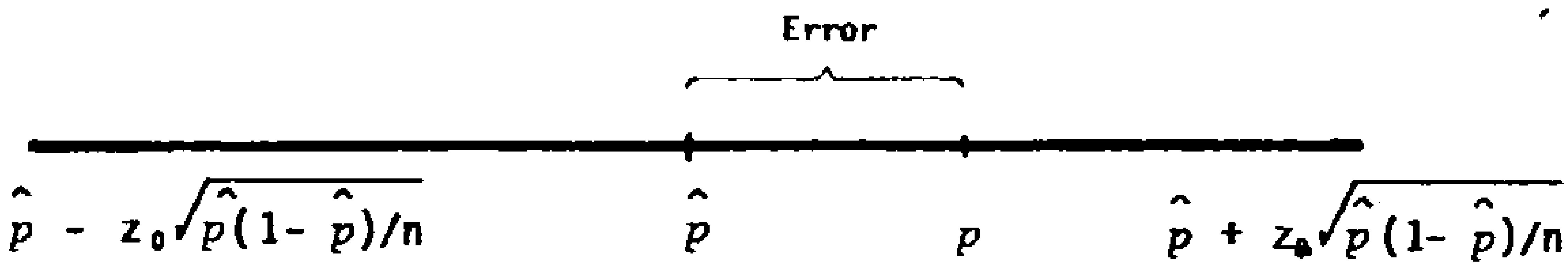


Fig. 8.3.4

El tamaño n de la muestra tal que el error en la estimación de p sea menor que un valor especificado E, por lo expresado en el párrafo anterior se debe escoger de tal manera que

$$z_o \sqrt{\hat{p}(1-\hat{p})/n} = E\ , \quad \text{de donde}$$

$$\boxed{n = \frac{z_o^2\ \hat{p}(1-\hat{p})}{E^2}}$$

Observe que se puede establecer una cota superior para n, para cualquier grado de confianza, de la siguiente manera. Desde que $\hat{p}$ está entre 0 y 1, $\hat{p}(1-\hat{p})$ es igual a lo más 1/4. Esto puede verificarse completando cuadrados, como sigue,

$$\begin{aligned} \hat{p}(1-\hat{p}) &= -\hat{p}^2 + \hat{p} = -(\hat{p}^2 - \hat{p}) \\ &= -(\hat{p}^2 - \hat{p} + \frac{1}{4} - \frac{1}{4}) \\ &= \frac{1}{4} - (\hat{p} - \frac{1}{2})^2 \end{aligned}$$

el cual es siempre menor que 1/4, excepto cuando $\hat{p} = \frac{1}{2}$, en este caso $\hat{p}(1-\hat{p}) = 1/4$. Por lo tanto, el tamaño de la muestra se puede escribir así,

$$n = \frac{z_o^2\ \hat{p}(1-\hat{p})}{E^2} = \frac{z_o^2 \cdot 1/4}{E^2}$$

de donde

$$\boxed{n = \frac{z_o^2}{4E^2}}$$

Esta fórmula se usa para hallar n cuando no se tiene conocimiento o información previa de la proporción verdadera p.

EJEMPLO 13 La oficina de planificación familiar de cierto distrito desea determinar la proporción de familias con un ingreso mensual inferior I/. 30,000.00. Estudios previos han indicado que esta proporción era de 20%.

(a) ¿Qué tamaño muestral se requiere para asegurar con confianza 0.95 que el error en la estimación de esta proporción no sobrepasará a 0.05?

(b) ¿En que forma variará el tamaño muestral requerido si el máximo error permisible es reducido a 0.01?

SOLUCION $\hat{p} = 0.2$ y $\gamma = 0.95$, entonces

$$P[Z \leq z_o] = \frac{1}{2}(1 + \gamma) = 0.975\ , \text{ de la tabla III, } z_o = 1.96.$$

(a) $E = 0.05$, error máximo permisible; luego,

$$n = \frac{(1.96)^2(0.2)(0.8)}{(0.05)^2} = 245.86$$

$$n \simeq 246.$$

(b) El error máximo permisible es $E = 0.01$, entonces,

$$n = \frac{(1.96)^2(0.2)(0.8)}{(0.01)^2} = 6146.56$$

$$n \simeq 6147.$$

8.3.6 TAMAÑO DE LA MUESTRA PARA POBLACIONES FINITAS

Para determinar el tamaño de la muestra cuando el muestreo es sin reemplazo, se debe utilizar también el factor de corrección para población finita. Así al estimar la media, el error muestral es

$$E = z_o \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}}$$

y al estimar las proporciones, el error muestral es

$$E = z_o \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \sqrt{\frac{N - n}{N - 1}}$$

Entonces, el tamaño de la muestra se puede obtener en dos etapas, de la siguiente manera:

1. Se determina el tamaño de la muestra como en 8.3.2 y 8.3.5. Esto es

(a) El tamaño de la muestra en la estimación de la media es

$$n_o = \frac{z_o^2 \sigma^2}{E^2}$$

(b) El tamaño de la muestra en la estimación de la proporción es

$$n_o = \frac{z_o^2 \hat{p}(1 - \hat{p})}{E^2} \quad \text{o} \quad n_o = \frac{z_o^2}{4E^2}$$

2. Se aplica el factor de corrección para población finita para obtener el tamaño final de la muestra. Así, el tamaño de la muestra en la estimación de la media

$$E = z_o \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}}$$

$$\frac{n}{\frac{N - n}{N - 1}} = \frac{z_o^2 \sigma}{E^2}$$

$$\frac{n(N - 1)}{N - n} = n_o \text{, de donde} \qquad n = \frac{n_o N}{n_o + (N - 1)}$$

Similarmente, se obtiene el tamaño de la muestra en la estimación de la proporción.

EJEMPLO 14 La desviación típica de la duración de los focos de una determinada fábrica de focos es 100 horas. Para un embarque de 2000 focos, el gerente de control de calidad de la fábrica desea determinar el tamaño de la muestra necesaria, para estimar la duración promedio con aproximación de ± 20 horas del promedio real con 95% de confianza.

SOLUCION Los datos son: $\sigma = 100$, $N = 2000$ $E = 20$ y $\gamma = 0.95$

1. De $P[Z \leq z_o] = \frac{1 + \gamma}{2} = 0.975$ y la tabla III, $z_o = 1.96$

2. $$n_o = \frac{z_o^2 \sigma^2}{E^2} = \frac{(1.96)^2 (100)^2}{(20)^2} = \frac{(3.8416)(100)^2}{400} = 96 .$$

3. $$n = \frac{n_o N}{n_o + (N - 1)} = \frac{96(2000)}{96 + 2000 - 1} = \frac{96}{2.095} = 92 .$$

$n = 92.$

EJEMPLO 15 Con referencia al ejemplo 12, determinar el tamaño de la muestra necesaria para estimar la proporción real con aproximación de ± 0.05, con 90% de confianza.

SOLUCION Los datos son; $N = 1000$, $E = 0.05$ y $\gamma = 0.90$.

1. De $P[Z \leq z_o] = \frac{1+\gamma}{2} = 0.95$ y la tabla III, $z_o = 1.645$.

2. Puesto que no hay información disponible de la proproción, usaremos la fórmula $z_o/4E^2$ para determinar n_o. Es decir

$$n_o = \frac{z_o^2}{4E} = \frac{(1.645)^2}{4(0.05)^2} = \frac{2.706025}{0.01} = 271.$$

3. $$n = \frac{n_o N}{n_o + (N-1)} = \frac{271\ (1000)}{271 + 1000 - 1} = \frac{27100}{127}$$

$$n = 214.$$

8.3.7 INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE PROPORCIONES

Consideremos dos muestras independientes de tamaños n_1 y n_2 seleccionados aleatoriamente de dos poblaciones binomiales con medias $n_1 p_1, n_2 p_2$ y varianzas $n_1 p_1 q_1$ y $n_2 p_2 q_2$ respectivamente. Denotemos la proporción de éxitos en cada muestra por $\hat{p}_1 = \frac{x}{n_1}$ y $\hat{p}_2 = \frac{y}{n_2}$. Un estimador puntual de la diferencia entre las dos proporciones, $p_1 - p_2$ es el estadísticos $\hat{P}_1 - \hat{P}_2$.

Un intervalo de confianza para $p_1 - p_2$ puede establecerse, considerando la la distribución muestral de $\hat{P}_1 - \hat{P}_2$. De la sección 7.2.5, para n_1 y n_2 suficientemente grande ($n \geq 30$), sabemos que $\hat{P}_1 - \hat{P}_2$ se distribuye aproximadamente normal con media $p_1 - p_2$ y varianza

$\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$. Es decir $\mu_{\hat{P}_1 - \hat{P}_2} = p_1 - p_2$ y

$\sigma^2_{\hat{P}_1 - \hat{P}_2} = \frac{p_1 q_1}{n} + \frac{p_2 q_2}{n}$. Entonces la variable aleatoria

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

tiene aproximadamente una distribución N(0,1)

La variable aleatoria Z depende de $p_1 - p_2$ pero su distribución no, entonces, elegido un γ podemos encontrar z_o tal que

$$P[-z_o \leq Z \leq z_o] = \gamma \tag{1}$$

ó

$$P[Z \leq z_o] = \frac{1}{2}(1+\gamma) \tag{2}$$

de (2) usando la tabla de la distribución normal se obtiene z_o.
Sustituyendo Z en la expresión (1), se escribe,

$$P[-z_o \leq Z \leq z_o] = P\left[-z_o \leq \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{p_1q_1}{n_1} \quad \frac{p_1q_2}{n_2}}} \leq z_o\right]$$

multiplicando cada término de la desigualdad por

$\sqrt{\frac{p_1q_1}{n_2} + \frac{p_2q_2}{n_2}}$, luego restando $\hat{P}_1 - \hat{P}_2$ y multiplicando por - 1, obtenemos,

$$P[-z_o \leq Z \leq z_o] = P\left[(\hat{P}_1 - \hat{P}_2) - z_o\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \leq p_1 - p_2 \leq (\hat{P}_1 - \hat{P}_2) + z_o\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\right] = \gamma$$

Si n_1 y n_2 ambos son grandes, reemplazamos p_1 y p_2 bajo el signo radical por sus estimados

$\hat{p}_1 = \frac{x}{n_1}$ y $\hat{p}_2 = \frac{y}{n_2}$. Entonces,

$$P\left[(\hat{P}_1 - \hat{P}_2) - z_o\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \leq p_1 - p_2 \leq (\hat{P}_1 - \hat{P}_2) + z_o\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}\right] = \gamma$$

Para dos muestras aleatorias independientes de tamaño n_1 y n_2, extraída de dos poblaciones binomiales, la diferencia muestral de proporciones, $\hat{p}_1 - \hat{p}_2$, se calcula y el intervalo del $100\gamma\%$ de confianza está dado por

$$\left[(\hat{p}_1 - \hat{p}_2) - z_o \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \;,\; (\hat{p}_1 - \hat{p}_2) + z_o \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}\right]$$

TEOREMA 8.3.4 INTERVALO DE CONFIANZA PARA $p_1 - p_2$; n_1 y $n_2 \geq 30$. El intervalo de confianza del 100γ % para la diferencia de dos parámetros binomiales, $p_1 - p_2$, es aproximadamente,

$$\left[(\hat{p}_1 - \hat{p}_2) - z_o \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \;,\; (\hat{p}_1 - \hat{p}_2) + z_o \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_1\hat{q}_2}{n_2}}\right], \text{ donde}$$

$\hat{p}_1$ y $\hat{p}_2$ son las proporciones de éxitos en la muestra aleatoria de tamaño n_1 y n_2 respectivamente, $\hat{q}_1 = 1 - \hat{p}_1$ y $\hat{q}_2 = 1 - \hat{p}_2$, y z_o es determinado por

$$P[Z \leq z_o] = \frac{1}{2}(1 + \gamma) .$$

EJEMPLO 16 Se ha encontrado que 25 de 250 cinescopios de televisión producidos por el proceso A son defectuosos y que 14 de 180 producidos por un proceso B son defectuosos. Suponiendo que el muestreo es aleatorio, determinar el intervalo del 99% de confianza para la diferencia verdadera en la proporción de defectuosos, de los dos procesos.

EJEMPLO 1 $\gamma = 0.99$.

2. $P[Z \leq z_o] = \frac{1}{2}(1 + \gamma) = \frac{1}{2}(1 + 0.99) = 0.995$, de la tabla de la distribución normal obtenemos $z_o = 2.58$

3. $n_1 = 250$, tamaño de la muestra de la producción del proceso A

   $x = 25$, número de éxitos (defectuosos) en la muestra

   $\hat{p}_1 = \frac{25}{250} = \frac{1}{10}$, la proporción muestral para la producción del proceso A y

   $\hat{q}_1 = \frac{9}{10}$

   $n_2 = 180$, tamaño de la muestra de producción del proceso B.

   $y = 14$, número de éxitos en la muestra

   $\hat{p}_2 = \frac{14}{180} = \frac{7}{90}$, la proporción muestral para la producción del proceso B y

   $\hat{q}_2 = \frac{83}{90}$

Luego $\hat{p}_1 - \hat{p}_2 = \frac{1}{10} - \frac{7}{90} = \frac{20}{900} = 0.02222$

4. $z_o\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = (2.58)\sqrt{\frac{9/100}{250} + \frac{581/8100}{180}} = 0.07095$

5. $(\hat{p}_1 - \hat{p}_2) - z_o\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 0.02222 - 0.07095 = -0.04873$

$(\hat{p}_1 - \hat{p}_2) + z_o\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 0.02222 + 0.07095 = 0.09317.$

Luego [- 0.049 , 0.093] es el intervalo de confianza del 99%, para la difencia de proporciones.

8.3.8 INTERVALO DE CONFIANZA PARA LA MEDIA CON VARIANZA DESCONOCIDA, MUESTRA PEQUEÑA

Sea X una variable aleatoria con distribución aproximadamente normal con media μ y varianza σ^2(desconocida). Además cuando σ^2 es desconocida se usa el estimador puntual S^2 . Entonces

1. Elégimos el nivel de confianza $\gamma = 1 - \alpha$
2. Consideremos una muestra aleatoria $X_1, X_2, \ldots, X_n$ de tamaño n($n < 30$), $\bar{X}$ y S la media y desviación típica muestral
3. Sabemos que $\bar{X}$ es adecuada para estimar μ, pero como σ^2 es desconocida usaremos la distribución muestral de la variable aleatoria $T = (\bar{X} - \mu)\sqrt{n}/S$ que tiene una dsitribución t con $n - 1$ grados de libertad
4. En el paso 3, aunque la variable aleatoria T depende de μ, su distribución no. Entonces elegido $\gamma = 1 - \alpha$, puede determinarse dos valores t_1 y t_2 tal que

$$P[t_1 \le T \le t_2] = \gamma = 1 - \alpha$$

5. Hay infinidad de formas de escoger t_1 y t_2 que cumplen tal condición, la más simple es tomar $t_2 = - t_1 = t_o$ (ver Fig. 8.3.5)

$$P[- t_o \le t_o] = \gamma = 1 - \alpha$$

6. Utilizando la simetría de la distribución t, tenemos

$$P[- t_o \le T \le t_o] = 2P[T \le t_o] - 1 = \gamma$$

Es decir

$$P[T \le t_o] = \frac{1}{2}(1+\gamma)$$

7. Del paso 6 y la tabla V, (de la distribución t), se encuentra el valor de t_o.

8. La expresión del paso 5, sustituyendo T, se escribe

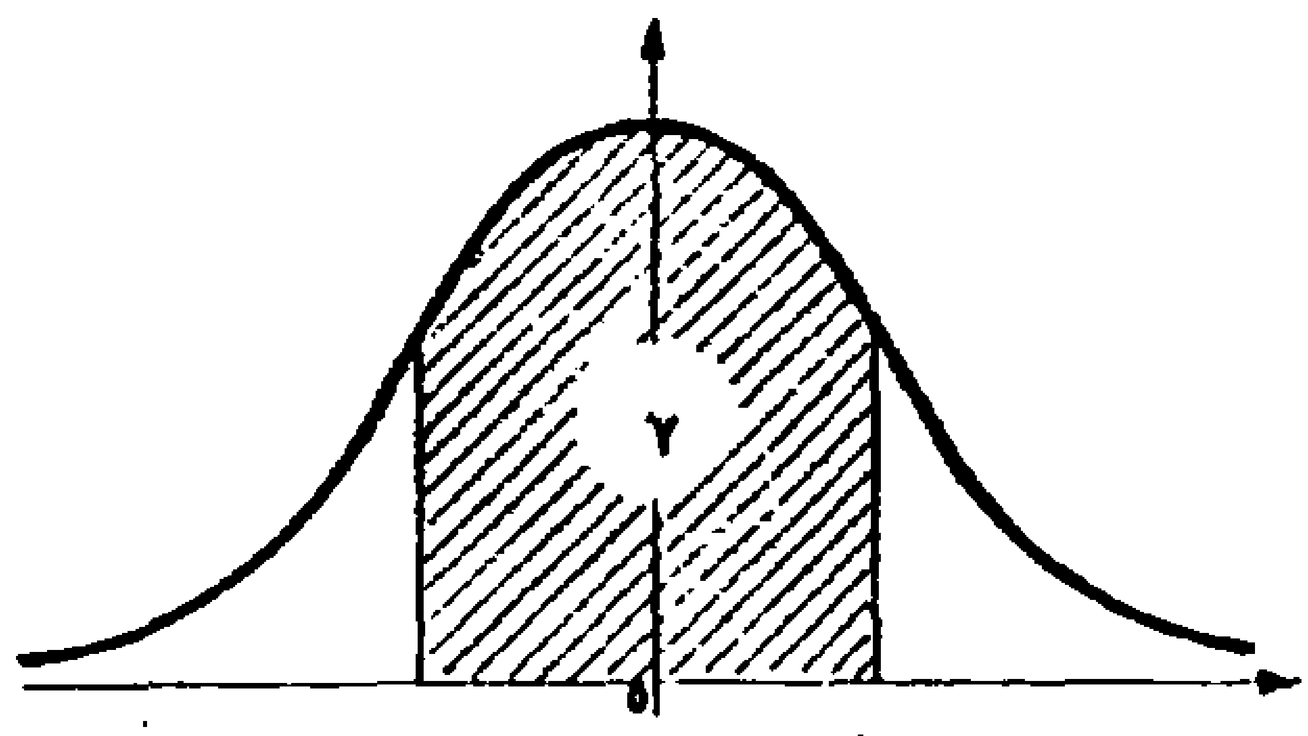


Fig. 8.3.5

$$P[-t_o \le T \le t_o] = P[-t_o \le \frac{\overline{X}-\mu}{S}\sqrt{n} \le t_o] = \gamma$$

$$= P[-\frac{t_oS}{\sqrt{n}} \le \overline{X}-\mu \le \frac{t_oS}{\sqrt{n}}] = \gamma$$

$$= P[-\overline{X}-\frac{t_oS}{\sqrt{n}} \le -\mu \le -\overline{X}+\frac{t_oS}{\sqrt{n}}] = \gamma$$

$$= P[\overline{X}-\frac{t_oS}{\sqrt{n}} \le \mu \le \overline{X}+\frac{t_oS}{\sqrt{n}}] = \gamma$$

Haciendo $\theta_1 = \overline{X} - \frac{t_oS}{\sqrt{n}}$, $\theta_2 = \overline{X} + \frac{T_oS}{\sqrt{n}}$

Obtenemos el intervalo aleatorio $[\overline{X} - \frac{t_oS}{\sqrt{n}}, \overline{X} + \frac{t_oS}{\sqrt{n}}]$ del $100\gamma\%$ de confianza para μ

9. Para un valor particular de la muestra $x_1, x_2, \ldots, x_n$, sustituyendo el valor $\bar{x}$ de $\overline{X}$ y s de S en el intervalo aleatorio, se obtiene el intervalo del $100\gamma\%$ de confianza para .

$$[\bar{x} - \frac{t_os}{\sqrt{n}} , \bar{x} + \frac{t_os}{\sqrt{n}}]$$

Este resultado puede sintetizarse en el siguiente teorema.

TEOREMA 83.5 INTERVALO DE CONFIANZA PARA μ, σ^2 DESCONOCIDO y $n < 30$. Si $\bar{x}$ y s son la media y la desviación típica de una muestra de tamaño n ($n < 30$) de una población aproximadamente normal con media μ (desconocida) y varianza σ^2 (desconocida), entonces

$$[\bar{x} - \frac{t_o s}{\sqrt{n}} \quad , \quad \bar{x} + \frac{t_o s}{\sqrt{n}}]$$

es el intervalo del 100 γ% de confianza para μ; donde t_o, es el valor de la distribución t, con n - 1 grados de libertad, determinado por

$$P[T \leq t_o] = \frac{1}{2}(1 + \gamma)$$

El procedimiento para encontrar el intervalo de confianza para μ con σ^2 desconocido puede resumirse en el siguiente cuadro.

1. Se escoge un nivel de confianza γ
2. Se determina t_o, de la ecuación

   $$P[T \leq t_o] = \frac{1}{2}(1 + \gamma)$$

   Con la tabla V de la distribución t con n - 1 grados de libertad.
3. Se calcula la media $\bar{x}$ y varianza s^2 de la muestra
4. Se calcula $\frac{t_o s}{\sqrt{n}}$.
5. Calcular $\bar{x} - \frac{t_o s}{\sqrt{n}}$, $\bar{x} + \frac{t_o s}{\sqrt{n}}$
6. Escribir el intervalo de confianza del 100 γ% para

   $$[\bar{x} - \frac{t_o s}{\sqrt{n}} \quad , \quad \bar{x} + \frac{t_o s}{\sqrt{n}}]$$

NOTA 1 Cuando el muestreo es sin reemplazo en poblaciones finitas, de acuerdo con la Nota 3 de 8.3.1, el intervalo del 100 γ % de confianza para la media poblacional es

$$[\bar{X} - t_o \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \quad , \quad \bar{X} + t_o \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}]$$

NOTA 2 Teniendo en cuenta la Nota 4 de 8.3.1, la estimación del intervalo del 100 γ % de confianza del total de la población es

$$[N\bar{X} - t_o \frac{NS}{\sqrt{n}} \quad , \quad N\bar{X} + t_o \frac{NS}{\sqrt{n}}]$$

Si el muestreo es sin reemplazo en poblaciones finitas, la estimación del intervalo de confianza del total es

$$[N\bar{X} - t_0 \frac{NS}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \quad , \quad \bar{X} + t_0 \frac{NS}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}]$$

EJEMPLO 17 El diámetro final de un cable eléctrico blindado es distribuído normalmente. Una muestra de tamaño 20 produce una media 0.790 y una desviación típica muestral 0.010. Encuentre un intervalo de confianza del 95 % para μ.

SOLUCION 1. $\gamma = 0.95$.

2. $P[T \le t_o] = \frac{1 + 0.95}{2} = \frac{1.95}{2} = 0.975$

grados de libertad 19. Luego, $t_o = 2.093$.

3. $\bar{x} = 0.790$, $s = 0.010$.

4. $\frac{t_o s}{\sqrt{n}} = \frac{(2.093)(0.01)}{\sqrt{20}} = \frac{0.02093}{4.472136} = 0.00468$.

$$\theta_1 = 0.790 - 0.00468 = 0.78532$$
$$\theta_2 = 0.790 + 0.00468 = 0.79468$$

por lo tanto [0.7853 , 0.795] es un intervalo con 95 % de confianza para la media μ.

EJEMPLO 18 A un laboratorio de ensayo de materiales se lleva una muestra de 10 cables para obtener sus cargas de rotura a la tracción. Los resultados obtenidos (en kg/cm^2) fueron 280, 295, 308, 320, 265, 350, 300, 310, 285, 310. Considerando que estas cargas poseen distribución de probabilidad normal, determinar el intervalo de confianza de 90 % para la media de la población.

SOLUCION 1 $\gamma = 0.90$.

2. De $P[T \le t_o] = \frac{1 + 0.90}{2} = 0.95$ y la tabla V de la distribución t, con 9 grados de libertad, $t_o = 1.833$.

3. Cálculo de $\bar{x}$ y s. Del siguiente cuadro

x_i	f_i	$f_i x_i$	$f_i(x_i - \bar{x})^2$
280	1	280	497.29
295	1	295	53.29
308	1	308	32.49
320	1	320	313.29
265	1	265	1391.29
350	1	350	2275.29
300	1	300	5.29
310	2	620	118.58
285	1	285	299.29
	10	3023	4986.10

$$\bar{x} = \frac{\Sigma f_i x_i}{n} = \frac{3023}{10} = 302.3$$

$$s^2 = \frac{\Sigma f_i(x_i - \bar{x})^2}{n - 1}$$

$$= \frac{4986.1}{9} = 554.011$$

luego, $s = 23.537$

4. $$\frac{t_o s}{\sqrt{n}} = \frac{(1.833)(23.537)}{\sqrt{10}} = \frac{43.143}{3.162} = 13.644$$

por lo tanto, $\theta_1 = 302.3 - 13.644 = 288.656$

$\theta_2 = 302.3 + 13.644 = 315.944$

Es decir, el intervalo con 90 % de confainza para μ es, [288.656 , 315.944]

EJEMPLO 19 Un ingeniero mecánico cree haber perfeccionado un programa de adiestramiento que puede reducir considerablemente el tiempo de montaje de ciertos mecanismos empleado por los trabajadores. Para comprobar esta creencia, escoge 10 trabajadores al azar y realiza estudios de tiempo antes y después del adiestramiento, obteniendo como resultado las siguientes reducciones de tiempo en minutos,

Trabajador	1	2	3	4	5	6	7	8	9	10
reducción de tiempo	3	5	-1	0	7	4	8	3	-1	2

(a) Halle un intervalo de confianza del 95 % para la media de la reducción de tiempo de los primeros 5 trabajadores.

(b) Halle un intervalo de confianza del 95 % para la media de la reducción de tiempo para los diez trabajadores.

SOLUCION (a) Asumiendo que la población tiene distribución normal.

1. $\gamma = 0.95$.

2. $P[T \leq t_o] = \dfrac{1 + 0.95}{2} = 0.975$,

de la tabla de la distribución t con $n - 1 = 4$ grados de libertad obtenemos $t_o = 2.776$.

3. Cálculo de $\bar{x}$ y s. Del cuadro siguiente:

x_i	$(x_i - \bar{x})^2$
3	0.04
5	4.84
-1	14.44
0	7.84
7	17.64
14	44.8

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{14}{5} = 2.8$$

$$s^2 = \frac{\Sigma (x_i - \bar{x})^2}{n - 1} = \frac{44.8}{4}$$

$$s^2 = 11.2$$

luego, $s = 3.347$,

4. $$\frac{t_o s}{\sqrt{n}} = \frac{(2.776)(3.347)}{\sqrt{5}} = \frac{9.291}{2.236} = 4.155$$

por lo tanto $\theta_1 = 2.8 - 4.155 = -1.355$

$\theta_2 = 2.8 + 4.155 = 6.955$

Es decir, el intervalo de confianza del 95% para μ es $[-1.355 , 6.955]$

(b) 1. $\gamma = 0.95$.

2. De $P[T \leq t_o] = \dfrac{1 + 0.95}{2} = 0.975$

y de la tabla de la distribución t con $n - 1 = 9$ grados de libertad obtenemos $t_o = 2.262$.

3. Cálculo de $\bar{x}$ y s .

x_i	f_i	$f_i x_i$	$(x_i - \bar{x})^2$
3	2	6	0
5	1	5	4
-1	2	-2	32
0	1	0	9
7	1	7	16
4	1	4	1
8	1	8	25
2	1	2	1
	10	30	88

$$\bar{x} = \frac{\Sigma f_i x_i}{n} = \frac{30}{10} = 3$$

$$s^2 = \frac{\Sigma f_i (x_i - \bar{x})^2}{n - 1}$$

$$s^2 = \frac{88}{9} = 9.778$$

y $s = 3.127$.

4. $$\frac{t_o s}{\sqrt{n}} = \frac{(2.262)(3.127)}{\sqrt{10}} = \frac{7.073}{3.162} = 2.237$$

luego $\theta_1 = 3 - 2.237 = 0.763$

$\theta_2 = 3 + 2.237 = 5.237$

Por lo tanto, el intervalo de confianza del 95 % para μ es

$$[0.763 \ , \ 5.237] \ .$$

EJEMPLO 20 Los ingresos del impuesto sobre ventas en una ciudad que tiene un total de 300 establecimientos comerciales se recogen cada trimestre. Los siguientes datos representan los ingresos (en miles de intis) cobrados durante el primer trimestre en una muestra de nueve establecimientos comerciales

16, 18, 11, 17, 13, 10, 22, 15, 16

(a) Determinar una estimación del intervalo con 95% de confianza de los ingresos trimestrales del impuesto sobre ventas en los establecimientos comerciales.

(b) Determine una estimación del intervalo con 95% de confianza de los ingresos totales por impuesto sobre ventas que recogerán este trimestre.

SOLUCION 1 $\gamma = 0.95$

2. De $P[T \le t_o] = \dfrac{1 + 0.95}{2} = 0.975$ y la tabla V de la distribución t con $n - 1 = 8$ grados de libertad, $t_o = 2.3060$.

3. Cálculo de $\bar{x}$ y s. De la tabla

x_i	f_i	$f_i x_i$	$f_i(\bar{x}_i - \bar{x})^2$
10	1	10	28444408.880
11	1	11	18777748.880
13	1	13	5444428.888
15	1	15	111108.8889
16	2	32	888897.7778
17	1	17	2777788.888
18	1	18	7111128.888
22	1	22	44444488.88
	9	138	107999999.9

$$\bar{X} = \frac{\Sigma f_i x_i}{n} = \frac{138000}{9} = 15333.33$$

$$S = \sqrt{\Sigma f_i(x_i - \bar{x})^2/(n-1)}$$

$$= \sqrt{\frac{107999999.9}{8}}$$

$$= \sqrt{13499999.98}$$

$$= 3674.2346$$

4. Cálculo de $t_o \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$, donde N = 300 .

$$t_o \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{(2.306)(3674.23)}{3} \sqrt{\frac{291}{299}} = \frac{8472.77438}{3} \times \frac{17.0587221}{17.29161646}$$

$$= (2824.258126)(0.986531371)$$

$$= 2786.22$$

5. $\bar{X} - t_o \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = 15333.33 - 2786.22 = 12547.11$

$$\bar{X} + t_o \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = 15333.33 + 2786.22 = 18119.55$$

(a) El intervalo del 95% de confianza para los ingresos trimestrales del impuesto sobre rentas en los establecimientos comerciales es [12547.11 , 18119.55] .

(b) Por la Nota 2, el intervalo del 95% de confianza de los ingresos totales por impuesto sobre ventas para el trimestre se obtiene multiplicando por N los límites superiores e inferior del intervalo anterior.

$$300(12547.11) = 3764133$$
$$300(18119.55) = 5435865$$

o sea $3764133 \leq \tau \leq 5435865$.

8.3.9 INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE MEDIAS CON VARIANZAS DESCONOCIDAS PERO IGUALES, MUESTRAS PEQUEÑAS

Sea X una variable aleatoria, distribuída normalmente con media μ_X y varianza σ^2(desconocida). Sea Y una variable aleatoria independiente (de X) distribuída normalmente con media μ_Y y varianza σ^2 (desconocida) . Sea $\bar{X}$ la media muestral de una muestra aleatoria de n ($n < 30$) observaciones de X y sea $\bar{Y}$ la media muestral de una muestra aleatoria de m ($m < 30$) observaciones de Y. En el capítulo anterior hemos visto que la variable aleatoria

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n} + \frac{1}{m}}\sqrt{\frac{(n - 1)S_X^2 + (m - 1)S_Y^2}{n + m - 2}}}$$

tiene una distribución t con $n + m - 2$ grados de libertad. Desde que la distribución de la variable aleatoria T no depende de $\mu_X - \mu_Y$; elegido γ podemos encontrar t_o , tal que:

$$P[- t_o \leq T \leq t_o] = \gamma \qquad (1)$$

por la simetría de la curva de la distribución T, (1) se escribe

$$P[T \leq t_o] = \frac{1}{2}(1 + \gamma) \qquad (2)$$

de (2) usando la tabla V, se encuentra el valor de t_o . Si denotamos por S_c la desviación típica combinada de las dos muestras, es decir

$$S_c = \sqrt{\frac{(n - 1)S_X^2 + (m - 1)S_Y^2}{n + m - 2}}$$

la expresión (1) sustituyendo la variable aleatoria T se escribe como sigue

$$\gamma = P[- t_o \leq T \leq t_o] = P\left[- t_o \leq \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n} + \frac{1}{m}}\; S_c} \leq t_o\right]$$

$$= P\left[- t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \leq (\bar{X} - \bar{Y}) - (\mu_X - \mu_Y) \leq t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}}\right]$$

$$= P\left[-(\bar{X} - \bar{Y}) - t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \leq -(\mu_X - \mu_Y) \leq -(\bar{X} - \bar{Y}) + t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \right]$$

$$= P\left[(\bar{X} - \bar{Y}) - t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \leq \mu_X - \mu_Y \leq (\bar{X} - \bar{Y}) + t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \right]$$

Haciendo

$$\theta_1 = (\bar{X} - \bar{Y}) - t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \qquad y$$

$$\theta_2 = (\bar{X} - \bar{Y}) - t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}}$$

Se obtiene el intervalo aleatorio

$$\left[(\bar{X} - \bar{Y}) - t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \;,\; (X - Y) + t_o S_c \sqrt{\frac{1}{n} + \frac{1}{m}} \right]$$

que con probabilidad γ contiene a la diferencia de medias $\mu_X - \mu_Y$.

Para dos muestras aleatorias independientes de tamaño n y m extraídas de dos poblaciones cuyas varianzas σ_X^2 y σ_Y^2 son desconocidas pero iguales ($\sigma_X^2 = \sigma_Y^2$); sea $\bar{x}$ e $\bar{y}$ las medias de las muestras observadas.

El intervalo de confianza del 100 γ% para $\mu_X - \mu_Y$, esta dado por

$$\left[(\bar{x} - \bar{y}) - t_o s_c \sqrt{\frac{1}{n} + \frac{1}{m}} \;,\; (\bar{x} - \bar{y}) + t_o s_c \sqrt{\frac{1}{n} + \frac{1}{m}} \right]$$

El teorema siguiente resume el resultado anterior

TEOREMA 8.3.6 INTERVALO DE CONFIANZA PARA $\mu_X - \mu_Y$ CON σ_X^2 y σ_Y^2 DESCONOCIDOS PERO $\sigma^2 = \sigma_X^2 = \sigma_Y^2$. Si $\bar{x}$ e $\bar{y}$ son las medias de dos muestras aleatorias independientes de tamaños n y m, respectivamente, extraídas de dos poblaciones distribuídas - aproximadamente normal con medias (desconocidas) μ_X, μ_Y y varianzas desconocidas pero iguales $\sigma_X^2 = \sigma_Y^2$. Entonces

$$\left[(\bar{x} - \bar{y}) - t_o s_c \sqrt{\frac{1}{n} + \frac{1}{m}} \;,\; (\bar{x} - \bar{y}) + t_o s_c \sqrt{\frac{1}{n} + \frac{1}{m}} \right]$$

es el intervalo de confianza del 10 γ% para la diferencia de medias $\mu_X - \mu_Y$.

Donde s_c es la desviación típica muestral común y t_o es el valor de la distribución t_o con $n + m - 2$ grados de libertad, determinado por

$$P[T \leq t_0] = \frac{1}{2}(1+\gamma)$$

El procedimiento para encontrar el intervalo de confianza para $\mu_X - \mu_Y$ con σ_X^2 y σ_Y^2 desconocidas pero iguales. Es decir $\sigma_X^2 = \sigma_Y^2$; se resume en el cuadro siguiente,

1. Se escoge el nivel de confianza γ
2. Se determina t_0 de la ecuación
$$P[T \leq t_0] = \frac{1}{2}(1+\gamma)$$
y la tabla V, de la distribución t, con $n + m - 2$ grados de libertad.
3. Cálculo de las medias $\bar{x}$, $\bar{y}$ muestrales observadas y la desviación típica muestral combinada s_c
4. Cálculo de $t_0 s_c \sqrt{\frac{1}{n} + \frac{1}{m}}$.
5. Se escribe el intervalo de confianza.
$$\left[(\bar{x} - \bar{y}) - t_0 s_c \sqrt{\frac{1}{n} + \frac{1}{m}} \ , \ (\bar{x} - \bar{y}) + t_0 s_c \sqrt{\frac{1}{n} + \frac{1}{m}}\right]$$

EJEMPLO 21 Las siguientes son 16 determinaciones independientes del punto de fusión de un compuesto, ocho hechas por un analista A y ocho hechas por un analista B. Los datos se dan en grados centígrados.

Analista A(X)	164.4, 169.7, 169.2, 169.5, 161.8, 168.7, 169.5, 163.9
Analista B(Y)	163.5, 162.8, 163.0, 163.2, 160.7, 161.5, 160.9, 162.1

Hallar un intervalo de confianza del 90% para la diferencia media entre analistas, suponiendo varianzas iguales pero desconocidas.

SOLUCION 1. $\gamma = 0.90$.

2. $P[T \leq t_0] = \frac{1}{2}(1 + 0.90) = 0.95$

y la tabla V de la distribución t, con $n + m - 2 = 8 + 8 - 2 = 14$ grados de li_ bertad obtenemos $t_0 = 1.761$.

3.

x_i	$(x_i - \bar{x})^2$
164.4	7.22265625
169.7	6.82515625
169.2	4.46265625
169.5	5.82015625
161.8	27.95765625
168.7	2.60015625
169.5	5.82015625
163.9	10.16015625
1336.7	70.86875

y_i	$(y_i - \bar{y})^2$
163.5	1.65765625
162.8	0.34515625
163.0	0.62015625
163.2	0.95515625
166.7	2.28765625
161.5	0.50765625
160.9	1.72265625
162.1	0.01265625
1297.7	8.12875

$$\bar{x} = \frac{\Sigma x_i}{8} = \frac{1336.7}{8} = 167.0875 ,$$

$$\bar{y} = \frac{\Sigma y_i}{8} = \frac{1297.7}{8} = 162.2125 ,$$

luego $\bar{x} - \bar{y} = 167.0875 - 162.2125 = 4.875$.

y $$S_c = \sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}} = \sqrt{\frac{\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2}{n+m-2}}$$

$$S_c = \sqrt{\frac{70.86875 + 8.12875}{14}} = 2.3754 .$$

4. $$t_0 S_c \sqrt{\frac{1}{n} + \frac{1}{m}} = (1.76)(2.3754)\sqrt{\frac{1}{8} + \frac{1}{8}} = 2.0915 ,$$

luego $\theta_1 = 4.875 - 2.0915 = 2.7835$

$\theta_2 = 4.875 + 2.0915 = 6.9665$

por lo tanto, [2.78 , 6.97] es un intervalo de confianza del 90 % para la diferencia media entre analistas.

8.3.10 INTERVALO DE CONFIANZA PARA LA VARIANZA

Sea **X** una variable aleatoria distribuída normalmente con media μ y varianza

σ^2 desconocidos. Consideremos una muestra aleatoria $X_1, X_2, \ldots, X_n$, de tamaño n de X. Sea $\overline{X}$ la media muestral y S^2 la varianza muestral. En el teorema 7.3.5.b hemos visto, que la variable aleatoria,

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2}$$

tiene una distribución chi-cuadrado con $n - 1$ grados de libertad. Observe que si bien la variable aleatoria χ^2 depende de σ^2, su distribución no; por lo tanto, elegido un γ, se puede encontrar a y b de la tabla IV, tal que

$$P[a \leq \chi^2 \leq b] = \gamma \qquad (1)$$

ó $$P[\chi^2 \leq b] - P[\chi^2 < a] = \gamma$$

una manera de hacer esto, es escogiendo a y b tal que

$$P[\chi^2 < a] = \alpha/2, \qquad (2)$$

y $$P[\chi^2 \leq b] = 1 - \alpha/2 \qquad (3)$$

donde $\alpha = 1 - \gamma$. Entonces de acuerdo a la notación dada en 7.3.1, $a = \chi^2_{\alpha/2}$, $b = \chi^2_{1-\alpha/2}$ (Ver Fig. 8.3.6) .

Observe que, restando (2) de (3) obtenemos la expresión (1)

$$P[a \leq \chi^2 \leq b] = P[\chi^2 \leq b] - P[\chi^2 < a]$$

$$= 1 - \frac{\alpha}{2} - \frac{\alpha}{2}$$

$$= 1 - \alpha$$

$$= \gamma$$

ó $$P[\chi^2_{\alpha/2} \leq \chi^2 \leq \chi^2_{1-\alpha/2}] =$$

$$= \gamma = 1 - \alpha \qquad (4)$$

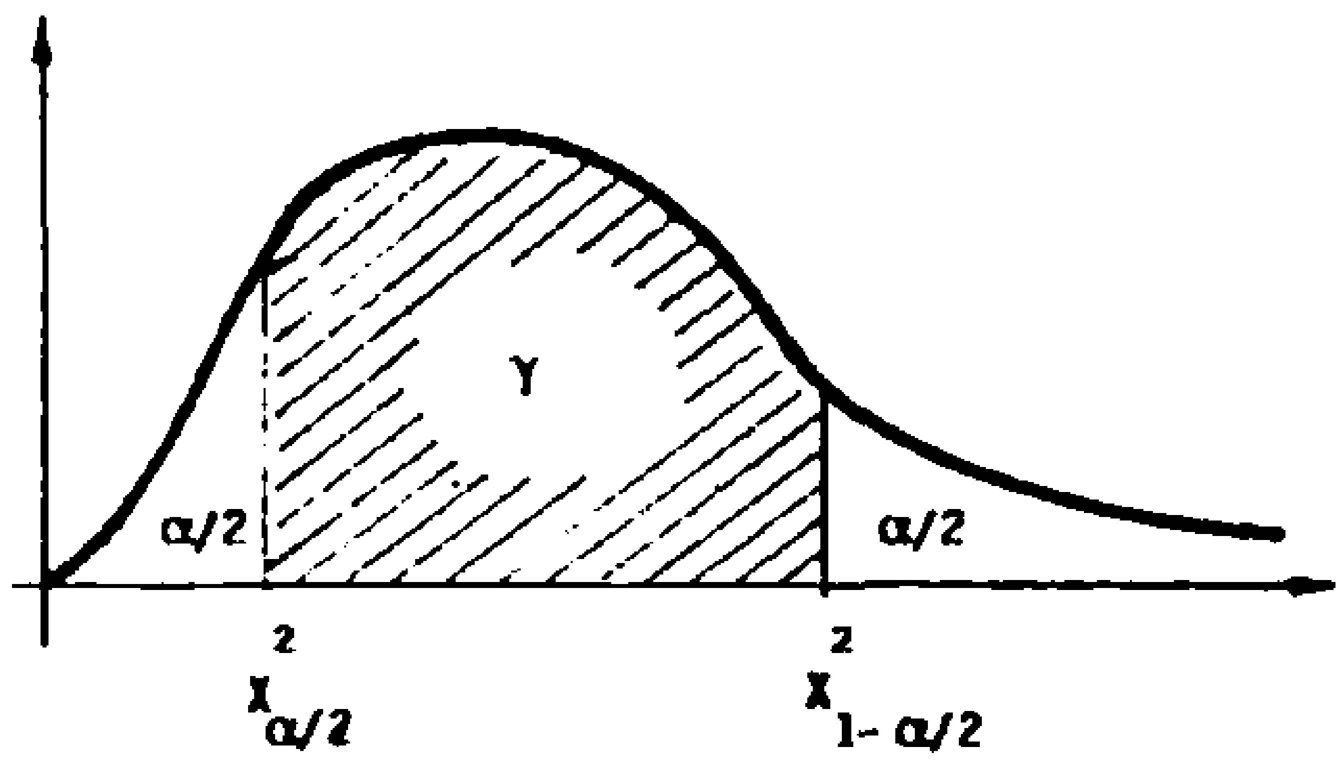


Fig. 8.3.6

donde $\chi^2_{\alpha/2}$ y $\chi^2_{1-\alpha/2}$ son valores de chi-cuadrado con $n - 1$ grados de libertad - tales que $P[\chi^2 \leq \chi^2_{\alpha/2}] = \alpha/2$ y $P[\chi^2 \leq \chi^2_{1-\alpha/2}] = 1 - \alpha/2$. Por otro lado, en la expresión (4) reemplazando χ^2 por $\dfrac{(n-1)S^2}{\sigma^2}$, tenemos,

$$P[\chi^2_{\alpha/2} \le \chi^2 \le \chi^2_{1-\alpha/2}] = P[\chi^2_{\alpha/2} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{1-\alpha/2}]$$

$$= P\left[\frac{1}{\chi^2_{1-\alpha/2}} \le \frac{\sigma^2}{(n-1)S^2} \le \frac{1}{\chi^2_{\alpha/2}}\right]$$

$$= P\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{\alpha/2}}\right]$$

$$= \gamma = 1 - \alpha$$

para la muestra particular de tamaño n, la varianza muestral s^2 es calculada, y el intervalo del 100γ % de confianza para σ^2 es,

$$\left[\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} , \frac{(n-1)s^2}{\chi^2_{\alpha/2}}\right]$$

TEOREMA 8.3.7 INTERVALO DE CONFIANZA PARA σ^2. Si S^2 es la varianza muestral, de una muestra de tamaño n , tomada de una población normal con media μ y varianza σ^2 (desconocida), se tiene que el intervalo del 100γ % de confianza para σ^2, está dado por

$$\left[\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} , \frac{(n-1)s^2}{\chi^2_{\alpha/2}}\right]$$

donde $\chi^2_{\alpha/2}$ y $\chi^2_{1-\alpha/2}$ son los valores de una distribución chi-cuadrado con -

$n - 1$ grados de libertad, dejando áreas de $\alpha/2$ y $1 - \alpha/2$ respectivamente, a la izquierda. (Calculadas en (2) y (3)).

CONSECUENCIA La estimación del intervalo del 100γ % de confianza para la desviación típica σ es

$$P\left[S\sqrt{\frac{n-1}{\chi^2_{1-\alpha/2}}} \le \sigma \le S\sqrt{\frac{n-1}{\chi^2_{\alpha/2}}}\right] = \gamma = 1 - \alpha$$

EJEMPLO 22 En el ejemplo 18, determinar el intervalo de confianza del 90 % para la desviación típica de la población

SOLUCION 1 $\gamma = 0.9$ y tamaño de la muestra, $n = 10$

2. $P[\chi^2 \leq \chi^2_{\alpha/2}] = \alpha/2 = \dfrac{1-0.9}{2} = 0.05$

y $P[\chi^2 \leq \chi^2_{1-\alpha/2}] = 1 - \alpha/2 = 1 - 0.05 = 0.95$

de la tabla IV con $n-1 = 9$ grados de libertad, obtenemos

$\chi^2_{0.05} = 3.325$, $\chi^2_{0.95} = 16.92$

3. $S^2 = 554.011$, luego $(n-1)S^2 = 9(554.011) = 4986.099$

4. $$\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} = \frac{4986.099}{16.92} = 294.6867$$

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} = \frac{4986.099}{3.325} = 1499.5786 \quad ;$$

el intervalo del 90 % de confianza para la varianza de población es,

[294.6867 , 1499.5786], y por la consecuencia del teorema anterior

el intervalo del 90 % de confianza para desviación típica de la población es

[17.1664, 38.7244] .

8.3.11 INTERVALO DE CONFIANZA PARA LA RAZON DE DOS VARIANZAS

Sea X una variable aleatoria distribuída normalmente con media μ_X y varianza σ_Y^2, y sea Y una variable aleatoria independiente (de X) distribuída normalmente con media μ_Y y varianza σ_Y^2. Sea S_X^2 la varianza muestral basada en una muestra aleatoria de tamaño n de X y sea S_Y^2 la varianza muestral basada en una muestra aleatoria de tamaño m de Y. En 7.3.7, hemos visto que la variable aleatoria,

$$F = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} = \frac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2}$$

Tiene una distribución F con $n-1$ y $m-1$ grados de libertad. Desde que la distribución de la variable aleatoria F no depende de σ_X^2 y σ_Y^2 , elegido γ podemos encontrar c y d, de la tabla VI tal que,

$$P[c \leq F \leq d] = \gamma \qquad (1)$$

$$P[F \leq d] - P[F \leq c] = \gamma$$

escogiendo, c y d tal que

$$P[F < c] = \alpha/2 \qquad (2)$$

y
$$P[F < d] = 1 - \alpha/2 \qquad (3)$$

donde $\alpha = 1 - \gamma$ ó $\gamma = 1 - \alpha$ (Ver Fig. 8.3.7)

o sea $c = f_{\alpha/2, n-1, m-1}$ y $d = f_{1-\alpha/2, n-1, m-1}$ que se obtienen de la tabla VI y las fórmulas (2) y (3) respectivamente con $n - 1$ y $m - 1$ grados de libertad. Es decir, (1) se puede escribir,

$$P[c \leq F \leq d] = P[f_{\alpha/2, n-1, m-1} \leq F \leq f_{1-\alpha/2, n-1, m-1}] = \gamma = 1 - \alpha$$

donde $f_{\alpha/2}$ y $f_{1-\alpha/2}$ son valores de la distribución F con $n - 1$ y $m - 1$ grados de libertad, dejando áreas bajo la curva de $\alpha/2$ y $1 - \alpha/2$ respectivamente a la izquierda; por otra parte en (1) sustituyendo F por $\dfrac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2}$, se escribe

$$P[f_{\alpha/2, n-1, m-1,} \leq F \leq f_{1,\alpha/2, n-1, m-1}] =$$

$$= P\left[f_{\alpha/2, n-1, m-1} \leq \frac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2} \leq f_{1-\alpha/2, n-1, m-1}\right]$$

$$= P\left[\frac{1}{f_{1-\alpha/2, n-1, m-1}} \leq \frac{\sigma_X^2 S_Y^2}{\sigma_Y^2 S_X^2} \leq \frac{1}{f_{\alpha/2, n-1, m-1}}\right]$$

$$= P\left[\frac{S_X^2}{S_Y^2} \cdot \frac{1}{f_{1-\alpha/2, n-1, m-1}} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{S_X^2}{S_Y^2} \cdot \frac{1}{f_{\alpha/2, n-1, m-1}}\right] = \gamma$$

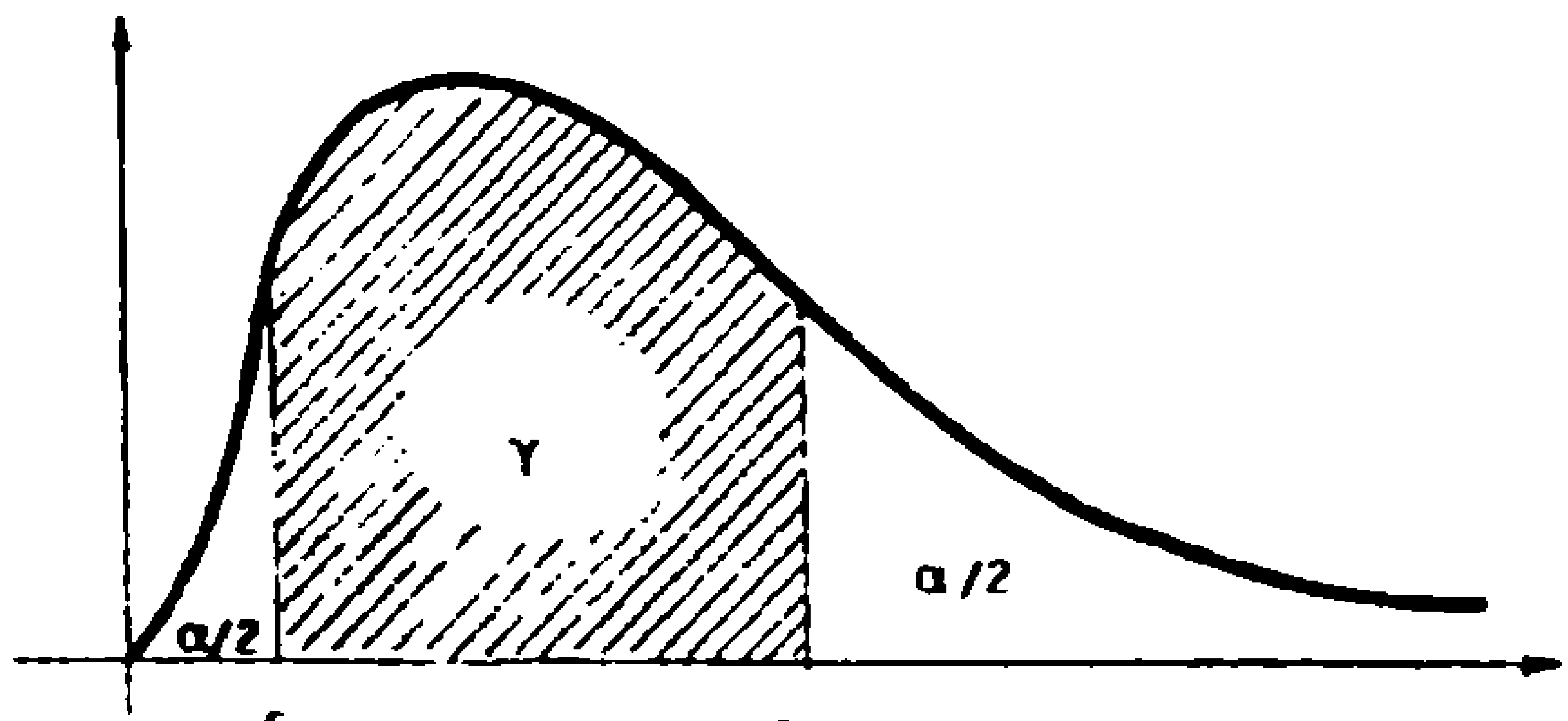


Fig. 8.3.7

reemplazando $f_{\alpha/2, n-1, m-1}$ por $\dfrac{1}{f_{1-\alpha/2,\, m-1,\, n-1}}$ tenemos

$$P\left[\frac{S_X^2}{S_Y^2} \frac{1}{f_{1-\alpha/2, n-1, m-1}} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{S_X^2}{S_Y^2} \cdot f_{1-\alpha/2, m-1, n-1} \right] = \gamma$$

y para las dos muestras independientes de tamaño n y m, seleccionadas de las dos poblaciones normales, la razón de las varianzas S_X^2/S_Y^2, se conoce. Entonces, el intervalo del $100\gamma\%$ σ_X^2/σ_Y^2 está dado por;

$$\left[\frac{S_X^2}{S_Y^2} \frac{}{f_{1-\alpha/2, n-1, m-1}} \ , \ \frac{S_X^2}{S_Y^2} f_{1-\alpha/2, m-1, n-1} \right]$$

TEOREMA 8.3.8 INTERVALO DE CONFIANZA PARA σ_X^2/σ_Y^2. Si S_X^2, S_Y^2 son las varianzas de muestras independientes de tamaños n y m de dos poblaciones X e Y distribuídos normalmente con media μ_X, μ_Y y varianzas σ_X^2 y σ_Y^2 respectivamente. El intervalo del $100\gamma\%$ de confianza para σ_X^2/σ_Y^2 está dado por,

$$\left[\frac{S_X^2}{S_Y^2} \cdot \frac{1}{f_{1-\alpha/2, n-1, m-1}} \ , \ \frac{S_X^2}{S_Y^2} f_{1-\alpha/2, m-1, n-1} \right]$$

donde $f_{1-\alpha/2, n-1, m-1}$ es el valor de la distribución F con $n - 1$ y $m - 1$ grados de libertad dejando una área de $1 - \alpha/2$ a la izquierda y similarmente

$f_{1-\alpha/2,m-1,n-1}$ es un valor f con $m-1$ y $n-1$ grados de libertad. Es decir

$$P[F \le f_{1-\alpha/2}] = 1 - \alpha/2$$

CONSECUENCIA El intervalo del 100 γ% de confianza para la razón de las desviaciones estándar σ_X/σ_Y está dado por,

$$\left[\frac{S_X}{S_Y}\sqrt{\frac{1}{f_{1-\alpha/2,n-1,m-1}}}\ ,\ \frac{S_X}{S_Y}\sqrt{f_{1-\alpha/2,m-1,n-1}} \right]$$

EJEMPLO 23 Se preparan dos diferentes tandas de cemento, y de cada una se fabrica gran número de tabiques. Se toma una muestra de seis tabiques de cada tanda, y la fuerza de tensión en lb/pulg^2, se mide para cada tabique de las dos muestras. Suponiendo que en la tanda i la fuerza de tensión está normalmente distribuída con varianza σ_i^2, $i = 1,2$. Obtenga un intervalo de confianza de 95% para la razón de las dos desviaciones típicas, σ_1/σ_2.

Tanda 1: 536, 492, 528, 572, 582, 506

Tanda 2: 555, 567, 550, 550, 535, 540

SOLUCION 1 $\alpha = 1 - \gamma = 1 - 0.95 = 0.05$; luego,

$1 - \alpha/2 = 0.975$

$n = m = 6$, por lo tanto $n - 1 = m - 1 = 5$; entonces se tiene

$f_{0.975,5,5} = 7.15$ (tabla V) (en este caso,

$f_{1-\alpha/2,n-1,m-1} = f_{1-\alpha/2,m-1,n-1}$)

2. tanda 1

x_i	$(x_i - \bar{x}_1)^2$
536	0
492	1936
528	64
572	12.96
582	2116
506	900
3216	6312

; tanda 2

x_i	$(x_i - \bar{x}_2)^2$
555	30.25
567	306.25
550	0.25
550	0.25
535	210.25
540	90.25
3297	637.5

$$\bar{x}_1 = \frac{3216}{6} = 536, \quad S_1^2 = 1262.4 \ ; \quad \bar{x}_2 = \frac{3297}{6} = 549.5,$$

$$S_2^2 = \frac{637.5}{5} = 127.5$$

3. $$\frac{S_1}{S_2} = \sqrt{\frac{S_1^2}{S_2^2}} = \sqrt{\frac{1262.4}{127.5}} = \sqrt{9.90} = 3.15 \qquad y$$

$$\sqrt{\frac{1}{f_{0.975,55}}} = \sqrt{\frac{1}{7.15}} = \sqrt{0.14} = 0.37$$

$$\sqrt{f_{0.975;55}} = \sqrt{7.15} = 2.67$$

4. $$\frac{S_1}{S_2}\sqrt{\frac{1}{f_{1-\alpha/2,\, n-1,\, m-1}}} = (3.15)(0.37) = 1.1655 = 1.2$$

$$y \quad \frac{S_1}{S_2}\sqrt{f_{1-\alpha/2,\, m-1,\, n-1}} = (3.15)(2.67) = 8.4105 = 8.4$$

por lo tanto, el intervalo de confianza de 95 % para σ_1/σ_2 es, $[1.2 \ , \ 8.4]$

PROBLEMAS 8.3

1. Sea X una variable aleatoria, uniformemente distribuído en el intervalo $[0,10]$. Determinar la probabilidad de que el intervalo aleatorio $\left\langle \frac{X}{2} \ , \ 2X \right\rangle$ incluya a la media de la variable aleatoria X.
2. Un ingeniero de la planta de purificación de agua mide el contenido de cloro diariamente en 100 muestras diferentes. Sobre un período de años, ha establecido que la desviación típica (error estándar) de la población es 1.2 miligramos de cloro por litro. Las ultimas muestras promediaron 4.8 miligramos de cloro por litro. Establezca el intervalo alrededor de la media muestra que incluirá la media de la población en un 68.7% de las veces.

3. A partir de una muestra aleatoria de 144 galones de leche, el gerente de procesos de la planta, calculó que el llenado medio es de 128.4 onzas líquidas. De acuerdo a las especificaciones del fabricante, la máquina llenadora tiene una desviación típica de 0.6 onzas. ¿Cuál es el intervalo alrededor de la media muestral que contendrá la media de la población en un 95.5% de las veces?

4. El equipo de control de calidad de una industria muestrea en forma rutinaria la línea de producción de determinado artículo y diariamente calcula un intervalo de confianza del 90% para la longitud media de las piezas producidas en el día. Se han calculado 100 intervalos de confianza del 90% despues de 100 días.
   (a) Sea **X** la variable aleatoria que representa el número de los intervalos que en efecto cubren la longitud media desconocida de las piezas producidas en el día. ¿Cuál es la distribución de probabilidad de **X**?
   (b) Calcular la probabilidad aproximada que exactamente 90 de los 100 intervalos cubran la media verdadera.

5. Una muestra aleatoria de 144 observaciones arroja una media muestral $\bar{x} = 160$ y una varianza muestral $s^2 = 100$
   (a) Determinar un intervalo de confianza del 95% para la media poblacional.
   (b) Determinar un intervalo de confianza de 90% para μ.
   (c) Si el investigador quiere tener un 95% de "seguridad" que su estimación se encuentre a una distancia de 1.2 en más o menos de la verdadera media poblacional, ¿cuántas observaciones adicionales deberá efectuar para corroborarlo?

6. Se toma una muestra al azar de 45 alumnos, sin reposición de una clase de estadística de 221 alumnos que da una media de 70 puntos y una desviación típica de 9 puntos en las calificaciones finales. Determinar el intervalo de confianza de 98% para la media de las 221 calificaciones.

7. Se ha medido el contenido de nicotina de 36 cigarrillos de una determinada marca. A continuación se resumen los resultados obtenidos. Sea
   x_i = contenido de nicotina de un cigarrillo, medido en miligramos.
   $\Sigma x_i = 756$ miligramos y $\Sigma (x_i - \bar{x})^2 = 315$ miligramos
   Determinar un intervalo de confianza de 95% para el contenido promedio de nicotina de los cigarrillos de esta marca.

8. Un grupo de 50 animales experimentales reciben una cierta clase de raciones por un período de 2 semanas. Sus aumentos de pesos arrojan los valores $\bar{x}$ = 420 gr. y s = 60 gr.

   (a) ¿Qué tamaño debe tomarse una muestra si se desea que $\bar{x}$ difiera de μ por menos de 10 gr. con 0.95 de probabilidad de ser correcto?

   (b) Encontrar el intervalo de confianza de 95% para μ.

9. En una granja de 1,000 pollos se va a experimentar con una nueva dieta de engorde. Si se sabe que la desviación típica del aumento de peso en un período de un mes es igual a dos onzas, ¿Qué tamaño debe tomarse una muestra que conduzca a una estimación del aumento de peso de la totalidad de la parvada, si se quiere que esta estimación no contenga un error mayor que 40 lb (una lb = 16 oz) con probabilidad de 0.95?

10. La asociación de ahorro y préstamo "La tacaña" desea determinar la cantidad promedio que tiene sus clientes en sus cuentas de ahorro. La desviación típica de todas las cuentas de ahorro es estimada por el gerente en I/. 4,000.00

    (a). ¿Qué tamaño de muestra se requiere para afirmar con confianza 0.95 que el error en la estimación no excede de I/. 200,00?

    (b) ¿Qué tamaño de muestra se requiere para afirmar con confianza 0.95 que el error en la estimación no excede de I/. 400.00?

11. Una tienda de departamentos desea estimar, con un nivel de confianza de 0.98 y un error máximo de 0.5, el verdadero valor medio en dólares de las compras a crédito por mes realizadas por sus clientes. Dado que la desviación típica es \$ 15, determinar el tamaño de la muestra.

12. Una muestra de 35 alambres de acero escogidos al azar de una remesa muy grande muestra una resistencia media a la tracción de 1,500 libras y una desviación típica de 20 libras.

    (a) Si se usa esta media de muestra para estimar la resistencia media a la tracción de toda la remesa, con un nivel de confianza de 95%, determinar el error de muestreo.

    (b) ¿Qué tamaño de muestra se requiere, si el error estimado en (a) ha de ser reducido a la mitad sin cambiar el nivel de confianza?.

13. De una orden especial de 1,500 taladros recibidos de la compañía Andina de máquinas y herramientas, se probó una muestra de 36 taladros. La muestra tuvo una vida de 1,800 horas y una desviación estándar de 150 horas.

    (a) Estime la desviación estándar de la población a partir de la desviación estándar de la muestra.

(b) Estime el error estándar de la muestra para esta población finita.

(c) Construya un intervalo de confianza de un 98% para la vida media de los taladros.

14. Para estimar la cantidad total de depósitos a la vista, un banco comercial selecciona una muestra aleatoria de 400 cuentas. La muestra da una media de I/. 5000 y una desviación típica de I/. 1000. Suponiendo que el banco tiene 12,000 cuentas a la vista, Determinar un intervalo de confianza del 99% para la cantidad total en depósitos.

15. La fábrica de calzado COMPRE-AHORA tiene una cadena de tiendas de venta al por menor en diversas ciudades del Perú. La política de COMPRE-AHORA es no establecer una tienda de ventas en ninguna ciudad a menos de tener una seguridad del 99% de que la venta total anual de calzado en la ciudad sea de por lo menos I/. 5 millones. La compañía está considerando la posibilidad de instalar una tienda de ventas en Huaral (Dpto. de Lima), que es una ciudad con 20,000 familias, para lo cual selecciona una muestra aleatoria de 49 familias, que da un gasto familiar anual medio en calzado de I/. 300 con una desviación típica de I/. 105. Con base en esta información, ¿debe COMPRE-AHORA abrir una tienda de ventas en Huaral?

16. La cooperativa de Huando, desea determinar el peso total de una partida de 10,000 naranjas. Como la cooperativa sólo tiene una balanza pequeña y además no hay tiempo para estas cosas, selecciona una muestra aleatoria de 16 naranjas, la cual da una media de 175 gramos y una desviación típica de 25 gramos. Determinar un intervalo de confianza del 95% para el peso total de la partida de naranjas.

17. El mantenimiento de cuentas de crédito puede resultar demasiado costoso si el promedio de compra por cuenta baja de cierto nivel. El gerente de un almacén desea estimar el promedio de cantidad comprada por mes por sus clientes que tienen cuenta de crédito con un error de no más de I/. 250 con una probabilidad aproximada de 0.95. ¿ Cuántas cuentas deben ser seleccionadas del archivo de la compañía si se sabe que la desviación típica de los balances mensuales de las cuentas de crédito es de I/. 750?.

18. La oficina de protección al consumidor, entre otras cosas, han ocasionado que las empresas se preocupen más por la aceptación de sus productos en el mercado. Una empresa, con dos productos en su línea, quiere estimar la diferencia entre el número de quejas ocurridas cada mes durante cuatro años (48 meses), que para cada productos arrojan los siguientes resultados,

Producto 1	$\bar{x}_1 = 17.2$	$s_1 = 4.6$
Producto 2	$\bar{x}_2 = 25.1$	$s_2 = 5.3$

Determinar un intervalo de confianza del 90% para la diferencia entre el número medio de quejas sobre sus productos. (suponga que las quejas sobre cada producto se pueden considerar como muestras aleatorias independientes).

19. Una muestra al azar de 200 pilas de la marca A para calculadoras muestra una vida media de 140 horas y una desviación típica de 10 horas. Una muestra al azar de 120 pilas de la marca B da una vida media de 125 horas y una desviación estándar de 9 horas. Determinar:

    (a) un intervalo de confianza del 95% para la diferencia de la vida media de las poblaciones A y B.

    (b) Un intervalo de confianza del 99% para la diferencia de la vida media de las poblaciones A y B.

20. Dos grupos escogidos al azar, de 50 alumnas de una escuela para secretarias, aprenden taquigrafía por dos sistemas diferentes y luego se las somete a pruebas de dictado. Se encuentra que el primer grupo obtiene en promedio 120 palabras por minuto con una desviación estándar de 11 palabras, mientras que el segundo grupo promedia 110 palabras por minuto con una desviación estándar de 10 palabras. Determinar un intervalo de confianza del 99% para la diferencia de medias de los dos métodos.

21. Un investigador desea comparar la efectividad de dos métodos de entrenamiento industrial para obreros que han de llevar a cabo un trabajo especial en una planta ensambladora. Los empleados seleccionados se dividen en dos grupos. El primer grupo recibe el método 1, el segundo el método 2. Cada uno realizará una operación de ensamblado y se registrará al tiempo que emplea en realizar el trabajo. Se espera que las observaciones en ambos grupos tengan un rango de 12 minutos (suponga también que se espera que la variabilidad de cada método sea la misma). Si se desea que el estimador de la diferencia en tiempo medio de armado sea correcta hasta por un minuto con probabilidad aproximada de 0.95, ¿cuántos trabajadores han de incluirse en cada grupo?

22. Se efectuó un estudio para determminar si cierto tratamiento para metales tiene algún efecto sobre la cantidad de metal desprendido durante la opera-

ción de decapado. Una muestra aleatoria de 100 piezas fue sumergida en un baño durante 24 horas sin el tratamiento, alcanzando un promedio de 12.2 mm de metal desprendido y una desviación típica de la muestra de 1.1 mm. Una segunda muestra de 200 piezas se sometió al tratamiento seguida por el baño durante 24 horas, obteniendose un desprendimiento promedio de 9.1 mm de metal con una desviación típica de la muestra de 0.9 mm. Determinar un intervalo del 98% de confianza para estimar la diferencia entre las medias de las poblaciones.

23. Sean $\bar{X}$, $\bar{Y}$ las medias de dos muestras aleatorias independientes cada una de tamaño n tomadas de las poblaciones normales $N(\mu_X, \sigma^2)$ y $N(\mu_Y, \sigma^2)$ respectivamente, donde la varianza común es conocida. Determinar n tal que,

$$P[\bar{X} - \bar{Y} - \frac{\sigma}{5} < \mu_X - \mu_Y < \bar{X} - \bar{Y} + \frac{\sigma}{5}] = 0.90$$

24. Sea $\bar{X}$ la media de una muestra aleatoria de tamaño n de una distribución normal con desviación típica $\sigma = 7$. Hallar n tal que

$P[\bar{X} - 0.8 < \mu < \bar{X} + 0.8] = 0.90$, aproximadamente.

25. Se selecciona una muestra aleatoria de 200 votantes y se halla que 114 estan contentos con el actual presidente. Hallar un intervalo de confianza del 95% para la fracción de votantes que estan a favor del actual presidente.

26. Se selecciona una muestra aleatoria de 500 fumadores y se encuentra que 86 tienen preferencia por una marca A. Determinar un intervalo de confianza del 90% para la fracción de la población de fumadores que prefieren la marca A.

27. Un auditor de una dependencia gubernamental de protección al consumidor quiere determinar la proporción de reclamos sobre pólizas de enfermedades que paga el seguro en un plazo de dos meses de haber recibido el reclamo. Se selecciona una muestra aleatoria de 200 reclamos y se determina que 80 fueron pagadas en un plazo de dos meses despues de recibirlas. Determine una estimación del intervalo con 99% de confianza de la proporción real de reclamos pagados dentro de ese plazo de dos meses.

28. El departamento de mantenimiento de una planta manufacturera de tejidos es responsable por 1,200 telares. El gerente del departamento ha determinado que el 45% de los daños en una muestra de 64 máquinas es por falta de mante-

nimiento. Determine un intervalo de confianza del 95% para esta proporción.

29. Mediante un muestreo al azar de 49 de 500 compradores en la exposición de muebles en la Feria del Hogar, el gerente de ventas de la compañía de muebles, encontró que el 80% de estos clientes se interesaron por una línea nueva de muebles contemporaneos. Establezca un intervalo de confianza del 96% para la proporción de compradores interesados por esta línea particular.

30. En una muestra al azar de 600 mujeres, 300 indican que están a favor de la ayuda del estado a los colegios privados. En una muestra al azar de 400 hombres, 100 indican que están a favor de lo mismo. Determinar un intervalo de confianza (a) del 55%, (b) 95% para la diferencia de proporciones de todas las mujeres y todos los hombres que favorecen tal ayuda.

31. Una empresa de estudios de mercado quiere estimar las proporciones de hombres y mujeres que conocen un producto promocionado a escala nacional. En una muestra aleatoria de 100 hombres y 200 mujeres se determina que 20 hombres y 60 mujeres están familiarizados con el artículo indicado. Calcular el intervalo de confianza de 95% para la diferencia de proporciones de hombres y mujeres que conocen el producto.

32. Cierto genetista quiere conocer la proporción de hombres y mujeres de cierta ciudad que padecen un desorden sanguineo menor. Una muestra aleatoria de 1000 mujeres arroja 250 afectadas, en tanto que 275 de 1000 hombres sufren desorden. Establezca un intervalo de confianza del 95% para la diferencia entre la proporción de hombres y mujeres que padecen tal desorden.

33. La toma de decisiones participativa ha sido una estrategia administrativa que se ha adoptado como un medio para mejorar la eficiencia y la participación de los individuos en las organizaciones. Se entrevistó a dos grupos de empleados, los cuales difieren substancialmente en el nivel de participación permitida por el gerente, y se les preguntó si estaban o no satisfechos con su empleo actual. De 110 empleados de un grupo en el cual se ha fomentado la participación del empleado, 77 afirmaron que estaban satisfechos con sus empleos. En tanto 52 de 125 empleados, de un grupo en el que no se permite la participación del empleado, afirmaron que estaban satisfechos con su empleo.

Encuentre un intervalo de confianza del 90% para la diferencia de la proporción de empleados satisfechos con su trabajo.

34. El auditor de un banco desea estimar la proporción de estados, de cuenta bancarias mensuales para los depositantes del banco que tendrán errores de varias clases, y especifica un coeficiente de confianza del 99% y un error máximo de 0.25%.

    (a) Determinar el tamaño de la muestra si no se dispone de información sobre la proporción verdadera de los estados de cuenta mensuales que tienen errores.

    (b) Determinar el tamaño de la muestra, si el auditor, por su experiencia, cree que la verdadera proporción de estados de cuenta con errores es 0.01.

35. Se efectua un estudio para estimar la proporción de amas de casa que poseen una secadora automática. ¿Cuál debe ser el tamaño de la muestra si se desea tener almenos una confianza del 99% que la,estimación difiera de la verdadera proporción en una cantidad menor de 0.01?

36. Un equipo de investigación médica está seguro sobre un suero que han desarrollado el cual curará cerca del 75% de los pacientes que sufren de ciertas enfermedades. ¿Qué tamaño debe ser la muestra para que el grupo pueda estar seguro en un 98% que la proporción muestral de los que se curan está dentro de más o menos 0.04 de la proporción de todos los casos que el suero curará?

37. En una muestra aleatoria de 25 presidentes de corporaciones sud-americanas, se encontró que 16 han cursado estudios superiores. Determinar un intervalo de confianza del 95% para la proporción de todos los presidentes de corporaciones que han cursado estudios superiores. (Sug. puesto que n es pequeño, introdusca el factor de corrección de $1/2\,n$ para $\hat{P}$ con el objeto de mejorar el intervalo de confianza para p).

$$(\hat{p} - \frac{1}{2n}) - z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < (\hat{p} + \frac{1}{2n}) + z_0\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

38. Sea Y/n la frecuencia relativa de éxitos de n ensayos independientes, con probabilidad de éxito igual a p. Si p es desconocido, pero está en una vecindad de 0.30. Determinar n tal que

$P[Y/n - 0.03 < p < Y/n + 0.03] = 0.08$ aproximadamente.

39. Si Y_1/n y Y_2/n son las frecuencias relativas respectivas de éxitos asociado con dos distribuciones binomiales independientes $b(n, p_1)$ y $b(n, p_2)$. Hallar n tal que,

$$P[Y_1/n - Y_2/n - 0.05 < p_1 - p_2 < Y_1/n + Y_2/n + 0.05] = 0.80$$

40. Doscientos cincuenta y seis pacientes que sufren de una cierta enfermedad fueron tratados con un nuevo medicamento. Este medicamento curó a 128 pacientes. ¿Con qué grado de confianza puede afirmarse que la efectividad del medicamento está entre 45% y 55%?

41. El director de Bienestar Estudiantil de una Universidad está considerando una nueva política en relación con hogares de estudiantes. Antes de tomar una decisión final, desea seleccionar una muestra aleatoria para estimar la proporción de estos que está en favor de esta nueva política. ¿Qué tamaño muestral se requiere para asegurar que el riesgo de sobrepasar un error de 0.10 es sólo 0.05?

42. Una muestra aleatoria de tamaño 400 seleccionada entre los alumnos que habían consultado el Servicio de Salud de una Universidad durante el año pasado indicó que 80 tenían una enfermedad de naturaleza psicosomática.
    (a) ¿Con qué grado de confianza puede afirmarse que 16% a 24% de todos los alumnos que consultaron el servicio de salud el año pasado tenían una enfermedad psicosomátia?
    (b) Supóngase que 2000 alumnos consultaron el servicio de salud el año pasado. ¿Cuántos de estos alumnos tenían una enfermedad psicosomática? ¿Qué grado de confianza tiene su estimación?

43. Una muestra de 5 tarros de café instanténeo seleccionados de un proceso de producción, dió los siguientes valores para el contenido medio en gramos: 285; 291; 279; 288; 282. Determine un intervalo de confianza del 95% para estimar el peso neto medio de los frascos producidos por este proceso.

44. Se prueba una muestra aleatoria de 5 fusibles de cierta marca para determinar el punto medio de ruptura. Los puntos de ruptura medidos en amperes fueron:

18; 22; 20; 14; 26.

¿Con qué grado de confianza puede afirmarse que el punto medio de ruptura para esta marca de fusibles está entre 15.736 y 24.264?

45. La resistencia a la rotura, expresada en libras, de cinco ejemplares de cuerda, cuyos diámetros sean 3/16 de pulgada, es de: 660, 460, 540, 580, 550. Estímese la resistencia media a la rotura mediante un intervalo de confianza del 95%, suponiendo distribuciones normales.

46. Las cajas de un cereal producidos por una fábrica deben tener un contenido de 16 onzas. Un inspector tomó una muestra que errojo los siguientes pesos en onzas:

    15.7, 15.7, 16.3, 15.8, 16.1, 15.9, 16.2, 15.9, 15.8, 15.6

    Calcule intervalos de confianza del 90% para la media poblacional y la varianza poblacional de los pesos.

47. Los pesos de 10 cajas de cereal son: 10.2, 9.7, 10.1, 10.3, 10.1, 9.8, 9.9, 10.4, 10.3, 9.8 onzas. Hallar un intervalo de confianza del 99% para la media de todas las cajas de cereal, asumiendo una distribución aproximadamente normal.

48. Una muestra aleatoria de 8 cigarrillos de cierta marca da un promedio de 18.16 miligramos de contenido de nicotina y una desviación estándar de 2.4 miligramos. Construir un intervalo de confianza del 99% para el verdadero promedio del contenido de nicotina de esta marca particular de cigarrillo, asumiendo una distribución aproximadamente normal.

49. Los pesos netos (grs) de ocho latas de conserva fueron los siguientes: 121, 119, 124, 123, 119, 121, 124, 120. Obtener un intervalo de confianza del 99% para el peso neto medio de las conservas.

50. Los tiempos de encendido en segundos de crisoles de humo flotante de dos tipos diferentes son los siguientes:

    Tipo I : 481, 506, 527, 661, 501, 572, 561, 501, 487, 524.

    Tipo II : 526, 511, 556, 542, 491, 537, 582, 605, 558, 578.

    Determine un intervalo de confianza del 95% para la diferencia media en tiempos de encendido, suponiendo varianza iguales pero desconocidas.

51. Dos analistas tomaron lecturas repetidas en la dureza del agua de la ciudad. Determine un intervalo de confianza del 95% para la diferencia entre analistas, suponiendo varianzas iguales pero desconocidas.

MEDIDA DE DUREZA CODIFICADA

Analista A:	0.46, 0.62, 0.37, 0.40, 0.44, 0.58, 0.48, 0.53
Analista B:	0.82, 0.61, 0.89, 0.51, 0.33, 0.48, 0.23, 0.25, 0.67, 0.88

52. El gerente de una cadena de 200 super-mercados en una ciudad grande, reune los datos de ventas diarias de 5 tiendas escogidas al azar. Ellos son, en miles de intis, 18, 24, 22, 26, 16.
   (a) Estime la desviación típica de la población.
   (b) Construya un intervalo de confianza de un 98% para las ventas medias.

53. Al examinar los registros de facturación mensual de una empresa editora con ventas por correo, se encuentra un total de 250 facturas no pagadas, el auditor toma una muestra de 10 facturas no pagadas. Las sumas que se adeudan a la compañía en miles de intis son: 4, 18, 11, 7, 7, 10, 5, 33, 9 y 12.
   (a) Determine una estimación del intervalo con 99% de confianza de la cantidad promedio de facturas no pagadas.
   (b) Establezca una estimación del intervalo con 99% de confianza de la cantidad total que se adeuda a esa empresa.

54. Los siguientes datos representan los tiempos de duración (en minutos) de las pelícülas producidas por dos compañías cinematográficas

   Compañía A : 103, 94, 110, 87, 98

   Compañía B : 97, 82, 123, 92, 175, 88, 118

   Determinar el intervalo de confianza del 90% para la diferencia entre los tiempos promedios de las películas producidas por las compañías, asumiendo que los tiempos de duración tienen una distribución aproximadamente normal.

55. Una compañía productora de semillas de maíz híbrico planta dos nuevas hileras de maiz híbrico en cinco granjas diferentes. Las producciones en búshels por acre fueron:

Híbrico I :	90	85	95	76	80
Híbrico II :	84	87	90	92	90

   (a) Determine un intervalo de confianza del 90% para la diferencia entre las dos producciones medias.
   (b) ¿Con qué tipo de "población" trabaja la compañía en esta prueba?
   (c) ¿Qué suposiciones se hicieron para estimar el intervalo de confianza en (a)?

56. Se determina que las tensiones de rotura de una línea de pesca de prueba de 30 libras, para una muestra de 6 carreteles, son 34, 33, 26, 32, 28 y 27 libras. Determinar un intervalo de confianza del 95% para la varianza

poblacional.

57. Para una muestra de tamaño 15, $\Sigma(X_i - \bar{X})^2 = 17.8$. Hallar un intervalo de confianza del 95% para σ.

58. Se realiza un experimento para probar la diferencia en efectividad de dos métodos de cultivar trigo. Diez parceles se tratan con arado superficial y quince con arado profundo. El rendimiento medio por acre del primer grupo es 40.8 fanegas y la media para el segundo grupo es 44.7. Suponga que la desviación estándar de la producción del arado superficial es 0.6 fanegas y para el profundo es 0.8. Encuentre un intervalo de confianza del 90% para la diferencia de rendimiento.

59. En investigación sobre combustibles para cohetes orientada a reducir la demora entre la aplicación de la corriente de encendido y la explotación, se realizaron pruebas sobre un combustible T y un combustible C. La desviación estándar de ambos grados puede suponerse igual a 0.04. Si $\bar{C} = 0.261$ segundos y $\bar{T} = 0.250$ segundos para 14 observaciones en cada uno, determine un intervalo de confianza del 95% para el cambio en tiempo de encendido.

60. Para determinar la efectividad de un programa de seguridad industrial se recogieron los siguientes datos sobre el tiempo perdido por accidentes.

	1	2	3	4	5	6	7	8
antes del programa	38.0	69	15.4	10	121	47.6	79	54
después del programa	29.5	48.5	19	00	92.5	50	80.5	40

(Los números dados son las medias de horas-hombre pérdida por mes, en un período de 8 meses).

(a) Hallar intervalos de confianza del 95% para cada una de las medias, antes del programa y después del programa.

(b) ¿Qué podría decir acerca del programa en base a los resultados?

61. En el problema 59, determinar un intervalo de confianza del 95% para σ_T / σ_c .

62. Suponga que se hacen 15 ensayos en cada uno de dos tratamientos con la razón de desviaciones estándares muestrales $S_X / S_Y = 3.5$. Determinar un intervalo de confianza del 90% para σ_X / σ_Y

63. Suponga que se hacen 15 ensayos en cada uno de los tres tratamientos con las siguientes varianzas muestrales $S_X^2 = 12$, $S_Y^2 = 30$ y $S_Z^2 = 39$. Determinar intervalos de confianza del 85% para todos los pares de razones de varianzas.

9

PRUEBA DE HIPOTESIS

9.1 HIPOTESIS ESTADISTICA

La prueba de hipótesis estadística es quizas el área más importante de la teoría de decisión. Primero definiremos, que entendemos por una hipótesis estadística.

DEFINICION 9.1.1 *Una hipótesis estadística*, es una aseveración que se hace a cerca de la distribución de una o más variables aleatorias (o poblaciones).

Se puede especificar una hipótesis, dando el tipo de distribución y el valor o valores del parámetro o los parámetros que la definen. Hipótesis de este tipo serán, por ejemplo:

(a) **X**, tiene una distribución binomial con $p = \frac{1}{3}$.

(b) **X**, tiene una distribución normal con $\mu = 50$, $\sigma = 5$.

En la práctica, la distribución de la población, generalmente se asume. Por lo tanto una hipótesis se especifica con el valor o los valores del parámetro. Los siguientes, son ejemplos de este tipo:

(c) El promedio salarial por semana de los obreros en la industria textil es I/. 600.00. Es decir, $\mu = 600.00$ por semana.

(d) El porcentaje de personas atacadas por cierta epidemia en una ciudad grande es el 10%. Es decir, $p = 0.1$.

Si la hipótesis estadística, especifica completamente la distribución, es decir específica su forma funcional y los valores de todos los parámetros, se llama *una hipótesis estadística simple*. En caso contrario se llama una *hipótesis estadística compuesta*.

El ejemplo (b) es una hipótesis simple; (a) es compuesta, pues $p = 1/3$ no define completamente la distribución de X sin especificar n.

DEFINICION 9.1.2 *La prueba estadística de una hipótesis*, es una regla que cuando los valores experimentales son observados nos conducen a una decisión; no rechazar (aceptar) o rechazar la hipótesis bajo consideración.

Una forma cómoda de especificar lo que se quiere de un procedimiento de prueba es concentrar la atención en dos hipótesis estadística, llamadas respectivamente, *Hipótesis nula* denotado por H_0 (que es la hipótesis que se quiere probar) y *la hipótesis alternativa*, denotado por H_1 (que es una suposición contraria a la que se quiere probar), que se acepta en caso de que la primera sea rechazada.

La hipótesis nula generalmente es especificada en una forma exacta. Es decir, la hipótesis nula "anula" el efecto de un tratamiento y corresponde a la ausencia de efectos de la variable que se investiga. La hipótesis alternativa generalmente es formulada con menos precisión; a menudo se especifica como una variación de valores que prevalecería si la variable que se estudia ejerció algún efecto. Así la hipótesis nula se especifica con frecuencia en una forma opuesta a la que se supone cierta y la alternativa es expresada como la opuesta a la hipótesis nula.

En una terminología de prueba, hablamos de probar la hipótesis nula contra una alternativa en el supuesto tentativo que la hipótesis nula es cierta. Pero debemos comprender que realmente estamos tomando una decisión entre dos acciones o entre H_0 y H_1.

Hay tres tipos principales de pruebas, cada uno de los cuales es identificado por la forma en que se formulan H_0 y H_1. Las pruebas unilaterales y la prueba bilateral.

1. *Prueba de una cola o unilateral*, estas pueden ser:

   (a) *Prueba de cola inferior o prueba del lado izquierdo*: para la cual las hipótesis toman las siguientes formas

$$H_0: \theta = \theta_0 \; ; \; H_1: \theta < \theta_0$$

Este tipo de prueba se emplea cuando se tiene alguna evidencia que el parámetro no es igual a θ_0 ,si no que debe ser menor. También cuando el valor de un parámetro no es demasiado pequeño para algún objetivo.

(b) *Prueba de la cola superior o prueba de la cola derecha.* En este caso las hipótesis suelen expresarse de la siguiente forma,

$$H_0: \theta = \theta_0 \; ; \; H_1: \theta > \theta_0$$

Este tipo de prueba se emplea, en problemas, en que se tiene algún indicio que el parámetro no es igual al valor postulado, debe ser mayor que el postulado. También cuando el valor de un parámetro sea bastante grande para algún objetivo específico.

2. *Prueba de dos colas o bilaterales.* En este caso las hipótesis toman la siguiente forma,

$$H_0: \theta = \theta_0 \; ; \; H_1: \theta \neq \theta_0$$

Este tipo de prueba se emplea, en el caso que el valor que se prueba no sea verdadera; entonces, todos los demás valores son posibles. También cuando el valor de un parámetro, es demasiado pequeño o demasiado grande para algún objetivo específico.

9.1.1 TIPOS DE ERRORES

En la prueba de hipótesis pueden cometerse dos tipos posibles de errores: rechazar la hipótesis H_0 , cuando es en realidad verdadera o aceptarla (no rechazarla) cuando es falsa. El rechazo de una hipótesis nula cuando es verdadera, *se llama error de tipo I.* La aceptación de una hipótesis nula cuando es falza, *se llama error de tipo II.*

H_0 es la hipótesis nula (sometida a prueba) y H_1 , la hipótesis alternativa. Estas dos posibilidades incorrectas junto con las dos posibilidades de decisiones correctas, aparecen en la tabla I.

TABLA I

Decisión	H_o verdadera	H_1 verdadera
Aceptar, H_o	decisión correcta	Error de tipo II
Aceptar, H_1	error de tipo I	Decisión correcta

Es obvio que, quien toma las decisiones, quiera reducir al máximo las probabilidades de cometer cualquiera de estos dos tipos de errores, esto no es fácil, pues las probabilidades de cometer errores de tipos I y II son inversamente proporcionales, para cualquier prueba dada. De ahí que, cuanto menor es el riesgo de cometer un error de tipo I, tanto mayor es la probabilidad de cometer un error del tipo II, y viceversa. Sin embargo dada la regla de decisión, es posible reducir ambos tipos de errores en forma simultánea, aumentando el tamaño de la muestra.

DEFINICION 9.1.3 *El nivel de significación*, se representa por α y se define como la probabilidad de cometer un error de tipo I. Es decir,

$$\alpha = P[\text{error tipo I}] = P[\text{rechazar } H_o \mid H_o \text{ es verdadera}]$$
$$= P[\text{aceptar } H_1 \mid H_1 \text{ es falsa}]$$

La probabilidad de cometer un error de tipo II, se representa por β , es decir,

$$\beta = P[\text{error tipo II}] = P[\text{aceptar } H_o \mid H_o \text{ es falsa}]$$
$$= P[\text{rechazar } H_1 \mid H_1 \text{ es verdadera}]$$

Los valores que frecuentemente se emplean para α son 0.01 ó 0.05. Observe, que si $\alpha = 0.01$ ó 0.05, se arriesga alrededor de una vez en 100 casos o en 20 casos respectivamente, el rechazo de una hipótesis aún cuando sea cierta.

Debido a que la decisión de no rechazar o rechazar H_o (o sea la elección entre H_o y H_1) se hace basándose en pruebas de muestras, debemos escoger una función $\hat{\theta}$ de las n observaciones,

$$\hat{\theta} = G(X_1, X_2, \ldots, X_n)$$

como *estadística de prueba*, cuya distribución por muestreo sea conocida en el supuesto (tentativo) que la hipótesis nula $H_o: \theta = \theta_o$ es verdadera. Las reglas de decisión sobre la aceptación o rechazo de H_o , se hace respecto al rango de $\hat{\Theta}$ y un resultado particular $\hat{\theta}$ de la muestra. Esto se hace hallando un valor $\hat{\theta}_c$ llamado *valor crítico* de la estadística de prueba (a veces hay más de un valor crítico) la cual divide al rango de $\hat{\Theta}$ en dos regiones; región crítica o de rechazo (R.C) y la región de aceptación (R.A). Si $\hat{\theta} \in$ R.C rechazamos H_o. Si $\hat{\theta} \in$ R.A, no rechazamos H_o.

DEFINICION 9.1.4 *Región Crítica o región de rechazo*, es la región, del rango de $\hat{\Theta}$ que de acuerdo con una prueba prescrita, conduce al rechazo de la hipótesis bajo consideración. En otras palabras, región crítica o de rechazo es la región que contiene los valores para los cuales se rechaza la hipótesis bajo consideración.

DEFINICION 9.1.5 *Región de Aceptación*, es la región que contiene los valores para los cuales no se rechaza la hipótesis bajo consideración.

Los pasos para la prueba de hipótesis, relativa al parámetro θ de una población, puede resumirse como sigue:

1. Formular la hipótesis nula y alternativa de acuerdo al problema.

   H_o : $\theta = \theta_o$

   H_1 : cualquiera de las alternativas, $\theta < \theta_o$, $\theta > \theta_o$ ó $\theta \neq \theta_o$
2. Escoger un nivel de significación o riesgo α.
3. Escoger la estadística de prueba apropiada, cuya distribución por muestreo sea conocida en el supuesto de que H_o es cierta.
4. Establecer la región crítica. Es decir, determinar el valor (o valores) crítico.
5. Calcular los valores de la prueba estadística de una muestra aleatoria de tamaño n.
6. Conclusión: Rechazar H_o si la estadística tiene un valor en la región crítica y no rechazar (aceptar), en otro caso.

9.2. PRUEBAS RELATIVAS A MEDIAS Y VARIANZAS

9.2.1 PRUEBA UNILATERAL DE UNA HIPOTESIS SOBRE LA MEDIA

PRIMER CASO Comenzaremos considerando la siguiente hipótesis

1. H_o : $\mu = \mu_o$

   H_1 : $\mu < \mu_o$

   donde μ_o es un valor de la media poblacional

2. Escogemos el nivel de significación α .
3. Una estadística para la media de la población, es la media muestral $\bar{X}$. Si la población es normal (o si la muestra es grande $n \geq 30$, aún cuando la población no es normal) la distribución de $\bar{X}$ es $N(\mu,\sigma^2/n)$ y la variable aleatoria $Z = \dfrac{\bar{X} - \mu_o}{\sigma / \sqrt{n}}$, tiene una distribución N(0,1).
4. Es razonable, rechazar H_o en favor de H_1 , si la media muestral $\bar{X}$ es demasiado pequeña con respecto a μ_o . Entonces, una región crítica podría obtenerse, seleccionando un valor $\bar{x}_c$ de la media muestral de manera que, R.C = $\langle -\infty, \bar{x}_c \rangle$ donde $\bar{x}_c$ es tal que $P[\bar{X} < \bar{x}_c/H_o] = \alpha$

$$ó \qquad P[\frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} < \frac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}}] = \Phi[\frac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}}] = \alpha$$

en la tabla de la distribución normal se encuentra z_α que corresponde al area α (denotado por z_α) (ver fig. 9.2.1)

Es decir, $\dfrac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}} = z_\alpha$ de donde $\bar{x}_c = \mu_o + \dfrac{z_\alpha \sigma}{\sqrt{n}}$

luego, $R.C = \langle -\infty, \mu_o + \dfrac{z_\alpha \sigma}{\sqrt{n}} \rangle$ ó $\bar{X} < \mu_o + \dfrac{z_\alpha \sigma}{\sqrt{n}}$.

5. Calcular $\bar{x}$ a partir de la muestra observada.
6 Conclusión: si $\bar{x} \in R.C = \langle -\infty, \mu_o + \dfrac{z_\alpha\sigma}{\sqrt{n}} \rangle$, rechazar H_o. Si $\bar{x} \notin R.C$, no rechazar H_o , (aceptar).

OBSERVACION En el paso 4, en vez de R.C, se puede obtener la región de aceptación R.A. Esto es,

$$R.A = \langle \bar{x}_c, \infty \rangle \quad \text{tal que } P[\bar{X} > \bar{x}_c/H_o] = 1 - \alpha$$

$$ó \qquad P[\frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} > \frac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}}] = 1 - \alpha$$

utilizando la tabla normal estándar, obtenemos z_α y

$$\bar{x}_c = \mu_o + \frac{z_\alpha \sigma}{\sqrt{n}}$$

luego $\quad R.A = \left\langle \mu_o + \frac{z_\alpha \sigma}{\sqrt{n}}, \infty \right\rangle$

Entonces, el paso 6, será: Aceptar H_o, si $\bar{x} \in R.A$ y rechazar H_o si $\bar{x} \notin R.A.$

METODO ALTERNATIVO

Un método alternativo práctico es trabajar directamente en la escala z (ver fig. 9.2.1), de la siguiente manera:

1. $H_o: \mu = \mu_o$ contra $H_1: \mu < \mu_o$
2. Escoger el nivel de significación α.
3. La estadística de prueba para $n \geq 30$ es $z = \frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}}$ con distribución normal estándar
4. La región crítica es $R.C = \langle -\infty, z_\alpha \rangle$ donde z_α es tal que
$$P[Z < z_\alpha] = \alpha$$
5. Calcular $\bar{x}$ de los datos, luego obtener z por la fórmula
$$z = \frac{\bar{x} - \mu_o}{\sigma/\sqrt{n}}$$
6. CONCLUSION: se compara z con z_α. Si $z < z_\alpha (z \in \langle -\infty, z_\alpha \rangle)$ se rechaza H_o.

   si $z > z_\alpha$ $(z \in \langle z_\alpha, \infty \rangle)$ no se rechaza H_o.

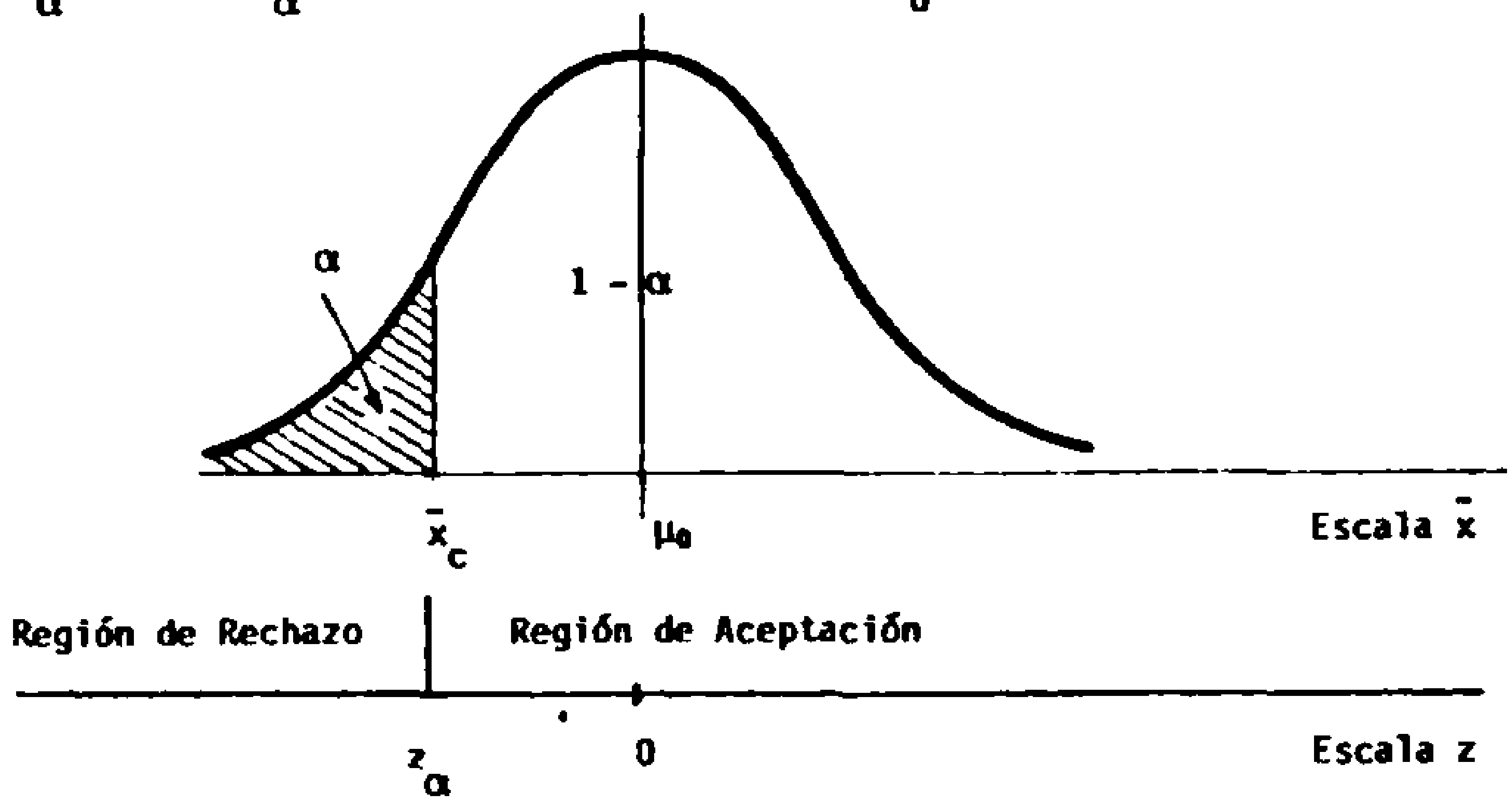


Fig. 9.2.1

El cálculo de β para pruebas estadísticas de una sola cola puede facilitarse si se considera la Fig. 9.2.2. La distribución de $\bar{X}$ bajo el supuesto que - $H_o : \mu = \mu_o$ se ilustra en la Fig. 9.2.2a con una curva normal punteada; se usa para localizar la región crítica o de rechazo y el valor crítico $\bar{x}_c$ para la estadística de prueba $\bar{X}$. Cuando H_o es falza, esto es $\mu = \mu_1$, donde digamos - $\mu_1 < \mu_o$, entonces la estadística de prueba $\bar{X}$ tiene una distribución normal - con media μ_1 en lugar de μ_o. La distribución de $\bar{X}$ suponiendo $\mu = \mu_1$ se ilustra en la Fig. 9.2.2b con una curva sólida. El punto $\bar{x}_c$ es aquel donde el área de la izquierda de la curva (a) es igual al nivel de significación α de la prueba, β es igual al área de la cola derecha de la curva (b) observe que está localizada en la región de aceptación.

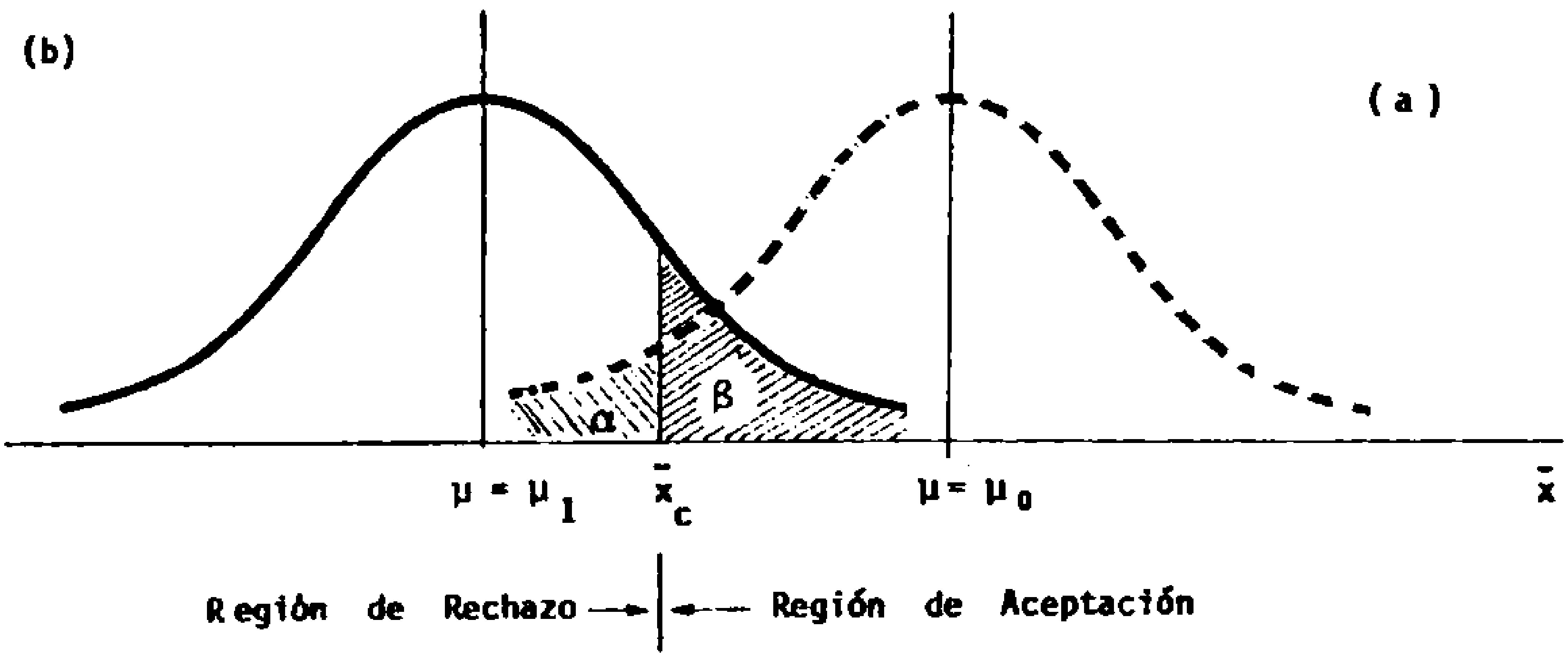


Fig. 9.2.2

EJEMPLO 1 Un comprador de ladrillos cree que la calidad de los ladrillos está disminuyendo. De experiencias anteriores, la resistencia media al desmoronamiento de tales ladrillos es 200 kg, con una desviación típica de 10 kg. Una muestra de 100 ladrillos arroja una media de 195 kg. Probar la hipótesis, la calidad media no ha cambiado, contra la alternativa que ha disminuído.

SOLUCION Primera forma:

1. $H_o : \mu = 200$ kg. y $H_1 : \mu < 200$ kg.
2. Escogemos el nivel de significación $\alpha = 0.05$
3. La estadística de prueba es $\bar{X}$. Desde que la muestra es grande $n = 100$, la

la distribución de $\bar{X}$ es

$$N(200, \frac{10^2}{100}) = N(200,1)$$

(Teorema central del límite). Luego,

$$Z = \frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} \quad \text{es} \quad N(0,1)$$

4. R.C = $\langle -\infty, \bar{x}_c \rangle$ donde $\bar{x}_c$ tal que $P[\bar{X} < \bar{x}_c / H_o] = \alpha$

ó $P[\frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} < \frac{\bar{x}_c - 200}{10/10}] = P[Z < \bar{x}_c - 200] = 0.05$

de donde $z_\alpha = \bar{x}_c - 200 = - 1.64$, luego $\bar{x}_c = 198.36$

R.C = $\langle -\infty, 198.36 \rangle$.

5. Cálculo de la media muestral: del enunciado una muestra de $n = 100$, da $\bar{x} = 195$.

6. Conclusión: Puesto que $\bar{x} = 195 \in$ R.C = $\langle -\infty, 198.36 \rangle$. Rechazamos H_o,

METODO ALTERNATIVO Una solución alternativa al problema se obtiene como sigue

(*i*) En el paso 4. De $P[Z < z_\alpha] = \alpha$ se obtiene $z_\alpha = - 1.64$. Entonces R.C = $\langle -\infty, - 1.64 \rangle$.

(*ii*) En el paso 5, una vez calculado $\bar{x}$, obtener z por la fórmula

$$Z = \frac{\bar{x} - \mu_o}{\sigma/\sqrt{n}} = \frac{195 - 200}{10/\sqrt{100}} = - 5$$

Conclusión: Comparar $z = - 5$ con $z_\alpha = - 1.64$. Puesto que $z = - 5 < - 1.64 = z_\alpha$ es decir,

$$z = - 5 \in \text{R.C} = \langle -\infty, - 1.64 \rangle \text{, rechazamos } H_o .$$

EJEMPLO 2 Con referencia al ejemplo 1 calcule la probabilidad β de aceptar H_o cuando en realidad μ es igual a 196 kg.

SOLUCION La región de aceptación para la prueba del ejemplo 1 se localiza en el intervalo

$$\langle \mu_o + \frac{z_\alpha \sigma}{\sqrt{n}}, \infty \rangle$$

si se sustituyen los valores numéricos correspondientes se obtiene

$$200 + (-\ 1.64)\ \frac{10}{\sqrt{100}} = 198.36$$

R.A = $\langle 198.36, \infty \rangle$ (es el complemento de R.C)

La probabilidad de aceptar H_0 cuando es falsa, en este caso cuando $\mu = 196$ es igual al área bajo la curva de la distribución de la estadística de prueba $\bar{X}$ en el intervalo $\langle 198.36, \infty \rangle$. Puesto que $\bar{X}$ se distribuye normalmente con media 196 y $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, β es igual a (ver Fig. 9.2.3)

$$\begin{aligned}
\beta &= P[\text{aceptar } H_0 \text{ cuando } H_0 \text{ es falsa}] \\
&= P[\text{aceptar } H_0 \text{ cuando } \mu = 196] \\
&= P[\bar{X} > 198.36 \text{ cuando } \mu = 196] \\
&= P[\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} > \frac{198.36 - 196}{10/\sqrt{100}}] = P[Z > 2.36] \\
&= 0.0091
\end{aligned}$$

Entonces la probabilidad de aceptar H_0, dado que en realidad $\mu = 196$ es 0.0091 o aproximadamente 1 de cada 100.

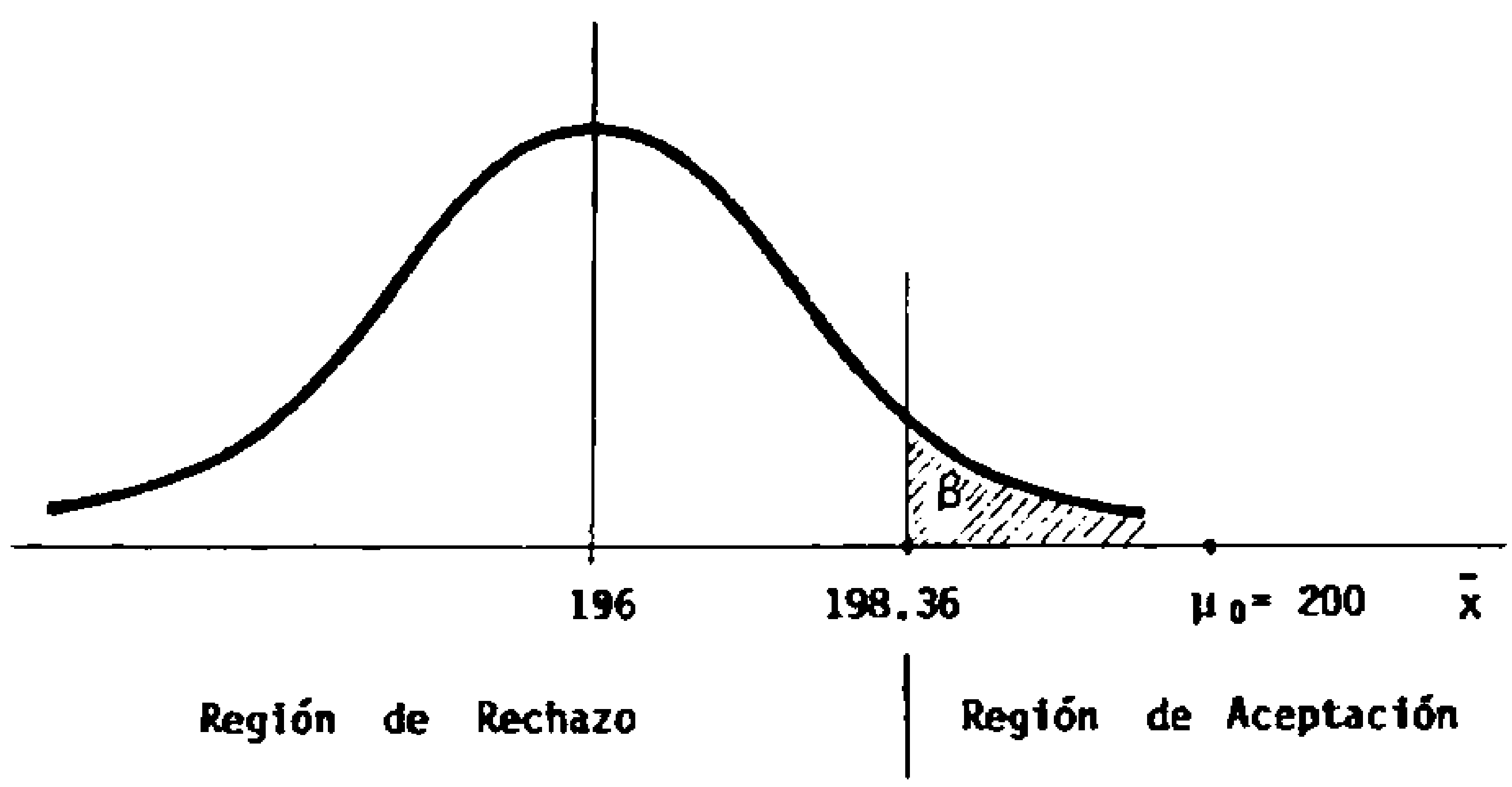


Fig. 9.2.3

Si se calcula μ para un valor de μ_1 más cerca a $\mu_0 = 200$, la distribución que se muestra en la Fig. 9.2.3 se desplazaría a la derecha y β aumentaría. Por el contrario cuanto mayor sea la distancia entre μ_1 y μ_0, menor será el valor de β. Es decir se cometen menos errores en decidir que la media será menor de μ_0 (H_0 es falsa) cuando la diferencia entre μ_1 y μ_0 es grande.

MUESTRAS PEQUEÑAS

Para muestras pequeñas de poblaciones con distribución aproximadamente normal, se recurre a la distribución t para prueba de hipótesis sobre la media. Entonces

1. $H_0 : \mu = \mu_0$; $H_1 : \mu < \mu_0$
2. Escoger el nivel de significación α .
3. La estadística de prueba es $\bar{X}$, para muestras pequeñas, usamos la variable aleatoria
$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$
que tiene una distribución t con $n - 1$ grados de libertad
4. La región crítica es
$$R.C = \langle -\infty , \bar{x}_c \rangle \quad \text{donde } \bar{x}_c \text{ es tal que } P[\bar{X} < \bar{x}_c / H_0] = \alpha$$
$$P[\frac{\bar{X} - \mu_0}{S/\sqrt{n}} < \frac{\bar{x}_c - \mu_0}{S/\sqrt{n}}] = P[T < \frac{\bar{x} - \mu_0}{S/\sqrt{n}}] = P[T < t_\alpha] = \alpha$$
De la tabla V obtenemos $\dfrac{\bar{x}_c - \mu_0}{S/\sqrt{n}} = t_\alpha$, o sea $\bar{x}_c = \mu_0 + \dfrac{t_\alpha S}{\sqrt{n}}$
Luego, $R.C = \langle -\infty , \mu_0 + \dfrac{t_\alpha S}{\sqrt{n}} \rangle$ (o $R.C = \langle -\infty , t_\alpha \rangle$ método alternativo)
5. Calcular $\bar{x}$ y s de la muestra. Luego $t = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}}$ si se usa el método alternativo.
6. Conclusión: Si $\bar{x} \in \langle -\infty , \mu_0 + \dfrac{t_\alpha S}{\sqrt{n}} \rangle$ rechazar H_0 . Para el método alternativo si $t \in \langle -\infty , t_\alpha \rangle$ rechazar H_0 .

EJEMPLO 3 En una muestra aleatoria de 10 latas de maracuya de un proveedor B, el peso medio por lata de maracuya fue $\bar{x} = 9.4$ con desviación típica, $s = 1.8$ onz, ¿contiene esta muestra suficiente evidencia para indicar que el peso medio es menor que 10 onzas a un nivel $\alpha = 0.1$? Y encontrar un intervalo de confianza del 98% para μ .

SOLUCION Usaremos el método alternativo para la solución de este problema

1. $H_0 : \mu = 10$ y $H_1 : \mu < 10$
2. $\alpha = 0.1$
3. Desde que $n = 10$ es pequeño, la estadística de prueba es

   $T = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$, que tiene una distribución t con $n - 1 = 9$ grados de libertad. Suponiendo que la población tiene una distribución aproximadamente normal.
4. Región Crítica o de rechazo: $P[T < t_\alpha] = \alpha$ de donde $t_\alpha = -1.383$; es decir, $R.C = \langle -\infty , -1.383\rangle$

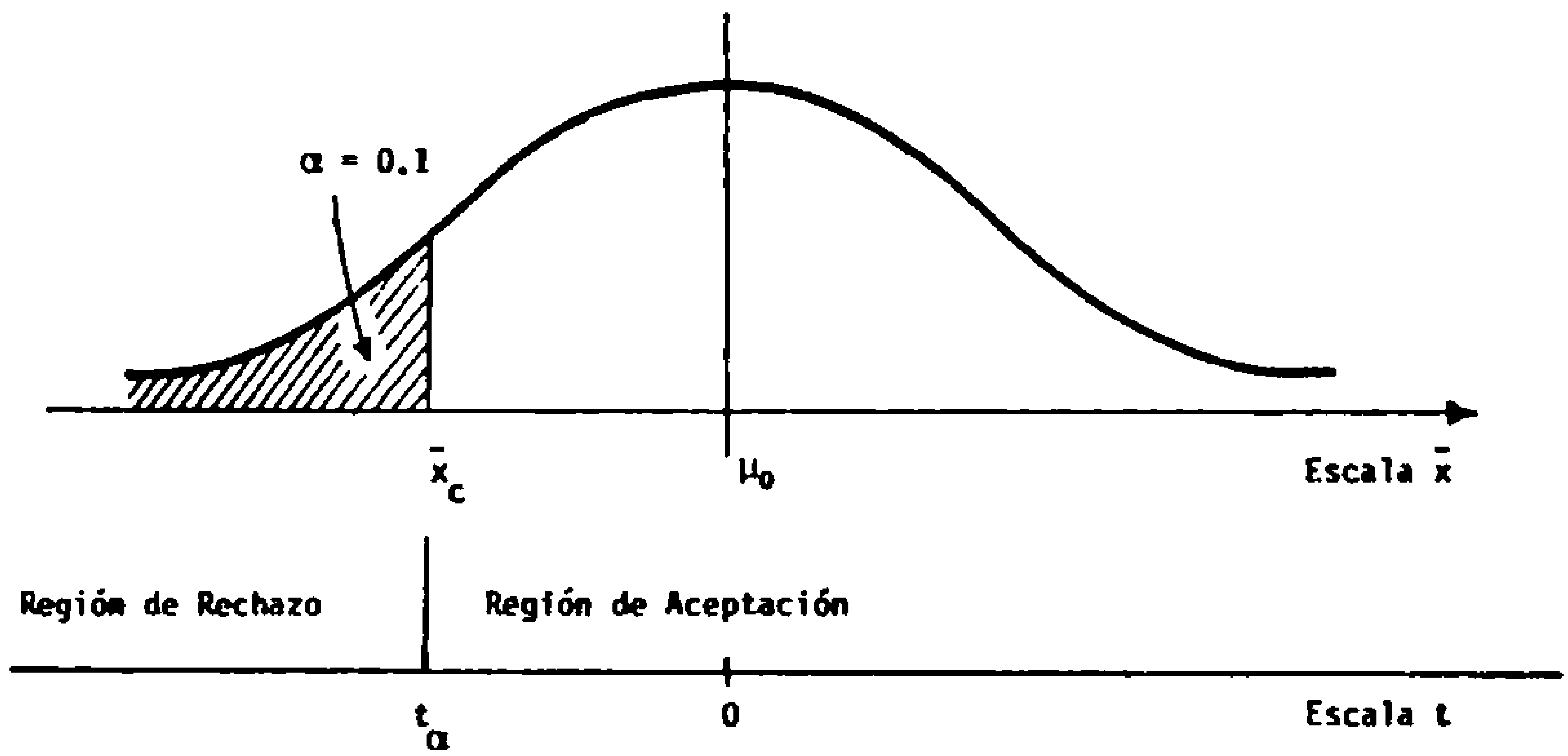


Fig. 9.2.4

5. Del enunciado del problema se tiene

   $\bar{x} = 9.4$, $s = 1.8$, $n = 10$

   luego, $t = \dfrac{9.4 - 10}{1.8/\sqrt{10}} = -1.054$
6. Conclusión: $t = -1.054 \notin \langle -\infty, -1.383\rangle$, no se rechaza H_0 ; es decir, esta muestra no contiene suficiente evidencia para indicar que el peso medio es menor que 10 onzas a un nivel de significación 0.1.

   La parte (b) queda como ejercicio para el lector.

EJEMPLO 4 Supóngase que en cierto proceso para producir alambre, la resistencia a la ruptura del alambre es una variable aleatoria normal con media 90.80 kg. Para reducir los costos de producción, se prueba otro proceso. Una muestra de 10 valores obtendidos bajo el nuevo proceso dio una media de 85.352 kg. y una desviación típica de 2.724 kg. ¿El nuevo proceso tiene un efecto negativo sobre el alambre? use $\alpha = 0.05$.

SOLUCION

1. $H_0 : \mu = 90.80$ y $H_1 : \mu < 90.8$
2. $\alpha = 0.05$
3. Desde que $n = 10$ es pequeño usamos la estadística de prueba $T = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$

   que tiene una distribución **t** con $n - 1 = 9$ grados de libertad. Suponiendo que la población tiene una distribución aproximadamente normal.
4. Región crítica o de rechazo: $P[T < t_\alpha] = \alpha$ de donde $t_\alpha = -1.833$; es decir R.C = $\langle -\infty, -1.833\rangle$ o $t < -1.833$
5. Del enunciado del problema se tiene que la media muestral es $\bar{x} = 85.352$ y $s = 2.724$

$$t = \frac{85.352 - 90.80}{2.724/\sqrt{10}} = -6.324$$

6. Conclusión: desde que $t = -6.324 \in \langle -\infty, -1.833\rangle$, se rechazo H_0 , es decir el nuevo proceso tiene un efecto negativo cobre el alambre a un nivel de significación 0.05.

SEGUNDO CASO Consideremos ahora la siguiente prueba

1. $H_0 : \mu = \mu_0$

   $H_1 : \mu > \mu_0$
2. Escoger el nivel de significación o de riesgo α.
3. La estadística para la media poblacional es la media muestral $\bar{X}$. Si la población es normal o si la muestra es grande $n > 30$ la variable aleatoria $\bar{X}$ tiene una distribución $N(\mu , \frac{\sigma^2}{n})$. Cuando la muestra es pequeña y la población es aproximadamente normal se usa la estadística $T = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$ que tiene una distribución **t** con $n - 1$ grados de libertad.
4. La región crítica; en este caso es razonable rechazar H_0 en favor de H_1 si la media muestral es demasiado grande con relación a μ_0. Por lo tan-

to, podría obtenerse una región crítica, escogiendo un valor $\bar{x}_c$ de la media muestral tal que

$$R.C. = \langle \bar{x}_c, \infty \rangle \quad \text{tal que} \quad P[\bar{X} > \bar{x}_c / H_o] = \alpha$$

ó
$$P\left[\frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} > \frac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}}\right] = P\left[Z > \frac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}}\right] = \alpha$$

de donde, $\dfrac{\bar{x}_c - \mu_o}{\sigma/\sqrt{n}} = z_{1-\alpha}$ ó $\bar{x}_c = \mu_o + \dfrac{z_{1-\alpha}\sigma}{\sqrt{n}}$

luego $R.C = \left\langle \mu_o + \dfrac{z_{1-\alpha}\sigma}{\sqrt{n}}, \infty \right\rangle$

ó simplemente $R.C = \langle z_{1-\alpha}, \infty \rangle$ (Método alternativo).

Cuando, se usa la distribución **t**, la región de rechazo es

$R.C = \langle t_{1-\alpha}, \infty \rangle$

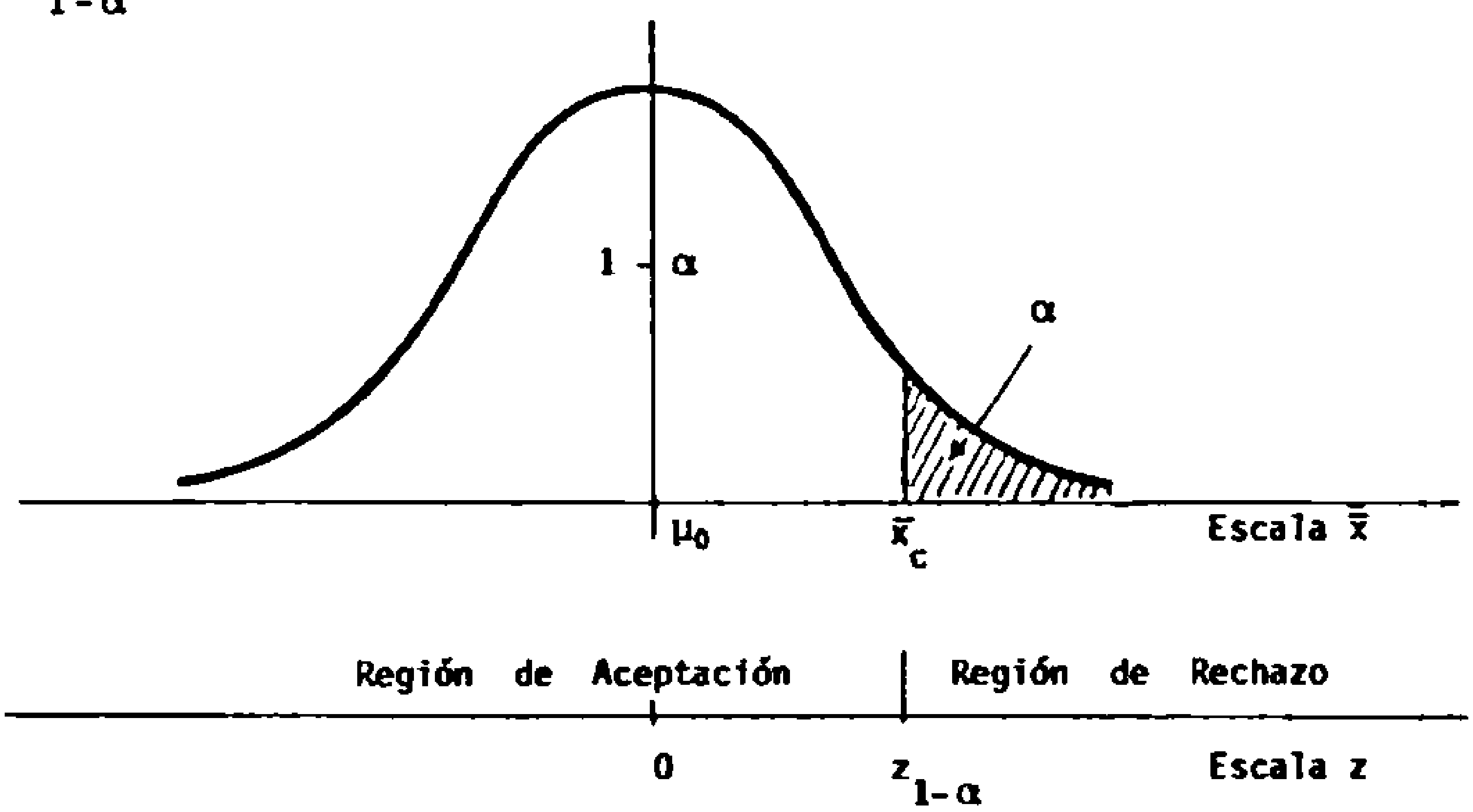


Fig. 9.2.5

5. Calcular $\bar{x}$ (y s si n es pequeño) a partir de una muestra aleatoria de tamaño n. O también se calcula z o **t** según sea el caso si se usa el método alternativo.

6. Conclusión: si $\bar{x} \in \left\langle \mu_o + \dfrac{z_{1-\alpha}\sigma}{\sqrt{n}}, \infty \right\rangle$ se rechaza H_o. Para el método

alternativa si $z \in \langle z_{1-\alpha}, \infty \rangle$ o $t \in \langle t_{1-\alpha}, \infty \rangle$ según el caso se rechaza H_0. En otro caso se acepta H_0

EJEMPLO 5 La experiencia de muchos años respecto a un examen de inglés, para obtener admisión en la Universidad, arroja una calificación media de 64 puntos con una desviación típica de 9 puntos. Todos los estudiantes de una ciudad, de la cual hay 54 en el exámen, han obtenido una calificación media de 68 puntos. ¿Puede tenerse la certeza que los estudiantes de esta ciudad saben más inglés?

SOLUCION

1. $H_0 : \mu = 64$ y $H_1 : \mu > 64$
2. Escogemos $\alpha = 0.05$
3. La estadística de prueba es $\bar{X}$. Desde que la muestra es grande $n = 54$ y $\sigma = 9$, la variable aleatoria $\bar{X}$ tiene distribución $N(64, \frac{81}{54})$. (teorema central del límite)
4. $R.C = \langle \bar{x}_c , \infty\rangle$ donde $\bar{x}_c$ es tal que $P[\bar{X} > \bar{x}_c / H_0] = 0.05$

   o $$P\left[\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{\bar{x}_c - 64}{9/\sqrt{54}}\right] = P\left[Z > \frac{\bar{x}_c - 64}{9/\sqrt{54}}\right] = 0.05$$

   o $$P\left[Z < \frac{\bar{x}_c - 64}{9/\sqrt{54}}\right] = \Phi\left[\frac{\bar{x}_c - 64}{9/\sqrt{54}}\right] = 0.95$$

   de donde $$z_{1-\alpha} = 1.64 = \frac{\bar{x}_c - 64}{9/\sqrt{54}}$$

   luego, $$\bar{x}_c = 64 + \frac{(1.64)9}{\sqrt{54}} = 66$$

   entonces $R.C = \langle 66, \infty\rangle$
5. Cálculo de $\bar{x}$: del enunciado $\bar{x} = 68$
6. Conclusión $\bar{x} = 68 \in \langle 66, \infty\rangle$, rechazamos H_0 , es decir los estudiantes de esta ciudad saben más inglés con un nivel de significación 0.05.

SOLUCION ALTERNATIVA

En el paso 4, de $P[Z > z_\alpha] = 1 - P[Z \leq z_\alpha] = \alpha$ se obtiene $z_{1-\alpha} = 1.64$. Es decir la región de rechazo es

$$R.C = \langle 1.64, \infty\rangle$$

En el paso 5, una vez conocida $\bar{x}$ calcular

$$z = \frac{\bar{x} - \mu_o}{\sigma/\sqrt{n}} = \frac{68 - 64}{s/\sqrt{54}} = 3.27$$

Conclusión: puesto que $z = 3.27 \in \langle 1.64, \infty\rangle$, se rechaza H_o .

EJEMPLO 6 Se han realizado 26 experimentos para estudiar el contenido de productos embasados, encontrándose que la media es de 240 gr. y s = 10 gr. ¿Es suficiente prueba para pensar que el verdadero promedio sea mayor que 238 gr? use $\alpha = 0.05$.

SOLUCION

1. $H_o : \mu = 238$ y $H_1 : \mu > 238$
2. $\alpha = 0.05$
3. Puesto que $n = 26$ es pequeño, usaremos la distribcuión **t**. Es decir

   $T = \dfrac{\bar{X} - \mu_o}{s/\sqrt{n}}$ tiene una distribcuión **t** con $n - 1 = 25$ grados de libertad.
4. Región de rechazo $T > t_{1-\alpha} = 1.708$, o sea

   $$R.C = \langle 1.708, \infty\rangle .$$
5. Cálculo de $\bar{x} = 240$, $s = 10$, $n = 26$

   y $$t = \frac{\bar{x} - \mu_o}{s/\sqrt{n}} = \frac{240 - 238}{10/\sqrt{26}} \approx 1.02$$
6. Conclusión: ya que $t = 1.02 \notin R.C = \langle 1.708, \infty\rangle$, no se rechaza H_o ; es decir no es suficiente prueba para penzar que el verdadero promedio sea mayor que 238.

EJEMPLO 7 Una fábrica produce clavos cuya longitud media es de 1 pulgada. Después de efectuadas algunas modificaciones en los dispositivos de las máquinas de dicha fábrica y con respecto a la producción de clavos durante los últimos meses se han recibido continuos reclamos de los compradores quienes han manifestado que los clavos presentan un incremento en más de 0.1 pulgadas en su longitud, lo que perjudica a los usuarios; para verificar lo manifestado por los compradores, el fabricante tomó una m.a. de 10 clavos cuyas longitudes resultaron:

1.14, 1.12, 1.11, 1.10, 1.16, 1.09, 1.08, 1.12, 1.11, 1.10

(a) Usando $\alpha = 0.05$, podrá el fabricante aceptar lo manifestado por los compradores?

(b) Construir un intervalo de confianza del 95% para la longitud media de los clavos fabricados después de las modificaciones efectuadas en los dispositivos de las máquinas.

SOLUCION

1. $H_0 : \mu = 1$ y $H_1 : \mu > 1$
2. $\alpha = 0.05$
3. Debido a que $n = 10$ es pequeño, empleamos como estadística de prueba la variable aleatoria $T = \dfrac{\bar{X} - \mu_o}{s/\sqrt{n}}$, que tiene una distribución t con $n - 1 = 9$ grados de libertad.
4. Región Crítica: $T > t_{1-\alpha} = 1.833$

   luego, $R.C = \langle 1.833, \infty \rangle$
5. Cálculo de $\bar{x}$ y s

$$\bar{x} = \frac{\Sigma f_i x_i}{n} = \frac{11.13}{10} = 1.113$$

$$s^2 = \frac{0.00501}{9}$$

$$s \simeq 0.0236$$

$$y \quad t = \frac{1.113 - 1}{0.0236/\sqrt{10}} \simeq 15.14$$

x_i	f_i	$f_i x_i$	$f_i(x_i - \bar{x})^2$
1.08	1	1.08	0.001089
1.09	1	1.09	0.000529
1.10	2	2.20	0.000338
1.11	2	2.22	0.000018
1.12	2	2.24	0.000098
1.14	1	1.14	0.000729
1.16	1	1.16	0.002209
	10	11.13	0.00501

6. Conclusión: desde que $t = 15.14 \in R.C$, rechazamos H_0 es decir el fabricante debe aceptar lo manifestado por los compradores.

   (b) Queda como ejercicio para el lector.

9.2.2 PRUEBA BILATERAL DE UNA HIPOTESIS SOBRE LA MEDIA

Consideremos ahora la prueba bilateral.

1. $H_0 : \mu = \mu_0$

   $H_1 : \mu \neq \mu_0$

   donde μ_0 es un valor de la media poblacional.
2. Escogemos el nivel de significación o de riesgo α.

3. La estadística de prueba para el parámetro μ es $\bar{X}$. Si la población es normal o la muestra es grande $(n \geq 30)$ aún si la población no es normal, la variable aleatoria $\bar{X}$ tiene una distribución $N(\mu, \sigma^2/n)$ o aproximadamente $N(\mu, \sigma^2/n)$.
4. En este caso es razonable rechazar H_0 a favor H_1 si la media muestral es demasiado grande o demasiado pequeña con relación a μ_0. Es decir la región de rechazo se divide en dos partes. La región más pequeña de aceptación de H_0, se obtiene cuando la probabilidad α se divide en dos partes iguales entre los extremos (colas) de la distribcuión de $\bar{X}$. En la figura 9.2.6 se representa esta división. Por la simetría de la curva normal, los valores críticos a y b son simétricos con respecto a μ_0. Entonces

   R.A = $\langle a,b \rangle$ donde a y b son tal que $P[a < \bar{X} < b] = 1 - \alpha$ ó

$$P\left[\frac{a-\mu_0}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < \frac{b-\mu_0}{\sigma/\sqrt{n}}\right] = 1 - \alpha$$

   de donde $a = \mu_0 - z_{\alpha/2}\,\sigma/\sqrt{n}$

   $b = \mu_0 + z_{\alpha/2}\,\sigma/\sqrt{n}$

   Luego, la R.A $= \left\langle \mu_0 - z_{\alpha/2}\,\sigma/\sqrt{n}\,,\ \mu_0 + z_{\alpha/2}\,\sigma/\sqrt{n} \right\rangle$ (*)

5. Cálculo de $\bar{x}$ a partir de una muestra observada (también s, si no se cononoce σ^2 y n es grande).
6. Conclusión: si $\bar{x} \notin$ R.A se rechaza H_0, y si $\bar{x} \in$ R.A, aceptar H_0

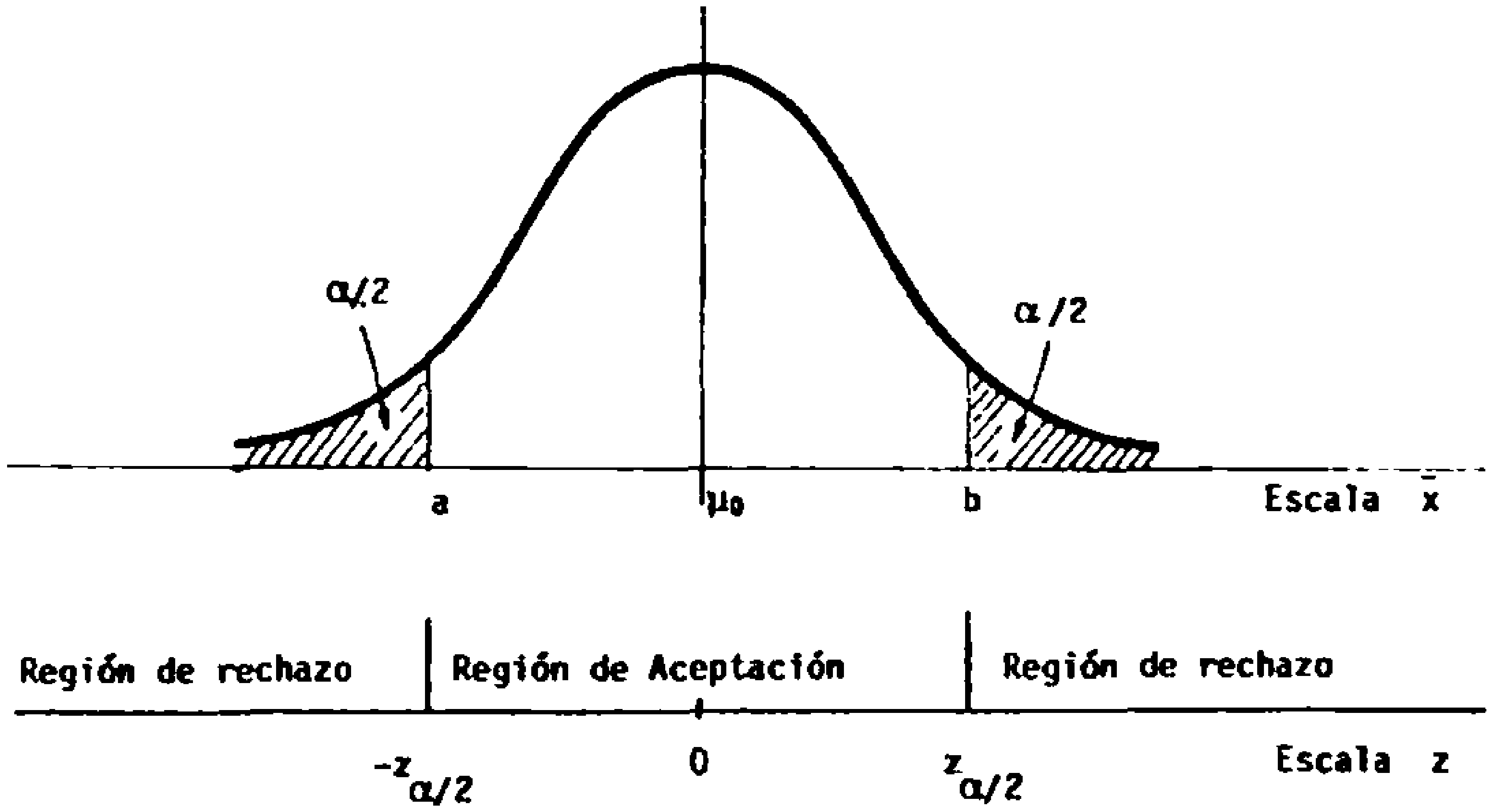


Fig. 9.2.6

NOTA Un método alternativo, es el siguiente:

(a) en el paso 4 considerar R.A = $\langle -z_{\alpha/2}, z_{\alpha/2}\rangle$ (ver fig. 9.2.6)

(b) en el paso 5, una vez conocido $\bar{x}$, calcular $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

(c) en el paso 6, será, aceptar H_0, si $z \in \langle -z_{\alpha/2}, z_{\alpha/2}\rangle$ en caso contrario rechazar H_0.

$$* \quad R.A = \{(X_1, X_2, \ldots, X_n): \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}$$

$$= \{(X_1, X_2, \ldots, X_n): -z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu_0 < z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}$$

$$= \{(X_1, X_2, \ldots, X_n): \bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu_0 < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}$$

Es decir, la región de aceptación (de no rechazo) es equivalente a este último conjunto. El lector podrá observar que este último conjunto es igual al intervalo de confianza construido para la media en el capítulo anterior, esto nos lleva a preguntarnos, si es posible probar hipótesis vía intervalos de confianza. La respuesta es afirmativa. Si queremos probar $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$ a un vivel α, construimos un intervalo de confianza de 100(1 - α)% -

para μ_o . Si $\mu_o \epsilon$ I.C. entonces rechazamos la hipótesis. Recíprocamente, teniendo la región crítica para la prueba bilateral ($H_o : \mu = \mu_o$ contra $H_1 : \mu \neq \mu_o$) a partir de esta región crítica podemos construir un intervalo de confianza para μ. Estas mismas conclusiones pueden obtenerse para los otros casos de pruebas de hipótesis con los correspondientes intervalos de confianza. El lector interesado en profundizar sobre estos aspectos puede ver HOGG-CRAIG. prueba de la razón de verosimilitud.

MUESTRAS PEQUEÑAS

Si se trabaja con muestras pequeñas de poblaciones que tienen una distribución aproximadamente normal, se usa la variable aleatoria,

$$T = \frac{\bar{X} - \mu_o}{s/\sqrt{n}}$$

que tiene una distribución t con $n - 1$ grados de libertad. En este caso la región de aceptación es,

$$R.A = \langle a,b \rangle \quad \text{tal que} \quad P[a < \bar{X} < b] = 1 - \alpha$$

$$P\left[\frac{a - \mu_o}{s/\sqrt{n}} < \frac{\bar{X} - \mu_o}{s/\sqrt{n}} < \frac{b - \mu_o}{s/\sqrt{n}}\right] = P[- t_{\alpha/2} < T < t_{\alpha/2}] = 1 - \alpha$$

Suponiendo que a y b son puntos simétricos respecto de μ_o

luego, $a = \mu_o - t_{\alpha/2}\, s/\sqrt{n}$

$b = \mu_o + t_{\alpha/2}\, s/\sqrt{n}$

y $\quad R.A = \langle \mu_o - t_{\alpha/2}\, s/\sqrt{n} \; , \; \mu_o + t_{\alpha/2}\, s/\sqrt{n} \rangle$

ó $\quad R.A = \langle - t_{\alpha/2} \, , \, t_{\alpha/2} \rangle$

la región de rechazo será; R.C: $t < - t_{\alpha/2}$ ó $t > t_{\alpha/2}$.

EJEMPLO 8 Los registros del puntaje de un test de aptitud académica tomado a los aspirantes a primer año de la Universidad revelaron que el puntaje medio aritmético era 110 y la desviación típica (o standar) es 10. Aplicado el test a los nuevos aspirantes se vió que para una muestra de 400 el puntaje medio era 106.

(a) Hay razón para creer que el reendimiento medio de estos nuevos aspirantes ha cambiado?. Use $\alpha = 0.01$.

(b) Encuentre un intervalo de confianza del 95% para el puntaje medio de estos nuevos aspirantes.

SOLUCION (a) Resolveremos usando el método alternativo.

1. $H_0 : \mu = 110$ y $H_1 : \mu \neq 110$
2. $\alpha = 0.01$
3. Puesto que la muestra es grande $n = 400$ y se conoce la desviación típica, la distribución de la estadística

   $\bar{X}$ es $N(110, \frac{100}{400})$

   luego $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ es $N(0,1)$

4. Región Crítica: $Z < - z_{\alpha/2} = 2.58$ y $Z > z_{\alpha/2} = 2.58$

   o $R.A = \langle - 2.58, 2.58 \rangle$.

5. Cálculo de $\bar{x}$. De los datos $\bar{x} = 106$, y $\sigma = 10$, luego,

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{106 - 110}{10/\sqrt{400}} = - 8$$

6. Conclusión: puesto que $z = - 8 \notin R.A = \langle - 2.58, 2.58 \rangle$, rechazamos H_0 ; es decir que hay razón para creer que el rendimiento ha cambiado.

SOLUCION ALTERNATIVA

En el paso 4 se calcula, la región de aceptación

$$R.A = \langle a,b \rangle \text{ tal que } P[a < \bar{X} < b] = 0.99$$

donde $a = \mu_0 - z_{\alpha/2}\sigma/\sqrt{n} = 110 - \dfrac{(2.58)10}{20} = 108.71$

$$b = \mu_0 + z_{\alpha/2}\,\alpha/\sqrt{n} = 110 + \frac{(2.58)10}{20} = 111.29$$

del paso 5. es $\bar{x} = 106 \notin \langle 108.71, 11.29 \rangle$ luego rechazamos H_0 .

(b) Queda como ejercicio para el lector verificar que el intervalo de confianza del 95% para el puntaje medio es $\langle 105.02, 106.98 \rangle$. Observe que $110 \notin \langle 105.02, 106.98 \rangle$

EJEMPLO 9 La estatura media de los alumnos de cierta universidad es de 1.70 m. con una desviación típica de 7 cm. ¿Hay razón para creer que se ha producido un cambio en la estatura promedio, si una muestra de 50 estudiantes dió una estatura promedio de 1.74 m? Use $\alpha = 0.02$.

SOLUCION

1. $H_0 : \mu = 1.70$ y $H_1 : \mu \neq 1.70$
2. $\alpha = 0.02$
3. Desde que la muestra es grande $n = 50$ y se conoce la desviación estandar, la distribución del estadístico $\bar{X}$ es $N(1.70, \frac{49}{50})$ y la variable aleatoria - $Z = \frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}}$ tiene una distribución $N(0,1)$.
4. Región Crítica: $Z < -2.34$ y $Z > 2.34$. Es decir

$$R.A = \langle -2.34, 2.34 \rangle .$$

5. $\bar{x} = 1.74$ m $\sigma = 0.07$ m, entonces

$$z = \frac{\bar{x} - \mu_o}{\sigma/\sqrt{n}} = \frac{1.74 - 1.70}{0.07/\sqrt{50}} \approx 4.04$$

6. Conclusión: Desde que $z = 4.04 \notin R.A$, rechazamos H_o ; es decir hay razón para crear que se ha producido un cambio en la estatura promedio.

EJEMPLO 10 Una máquina para enlatar conservas de pescado ha sido regulada para que el contenido de cada lata sea de 16 onzas. Usando $\alpha = 0.05$, ¿Diría Ud. que la máquina ha sido adecuadamente regulada, si una muestra de 20 latas dió un peso medio de 16.05 onzas y una desviación típica de 1.5 onzas?
Con los datos anteriores, calcule también un intervalo de confianza de 95% para el contenido medio.

SOLUCION

1. $H_o : \mu = 16$ y $H_1 : \mu \neq 16$
2. $\alpha = 0.05$
3. Puesto que $n = 20$ es pequeño y suponiendo que la población tiene distribución aproximadamente normal, usamos la variable aleatoria

$T = \frac{\bar{X} - \mu_o}{s/\sqrt{n}}$ que tiene una distribución t con $n - 1 = 19$ grados de libertad como estadística de prueba.

4. Región Crítica: $T < -t_{\alpha/2} = -2.093$ ó $T > t_{\alpha/2} = 2.093$

con $\alpha/2 = 0.025$, para buscar en la tabla tomamos

$1 - \alpha/2 = 0.975$ luego, R.A = $\langle -2.093, 2.093 \rangle$

5. De los datos $\bar{x} = 16.05$, $s = 1.5$ para $n = 20$, entonces

$$t = \frac{\bar{x} - \mu_o}{s/\sqrt{n}} = \frac{16.05 - 16}{1.5/\sqrt{20}} \simeq 0.148$$

6. Conclusión: desde que $t = 0.148 \in$ R.A, aceptamos H_o ; es decir se acepta que la máquina ha sido adecuadamente regulada.
   La obtención del intervalo de confianza del 95% queda como ejercicio para el lector.

EJEMPLO 11 Durante los tres primeros meses de vida el aumento de peso registrado por cierto animal fué de 65 gr. Desde el nacimiento hasta los 3 meses de edad; una docena de estos animales fueron alimentados con determinada dieta y los aumentos de peso observados fueron los siguientes (en gramos).

61, 62, 67, 59, 62, 60, 63, 65, 58, 54, 62, 55

¿Hay razón para creer, al nivel de significación del 5% que la dieta originó cambio en el aumento de peso?

SOLUCION

1. $H_o : \mu = 65$ y $H_1 : \mu \neq 65$
2. $\alpha = 0.05$
3. Puesto que el tamaño de la muestra $n = 12$ es pequeña y no se conoce la desviación estándar de la población asumiendo que tiene una distribución aproximadamente normal utilizamos la variable aleatoria,

$$T = \frac{\bar{X} - \mu_o}{s/\sqrt{n}}$$

   que tiene una distribución t con $n - 1 = 11$ grados de libertad.
4. Región crítica: $T < -t_\alpha = -2.281$, ó $T > t_\alpha = 2.281$

   luego, R.A = $\langle -2.281, 2.281 \rangle$
5. Cálculo de $\bar{x}$ y s .

$$\bar{x} = \frac{\Sigma f_i x_i}{n} = \frac{728}{12} \simeq 60.66 .$$

$$s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2 f_i}{n - 1}}$$

$= \sqrt{\dfrac{156.6672}{11}} \approx 3.77$

x_i	f_i	$f_i x_i$	$f_i(x_i - \bar{x})^2$
54	1	54	44.3556
55	1	55	32.0356
58	1	58	7.0756
59	1	59	2.7556
60	1	60	0.4356
61	1	61	0.1156
62	3	186	5.3868
63	1	63	5.4756
65	1	65	18.8356
67	1	67	40.1956
	12	728	156.6672

luego, $t = \dfrac{\bar{x} - \mu_o}{s/\sqrt{n}} = \dfrac{60.66 - 65}{3.77/\sqrt{12}} \approx -3.99$

6. Conclusión: desde que $t = -3.99 \notin$ R.A, rechazamos H_o ; es decir la dieta origino un cambio en el aumento de peso, con un nivel de significación del 5%.

EJEMPLO 12 Un ingeniero industrial afirma que la temperatura de fusión de cierto material de fierro es 1250°C. Se toma una muestra aleatoria de tamaño 4 de este material y se obtiene como temperatura de fusión los siguientes resultados, 1060°C, 1260°C, 1380°C, 1200°C, se pregunta

(a) La afirmación hecha por el ingeniero está de acuerdo con los resultados obtenidos, al 5% de nivel de significación?

(b) construir un intervalo de confianza del punto medio de fusión con un grado de confiabilidad del 99%.

SOLUCION (a)

1. $H_o : \mu = 1250$ y $H_1 : \mu \neq 1250$
2. $\alpha = 0.05$
3. Desde que la muestra es pequeña $n = 4$, asumiendo que la población tiene aproximadamente una distribución normal, usamos la variable aleatoria,

$T = \dfrac{\bar{X} - \mu_o}{s/\sqrt{n}}$, que tiene una distribución t con $n - 1 = 3$ grados de libertad.

4. Región Crítica: $T < -t_{\alpha/2} = -3.182$, $T > 3.182$, luego,

R.A = $\langle -3.182, 3.182 \rangle$

5. Cálculo de $\bar{x}$ y s, con $n = 4$.

x_i	1060	1260	1380	1200	4900
$(x_i - \bar{x})^2$	27225	1225	24025	625	53100

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{4900}{4} = 1225 .$$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{53.100}{3}} \simeq 133 .$$

$$\text{luego,} \quad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{1225 - 1250}{133/2} \simeq -0.376 .$$

6. Conclusión: desde que $t = -0.376 \in$ R.A, no rechazamos H_0 ; es decir no hay razón para rechazar la afirmación del ingeniero.

(b) queda como ejercicio para el lector verificar que el intervalo de confianza del 99% para la media es

$$\langle 836.57, 1613.43 \rangle$$

observe que $\mu_0 = 1250 \in \langle 836.57, 1613.43 \rangle$

9.2.3 PRUEBA DE HIPOTESIS SOBRE LA DIFERENCIA ENTRE MEDIAS

En muchos problemas prácticos se está interesado en determminar si existe o no una diferencia significativa entre las medias μ_X y μ_Y de dos poblaciones o variables aleatorias X y Y. La prueba de hipótesis que comprenden dos medias, son las mismas que la de una sola media, salvo que se necesitan dos muestras, una de cada población. La hipótesis nula suele escribirse así,

$$H_0 : \mu_X = \mu_Y \quad \text{ó} \quad H_0 : \mu_X - \mu_Y = 0$$

y la hipótesis alternativa toma una de las siguientes formas:

(a) $H_1 : \mu_X \neq \mu_Y$ ó $H_1 : \mu_X - \mu_Y \neq 0$

(b) $H_1 : \mu_X < \mu_Y$ ó $H_1 : \mu_X - \mu_Y < 0$

(c) $H_1 : \mu_X > \mu_Y$ ó $H_1 : \mu_X - \mu_Y > 0$

Si H_1 toma la forma (a) se utiliza prueba bilateral, en los otros casos se emplea la prueba unilateral.

DESVIACIONES TIPICAS σ_X Y σ_Y CONOCIDAS, MUESTRAS GRANDES

A. PRUEBA UNILATERAL

PRIMER CASO 1 Consideremos la siguiente prueba:

$$H_0 : \mu_X = \mu_Y \quad \text{ó} \quad \mu_X - \mu_Y = 0$$
$$H_1 : \mu_X > \mu_Y \quad\quad \mu_X - \mu_Y > 0$$

2. Escogemos el nivel de significación α.
3. La estadística para la diferencia de medias poblacionales $\mu_X - \mu_Y$ es la diferencia de medias muestrales $\bar{X} - \bar{Y}$. Si la población tiene distribución normal con desviaciónes estándar σ_X y σ_Y conocidas (o si las muestras son grandes $n \geq 30$, $m \geq 30$ aún cuando la población no tenga distribución normal), la distribución de $\bar{X} - \bar{Y}$ es normal con media $\mu_X - \mu_Y$ y varianza $\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}$. Por lo tanto la variable aleatoria,
$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}}$$
tiene una distribución normal estándar.
4. Es razonable rechazar H_0 en favor de H_1, si la diferencia de las medias muestrales es demasiado grande con respecto a 0. Una región crítica puede obtenerse escogiendo el valor crítico $\bar{x}_c$ de la diferencia de medias de manera que,
$$R.C = \langle \bar{x}_c , \infty \rangle \text{ tal que, } P[\bar{X} - \bar{Y} > \bar{x}_c | H_0] = \alpha \quad \text{ó}$$
$$P\left[\frac{\bar{X} - \bar{Y} - 0}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} > \frac{\bar{x}_c - 0}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}}\right] = P\left[Z > \frac{\bar{x}_c}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}}\right] = \alpha$$
en el supuesto de que H_0 es verdadero, es decir $\mu_X - \mu_Y = 0$. De la tabla obtenemos, $z_{1-\alpha}$ (ver fig. 9.2.7), luego,

$$\bar{x}_c = z_{1-\alpha}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$$

5. Cálculo de la diferencia de medias muestrales $\bar{x} - \bar{y}$ y $s_{\bar{x}-\bar{y}}$ si no se conoce σ_X^2 , σ_Y^2. Para muestras grandes se utiliza $s_{\bar{X}-\bar{Y}}$
6. Conclusión: si $\bar{x} - \bar{y} \in$ R.C, se rechaza H_o , en caso contrario se acepta.

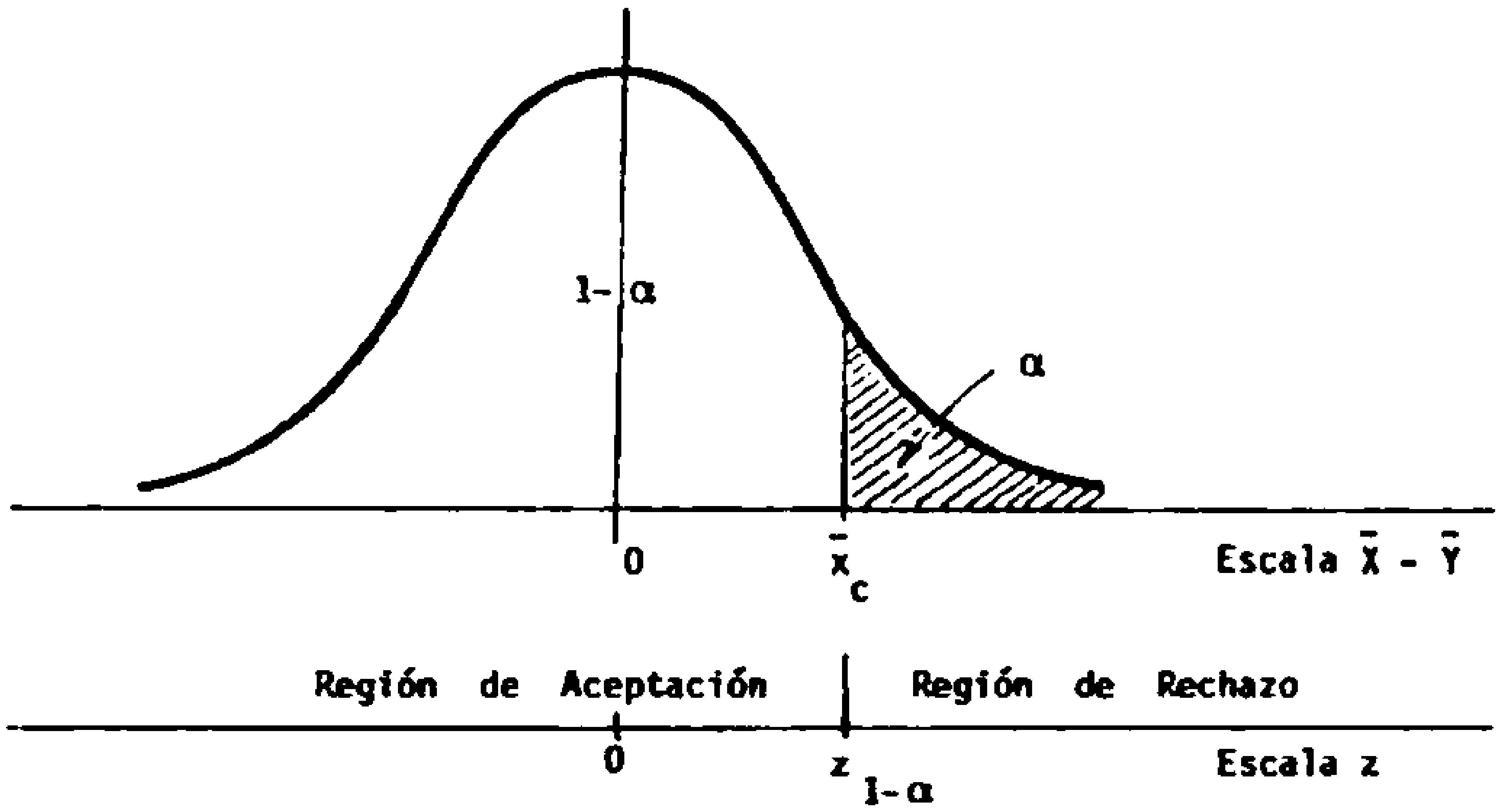


Fig. 9.2.7

METODO ALTERNATIVO Un método alternativo consta de los siguientes pasos:

(a) en el paso 4 considerar la región crítica en la escala z, (ver Fig. 9.2.7).

$$R.C = \langle z_{1-\alpha}, \infty \rangle$$

(b) En el paso 5, una vez calculado $\bar{x} - \bar{y}$, calcular z tal que

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

(c) El paso 6, si $z \in R.C = \langle z_{1-\alpha}, \infty \rangle$, rechazar H_o , en caso contrario - aceptar.

SEGUNDO CASO Si la prueba es $H_o : \mu_X = \mu_Y$, $H_1 : \mu_X < \mu_Y$, se sigue el - mismo procedimiento anterior. Excepto que $R.C = \langle -\infty, \bar{x}_c \rangle$ tal que $P[\bar{X} - \bar{Y} < \bar{x}_c | H_o] = \alpha$

B. **PRUEBA BILATERAL**

1. Consideremos ahora la prueba bilateral

$$H_0 : \mu_X = \mu_Y \quad ó \quad \mu_X - \mu_Y = 0$$
$$H_1 : \mu_X \neq \mu_Y \quad ó \quad \mu_X - \mu_Y \neq 0$$

2. Elegir el nivel de significación α.

3. La estadística de prueba es la misma que en los casos anteriores.

4. La región crítica. Obtendremos la región de aceptación por facilidad

$$R.A = \langle a,b \rangle \text{ tal que, } P[a < \bar{X} - \bar{Y} < b \mid H_0] = 1 - \alpha$$

$$ó \quad P\left[\frac{a - 0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} < \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} < \frac{b - 0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}\right] = 1 - \alpha$$

en el supuesto de que H_0 es verdadera, es decir $\mu_X - \mu_Y = 0$. El caso más simple es cuando $a = -b$, luego las áreas en las colas es $z_{\alpha/2}$. Es decir se tiene.

$$P[-z_{\alpha/2} < Z < z_{\alpha/2}] = 1 - \alpha$$

de donde, $$a = -z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$$

$$b = z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$$

5. Calcular $\bar{x} - \bar{y}$ de la muestra observada de tamaños n y m. También s_X^2 y s_Y^2 si no se conocen σ_X^2 y σ_Y^2 y las muestras son grandes.

6. Conclusión: si $\bar{x} - \bar{y} \in R.A = \langle a,b \rangle$, aceptar H_0, en caso contrario rechazar.

METODO ALTERNATIVO El método alternativo consiste en:

(a) considerar en el paso 4, la región de aceptación,

$$R.A = \langle -z_{\alpha/2}, z_{\alpha/2} \rangle$$

o región de rechazo $R.C$; $z < -z_{\alpha/2}$ ó $z > z_{\alpha/2}$.

(b) En el paso 5, calcular $z = (\bar{x} - \bar{y})\Big/\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$, una vez conocido $\bar{x} - \bar{y}$.

(c) Rechazar, H_o , si $z \notin \langle -z_{\alpha/2}, z_{\alpha/2} \rangle$ en caso contrario aceptar.

EJEMPLO 13 Dos grupos de 50 niños de una escuela elemental, han sido enseñados a leer por dos métodos diferentes. Una vez terminada la instrucción, una prueba de lectura da los siguientes resultados.

$$\bar{x} = 73.4, \quad \bar{y} = 70.2, \quad s_X = 9, \quad s_Y = 10$$

probar la hipótesis $\mu_X = \mu_Y$.

SOLUCION

1. $H_o : \mu_X = \mu_Y$ contra la alternativa; $H_1 : \mu_X > \mu_Y$.
2. Escogemos $\alpha = 0.05$.
3. Debido a que las muestras son grandes, entonces $\bar{X} - \bar{Y}$ tiene una distribución normal. Es decir la variable aleatoria,

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \quad \text{es } N(0,1)$$

4. La región crítica es $z > z_\alpha$, $\alpha = 0.05$, da la tabla $z_\alpha = 1.64$ entonces R.C = $\langle 1.64, \infty \rangle$.
5. De los datos, $\bar{x} = 73.4$, $\bar{y} = 70.2$, $s_X = 9$, $s_Y = 10$ y $n = m = 50$, luego

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}} = \frac{73.4 - 70.2}{\sqrt{\frac{81}{50} + \frac{100}{50}}} = 1.68$$

6. Conclusión: desde que $z = 1.68 \in$ R.C = $\langle 1.64, \infty \rangle$, rechazamos H_o .

EJEMPLO 14 Para determinar el impacto de las escuelas sin ventanas sobre el desarrollo psicológico de los escolares, se sometió a una misma prueba de ansiedad a un grupo de 40 niños de una escuela sin ventanas y a un grupo de 30 niños de una escuela con ventanas. Los resultados de la prueba aparecen a continuación.

Escuelas sin ventanas: $\bar{x} = 117$, $s_X = 10$, $n = 40$

Escuelas con ventanas: $\bar{y} = 112$, $s_Y = 12$, $m = 30$

Si ud. está dispuesto a rechazar una hipótesis verdadera no más de 5 veces en 100 casos, ¿puede concluir que el impacto de los dos tipos de escuela sobre la ansiedad de los niños no es el mismo?

SOLUCION

1. $H_0 : \mu_X = \mu_Y$ ó $\mu_X - \mu_Y = 0$
2. $H_1 : \mu_X \neq \mu_Y$ ó $\mu_X - \mu_Y \neq 0$
3. $\alpha = 0.05$
3. Desde que las muestras son grandes $n = 40$ y $m = 30$, usamos la distribución normal. Es decir $\bar{X} - \bar{Y}$ tiene una distribución normal y

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}} \quad \text{es } N(0,1)$$

4. La región crítica es

$$z < -z_{\alpha/2} = -1.96 \quad \text{ó} \quad z > z_{\alpha/2} = 1.96$$

ya que $\alpha = 0.05$, $\alpha/2 = 0.025$. Luego la región de aceptación es,

R.A = $\langle -1.96, 1.96 \rangle$.

5. De los datos $\bar{x} = 117$, $s_X = 10$, $n = 40$

$\bar{y} = 112$, $s_Y = 12$, $m = 30$

luego $$z = \frac{\bar{x} - \bar{y} - (\mu_X - \mu_Y)}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}} = \frac{117 - 12 - 0}{\sqrt{\frac{100}{40} + \frac{144}{30}}} \simeq 1.85$$

bajo el supuesto que H_0 es verdadera, es decir $\mu_X - \mu_Y = 0$

6. Conclusión: desde que $z = 1.85 \in$ R.A, aceptamos H_0 .

DESVIACIONES TIPICAS σ_X Y σ_Y DESCONOCIDAS, MUESTRAS PEQUEÑAS

Si se quiere probar hipótesis sobre la diferencia de medias, bajo el supuesto que H_0 es verdadera; es decir $\mu_X - \mu_Y = 0$, cuando los tamaños de las muestras son pequeños y las poblaciones tienen distribuciónes normales, con desviaciones estándar iguales, se utiliza la variable aleatoria

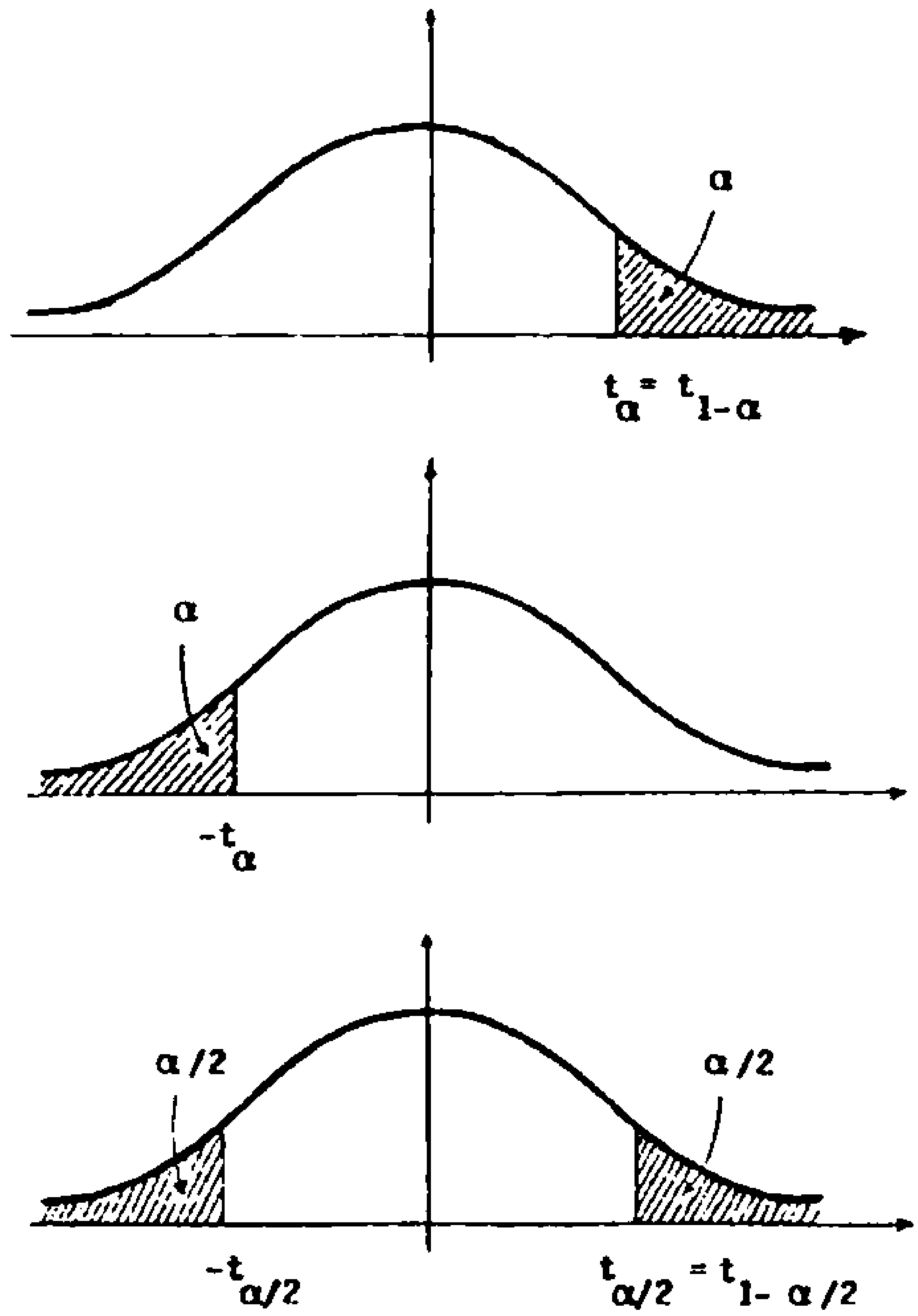
$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{(n-1)s_X^2 + (m-1)s_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}$$

que tiene una distribución t con $n + m - 2$ grados de libertad.

Cuando la hipótesis alternativa es

$H_1 : \mu_X - \mu_Y > 0$, se determina la región crítica, calculando $t_{1-\alpha}$ tal que

$P[T < t_{1-\alpha}] = 1 - \alpha$



luego, R.C = $\langle t_{1-\alpha}, \infty \rangle$

si la hipótesis alternativa es, $H_1 : \mu_X - \mu_Y < 0$, la región crítica se determina obteniendo t_α tal que:

$$P[T < t_\alpha] = \alpha$$

Es decir, R.C = $\langle -\infty, t_\alpha \rangle$

y finalmente cuando, la hipótesis alternativa es,

$H_1 : \mu_X - \mu_Y \neq 0$

la región crítica se determina obteniendo los valores

$-t_{\alpha/2}$ y $t_{\alpha/2}$ tal que

$$P[-t_{\alpha/2} < T < t_{\alpha/2}] = 1 - \alpha$$

$$R.C = \langle -\infty, -t_{\alpha/2} \rangle \cup \langle t_{1-\alpha/2}, \infty \rangle$$

Fig. 9.2.8

EJEMPLO 15 Diez barras de acero fabricadas por un proceso A tienen una fuerza de ruptura media de 50, con desviación estándar muestral de 10, mientras que ocho fabricadas por un proceso B tienen una fuerza de ruptura media de 55, con desviación estándar muestral de 12. Supóngase la población de fuerzas de ruptura normal con la misma desviación estándar. Pruébase con nivel de significación del 5% la hipótesis que los dos procesos producen acero de la misma fuerza en contra de la posibilidad que no es así.

SOLUCION

1. $H_0 : \mu_A = \mu_B$ y $H_1 : \mu_A \neq \mu_B$
2. $\alpha = 0.05$
3. Desde que las muestras son pequeñas y las poblaciones son normales, usamos la variable aleatoria.

$$T = \frac{\bar{X}_A - \bar{Y}_B - 0}{\sqrt{(n-1)s_A^2 + (m-1)s_B^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}$$

3. que tiene una distribución t, con $n + m - 2$ grados de libertad bajo el supuesto que H_0 es verdadera, es decir

$$\mu_A - \mu_B = 0$$

4. Región Crítica: $t < -t_{\alpha/2}$ ó $t > t_{\alpha/2} = t_{1-\alpha/2}$

$\alpha = 0.05$, $\alpha/2 = 0.025$, $1 - \frac{\alpha}{2} = 0.975$, de la tabla de la distribución **t**, con $10 + 8 - 2 = 16$ grados de libertad, obtenemos $t_{1-\alpha/2} = 2.12$. Por lo tanto $-t_{\alpha/2} = -2.12$

Es decir, R.A = $\langle -2.12, 2.12 \rangle$

5. De los datos $\bar{x}_A = 50$, $s_A = 10$, $n = 10$, $\bar{x}_B = 55$, $s_B = 12$ y $m = 8$.

luego $$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{(n-1)s_A^2 + (m-1)s_B^2}} \sqrt{\frac{n\,m(n+m-2)}{n+m}}$$

$$= \frac{55 - 50}{\sqrt{9 \times 10^2 + 7 \times 12^2}} \sqrt{\frac{80(16)}{18}} = 0.965.$$

6. Conclusión: desde que $t = 0.97 \in \langle -2.12, 2.12 \rangle$, no rechazamos H_0, es decir no hay razón para creer que los dos procesos producen acero con fuerzas diferentes.

EJEMPLOS 16 Como psicólogo de un hospital para enfermos mentales el lector obtiene calificaciones para una prueba visual-motora para cada uno de dos grupos de pacientes. La calificación media para el grupo **A** (10 pacientes) es 80 con desviación estándar 18, y la correspondiente al grupo **B** (15 pacientes) es 70 con desviación estándar 22. El lector cree tener suficiente razones para considerar las desviaciones estándar de población iguales. ¿Difieren significativamente las calificaciones con nivel del 10%?

SOLUCION

1. $H_0 : \mu_A = \mu_B$ y $H_1 : \mu_A \neq \mu_B$

2. $\alpha = 0.10$

3. Se usa la distribución t, debido a que las muestras son pequeñas.

4. La región crítica:

$$t < -t_{\alpha/2}, \qquad ó \qquad t > t_{\alpha/2} = t_{1-\alpha/2}$$

$\alpha = 0.10$, $\alpha/2 = 0.05$, $1 - \alpha/2 = 0.95$, con este valor y con $10 + 15 - 2 = 23$ grados de libertad, se encuentra

$t_{1-\alpha/2} = -t_{\alpha/2} = 1.714$

luego R.A = $\langle -1.714, 1.714 \rangle$; entonces,

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{(n-1)s_A^2 + (m-1)s_B^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}$$

$$= \frac{80 - 70}{\sqrt{9(18)^2 + 14(22)^2}} \sqrt{\frac{150(23)}{50}} = 1.193.$$

5. Conclusión: desde que,

$t = 1.193 \in \langle -1.714, 1.714 \rangle$, no se rechaza H_0, es decir la diferencia no es significativa a un nivel del 10 %.

EJEMPLO 17 Los siguientes datos dan el aumento de peso de 20 conejos, de las cuales la mitad recibió su proteína de maní crudo y la otra mitad de maní tostado. Probar si el tostado de maní ha tenido un menor efecto en el aumento del peso de los conejos. Los aumentos de peso entán registrados en gramos. Use $\alpha = 0.05$

CRUDO	61	60	56	63	56	63	59	56	44	61
TOSTADO	55	54	47	59	51	61	57	54	62	58

SOLUCION

1. $H_0 : \mu_C = \mu_T$ y $H_1 : \mu_T < \mu_C$

2. $\alpha = 0.05$

3. Se usa la distribución t, con $10 + 10 - 2 = 18$ grados de libertad, debido a que las muestras son pequeñas.

4. Región Crítica: $t < -t_{\alpha}$

4. $\alpha = 0.05$, $1 - \alpha = 0.95$ con este valor y 18 grados de libertad se encuentra que $-t_\alpha = -1.734$.

$$R.C = \langle -\infty, -1.734\rangle$$

5. Cálculo de $\bar{x}_C$, $\bar{x}_T$, s_C^2, s_T^2.

$$\bar{x}_T = \frac{\Sigma f_i x_i}{n} = \frac{558}{10} = 55.8 \; ; \qquad \bar{x}_C = \frac{\Sigma f_i x_i}{n} = \frac{579}{10} = 57.9$$

$$s_T^2 = \frac{\Sigma f_i (x_i - \bar{x})^2}{n-1} = \frac{189.6}{9} \; ; \qquad s_C^2 = \frac{\Sigma f_i (x_i - \bar{x})^2}{n-1} = \frac{280.9}{9}$$

luego, $$t = \frac{55.8 - 57.9}{\sqrt{9(\frac{189.6}{9}) + 9(\frac{280.9}{9})}} \sqrt{\frac{100(18)}{20}} = 0.92$$

TOSTADO

x_i	f_i	$x_i f_i$	$f_i(x_i - \bar{x}_T)^2$
47	1	47	77.44
51	1	51	23.04
54	2	108	6.48
55	1	55	0.64
57	1	57	1.44
58	1	58	4.84
59	1	59	10.24
61	1	61	27.04
62	1	62	38.44
	10	558	189.60

CRUDO

x_i	f_i	$f_i x_i$	$(x_i - \bar{x}_T)^2 f_i$
44	1	44	93.21
56	3	168	10.83
59	1	59	1.21
60	1	60	4.41
61	2	122	19.22
63	2	126	52.02
	10	579	280.9

6. Conclusión: desde que $t = 0.92 > -1.734$, no se rechaza, H_0, es decir, el maní tostado no ha tenido efecto negativo en el peso de conejos a un nivel significación del 5 %.

EJEMPLO 18 Se estudia el contenido de nicotina en los cigarrillos de dos marcas *A* y *B*, obteniéndose los siguientes resultados.

A: 17, 20, 20, 23

B: 18, 20, 21, 22, 24

con $\alpha = 0.05$. Determinar si es posible llegar a la conclusión que el contenido de nicotina en ambas marcas es diferente.

SOLUCION

1. $H_o : \mu_A = \mu_B$ y $H_1 : \mu_A \neq \mu_B$
2. $\alpha = 0.05$
3. Se usa la distribución t, desde que las muestras son pequeñas.
4. La región crítica, es

$$t < -t_{\alpha/2} = -2.365 \quad ó \quad t > t_{\alpha/2} = 2.365$$

donde $\alpha = 0.05$, $\alpha/2 = 0.025$, $1 - \alpha/2 = 0.975$, con este valor y $m + n - 2 = 7$ grados de libertad, se busca en la tabla de t, encontrándose, $t_{1-\alpha/2} = 2.365$

5. Cálculo de $\bar{x}_A$, $\bar{x}_B$, s_A^2 y s_B^2.

MARCA A

x_i	$(x_i - \bar{x}_A)^2$
17	9
20	0
20	0
23	9
80	18

MARCA B

x_i	$(x_i - \bar{x}_B)^2$
18	9
20	1
21	0
22	1
24	9
105	20

$$\bar{x}_A = \frac{80}{4} = 20 \qquad \bar{x}_B = \frac{105}{5} = 21$$

$$s_A^2 = \frac{18}{3} \qquad s_B^2 = \frac{20}{4}$$

luego, $$t = \frac{21 - 20}{\sqrt{3(\frac{18}{3}) + 4(\frac{20}{5})}} \sqrt{\frac{20(7)}{9}} = 0.641 \text{ ,}$$

6. Conclusión: desde que

$t = 0.641 \in R.A = \langle -2.365, 2.365 \rangle$, no se rechaza, H_0 . Es decir la diferencia no es significativa con nivel de significación del 5%.

9.2.4 PRUEBA DE DIFERENCIA PAREADA

En la sección 9.2.3 se discutió la prueba de la diferencia entre las medias de dos poblaciones independientes. Por lo tanto las dos muestras (o grupos de datos) también eran independientes. En esta sección, se desarrollará un procedimiento para analizar la prueba de dos muestras (o dos grupos de datos) que estan *relacionados* ; es decir, los resultados de la primera muestra no son independientes de la segunda. Esta característica de dependencia de las muestras ocurre ya sea, por que se obtienen mediaciones repetidas con el mismo grupo de artículos o individuos o por que los artículos o individuos estan apareados según alguna característica. Por lo tanto no se puede usar como estadística de prueba la diferencia de medias muestrales, pues no podemos obtener la varianza de la distribución de medias muestrales. En consecuencia se trabajará en cualquiera de estos casos con la ***diferencia*** entre los valores de las observaciones, es decir, trabajaremos con un nuevo conjunto de datos formado por las diferencias apareadas de las muestras tal como se muestra en tabla siguiente. Debe notarse que el apareamiento de los datos ocurrió al planearse el experimento y no despues que los datos fueron seleccionados.

	Muestra		Diferencia
Observación	1	2	D_i
1	X_{11}	X_{21}	$D_1 = X_{11} - X_{21}$
2	X_{12}	X_{22}	$D_2 = X_{12} - X_{22}$
.	.	.	.
.	.	.	.
.	.	.	.
i	X_{1i}	X_{2i}	$D_i = X_{1i} - X_{2i}$
.	.	.	.
.	.	.	.
.	.	.	.
n	X_{1n}	X_{2n}	$D_n = X_{1n} - X_{2n}$

donde X_{1i} es el i-ésimo valor en la muestra 1.

X_{2i} es el i-ésimo valor en la muestra 2.

$D_i = X_{1i} - X_{2i}$, la diferencia entre el i-ésimo valor en la muestra 1 y el i-ésimo valor en la muestra 2.

$$\bar{D} = \sum_{i=1}^{n} D_i/n \quad , \quad S_D = \sqrt{n\sum_{i=1}^{n} D_i^2 - (\sum_{i=1}^{n} D_i)^2/n(n-1)}$$

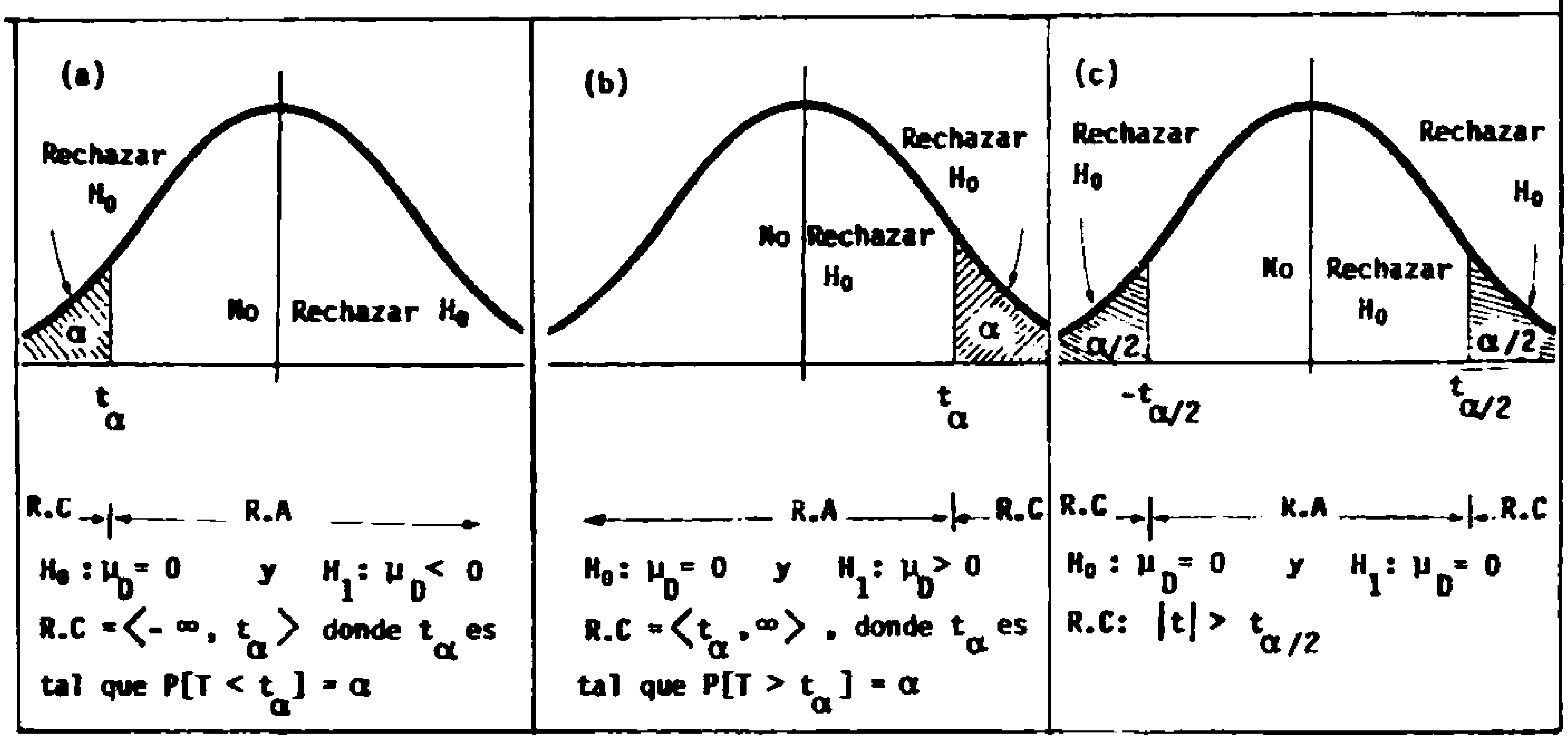
Para n suficientemente grande, por el teorema central del límite, la diferencia promedio $\bar{D}$ sigue una distribución normal cuando la desviación típica de la población de la diferencia σ_D es conocida. Sin embargo por lo general, sólo se dispone un número pequeño de datos y de la desviación típica muestral de la diferencia S_D , para poblaciones aproximadamente normal la variable aleatoria

$$T = \frac{\bar{D}\sqrt{n}}{S_D}$$

tiene una distribución t con $n - 1$ grados de libertad.
Por lo tanto el problema se reduce a probar una media de población hipotética, que ya hemos visto. Se empieza estableciendo:

1. H_o : $\mu_D = 0$ No hay diferencia alguna entre las dos muestras relacionados y H_1 : toma las formas $\mu_D < 0$, $\mu_D \neq 0$, $\mu_D > 0$
2. escoger el nivel significación α .
3. La estadística de prueba es $T = \dfrac{\bar{D}\sqrt{n}}{S_D}$ que tiene distribución t con $n - 1$ grados de libertad.
4. La región de rechazo para cada posible prueba de una cola y de dos colas se ven en la fig. 9.2.9

Fig. 9.2.9



5. Cálculo de $t = \dfrac{\bar{D}\sqrt{n}}{S_D}$

6. Conclusión: si $t \in R.C$ rechazar H_o .

EJEMPLO 19 Para comparar la efectividad de un programa de seguridad en el trabajo, se observó en 6 distintas plantas el número de accidentes por mes antes y después del programa. Los datos aparecen en la tabla siguiente. ¿Proporcionan estos datos evidencia suficiente para indicar que el programa ha sido efectivo al reducir el número de accidentes laborales por mes? Haga la prueba al $\alpha = 0.1$ de nivel de significación.

	PLANTA NUMERO					
	1	2	3	4	5	6
Antes del programa	38	64	42	70	58	30
Después del programa	31	58	43	65	52	29

SOLUCION

1. $H_o : \mu_D = 0$, $H_1 : \mu_D > 0$
2. El nivel de significación $\alpha = 0.1$
3. La estadística de prueba es $T = \dfrac{\bar{D}\sqrt{n}}{S_D}$ que tiene una distribución t con 5 grados de libertad.
4. La región de rechazo: es $t > t_\alpha$, donde t_α es tal que

   $P[T > t_\alpha] = 0.1$, de la tabla $t_\alpha = 1.4759$

   luego R.C: $t > 1.4759$
5. Cálculo de $\bar{D}$, S_D y t.

Observaciones	Muestra 1	Muestra 2	Diferencia D_i	D_i^2
1	38	31	7	49
2	64	58	6	36
3	42	43	-1	1
4	70	65	5	25
5	58	52	6	36
6	30	29	1	1
Total			24	148

$$\bar{D} = \frac{1}{6}\sum_{i=1}^{6} D_i = \frac{24}{6} = 4 \ .$$

$$S_D = \sqrt{\frac{n\sum_{i=1}^{n} D_i^2 - (\sum_{i=1}^{6} D_i)^2}{n(n-1)}} = \sqrt{\frac{6(148) - (24)^2}{6 \times 5}} = 2\sqrt{13}/\sqrt{5} \ .$$

$$t = \frac{\bar{D}\sqrt{n}}{S_D} = \frac{4\sqrt{6}}{\frac{2\sqrt{13}}{\sqrt{5}}} = \frac{2\sqrt{30}}{\sqrt{13}} = \frac{2(5.477)}{3.606} = 3.04.$$

6. Conclusión: 3.04 > 1.4759, se rechaza H_0 ; es decir, si hay evidencia

9.2.5 PRUEBA DE HIPOTESIS RELATIVA A LA VARIANZA DE UNA POBLACION

En principio, las pruebas de hipótesis relativas a la varianza de una población son iguales a las relaciones con las medias de la población.

La hipótesis nula toma la forma,

$$H_0 : \sigma^2 = \sigma_0^2$$

donde σ_0^2 es un valor conocido de la varianza poblacional. Y la hipótesis alternativa toma una de las siguientes formas.

(a) $H_1 : \sigma^2 < \sigma_0^2$

(b) $H_1 : \sigma^2 > \sigma_0^2$ pruebas unilaterales

(c) $H_1 : \sigma^2 \neq \sigma_0^2$, prueba bilateral

Veamos la prueba bilateral, los otros casos son completamente similares. Se siguen los siguientes pasos:

1. Formulación de las hipótesis:

   $H_0 : \sigma^2 = \sigma_0^2$

   $H_1 : \sigma^2 \neq \sigma_0^2$

2. Se escoge un nivel de significación α.

3. La estadística de prueba de la varianza poblacional σ^2, es la varianza muestral S^2. Hemos visto en el teorema 7.3.5.b , si extraemos una muestra aleatoria de una población normal, el cociente,

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2}$$

tiene una distribución chi-cuadrado con $n - 1$ grados de libertad

4. Región Crítica (Región de no aceptación): puesto que estamos considerando una prueba bilateral, la región de aceptación será:

R.A = $\langle a,b \rangle$ tal que, $P[a < S^2 < b] = 1 - \alpha$, (Fig. 9.2.10)

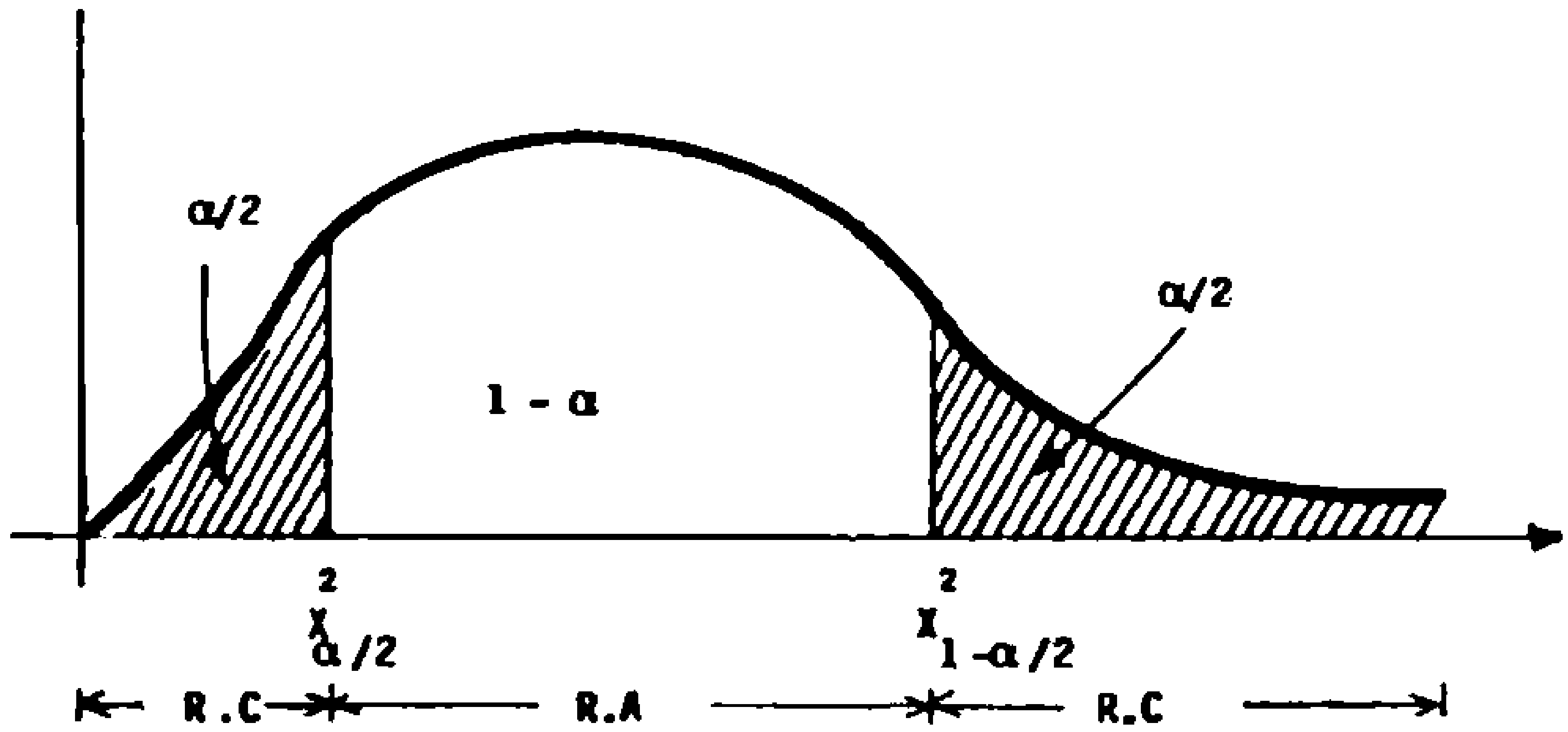


Fig. 9.2.10

ó $$P\left[\frac{(n-1)a}{\sigma_0^2} < \frac{(n-1)S^2}{\sigma_0^2} < \frac{(n-1)b}{\sigma_0^2}\right] = 1 - \alpha$$

$$P[\chi^2_{\alpha/2} < \chi^2 < \chi^2_{1-\alpha/2}] = 1 - \alpha$$

de donde se determinan a y b ,

$a = \chi^2_{\alpha/2}\,\sigma_0^2/n - 1$, $b = \chi^2_{1-\alpha/2}\,\sigma_0^2/n - 1$. Bajo el supuesto de que H_0 es verdadera.

5. Se calcula la varianza muestral s^2 , a partir de la muestra aleatoria observada.

6. Conclusión: si $s^2 \in \langle a,b \rangle$, se acepta H_0 , en otro caso se rechaza.

NOTA Una solución alternativa, es

(a) en el paso 4, considerar la región de aceptación

R.A = $\langle \chi^2_{\alpha/2} \, , \, \chi^2_{1-\alpha/2} \rangle$, o equivalente,

R.C : $\chi^2 < \chi^2_{\alpha/2}$ ó $\chi^2 > \chi^2_{1-\alpha/2}$

(b) se calcula, a partir de la varianza muestral obtenida

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

(c) Conclusión: si $\chi^2 < \chi^2_{\alpha/2}$ ó $\chi^2 > \chi^2_{1-\alpha/2}$, se rechaza H_0 , en otro caso se acepta.

EJEMPLO 20 Se ha puesto un examen durante varios años con $\mu = 70$ y $\sigma^2 = 9$. Una escuela que utiliza por primera vez este examen lo puso para 25 alumnos, que obtuvieron $\bar{x} = 71$ y una verianza $s^2 = 12$ ¿hay razón para creer que las calificaciones de todos los estudiantes de escuela tuvieron una varianza de 9 con un nivel de significación del 10%?

SOLUCION

1. planteamos las hipótesis

   $H_0 : \sigma^2 = 9$ y $H_1 : \sigma^2 \neq 9$

2. $\alpha = 0.1$

3. La estadística de prueba es S^2 . La distribución de $\frac{(n-1)S^2}{\sigma_0^2}$ es una chi-cuadrado con $n - 1$ grados de libertad.

4. La región crítica: $\chi^2 < \chi^2_{\alpha/2}$ ó $\chi^2 < \chi^2_{1-\alpha/2}$

   (ver nota anterior), $\alpha = 0.10$, $\alpha/2 = 0.05$, con este valor y $n - 1 = 24$ grados de libertad en la tabla encontramos $\chi^2_{\alpha/2} = 13.85$.

   $1 - \alpha/2 = 0.95$, luego, con este valor y 24 grados de libertad, se encuentra en la tabla $\chi^2_{1-\alpha/2} = 36.42$. Por lo tanto

   R.C: $\chi^2 < 13.85$ o $\chi^2 > 36.42$,

5. $s^2 = 12$, $n = 25$. Luego, $\chi^2 = \dfrac{24 \times 12}{9} = 32$

6. Conclusión: puesto que $\chi^2 = 32 \notin$ R.C, aceptamos H_0 . Es decir no hay razón para dudar que la varianza de las calificaciones de todos los estudiantes, de la escuela sea diferente de 9.

   La solución alternativa se obtiene, calculando en el paso

(4) $a = 13.85 \dfrac{9}{24} = 5.194$; $b = 36.42 \dfrac{9}{24} = 13.66$

   en el paso 5, $s^2 = 12 \in$ R.A $= \langle 5.194, 13.66 \rangle$, se acepta H_0 .

EJEMPLO 21 Suponga que al operar un equipo electrónico con energía de baterías, es menos costoso reemplazar todas las baterías a intervalos fijos que reemplazar cada batería individualmente cuando se ha gastado, siempre y cuando la desviación estándar del tiempo de vida sea menor que cierto límite, digamos, menor de 5 horas. Suponiendo normalidad y usando una muestra de 28 valores de tiempo de vida con desviación estándar s = 3.5 horas, probar la hipótesis $\sigma^2 = \sigma_0^2 = 25$ contra la alternativa $\sigma^2 < 25$.

SOLUCION

1. $H_0 : \sigma^2 = 25$ y $H_1 : \sigma^2 < 25$
2. Escogemos $\alpha = 0.05$
3. La estadística de prueba es $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$ se tiene distribución chi-cuadrado con $n - 1 = 27$ grados de libertad, supuesto que H_0 es verdadera.
4. La región Crítica es $\chi^2 < \chi^2_\alpha$

   $\alpha = 0.05$, con este valor y 27 grados de libertad, se encuentra en la tabla $\chi^2_\alpha = 16.15$
5. $s^2 = (3.5)^2$. Luego $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2} = \dfrac{27(3.5)^2}{25} = 13.23$
6. Conclusión: desde que $13.23 < 16.15$, se rechaza H_0 . Es decir será menos caro reemplazar todas las baterías simultáneamente.

9.2.6 PRUEBA DE HIPOTESIS RELATIVA A LAS VARIANZAS DE DOS POBLACIONES

Cuando se prueban hipótesis sobre la diferencia entre dos varianzas, la hipótesis nula sometida a prueba tiene la forma,

$$H_o : \sigma_X^2 = \sigma_Y^2$$

donde σ_X^2 y σ_Y^2 son las varianzas de dos poblaciones. La hipótesis alternativa, toma una de las siguientes formas

(a) $H_1 : \sigma_X^2 < \sigma_Y^2$, (b) $H_1 : \sigma_X^2 > \sigma_Y^2$, pruebas unilaterales

(c) $H_1 : \sigma_X^2 \neq \sigma_Y^2$ prueba bilateral

El procedimiento en la prueba de comparación de varianzas es el mismo que las pruebas de una sola varianza. Excepto que la estadística de prueba es la variable aleatoria

$$F = \frac{S_X^2}{S_Y^2}$$

que tiene una distribución F con $(n - 1)$ y $(m - 1)$ grados de libertad, donde S_X^2 y S_Y^2 son las varianzas muestrales obtenidas de muestras *independientes* de ambas poblaciones. Si las muestras son grandes, las varianzas muestrales deverían encontrarse próximos a los verdaderos valores de σ_X^2 y σ_Y^2 respectivamente, en cuyo caso el valor de F tendría que estar cerca de 1, si H_o fuera verdadera. Por lo tanto los valores observados de F cercanos a 0 ó mucho mayores que 1 conducirían al rechazo de H_o . Las tablas de la distribución F dan valores grandes por lo que es necesario poner en el numerador la mayor varianza muestral S_j , j = X,Y .

EJEMPLO 22 En el ejemplo 16 de 8.2.8 pruébese con el nivel de 0.05 la hipótesis $H_o : \sigma_1^2 = \sigma_2^2$ contra $H_1 : \sigma_1^2 \neq \sigma_2^2$ ..

SOLUCION

1. Formulación de las hipótesis $H_o : \sigma_1^2 = \sigma_2^2$ y $H_1 : \sigma_1^2 \neq \sigma_2^2$ prueba bilateral
2. Nivel de significación $\alpha = 0.05$.
3. La estadística de prueba, es la variable aleatoria $F = \dfrac{S_1^2}{S_2^2}$ que tiene una distribución F con $(n - 1) = 5$ y $(m - 1) = 5$ grados de libertad.
4. Región de rechazo: $F < f_{\alpha/2}$ o $F > f_{1-\alpha/2}$

   (ver Fig. 9.2.11), $\alpha = 0.05$, $\alpha/2 = 0.025$, entonces,

$$f_{\alpha/2,\,5,5} = \frac{1}{f_{1-\alpha/2,5,5}} \quad , \quad f_{1-\alpha/2,5,5} = 7.15$$

$f_{\alpha/2} = \dfrac{1}{7.15} = 0.14$, luego: R.C: $F < 0.14$, ó $F > 7.15$

ó R.A = $\langle 0.14,\ 7.15 \rangle$

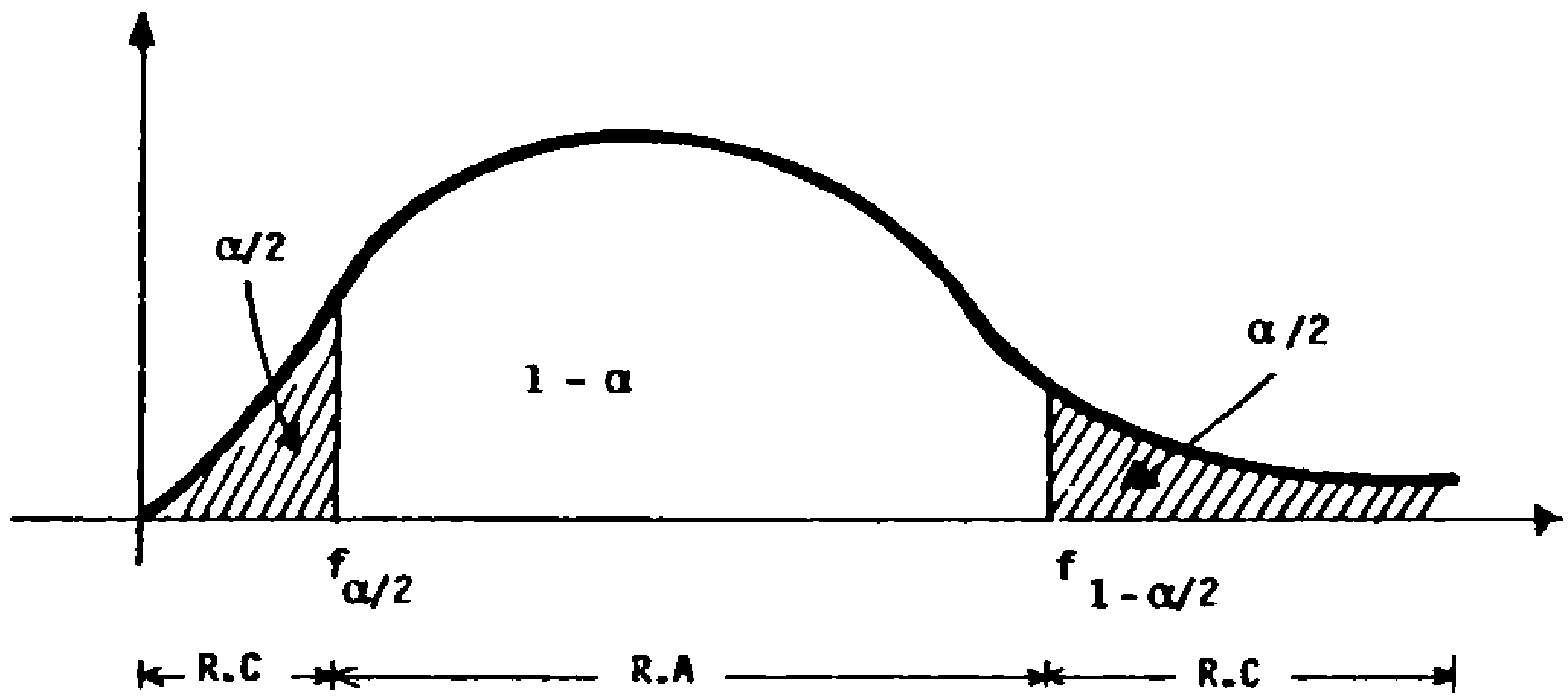


Fig. 9.2.11

5. Cálculo de s_1^2 y s_2^2 a partir de las muestras aleatorias de tamaños $n = 6$ y $m = 6$.
   De la solución del ejemplo 16, obtenemos:

$$s_1^2 = \frac{6312}{5} \ ; \ s_2^2 = \frac{637.5}{5} \ ; \ \text{luego} \ f = \frac{s_1^2}{s_2^2} = \frac{6312/5}{637.5/5} = 9.9 \ .$$

6. Conclusión: desde que $f = 9.9 > 7.15$, se rechaza H_0 . Es decir, hay razón para creer que las varianzas son diferentes.

9.2.7 PRUEBAS RELATIVAS A PROPORCIONES

Las pruebas de hipótesis con relación a proporciones son básicamente iguales a las relativas con medias. Consideremos el problema de probar la hipótesis que la proporción de éxitos en un proceso de Bernoulli (experimento binomial) es igual a un valor específico. Es decir, queremos probar la hipótesis nula

$$H_0 : p = p_0$$

donde p es el parámetro de la distribución binomial. La hipótesis alternativa toma una de las siguientes formas,

(a) $H_1 : p < p_0$; (b) $H_1 : p > p_0$, prueba unilateral

(c) $H_1 : p \neq p_0$, prueba bilateral

La estadística apropiada sobre la cual debemos basar nuestro criterio de decisión es la variable aleatoria binomial X que tiene una distribución binomial. Alternativamente puede usarse el estadístico $\hat{p} = \frac{X}{n}$. Los valores de X que están distante de la media $\mu = np_0$ conducirá al rechazo de la hipótesis nula.

(a) para probar la hipótesis.

$H_0 : p = p_0$
$H_1 : p > p_0$

Se usa la distribución binomial con $p = p_0$ y $q = 1 - p_0$ y se determina $P[X < x | H_0$ es verdadera$]$. El valor x es el número de éxitos en la muestra de tamaño n.

Si $P[X < x | H_0$ es verdadera$] < \alpha$, la prueba es significativa al nivel de significación α, se rechaza H_0 a favor de H_1. Similarmente, para probar la hipótesis,

$H_0 : p = p_0$
$H_1 : p > p_0$

al nivel de significación α se halla $P[X > x | H_0$ es verdadera$]$, si $P[X > x | H_0$, es verdadera$] < \alpha$ se rechaza H_0 en favor de H_1.

Analogamente, para probar la hipótesis

$$H_0 : p = p_0 \quad ; \quad H_1 : p_1 \neq p_0$$

al nivel de significación α ; se rechaza H_0 a favor de H_1 cuando $x < np_0$ y $P[X < x | H_0$ es verdadera$] < \alpha/2$ ó cuando $x > np_0$ y $P[X > x | H_0$ es verdadera$] < \alpha/2$.

El procedimiento para probar hipótesis a cerca de una proporción se puede resumir en los siguientes pasos:

1. $H_0 : p = p_0$

   H_1 : es una de las siguientes alternativas,

   $$p < p_0 \ , \quad p > p_0 \quad ó \quad p \neq p_0$$

2. Escoger el nivel de significación α
3. La estadística de prueba es la variable aleatoria binomial X que tiene una distribución binomial, cuando n es pequeño se usa esta distribución.
4. Región crítica

(a) Para la alternativa $p < p_0$, son todos los x tal que

$P[X < x | H_0$ es verdadera$] < \alpha$

(b) Para la alternativa $p > p_0$, son todos los x tal que

$P[X > x | H_0$ es verdadera$] < \alpha$

(c) Y para la alternativa $p \neq p_0$, son todos los x tal que

$P[X < x | H_0$ es verdadera$] < \alpha/2$ cuando $x < np_0$ todos los x tal que

$P[X > x | H_0$ es verdadera$] < \alpha/2$ cuando $x > np_0$.

5. Hallar x y calcular la probabilidad apropiada
6. Conclusión: rechazar H_0 si x pertenece a la región crítica, en otro caso aceptar H_0.

EJEMPLO 23 Un fabricante de cigarrillos asegura que el 20% de los fumadores de cigarrillos prefieren A, para probar esta aseveración toma una muestra aleatoria de 20 fumadores de cigarrillos y se les pregunta por la marca que prefieren. Si de los 20 fumadores, 6 prefieren la marca A, ¿Qué concluye? Use $\alpha = 0.01$.

SOLUCION

1. $H_0 : p = 0.2$; $H_1 : p \neq 0.2$
2. $\alpha = 0.01$
3. Se usa la distribución binomial, pues n es pequeña.
4. Región crítica: todos los x tal que $P[X > x | H_0$ es verdadera$] < 0.005$, pues $np_0 = 20(0.2) = 4 < x = 6$
5. Cálculo: $x = 6$, $n = 20$, y usando la tabla I. (de la binomial)

$$P[X > 6 | p = 0.2] = \sum_{x=6}^{20} b(x,20,0.2)$$

$$= \sum_{x=6}^{20} (0.2)^x (0.8)^{20-x}$$

$$= 0.1958 > 0.005$$

6. Conclusión: aceptamos H_0, es decir no hay razón para dudar que el 20% de los fumadores prefieren la marca A a un nivel del 1%.

En el capítulo 5 hemos visto que cuando n es grande y p está cercano a cero ó 1, aproximamos la distribución binomial por la distribución de Poisson con parámetro $\lambda = np_o$. También hemos visto (en el cap. 6) que cuando n es grande y p_o no está cerca de cero ni de 1, se usa la distribución normal como aproximación de la binomial. Por lo tanto usando la aproximación normal, el criterio de decisión se basa en la variable aleatoria

$Z = \dfrac{\hat{P} - p_o}{\sqrt{p_o q_o / n}} = \dfrac{X - np_o}{\sqrt{np_o q_o}}$, siempre que H_o es verdadera, cuyos valores están dados por,

$$z = \frac{\hat{p} - p_o}{\sqrt{p_o q_o / n}} = \frac{x - np_o}{\sqrt{np_o q_o}}$$

Luego, para una prueba bilateral al nivel de significación α , la región crítica es $Z < -z_{\alpha/2}$ ó $Z > z_{\alpha/2} = z_{1-\alpha/2}$. Para la alternativa unilateral $p < p_o$ la región crítica es $Z < -z_\alpha$. Y para la alternativa $p > p_o$, la región crítica es $Z > z_\alpha = z_{1-\alpha}$

Para probar una hipótesis acerca de una proporción usando la aproximación normal, se sigue los siguientes pasos:

1. $H_o : p = p_o$

   H_1 : Toma una de las siguientes alternativas,

   $$p < p_o \;,\; p > p_o \quad ó \quad p \neq p_o$$

2. Escoger un nivel de significación α .
3. Se usa la aproximación normal, cuando n es grande.
4. Región crítica:

   $Z < -z_\alpha$, para la alternativa $p < p_o$

   $Z > z_\alpha = z_{1-\alpha}$, para la alternativa $p > p_o$

   $Z < -z_{\alpha/2}$ ó $Z > z_{\alpha/2}$ para la alternativa $p \neq p_o$

5. Cálculos: hallar x de la muestra de tamaño n , y calcular

   $$z = \frac{x - np_o}{\sqrt{np_o q_o}}$$

6. Conclusión: Rechazar H_o , si z pertenece a la región crítica, en otros casos acepte H_o .

EJEMPLO 24 La oficina de relaciones familiares informa que el 50% de los matrimonios que viven en la ciudad A, llegan a la corte de divorcios dentro de su primer año de casados. ¿Qué conclusión puede sacarse acerca de la valides de este informe si de una muestra aleatoria de 400 matrimonios, sólo 193 fueron a una corte de divorcios dentro de su primer año de casados? Utilice un nivel de significación $\alpha = 0.01$.

SOLUCION

1. $H_0 : p = 0.50$ y $H_1 : p < 0.50$
2. $\alpha = 0.01$
3. Se usa la distribución normal como aproximación de la binomial, $n = 400$.
4. Región Crítica: $Z < - z_\alpha = - 2.33$, ya que $\alpha = 0.01$
5. $x = 193$, $p_0 = 0.5$, $q_0 = 0.5$, $np_0 = 400(0.5)$, luego

$$z = \frac{193 - 400(0.5)}{\sqrt{400(0.5)(0.5)}} \simeq - 0.7$$

6. Conclusión: desde que $z = - 0.7 > - 2.33$, o sea $z \in$ R.A, se acepta H_0, es decir no hay razón para dudar el informe a un nivel de significación del 1%.

EJEMPLO 25 En una conferencia de prensa, una alta autoridad anuncia que el 90% de los habitantes adultos del país están a favor de cierto proyecto económico del gobierno. Una muestra aleatoria de 625 adultos indica que 550 están a favor del proyecto. Si usted desea rechazar una hipótesis verdadera no más de una vez en 100, ¿Concluiría que la popularidad del proyecto ha sido exagerado por la autoridad?

SOLUCION

1. $H_0 : p = 0.90$; $H_1 : p < 0.90$
2. $\alpha = 0.01$
3. Se usa la distribución normal como aproximación a la binomial, pues n es grande.
4. Región crítica $Z < - z_\alpha = - 2.33$, pues $\alpha = 0.01$
5. Cálculos: $x = 550$, $p_0 = 0.90$, $q_0 = 0.10$, $n = 625$, $np_0 = 625(0.9)$,

luego $$z = \frac{550 - 625(0.9)}{\sqrt{625(0.9)(0.1)}} = - 1.67$$

6. Conclusión: desde que $-1.67 > -2.33$, aceptamos H_o. Es decir no hay razón para decir que la popularidad del proyecto ha sido exagerado a un nivel de significación del 1%.

9.2.8 PRUEBA PARA LA DIFERENCIA ENTRE DOS PROPORCIONES

En este caso la hipótesis nula toma la forma

$$H_o : p_1 = p_2 = p$$

la hipótesis alternativa toma una de las siguientes formas; $p_1 < p_2$ ó $p_1 > p_2$ ó $p_1 \neq p_2$. Los parámetros p_1 y p_2 son las proporciones de éxitos de las dos poblaciones bajo estudio. La estadística de prueba en la cual se base los criterios de decisión, es la variable aleatoria $\hat{P}_1 - \hat{P}_2$ que tiene una distribución aproximadamente normal cuando las muestras son grandes y la variable aleatoria

$$Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1 q_2}{n_1} + \frac{p_1 q_2}{n_2}}}$$

es aproximadamente una normal estándar.

Seleccionando muestras aleatorias independientes de tamaño n_1 y n_2 de cada población binomial respectivamente se calcula la proporción de éxitos $\hat{P}_1$ y $\hat{P}_2$, de cada muestra, $\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{p}_2 = \frac{x_2}{n_2}$, donde x_1 es el número de éxitos - en la muestra de tamaño n_1 y x_2 en la de tamaño n_2. Entonces

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_2}{n_1} + \frac{p_1 q_2}{n_2}}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

es un valor de la normal estándar Z, cuando H_o es verdadera y n_1, n_2 grandes. Para calcular z, se debe estimar el valor de p que aparece dentro del radical. El estimador de p es

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Luego el valor de la estadística Z, es

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}(\frac{1}{n_1} + \frac{1}{n_2})}}$$

donde $\hat{q} = 1 - \hat{p}$. Luego la región crítica para cada hipótesis alternativa se hace como en las pruebas, anteriores usando los puntos críticos de la curva normal estándar.

Para probar hipótesis de dos proporciones, cuando las muestras son grandes, se siguen los pasos siguientes;

1. H_0 : $p_1 = p_2$

   H_1 : puede ser una de lasalternativas,

   $$p_1 < p_2 , \quad p_1 > p_2 \quad ó \quad p_1 \neq p_2$$

2. Escoger un nivel de significación α.
3. La estadística de prueba es la variable aleatoria $\hat{p} - \hat{p}_2$ que tiene una distribución aproximadamente normal cuando n_1 y n_2 son grandes. Es decir, la variable aleatoria

   $$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

   supuesto que H_0 es verdadera.
4. Región Crítica,

   $Z < -z_\alpha$, para la hipótesis alternativa $p_1 < p_2$

   $Z > z_\alpha = z_{1-\alpha}$, para la hipótesis alternativa $p_1 > p_2$

   $Z < -z_{\alpha/2}$ ó $Z > z_{\alpha/2} = z_{1-\alpha/2}$, para la hipótesis alternativa $p_1 \neq p_2$.

5. Cálculos: calcular $\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{p}_2 = \frac{x_2}{n_2}$, $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$

   y luego hallar $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\,[\frac{1}{n_1} + \frac{1}{n_2}]}}$

6. Conclusión: Rechazar H_0 , si z pertenece a la región crítica; en otro caso aceptar H_0 .

EJEMPLO 26 Un investigador selecciona muestras aleatorias de 120 psicólogos y 80 psiquiatras para investigar sus opiniones acerca de si la esquizofrenia es causado por anormalidad bioquímica o una inadaptación originado en la niñez. La tabla que sigue da los resultados de esta investigación.

	Psicólogos	Psiquiatras
Anormalidad bioqufmica	60	50
Inadaptación de la niñez	60	30
Total	120	80

Si usted está dispuesto a rechazar una hipótesis verdadera no más de una vez en 100, ¿rechazaría la hipótesis que las opiniones de psicólogos y psiquiatras acerca de las causas de la esquizofrenia son las mismas?

SOLUCION

1. $H_0 : p_1 = p_2$, p_1 = proporción de anormalidad bioquímica para Psicólogos

   $H_1 : p_1 \neq p_2$, p_2 = proporción de anormalidad bioquímica para Psiquiatras.

2. $\alpha = 0.01$

3. Debido a que $n_1 = 120$ y $n_2 = 80$ son grandes, usamos la estadística $\hat{P}_1 - \hat{P}_2$ con distribución normal.

4. Región crítica, $Z < -z_{\alpha/2} = -2.58$ ó

   $Z > z_{\alpha/2} = 2.58$ ya que, $\alpha = 0.01$, $\alpha/2 = 0.005$,

   $z_{1-\alpha/2} = z_{\alpha/2} = 2.58$

5. Cálculos: $\hat{p}_1 = \dfrac{x_1}{n_1} = \dfrac{60}{120} = 0.5$

   $$\hat{p}_2 = \frac{x_2}{n_2} = \frac{50}{80} = 0.625$$

   $$\hat{p} = \frac{60 + 50}{120 + 80} = 0.55 , \quad \hat{q} = 0.45$$

   luego $$z = \frac{0.5 - 0.625}{\sqrt{(0.55)(0.45)[\frac{1}{120} + \frac{1}{80}]}} = -1.74$$

6. Conclusión: desde que - 1.74 > - 2.58, aceptamos H_o ; es decir la diferencia no es significativa por lo tanto no hay razón para rechazar que las opiniones de psicólogos y psiquiatras a acerca de las causas de la esquizofrenia son las mismas a un nivel de significación del 1%.

EJEMPLO 27 Una firma fabricante de cigarrillos distribuye dos marcas de cigarrillos. En una encuesta se encuentra que 56 de 200 fumadores prefieren la marca A y que 29 de 150 fumadores encuestados prefieren la marca B. ¿Se puede concluir al nivel de significación 0.06 que la marca A se vende más rápidamente que la marca B?

SOLUCION

1. $H_o : p_A = p_B$ y $H_1 : p_A > p_B$
2. $\alpha = 0.06$
3. Se usa la distribución Z, ya que $n_1 = 200$ y $n_2 = 150$ son grandes
4. Región crítica $Z > z_\alpha = 1.55$
5. Cálculos: $\hat{p}_A = \dfrac{56}{200} = 0.28$

$$\hat{p}_B = \frac{29}{150} \simeq 0.193$$

$$\hat{p} = \frac{56 + 29}{200 + 150} = \frac{85}{350} \simeq 0.242, \quad \hat{q} = 0.757$$

luego,

$$z = \frac{0.28 - 0.193}{\sqrt{(0.243)(0.757)[\frac{1}{200} + \frac{1}{150}]}} \simeq 1.875$$

6. Conclusión: ya que 1.875 > 1.55, se rechaza H_o , es decir la diferencia es significativa.

PROBLEMAS 9.2

1. En un experimento con el que se pretende determinar el efecto de un nuevo medicamento, este se aplica a 400 pacientes con cierta enfermedad. Si más de 300 pero menos de 340 pacientes se curan. Se concluíra que el medicamento es eficaz en un 80%? ¿Encuentre la probabilidad de cometer un error de tipo I. ¿Cuál es la probabilidad de cometer un error tipo II si el nuevo medicamento es eficaz en un 70%?
2. Se ha encontrado un aditivo para cierto tipo de cemento, con la que se obtiene una resistencia a la comprensión de 5000 kilogramos por centímetro

cuadrado y una desviación típica de 120. Para probar la hipótesis que $\mu=5000$ contra la alternativa que $\mu < 5000$, se prueba una muestra aleatoria de 50 piesas de cemento. La región crítica esta definida por $\overline{X} < 4970$. Determine la probabilidad de cometer un error de tipo I. Evalúe β para las alternativas $\mu = 4970$ y $\mu = 4960$.

3. Un fabricante de lavadoras automáticas produce un modelo en tres colores diferentes A, B y C. De las primeras 1000 lavadoras vendidas se observa que 400 fueron de color A. ¿concluíra usted que más de 1/3 de todos los clientes tienen preferencia por el color A? use $\alpha = 0.01$.

4. El fabricante de cierta marca de cigarrillos sostiene que sus cigarrillos contienen en promedio 18 miligramos de nicotina por cigarrillo. Un organismo de control examina una muestra de 100 cigarrillos. Utilizando un nivel de significación 0.01, ¿puede el organismo concluir que el fabricante subestima el contenido medio de nicotina de sus sigarrillos si el contenido medio de la muestra es de 19.2 miligramos con una desviación estándar de 2 miligramos?

5. El organismo de control de cierto Concejo Municipal analiza una muestra de 36 paquetes de carne molida que produce la fábrica de embutidos "LA UNICA". El rótulo en cada paquete dice "contiene no más de 25% de grasa". ¿Puede el organismo de control concluir que la carne que produce dicha fábrica tiene más de 25% de grasa, si la muestra da un contenido medio de grasa de 0.265 y una desviación estándar de 0.030? use $\alpha = 0.05$.

6. Un fabricante de pilas para linterna afirma que la vida media de su producto excederá a 30 horas. Una compañía desea comprar un lote grande de pilas si la afirmación es cierta. Se toma al azar una muestra de 36 pilas, y se encuentra que la media de la muestra es 34 horas. Si la población de pilas tiene una desviación estándar de 5 horas, si H_0 es $\mu < 30$, ¿Para qué valores de α se adquirirán las pilas?

7. El director de cierto colegio muy famoso cree que, en parte debido al estatus económico de los padres, el porcentaje de los que han terminado secundaria que asisten a este colegio es mayor que el promedio de la ciudad. En el período de los cinco años precedentes, el 20% de todos los que terminaron secundaria de la ciudad entraron a la Universidad, mientras que en el mismo período, 350 de los 1,500 exalumnos de su colegio entraron a la Universidad. ¿Se justifica que el director diga que el porcentaje de sus exalumnos que entraron a la Universidad es significativamente mayor que 20%, con el nivel de significación del 1%?.

8. En una oficina gubernamental se investiga a un empacador de pescado congelado. Los empaques que utiliza indican que contienen 12 onzas de pescado, en tanto que se han recibido quejas de que ello no es cierto. La oficina adquiere 100 paquetes de pescado procesado por esta compañía, y representando con x_i el peso observado (en onzas) del i-ésimo paquete, i = 1,2,...., 100, encuentra que

$$\Sigma\, x_i = 1150, \quad \Sigma\, x_i^2 = 13{,}249.75$$

Suponiendo que los pesos verdaderos de los empaques que se expanden están distribuídos normalmente con media y varianza desconocida. Con base en esta muestra, con un nivel de significación $\alpha = 0.01$, ¿La oficina multará a la compañía?

9. Las cajas de un cereal producidas en una fábrica deben tener un contenido de 16 onzas. Un inspector tomó una muestra que arrojó los siguientes pesos en onzas. 15.7, 15.7, 16.3, 15.8, 16.1, 15.9, 16.2, 15.9, 15.8, 15.6, indicar si es razonable que el inspector, usando un nivel de significación del 5%, ordene se multe al fabricante.

10. Usando una dócima bilateral, con α = 0.05, determinar si es razonable admitir que la muestra.

55; 42; 52; 61; 76; 50; 56; 38; 71.

proviene de una población normal con media igual a 50. Obtener además, un intervalo de confianza del 95% para la media poblacional.

11. Se extrae una muestra de tamaño 10 de una poblacion normal con media y varianza desconocida. Si la media y la varianza muestrales son iguales a 10 y 2 respectivamente.

(a) Docimar la hipótesis que la media poblacional es igual a 9 contra la alternativa que no lo es tomando un nivel de significación de 5%.

(b) Encontrar un intervalo de confianza de 90% para la media poblacional.

12. Los pesos netos (en gramos) de las latas de conserva fueron los siguientes:

121, 119, 124, 123, 119, 121 124.

Hay razón para creer que el peso neto poblacional medio es mayor que 123.5 gr. Use α = 0.01

13. Una máquina produce ejes que, según las especificaciones, deben tener 100 mm. de diámetro. Para mantener la calidad requerida, todos los días

se examina una muestra de 10 ejes para determinar si es necesario detener la producción y reajustar la máquina. Un día determinado la muestra dá los siguientes resultados:

101, 101, 102, 100, 99, 99, 102, 102, 100, 120 (mm)

Tomando $\alpha = 0.05$ indique, mediante un análisis estadístico, si es necesario reajustar la máquina.

14. El fabricante de un cierto modelo de automóvil afirma que el kilometraje medio de este modelo es de 12 kilometros por litro de gasolina corriente. Un organismo de defensa del consumidor piensa que ese kilometraje promedio ha sido exagerado por el fabricante. Nueve autómoviles de este modelo son conducidos del mismo modo con un litro de gasolina corriente. Los kilómetros recorridos por los diversos automóviles son:

12; 11; 10; 10.5; 11.5; 11; 12.5; 10; 10.5

Si el organismo desea rechazar una afirmación verdadera no más de una vez en 100, ¿Rechazará la afirmación del fabricante?

15. Un grupo del primer grado formado por 20 alumnos, recibe enseñanza de lectura en la forma tradicional durante un año; al final del período deben contestar un examen. La suma de sus calificaciones y la suma de los cuadrados de las mismas son

$$\Sigma\, x_i = 1950 \quad ; \quad \Sigma\, x_i^2 = 192.861$$

Un segundo grupo del primer grado también de 20 estudiantes, se le enseña a leer usando un nuevo método, diferente. Al final del año presentan el mismo examen, y se obtienen la suma de sus calificaciones y la de los cuadrados de las mismas, que son

$$\Sigma\, y_i = 2020 \quad y \quad \Sigma\, y_i^2 = 205{,}920$$

Suponiendo que los dos son muestras independientes de distribuciones normales con la misma varianza. Usando un nivel de significación de 0.05, ¿concluye que ambos métodos dieron resultados diferentes?

16. Se usó gasolina de marca A en 9 automóviles semejantes bajo idénticas condiciones. La muestra correspondiente de 9 valores (kilómetros por litro) tienen una media 8.565 y una desviación estándar 0.212. Bajo las mismas condiciones, la gasolina de alta potencia de marca B da una muestra de 10 valores con media 9.245 y desviación estándar 0.254. Probar la hipótesis que A y B son de igual calidad con respecto al kilometraje, contra de que B es mejor.

17. ¿Es significativa la diferencia de aumento de peso (kilogramos por mes) de dos grupos de cerdos, si los dos grupos se alimentaron bajo dos dietas diferentes? use α = 0.05

GRUPO I	33	66	26	43	46	55	54
GRUPO II	53	53	37	73	58	61	38

18. Un auditor de una cadena de tiendas desea comparar la eficiencia de dos técnicas de auditoría diferentes. Para este fín selecciona una muestra de 9 cuentas de las tiendas y les aplica la técnica A y selecciona otras 9 cuentas de las tiendas y les aplica la técnica B. En la tabla se presenta el número de errores encontrados en las cuentas de las tiendas.

TECNICA A :	125	116	133	115	123	120	132	128	121
TECNICA B :	89	101	97	95	94	102	98	106	98

Determine si existe evidencia de una diferencia en el número medio de errores detectados por las dos técnicas de auditoría. α = 0.10

19. Los siguientes son porcentajes de grasa encontrados en muestras de dos tipos de carne. ¿Tienen las carnes diferentes contenidos de grasa?

CARNE A :	30	26	30	19	25	37	27	38	26	31
CARNE B :	40	34	28	29	26	36	28	37	35	42

20. Un analista de sistemas está probando la posibilidad de usar un nuevo sistema de computadora. El analista cambiará el procesamiento al nuevo sistema sólo si hay pruebas que el nuevo sistema usa menos tiempo para el procesamiento que el sistema antiguo. A fin de tomar una decisión, se seleccionó una muestra de siete trabajos y se registro el tiempo de procesamiento, en segundos, en los dos sistemas, con los siguientes resultados.

TRABAJOS	ANTIGUO	NUEVO
1	8	6
2	4	3
3	10	7
4	9	8
5	8	5
6	7	8
7	12	9

(a) Al nivel de significación de 0.01, ¿usa el nuevo sistema menos tiempo para el procesamiento?

(b) ¿Qué suposición es necesaria para efectuar esta prueba?

21. El gerente de procesamiento de datos de una empresa desea estudiar el uso de la computadora de dos departamentos de la compañía: el departamento de contabilidad y el departamento de investigaciones. Se seleccionó una muestra aleatoria de cinco trabajos del departamento de contabilidad en el mes pasado y seis trabajos del departamento de investigación en el mes pasado y se registró el tiempo de procesamiento (en segundos) para cada trabajo, con los siguientes resultados

Contabilidad :	9	3	8	7	12	
INVESTIGACION :	4	13	10	9	9	6

Saque conclusiones en relación con cada uno de los siguientes:

(a) ¿Es la desviación estándar del tiempo de procesamiento del departamento de contabilidad mayor de un segundo?

(b) ¿Hay una diferencia en el tiempo del procesamiento entre el departamento de contabilidad y el departamento de investigación?

(c) ¿Qué suposiciones se deben hacer para lograr (b)?

(d) ¿Hay una diferencia en la varianza entre el departamento de contabilidad y el departamento de investigación?

(e) ¿Qué suposiciones se necesitan para lograr (d) ?

use α = 0.05 en todo el problema.

22. Una muestra de 80 alambres de acero producidos por la fábrica A da una resistencia media a la rotura de 1230 libras, con una desviación estándar de 120 libras. Una muestra de 100 alambres de acero producidos por la fábrica B da una resistencia media a la rotura de 1190 libras, con una desviación estándar de 90 libras. ¿Hay una diferencia real en la resistencia media de las dos marcas de alambres de acero.? α = 0.05.

23. Se estan comparando los rendimientos de dos tipos de trigo, se plantan 25 acres de cada clase y se exponen a condiciones de cultivo bastante uniformes. El rendimiento de cada acre es considerado como una observación. Los resultados son: variedad A, $\bar{x}_1$ = 35.8 bushels por acre, s_1^2 = 6, variedad B: $\bar{x}_2$ = 32.9 bushels por acre, s_2^2 = 10. ¿Es el rendimiento de la media de la variedad A significativamente diferente de la variedad B? Use α = 0.05.

24. Se supone que los diámetros de cierta marca de tubo de acero están normalmente distribuídos. Una muestra al azar de 10 observaciones dió una varianza de 0.12 pulg^2 . ¿Será rechazada la hipótesis $\sigma^2 = 0.06$ a favor que $\sigma^2 > 0.06$ a un nivel de significación $\alpha = 0.05$.?

25. Se sabe que dos marcas de neumáticos tienen una duración media de 45,000 millas. Pero puede haber alguna diferencia en la variabilidad del millaje alcanzado por las dos marcas. Se efectúa un experimento en el que se usan 16 neumáticos de cada marca. Se hacen rodar los neumáticos en condiciones similares hasta que se desgastan. Se encuentra que las desviaciones estándares de las dos muestras son $s_1^2 = 4500$ millas y $s_2^2 = 2200$ millas. ¿Rechazaría ud. la hipótesis $\sigma_1^2 = \sigma_2^2$ a favor de $\sigma_1^2 > \sigma_2^2$ en un nivel de significación $\alpha = 0.05 (\alpha = 0.10)$?

26. La tabla siguiente muestra los números de roturas de hilo (por 45.40 kg) correspondientes a dos diferentes enlaces por cada 2.54 centímetros. ¿Es significativa la diferencia entre las varianzas?. Use $\alpha = 0.05$

Enlaces/9.54 cm.	Roturas/45.40 kg							
1.69	6.0	9.7	7.4	11.5	17.9	11.9	10.2	7.8
1.78	6.4	8.3	7.9	8.8	10.1	11.5	8.7	9.7

27. Una industria lechera está estudiando la posibilidad de cambiar sus botellas para la leche por envases de plástico. Pero el cambio no se hará a no ser que por lo menos 70% de sus clientes lo prefieran. Cuando se ha hecho una encuesta a 200 de sus clientes, 120 de ellos están a favor del cambio. ¿Hará el cambio de envases a un nivel de significación 0.05?

28. Un laboratorio farmacéutico ha elaborado un medicamento para tratar la presion sanguinea alta. El laboratorio afirma que el medicamento efectivamente baja la presión en el 80% de los casos. Si 175 de 225 pacientes tratados con el medicamento experimentaron una disminución substancial de la presión sanguínea, ¿concluirá usted que el laboratorio ha exagerado la efectividad del medicamento? use $\alpha = 0.01$.

29. De un grupo de 8,000 estudiantes, se toma una muestra aleatoria de tamaño 100, y se les pregunta a cada uno su opinión sobre cierta proposición. Encontrándose que 75 de ellos están a favor y 25 en contra. Pruébese la hipótesis que la proporción que favorece la proposición es 2/3 en contra de la posibilidad que no lo es, con nivel de significación de 5%.

30. Una muestra aleatoria de 100 hombres fue tomada de cierta ciudad y se encontró que 60 se mostraron a favor de una ley sobre el divorcio. Una muestra al azar de 100 mujeres escogidos de la misma ciudad en la misma fecha reveló que 40 estaban a favor de tal nueva ley. ¿Es igual la proporción de hombres que de mujeres que favorecen una nueva ley sobre el divorcio en tal ciudad? Use $\alpha = 0.05$.

31. De 292 chicas que terminaron secundaria en cierto colegio del Callao, se encontro que 13 no quieren seguir estudios superiores; de 290 chicas que terminaron secundaria en un colegio de Comas en el mismo año escolar, se encontro que 12 no quieren seguir estudios superiores. Probar la hipótesis que la proporción de chicas del Callao que no quieren seguir estudios superiores es la misma que en Comas, en contra que no son iguales. Use $\alpha = 0.05$.

BIBLIOGRAFIA

1. FADIL. H. ZUWAYLIF. Estadística general aplicada; Fondo Educativo Interamericano, S.A.
2. YA-LUN CHOU. Análisis estadístico, (segunda Edición) Editorial Interamericana.
3. ROBERT. V. HOGG-ELLIOT. A. TANIS. Probability and statistical inference; Collier macmillan international editions.
4. RONALD. E. WALPOLE. Introducction to statistics (second edition); Collier macmillan international editions.
5. RICARD. L. MILLS. Estadística para economía y administración; Editorial Mcgraw-hill Latinoamericano S.A.
6. HAROLD. J. LARSON. Introducción a la teoría de probabilidades e inferencia estadística; Editorial Limusa.
7. WILLIAM MERRILL-KARL FOX. Introducción a la estadística económica; Amorrortu editores S.C.A. Buenos Aires.
8. PAUL. G. HOEL. Estadística elemental; Compañía editorial continental, S.A.
9. ERWIN KREYSZIG. Introducción a la estadística matemática. Principios y métodos; Editorial Limusa.
10. PAUL. L. MEYER. Probabilidad y aplicaciones estadísticas; Fondo Educativo interamericano S.A.
11. ALBERT. H. BOWKER - GERALD. J. LIEBERMAN. Estadística para ingenieros; printice/Hal Internacional.
12. WILLIAM W. HINES - DOUGLAS C. MONTGOMERY. Probability and stadistics in Engineering and Management Science; Jhon Willey & Sons.
13. MARTIN M. EIGEN - CAROLE A. EISEN. Probability an its Applications; Quantum Publishors, inc.
14. J. V. USPENSKY. Introductión to Mathematical probability; Mcgraw - hill Book Company.
15. M. L. BERENSON - D. M. LEVINE. Estadística para Administración y Economía; Interamericana.
16. W. MENDENHALL - J. E. REINMUTH. Estadística para Administración y Economía; Wadsworth Internacional Iberoamericana.
17. A. C. BAJPAI - I. M. CALUS - J. A. FAIRLEY. Métodos estadísticos para estudiantes de Ingeniería y Ciencias; Limusa.

TABLAS

TABLA I

DISTRIBUCION BINOMIAL

$P(X \geq r|n,p)$

$n = 1$

p	01	02	03	04	05	06	07	08	09	10
r 1	0100	0200	0300	0400	0500	0600	0700	0800	0900	1000
p	11	12	13	14	15	16	17	18	19	20
r 1	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000
p	21	22	23	24	25	26	27	28	29	30
r 1	2100	2200	2300	2400	2500	2600	2700	2800	2900	3000
p	31	32	33	34	35	36	37	38	39	40
r 1	3100	3200	3300	3400	3500	3600	3700	3800	3900	4000
p	41	42	43	44	45	46	47	48	49	50
r 1	4100	4200	4300	4400	4500	4600	4700	4800	4900	5000

$n = 2$

p	01	02	03	04	05	06	07	08	09	10
r 1	0199	0396	0591	0784	0975	1164	1351	1536	1719	1900
2	0001	0004	0009	0016	0025	0036	0049	0064	0081	0100
p	11	12	13	14	15	16	17	18	19	20
r 1	2079	2256	2431	2604	2775	2944	3111	3276	3439	3600
2	0121	0144	0169	0196	0225	0256	0289	0324	0361	0400
p	21	22	23	24	25	26	27	28	29	30
r 1	3759	3916	4071	4224	4375	4524	4671	4816	4959	5100
2	0441	0484	0529	0576	0625	0676	0729	0784	0841	0900
p	31	32	33	34	35	36	37	38	39	40
r 1	5239	5376	5511	5644	5775	5904	6031	6156	6279	6400
2	0961	1024	1089	1156	1225	1296	1369	1444	1521	1600
p	41	42	43	44	45	46	47	48	49	50
r 1	6519	6636	6751	6864	6975	7084	7191	7296	7399	7500
2	1681	1764	1849	1936	2025	2116	2209	2304	2401	2500

$n = 3$

p	01	02	03	04	05	06	07	08	09	10
r 1	0297	0588	0873	1153	1426	1694	1956	2213	2464	2710
2	0003	0012	0026	0047	0073	0104	0140	0182	0228	0280
3				0001	0001	0002	0003	0005	0007	0010
p	11	12	13	14	15	16	17	18	19	20
r 1	2950	3185	3415	3639	3859	4073	4282	4486	4686	4680
2	0336	0397	0463	0533	0608	0686	0769	0855	0946	1040
3	0013	0017	0022	0027	0034	0041	0049	0058	0069	0080
p	21	22	23	24	25	26	27	28	29	30
r 1	5070	5254	5435	5610	5781	5948	6110	6268	6421	6570
2	1138	1239	1344	1452	1563	1676	1793	1913	2035	2160
3	0093	0106	0122	0138	0156	0176	0197	0220	0244	0270
p	31	32	33	34	35	36	37	38	39	40
r 1	6715	6856	6992	7125	7254	7379	7500	7617	7730	7840
2	2287	2417	2548	2682	2818	2955	3094	3235	3377	3520
3	0298	0328	0359	0393	0429	0467	0507	0549	0593	0640

$n = 3$ $P(X \geq r|n,p)$

r \ p	41	42	43	44	45	46	47	48	49	50
1	7946	8049	8148	8244	8336	8425	8511	8594	8673	8750
2	3665	3810	3957	4104	4253	4401	4551	4700	4850	5000
3	0689	0741	0795	0852	0911	0973	1038	1106	1176	1250

$n = 4$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0394	0776	1147	1507	1855	2193	2519	2836	3143	3439
2	0006	0023	0052	0091	0140	0199	0267	0344	0430	0523
3			0001	0002	0005	0008	0013	0019	0027	0037
4									0001	0001

r \ p	11	12	13	14	15	16	17	18	19	20
1	3726	4003	4271	4530	4780	5021	5254	5479	5695	5904
2	0624	0732	0847	0968	1095	1228	1366	1509	1656	1808
3	0049	0063	0079	0098	0120	0144	0171	0202	0235	0272
4	0001	0002	0003	0004	0005	0007	0008	0010	0013	0016

r \ p	21	22	23	24	25	26	27	28	29	30
1	6105	6298	6485	6664	6836	7001	7160	7313	7459	7599
2	1963	2122	2285	2450	2617	2787	2959	3132	3307	3483
3	0312	0356	0403	0453	0508	0566	0628	0694	0763	0837
4	0019	0023	0028	0033	0039	0046	0053	0061	0071	0081

r \ p	31	32	33	34	35	36	37	38	39	40
1	7733	7862	7985	8103	8215	8322	8425	8522	8615	8704
2	3660	3837	4015	4193	4370	4547	4724	4900	5075	5248
3	0915	0996	1082	1171	1265	1362	1464	1569	1679	1792
4	0092	0105	0119	0134	0150	0168	0187	0209	0231	0256

r \ p	41	42	43	44	45	46	47	48	49	50
1	8788	8868	8944	9017	9085	9150	9211	9269	9323	9375
2	5420	5590	5759	5926	6090	6252	6412	6569	6724	6875
3	1909	2030	2155	2283	2415	2550	2689	2831	2977	3125
4	0283	0311	0342	0375	0410	0448	0488	0531	0576	0625

$n = 5$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0490	0961	1413	1846	2262	2661	3043	3409	3760	4095
2	0010	0038	0085	0148	0226	0319	0425	0544	0674	0815
3		0001	0003	0006	0012	0020	0031	0045	0063	0086
4						0001	0001	0002	0003	0005

r \ p	11	12	13	14	15	16	17	18	19	20
1	4416	4723	5016	5296	5563	5818	6061	6293	6513	6723
2	0965	1125	1292	1467	1648	1835	2027	2224	2424	2627
3	0112	0143	0179	0220	0266	0318	0375	0437	0505	0579
4	0007	0009	0013	0017	0022	0029	0036	0045	0055	0067
5				0001	0001	0001	0001	0002	0002	0003

r \ p	21	22	23	24	25	26	27	28	29	30
1	6923	7113	7293	7464	7627	7781	7927	8065	8196	8319
2	2833	3041	3251	3461	3672	3883	4093	4303	4511	4718
3	0659	0744	0836	0933	1035	1143	1257	1376	1501	1631
4	0081	0097	0114	0134	0156	0181	0208	0238	0272	0308
5	0004	0005	0006	0008	0010	0012	0014	0017	0021	0024

$P(X \geq r \mid n,p)$

$n = 5$

r \ p	31	32	33	34	35	36	37	38	39	40
1	8436	8546	8650	8748	8840	8926	9008	9084	9155	9222
2	4923	5125	5325	5522	5716	5906	6093	6276	6455	6630
3	1766	1905	2050	2199	2352	2509	2670	2835	3003	3174
4	0347	0390	0436	0486	0540	0598	0660	0726	0796	0870
5	0029	0034	0039	0045	0053	0060	0069	0079	0090	0102

r \ p	41	42	43	44	45	46	47	48	49	50
1	9285	9344	9398	9449	9497	9541	9582	9620	9655	9688
2	6801	6967	7129	7286	7438	7585	7728	7865	7998	8125
3	3349	3525	3705	3886	4069	4253	4439	4625	4813	5000
4	0949	1033	1121	1214	1312	1415	1522	1635	1753	1875
5	0116	0131	0147	0165	0185	0206	0229	0255	0282	0313

$n = 6$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0585	1142	1670	2172	2649	3101	3530	3936	4321	4686
2	0015	0057	0125	0216	0328	0459	0608	0773	0952	1143
3		0002	0005	0012	0022	0038	0058	0085	0118	0159
4					0001	0002	0003	0005	0008	0013
5										0001

r \ p	11	12	13	14	15	16	17	18	19	20
1	5030	5356	5664	5954	6229	6487	6731	6960	7176	7379
2	1345	1556	1776	2003	2235	2472	2713	2956	3201	3446
3	0206	0261	0324	0395	0473	0560	0655	0759	0870	0989
4	0018	0025	0034	0045	0059	0075	0094	0116	0141	0170
5	0001	0001	0002	0003	0004	0005	0007	0010	0013	0016
6										0001

r \ p	21	22	23	24	25	26	27	28	29	30
1	7569	7748	7916	8073	8220	8358	8487	8607	8719	8824
2	3692	3937	4180	4422	4661	4896	5128	5356	5580	5798
3	1115	1250	1391	1539	1694	1856	2023	2196	2374	2557
4	0202	0239	0280	0326	0376	0431	0492	0557	0628	0705
5	0020	0025	0031	0038	0046	0056	0067	0079	0093	0109
6	0001	0001	0001	0002	0002	0003	0004	0005	0006	0007

r \ p	31	32	33	34	35	36	37	38	39	40
1	8921	9011	9095	9173	9246	9313	9375	9432	9485	9533
2	6012	6220	6422	6619	6809	6994	7172	7343	7508	7667
3	2744	2936	3130	3328	3529	3732	3937	4143	4350	4557
4	0787	0875	0969	1069	1174	1286	1404	1527	1657	1792
5	0127	0148	0170	0195	0223	0254	0288	0325	0365	0410
6	0009	0011	0013	0015	0018	0022	0026	0030	0035	0041

r \ p	41	42	43	44	45	46	47	48	49	50
1	9578	9619	9657	9692	9723	9752	9778	9802	9824	9844
2	7819	7965	8105	8238	8364	8485	8599	8707	8810	8906
3	4764	4971	5177	5382	5585	5786	5985	6180	6373	6563
4	1933	2080	2232	2390	2553	2721	2893	3070	3252	3438
5	0458	0510	0566	0627	0692	0762	0837	0917	1003	1094
6	0048	0055	0063	0073	0083	0095	0108	0122	0138	0156

n = 7

$P(X \geq r|n,p)$

$n = 7$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0679	1319	1920	2486	3017	3515	3983	4422	4832	5217
2	0020	0079	0171	0294	0444	0618	0813	1026	1255	1497
3		0003	0009	0020	0038	0063	0097	0140	0193	0257
4				0001	0002	0004	0007	0012	0018	0027
5								0001	0001	0002

r \ p	11	12	13	14	15	16	17	18	19	20
1	5577	5913	6227	6521	6794	7049	7286	7507	7712	7903
2	1750	2012	2281	2556	2834	3115	3396	3677	3956	4233
3	0331	0416	0513	0620	0738	0866	1005	1154	1313	1480
4	0039	0054	0072	0094	0121	0153	0189	0231	0279	0333
5	0003	0004	0006	0009	0012	0017	0022	0029	0037	0047
6					0001	0001	0001	0002	0003	0004

r \ p	21	22	23	24	25	26	27	28	29	30
1	8080	8243	8395	8535	8665	8785	8895	8997	9090	9176
2	4506	4775	5040	5298	5551	5796	6035	6266	6490	6706
3	1657	1841	2033	2231	2436	2646	2861	3081	3304	3529
4	0394	0461	0536	0617	0706	0802	0905	1016	1134	1260
5	0058	0072	0088	0107	0129	0153	0181	0213	0248	0288
6	0005	0006	0008	0011	0013	0017	0021	0026	0031	0038
7					0001	0001	0001	0001	0002	0002

r \ p	31	32	33	34	35	36	37	38	39	40
1	9255	9328	9394	9454	9510	9560	9606	9648	9686	9720
2	6914	7113	7304	7487	7662	7828	7987	8137	8279	8414
3	3757	3987	4217	4447	4677	4906	5134	5359	5581	5801
4	1394	1534	1682	1837	1998	2167	2341	2521	2707	2898
5	0332	0380	0434	0492	0556	0625	0701	0782	0869	0963
6	0046	0055	0065	0077	0090	0105	0123	0142	0164	0188
7	0003	0003	0004	0005	0006	0008	0009	0011	0014	0016

r \ p	41	42	43	44	45	46	47	48	49	50
1	9751	9779	9805	9827	9848	9866	9883	9897	9910	9922
2	8541	8660	8772	8877	8976	9068	9153	9233	9307	9375
3	6017	6229	6436	6638	6836	7027	7213	7393	7567	7734
4	3094	3294	3498	3706	3917	4131	4346	4563	4781	5000
5	1063	1169	1282	1402	1529	1663	1803	1951	2105	2266
6	0216	0246	0279	0316	0357	0402	0451	0504	0562	0625
7	0019	0023	0027	0032	0037	0044	0051	0059	0068	0078

$n = 8$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0773	1492	2163	2786	3366	3904	4404	4868	5297	5695
2	0027	0103	0223	0381	0572	0792	1035	1298	1577	1869
3	0001	0004	0013	0031	0058	0096	0147	0211	0289	0381
4			0001	0002	0004	0007	0013	0022	0034	0050
5							0001	0001	0003	0004

$P(X \geq r | n,p)$ $n = 8$

r \ p	11	12	13	14	15	16	17	18	19	20
1	6063	6404	6718	7008	7275	7521	7748	7956	8147	8322
2	2171	2480	2794	3111	3428	3744	4057	4366	4670	4967
3	0487	0608	0743	0891	1052	1226	1412	1608	1815	2031
4	0071	0097	0129	0168	0214	0267	0328	0397	0476	0563
5	0007	0010	0015	0021	0029	0038	0050	0065	0083	0104
6		0001	0001	0002	0002	0003	0005	0007	0009	0012
7									0001	0001

r \ p	21	22	23	24	25	26	27	28	29	30
1	8483	8630	8764	8887	8999	9101	9194	9278	9354	9424
2	5257	5538	5811	6075	6329	6573	6807	7031	7244	7447
3	2255	2486	2724	2967	3215	3465	3718	3973	4228	4482
4	0659	0765	0880	1004	1138	1281	1433	1594	1763	1941
5	0129	0158	0191	0230	0273	0322	0377	0438	0505	0580
6	0016	0021	0027	0034	0042	0052	0064	0078	0094	0113
7	0001	0002	0002	0003	0004	0005	0006	0008	0010	0013
8									0001	0001

r \ p	31	32	33	34	35	36	37	38	39	40
1	9486	9543	9594	9640	9681	9719	9752	9782	9808	9832
2	7640	7822	7994	8156	8309	8452	8586	8711	8828	8936
3	4736	4987	5236	5481	5722	5958	6189	6415	6634	6846
4	2126	2319	2519	2724	2936	3153	3374	3599	3828	4059
5	0661	0750	0846	0949	1061	1180	1307	1443	1586	1737
6	0134	0159	0187	0218	0253	0293	0336	0385	0439	0498
7	0016	0020	0024	0030	0036	0043	0051	0061	0072	0085
8	0001	0001	0001	0002	0002	0003	0004	0004	0005	0007

r \ p	41	42	43	44	45	46	47	48	49	50
1	9853	9872	9889	9903	9916	9928	9938	9947	9954	9961
2	9037	9130	9216	9295	9368	9435	9496	9552	9602	9648
3	7052	7250	7440	7624	7799	7966	8125	8276	8419	8555
4	4292	4527	4762	4996	5230	5463	5694	5922	6146	6367
5	1895	2062	2235	2416	2604	2798	2999	3205	3416	3633
6	0563	0634	0711	0794	0885	0982	1086	1198	1318	1445
7	0100	0117	0136	0157	0181	0208	0239	0272	0310	0352
8	0008	0010	0012	0014	0017	0020	0024	0028	0033	0039

$n = 9$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0865	1663	2398	3075	3698	4270	4796	5278	5721	6126
2	0034	0131	0282	0478	0712	0978	1271	1583	1912	2252
3	0001	0006	0020	0045	0084	0138	0209	0298	0405	0530
4			0001	0003	0006	0013	0023	0037	0057	0083
5						0001	0002	0003	0005	0009
6										0001

$n = 9$ $P(X \geq r|n,p)$

r \ p	11	12	13	14	15	16	17	18	19	20
1	6496	6835	7145	7427	7684	7918	8131	8324	8499	8658
2	2599	2951	3304	3657	4005	4348	4685	5012	5330	5638
3	0672	0833	1009	1202	1409	1629	1861	2105	2357	2618
4	0117	0158	0209	0269	0339	0420	0512	0615	0730	0856
5	0014	0021	0030	0041	0056	0075	0098	0125	0158	0196
6	0001	0002	0003	0004	0006	0009	0013	0017	0023	0031
7						0001	0001	0002	0002	0003

r \ p	21	22	23	24	25	26	27	28	29	30
1	8801	8931	9048	9154	9249	9335	9411	9480	9542	9596
2	5934	6218	6491	6750	6997	7230	7452	7660	7856	8040
3	2885	3158	3434	3713	3993	4273	4552	4829	5102	5372
4	0994	1144	1304	1475	1657	1849	2050	2260	2478	2703
5	0240	0291	0350	0416	0489	0571	0662	0762	0870	0988
6	0040	0051	0065	0081	0100	0122	0149	0179	0213	0253
7	0004	0006	0008	0010	0013	0017	0022	0028	0035	0043
8			0001	0001	0001	0001	0002	0003	0003	0004

r \ p	31	32	33	34	35	36	37	38	39	40
1	9645	9689	9728	9762	9793	9820	9844	9865	9883	9899
2	8212	8372	8522	8661	8789	8908	9017	9118	9210	9295
3	5636	5894	6146	6390	6627	6856	7076	7287	7489	7682
4	2935	3173	3415	3662	3911	4163	4416	4669	4922	5174
5	1115	1252	1398	1553	1717	1890	2072	2262	2460	2666
6	0298	0348	0404	0467	0536	0612	0696	0787	0886	0994
7	0053	0064	0078	0094	0112	0133	0157	0184	0215	0250
8	0006	0007	0009	0011	0014	0017	0021	0026	0031	0038
9				0001	0001	0001	0001	0002	0002	0003

r \ p	41	42	43	44	45	46	47	48	49	50
1	9913	9926	9936	9946	9954	9961	9967	9972	9977	9980
2	9372	9442	9505	9563	9615	9662	9704	9741	9775	9805
3	7866	8039	8204	8359	8505	8642	8769	8889	8999	9102
4	5424	5670	5913	6152	6386	6614	6836	7052	7260	7461
5	2878	3097	3322	3551	3786	4024	4265	4509	4754	5000
6	1109	1233	1366	1508	1658	1817	1985	2161	2346	2539
7	0290	0334	0383	0437	0498	0564	0637	0717	0804	0898
8	0046	0055	0065	0077	0091	0107	0125	0145	0169	0195
9	0003	0004	0005	0006	0008	0009	0011	0014	0016	0020

$n = 10$

r \ p	01	02	03	04	05	06	07	08	09	10
1	0956	1829	2626	3352	4013	4614	5160	5656	6106	6513
2	0043	0162	0345	0582	0861	1176	1517	1879	2254	2639
3	0001	0009	0028	0062	0115	0188	0283	0401	0540	0702
4			0001	0004	0010	0020	0036	0058	0088	0128
5					0001	0002	0003	0006	0010	0016
6									0001	0001

$P(X \geq r \mid n,p)$ $n = 10$

r \ p	11	12	13	14	15	16	17	18	19	20
1	6882	7215	7516	7787	8031	8251	8448	8626	8784	8926
2	3028	3417	3804	4184	4557	4920	5270	5608	5932	6242
3	0884	1087	1308	1545	1798	2064	2341	2628	2922	3222
4	0178	0239	0313	0400	0500	0614	0741	0883	1039	1209
5	0025	0037	0053	0073	0099	0130	0168	0213	0266	0328
6	0003	0004	0006	0010	0014	0020	0027	0037	0049	0064
7			0001	0001	0001	0002	0003	0004	0006	0009
8									0001	0001

r \ p	21	22	23	24	25	26	27	28	29	30
1	9053	9166	9267	9357	9437	9508	9570	9626	9674	9718
2	6536	6815	7079	7327	7560	7778	7981	8170	8345	8507
3	3526	3831	4137	4442	4744	5042	5335	5622	5901	6172
4	1391	1587	1794	2012	2241	2479	2726	2979	3239	3504
5	0399	0479	0569	0670	0781	0904	1037	1181	1337	1503
6	0082	0104	0130	0161	0197	0239	0287	0342	0404	0473
7	0012	0016	0021	0027	0035	0045	0056	0070	0087	0106
8	0001	0002	0002	0003	0004	0006	0007	0010	0012	0016
9							0001	0001	0001	0001

r \ p	31	32	33	34	35	36	37	38	39	40
1	9755	9789	9818	9843	9865	9885	9902	9916	9929	9940
2	8656	8794	8920	9035	9140	9236	9323	9402	9473	9536
3	6434	6687	6930	7162	7384	7595	7794	7983	8160	8327
4	3772	4044	4316	4589	4862	5132	5400	5664	5923	6177
5	1679	1867	2064	2270	2485	2708	2939	3177	3420	3669
6	0551	0637	0732	0836	0949	1072	1205	1348	1500	1662
7	0129	0155	0185	0220	0260	0305	0356	0413	0477	0548
8	0020	0025	0032	0039	0048	0059	0071	0086	0103	0123
9	0002	0003	0003	0004	0005	0007	0009	0011	0014	0017
10								0001	0001	0001

r \ p	41	42	43	44	45	46	47	48	49	50
1	9949	9957	9964	9970	9975	9979	9983	9986	9988	9990
2	9594	9645	9691	9731	9767	9799	9827	9852	9874	9893
3	8483	8628	8764	8889	9004	9111	9209	9298	9379	9453
4	6425	6665	6898	7123	7340	7547	7745	7933	8112	8281
5	3922	4178	4436	4696	4956	5216	5474	5730	5982	6230
6	1834	2016	2207	2407	2616	2832	3057	3288	3526	3770
7	0626	0712	0806	0908	1020	1141	1271	1410	1560	1719
8	0146	0172	0202	0236	0274	0317	0366	0420	0480	0547
9	0021	0025	0031	0037	0045	0054	0065	0077	0091	0107
10	0001	0002	0002	0003	0003	0004	0005	0006	0008	0010

n = 11

$P(X \geq r|n,p)$

n = 11

p / r	01	02	03	04	05	06	07	08	09	10
1	1047	1993	2847	3618	4312	4937	5499	6004	6456	6862
2	0052	0195	0413	0692	1019	1382	1772	2181	2601	3026
3	0002	0012	0037	0083	0152	0248	0370	0519	0695	0896
4			0002	0007	0016	0030	0053	0085	0129	0185
5					0001	0003	0005	0010	0017	0028
6								0001	0002	0003

p / r	11	12	13	14	15	16	17	18	19	20
1	7225	7549	7839	8097	8327	8531	8712	8873	9015	9141
2	3452	3873	4286	4689	5078	5453	5811	6151	6474	6779
3	1120	1366	1632	1915	2212	2521	2839	3164	3494	3826
4	0256	0341	0442	0560	0694	0846	1013	1197	1397	1611
5	0042	0061	0087	0119	0159	0207	0266	0334	0413	0504
6	0005	0008	0012	0018	0027	0037	0051	0068	0090	0117
7		0001	0001	0002	0003	0005	0007	0010	0014	0020
8							0001	0001	0002	0002

p / r	21	22	23	24	25	26	27	28	29	30
1	9252	9350	9436	9511	9578	9636	9686	9730	9769	9802
2	7065	7333	7582	7814	8029	8227	8410	8577	8730	8870
3	4158	4488	4814	5134	5448	5753	6049	6335	6610	6873
4	1840	2081	2333	2596	2867	3146	3430	3719	4011	4304
5	0607	0723	0851	0992	1146	1313	1493	1685	1888	2103
6	0148	0186	0231	0283	0343	0412	0490	0577	0674	0782
7	0027	0035	0046	0059	0076	0095	0119	0146	0179	0216
8	0003	0005	0007	0009	0012	0016	0021	0027	0034	0043
9			0001	0001	0001	0002	0002	0003	0004	0006

p / r	31	32	33	34	35	36	37	38	39	40
1	9831	9856	9878	9896	9912	9926	9938	9948	9956	9964
2	8997	9112	9216	9310	9394	9470	9537	9597	9650	9698
3	7123	7361	7587	7799	7999	8186	8360	8522	8672	8811
4	4598	4890	5179	5464	5744	6019	6286	6545	6796	7037
5	2328	2563	2807	3059	3317	3581	3850	4122	4397	4672
6	0901	1031	1171	1324	1487	1661	1847	2043	2249	2465
7	0260	0309	0366	0430	0501	0581	0670	0768	0876	0994
8	0054	0067	0082	0101	0122	0148	0177	0210	0249	0293
9	0008	0010	0013	0016	0020	0026	0032	0039	0048	0059
10	0001	0001	0001	0002	0002	0003	0004	0005	0006	0007

$P(X \geq r|n,p)$

$n = 11$

r \ p	41	42	43	44	45	46	47	48	49	50
1	9970	9975	9979	9983	9986	9989	9991	9992	9994	9995
2	9739	9776	9808	9836	9861	9882	9900	9916	9930	9941
3	8938	9055	9162	9260	9348	9428	9499	9564	9622	9673
4	7269	7490	7700	7900	8089	8266	8433	8588	8733	8867
5	4948	5223	5495	5764	6029	6288	6541	6787	7026	7256
6	2690	2924	3166	3414	3669	3929	4193	4460	4729	5000
7	1121	1260	1408	1568	1738	1919	2110	2312	2523	2744
8	0343	0399	0461	0532	0610	0696	0791	0895	1009	1133
9	0072	0087	0104	0125	0148	0175	0206	0241	0282	0327
10	0009	0012	0014	0018	0022	0027	0033	0040	0049	0059
11	0001	0001	0001	0001	0002	0002	0002	0003	0004	0005

$n = 12$

r \ p	01	02	03	04	05	06	07	08	09	10
1	1136	2153	3062	3873	4596	5241	5814	6323	6775	7176
2	0062	0231	0486	0809	1184	1595	2033	2487	2948	3410
3	0002	0015	0048	0107	0196	0316	0468	0652	0866	1109
4		0001	0003	0010	0022	0043	0075	0120	0180	0256
5				0001	0002	0004	0009	0016	0027	0043
6							0001	0002	0003	0005
7										0001

r \ p	11	12	13	14	15	16	17	18	19	20
1	7530	7843	8120	8363	8578	8766	8931	9076	9202	9313
2	3867	4314	4748	5166	5565	5945	6304	6641	6957	7251
3	1377	1667	1977	2303	2642	2990	3344	3702	4060	4417
4	0351	0464	0597	0750	0922	1114	1324	1552	1795	2054
5	0065	0095	0133	0181	0239	0310	0393	0489	0600	0726
6	0009	0014	0022	0033	0046	0065	0088	0116	0151	0194
7	0001	0002	0003	0004	0007	0010	0015	0021	0029	0039
8					0001	0001	0002	0003	0004	0006
9										0001

r \ p	21	22	23	24	25	26	27	28	29	30
1	9409	9493	9566	9629	9683	9730	9771	9806	9836	9862
2	7524	7776	8009	8222	8416	8594	8755	8900	9032	9150
3	4768	5114	5450	5778	6093	6397	6687	6963	7225	7472
4	2326	2610	2904	3205	3512	3824	4137	4452	4765	5075
5	0866	1021	1192	1377	1576	1790	2016	2254	2504	2763
6	0245	0304	0374	0453	0544	0646	0760	0887	1026	1178
7	0052	0068	0089	0113	0143	0178	0219	0267	0322	0386
8	0008	0011	0016	0021	0028	0036	0047	0060	0076	0095
9	0001	0001	0002	0003	0004	0005	0007	0010	0013	0017
10						0001	0001	0001	0002	0002

$n = 12$ $P(X \geq r|n,p)$

r \ p	31	32	33	34	35	36	37	38	39	40
1	9884	9902	9918	9932	9943	9953	9961	9968	9973	9978
2	9256	9350	9435	9509	9576	9634	9685	9730	9770	9804
3	7704	7922	8124	8313	8487	8648	8795	8931	9054	9166
4	5381	5681	5973	6258	6533	6799	7053	7296	7528	7747
5	3032	3308	3590	3876	4167	4459	4751	5043	5332	5618
6	1343	1521	1711	1913	2127	2352	2588	2833	3087	3348
7	0458	0540	0632	0734	0846	0970	1106	1253	1411	1582
8	0118	0144	0176	0213	0255	0304	0359	0422	0493	0573
9	0022	0028	0036	0045	0056	0070	0086	0104	0127	0153
10	0003	0004	0005	0007	0008	0011	0014	0018	0022	0028
11				0001	0001	0001	0001	0002	0002	0003

r \ p	41	42	43	44	45	46	47	48	49	50
1	9982	9986	9988	9990	9992	9994	9995	9996	9997	9998
2	9834	9860	9882	9901	9917	9931	9943	9953	9961	9968
3	9267	9358	9440	9513	9579	9637	9688	9733	9773	9807
4	7953	8147	8329	8498	8655	8801	8934	9057	9168	9270
5	5899	6175	6443	6704	6956	7198	7430	7652	7862	8062
6	3616	3889	4167	4448	4731	5014	5297	5577	5855	6128
7	1765	1959	2164	2380	2607	2843	3089	3343	3604	3872
8	0662	0760	0869	0988	1117	1258	1411	1575	1751	1938
9	0183	0218	0258	0304	0356	0415	0481	0555	0638	0730
10	0035	0043	0053	0065	0079	0095	0114	0137	0163	0193
11	0004	0005	0007	0009	0011	0014	0017	0021	0026	0032
12				0001	0001	0001	0001	0001	0002	0002

$n = 13$

r \ p	01	02	03	04	05	06	07	08	09	10
1	1225	2310	3270	4118	4867	5526	6107	6617	7065	7458
2	0072	0270	0564	0932	1354	1814	2298	2794	3293	3787
3	0003	0020	0062	0135	0245	0392	0578	0799	1054	1339
4		0001	0005	0014	0031	0060	0103	0163	0242	0342
5				0001	0003	0007	0013	0024	0041	0065
6						0001	0001	0003	0005	0009
7									0001	0001

r \ p	11	12	13	14	15	16	17	18	19	20
1	7802	8102	8364	8592	8791	8963	9113	9242	9354	9450
2	4270	4738	5186	5614	6017	6396	6751	7080	7384	7664
3	1651	1985	2337	2704	3080	3463	3848	4231	4611	4983
4	0464	0609	0776	0967	1180	1414	1667	1939	2226	2527
5	0097	0139	0193	0260	0342	0438	0551	0681	0827	0991
6	0015	0024	0036	0053	0075	0104	0139	0183	0237	0300
7	0002	0003	0005	0008	0013	0019	0027	0038	0052	0070
8			0001	0001	0002	0003	0004	0006	0009	0012
9								0001	0001	0002

TABLA II

DISTRIBUCION POISSON

$$P(X \geq x) = \sum_{x}^{\infty} e^{-\lambda}\lambda^{x}/x!$$

x \ λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.0952	.1813	.2592	.3297	.3935	.4512	.5034	.5507	.5934	.6321
2	.0047	.0175	.0369	.0616	.0902	.1219	.1558	.1912	.2275	.2642
3	.0002	.0011	.0036	.0079	.0144	.0231	.0341	.0474	.0629	.0803
4	.0000	.0001	.0003	.0008	.0018	.0034	.0058	.0091	.0135	.0190
5	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014	.0023	.0037
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0006
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x \ λ	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.6671	.6988	.7275	.7534	.7769	.7981	.8173	.8347	.8504	.8647
2	.3010	.3374	.3732	.4082	.4422	.4751	.5068	.5372	.5663	.5940
3	.0996	.1205	.1429	.1665	.1012	.2166	.2428	.2694	.2963	.3233
4	.0257	.0338	.0431	.0537	.0656	.0788	.0932	.1087	.1253	.1429
5	.0054	.0077	.0107	.0143	.0186	.0237	.0296	.0364	.0441	.0527
6	.0010	.0015	.0022	.0032	.0045	.0060	.0080	.0104	.0132	.0166
7	.0001	.0003	.0004	.0006	.0009	.0013	.0019	.0026	.0034	.0045
8	.0000	.0000	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0011
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002

x \ λ	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.8775	.8892	.8997	.9093	.9179	.9257	.9328	.9392	.9450	.9502
2	.6204	.6454	.6691	.6916	.7127	.7326	.7513	.7689	.7854	.8009
3	.3504	.3773	.4040	.4303	.4562	.4816	.5064	.5305	.5540	.5768
4	.1614	.1806	.2007	.2213	.2424	.2640	.2859	.3081	.3304	.3528
5	.0621	.0725	.0838	.0959	.1088	.1226	.1371	.1523	.1682	.1847
6	.0204	.0249	.0300	.0357	.0420	.0490	.0567	.0651	.0742	.0839
7	.0059	.0075	.0094	.0116	.0142	.0172	.0200	.0244	.0287	.0335
8	.0015	.0020	.0026	.0033	.0042	.0053	.0066	.0081	.0099	.0119
9	.0003	.0005	.0006	.0009	.0011	.0015	.0019	.0024	.0031	.0038
10	.0001	.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0011
11	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

x \ λ	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9550	.9592	.9631	.9666	.9698	.9727	.9753	.9776	.9798	.9817
2	.8153	.8288	.8414	.8532	.8641	.8743	.8838	.8926	.9008	.9084
3	.5988	.6201	.6406	.6603	.6792	.6973	.7146	.7311	.7469	.7619
4	.3752	.3975	.4197	.4416	.4634	.4848	.5058	.5265	.5468	.5665
5	.2018	.2194	.2374	.2558	.2746	.2936	.3128	.3322	.3516	.3712
6	.0943	.1054	.1171	.1295	.1424	.1559	.1699	.1844	.1994	.2149
7	.0388	.0446	.0510	.0579	.0653	.0733	.0818	.0909	.1005	.1107
8	.0142	.0168	.0198	.0231	.0267	.0308	.0352	.0401	.0454	.0511
9	.0047	.0057	.0069	.0083	.0099	.0117	.0137	.0160	.0185	.0214
10	.0014	.0018	.0022	.0027	.0033	.0040	.0048	.0058	.0069	.0081
11	.0004	.0005	.0006	.0008	.0010	.0013	.0016	.0019	.0023	.0028
12	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0007	.0009
13	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

$$P(X \geq x) = \sum_{x}^{\infty} e^{-\lambda} \lambda^{x}/x!$$

x	λ = 4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9834	.9850	.9864	.9877	.9889	.9899	.9909	.9918	.9926	.9933
2	.9155	.9220	.9281	.9337	.9389	.9437	.9482	.9523	.9561	.9596
3	.7762	.7898	.8026	.8149	.8264	.8374	.8477	.8575	.8667	.8753
4	.5858	.6046	.6228	.6406	.6577	.6743	.6903	.7058	.7207	.7350
5	.3907	.4102	.4296	.4488	.4679	.4868	.5054	.5237	.5418	.5595
6	.2307	.2469	.2633	.2801	.2971	.3142	.3316	.3490	.3665	.3840
7	.1214	.1325	.1442	.1564	.1689	.1820	.1954	.2092	.2233	.2378
8	.0573	.0639	.0710	.0786	.0866	.0951	.1040	.1133	.1231	.1334
9	.0245	.0279	.0317	.0358	.0403	.0451	.0503	.0558	.0618	.0681
10	.0095	.0111	.0129	.0149	.0171	.0195	.0222	.0251	.0282	.0318
11	.0034	.0041	.0048	.0057	.0067	.0078	.0090	.0104	.0120	.0137
12	.0011	.0014	.0017	.0020	.0024	.0029	.0034	.0040	.0047	.0055
13	.0003	.0004	.0005	.0007	.0008	.0010	.0012	.0014	.0017	.0020
14	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006	.0007
15	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0002	.0002
16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

x	λ = 5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9939	.9945	.9950	.9955	.9959	.9963	.9967	.9970	.9973	.9975
2	.9628	.9658	.9686	.9711	.9734	.9756	.9776	.9794	.9811	.9826
3	.8835	.8912	.8984	.9052	.9116	.9176	.9232	.9285	.9334	.9380
4	.7487	.7619	.7746	.7867	.7983	.8094	.8200	.8300	.8396	.8488
5	.5769	.5939	.6105	.6267	.6425	.6579	.6728	.6873	.7013	.7149
6	.4016	.4191	.4365	.4539	.4711	.4881	.5050	.5217	.5381	.5543
7	.2526	.2676	.2829	.2983	.3140	.3297	.3456	.3616	.3776	.3937
8	.1440	.1551	.1665	.1783	.1905	.2030	.2159	.2290	.2424	.2560
9	.0748	.0819	.0894	.0974	.1056	.1143	.1234	.1328	.1426	.1528
10	.0356	.0397	.0441	.0488	.0538	.0591	.0648	.0708	.0772	.0839
11	.0156	.0177	.0200	.0225	.0253	.0282	.0314	.0349	.0386	.0426
12	.0063	.0073	.0084	.0096	.0110	.0125	.0141	.0160	.0179	.0201
13	.0024	.0028	.0033	.0038	.0045	.0051	.0059	.0068	.0078	.0088
14	.0008	.0010	.0012	.0014	.0017	.0020	.0023	.0027	.0031	.0036
15	.0003	.0003	.0004	.0005	.0006	.0007	.0009	.0010	.0012	.0014
16	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0004	.0004	.0005
17	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0002
18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x	λ = 6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9978	.9980	.9982	.9983	.9985	.9986	.9988	.9089	.9990	.9991
2	.9841	.9854	.9866	.9877	.9887	.9897	.9905	.9913	.9920	.9927
3	.9423	.9464	.9502	.9537	.9570	.9600	.9629	.9656	.9680	.9704
4	.8575	.8658	.8736	.8811	.8882	.8948	.9012	.9072	.9129	.9182
5	.7281	.7408	.7531	.7649	.7763	.7873	.7978	.8080	.8177	.8270
6	.5702	.5859	.6012	.6163	.6310	.6453	.6594	.6730	.6863	.6993
7	.4098	.4258	.4418	.4577	.4735	.4892	.5047	.5201	.5353	.5503
8	.2699	.2840	.2983	.3127	.3272	.3419	.3567	.3715	.3864	.4013
9	.1633	.1741	.1852	.1967	.2084	.2204	.2327	.2452	.2580	.2709
10	.0910	.0984	.1061	.1142	.1226	.1314	.1404	.1498	.1505	.1695
11	.0469	.0514	.0563	.0614	.0668	.9720	.0786	.0849	.0916	.0985
12	.0224	.0250	.0277	.0307	.0339	.0373	.0409	.0448	.0490	.0534
13	.0100	.0113	.0127	.0143	.0160	.0179	.0199	.0221	.0245	.0270
14	.0042	.0048	.0055	.0063	.0071	.0080	.0091	.0102	.0115	.0128
15	.0016	.0019	.0022	.0026	.0030	.0034	.0039	.0044	.0050	.0057
16	.0006	.0007	.0008	.0010	.0012	.0014	.0016	.0018	.0021	.0024
17	.0002	.0003	.0003	.0004	.0004	.0005	.0006	.0007	.0008	.0010
18	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003	.0004
19	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001

TABLA III DISTRIBUCION NORMAL

$$P(Z \leq z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997

TABLA IV

DISTRIBUCION CHI-CUADRADO

$$P(X \le x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw$$

	$P(X \le x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21
11	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72
12	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22
13	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69
14	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14
15	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58
16	5.812	6.908	7.962	9.312	23.54	26.30	28.84	32.00
17	6.408	7.564	8.672	10.08	24.77	27.59	30.19	33.41
18	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.80
19	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19
20	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57
21	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93
22	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
26	12.20	13.84	15.38	17.29	35.56	38.88	41.92	45.64
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89
40	22.16	24.43	26.51	29.05	51.80	55.76	59.34	63.69
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15
60	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38
70	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4
80	53.34	57.15	60.39	64.28	96.58	101.9	106.6	112.3

TABLA V
DISTRIBUCION t

$$P(T \leq t) = \int_{-\infty}^{t} \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)\left(1+w^2/r\right)^{(r+1)/2}}\,dw$$

$$[P(T \leq -t) = 1 - P(T \leq t)]$$

	$P(T \leq t)$				
r	0.90	0.95	0.975	0.99	0.995
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750

TABLA VI

DISTRIBUCION F

$$P(F \le f) = \int_0^f \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)(1+r_1 w/r_2)^{(r_1+r_2)/2}} dw$$

$P(F \le f)$	r_2	r_1: 1	2	3	4	5	6	7	8	9	10	12	15
0.95	1	161	200	216	225	230	234	237	239	241	242	244	246
0.975		648	800	864	900	922	937	948	957	963	969	977	985
0.99		4052	4999	5403	5625	5764	5859	5928	5982	6023	6056	6106	6157
0.95	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43
0.975		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43
0.99		98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43
0.95	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70
0.975		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25
0.99		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87
0.95	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86
0.975		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66
0.99		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20
0.95	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62
0.975		10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43
0.99		16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72
0.95	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94
0.975		8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27
0.99		13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56

0.95	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51
0.975		8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57
0.99		12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31
0.95	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22
0.975		7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10
0.99		11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52
0.95	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01
0.975		7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77
0.99		10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96
0.95	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85
0.975		6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52
0.99		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56
0.95	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62
0.975		6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18
0.99		9.33	5.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01
0.95	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40
0.975		6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86
0.99		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52

RESPUESTA A LOS PROBLEMAS PROPUESTOS

CAPITULO I

PROBLEMAS 1.2

2. (a) Si C = corazones, D = Diamantes, E = Espada y T = Trebol.

$\Omega = \{C_1, C_2, \ldots, C_{13}, D_1, D_2, \ldots, D_{13}, E_1, E_2, \ldots, E_{13}, T_1, T_2, \ldots, T_{13}\}$

(b) Si 0 = apagado y 1 = prendido, $\Omega_i = \{0,1\}$, $i = 1,2$; $\Omega = \{0,1\}^2$

(c) $\Omega_i = \{0,1\}$, $i = 1,2,\ldots,10$, $\Omega = \{0,1\}^{10}$

(d) $\Omega = \{0,1,2,\ldots, n\}$

(e) Si denotamos lero de Enero, 2 de Enero, ..., 30 de Diciembre, 31 de Diciembre por 1,2,3,4,5,...,365, 366 respectivamente.

$\Omega = \{1,2, \ldots, 365,366\}$

(f) $\Omega = \{0,1,2, \ldots\}$

(g) $\{1,2,\ldots, 365,366\}^n$

(h) $\Omega = \{(x, y) / x^2 + y^2 \leq r^2\}$

(i) Si $a_1, a_2, \ldots, a_{12}$ son las diferentes bolas,

$\Omega = \{a_1, a_2, \ldots, a_{12}\}^5$

3. Si {A,B,C,D,E} es el conjunto de 5 opciones,

$\Omega = \{AB, AC, AD, AE, BC, BD, BE, CD, CE, DE\}$

4. $\Omega = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}$

5. Si A_i = la persona A está en la oficina $i = 1,2,3$,

B_j = la persona B está en la oficina $j = 1,2,3$.

$$\Omega = \left\{ \begin{array}{l} A_1B_1, A_1B_2, A_1B_3 \\ A_2B_1, A_2B_2, A_2B_3 \\ A_3B_1, A_3B_2, A_3B_3 \end{array} \right\}$$

7. $\Omega = \{BBB, BBD, BDB, DBB, BDD, DBD, DDB, DDD\}$,

B = artículo bueno, D = artículo defectuoso

8. $\Omega = \{0,1,2,3, \ldots\}$

9. $\Omega = \{w \text{ tal que } 0 \leq w \leq 1{,}000\}$

10. $\Omega = \{w \text{ tal que } 0 \leq w \leq U\}$. 11. $\Omega = \{\omega \text{ tal que } 0 \leq \omega < \infty\}$

12. (a) $\Omega = \{w = (x_1, x_2)$ tal que $0 \le x_1, x_2 < \infty\}$

donde x_1, indica el tiempo de falla del transistor designado como número 1 y x_2 indica el tiempo de falla del transistor designado como número 2.

(b) $\Omega = \{w = (x_1, x_2, x_3, x_4, x_5) \; / \; 0 \le x_1, x_2, x_3, x_4, x_5 < \infty\}$

donde x_i indica el tiempo de falla del transistor designado como número i.

13.
$$\Omega = \left\{\begin{matrix} (2,1), & (2,3), & (2,5), & (2,7), & (2,9) \\ (4,1), & (4,3), & (4,5), & (4,7), & (4,9) \\ (6,1), & (6,3), & (6,5), & (6,7), & (6,9) \\ (8,1), & (8,3), & (8,5), & (8,7), & (8,9) \end{matrix}\right\}$$

16. $\Omega = \{4, *4, **4, ***4, \ldots\}$, donde $*$ = obtener un número diferente de 4.

17. A = {CCS, CSC, SCC, CCC} ; B = {CCS, CSS, SCS, SSS}

C = {SSS, CSS, SCS, SSC}

21. $\{****4, *****4, \ldots\}$

22. (a) $\Omega = \{D,R,B,E\}^{10} = \{(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}) / x_i = D, R, B, E,$ $i = 1,2,\ldots,10\}$

(b) A = {(D,D,D,D,D,D,D,D,D,D) , (E,E,E,E,E,E,E,E,E,E)}

$B = \{D,R,B\}^9 \times \{E\}$

PROBLEMAS 1.3

1. (a) $AB\bar{C} \cup A\bar{B}C \cup \bar{A}BC$; (b) $A \cup B \cup C$

(c) $\bar{A}\bar{B}\bar{C} \cup A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C \cup AB\bar{C} \cup A\bar{B}C \cup \bar{A}BC$; (d) ABC ;

(e) $\bar{A}\bar{B}\bar{C}$ o $\overline{A \cup B \cup C}$; (f) $\bar{A}BC \cup A\bar{B}C \cup AB\bar{C}$

(g) $A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C$; (h) $\bar{A}\bar{B}\bar{C} \cup A\bar{B}\bar{C} \cup AB\bar{C} \cup \bar{A}\bar{B}C$.

2. (a) Ω ; (b) $\{w$ tal que $w \ge 2{,}000\}$

(c) $\{w = (x_1,x_2,x_3,x_4,x_5) \; / \; x_1 \ge 1{,}000, \; x_2 \ge 1{,}500, \; x_3, x_4, x_5 \le 1{,}600\}$

4. (a) F (b) F, (c) V, (d) V, (e) F , (f) V

5. (a) $A = [A \cap B] \cup [A \cap \bar{B}]$ $(B = [A \cap B] \cup [\bar{A} \cap B]$

(b) $\bar{A} = [\bar{A} \cap B] \cup [\bar{A} \cap \bar{B}]$ $(\bar{B} = [A \cap \bar{B}] \cup [\bar{A} \cap \bar{B}])$

6. (a) Son mutuamente excluyentes porque ninguno de los dos pares de eventos A_i y A_j , contienen la misma tasa de desempleo (predicha).
Son colectivamente exhaustivos porque todas las predicciones posibles están comprendidas en esos cinco eventos.

(b) (1) El desempleo predicho fue del 8% al 10% mientras el desempleo real fue del 6% al 8%.

(2) El desempleo fue predicho como 4% o más pero menos del 8%

(3) Una predicción correcta.

(4) El desempleo real fue más alto de lo predicho.

7. (a) (1) Más de la mitad de las acciones subirán de precio o más de la mitad de las acciones no cambiarán de precio.

(2) Un evento imposible.

(b) Los tres eventos son mutuamente excluyentes

(c) No

PROBLEMAS 1.4

1. 324 . 2. 48 3. 648 . 4. 25! . 5. 6! x 5! .

6. (a) 100; (b) 48; (c) 48. 7. $4! \times (3!)^2 \times (2!)^2$.

8. 34,650 . 9. 27,720 .

10. (a) 84 (b) 7 ; 35 (c) 24 . 11. 63 cantidades diferentes

12. $\binom{12}{5}\binom{10}{6}$. 13. 14.

15. $\dfrac{10!}{(3!)^2 4!}$. 16. 144 . 17. 5,600

18. (a) $C_{10}^{5} \times C_{5}^{3} + C_{10}^{4} \times C_{5}^{4} + C_{10}^{5}$; (b) C_{13}^{6} . 19. $\dfrac{10!}{(3!)^2 4!}$.

20. 21. (a) 91 ; (b) 420 .

22. (a) 816 ; (b) $1.C_{17}^{2} = 136$. 23.

24. 10 . 25. (a) no ; (b) debería usar 4 letras

26. $(2!)^2(25!)$. 27. (a) 3,280 ; (b) resultará mas perjudicial perder 3 banderas rectangulares iguales.

28. 656 . 29. (a) 252 , (b) 120 . 30. $\frac{10!}{3!3!4!}$

PROBLEMAS 1.5

1. (a) Clásica ; (b) frecuencia relativa ; (c) subjetiva ;
   (d) frecuencia relativa ; (e) subjetiva .
2. (a) $\Omega = \{0,1,2,\ldots\}$, los eventos elementales de Ω no son igualmente posibles, pués hay pocos panetones con cero pasas
   (b) $\Omega = \{0,1,2,3,4\}$, los eventos elementales de Ω no son igualmente posibles.
   (c) $\Omega = \{0,1,2\}$, no son igualmente posibles
   (d) $\Omega = \{0,1,2,\ldots,200\}$, no son igualmente posibles
   (e) $\Omega = \{2,3,4,5,\ldots\}$ no son igualmente posibles.
3. (a) $\Omega = \{A_1 B_1, A_1 B_2, A_2 B_1, A_2 B_2\}$, cada evento elemental son igualmente - posibles (A, B objeto, 1 y 2 cajas, $A_i B_j$: el objeto A en la caja i y el objeto B en la caja j. $i,j = 1,2$). Se aplica la definición clásica P[una se queda vacia] $= \frac{1}{2}$.
   (b) Ω = {MMM, MMH, MHM, HMM, MHH, HMH, HHM, HHH} los eventos elementales son igualmente posibles. Se aplica la definición clásica. (M = mujer - H = hombre).
   P[sin varones] $= \frac{1}{8}$.
   (c) $\Omega = \{E_1, E_2, \ldots, E_{13}, C_1, C_2, \ldots, C_{13}, O_1, O_2, \ldots O_{13}, T_1, T_2, \ldots, T_{13}\}$
   E_i : la i-ésima espada; C_i : la i-ésimo corazón; O_i : el i-ésimo oro; T_i : el i-ésimo trebol,los eventos elementales son igualmente posibles. Se aplica la definicón clásica
4. (a) subjetiva; (b) frecuencia relativa; (c) frecuencia relativa, (d) clásica.
5. (a) frecuencia relativa; (b) P[de recibir una respuesta] = 0.102 .
6. $\dfrac{(C_4^1)^3}{C_{52}^3}$. 7. 3/8 . 8. $\binom{26}{3}\binom{26}{3}\Big/\binom{52}{6}$.
9. 0.079 . 10. (a) 5/18 ; (b) 7/42 . 11. $\dfrac{2^4 C_6^2}{3^6}$. 12. $\dfrac{2^2 C_5^2 C_3^1}{4^5}$.

13. (a) $\dfrac{12!}{(2!)^6} \times \dfrac{1}{6^{12}}$; (b) $\dfrac{6!}{2!} \times \dfrac{12!}{(2!)^2 3!4!} \times \dfrac{1}{6^{12}}$. 14. $\dfrac{4!(3!)^2}{8!}$.

15. $\dfrac{C_4^1\, C_{13}^3\, C_3^2\, C_{13}^1\, C_{13}^1}{C_{52}^5}$. 16. (a) $\dfrac{2}{3}$, (b) $\dfrac{1}{3}$. 17. $\dfrac{10!4!3!}{15!}$.

18. 4/81 .

19. (a) 1/8 ; (b) 1/2 ; (c) 3/8 ; (d) 1/4 ; (e) 1 .

20. (a) 1/9 ; (b) 2/9 ; (c) 1/72 ; (d) 4/9 ; (e) 2/72 .

21. (a) 0.4865 ; (b) 0.12137 .

22. $\dfrac{55 \times 2^{11}}{3^{13}}$. 23. $\dfrac{\binom{n}{r}(N-1)^{n-r}}{N^n}$. 24. $\dfrac{\binom{n}{2}n!}{n^n}$.

25. $\dfrac{\binom{4}{2}\binom{48}{3} + \binom{4}{3}\binom{48}{2} + \binom{4}{4}\binom{48}{1}}{\binom{52}{5}}$. 26. 0.3 .

28. (a) $\dfrac{9.8.7.2^6}{1.2.3.3^9}$; (b) $\dfrac{9!}{(3!)^3.3^9}$. (c) $\dfrac{3.2\binom{9}{2}\binom{7}{3}}{3^9}$.

29. $P[A] = \dfrac{1}{6}$, $P[B] = \dfrac{1}{2}$, $P[C] = \dfrac{1}{3}$.

30. (a) $\dfrac{916}{1001}$; (b) $\dfrac{160}{1001}$.

31. 11/21 . 32. (a) $\dfrac{55}{147}$; (b) $\dfrac{22}{147}$.

33. 1/8 , 1/8 . 34. 0.05 . 35. $\dfrac{1}{13}$, $\dfrac{4}{13}$, $\dfrac{8}{13}$. 36. 1/3 .

37. (a) $\dfrac{43}{216}$; (b) $\dfrac{20}{216}$. 38. $\dfrac{35}{108}$. 39. 0.1536

44. (a) 1/216 ; (b) 5/18 .

45. (a) P[300,000 ó 350,000] = $\dfrac{8}{13}$; (b) P[más de 400,000 ó menos de 200,000] = 0

46. (a) 1/15 ; (b) 7/15 ; (c) 14/15 .

47. Las apuestas en contra de A son de 13 a 5 .

48. 1/3 . 49. 1/12 . 50. 80/189 .

51. (a) $\frac{9!}{(3!)^3 3^9}$, (b) $\frac{9!}{4!3!3^9}$

52. (a) 5/16 , (b) 15/32 , (c) 3/16 .

53. 22/425 .

PROBLEMAS 1.6

1. (a) 0.7 ; (b) 0.1 ; (c) 0.1 ; (d) 0.5 ; (e) 0.9
2. 0.1 .
3. (a) 0.9 ; (b) 0.8
4. (a) 0.6 ; (b) 0.1 ; (c) 0.7
5. (a) 0.3 ; (b) 0.7 .
6. 1/6 , 1/3 , 1/6 , 0.0 , 5/12 .
7. no . 8. no . 9. no .

12. (a) (b) (c) 0.5 .
13. (a) 0.89 ; (b)

17. 1/4 . 18. 0.6 . 19. 2/25, 71/500 .
20. 1401/50,000 . 21. 0.4 . 22. 7/15 .
23. 2/3 . 24. 0.48 . 25. 0.6 .

27. 0.3 . 28. (a) 0.945 ; (b) 0.395 .
29. 0.95 .
30. (a) 25/26 ; (b) 21/26 .
31. (a) 21/52 ; (b) 9/52 .
32. (a) 41 % ; (b) 12 % ; (c) 11% (d) 89% .
33. 0.155 .
34. (a) 29/50 ; (b) 42/50 ; (c) 1/10 .
35. (a) 12/25 ; (b) 9/50 ; (c) 195/250 .

PROBLEMAS 1.7

1. 0.4 . 2. 1/8 .
3. 7/12, 3/4, 1/2, 1 4. 2/3, 3/4, 2/3 .
5. (a) falso, ya que $P[AB] = \frac{1}{8}$;

(b) falso; (c) verdadero ; (d) falso .

7. $P[B_1]P[A|B_1] + P[B_2]P[A|B_2]$.

8. (a) si ; (b) si .

9. (a) 0.05 ; (b) 0.25 . 10. 7/40 .

11. (a) $\frac{455}{1140}$; (b) $\frac{525}{1140}$; (c) $\frac{150}{1140}$; (d) $\frac{10}{1140}$.

12. (a) 1/5 ; (b) 2/5 .

13. 1/2 . 14. 1/3 . 15. 0.26 .

16. 1/6, 1/3, 1/2 . 17. 0.846 .

18. $\frac{10}{11}$. 19. 0.015 .

21. 5/27 . 22. 97/324 .

23. 24. (a) 1/14 ; (b) 1/56 .

25. 2/7 . 26.

27. 28. 0.5658, 0.2914, 0.1427 .

29. 30. 5/26 .

31. 0.55 . 32. $\frac{77}{165}$; $\frac{53}{165}$; $\frac{35}{165}$

33. (a) 0.016 ; (b) 0.008 .

PROBLEMAS 1.8

1. 55/132 . 2. (a) 9/25; (b) 16/25 . 3. 69/140

4. 3/44 . 5. (a) 61/216; (b) 371/1296 . 6. 0.66 .

7. 31/60 . 8. 0.569 . 9. 13/108 .

10. 0.12 . 11. 1/3 . 12.

13. 0.7 . 14. 0.025 . 15. (a) 1/2; (b) 2/3; (c) 1/12 .

16. 1/3 . 17. 17/40 . 18. 2/5 .

19. 0.948 . 20. 53/140 . 21. 1/5 .

22. 4/13 . 23. (a) 58/135 ; (b) 9/29 .

24. (a) 12/29 ; (b) 9/29 ; (c) 8/29 . 25. 1/2 .

26. 1/3 . 27. 1/2 .

28. (a) 0.75 ; (b) 0.2 . 29. 10/19. 30. (a) $\frac{4}{5}$; (b) $\frac{1}{9}$.

31. 12 . 32. (a) 1/2 ; (b) 1/3 .

33. (a) 0.6 ; (b) 0.4 ; (c) 3/8 . 34.

35. (a) 0.2 ; (b) 0.12 . 36. (a) 1/4 ; (b) 1/2 .

37. (a) 0.44 ; (b) 3/22 .

38. (a) 0.22 ; (b) 7/11 . 39. 0.20 .

40. 0.47 . 41. 3/4 . 42. 8/11 .

43. 8/17 . 44. 45. 0.51 .

46. 63/139 . 47. 0.33 . 48. 5/8 .

49. 2/7 . 50. 4/9; 1/5; 2/15; 2/9 . 51. 1/2 .

52. 20/63 . 53. 5/8 . 54. 0.80 .

55. 8/53 .

PROBLEMAS 1.9

1. Si, $P[AB] = \frac{1}{18} = \frac{1}{6}\frac{1}{3} = P[A]P[B]$.

3. A y B son independientes ; A y C son independientes
B Y C no son independientes .

4. 5/6 . 5. 1/2 . 6. 0.4 . 7. 3/4 .

8. (a) 0.44 ; (b) 0.38 ; (c) 0.56 ; (d) 0.06 .

9. 1/3 .

10. 0.2035 .

11. (a) 0.063 (b) 0.9 .

13. 6/13 .

14. $1 - (\frac{5}{6})^6$.

15. (a) 1/16 ; (b) 47/48 ; (c) $n = 3$.

16. (a) $(1/6)^4$ (b) $(1/2)^4$; (c) $(1/6)^3$.

17. (a) $(1/2)^n$; (b) $(1/2)^n$; (c) $1 - (1/2)^n$.

18. (a) No ; (b) 1 .

19. 7 . 20. 29 .

21. (a) 0.585625; (b) 0.434375 ·

22. $\binom{10}{3}(0.6)^3(0.4)^7$; $1-(0{,}4)^{10}-10(0.6)(0.4)^9-\binom{10}{2}(0.6)^2(0.4)^8$.

23. (a) 0.02 ; (b) 0.076 ·

24. (a) 1/2 ; (b) 1/2 ; (c) 1/4 ; (d) 1/6 .

25. 53/80·

26.

x	0	1	2	3	4	5
$P(x)$	0.132	0.329	0.329	0.165	0.041	0.004

27. 0.74 .

28. 0.9605 . 29. (a) 0.1296 ; (b) 0.8704 ; (c) $n=6$.

30. (i) 32/75 ; (iii) 44/75 ·

31. (a) $(7/22)^3(35/66)^2$; (b) $3(7/22)^3(10/66)^2$. 32. 0.933 .

33. $6!(1/6)^6$. 34. $6\,\dfrac{7!/2!}{6^7}$. 35. $0.021=\dfrac{1296}{65771}$

37. $(\frac{4}{9})^5+5(\frac{4}{9})^4(\frac{5}{9})$ 38. 0.1 , 0.2 y 0.5 . 39. 0.039 .

40. 0.0776 . 41. 0.0016 . 42. 0.16 . 43. 59/64 . 44. 56/188

45. (i) 2/9 ; (ii) 7/30 . 46. 0.054271 . 47. 0.864 .

48. (a) 2 ; (b) 3 ; (c) 4 . 49. 1/8 ; 1/4 ; 5/16 .

50. 31/324 . 51. 4/81 . 52. 3/8 .

PROBLEMAS 1.10

1. 2/3, 1/3 . 2. $P(A)=\dfrac{30}{61}$, $P(B)=\dfrac{31}{61}$.

3. 4/9 · 4. (a) 9/64 ; (b) $11/2^4$; (c) $29/2^6$.

5. (a) 1/2 ; (b) 1/3 · 6. 6/17 . 7. $\dfrac{244}{495}$.

8. (a) $(0.9)^4$; (b) $(0.9)^{10}$.

9. $\frac{13}{25}$; $\frac{12}{25}$. 10. (a) 0.973 ; (b) 0.09 ; (c) 0.23 .

11. $P[G_A] = \frac{5}{14}$; $P[G_B] = \frac{5}{14}$; $P[G_C] = \frac{4}{14}$; $P[par] =$

12. $\frac{32}{95}$; $\frac{65}{95}$. 13. $\frac{9}{19}$; $\frac{6}{19}$; $\frac{4}{19}$.

14.
$$P[G_A] = \begin{cases} \frac{30}{61}\left[1 - \left(\frac{155}{216}\right)^m\right], & \text{si } n \text{ par}, (n = 2m) \\ \frac{30}{61}\left[1 - \left(\frac{31 \times 30}{36^2}\right)^{m+1}\right], & \text{si } n \text{ impar}, (n = 2m + 1) \end{cases}$$

$$P[G_B] = \begin{cases} \frac{31}{61}\left[1 - \left(\frac{155}{216}\right)^m\right], & \text{si } n \text{ par } (n = 2m) \\ \frac{31}{61}\left[1 - \left(\frac{155}{216}\right)^m\right], & \text{si } n \text{ impar } (n = 2m + 1) \end{cases}$$

$$P[E] = \begin{cases} \left(\frac{155}{216}\right)^m & \text{si } n = 2m \\ \frac{31}{36}\left(\frac{155}{216}\right)^m & \text{si } n = 2m + 1 \end{cases} .$$

15. 2/5 · 16. 0.369 . 21. $1 + \frac{\pi}{3} - \sqrt{3} = 0.3151$

CAPITULO 2

PROBLEMAS 2.1

1. (a) $\Omega = \{1,2,3,\ldots,12\}$, (b)

   (c) $R_X = \{1,2,3,4,6\}$

   (d) $A = \{\omega \in \Omega \ / \ X(\omega) \in E_X\} = \{2\}$,

   $B = \{3\}$, $C = \{4\}$, $D = \{1,6\}$

2. (a) $\Omega = \{(i,j) \ / \ i,j = 1,2,3,\ldots,6\}$

   (b) $R_X = \{1,2,3,4,5,6\}$

3. (a) Si, (b) $\{0,1,2,\ldots 10\}$

4. (a) {3,4,5,6,7,8,9} (b) $\frac{1}{10}$, $\frac{2}{10}$, $\frac{5}{10}$

5. (a) {2,3,4,5}

(b) {NNDD, NDND, DNND} ; (c) {NDD, DND}

6. {- 1, 9, 99, 499}

PROBLEMAS 2.2

1. (a) $k = 10$, (b) $k = \frac{1}{78}$ (c) $k = 4/5$, (d) $k = \frac{1}{29}$

2. (a)

R_X	0	1	5
$p(x)$	$\frac{10}{20}$	$\frac{3}{20}$	$\frac{7}{20}$
$F(x)$	$\frac{10}{20}$	$\frac{13}{20}$	1

(b)

3. (a) $\Omega = \{\bar{B}\,\bar{B}, \bar{B}\,B, B\bar{B}, BB\}$ donde B = bola blanca y
$\bar{B}$ = bola diferente de blanca

(b) R = {0,1,2}

(c)

x	0	1	2
$p(x)$	$\frac{1}{15}$	$\frac{8}{15}$	$\frac{6}{15}$

(d) $F(x) = \begin{cases} 0 & , \ x < 0 \\ 1/15 & , \ 0 \leq x < 1 \\ 9/15 & , \ 1 \leq x < 2 \\ 1 & , \ x \geq 2 \end{cases}$

4.

R_X	0	1	2
$p(x)$	$\frac{15}{70}$	$\frac{40}{70}$	$\frac{15}{70}$
$F(x)$	$\frac{15}{70}$	$\frac{55}{70}$	1

$$p(x) = \frac{\binom{2}{x}\binom{6}{4-x}}{\binom{8}{4}}$$

$x = 0,1,2.$

5. (a) $d = \frac{1}{12}$, (b) $F(3.6) = \frac{5}{12}$;

(c) $P(3 \leq X < 5) = \frac{1}{3}$

6.

R_X	-10	90	990	$4{,}990$
$p(x)$	$\frac{990}{1{,}000}$	$\frac{5}{1{,}000}$	$\frac{4}{1{,}000}$	$\frac{1}{1{,}000}$

7. (a)

R_X	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

(b)

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8, & 0 \leq x < 1 \\ 1/2, & 1 \leq x < 2 \\ 7/8, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

8.

R_X	-750	500	1750	$3{,}000$
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

9. (a) $f(x) = \frac{1}{10}$, $\quad x = 1, 2, \ldots, 10$

(b) $$F(x) = \begin{cases} 0 & x < 1 \\ \frac{[\![x]\!]}{10}, & 1 \leq x < 10 \\ 1, & x \geq 10 \end{cases}$$

(c) $f(x) = \frac{1}{n}$, $\quad = 1, 2, \ldots, n$.

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{[\![x]\!]}{n}, & 1 \leq x < n \\ 1, & x \geq n \end{cases}$$

12.

x	1	2	3	4
$p(x)$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
$F(x)$	$\frac{4}{10}$	$\frac{7}{10}$	$\frac{9}{10}$	1

15. Ω = {1C, 1S, 2CC, 2CS, 2SC, 2SS, 3CCC, 3CCS, 3CSC, 3SCC, 3CSS, 3SCS, 3SSC, 3SSS}

x	0	1	2	3
$p(x)$	$\frac{7}{24}$	$\frac{11}{24}$	$\frac{5}{24}$	$\frac{1}{24}$

$$F(x) = \begin{cases} 0 & x < 0 \\ 7/24 \ , & 0 \leq x < 1 \\ 12/24 \ , & 1 \leq x < 2 \\ 23/24 \ , & 2 \leq x < 3 \\ 1 \ , & x \geq 3 \end{cases}$$

16. (a) $\Omega = \{(x, y) / x, y = 1,2,\ldots,8\}$, (b) $R_X = \{1,2,3,4,5,6,7,8\}$

$$p(x) = \frac{15 - 2(x-1)}{64} \quad , \quad x = 1,2,\ldots,8$$

(c) $$F(x) = \begin{cases} 0 & , \quad x < 1 \\ \dfrac{[\![x]\!]\,[16 - [\![x]\!]]}{64} & , \quad 1 \leq x < 8 \\ 1 & , \quad x \geq 8 \end{cases}$$

(d) $[X \leq 2.5] = \{1,2\}$, $[X \leq 5.75] = \{1,2,3,4,5\}$

$[X > 6.90] = \{7,8\}$

(e) $P[X \leq 5.75] = \frac{55}{64}$, $P[1.75 \leq X \leq 6.75] = \frac{45}{64}$

17. Ω = todas las familias de la ciudad.

x	0	5	10	15	20
$p(x)$	1/3	1/3	1/6	1/12	1/12
$F(x)$	1/12	8/12	10/12	11/12	1

18.

x	1	2	3	4	5	6	7	8	9
$p(x)$	9/81	12/81	16/81	12/81	12/81	10/81	6/81	3/81	1/81

19. (a) $k = 4/15$; (b) 5/6 .

PROBLEMAS 2.3

2. $a = \frac{1}{2\pi}$.   3. (a) $a = \frac{1}{\pi}$;   (b) $\frac{1}{3}$.

4. (a) a =0.5 ;   (b) $F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}(1 - \cos X) , & 0 \leq x < \pi \\ 1 , & x \geq \pi \end{cases}$

5. (a) a = 3/2 ;   (b) $F(x)$ =   6. C = 2 .

7. (a) a = 0.5 ;   (b) 500.   8. (a) 0.6458 ;   (b) 15 500

9. (a) 0.03078 ;   (b) 0.14946

10. (a) 0.973 ;   (b) 0.1846 ;   (c) 0.50

11. (a) 0.9596 ;   (b) 0.2789.   12. 1/3

13. (a) 0.8376 ;   (b) 19 600 ;   (c) 0.9294.   14. k = 1/4

15. (a) a = 1/25 ;   (b) $F(x) = \begin{cases} 0 & x < -5 \\ \frac{1}{2} - \frac{x^2}{50} , & -5 \leq x < 0 \\ \frac{1}{2} + \frac{x^2}{50} , & 0 \leq x < 5 \\ 1 . & x \geq 5 \end{cases}$

16. (a) $\frac{9 - 2\sqrt{6}}{3}$;   (b) 0.625.   17. (a) $f(x) = \begin{cases} x - 1 , & 1 \leq x \leq 2 \\ 3 - x , & 2 \leq x \leq 3 \end{cases}$

18. (a) c = 10.7 ,   (b) d = 13.   19. a = 500 , b = 776.4

22. $m = \sqrt{1/2}$.   23. $m = \frac{\pi}{2}$.   24. $c = (0.95)^{1/4}$.

25. (a) $\frac{9}{16}$,   (b) $\frac{27}{64}$.   26. $(0.1)^{1/4}$.

28. $1 - (0.01)^{1/5}$ millares de galones

29 . (a) $F(x) = \begin{cases} 0 , & x \leq \frac{\pi}{6} \\ -\cos 3x , & \frac{\pi}{6} < x \leq \frac{\pi}{3} \\ 1 , & x > \frac{\pi}{3} \end{cases}$

31. (a) $F(x) = \begin{cases} 0 , & x < -1 \\ \frac{x+1}{2} , & -1 \leq x \leq 1 \\ 1 , & x > 1 \end{cases}$

(b) 0.75 ; (c) 0.5 ; (d) 0.45 ; (e) 1 ;
(f) 0 ; (g) 0.7 .

32. (a) $\frac{3}{4}e^2$, (b) $\frac{1}{4}$; (c) $\frac{1}{4}$; (d) 0 ;
(e) $1 - \frac{3}{4}e^3$.

CAPITULO 3

PROBLEMAS 3.2

1. I/. 0.29 . 2. I/. 2 . 3. 3/4 .

5. (a) $\frac{395}{1976}$; (b) E(X) = 0.375 . 6. $E(X) = \frac{5}{2}$.

7. E(X) = 1.7 . 8. - 0.0965 , - 0.965

9. (a) no , (b) I/. 6 .

13. E(X) = 2 lanzamientos.

15. E(X) = 37.45 tubos por cada caja.

21. Debe comprar al proveedor A.

22. 5 panes por día.

23. S/. 3,000 ganancia esperada.

25. $E(X_A) = 32$, $E(X_B) = 24$, $E(X_C) = 18$.

28. 2,333.33 dólares por noche .

32. (a) I/. 120, I/. 30 , I/. 60 ; (B) I/. 75 , I/. 75, I/. 60 .

33. 3.984 intis por hora .

34. (a) 5, $\frac{3}{2}$, 6 , $\frac{21}{2}$. (b) $x_{ma} = 1,2,3,$ $x_{me} =$

CAPITULO 4

PROBLEMAS 4.3

1.

y \ x	1	2	3
1	0	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	0	$\frac{1}{6}$
3	$\frac{1}{6}$	$\frac{1}{6}$	0

2.

x \ y	0	1	2	3	$P[X=x]$
0	0	$\frac{6}{27}$	0	0	$\frac{2}{9}$
1	$\frac{6}{27}$	$\frac{6}{27}$	$\frac{6}{27}$	0	$\frac{6}{9}$
2	$\frac{2}{27}$	0	0	$\frac{1}{27}$	$\frac{1}{9}$
$P[Y=y]$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$	1

X e **Y** son independientes

3.

(a)

x \ y	1	2	3	4	5	$P[X=x]$
1	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{5}$
2	0	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{5}$
3	0	0	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{5}$
4	0	0	0	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{5}$
5	0	0	0	0	$\frac{1}{5}$	$\frac{1}{5}$
$P[Y=y]$	$\frac{12}{300}$	$\frac{27}{300}$	$\frac{47}{300}$	$\frac{77}{300}$	$\frac{137}{300}$	1

(c) $P[Y = y \mid X = 3] = \frac{1}{3}$, para $y = 3,4,5$

(d) $P[X = x \mid Y = 3] = \frac{12}{47}, \frac{15}{47}, \frac{20}{47}$, para $x = 1,2,3$ respectivamente

(e) $\frac{7}{15}$; $\frac{163}{300}$

4. (a) $p_X(x) = \dfrac{2x^2+1}{32}$, $x = 0,1,2,3$

$p_Y(y) = \dfrac{2y^2+7}{16}$, $y = 0,1$

(b) $p(x|y) = P[X|Y = y] = \dfrac{x^2+y^2}{2(2y^2+7)}$

para $x = 0,1,2,3,$ e $y = 0,1$

(c) $p(y|x) = P[Y|X = x] = \dfrac{x^2+y^2}{2x^2+1}$

para $x = 0,1,2,3,$ e $y = 0,1$.

5. Dividir anotaciones entre 2652.

$y \backslash x$	0	1	2	$P[Y = y]$
0	1260	216	6	1482
1	864	144	6	1014
2	132	24	0	156
$P[X = x]$	2256	384	12	

$P[X > Y] = \dfrac{228}{2652}$

6. Dividir anotaciones entre 90 .

$y \backslash x$	0	1	2	
0	0	0	6	6
1	0	42	0	42
2	42	0	0	42
	42	42	6	

$P[X \leq Y] = \dfrac{84}{90}$

8. (a) 7/30 , (b) 0.79 , (c) son dependientes

9\. (a)

$y \backslash x$	0	1	2	3	4
0	$\frac{1}{16}$	0	0	0	0
1	0	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
2	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	0
3	0	$\frac{1}{16}$	$\frac{1}{16}$	0	0
4	0	$\frac{1}{16}$	0	0	0

(b) $\frac{1}{16}$; (c) $\frac{11}{16}$; (d) $\frac{5}{16}$; (e) $\frac{1}{4}$

(f) 0 ; (g) $\frac{3}{8}$.

10\. (a) $p_X(x) = \frac{x}{6}$, $x = 1,2,3$

$p_Y(y) = \frac{y+1}{9}$, $y = 1,2,3$

(b) $p(X|1) = \frac{3x-2}{12}$, $x = 1,2,3$

$p(X|2) = \frac{3x}{18}$, $x = 1,2,3$

$p(X|3) = \frac{3x+2}{24}$, $x = 1,2,3$

(c) $p(y|1) = \frac{2y-1}{9}$, $y = 1,2,3$

$p(y|2) = \frac{y+1}{9}$, $y = 1,2,3$

$p(y|3) = \frac{2y+5}{27}$, $y = 1,2,3$

11\. $E(X) = 1$; $E(Y) = 0.6$; $E(X+Y) = 1.6$; $E(XY) = 0.3$

14. (a) 0.4 ; (b) 1.4 ; (c) 6.8 ; (d) 2.5 ; (e) 0.54 ; (f) 0.65
(g) 1/3 ; (h) 1/3 ; (i) 5/3 ; (j) 5/3 ; (k) 1/3 .

15. (a) 2.2 ; (b) 1 ; (c) 4.2 ; (d) 1.2 ; (e) 2.6 ;
(f) Var(X) ; (g) - cov(X,Y) ; (h) 1 ; (i) - P(X,Y) .

16. (a)

x	0	1	2	3
$p_X(x)$	3/16	2/16	5/16	6/16

y	1	2
$p_Y(y)$	8/16	8/16

(b) 54/16 ; (c) 79/64 ; (d) 29/5

PROBLEMAS 4.4

1. $k = 2/\log_e 2$

2. $f_X(x) = 1$, $0 < x < 1$, $f_Y(y) = -\log_e y$, $0 < y < 1$

5. 62.5 ; $195\frac{5}{6}$; 260.42 ; $\frac{1}{3}$; $8\frac{1}{3}$; 0.894 6. Si

8. (a) $f_X(x) = \frac{2}{100} e^{-2x/100}$;

$$f_Y(y) = \frac{2}{100} e^{-5/100}(1 - e^{5/100})$$

(b) No

(c) $$f_{Y|X=60}(y|x) = \begin{cases} \frac{1}{100} e^{60/100} e^{-y/100}, & \text{para } y > 60 \\ 0 & \text{, en otros casos} \end{cases}$$

10. 0.8452 . 11. 0.5269 . 12. 0.4582 .

CAPITULO 5

PROBLEMAS 5.2

1. $n = 25$, $P = \frac{1}{5}$, 5 ; 2. $p(x) = \binom{6}{x}(\frac{3}{4})^x(\frac{1}{4})^{6-x}$ $x = 0,1,2,3,4,5,6$;
$p(4) \approx 0.30$; 3. 0.05853 ; 4. (a) 0.2036 ; (b) 0.2262

5. (a) $\binom{6}{2}(0.2)^2(0.8)^4$, (c) $1 - 6(0.2)^5(0.8) - (0.2)^6$; (d) $\mu = 1.2$
(b) $1 - (0.8)^6$.

6. (a) $(0.75)^6$, (b) $(0.25)^6$, (c) $\binom{6}{4}(0.75)^4 (0.25)^2 = 0.296631$

(d) $\binom{6}{4}(0.75)^4 (0.25)^2 + \binom{6}{5}(0.75)^5 (0.25) + (0.75)^6 = 0.8306$.

7. $n = 20$, $p = 0.60$, (a) 0.998 , (b) 0.126 , (c) 0.245 .

8. $\binom{10}{3}(0.1)^3 (0.9)^7$.

9. (a) $(0.9995)^6$, (b) $(0.9995)^6 + 6(0.0005)(0.9995)^5$

(c) $\binom{6}{2}(0.0005)^2 (0.9995)^4$.

10. $P(X \geq 2) = 0.0063$, $P(Y \geq 3) = 0.00873$, el avión de 6 motores tiene - mayor probabilidad de efectuar un vuelo exitoso.

11. (a) 0.0047 , 0.2186 , 0.0115 ; (b) $np = 4.5 \simeq 5$ tubos ;

(c) 4.

12. (a) 3/8 , (b) 7/8 , (c) $n = 4$.

13. $\binom{10}{3}(0.95)^7 (0.05)^3$.

14. (a) 0.068 , (b) 0.605 .

15. 99 plantas .

16. $(0.9)^{10}$, 1 .

17. (a) $P(X \geq 11) = \sum_{x=4}^{20} \binom{20}{x} (\frac{4}{5})^{20-x} (\frac{1}{5})^x$; (b) $np = 4$.

18. 270 , 15 .

20. 0.061 .

21. 3/4 .

22. (a) 0.3125 ; (b) 0.2373 .

23. 0.68256.

24. (a) 27/64 ; (b) 27/32 .

25. (a) 0.0064 ; (b) aproximadamente 1 alumno .

26. en ambos casos la distribución de probabilidad es $\binom{10}{x}(\frac{1}{2})^{10}$ $x = 0,1,\ldots,10$

$P(X \geq 9) + P(Y \geq 9) = \frac{11}{2^9}$.

27. $P(X > 1) = \frac{3}{4}$; (a) $\binom{10}{4}(\frac{3}{4})^4 (\frac{1}{4})^6$; (b) $\sum_{x=4}^{10} \binom{10}{x}(\frac{3}{4})^x (\frac{1}{4})^{10-x}$.

28. S/. 6,223.

30. (a) $\binom{10}{9}(\frac{2}{3})^9(\frac{1}{3})$; (b) 0 ; (c) $\binom{10}{4}(\frac{2}{3})^4(\frac{1}{3})^6$.

31. X = precio final de venta de cada paquete ; Y = número de artículos de clase A en la muestra de tamaño 10. La v.a Y tiene distribución binomial

$n = 10$, $p = \frac{1}{3}$, $X = 200 + 10Y$

(a) $P(X = 270) = \binom{10}{7}(\frac{1}{3})^7(\frac{2}{3})^3$; (b) $P(X \leq 230) = \sum_{y=0}^{3}\binom{10}{y}(\frac{1}{3})^y(\frac{2}{3})^{10-y}$

(c) $E(X) = 200 + 10\,E(Y) = 200 + 10(\frac{10}{3}) = \frac{700}{3}$.

32. X = número de clientes que no pagaron (de los 10);
X es binomial con parametro $n = 0$, $p = 0.1$
Y = utilidad obtenida en estas 10 ventas;
$Y = 1000 - 300X$. De aqui se obtiene la distribución de Y .
$E(Y) = 1000 - 300\,E(X) = 700$.

33. $P(d(A,B) = 5) = P(X = 4) + P(X = 8) = \binom{12}{8}(\frac{2}{3})^4(\frac{1}{3})^4[(\frac{2}{3})^4 + (\frac{1}{3})^4]$

donde X = número de veces que aparece un número primo.

X es binomial con $n = 12$ y $p = \frac{2}{3}$

34. $P(X > 5) = \sum_{x=6}^{10}\binom{10}{x}(\frac{1}{5})^x(\frac{4}{5})^{10-x}$.

35. $\sum_{x=0}^{2}\binom{10}{x}(\frac{1}{6})^x(\frac{5}{6})^{10-x}$.

37. (a) 0.17709 ; (b) 0.7869 .

38. (a) 0.6815 ; (b) 0.1153 .

39. (a) 0.896 ; (b) 0.47 .

40. 6 % menos que el estándar .

41. $\binom{20}{x}(0.02)^x(0.98)^{20-x}$.

PROBLEMAS 5.3

1. $(\frac{5}{6})^6$ 2. $\mu = 4$.

3. (a) $\frac{6}{25} + \frac{12}{125} + \frac{24}{725}$; (b) $1 - [\frac{6}{25} + \frac{12}{125}]$.

4. $\frac{3}{4}$, $\frac{1}{4}$.

5. $\mu = 50$ ó $\mu = 500$ intis .

6. $\mu = 6$.

7. (a) 0.16 , (b) 0.488 .

8. 0.942 .

10. (a) $P[X = x] = \binom{x-1}{9}(0.1)^{10}(0.9)^{x-10}$, $= 10,11,\ldots$

(b) $P[X = 15] = \binom{14}{9}(0.1)^{10}(0.9)^{5}$

12. $ 19,125 .

PROBLEMAS 5.4

1. 0.058 .

2. (a) $p(9.0,0) = \frac{9!}{9!0!0!}(0.5)^9(0.4)^0(0.1)^0 = (0.5)^9$

(b) $p(5,3,1) = \frac{9!}{5!3!1!}(0.5)^5(0.4)^3(0.1)$

(c) $p(3,3,3) = \frac{9!}{3!3!3!}(0.5)^3(0.4)^3(0.1)^3$.

4. (a) 0.064 .

PROBLEMAS 5.5

1. $p(x) = \binom{26}{x}\binom{26}{13-x}/\binom{52}{13}$, $x = 0,1,\ldots,13$; $\frac{13}{2}$; $\frac{169}{68}$.

2. 0.9517 .

3. (a) $p(x) = \binom{8}{x}\binom{2}{5-x}/\binom{10}{5}$, $x = 3,4,5$;

(b) $p(x) = \binom{2}{x}\binom{8}{5-x}/\binom{10}{5}$, $x = 0,1,2$

4. $\frac{22}{35}$.

5. 0.9593 (Aproximación binomial 0.9672).

7. (a) 77/115 ; (b) 3/25 .

9. (a) $\dfrac{\binom{4}{x}\binom{96}{5-x}}{\binom{100}{5}}$, N = 100 , M = 4 , n = 5, E(X) = 0.16,

$$\text{Var}(X) = \frac{16}{100} \times \frac{96}{100} \times \frac{96}{99} \quad ; \quad 0.9997765$$

(b) $\dfrac{n}{N} = \dfrac{5}{100} = 0.05 < 0.1$ se aproxima a la binomial con

$n = 4$, $p = \dfrac{4}{100} = 0.04$; 0.9998

11. (a) 0.9929 ;

12. (a) $P[X = x] = \dfrac{\binom{10}{x}\binom{15}{4-x}}{\binom{25}{4}}$, $x = 0,1,2,3,4$.

(b) es el modelo discreto hipergeométrico

(c) 0.11 ; (d) 0.02 ; (e) 1.6 ; 0.91

13. $(\frac{4}{5})^6$. $n = 6$, $\dfrac{n}{N} = \dfrac{6}{100} = 0.06 < 0.1$, se aproxima por una binomial, con $n = 6$ y $p = \dfrac{80}{100} = \dfrac{4}{5}$.

PROBLEMAS 5.6

1. E(X) = 1 . 2. E(X) = $\ln 2$. 3. (a) $e^{-4/3}$. 4. (a) $\lambda = 4$,

(b) e^{-12} .

5. $\alpha = 2$, P(X = 1 ó 2) = $4e^{-2}$.

6. $e^{-5} = 0.0067$.

7. (a) $\alpha = 6$, al año ; (b) $\sqrt{6}$, (c) $e^{-0.5} = 0.60653$.

9. (a) $np = 10$; (b) $P(X = x \mid B : 100{,}000,\ 0.0001) \approx \dfrac{10^x e^{-10}}{x!}$;

$$x = 0,1,\ldots,100{,}000$$

10. (a) $np = 5$; (b) $P(X = 10) \approx \dfrac{5^{10}e^{-5}}{10!}$. 11. $e^{-86/117}$

12. (a) $\binom{10}{2}(0.607)^2(0.393)^8 = 0.00943$; (b) 607 páginas ,

(c) 304 páginas.

14. $e^{0.005} = 0.995$. 15. 232 disparos .

20. 0.368 ; 0.264 .

22. (a) binomial de parámetros $n = 5 \times 10^4$ y $p = 10^{-4}$,

(b) aproximación de poisson con $np = 5$, $P(X = 0) = e^{-5}$

23. (a) Hipergeométrica $\binom{500}{x}\binom{9500}{100-x} / \binom{10000}{100}$

(b) use aproximación de poisson (porque?) ; 0.1403

24. $e^{-5} = 0.00674$; 7 panetones.

25. 1060 frutillas aproximadamente. 26. 0.3851 . 27. 0.48467 .

28. Se debe tomar una muestra de 1.535 cm^3 o más .

29. (a) $\frac{7^7}{7!} e^{-7} = 0.130$, (b) $e^{-7} / \frac{7^7}{7!} e^{-7} = \frac{7!}{7^7} = 0.006$.

30. 0.7860 .

31. (a) 0.1494 , (b) 0.0498 , (c) 0.4232 .

32. (a) Es una binomial con parámetros $n = 50$, $p = 0.01$; (b) 0.0126 ; (c) 0.0018 ;
(d) 0.9856 .

33. (a) binomial, con $n = 300$ y $p = \frac{1}{2n}$, (b) $(5999/6000)^{300}$ $e^{-1/2} = 0.607$

(c) $(0.607)^{16} \approx 0.0003$.

34. $e^{-4} = 0.0183$. 35. $4e^{-3} = 0.199$

36. (a) $e^{-n/400}$; (b) $400e^{-n/400}$; (c) $n = 832$.

37. \$269

38. $e^{-1}(1 + 3e^{-2}) \approx 0.517$.

39. (a) binomial con parámetros $n = 20$ y $p = 0.02$; (b) $E(X) = 0.4$.

(c) $P[X \geq 1] = 1 - P[X = 0] = 1 - (0.98)^{20}$

(d) Puesto que $np = 0.4 < 5$, aproximamos la binomial a la distribución de poisson con $\lambda = np = 0.4$.

$P[X \geq 1] = 1 - P[X = 0] = 1 - e^{-0.4}$.

PROBLEMAS 6.1

1. $\frac{1}{4}$; 2. $\frac{1}{4}$, $\frac{1}{2}$, 1 . 3. (a) 2, (b) 3 , (c) $\frac{5}{4}$.

4. $-\sqrt{3}$, $\sqrt{3}$. 5. $\frac{3}{5}$. 6. (a) $F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{30} , & 0 < x < 30 \\ 1 , & x \geq 30 \end{cases}$, (b) $\frac{1}{2}$, (c) $\frac{2}{3}$.

7. $P(X < \frac{1}{3}) = \frac{2}{3}$. 8. (a) $P(X < 0.04) + P(0.16 < X < 0.2) = 0.4$

(b) $P(0.05 < X < 0.15) = 0.5$.

9. $0.2C_1 + 0.4C_2 + 0.4C_3 - C$. 10. (a) $\frac{2}{5}$; (b) $\frac{3}{5}$, (c) 7.5

PROBLEMAS 6.2

5. (a) $e^{-0.5} - e^{1.5} = 0.3834$; (b) $e^{-2} = 0.1353$; (c) $1 - (1 - e^{-2})^3 = 0.35$.

6. \$ 1.30 .

9. $\sum_{x=4}^{6} \binom{6}{x} [1 - e^{-3/4}]^x \, [e^{-3/4}]^{6-x} = 0.3968$.

10. $P(X \geq 2) = 0.2627$. 11. 0.2865 ; 0.5654 .

12. (a) 0.08208 ; (b) 0.7576 .

13. (a) 0.0498 ; (b) 2 .

15. (a) 0.7866 ; (b) 2.63 años . 16. 0.3679 .

PROBLEMAS 6.3

2. $c = 12$. 3. 0.1336 .

4. (a) 0.0082 ; (b) 0.8904 . 6. $c = 2.12$.

7. $(\mu , \sigma^2) = (8,4)$ o $(\mu , \sigma^2) = (2,64)$.

8. 8.08 % . 9. 68.26 % . 10. 10.5 puntos .

13. $\sigma = 549$ horas . 14. $E(X) = 455$. 16. 0.1151 .

17. 0.2302 .

18. (a) 0.8164 ; (b) 2585 ; (c) 59.72 seg.

20. 0.34545 .

23. (a) 524.64 ; (b) 510.88 .

24. (a) 6.55 % ; (b) 307.5 ; (c) 307.7 .
26. 18 días .
28. $(0.9772)^5$. 29. 0.06356 . 31. σ = 87.463 .
32. Será preferible usar tubos de tipo B, pues $P[X_A \geq 35] < P[X_B \geq 35]$.
33. (a) usamos el procedimiento A, pues $P[X_A < 30] > P[X_B < 30]$; (b) se elige el - proced. B . 34. (a) 16 ; (b) 551 ; (c) 27 ; (d) 29 .
35. 62 . 37. 0.2514 . 38. 0.9050 .
39. 6.3 % . 40. si, 97.72% tienen durabilidad superior a 400 horas
41. (a) 334; (b) 2881 ; (c) 749 .
42. (a) 5.0% ; (b) 2.4% ; (c) 7.3 % .
43. Costo promedio por varilla utilizable = \$ 13.36 . 44. \$ 23.40 .
45. (a) distribución normal, μ = 288 onzas, σ = 0.4898 onzas; (b) 0.5 .
46. (a) 0.0207 ; (b) 0.4443 .
49. (a) μ = 14, σ^2 = 25; (b) 0.2881 . 50. 0.6768 .
51. 10.56 % . 52. 5.48 % .
53. (a) 0.9981 ; (b) 0.0003 . 54. 58.4 y 97.6 kg.; 0.0227.
55. 0.0359 . 56. 0.0307 . 57. (a) 0.1802 ;
58. 0.1920 . 59. 0.0793 . 60. 16 . 61. 0.1587 .

PROBLEMAS 6.4

1. (a) 0.7852 ; (b) 0.0853 ; (c) 0.6779 .
2. 0.0150 .
3. (a) 0.00 ; (b) 0.9328 .
4. (a) Tiene una distribución binomial con n = 72 y $p = \frac{1}{3}$; los parámetros son - μ = 48 , σ^2 = 16 (b) 0.3520
5. 0.07221 . 6. (a) 0.2327 ; (b) 0.0179 .
7. (a) 0.3085 ; (b) 0.7454 .
8. (a) 0.5548 ; (b) 0.3823 ; (c) 0.5802 .
9. 0.3300 . 10. (a) 0.1056 ; (b) 0.1251 .
11. 0.9386 . 12. (a) 0.0268 , ; (b) 0.0268 ; (c) 0.9616 .
13. (a) 6 (b) 125 .
14. (a) 6 , (b) 9 , (c) 2 , (d) 12

17. (a) 0.9793 ; 18. (a) 0.9662 ; (b) 24 .

19. $n(\frac{M}{N}) = 30(\frac{500}{600}) = 25 > 5$, $n = 30 = (0.05)600 = 0.05N$.

X = número de focos buenos en la muestra de 30. Se aproxima por la normal.

$P[X = 30] = 0.0101$.

CAPITULO 7

PROBLEMAS 7.1

1. 0.8051 . 2. (a) 0.0228 ; (b) 0.0668 .
3. (a) 14/3 ; (b) 10/9 . 4. $\mu = 1$, $\sigma^2 = \frac{1}{10}$
5. $p^{10}q^{\sum_{i=1}^{10} x_i - 10}$ 6. $f(x_1, \ldots, x_{10}) = 1$, $0 < x_i < 1$, $i = 1,2,\ldots,10$
8. (a) 0.7948 ; (b) 0.2052 .

PROBLEMAS 7.2

1. (a) 5 , $\sqrt{5}$; (b) {(2,4), (2,6), (2,8), (4,6), (4,8), (6,8)} ;
   (c) 5 , $\sqrt{5/3}$.
2. $\mu = 4.5$. $\sigma = 2.872281$; (a) $\mu_{\bar{X}} = 4.5$, $\sigma_{\bar{X}} = 2.03101$;
   (b) $\mu_{\bar{X}} = 4.5$, $\sigma_{\bar{X}} = 1.914854$
3. (a) 4, $\sqrt{8/3}$; (b) 4, 2/9 ; (c) 0.2905 .
4. $\mu = 6$, $\sigma^2 = 4$;

(a)

$\bar{x}$	5.00	7.50	10.00
$p(\bar{x})$	0.64	0.32	0.04

, (b) $E(\bar{X}) = 6$, $var(\bar{X}) = 2$. Se cumple que $E(\bar{X}) = E(X)$ y $Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{4}{2} = 2$.

(c)

$\bar{x}$	5.00	7.50	10.00
$p(\bar{x})$	56/90	32/90	2/90

(d) $E(\bar{X}) = 6$, $Var(\bar{X}) = 16/9$. Se verifica que $E(\bar{X}) = 6$ y

$$Var(\bar{X}) = \frac{\sigma^2}{n}\,\frac{N-n}{N-1} = \frac{4}{2}\,\frac{10-2}{10-1}$$

5. 0.6818 . 6. (a) 0.7066 ; (b) 0.3802 .
7. 0.4772 .
8. (a) $\mu_{\bar{X}} = 6.25$, $\sigma_{\bar{X}} = 0.617$; (b) 0.0212 .
9. (a) 6.726 ; (b) No , la población tiene distribución normal.

10. (a) 110.5 , (b) 0.9876 ; (c) Debe cumplirse el teorema central del límite.

11. (a) 0.0228 ; (b) 0.3707 ; (c) 0.9082 ; (d) a las 2. p.m.

12. (a) 0.7734 ; (b) entre 4804 y 5196 inclusive.

13. 0.6950 . 14. (a) 0.0228 ; (b) 0.1587 ; (c) 0.50

15. $t = 407$. 16. (a) 0.6892 ; (b) 0.2881

17. Cero 18. (a) Entre 1.99804 y 2.00196 ; (b) 0.0668

21. (a) 0.0392 ; (b) 0.8106 ; (c) 0.1502

22. (a) 0.2164 ; (b) 0.0015

23. 0.0258

24. (a) 0.8621 ; (b) 0.8621 ; (c) 0.03 (d) 0.037 .

25. si, pués $P(\bar{X} < 0.5) = P(z < -6) = 0$

26. (a) 1 ; (b) 15 27. $n = 41$. 28. $n = 68$

29. $n = 16$. 30. $n = 136$.

31. $\mu_{\bar{X}_1 - \bar{X}_2} = \frac{7}{3}$, $\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{25}{9}$ 32. 0.7064 33. (a) 0.9582 ;

34. $P(\bar{X}_E - \bar{X}_V < 0) = 0$. 35. $1.96\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$

36. (a) 0.0802 ; (b) 0.6212 37. 0.0188 .

38. $n = 13$; $n = 14$. 39. $n = 7$.

40. (a)

$\bar{P}$	0/5 = 0.0	1/5=0.2	2/5 = 0.4	3/5 = 0.6	4/5 = 0.8	5/5 = 1.0
$p(\bar{P})$	0.5905	0.3280	0.0729	0.0081	0.0004	0.0000

(b) $E(\bar{P}) = 0.1 = p$, $Var(\bar{P}) = \frac{0.09}{5} = 0.0180 = \frac{p(1-p)}{5}$

(c) Desde que $p = 0.1 < 0.5$, $\bar{p}$ es sesgado a la derecha ó sesgado positivamente

41. $E(\bar{P}) = 0.5$ y $\sigma_{\bar{P}} = 0.13608$. 42. (a) $\mu_{\bar{P}} = 1/3$; (b) 0.8664

43. (a) 0.4854 , (b) entre 21.02 % y 38.98 %

44. (a) 0 , (b) 0.9938 .

45. (a) 6 ; (b)

$\bar{P}$	0	0.5	1
$p(\bar{P})$	0.1	0.6	0.3

(c) (1) 0.3 ; (2) 0.1 ; (d) 0.6

(e) (1) 0.6 ; (2) 0.3 .

46. 0.6 .   47. (a)

$\bar{P}$	0	1/7	2/7	3/7	4/7	5/7	6/7	1
$p(\bar{P})$	0.152	0.0872	0.2140	0.2918	0.2388	0.1172	0.0320	0.0037

(b) (1) 0.2388 , (2) 0.3164 ; (c) $\mu_{\bar{P}} = 0.45$, $\sigma_{\bar{P}} = 0.188$

48. (a) $\mu_{\bar{P}} = 0.15$; $\sigma_{\bar{P}} = 0.0357$; (b) 0.8384 , 0.7187 ;
(c) 0.091 a 0.209 ; (d) 0.121 a 0.179 ; (e) 0.5 .

49. n = 469 .   50. n = 900

PROBLEMAS 7.3

1. (c) 13.28 ; 3. (a) 0.6 aprox. ; (b) 0.72 aprox. ; (c) 0.32 aprox.

4. 0.94 ; 5. 0.95 ; 6. 0.25

9. (c) 1.714 ; 10. (a) 0.05 ; (b) 0.94

11. (a) 0.15 aprox. , (b) 0.985 aprox. (c) 0.99 aprox.

15. (a) 0.1 aprox.

17. 0.94 .   18. 0.025 .

20. (a) 0.36 ,   (b) 0.49

21. (a) es menor que 0.01 ,(b) entre 0.01 y 0.05 pero mas cerca a 0.01 .

CAPITULO 8

PROBLEMAS 8.2

1. 0.75 ; 5. (a) $157.65 < \mu < 162.35$;
(b) $158.03 < \mu < 162.35$; (c) 284 observaciones adicionales

6. $67.2 < \mu < 72.8$ puntos ; 7. $20.02 < \mu < 21.98$

8. (a) 138 ; (b) $40.3 < \mu < 43.7$

9. 38 10. (a) 1537 ; (b) 385 ; 11. 49

12. (a) 6,625 fibras o menos ; (b) 140.

14. entre 9.58 , 452.00 y S/. 61,548,000

15. entre S/. 5,226,000 y S/. 6,774,000

16. 1,616,812.5 y 1,883, 187.5

19. (a) $13 < \mu_1 - \mu_2 < 17$; (b) $12 < \mu_1 - \mu_2 < 18$

20. $4.6 < \mu_1 - \mu_2 < 15.4$

24. $n = 208$ 25. $0.49 < p < 0.642$; 26. $0.1442 < p < 0.1998$

30. (a) $0.19 < p_1 - p_2 < 0.31$; (b) $0.17 < p_1 - p_2 < 0.33$

31. $-0.0054 < p_1 - p_2 < 0.2054$

34. (a) $n = 266{,}256$; (b) 10,544 · 37. $0.37 < p < 0.91$ · 38. 384 ·

40. 0.8904 ; 41. $n = 97$ 42. (a) 0.9544

(b) varias respuestas posibles

43. $279 < \mu < 291$; 44. 90 %

47. $9.81 < \mu < 10.31$; 48. $15.63 < \mu < 21.57$

55. (a) $\bar{X}_1 = 85.2$, $\bar{X}_2 = 88.6$; $S^2 = 33.75$

$T = 1.8595$, $-9.23 < \mu_1 - \mu_2 < 3.43$

(b) Se supone que las poblaciones de las que se extraen las muestras son aproximadamente normales

(c) Se supone que las distribuciones de las medias muestrales son aproximadamente del tipo normal, con varianzas iguales .

56. $S^2 = 11.6$; $4.521 \leq \sigma^2 \leq 69.795$

CAPITULO 9

PROBLEMAS 9.2

1. $\alpha = 0.0146$; $\beta = 0.0126$. 2. $\alpha = 0.0386$; $\beta = 0.2779$; $\beta = 0.5$

3. Si ; $z = 4.67$. 4. $z = 6.6$, $66 > 2.33$, rechazar H_o

5. $z = 3.3$; $3.3 > 2.58$; rechazar H_o

7. Si ; 8. Si ; 9. No 10. Si

11. No rechace H_o

14. $t = -3.47$; $-3.47 < -2.896$; rechazar H_o

15. aceptar H_o ; 17. No ; 23. Si ; 24. Si

25. No , (si, cuando $\alpha = 0.1$) ; 26. Si

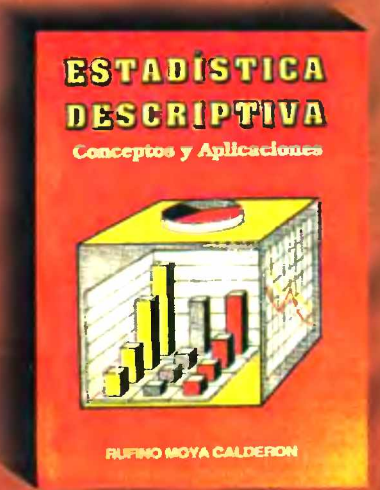
27. No ; 28. $z = -0.824 < -2.83$; no se rechaza

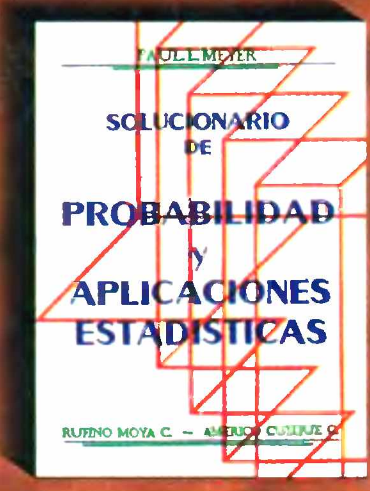
29. No se rechaza H_o ; 31. No rechace H_o

Impreso en los Talleres Gráficos de
Editorial "San Marcos"
RUC 11029221

PEDIDOS:

Av. Garcilaso de la Vega 911-Of. 404 - Lima
Teléf. 424-6563
Jr Natalio Sánchez 220-Of. 304 - Jesús María
(Alt. Cdra. 5 de Av. Arenales)
Teléf. 330-8553 / Teléfax 433-7542





EDITORIAL SAN MARCOS

Jr. Natalio Sánchez 220 - Of. 304. Jesús María
(Alt. Cdra. 5 de Av. Arenales) - Telf.: 330-8553 Telefax: 433-7542
Av. Garcilaso de la Vega 911 - Of. 404 - Lima. Teléfono: 424-6563