REVISIÓN DEL CAPÍTULO

▪ TÉRMINOS IMPORTANTES ▪

Los términos utilizados en el capítulo se han mezclado para proporcionarle una mejor práctica de revisión. Proporcione una definición de cada uno con sus propias palabras, después compare sus respuestas con las dadas en el capítulo.

método clásico de probabilidad
tabla de probabilidad
eventos compuestos
variable aleatoria
probabilidad condicional
frecuencia relativa
variable aleatoria continua
espacio muestral
variable aleatoria discreta
eventos simples
probabilidad empírica
subconjunto
conjunto vacío
regla de la suma
resultados igualmente probables
diagrama de árbol
evento
varianza de una variable aleatoria
valor esperado de una variable
permutación
combinación
experimento
resultados favorables
diagrama de Venn
eventos independientes
selección con remplazo
media de una variable aleatoria
selección sin remplazo
selección con remplazo
regla de la multiplicación
n factorial
regla de la permutación
simulación por computadora
regla de multiplicación
método subjetivo
método objetivo
desviación estándar de una variable
eventos dependientes
función de probabilidad
histograma de probabilidad
regla del producto
triángulo de Pascal
eventos mutuamente excluyentes
probabilidad
oportunidades

▪ SÍMBOLOS IMPORTANTES ▪

S, espacio muestral

$\overline{E}$, el evento no E

$(E \cup F)$, evento E o F

$(E \cap F)$, evento E y F

$P(E)$, la probabilidad del evento E

$a{:}b$, las oportunidades son como a a b

P^n_n, el número de permutaciones de n objetos tomados de r en r

$n!$, n factorial

P^n_r, el número de permutaciones de n objetos tomados de r en n

$\binom{n}{r}$, coeficiente binomial, el número de combinaciones de n objetos tomados de r en r

$P(E \mid F)$, probabilidad condicional

μ_x', media de una variable aleatoria X

σ_x^2, varianza de una variable aleatoria X

σ_x, desviación estándar de una variable aleatoria X

$E(X)$, valor esperado de X

▪ FÓRMULAS Y HECHOS IMPORTANTES ▪

Si S es un espacio muestral de resultados igualmente probables y E es un evento, entonces $P(E) = f/n$ donde f es el número de resultados contenidos en E y n es el número de resultados en S. (5.1)

Oportunidades a favor de E: $\text{Ops}(E) = \dfrac{P(E)}{P(\overline{E})}$

Oportunidades en contra de E: $\text{Ops}(\overline{E}) = \dfrac{P(\overline{E})}{P(E)} = \dfrac{1}{\text{Ops}(E)}$

Si $\text{Ops}(E) = a{:}b$, entonces $P(E) = \dfrac{a}{a+b}$

Teorema fundamental del conteo. Si un evento puede ocurrir de m formas distintas y, si después de que ha sucedido, otro evento puede ocurrir en n formas distintas, entonces los dos pueden suceder, uno después del otro, de $m \cdot n$ formas distintas.

n factorial: $n! = n(n-1)(n-2)(n-3)\ldots(2)(1)$.

El número de permutaciones de n objetos tomados de n en n: $P^n_n = n!$ (5.3)

El número de permutaciones de n objetos tomados de r en r:

$$P^n_r = \frac{n!}{(n-r)!}. \quad (5.4)$$

Coeficiente binomial: $\binom{n}{r} = \frac{P^n_r}{r!}$. (5.5)

El número de combinaciones de n objetos tomados de r en r $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ (5.6)

Regla de la suma: $P(E \cup F) = P(E) + P(F) - P(E \cap F)$. (5.7)

Si E y F son eventos mutuamente excluyentes,

$P(E \cup F) = P(E) + P(F)$. (5.8)

Probabilidad de no E: $P(\overline{E}) = 1 - P(E)$. (5.9)

Regla de probabilidad condicional:

$P(E \mid F) = P(E \cap F)/P(F)$, si $P(F) \neq 0$. (5.10)

Regla del producto: $P(E \cap F) = P(F)\, P(E|F)$. (5.12)

E y F son eventos independientes si $P(E|F) = P(E)$. (5.14)

Regla de la multiplicación para eventos independientes:

$P(E \cap F) = P(E)P(F)$. (5.15)

Media de una variable aleatoria X:
$\mu_X = E(X) = \Sigma(xP(x))$. (5.16)

Varianza de una variable aleatoria
$\sigma^2_X = \Sigma[(x-\mu)^2 P(x)]$. (5.17)

Desviación estándar de una variable aleatoria X:
$\sigma_X = \sqrt{\text{varianza}}$

Fórmula para calcular la varianza de una variable aleatoria X:

$\sigma^2_x = \Sigma[x^2 P(x)] - \mu^2_x$. (5.18)

EJERCICIOS DE REPASO

1. Los registros de un hospital particular indican que el 18% de sus pacientes ingresan a cirugía, el 30% a obstetricia y el 5% tanto a una como a otra especialidad.
a) ¿Cual es la probabilidad de que un paciente elegido al azar ingrese ya sea en cirugía, en obstetricia o en ambos?
b) ¿Cuál es la probabilidad de que un paciente elegido al azar no ingrese en cirugía?
c) ¿Qué probabilidad hay de que un enfermo elegido al azar sea un paciente de cirugía que no requiere tratamiento obstétrico?

2. Suponga que E y F son dos eventos con $P(E) = 1/2$ y $P(F) = 1/3$
a) Si E y F son eventos mutuamente excluyentes, encuentre $P(E \cup F)$.
b) Si E y F son independientes, encuentre $P(E \mid F)$, $P(F \mid E)$, $P(E \cap F)$, $P(E \cup F)$.
c) ¿Puede ser cierto $P(E \cup F) = 11/12$? Explique
d) ¿Puede resultar $P(E \cup F) = 1/6$? ¿Por qué?
e) ¿Puede suceder $P(E \cap F) = 0.6$? Explique.

3. Se encuestó a setenta y cinco estudiantes pidiéndoles mencionar su bebida favorita. Sus respuestas fueron:

cerveza	13	vino	2
refresco	40	agua	7
té helado	4	té caliente	1
café	7	whisky	1

Se elige al azar a uno de los encuestados.
a) Encuentre la probabilidad de que el estudiante en cuestión mencione cerveza o vino.
b) Calcule la probabilidad de que responda café.

4. Los empleados de una empresa particular se clasificaron según su sexo y su afiliación política. Los resultados aparecen a continuación:

	Afiliación política		
Sexo	Demócrata	Republicano	Independiente
Hombre	40	50	5
Mujer	18	8	4

Si se escoge a un empleado al azar, encuentre la probabilidad de que sea:

a) hombre.
b) republicano.
c) mujer y demócrata.
d) republicana dado que es mujer.
e) hombre cuando es republicano.

5. Si se lanza un dado, denotemos por X la variable aleatoria correspondiente al número de arriba.
a) Construya una tabla de probabilidad para X.
b) Trace una gráfica de probabilidad para X.
c) Encuentre $E(X)$ e interprete sus resultados.
d) Encuentre σ_X.

6. Una persona lanza tres monedas; si obtiene tres "soles" o tres águilas, gana 10 dólares; si no obtiene tres águilas o tres "soles", paga 10 dólares. ¿Cuál es la ganancia esperada?

7. El señor Pérez puede obtener 5,000 dólares con una probabilidad de 0.4 o perder 2,000, con una probabilidad de 0.6, si invierte en acciones de la Ajax Company. ¿Cuál es su ganancia esperada? ¿Cuál la desviación estándar de su ganancia?

8. Si se lanzan tres dados, denotemos por X la variable aleatoria correspondiente a los números dos obtenidos.
a) Construya una tabla de probabilidad para X.
b) Dibuje una gráfica de probabilidad para X.
c) Encuentre $E(X)$.
d) Calcule σ_X.

9. Una moneda está adulterada en forma tal que los "soles" tienen una frecuencia del doble de las águilas cuando se lanza la moneda; representemos por X la variable aleatoria correspondiente al número de "soles" obtenidos.
a) Construya una tabla de probabilidad para X.
b) Encuentre μ_x y σ_x.
c) Suponga que la variable aleatoria Y denota el número de águilas obtenidas. Encuentre μ_y y σ_y.

10. Las acciones del Ace Bank suelen venderse a 10 dólares el paquete; un inversionista planea comprar de esas acciones y conservarlas durante un año. Denotemos por X el precio del paquete después de un año. La tabla de probabilidad para X se muestra aquí.

x	$P(x)$
10	0.35
11	0.25
12	0.20
13	0.15
14	0.05

a) ¿Cuál es el precio esperado del paquete al transcurrir un año?
b) ¿Cuál es la ganancia esperada por paquete de aciones para el periodo de un año?
c) ¿Qué porcentaje de la inversión corresponde a la ganancia esperada por las acciones
d) Encuentre la varianza en el precio del paquete para el periodo de un año.
e) Encuentre la varianza en la ganancia por paquete de acciones en un año.
f) Construya una gráfica para la función de probabilidad.

Aplicaciones de computación

1. Use un programa computacional para simular el problema del cumpleaños: estime la probabilidad de que entre 25 personas elegidas al azar, al menos dos cumplan años el mismo día. Use órdenes análogas a

```
RANDOM 25 C1;
INTEGERS 1 365
SORT C1 C2
PRINT C2
```

Repita este procedimiento diecinueve veces más y estime la probabilidad de que dos personas tengan el mismo compleaños. (*Sugerencia:* con MINITAB use la orden STORE para crear un archivo que pueda ejecutarse 20 veces.)

2. Use simulación por computadora para hacer el ejercicio 21 del grupo de ejercicios 5.2; tire un par de dados 100 veces y anote la suma de los números mostrados en cada tirada; determine la frecuencia relativa de una suma de 8. (Si los dados son de uso común, la probabilidad de tirar una suma de 9 es, aproximadamente, 0.14.)

3. Simule en la computadora para hacer el ejercicio 23 del grupo de ejercicios 5.2; si un paquete de

52 cartas se baraja bien, ¿cuál es la probabilidad estimada por usted de que las tres cartas de arriba contengan un rey o una reina, o de que un rey y una reina estén juntos en algún lugar de las 49 cartas restantes? Verifique si adivinó simulando el experimento 20 o más veces.

▪ *EXAMEN DE CONOCIMIENTOS DEL CAPÍTULO* ▪

1. Un restaurante ofrece seis tipos de emparedados, diez clases de bebidas, cuatro variedades de sopas y tres postres en su menú. ¿Cuántos almuerzos distintos se puede ordenar si se desea un tipo de emparedado, una bebida, una sopa y un postre?

2. Se realizó un estudio sobre afiliación religiosa y partido político, obteniéndose los resultados siguientes:

	Religión		
Partido político	Protestante	Católico	Judío
Demócrata	10	15	25
Republicano	20	30	40
Independiente	5	15	5

Se elige una persona al azar del grupo en cuestión; D, R e I denotan demócrata, republicano e independiente respectivamente, mientras que P, C y J significan protestante, católico y judío, respectivamente. Encuentre:

a) $P(\mathrm{R})$.
b) $P(\mathrm{J})$.
c) $P(\mathrm{R} \cap \mathrm{J})$.
d) $P(\mathrm{R} \cup \mathrm{J})$.
e) $P(\mathrm{R} \mid \mathrm{J})$.
f) $P(\mathrm{C} \mid \mathrm{D})$.
g) ¿Son eventos independientes J y R? Explique.
h) Y si son eventos mutuamente excluyentes, ¿por qué?

3. Se ofrecen dos puestos de profesor y dos hombres y tres mujeres presentan su solicitud; como son todos los aspirantes y están igualmente calificados, se escogen dos al azar para ocupar los empleos.

a) Enliste un espacio muestral para el experimento.
b) ¿De cuántas formas puede escogerse a dos hombres?
c) ¿De cuántas formas pueden seleccionarse dos mujeres?
d) Encuentre la probabilidad de que se emplee a dos mujeres.
e) Calcule la probabilidad de que se emplee a un hombre y a una mujer.

4. La tabla siguiente registra las muertes anuales de pacientes masculinos cancerosos de más de 65 años para los cinco tipos principales de cáncer:

Localización	Núm. de muertes
Colon	8,000
Pulmón	12,500
Páncreas	3,000
Próstata	10,800
Estómago	3,200

Si se elige al azar un paciente muerto por cáncer, ¿cuál es la probabilidad de que haya muerto por:

a) cáncer pulmonar?
b) alguna de las otras causas principales?
c) cáncer de colon o de páncreas?
d) un cáncer estomacal?

5. Usted paga 0.50 dólares por jugar a una ruleta de 38 posiciones espaciadas; si la rueda se detiene en la posición elegida por usted, gana 5 dólares; de otro modo, usted pierde. Si juega una vez, encuentre la ganancia esperada.

6. Si se eligen dos cartas con reemplazo de una paquete de baraja común, denotemos por X la variable aleatoria número de corazones. Encuentre μ_x y σ_x.

6 Distribuciones discretas

DESCRIPCIÓN

OBJETIVOS DEL CAPÍTULO

En este capítulo estudiaremos:

- *Experimentos binomiales.*
- *El lenguaje asociado con experimentos binomiales.*
- *Cómo calcular coeficientes binomiales.*
- *Cómo calcular probabilidades binomiales.*
- *Cómo usar tablas binomiales para calcular probabilidades binomiales.*
- *Cómo calcular la media de una variable aleatoria binomial.*
- *Cómo calcular la varianza y la desviación estándar de una variable aleatoria binomial.*
- *Cómo construir gráficas de distribuciones de probabilidad para variables aleatorias binomiales.*
- *Experimentos multinomiales.*
- *Cómo calcular coeficientes multinomiales.*
- *Cómo calcular probabilidades multinomiales.*
- *Experimentos hipergeométricos.*
- *Cómo calcular probabilidades hipergeométricas.*
- *Experimentos de Poisson.*
- *Cómo calcular probabilidades de Poisson.*

MOTIVADOR 6

Algunas líneas aéreas acostumbran vender más boletos que el número de asientos disponibles; la razón de esto es que un cierto porcentaje de quienes hacen reservaciones no se presentan a tiempo para el vuelo. Suponga que los planes de transportación de una aerolínea son acomodar 35 pasajeros y se ha estimado que, en promedio, un 5% de quienes hicieron reservaciones no se presentarán a tiempo; con base en esto, la aerolínea vendió 37 boletos para un cierto vuelo, ¿cuál es la probabilidad de que se presenten los 37 pasajeros?, ¿y la de que todos los que se presenten para el vuelo encuentren lugar? Después de estudiar la sección 6.2, usted será capaz de responder estas dos preguntas.

Panorama del capítulo

Las distribuciones binomiales forman una clase importante de distribuciones discretas en estadística; se usan para describir una amplia variedad de procesos de muchas formas, y resultan de la repetición de experimentos binomiales; en este capítulo examinaremos las propiedades y algunas de las aplicaciones de los experimentos binomiales. Los experimentos multinomiales e hipergeométricos están relacionados con los experimentos binomiales porque permiten aplicar, sin mucho rigor, ciertas propiedades de éstos y, por tanto, adaptarlos a una gran variedad de situaciones. Finalmente, exploraremos los experimentos de Poisson como otra variedad de los experimentos binomiales; estos cuatro tipos de experimentos dan lugar a variables aleatorias discretas que tienen distribuciones discretas de probabilidad. Aprenderemos a calcular las probabilidades asociadas con cada tipo de variable aleatoria.

SECCIÓN 6.1 Distribuciones binomiales

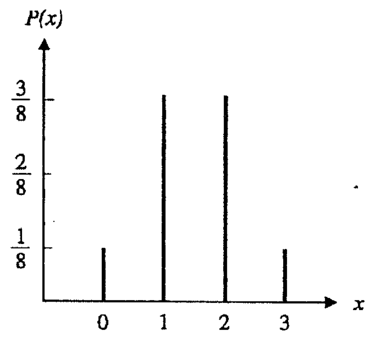
Considere el experimento de lanzar tres veces una moneda y observar el número de "soles" que resulten; cada lanzamiento se denomina un intento. Este experimento posee las características siguientes:

1. Consiste de tres intentos idénticos de lanzar una moneda.
2. Cada intento da lugar exactamente a uno de dos resultados.
3. Los intentos son independientes; el resultado de un intento no afecta el de otro.
4. La probabilidad de obtener un "sol" permanece constante de un intento al otro, siendo esa probabilidad de 0.5.

Hay ocho posibilidades distintas, igualmente probables, que pueden resultar de lanzar una moneda tres veces. Se enlistan como sigue:

SSA ASA SAS AAS SSS ASS SAA AAA

FIGURA 6.1
Distribución de probabilidad para X = número de "soles" en tres lanzamientos de una moneda



Si la variable aleatoria discreta X denota el número de "soles" obtenidos, entonces puede usarse la siguiente tabla de distribución de probabilidades para organizar los resultados.

x	$P(x)$
0	1/8
1	3/8
2	3/8
3	1/8

Esta distribución de probabilidad tiene la gráfica mostrada en la figura 6.1. Note que la gráfica es simétrica respecto a $x = 1.5$ "soles". El experimento que acabamos de describir es ejemplo de una clase llamada experimentos binomiales.

En general, un **experimento binomial** es aquel que posee las cuatro propiedades siguientes:

Propiedades de un experimento binomial

1. El experimento consiste de n intentos idénticos.
2. Cada intento da lugar a exactamente dos resultados, llamados éxito o fracaso. Un éxito se denota por E y un fracaso por F.
3. Los n intentos son independientes.
4. La probabilidad p de un éxito permanece constante de un intento al otro.

La distribución de probabilidad para el número de éxitos se denomina **distribución binomial**.

EJEMPLO 6.1

Cada una de las situaciones siguientes puede modelarse usando un experimento binomial.

- Administrar una medicina para el resfriado que tiene un promedio de cura de 0.90 a diez personas resfriadas, y observar cuántas personas mejoran
- Bailar una perinola ocho veces y cuantificar el número de veces que se detiene con el vástago hacia arriba
- Responder adivinando un examen de verdadero-falso y contar el número de respuestas correctas
- Lanzar un dado seis veces observando cuántas veces resulta el número tres
- Al observar a un beisbolista con un promedio de bateo de 0.400 y al contar la cantidad de hits que consigue en los tres siguientes turnos al bat.

Los símbolos siguientes se usan en la descripción de experimentos binomiales:

$$\begin{aligned} E &= \text{éxito} \\ F &= \text{fracaso} \\ p &= P\ (\text{éxito}) \\ 1-p &= P\ (\text{falla}) \\ n &= \text{número de intentos} \\ x &= \text{número de éxitos} \end{aligned}$$

Nota: Para un experimento binomial con n intentos, x puede tomar uno de los $(n + 1)$ valores $0, 1, 2, 3, \ldots, n$. Considere las aplicaciones 6.1 y 6.2; aunque éxito es el término usado para describir el resultado de interés para un experimento binomial, esto no necesariamente corresponde a un "buen" evento, como lo muestra la aplicación 6.3.

APLICACIÓN 6.1

Una fertilización cruzada de especies afines de plantas, con flores blancas unas y azules otras, produce vástagos de los cuales 20% tienen flores blancas. Se cruzaron seis plantas de flores azules con seis de flores blancas, y se encontró que había dos plantas de flores blancas entre sus vástagos. ¿Es este un experimento binomial? Si lo es, identifique en este contexto lo que es un intento, un éxito y los valores de p, n y x.

Solución: Este experimento es una aplicación de la teoría de Mendel sobre caracteres hereditarios. Un intento consiste en cruzar una planta de flores azules con una de flores blancas; un éxito es obtener un vástago de flores blancas y $p = P(E) = 0.20$. De acuerdo con la teoría de Mendel, los intentos son independientes; en consecuencia, el experimento es binomial con $n = 6$ y $x = 2$. ■

APLICACIÓN 6.2

María lanzó un dado diez veces para determinar el número de unos que resulta; obtuvo tres unos. ¿Es éste un experimento binomial? Si lo es, ¿qué constituye el intento, qué el éxito y cuál sería el fracaso? ¿cuáles son los valores de n, p y x?

Solución: Es un experimento binomial en el cual un intento consiste en lanzar un dado, un éxito en obtener un uno y un fracaso en lograr algún resultado distinto de un 1; la probabilidad de éxito es $p = 1/6$, mientras que la de fallar es 5/6; hay $n = 10$ intentos y x, el número de éxitos, es tres. Si en lugar de anotar el número de unos que resultaron, María hubiera registrado las veces que cada valor aparece en la cara superior del dado, el experimento no hubiera sido binomial porque cada intento habría dado como resultado alguno de seis números. ■

APLICACIÓN 6.3

Cierto tipo de medicamento no causa reacción en la piel en el 90% de las personas que lo usan. Estamos interesados en saber cuántas personas presentan reacción en la piel de las próximas cinco que lo usen; identifique un intento, un éxito, los valores de p y de n y los posibles valores de x.

Solución: Un intento es tratar a una persona con el medicamento, y un éxito que tenga reacción en la piel; por tanto, $p = 1 - 0.90 = 0.10$, $n = 5$ y x puede ser cualquiera de los seis valores 0, 1, 2, 3, 4, o 5. ■

La hipótesis de independencia de los intentos para un experimento binomial, implica que la probabilidad de éxito p permanece constante de intento en intento, sin embargo, el recíproco no es cierto. La probabilidad de éxito puede permanecer constante de intento en intento sin que éstos sean independientes; la hipótesis de independencia no se cumple en muchas situaciones, particularmente en aquellas donde el muestreo se hizo sin remplazo. Es preciso tener cuidado porque el tratamieno binomial no siempre es aconsejable.

EJEMPLO 6.2

Suponga que una ciudad tiene cinco restaurantes con licencia, dos de los cuales suelen cometer al menos una violación seria al código sanitario; hay dos inspectores, cada uno inspeccionará un restaurante la semana próxima; se escriben los nombres de los establecimientos en tiras distintas de papel que luego se mezclan; cada inspector toma al azar una de las tiras, sin reemplazar la primera antes de que se escoja la segunda. Un intento es inspeccionar un restaurante y el número de intentos es 2, un éxito E se interpreta como observar que no hay violaciones al código sanitario. Denotemos por

E_1 un éxito en el primer intento, por E_2 un éxito en el segundo intento y por F_1 un fracaso en primer intento, entonces la probabilidad de obtener un éxito en el primer intento es:

$$P(E_1) = \frac{3}{5} = 0.6$$

Hay dos formas de obtener un éxito en el segundo intento:

$$E_2 = (E_1 \cap E_2) \text{ o } (F_1 \cap E_2)$$

Si usamos la regla de la suma, tenemos:

$$P(E_2) = P(E_1 \cap E_2) + P(F_1 \cap E_2)$$

Mediante la regla del producto obtenemos:

$$\begin{aligned} P(E_2) &= P(E_1 \cap E_2) + P(F_1 \cap E_2) \\ &= \frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{3}{4} \\ &= \frac{3}{10} + \frac{3}{10} = 0.6 \end{aligned}$$

Sin embargo, los intentos no son independientes, porque

$$\begin{aligned} P(E_2 | E_1) &= \frac{P(E_1 \cap E_2)}{P(E_1)} \\ &= \frac{3/10}{6/10} = 0.5 \end{aligned}$$

y

$$P(E_2 | E_1) \neq P(E_2)$$

Coeficientes binomiales

Se ha determinado, con base en un periodo largo, que el 80% de las personas que hacen reservaciones aéreas llevan a cabo sus planes de vuelo, mientras que un 20% de los clientes potenciales no se presentan a tiempo para el vuelo. Si cuatro personas hacen reservaciones y estamos interesados en el número de pasajeros que llegan a tiempo para el vuelo, tenemos un ejemplo de experimento binomial con $n = 4$ y $p = 0.80$; el número de pasajeros puntuales es una variable aleatoria con valores $x = 0, 1, 2, 3$ y 4. Etiquetemos los resultados "llegó a tiempo" como un éxito (E), y "no llegó a tiempo" como un fracaso (F); si x denota el número de éxitos, el resultado puede ser uno de los siguientes 16:

$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$
FFFF	*EFFF*	*EEFF*	*EEEF*	*EEEE*
	FEFF	*EFEF*	*EEFE*	
	FFEF	*EFFE*	*EFEE*	
	FFFE	*FEEF*	*FEEE*	
		FEFE		
		FFEE		

En resumen, vemos que el número de formas en que puede obtenerse cada valor de x es como sigue:

Valor de x (número Es)	0	1	2	3	4
Número de resultado posibles	1	4	6	4	1

Para un valor dado de x, el número de resultados posibles que contienen x éxitos está dado por el **coeficiente binomial** $\binom{n}{x}$. Recuerde de la sección 5.3 que $\binom{n}{x}$ representa el número de combinaciones de n resultados que contienen x éxitos. Para nuestro ejemplo tenemos:

x	0	1	2	3	4
$\binom{4}{x}$	1	4	6	4	1

Vea que esto concuerde con nuestro resumen anterior.

En la sección 5.3, la fórmula para el coeficiente binomial $\binom{n}{x}$ estuvo dada por

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{6.1}$$

donde n es el número de intento, x el de éxitos y $(n-x)$ el de fracasos. Note que el número de éxitos más el número de fracasos es igual al número de intentos.

APLICACIÓN 6.4

Los árboles de un bosque están infestados por un parásito específico. Si se seleccionan al azar 15 árboles para estudiarlos, ¿cuántos resultados pueden obtenerse con:

a) 3 árboles infestados?
b) ningún árbol dañado?
c) 15 árboles con la plaga?
d) a lo más 2 árboles infestados?

Solución: Si usamos el triángulo de Pascal o la fórmula (6.1), tenemos:

a) $\binom{15}{3} = \frac{(15)(14)(13)(12!)}{(3)(2)(1)(12!)} = 455.$

b) $\binom{15}{0} = 1.$

c) $\binom{15}{15} = 1.$

d) $\binom{15}{0} + \binom{15}{1} + \binom{15}{2} = 1 + 15 + 105 = 121.$ ■

GRUPO DE EJERCICIOS 6.1

Habilidades básicas

1. Interprete los coeficientes binomiales siguientes, sin calcularlos, en términos de un experimento binomial.

   a) $\binom{8}{4}$ b) $\binom{6}{0}$

2. Haga lo mismo que en el anterior.

   a) $\binom{10}{1}$ b) $\binom{7}{7}$

Para los ejercicios 3 al 8, determine si el experimento es binomial. Si lo es, indique los valores para n, p y x e identifique un intento y un éxito.

3. Cuarenta por ciento de las personas que acampan en un lugar de veraneo sufren de irritación en la piel debido a una ortiga; de ocho estudiantes que acuden allí, estamos interesados en conocer el número de los que padecen la irritación cutánea.
4. Un estudio de residentes en una cierta ciudad, indicó que 30% de ellos están a favor de construir un centro comunitario y que el 70% se opone; se eligen diez personas al azar y se les pregunta si están a favor del nuevo centro comunitario.
5. Se lanza un dado cinco veces y se determina la suma de las cinco caras mostradas hacia arriba.
6. Se quiere calcular el número de hombres que entre los siguientes diez bebés que nazcan en el Memorial Hospital. (Suponga que los nacimientos de hombres y de mujeres son igualmente posibles.)
7. En una universidad se titula un 40% de los estudiantes que ingresan; determine el número de estudiantes que eventualmente se graduará de entre 30 que ingresarán el próximo semestre.
8. Cuatro candidatos están compitiendo para una gubernatura; se realiza un estudio para estimar los votantes potenciales para cada uno de los cuatro.
9. Obtenga cada uno de los números siguientes:

   a) $5!$ b) $\binom{18}{7}$ c) $0!$

   d) $\binom{12}{5}$ e) $\binom{4}{0}$

10. Calcule cada uno de los números siguientes:

    a) $\binom{5}{5}$ b) $\binom{25}{23}$ c) $\binom{11}{4}$

    d) $\binom{4}{2}(0.2)^2\,(0.8)^2$ e) $\binom{12}{7}$

Un paso más allá

11. Demuestre que $\binom{n}{x} = \binom{n}{n-x}$. ¿Cómo se relaciona este hecho con el triángulo de Pascal? Explique.
12. Para un entero positivo n, encuentre la suma de todos los coeficientes binomiales.
13. Demuestre que $\binom{n+1}{x} = \binom{n}{x} + \binom{n}{x-1}$. De qué manera se relaciona este hecho con el triángulo de Pascal? Explique.

SECCIÓN 6.2 Cálculo de probabilidades binomiales

En esta sección, nos interesaremos en el cálculo de las probabilidades asociadas con los resultados de un experimento binomial.

Fórmula de probabilidad binomial

Regresemos al ejemplo de las reservaciones aéreas de la sección 6.1, un experimento binomial con $n = 4$ y $p = 0.80$. Recuerde las 16 posibilidades referidas:

$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$
FFFF	*EFFF*	*EEFF*	*EEEF*	*EEEE*
	FEFF	*EFEF*	*EEFE*	
	FFEF	*EFFE*	*EFEE*	
	FFFE	*FEEF*	*FEEE*	
		FEFE		
		FFEE		

Esto puede resumirse como sigue

Valor de x (número Es)	0	1	2	3	4
Núm. de resultados posibles	1	4	6	4	1

Nota: Cada resultado asociado con un valor dado de x tiene la misma probabilidad porque los intentos son independientes. Por ejemplo, si $x = 2$,

$$P(EEFF) = P(E)\,P(E)\,P(F)\,P(F) = (0.80)^2(0.20)^2$$

y

$$P(EFEF) = P(E)\,P(F)\,P(E)\,P(F) = (0.80)^2(0.20)^2$$

Las otras cuatro posibilidades con $x = 2$ éxitos, tienen también probabilidades iguales a $(0.80)^2(0.20)^2$. Las probabilidades asociadas con los cinco valores posibles de x se resumen en la tabla 6.1.

TABLA 6.1

Probabilidad de los resultados individuales para un experimento binomial con $n = 4$ y $p = 0.80$

x	Probabilidad de cada resultado
0	$(0.80)^0(0.20)^4$
1	$(0.80)^1(0.20)^3$
2	$(0.80)^2(0.20)^2$
3	$(0.80)^3(0.20)^1$
4	$(0.80)^4(0.20)^0$

Cuatro observaciones relativas a las entradas de la tabla son:

1. El exponente de $p = P(E)$ es igual al número de éxitos x.
2. El exponente de $(1 - p) = P(F)$ es igual al número de fracasos $(n - x)$.
3. Las dos bases, p y $(1 - p)$, deben sumar 1.
4. La suma de los exponentes, x y $(n - x)$, debe ser n.

Como para cada valor de x hay $\binom{4}{x}$ resultados posibles que contienen x éxitos, veamos la tabla de probabilidades 6.2 para el número de éxitos x.

TABLA 6.2

Probabilidades para un experimento binomial con $n = 4$ y $p = 0.80$

Núm. de sucesos x	Núm. de resultados que tiene sucesos x	Probabilidad de cada resultado	Probabilidad de sucesos, $P(x)$
0	1	$(0.80)^0(0.20)^4$	$(1)(0.80)^0(0.20)^4 = 0.002$
1	4	$(0.80)^1(0.20)^3$	$(4)(0.80)^1(0.20)^3 = 0.026$
2	6	$(0.80)^2(0.20)^2$	$(6)(0.80)^2(0.20)^2 = 0.154$
3	4	$(0.80)^3(0.20)^1$	$(4)(0.80)^3(0.20)^1 = 0.410$
4	1	$(0.80)^4(0.20)^0$	$(1)(0.80)^4(0.20)^0 = 0.410$

Recuerde del álgebra que $p^0 = 1$ para todos los valores no cero de p. Observe de la tabla que las probabilidades siguen el patrón siguiente:

$$\binom{n}{x} p^x (1 - p)^{n - x}$$

Una fórmula general para calcular $P(x)$; la probabilidad de obtener x éxitos en un experimento binominal teniendo n intentos con $P(E) = p$, se conoce como **fórmula de probabilidad binomial**.

Fórmula de probabilidad binomial

$$P(x) = \binom{n}{x} p^x (1 - p)^{n-x} \tag{6.2}$$

EJEMPLO 6.3

Encontremos la probabilidad de $x = 2$ éxitos cuando $n = 5$ y $p = 0.3$, con el uso de la fórmula (6.2) tenemos:

$$\begin{aligned} P(x) &= \binom{n}{x} p^x (1 - p)^{n-x} \\ P(2) &= \binom{5}{2} (0.3)^2 (1 - 0.3)^{5-2} \\ &= \binom{5}{2} (0.3)^2 (0.7)^3 \\ &= (10)(0.09)(0.343) = 0.3087 \approx 0.309 \end{aligned}$$

Nota: En este capítulo redondearemos las probabilidades a la tercera cifra decimal.

Podemos usar MINITAB para determinar probabilidades binomiales, en la pantalla 6.1 vemos el uso de la orden PDF (del inglés "probability density function") para obtener $P(2)$ para un experimento binomial con $n = 5$ y $p = 0.3$.

Pantalla 6.1

```
MTB > PDF2;
SUBC > BINOMINAL WITH N = 5, P = 0.3
      K      P(X = K)
      2.00    0.3087
MTB >
```

El procedimiento siguiente probará ser útil para determinar probabilidades asociadas con un experimento binomial:

1. Determine qué constituye un intento y un éxito.
2. Calcule la probabilidad de éxito, p.
3. Cuantifique el número de intentos, n.
4. Encuentre los valores posibles de x, el número de éxitos.
5. Use la fórmula de probabilidad binomial para encontrar $P(x)$:

$$P(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

APLICACIÓN 6.5

Un estudio reciente mostró que el 60% de los estudiantes universitarios fuman. ¿Cuál es la probabilidad de que al elegir a cinco estudiantes, tres de ellos fumen?

Solución: Si usamos el procedimiento anterior de cinco pasos, tenemos:
Paso 1 Un intento consiste en determinar si un estudiante universitiario fuma, un éxito es encontrar que un estudiante fuma
Paso 2 $p = 0.6$
Paso 3 $n = 5$
Paso 4 $x = 3$
Paso 5

$$\begin{aligned} P(x) &= \binom{n}{x} p^x (1-p)^{n-x} \\ P(3) &= \binom{5}{3} (0.6)^3 (1-0.6)^{5-3} \\ &= \binom{5}{3} (0.6)^3 (0.4)^2 \\ &= (10)(0.216)(0.16) = 0.3456 \approx 0.346 \quad \blacksquare \end{aligned}$$

APLICACIÓN 6.6

Si $P(12) = \dfrac{18!}{12!6!}(0.7)^{12}(0.3)^6$ representa una probabilidad binomial, encuentre:

a) el número de éxitos.
b) $P(E)$.
c) el número de fracasos.
d) $P(F)$.

Solución:

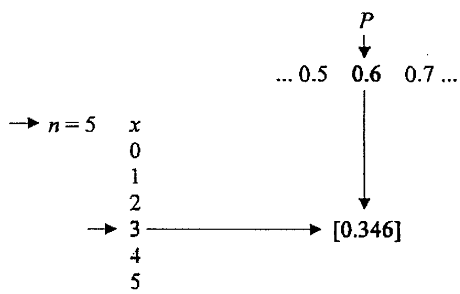
a) $P(12)$ indica que $x = 12$.
b) $P(E) = p = 0.7$, la base que tiene el exponente 12.
c) $n - x = 6$, el exponente de la base 0.3.
d) $P(F) = 1 - 0.7 = 0.3$, la otra base. ■

Tablas de probabilidad binomial

La tabla 1 del apéndice B da valores de $P(x)$ para valores selectos de p y valores de n hasta 25; para usar la tabla, localice la sección apropiada para el valor de n, después, bajo la columna etiquetada con p y a lo largo del renglón x encuentre el valor de $P(x)$. Considere la aplicación 6.7.

APLICACIÓN 6.7

Si un jugador de beisbol con un porcentaje de bateo de 0.600 va al bat cinco veces durante un juego, ¿cuál es la probabilidad de que logre tres hits?
Solución: Se identifican los valores siguientes: $n = 5$, $p = 0.6$ y $x = 3$. Mediante las tablas de probabilidad binomial, encontramos:



Por tanto, $P(3) = 0.346$.

Si usamos la fórmula de probabilidad binomial, obtenemos:

$$\begin{aligned} P(3) &= \binom{5}{3}(0.6)^3(0.4)^2 \\ &= (10)(0.6)^3(0.4)^2 \\ &= 0.3456 \approx 0.346 \end{aligned}$$

lo cual concuerda con la respuesta obtenida de las tablas de probabilidad binomial. ■

Nota: Para valores dados de n y p, la suma de todos los de la tabla puede no ser exactamente 1.000. Por ejemplo, para $n = 4$ y $p = 0.8$, la suma de los valores de la tabla es $0.002 + 0.026 + 0.154 + 0.410 + 0.410 = 1.002$; esto se debe a que las entradas de la tabla se han redondeado a la tercera cifra decimal, una entrada 0.000 en la tabla significa que es 0 al redondearla a tres lugares decimales.

Se puede usar MINITAB para generar tablas de probabilidad binomial. La pantalla 6.2 contiene las órdenes y las respuestas de MINITAB para las probabilidades binomiales listadas en la tabla 6.2.

Pantalla 6.2

```
MTB >PDF;
SUBC > BINOMINAL N = 4 P = 0.8

  BINOMINAL WITH N = 4 P = 0.800000
         K      P(X = K)
         0       0.0016
         1       0.0256
         2       0.1536
         3       0.4096
         4       0.4096
MTB >
```

En general, para calcular probabilidades binomiales es mucho más sencillo usar un paquete estadístico, como MINITAB, que la fórmula de probabilidad binomial; la tabla binomial en el apéndice B fue generada por computadora, podemos usarla para buscar muchas de las probabilidades binomiales usadas en este libro; la computadora resulta más manejable que la fórmula de probabilidad binomial cuando la probabilidad de éxito no es uno de los valores listados en la tabla o cuando el número de intentos excede de 25. Por ejemplo, si queremos encontrar $P(5)$ para un experimento binomial donde $p = 0.178$ y $n = 10$, no podríamos usar la tabla binomial del apéndice B porque no hay una columna encabezada por $p = 0.178$, en ese caso debemos usar un paquete computacional o la fórmula de probabilidad binomial. La pantalla 6.3 muestra los resultados de usar MINITAB para generar la tabla; se encuentra que el valor de $P(5)$ es 0.0169. Creo que usted estará de acuerdo en que el uso de MINITAB en este caso es más sencillo que el de la fórmula de probabilidad binomial.

Pantalla 6.3

```
MTB > PDF;
SUBC > BINOMINAL N = 10, P = 0.178

 BINOMINAL WITH N = 10 P = 0.178000
          K      P(X = K)
          0        0.1408
          1        0.3050
          2        0.2972
          3        0.1716
          4        0.0650
          5         0.169
          6        0.0030
          7        0.0004
          8        0.0000

MTB >
```

GRUPO DE EJERCICIOS 6.2

Más aplicaciones

Para cada uno de los problemas siguientes, usted debe verificar que involucran un experimento binomial, identificando los valores de p, n y x antes de resolverlos. Sus respuestas a los problemas de probabilidad pueden diferir de las dadas en el libro, dependiendo de que use la fórmula de probabilidad binomial o tablas de probabilidad, esas diferencias se deben a errores de redondeo en las tablas.

1. Se encontró que 40% de las personas que acampan en un lugar de veraneo, sufren de irritación cutánea debido a una ortiga. Si ocho estudiantes acuden a dicho campamento este verano, encuentre la probabilidad de que:
   a) todos padezcan de irritación cutánea.
   b) dos padezcan la molestia cutánea.
   c) a lo más tres sufran los efectos de la ortiga.
   d) por lo menos siete presenten irritación de la piel.
2. Un estudio de los residentes de cierta ciudad mostró que 30% prefieren la marca X de pasta dentrífica; si se analiza a diez de esos residentes al azar en una miscelánea, ¿cuál es la probabilidad de que:
   a) ninguna prefiera la marca X?;
   b) a cuatro sí les guste la marca X?;
   c) al menos ocho prefieran la marca X?;
   d) a lo más dos se decidan por la marca X?;
   e) todas se inclinen por la marca X?
3. Se ha observado que Ricardo logra 65% de sus tiros libres en los partidos de basquetbol, calcule la probabilidad de que Ricardo logre:
   a) tres de los siguientes seis tiros.
   b) cinco de los próximos diez tiros.
   c) todos los cuatro tiros siguientes.
4. En una ciudad en particular, 40% de los votantes registrados son demócratas; si se elige a nueve votantes al azar, encuentre la probabilidad de que:
   a) dos de ellos sean demócratas;
   b) al menos uno de ellos resulte demócrata;
   c) al menos ocho digan ser demócratas;
   d) a lo mucho tres sean demócratas.
5. Suponga que hombres y mujeres tienen la misma probabilidad de nacer, técnicamente esto no es correcto, porque nacen aproximadamente 100 mujeres por cada 106 hombres. Encuentre la probabilidad de que

entre los siguientes seis bebés que nazcan:
a) haya tres niños;
b) haya cuando mucho tres niños;
c) estén cuatro niñas;
d) ninguno sea niño.

6. En cierta universad se gradúa el 35% de los estudiantes que ingresan; ¿cuál es la probabilidad de que cinco principiantes que recién ingresan:
a) se gradúen todos?;
b) cuatro reciban su título?;
c) no se titulen tres?;
d) se gradúen al menos cuatro?;
e) ninguno termine la educación universitaria?

7. Una semilla tiene un porcentaje de germinación del 83%, si se siembran 32 semillas, encuentre la probabilidad de que:
a) germinen todas;
b) germinen 10;
c) broten 11;
d) salgan a lo más 2;
e) germinen al menos 10.

8. Si se lanzan cinco dados, ¿cuál es la probabilidad de que:
a) tres de ellos muestren un uno?;
b) todos muestren un uno?;
c) al menos tres muestren un uno?;
d) cuatro no muestren un uno?

9. Un examen consiste en diez preguntas de opción múltiple con cinco respuestas posibles. Si una persona responde siempre adivinando, ¿cuál es la probabilidad de que:
a) responda correctamente a todas las preguntas?;
b) acierte a lo más a tres preguntas?;
c) conteste correctamente cinco preguntas?;
d) le atine a siete preguntas?

10. Si se lanzan 20 monedas, encuentre la probabilidad de que:
a) salgan 11 "soles";
b) salgan 15 "soles";
c) al menos salgan 16 "soles";
d) no salgan "soles" o únicamente salgan "soles".

Un paso más allá

11. El cálculo de probabilidades binomiales puede obtenerse recursivamente; representemos por $b(x;\, n,\, p)$ la probabilidad de obtener x éxitos en n intentos, si p es la probabilidad de éxito en un solo intento.
a) Demuestre que: $b(x + 1;\, n, p) = p(n - x)/[(x + 1)(1 - p)];\, b(x;\, n,\, p)$. Es decir, si conocemos la probabilidad de obtener x éxitos $b(x;\, n,\, p)$, podemos obtener la probabilidad de obtener $x + 1$ éxitos, $b(x + 1;\, n,\, p)$ multiplicando $b(x;\, n,\, p)$ por $p(n - x)/[(x + 1)(1 - p)]$.
b) Use la fórmula recursiva del inciso a) para calcular las probabilidades binomiales para $n = 5$ y $p = 0.3$.

12. Suponga que permitimos que el número de intentos varíe en un experimento binomial; es decir, que repetimos el experimento hasta que ocurra un éxito; la variable aleatoria X representa el número de veces necesarias para lograr un éxito, incluyendo este primer éxito. La fórmula de probabilidad para X es,

$$P(X = k) = (1 - p)^{k-1}p$$

donde $k = 1, 2, 3, \ldots$. La variable aleatoria X se denomina una *variable aleatoria geométrica*. Demuestre que la suma de las probabilidades es igual a uno.

13. Cuando se lanza una moneda, ¿cuál es la probabilidad de que el primer "sol" aparezca en el cuarto lanzamiento?

14. Un dado ordinario se lanzará hasta que aparezca un 6 por primera vez. ¿Cuál es la probabilidad de que eso suceda en:
a) el tercer lanzamiento?;
b) el quinto lanzamiento?

15. Resuelva los problemas propuestos en el motivador 6.

SECCIÓN 6.3 Cálculo de parámetros para distribuciones binomiales

Recordemos que la distribución de probabilidad para el número x de éxitos que se obtienen de un experimento binomial se conoce como *distribución binomial*. En esta sección queremos encontrar la media y la varianza de una distribución binomial y examinar la forma de su gráfica de probabilidad.

Una distribución de probabilidad binomial puede mostrarse como una tabla o como una colección de $(n + 1)$ pares ordenados $(x, P(x))$ cuyo primer componente es el número de éxitos y su segundo componente, la probabilidad asociada al número dado de éxitos; esta forma de mostrarla puede usarse para construir una gráfica de la distribución.

EJEMPLO 6.4

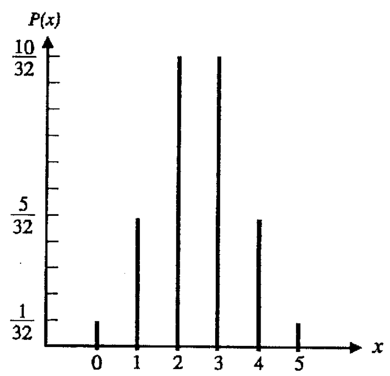
Considere el experimento binomial de lanzar una moneda legal cinco veces. La tabla 6.3 es una tabla de distribución de probabilidad para el número de "soles" obtenidos.

TABLA 6.3
Tabla de distribución de probabilidad para el número de "soles" obtenidos en cinco lanzamientos de una moneda balanceada

x	$\binom{5}{x}$	$P(x)$
0	1	1/32
1	5	5/32
2	10	10/32
3	10	10/32
4	5	5/32
5	1	1/32
		1

Notemos que como $p = 0.5$, la probabilidad de obtener cualquier selección particular de cinco resultados es (1/2)(1/2)(1/2)(1/2)(1/2) = 1/32. Por tanto, $P(x) = \binom{5}{x}\left(\frac{1}{32}\right)$ para $x = 0, 1, 2, 3, 4$ y 5. Esta información puede usarse para construir la gráfica de distribución de probabilidad para el número de éxitos (véase la figura 6.2). Advierta que la distribución de probabilidad es simétrica respecto a $x = 2.5$, el cual parece ser el "centro" de la distribución. Cuando $p = 0.5$, la distribución de probabilidad será siempre simétrica.

FIGURA 6.2
Gráfica de distribución de probabilidad para el número de soles en cinco lanzamientos de una moneda

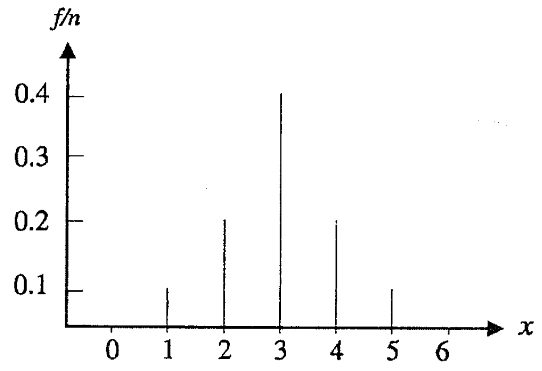


Una distribución binomial describe cuidadosamente, hasta un grado razonable, la repetición de experimentos binomiales consistentes en intentos repetidos, cada uno de los cuales tiene dos resultados; también nos presenta todos los resultados posibles y las probabilidades correspondientes para un experimento binomial; es muy similar a una distribución de frecuencias relativas, que presenta lo que ha ocurrido, pero difiere de dicha distribución porque se proyecta a futuro y nos dice - utilizando la probabilidad - lo que teóricamente debe ocurrir.

Si un experimento binomial se repite un número fijo de veces, la distribución de frecuencia relativa resultante para cada uno de los $(n + 1)$ valores posibles para x éxitos se denomina **distribución binomial empírica**; cuando el número de repeticiones del experimento binomial crece, las distribuciones empíricas se aproximan a la distribución binomial teórica. Estas dos distribuciones se ilustran en la figura 6.3 para el experimento binomial de lanzar una moneda seis veces y anotar el número de "soles" obtenidos. Para la distribución empírica el experimento se repitió 100 veces. Vea que la gráfica de la distribución se parece a la de la distribución teórica.

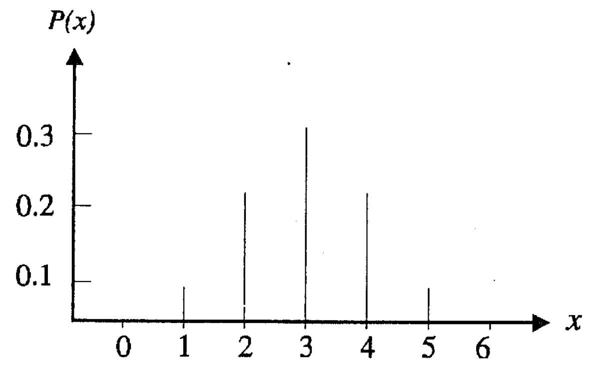
FIGURA 6.3
Comparación de distribuciones binomiales empírica y teórica

x	f/n
0	0.01
1	0.09
2	0.21
3	0.41
4	0.20
5	0.08
6	0.00



(a) Distribución binominal empírica (n = 100)

x	$P(x)$
0	0.016
1	0.094
2	0.234
3	0.313
4	0.234
5	0.094
6	0.016



(b) Distribución binominal teórica

Mucha gente confunde los experimentos binomiales con las distribuciones binomiales, pero hay una diferencia; esto es, un experimento binomial consistente en n intentos da lugar a exactamente un resultado de los $(n + 1)$ posibles para la variable binomial aleatoria asociada. Por otro lado, una distribución binomial describe las probabilidades asociadas con los $(n + 1)$ valores (resultados) de la variable aleatoria X que denota el número de éxitos que puede obtenerse.

Ahora queremos obtener el promedio de una distribución binomial; esto es, encontrar la media de una variable binomial aleatoria. Para una distribución binomial, la media μ es el valor esperado $E(x)$ para el número de éxitos x. Recuerde de la sección 5.7 que:

$$E(X) = \Sigma\,[xP(x)] \tag{6.3}$$

EJEMPLO 6.5

Para el experimento de lanzar una moneda cinco veces, la media de la variable aleatoria X que denota el número de "soles" es:

$$\begin{aligned}\mu &= \Sigma\,[xP(x)]\\ &= 0P(0) + 1P(1) + 2P(2) + 3P(3) + 4P(4) + 5P(5)\\ &= 0 + 1\left(\frac{5}{32}\right) + 2\left(\frac{10}{32}\right) + 3\left(\frac{10}{32}\right) + 4\left(\frac{5}{32}\right) + 5\left(\frac{1}{32}\right)\\ &= \frac{80}{32} = 2.5\end{aligned}$$

¿Cómo debe interpretarse esta respuesta? Es claro que no significa que si lanzamos una moneda cinco veces obtendremos 2.5 "soles", sino que, si para un gran número de personas, donde cada una lanza cinco veces una moneda, anotamos el número de "soles" obtenidos, el promedio sería cercano a 2.5.

La fórmula (6.3) para calcular la **media μ de una distribución binomia** puede requerir cálculos considerables. Considere que intuitivamente es razonable que una máquina que produce partes defectuosas el 1% del tiempo, al terminar lotes de 500 partes por un periodo largo, debería producir, en promedio un número de piezas defectuosas por lote igual a (0.01)(500) = 50; de forma similar, es razonable esperar que en un número grande de experimentos repetidos de lanzar una moneda 300 veces se obtengan (0.5)(300)= 150 "soles". En ambas situaciones, el valor esperado, promedio, se obtuvo multiplicando el número de intentos n por la probabilidad de éxito p; como consecuencia, para encontrar la media de una distribución binomial tenemos la fórmula siguiente que matemáticamente es equivalente a la (6.3) pero más fácil de usar.

Media de una distribución binomial

$$\mu = np \tag{6.4}$$

EJEMPLO 6.6

Para el ejemplo 6.5, la media es:

$$\begin{aligned}\mu &= np\\ &= (5)(0.5) = 2.5\end{aligned}$$

Vea que esta respuesta concuerda con la encontrada usando la fórmula 6.3.

APLICACIÓN 6.8

La probabilidad de que un paciente se recupere de una cirugía de pulmón es 0.95; si 25 personas se someten a esta cirugía, encuentre el número de la media de recuperaciones e interprete el resultado.

Solución: Si usamos la fórmula (6.4), tenemos

$$\begin{aligned}\mu &= np\\ &= (25)(0.95) = 23.75\end{aligned}$$

Interpretamos esto así: si en cada uno de un gran número de hospitales se realiza cirugía de pulmón a 25 pacientes y se registra la cantidad de los que se recuperan, el promedio de recuperaciones en todos los hospitales estudiados será cercano a 23.75. ■

Varianza de una distribución binomial

Recuerde de la sección 5.7 que la varianza de una variable aleatoria X está dada por

$$\sigma_X^2 = \Sigma\,[(x - \mu)^2 P(x)] \tag{6.5}$$

La aplicación 6.9 deja ver cómo encontrar la varianza de una distribución binomial.

APLICACIÓN 6.9

David y Ricardo están jugando a tirar un dado que si muestra un 2, un 3 o un 4, gana David, de lo contrario gana Ricardo; si el dado se lanza tres veces, encuentre la media y la varianza para la distribución del número de veces que gana David.

Solución: El experimento es binomial con $n = 3$ y $p = 1/2$. Mediante la fórmula (6.4) obtenemos.

$$\begin{aligned}\mu &= np\\ &= (3)\left(\frac{1}{2}\right) = 1.5\end{aligned}$$

Para encontrar la varianza, necesitamos la siguiente tabla de distribución de probabilidad para x, el número de veces que gana David:

x	$P(x)$
0	1/8
1	3/8
2	3/8
3	1/8

A fin de facilitar los cálculos necesarios para obtener σ^2, usamos la tabla

x	$P(x)$	$x-\mu$	$(x-\mu)^2$	$(x-\mu)^2 P(x)$
0	1/8	− 1.5	2.25	0.28125
1	3/8	− 0.50	0.25	0.09375
2	3/8	0.5	0.25	0.09375
3	1/8	1.5	2.25	0.28125
				0.75

En consecuencia, $\sigma_x^2 = 0.75$.

Para una distribución binominal, la fórmula (6.5) es matemáticamente equi-valente a ésta, más fácil de usar:

Varianza de una distribución binomial

$$\sigma^2 = np(1-p) \tag{6.6}$$

Preferida para calcular la **varianza σ^2 para una distribución binomial** porque requiere menos cálculos.

EJEMPLO 6.7

Si usamos la fórmula (6.6) para calcular la varianza σ^2 de los datos binomiales de la aplicación 6.9, conseguiremos:

$$\begin{aligned}\sigma^2 &= np(1-p)\\ &= (3)(0.5)(0.5) = 0.75\end{aligned}$$

que concuerda con el resultado que obtenemos usando la fórmula (6.5).

La **desviación estándar de una distribución binomial** es la raíz cuadrada de su varianza.

Desvación estándar de una variable aleatoria binomial

$$\sigma = \sqrt{np\,(1-p)} \tag{6.7}$$

Consideremos su uso en las aplicaciones 6.10 y 6.11.

APLICACIÓN 6.10

Refiérase a la aplicación 6.9 y suponga que David gana si el dado muestra un 2 o un 4; encuentre la media y la desviación estándar para la distribución del número de veces que gana David.

Solución: Para este experimento binomial, $p = 1/3$. Usando la fórmula (6.4) obtenemos:

$$\begin{aligned}\mu &= np\\ &= (3)\left(\frac{1}{3}\right) = 1\end{aligned}$$

Por tanto, David espera ganar una vez de cada tres intentos. Para calcular la varianza aplicamos la fórmula (6.6) para obtener:

$$\begin{aligned}\sigma^2 &= np(1-p)\\ &= (3)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{3}\end{aligned}$$

Así, usando la fórmula (6.7) encontramos que la desviación estándar σ es:

$$\sigma = \sqrt{\frac{2}{3}} \approx 0.82 \quad \blacksquare$$

APLICACIÓN 6.11

Un estudiante presenta un examen de opción múltiple con 50 preguntas, cada una de ellas con cinco elecciones posibles, si responde cada pregunta adivinando, encuentre la media y la desviación estándar de la distribución del número de preguntas respondidas correctamente, así como la media y la desviación estándar para la distribución del número de preguntas en que falla el estudiante.

Solución: Para este experimento binomial, $n = 50$ y $p = 0.20$; según la fórmula (6.4), la media es:

$$\begin{aligned}\mu &= np\\ &= (50)(0.2) = 10\end{aligned}$$

El estudiante debe esperar obtener nada más adivinando, 10 respuestas correctas. Para determinar la desviación estándar, usamos la fórmula (6.7):

$$\begin{aligned}\sigma &= \sqrt{np(1-p)}\\ &= \sqrt{(50)(0.2)(0.8)} \approx 2.83\end{aligned}$$

El número de respuestas incorrectas forma una distribución binomial con $n = 50$ y $p = 0.80$. Así, la media será:

$$\begin{aligned}\mu &= np\\ &= (50)(0.80) = 40\end{aligned}$$

y la desviación estándar:

$$\sigma = \sqrt{(50)(0.80)(0.20)} \approx 2.83$$

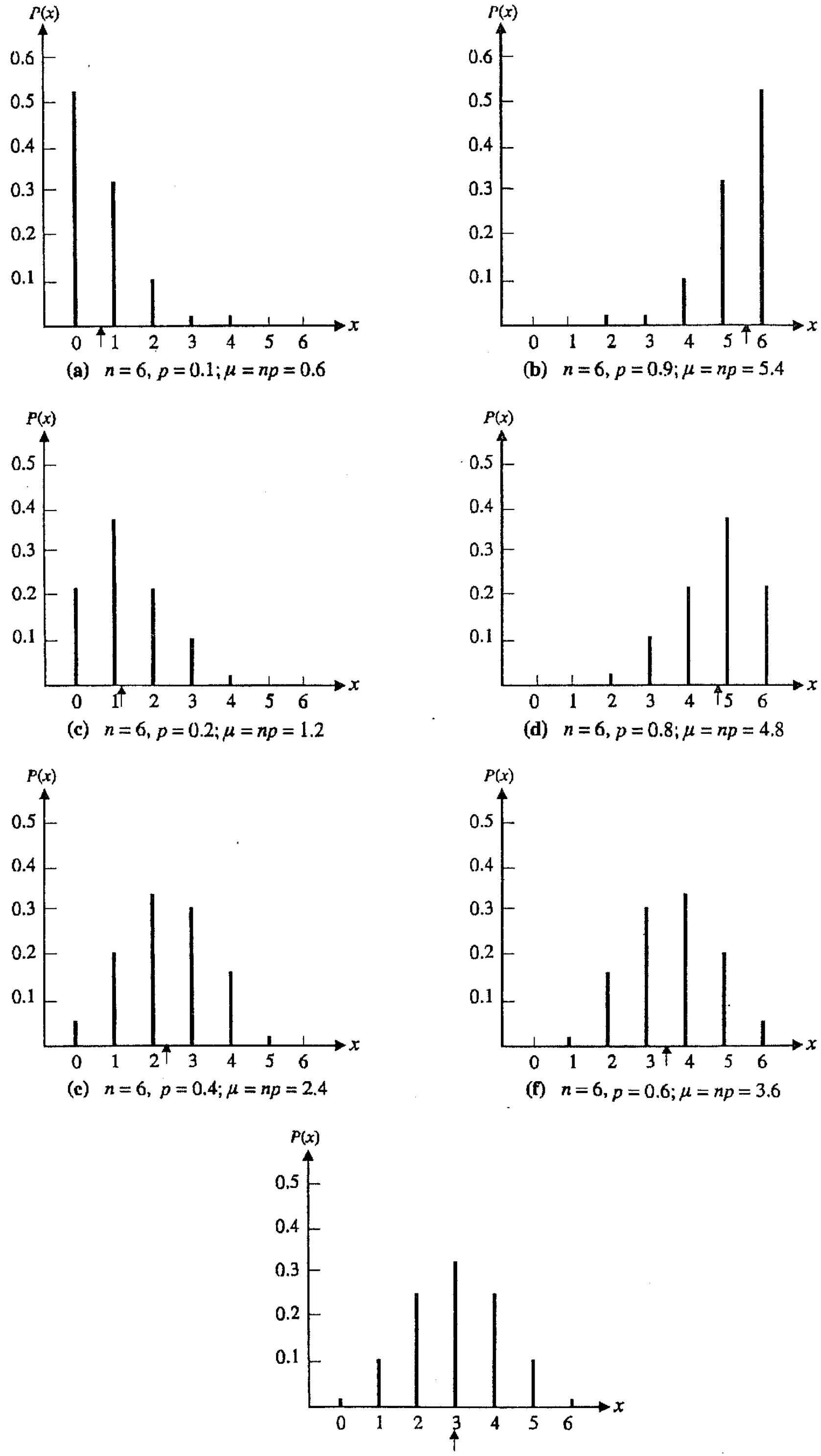
Nota: Las desviaciones estándar son iguales y la suma de las dos medias es igual a n. ▪

Formas de las gráficas de distribuciones binomiales

Los valores de n y p determinan la forma de una distribución binomial. Las gráficas de la figura 6.4 (a-g), indican cómo cambian las gráficas de las distribuciones binomiales para valores distintos de p cuando $n = 6$. Se usaron siete valores distintos de p. Advierta que en la figura 6.4 g, la gráfica es simétrica respecto a su media $\mu = 3$. En general, si $p = 1/2$ tenemos:

$$\mu = np$$
$$= \frac{n}{2}$$

FIGURA 6.4
Gráficas de distribuciones binomiales para $n = 6$ y diferentes valores de p



(a) $n = 6, p = 0.1; \mu = np = 0.6$

(b) $n = 6, p = 0.9; \mu = np = 5.4$

(c) $n = 6, p = 0.2; \mu = np = 1.2$

(d) $n = 6, p = 0.8; \mu = np = 4.8$

(e) $n = 6, p = 0.4; \mu = np = 2.4$

(f) $n = 6, p = 0.6; \mu = np = 3.6$

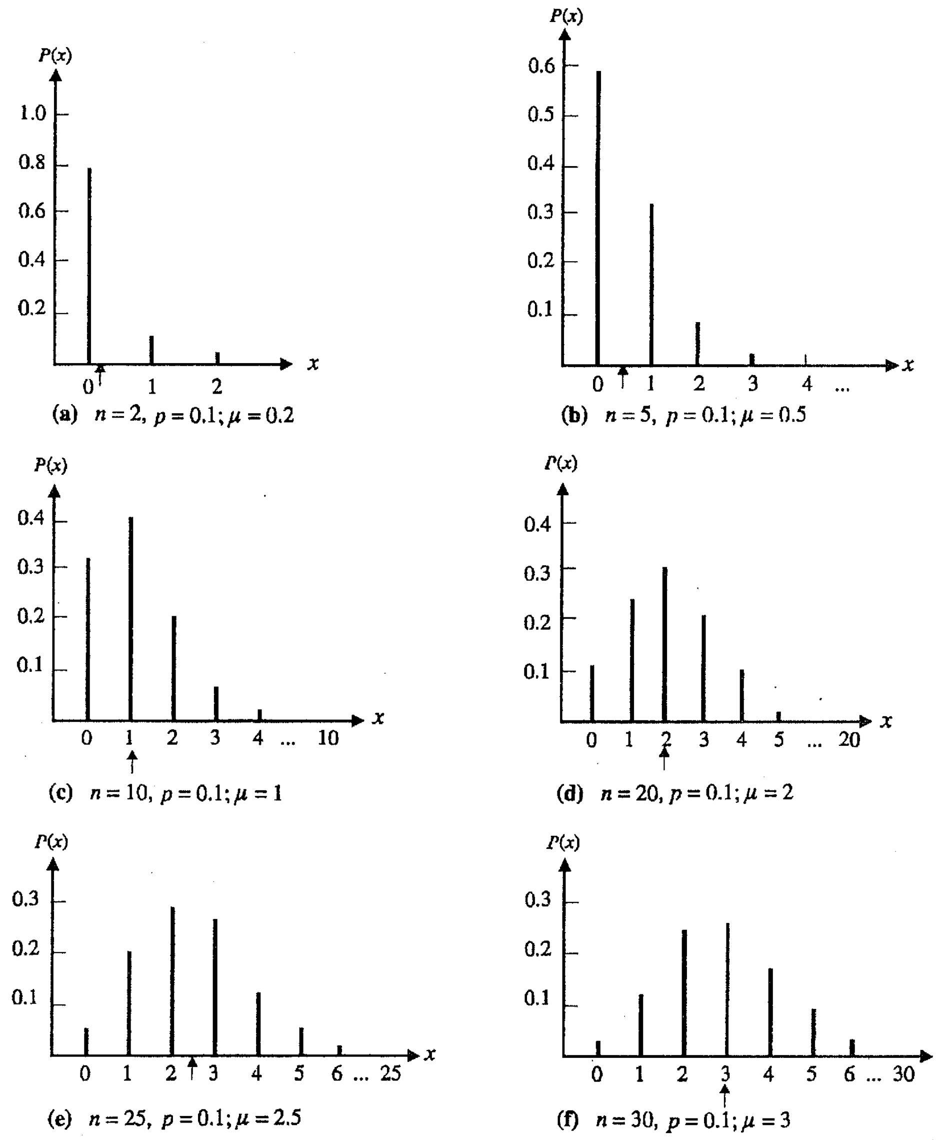
(g) $n = 6, p = 0.5; \mu = np = 3$

y la gráfica de la distribución binomial es simétrica respecto a su media $\mu = n/2$. Y si $p \neq 0.5$, note que la gráfica de una distribución binomial no es simétrica respecto a su media. En las figuras 6.4a, c, e y g puede apreciarse que cuando el valor de p se aproxima por la izquierda al valor 0.5, la media, marcada por una flecha, casi alcanza por la izquierda al valor $n/2 = 3$; al examinar las gráficas de las figuras 6.4b, d, f y g, vemos también que cuando los valores de p se aproximan por la derecha al valor 0.5, los valores de μ llegan cerca, por la derecha, al valor $n/2 = 3$.

Las gráficas de la figura 6.5 (a-f), ilustran el efecto de variar en una distribución binomial los valores de n para un valor fijo de $p = 0.1$. Las gráficas de probabilidad están dadas por $n = 2, 5, 10, 20, 25$ y 30. Note que cuando el tamaño n de la muestra crece, la media, indicada por una flecha, también crece y parece que las distribuciones se vuelven más simétricas respecto a sus medias.

FIGURA 6.5

Gráficas de distribuciones binomiales para $p = 0.1$ y diferentes valores de n



APLICACIÓN 6.12 Construya una gráfica para la distribución binomial que tiene $n = 10$ y $p = 0.6$, y encuentre μ.

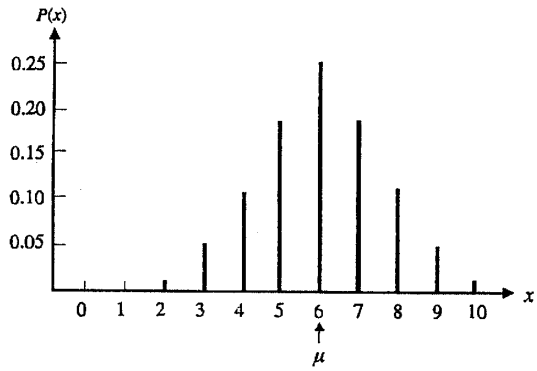
Solución: Las probabilidades siguientes se encontraron usando la tabla de probabilidades binomial (tabla 1 del apéndice B).

x	0	1	2	3	4	5	6	7	8	9	10
$P(x)$	0.000	0.002	0.011	0.042	0.111	0.201	0.251	0.215	0.121	0.040	0.006

La gráfica correspondiente se muestra en la figura 6.6. La media μ es

$$\begin{aligned}\mu &= np\\ &= (10)(0.6)\\ &= 6\end{aligned}$$

FIGURA 6.6
Distribución binomial con $n = 10$ y $p = 0.6$



Aunque la gráfica no es simétrica respecto a $\mu = 6$, deja ver que la distribución no está muy sesgada, lo que no es sorprendente porque el valor $p = 0.6$ es cercano al valor $p = 0.5$. ■

GRUPO DE EJERCICIOS 6.3

Habilidades básicas

1. Suponga que las condiciones para un experimento binomial son válidas en las situaciones siguientes. Para cada una, encuentre μ, σ^2 y σ.
   a) $n = 10, p = 0.6$ b) $n = 20, p = 0.85$
   c) $n = 40, p = 0.3$ d) $n = 5, p = 0.2$
2. Para cada una de las situaciones siguientes encuentre μ, σ^2 y σ.
   a) $n = 15, p = 0.9$ b) $n = 6, p = 0.45$
   c) $n = 500, p = 0.1$ d) $n = 4, p = 0.60$

Más aplicaciones

3. Si el 2% de todos los radios fabricados por la AJAX Company son defectuosos, encuentre la media y la desviación estándar de radios defectuosos en lotes de 50.
4. Encuentre la media y la varianza del número de niñas en familias con seis hijos, si se sabe que nacen 100 niñas por cada 106 niños.
5. ¿Cuáles son la media y la desviación estándar del número de "soles" que se obtiene cuando se lanzan 1000 monedas al aire?
6. La probabilidad de que José acierte a una claraboya con su rifle es 68%; si dispara en rondas de diez tiros, encuentre la media y la desviación estándar del número de aciertos por ronda.
7. Con base en un periodo largo, se ha determinado que 90% de todos los estudiantes inscritos en el curso 209 de matemáticas lo aprueban; si toman el curso grupos de 30 estudiantes, determine la media y la desviación estándar del número de estudiantes aptos por curso.

8. Construya gráficas similares a las de la figura 6.5 para las distribuciones de probabilidad siguientes, asumiendo que se cumplen las condiciones binomiales, localice μ en cada gráfica.
   a) $n = 5,\ p = 0.2$
   b) $n = 6,\ p = 0.45$
   c) $n = 4,\ p = 0.60$

9. Trace una gráfica de la distribución de probabilidad para cada una de las distribuciones binomiales siguientes.
   a) $n = 20, p = 0.5$
   b) $n = 20, p = 0.4$
   c) $n = 20, p = 0.6$
   d) $n = 20, p = 0.3$
   e) $n = 20, p = 0.8$
   f) $n = 20, p = 0.2$

Un paso más allá

10. Suponga que hemos obtenido una distribución binomial con n intentos y probabilidad de éxito p. Ahora considere la distribución binomial con n intentos en la cual la probabilidad de éxito es $1 - p$.
    a) ¿Cómo se comparan las medias de las distribuciones binomiales?
    b) ¿Cómo se comparan sus desviaciones estándar?

11. Dada una distribución binomial con una valor fijo de n, ¿existen valores de p para los cuales $\sigma^2 = 0$? Explique.

12. Dada una distribución binomial con un valor fijo de n, ¿cuál es el valor de p para el que el valor de σ^2 es mayor? Explique.

13. Si $b(x; n, p)$ denota la probabilidad de x éxitos para un experimento binomial con n intentos y $P(E) = p$, demuestre que:

$$b(x; n, 1 - p) = b(n - x; n, p)$$

SECCIÓN 6.4 *Distribuciones multinomiales*

En las secciones de la 6.1 a la 6.3 examinamos experimentos binomiales y sus correspondientes distribuciones de probabilidad, mismos que se caracterizaron por las propiedades básicas siguientes:

1. El experimento consiste de n intentos idénticos.
2. Cada intento da lugar a uno de dos resultados llamados *éxito* y *fracaso*.
3. Los n intentos son independientes.
4. La probabilidad de éxito permanece constante de un intento a otro.

Si modificamos la propiedad 3 para permitir más de dos resultados en cada intento, el experimento se llama **experimento multinomial**; como un experimento multinomial contiene más de dos resultados, no usamos los términos éxito y fracaso para describirlos, y la probabilidad de cada resultado permanece constante de un intento a otro. Un experimento donde cada intento da lugar a uno de tres resultados posibles, se denomina **experimento trinomial**.

Experimentos trinomiales

EJEMPLO 6.8

Considere este ejemplo de un experimento trinomial: una clínica de salud se especializa en el uso de un método para pacientes artríticos; los registros demuestran que la mitad de todos los pacientes tratados ahí se benefician, que la tercera parte no presenta efecto alguno, y que una sexta parte tiene efectos colaterales adversos; se va a realizar un experimento con dos pacientes nuevos que sufren de artritis.

Un procedimiento es tratar a un paciente que sufre artritis. Como hay dos pacientes, hay $n = 2$ procedimientos. Éstos son independientes, porque el tratamiento de un paciente no tendrá influencia en el del otro enfermo. Cada procedimiento da lugar a uno de los tres resultados siguientes:

E_1: el paciente sale beneficiado;
E_2: el paciente no se beneficia;
E_3: el enfermo tiene una reacción adversa.

Al tratar a un solo paciente, las probabilidades resultantes son:

$$p_1 = P(E_1) = \frac{1}{2}$$

$$p_2 = P(E_2) = \frac{1}{3}$$

$$p_3 = P(E_3) = \frac{1}{6}$$

Note que como $p_1 + p_2 + p_3 = 1/2 + 1/3 + 1/6 = 1$, la probabilidad $P(E_3)$ se puede escribir como $p_3 = (1 - p_1 - p_2)$. Estas tres probabilidades permanecen constantes de un procedimiento a otro.

Para $n = 2$ procedimientos, el experimento puede dar lugar a uno de nueve resultados, estos se listan en la tabla 6.4.

TABLA 6.4
Resultados del experimento trinomial con $n = 2$

(E_1, E_1)	(E_2, E_1)	(E_3, E_1)
(E_1, E_2)	(E_2, E_2)	(E_3, E_2)
(E_1, E_3)	(E_2, E_3)	(E_3, E_3)

Por ejemplo, el resultado (E_1, E_3) indica que el primer paciente resulta beneficiado con el tratamiento y que el segundo presenta una reacción adversa. Se puede calcular la probabilidad para cada uno de los nueve resultados.

A fin de deducir la fórmula general para la distribución de probabilidad de un experimento trinomial, procedemos como lo hicimos para la distribución binomial estudiada en la sección 6.2. Recuerde la fórmula de probabilidad binomial:

$$P(x) = \binom{n}{x} p^x (1 - p)^{n - x}$$

El lado derecho de la fórmula de probabilidad binomial consta de dos factores: el coeficiente binomial $\binom{n}{x}$ y $p^x(1-p)^{n-x}$. El coeficiente binomial cuenta el número de formas en que pueden obtenerse x éxitos y $(n - x)$ fracasos en n intentos. El segundo factor, $p^x(1-p)^{n-x}$, representa la probabilidad de obtener cualquiera de los $\binom{n}{x}$ resultados.

Una probabilidad para un experimento trinomial constará también de dos factores, el primero es el coeficiente trinomial, si cada uno de n intentos puede dar lugar a tres resultados posibles, E_1, E_2 y E_3, entonces **el coeficiente**

trinomial $\binom{n}{x_1\ x_2\ x_3}$ proporciona el número de maneras en que, en n ensayos, se puede obtener x_1 resultados de tipo E_1, x_2 de tipo E_2 y x_3 resultados de tipo E_3. La fórmula siguiente, llamada la **fórmula del coeficiente trinomial,** proporciona el coeficiente trinomial:

Fórmula del coeficiente trinomial

$$\binom{n}{x_1\ x_2\ x_3} = \frac{n!}{x_1!x_2!x_3!} \tag{6.8}$$

EJEMPLO 6.9

Refiérease al ejemplo 6.8, las formas en que dos pacientes pueden dividirse en tres categorías de manera que la primera categoría tenga $x_1 = 0$ pacientes, la segunda tenga $x_2 = 1$ paciente y la tercera tenga $x_3 = 1$ paciente, pueden representarse por el coeficiente trinomial $\binom{2}{0\ 1\ 1}$; podemos determinar su valor mediante la fórmula (6.8).

$$\binom{2}{0\ 1\ 1} = \frac{2!}{0!\ 1!\ 1!} = 2$$

Este resultado se puede verificar examinando el ejemplo 6.4, los dos resultados son (E_2, E_3) y (E_3, E_2).

APLICACIÓN 6.13

Calcule los coeficientes trinomiales siguientes:

a) $\binom{4}{1\ 2\ 1}$

b) $\binom{6}{1\ 2\ 3}$

Solución: Usaremos la fórmula del coeficiente trinomial 6.8:

a) $\binom{4}{1\ 2\ 1} = \frac{4!}{1!2!1!} = \frac{24}{2} = 12$

b) $\binom{6}{1\ 2\ 3} = \frac{6!}{1!2!3!} = \frac{(6)(5)(4)(3)(2)(1)}{(2)(6)} = 60$ ■

Nota: En general, el coeficiente trinomial proporciona el número de formas en que se pueden colocar n objetos en tres categorías para las cuales hay x_1 objetos en la primera categoría, x_2 objetos en la segunda y x_3 objetos en la tercera $(x_1 + x_2 + x_3 = n)$.

APLICACIÓN 6.14

¿De cuántas formas pueden colocarse cinco pelotas distintas en tres cajas de tal forma que la primera caja contenga dos pelotas, la segunda una pelota y la tercera dos pelotas?

Solución: El coeficiente trinomial $\binom{5}{2\ 1\ 2}$ proporciona la respuesta mediante la fórmula (6.8).

$$\binom{5}{2\ 1\ 2} = \frac{5!}{2!\,1!\,2!} = 30 \quad \blacksquare$$

Como para un experimento trinomial los n intentos son independientes, cualquier orden específico que de x_1 resultados del tipo E_1, x_2 resultados del tipo E_2 y x_3 resultados del tipo E_3, ocurrirá con probabilidad $p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3}$. El número total de ordenamientos de los resultados que proporciona esta probabilidad está dado por el coeficiente trinomial $\binom{n}{x_1\ x_2\ x_3}$. En consecuencia, el producto de estos dos factores proporciona la probabilidad trinomial; la fórmula siguiente, llamada **fórmula de probabilidad trinomial**, puede usarse para determinar probabilidades asociadas con una experimento trinomial.

Fórmula de probabilidad trinomial

$$P(x_1, x_2, x_3) = \binom{n}{x_1\ x_2\ x_3} p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3} \qquad (6.9)$$

donde $x_1 + x_2 + x_3 = n$ y $p_1 + p_2 + p_3 = 1$

EJEMPLO 6.10

Refiérase al ejemplo 6.8; suponga que la variable aleatoria X_1 representa el número de pacientes que se benefician con el tratamiento, que la variable aleatoria X_2 es la cantidad de pacientes que no se benefician, y que la variable aleatoria X_3 son los pacientes que presentan una reacción adversa. Vea: El número de pacientes que presentan reacción adversa puede escribirse como $(2 - X_1 - X_2)$. Las tres variables aleatorias pueden tomar cualquiera de los valores 0, 1 y 2 si satisfacen la desigualdad $0 \le 2 - X_1 - X_2 \le 2$. Si usamos la fórmula de probabilidad trinomial, podemos obtener la distribución de probabilidad para el experimento, las probabilidades se listan en la tabla 6.5. Advierta que la suma de las probabilidades es igual a 1.

TABLA 6.5

Distribución de probabilidad trinomial para $n = 2$, $p_1 = 1/2$, $p_2 = 1/3$ y $p_3 = 1/6$

(x_1, x_2, x_3)	$P(x_1, x_2, x_3)$
(0, 0, 2)	$\frac{2!}{0!\,0!\,2!}\left(\frac{1}{2}\right)^0\left(\frac{1}{3}\right)^0\left(\frac{1}{6}\right)^2 = \frac{1}{36}$
(0, 1, 1)	$\frac{2!}{0!\,1!\,1!}\left(\frac{1}{2}\right)^0\left(\frac{1}{3}\right)^1\left(\frac{1}{6}\right)^1 = \frac{1}{9}$
(0, 2, 0)	$\frac{2!}{0!\,2!\,0!}\left(\frac{1}{2}\right)^0\left(\frac{1}{3}\right)^2\left(\frac{1}{6}\right)^0 = \frac{1}{9}$
(1, 0, 1)	$\frac{2!}{1!\,0!\,1!}\left(\frac{1}{2}\right)^1\left(\frac{1}{3}\right)^0\left(\frac{1}{6}\right)^1 = \frac{1}{6}$
(1, 1, 0)	$\frac{2!}{1!\,1!\,0!}\left(\frac{1}{2}\right)^1\left(\frac{1}{3}\right)^1\left(\frac{1}{6}\right)^0 = \frac{1}{3}$
(2, 0, 0)	$\frac{2!}{2!\,0!\,0!}\left(\frac{1}{2}\right)^2\left(\frac{1}{3}\right)^0\left(\frac{1}{6}\right)^0 = \frac{1}{4}$
	1.0

APLICACIÓN 6.15

Del ejemplo 6.8, si cada uno de cinco pacientes nuevos recibe el tratamiento, ¿cuál es la probabilidad de que dos de ellos se beneficien y uno presente efectos colaterales adversos?

Solución: Si dos se benefician con el tratamiento y otro tiene complicaciones, entonces debe haber dos pacientes que no mejoran; por tanto, queremos determinar la probabilidad $P(2, 2, 1)$. Note que $x_1 = 2$, $x_2 = 2$ y $x_3 = 1$. Si usamos la fórmula (6.9) tendremos:

$$
\begin{aligned}
P(x_1, x_2, x_3) &= \binom{n}{x_1\ x_2\ x_3} p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3} \\
P(2, 2, 1) &= \binom{5}{2\ 2\ 1} \left(\frac{1}{2}\right)^2 \left(\frac{1}{3}\right)^2 \left(\frac{1}{6}\right)^1 \\
&= \frac{5!}{2!\,2!\,1!} \left(\frac{1}{4}\right)\left(\frac{1}{9}\right)\left(\frac{1}{6}\right) \\
&= (30)\left(\frac{1}{216}\right) = \frac{5}{36} \approx 0.139 \quad \blacksquare
\end{aligned}
$$

Experimentos multinomiales

Si un intento en un experimento multinomial tiene más de tres resultados, las fórmulas (6.8) y (6.9) pueden generalizarse fácilmente a fórmulas para el **coeficiente multinomial** y la **fórmula de probabilidad multinomial.**

Coeficiente multinomial

$$\binom{n}{x_1\ x_2\ x_3 \cdots x_k} = \frac{n!}{x_1!\,x_2!\,x_3! \cdots x_k!} \qquad (6.10)$$

donde $x_1 + x_2 + x_3 + \ldots + x_k = n$

Fórmula de probabilidad multinomial

$$P(x_1, x_2, \ldots, x_k) = \binom{n}{x_1\ x_2\ x_3 \cdots x_k} p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3} \cdots p_k^{x_k} \qquad (6.11)$$

donde $x_1 + x_2 + \ldots + x_k = n$ y $p_1 + p_2 + \ldots + p_k = 1$

APLICACIÓN 6.16

Los miembros de una organización educativa se dividen de acuerdo con una de cuatro clasificaciones políticas: republicano registrado, demócrata registrado, independiente registrado y no registrado. Los porcentajes de educadores republicanos registrados, demócratas registrados, independientes registrados y no registrados son: 30%, 40%, 10% y 20%, respectivamente. ¿Cuál es la probabilidad de que en una muestra aleatoria de diez educadores haya cinco republicanos registrados, cuatro demócratas registrados y un independiente registrado?

Solución: Este es un experimento multinomial con cuatro categorías de resultados. Para $n = 10$, representemos por la variable aleatoria X_1 el número de republicanos registrados, por X_2 a los demócratas registrados, por X_3 a los independientes registrados y por X_4 a educadores no registrados; también, sea $p_1 = 0.3$, $p_2 = 0.4$, $p_3 = 01$ y $p_4 = 0.2$. Necesitamos determinar la probabilidad de $P(X_1 = 5, X_2 = 4, X_3 = 1, X_4 = 0)$.

$$P(5, 4, 1, 0) = \binom{10}{5\ 4\ 1\ 0}(0.3)^5(0.4)^4(0.1)^1(0.2)^0$$
$$= \frac{10!}{5!\,4!\,1!\,0!}(0.00243)(0.0256)(0.1)(1) = 0.0078 \quad \blacksquare$$

GRUPO DE EJERCICIOS 6.4

Habilidades básicas

1. Evalúe los coeficientes multinomiales siguientes:
   a) $\binom{7}{3\ 2\ 0\ 2}$
   b) $\binom{8}{1\ 2\ 3\ 1\ 1}$
2. Haga lo mismo que en el anterior:
   a) $\binom{9}{1\ 0\ 0\ 8}$
   b) $\binom{10}{0\ 0\ 5\ 3\ 2}$
3. ¿Es válido $\binom{9}{1\ 2\ 0\ 7}$ como coeficiente multinomial? ¿Por qué?
4. ¿Es válido $\binom{10}{1\ 2\ 3\ 2}$ como coeficiente multinomial? ¿Por qué?
5. Suponga que X_1 y X_2 son variables aleatorias trinomiales con $p_1 = 0.3$ y $p_2 = 0.2$. Si $n = 6$, encuentre las probabilidades siguientes:
   a) $P(X_1 = 3 \text{ y } X_2 = 2)$
   b) $P(X_1 = 1 \text{ y } X_2 = 1)$
   c) $P(X_1 = 0 \text{ y } X_2 = 0)$
   d) $P(X_1 = 2 \text{ y } X_2 = 4)$
6. Suponga que X_1 y X_2 son variables aleatorias trinomiales con $p_1 = 0.4$ y $p_2 = 0.3$. Si $n = 7$, encuentre las probabilidades siguientes:
   a) $P(X_1 = 2 \text{ y } X_2 = 2)$
   b) $P(X_1 = 3 \text{ y } X_2 = 1)$
   c) $P(X_1 = 0 \text{ y } X_2 = 0)$
   d) $P(X_1 = 3 \text{ y } X_2 = 4)$

Más aplicaciones

7. Un dado se lanza cinco veces. Encuentre la probabilidad de obtener:
   a) tres unos, un 1 y un 6;
   b) tres cuatros y dos números tres;
   c) únicamente unos.
8. Un dado se lanza siete veces. Encuentre la probabilidad de obtener:
   a) dos unos, dos números dos, un 3 y dos cuatros.
   b) un 4, un 2, un 3, un 4, un 5 y dos números seis.
   c) dos unos, un 2, un 3, un 4 y dos números seis.
9. Una caja contiene un gran número de canicas, 50% de ellas son blancas, 30% rojas y 20% negras. Suponga que se eligen al azar diez canicas de la caja. ¿Cuál es la probabilidad de que:
   a) haya cinco canicas rojas, dos blancas y tres negras?
   b) tengamos tres canicas blancas y cinco rojas?
   c) logremos cinco canicas blancas y cinco canicas rojas?
10. Las llamadas para reparar refrigeradores caen en las siguientes cuatro categorías: escurrimiento de freón, fallas en el compresor, fusibles fundidos y varios. Con base en los registros las probabilidades asociadas con esas cuatro categorías son 0.4, 0.2, 0.3 y 0.1, respectivamente. En las siguientes seis solicitudes de servicios, ¿cuál es la probabilidad de que:
    a) la mitad involucre lugar de freón y la otra mitad fusibles fundidos?;
    b) dos órdenes de servicio impliquen fallas del compresor, una a fusibles fundidos y tres a problemas diversos?;
    c) cuatro tengan que ver con fallas del compresor y dos con fusibles fundidos?

11. Los miembros de una organización educativa se clasifican de acuerdo con una de tres clases políticas: republicanos registrados, demócratas registrados y no registrados; los porcentajes de educadores republicanos registrados, demócratas registrados y no registrados son 50%, 40% y 10%, respectivamente. ¿Cuál es la probabilidad de que en una muestra aleatoria de cinco educadores haya:
    a) cinco republicanos registrados?;
    b) tres republicanos registrados, un demócrata registrado y dos no registrados?
    c) tres republicanos registrados y dos demócratas registrados?
12. Una clínica de salud se especializa en usar un método particular para tratar pacientes artríticos; sus registros muestran que la mitad de todos los pacientes se benefician con el tratamiento, que la tercera parte no presentan reacciones al mismo, y que una sexta parte sufren complicaciones colaterales adversas. Entre los siete pacientes siguientes, ¿cuál es la probabilidad de que:
    a) tres se beneficien con el tratamiento y cuatro no presenten efectos?;
    b) todos se alivien?;
    c) cuatro se alivien y tres no tengan reacción?
13. Suponga que se tiran 18 veces un dado rojo y uno verde.
    a) Encuentre la probabilidad de que siete de las veces el dado rojo tenga un número mayor y que cuatro veces empaten.
    b) Calcule la probabilidad de que empaten cuatro veces.
14. En una universidad grande se va a escoger un comité de 12 miembros, supongamos que es un muestreo con remplazo. La universidad tiene un 30% de estudiantes de raza negra, 40% blancos y 30% latinos.
    a) Encuentre la probabilidad de que el comité tenga tres latinos y cuatro blancos.
    b) ¿Cuál es la probabilidad de que haya el mismo número de latinos, de blancos y de raza negra.
15. Si un par de dados se lanza ocho veces, ¿cuál es la probabilidad de obtener tres veces una suma de 7, un par una vez y cualquier otra combinación cuatro veces?
16. Una universidad juega diez partidos de futbol americano durante una temporada. ¿De cuántas formas puede terminar el equipo de temporada con cuatro ganados, cinco perdidos y un empate?
17. ¿De cuántas maneras pueden viajar once personas usando tres coches en los que caben 2, 4 y 5 pasajeros, respectivamente?
18. Si se lanzan cuatro dados de uso común, ¿cuál es la probabilidad de que todos presenten un número par?
19. Para el supuesto anterior, ¿cuál es la probabilidad de que dos dados muestren al caer un número par y dos un número impar?

Un paso más allá

20. Demuestre que el coeficiente binomial $\binom{n}{x_1}$ puede escribirse como $\binom{n}{x_1\ x_2}$, donde $x_1 + x_2 = n$.

SECCIÓN 6.5 *Distribuciones hipergeométricas*

En esta sección queremos considerar experimentos que obedezcan tres de las cuatro propiedades de un experimento binomial; se debilitará la propiedad de independencia entre los intentos, es decir, los intentos individuales se considerarán dependientes, el experimento resultante se llamará **experimento hipergeométrico.** Los experimentos hipergeométricos se usan comúnmente cuando el muestreo se hace sin remplazo; considere el ejemplo siguiente de un experimento hipergeométrico.

EJEMPLO 6.11

Suponga que una urna contiene dos tablitas rojas y tres negras; un experimento consiste en sacar de la urna tres tablitas al azar, sin remplazo, donde el experimentador se interesa en el número de tablitas rojas elegidas.

Si la variable aleatoria X denota el número de tablitas rojas seleccionadas, entonces sus valores son $\{0, 1, 2\}$; determinemos la distribución de probabilidad de X: el número de formas de seleccionar una muestra de tres tablitas de una colección de cinco es precisamente, $\binom{5}{3} = 10$. Por tanto, un espacio muestral consta de diez resultados igualmente posibles.

a) Determinemos el número de muestras de tres tablitas que no contienen tablitas rojas y sí tres negras; solo hay una manera de no sacar tablitas rojas de entre dos tablitas rojas $\binom{2}{0} = 1$, y también una sola forma de sacar tres tablitas negras de entre tres tablitas negras, $\binom{3}{3} = 1$.

   Por tanto, la probabilidad de sacar una muestra de tamaño tres sin tablitas rojas es:

$$P(X = 0) = \frac{\binom{2}{0}\binom{3}{3}}{\binom{5}{3}} = \frac{1}{10}$$

b) El número de muestras de tamaño tres que contienen una tablita roja, se encuentra multiplicando el número de formas de sacar una tablita roja de dos, y el número de formas de sacar dos tablitas negras de tres.

   Las maneras de seleccionar una tablita roja de dos es $\binom{2}{1} = 2$ y el número de formas de sacar dos tablitas negras de tres es $\binom{3}{2} = 3$; en consecuencia, hay (2)(3) = 6 muestras que tienen exactamente una tablita roja.

   Por tanto, la probabilidad de obtener una muestra de tamaño tres con una tablita roja es:

$$P(X = 1) = \frac{\binom{2}{1}\binom{3}{2}}{\binom{5}{3}} = \frac{6}{10} = \frac{3}{5}$$

c) El número de formas de obtener dos tablitas rojas es igual al número de formas de obtener las mismas tablitas multiplicando dicho número de tablitas del mismo color por el número de formas de obtener una tablita negra de tres tablitas de este color.

   El número de formas de seleccionar dos tablitas rojas es $\binom{2}{2} = 1$, y de seleccionar una tablita negra es $\binom{3}{1} = 3$.

   Así, el número de formas de sacar una muestra de tres tablitas con exactamente dos tablitas rojas es $1{\cdot}3 = 3$.

   La probabilidad de obtener una muestra de tres tablitas, sin remplazo, en la que haya dos tablitas rojas, es:

$$P(X = 2) = \frac{\binom{2}{2}\binom{3}{1}}{\binom{5}{3}} = \frac{3}{10}$$

Note que: $P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{10} + \frac{6}{10} + \frac{3}{10} = \frac{10}{10} = 1.$

En la tabla 6.6 vemos la distribución de probabilidad para este experimento hipergeométrico. Advierta el patrón con las entradas numéricas.

TABLA 6.6

Distribución de probabilidad para la distribución hipergeométrica que contiene dos tablitas rojas y tres tablitas negras

x	$P(X = x)$
0	$\dfrac{\binom{2}{0}\binom{3}{3}}{\binom{5}{3}} = \dfrac{1}{10}$
1	$\dfrac{\binom{2}{1}\binom{3}{2}}{\binom{5}{3}} = \dfrac{6}{10}$
2	$\dfrac{\binom{2}{2}\binom{3}{1}}{\binom{5}{3}} = \dfrac{3}{10}$

Para generalizar el método que usamos en el ejemplo 6.11, supongamos que vamos a escoger n_1 objetos de una clase, llamados *éxitos*, y n_2 objetos de otra clase llamados *fracasos*. La selección es sin remplazo, y estamos interesados en el número de éxitos escogidos; el número total de formas de escoger n objetos de $n_1 + n_2$ objetos es el coeficiente binomial $\binom{n_1 + n_2}{n}$ y el número de formas de escoger x éxitos y $(n - x)$ fracasos es el producto $\binom{n_1}{x}\binom{n_2}{n - x}$. En consecuencia, la probabilidad de elegir x éxitos en n intentos está dada por la fórmula siguiente, llamada **fórmula de probabilidad hipergeométrica.**

Fórmula de probabilidad hipergeométrica

$$P(x) = \frac{\binom{n_1}{x}\binom{n_2}{n - x}}{\binom{n_1 + n_2}{n}} \quad \text{para x = 0, 1, 2, 3, ...o } n \qquad (6.12)$$

donde $n \leq n_1 + n_2$. La fórmula (6.12) se sigue del patrón establecido en la tabla 6.6.

APLICACIÓN 6.17

Se embarcan abanicos eléctricos en lotes de diez; antes de aceptar un lote, un inspector elige tres de esos abanicos y los inspecciona, si ninguno de los abanicos probados está defectuoso, el lote se acepta; si uno o más salen con defectos, revisan todo el lote. Suponga que hay dos abanicos deficientes. ¿Cuál es la probabilidad de que se necesite un 100% de inspección?

Solución: Denotemos por la variable X el número de abanicos defectuosos en un lote de diez. Usemos la fórmula (6.12) para determinar $P(X = 0)$. Notemos que $x = 0$, $n_1 = 2$, $n_2 = 8$ y $n = 10$.

$$P(x) = \frac{\binom{n_1}{x}\binom{n_2}{n-x}}{\binom{n_1+n_2}{n}}$$

$$P(0) = \frac{\binom{2}{0}\binom{8}{3}}{\binom{10}{3}}$$

$$= \frac{56}{120} = 0.467$$

Se necesitará un cien por ciento de inspección si $x \geq 1$. Como $P(X \geq 1) = 1 - P(X = 0)$, dicha probabilidad es de $1 - 0.467 = 0.533$. ■

GRUPO DE EJERCICIOS 6.5

Habilidades básicas

1. Suponga que una urna contiene cinco tablitas rojas y siete negras; un experimento consiste en elegir tres tablitas al azar, sin remplazo. Denotando con la variable aleatoria X el número de tablitas rojas elegidas, construya la distribución de probabilidad para X.

2. Una urna contiene ocho tablitas rojas y siete negras; un experimento consiste en sacar cinco tablitas al azar, sin remplazo. Designamos con la variable aleatoria X el número de tablitas rojas seleccionadas; construya la distribución de probabilidad para X.

3. Se debe elegir un comité de seis miembros de entre cuatro matemáticos y seis expertos en computación; localice la distribución de probabilidad para la cantidad de matemáticos que habrá en el comité.

4. Se va a elegir aleatoriamente un comité de cuatro personas de entre ocho matemáticos y siete expertos en computación. ¿Cuál será la distribución de probabilidad para el número de matemáticos integrantes del comité?

Más aplicaciones

5. Una caja con 24 calculadoras contiene 4 que están defectuosas; si se eligen 4 al azar de esa caja, ¿cuál es la probabilidad de que:
   a) tres estén defectuosas?
   b) a lo mucho una resulte defectuosa?
   c) las 4 estén defectuosas?

6. De un equipo de basquetbol de 12 miembros se elige un conjunto de 5 para representarlo en una competencia, como lo establecen las reglas.
   a) ¿Cuál es la probabilidad de que entre los 5 elegidos se incluyan los dos mejores jugadores del equipo?
   b) ¿Cuál es la probabilidad que los 5 elegidos sean los novatos?
   c) ¿Qué probabilidad existe de que ninguno de los seleccionados sea novato?

7. Un inspector de control de calidad examina una muestra aleatoria de cinco baterías de cada caja con 24 piezas que sale de la línea de ensamble; si, de hecho, una caja contiene 4 baterías defectuosas, encuentre la probabilidad de que en una muestra aleatoria de 5 piezas:
   a) ninguna esté defectuosa.
   b) haya 2 baterías en mal estado.
   c) haya almenos 1 pieza mala.

8. Un auditor de recaudación de impuestos selecciona al azar 4 solicitudes de devolución de impuestos de entre 15 presentadas, si 5 solicitudes contienen deducciones ilegales, ¿cuál es la probabilidad de que el auditor examine:
   a) exactamente una solicitud legal?;
   b) 4 solicitudes ilegales?;
   c) 2 solicitudes no válidas?

9. Un jardinero planta cinco bulbos seleccionados al azar de una caja que contiene seis de narcisos y cinco de tulipanes. ¿Cuál es la probabilidad de que plante:

a) tres bulbos de narcisos y dos de tulipanes?;
b) cinco bulbos de tulipanes?

10. Suponga que en un almacén hay 20 llantas de las que 3 están defectuosas. Si se toman aleatoriamente 5 llantas de ese almacén, ¿cuál es la probabilidad de que:
a) al menos 2 de ellas estén defectuosas?;
b) se elija una llanta en mal estado?

11. Un consejo de estudiantes está integrado por dos estudiantes de primer año, cinco de segundo, ocho de tercero y diez estudiantes de cuarto año; si 4 de sus miembros son elegidos al azar para asistir a una convención nacional, ¿cuál es la probabilidad de que cada una de las clases sea representada en la convención?

Un paso más allá

12. La *media de una variable aleatoria hipergeométrica* es $\mu = n\,n_1/(n_1+n_2)$, donde n es el tamaño de la muestra, n_1 el total de éxitos y n_2 representa a todos los fracasos. Verifique esto para la variable aleatoria dada en el ejercicio 1.

13. Refiérase al ejercicio 1. Use la fórmula (5.18) de la sección 5.7 para determinar la varianza de X.

14. La *varianza de una variable aleatoria hipergeométrica* está dada por:

$$\sigma_X^2 = \frac{n\,n_1\,n_2(n_1+n_2-n)}{(n_1+n_2)^2(n_1+n_2-1)}$$

Use esta fórmula para determinar la varianza de la variable aleatoria dada en el ejercicio 1, y contraste la respuesta con la obtenida en el ejercicio 13.

15. La distribución binomial puede usarse para aproximar la **distribución hipergeométrica** si $20n \leq n_1 + n_2$; se va a escoger una muestra de cinco estudiantes de entre un conjunto de mil, si en su conjunto hay un 25% de estudiantes de ingeniería, ¿cuál es la probablidad de que dos de los cinco alumnos sean estudiantes de ingeniería si la muestra se obtuvo:
a) sin remplazo;
b) con remplazo;
c) compare sus respuestas a) y b).

16. Suponga que en una ciudad con 10,000 votantes, 6,000 de ellos ven cierto programa de televisión; si se saca una muestra aleatoria de cinco votantes, ¿cuál es la probabilidad de que tres de los cinco vean el programa?

SECCIÓN 6.6 *Distribución de Poisson*

El cálculo de probabilidades binomiales puede ser tedioso, especialmente si el número de intentos es grande. Cuando el número de intentos es grande ($n \geq 100$) y $\mu = np$ es pequeño ($\mu < 10$), las probabilidades binomiales pueden aproximarse mediante una forma particular de la función de probabilidad de Poisson. La **distribución de probabilidad de Poisson** se define por la **fórmula de probabilidad de Poisson:**

Fórmula de probabilidad de Poisson

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad (6.13)$$

donde el parámetro $\lambda > 0$, $e \approx 2.71828$ y $x = 1, 2, 3, \ldots$. La media de una variable aleatoria es $\mu = np$; si en la fórmula (6.13) sustituimos np en lugar de λ, que es la **media de una variable aleatoria de Poisson**, entonces obtenemos la siguiente fórmula, que puede usarse para aproximar probabilidades binomiales.

$$P(x) = \frac{(np)^x e^{-np}}{x!} \tag{6.14}$$

APLICACIÓN 6.18

Si la probabilidad de que un aeroplano comercial choque durante el vuelo es de 0.0001, ¿cuál es la probabilidad de que choquen 10 de los siguientes 30,000 aeroplanos?

Solución: Consideremos un aeroplano como un intento y un choque como un éxito, luego determinemos la probabilidad binomial $P(10)$, donde $n = 30{,}000$ y $p = 0.0001$. Si usamos la fórmula de probabilidad binomial (6.2) obtenemos:

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$P(10) = \binom{30{,}000}{10} (0.0001)^{10} (0.9999)^{29{,}990}$$

Esta fórmula es extremadamente difícil para hacer cálculos sin la ayuda de una computadora. Mediante MINITAB, determinamos $P(10) = 0.0008$, vea la pantalla 6.4 para conocer una lista de las órdenes usadas y la respuesta correspondiente.

Pantalla 6.4

```
MTB > PDF 10;
SUBC > BINOMINAL N=30000, P=0.0001.
      K        P(X = K)
    10.00       0.0008
```

Como $n = 30{,}000$ y $\mu = np = (30{,}000)(0.0001) = 3$, podemos usar la distribución de probabilidad de Poisson con $\lambda = 3$ para aproximar $P(10)$; con la ayuda de una calculadora encontramos que:

$$P(x) = \frac{(np)^x e^{-np}}{x!}$$

$$P(10) = \frac{(3)^{10} e^{-3}}{10!} = 0.00081$$

lo cual, hasta la cuarta cifra decimal, concuerda con el resultado encontrado antes por medio de MINITAB.

Para evaluar $P(10) = (3)^{10} e^{-3}/10!$ directamente, podemos usar la tabla 2 del apéndice B o MINITAB; si optamos por la tabla, localizamos la columna etiquetada por $\lambda = 3$ a lo largo de la parte superior y el renglón marcado con $X = 10$ a lo largo del lado izquierdo, la probabilidad es el valor en el que se intersecan el renglón y la columna; ese valor es 0.0008. Si usamos MINITAB, la pantalla 6.5 contiene las órdenes usadas para obtener $P(10)$ y la respuesta correspondiente.

Pantalla 6.5

```
MTB > PDF 10;
SUBC > POISSON 3.
        K       P(X = K)
      10.00     0.0008
      MTB>
```

■

Las impresiones de MINITAB siguientes muestran qué tan cercana a la distribución de probabilidad binomial, con $n = 100$ y $p = 0.05$, es la aproximación dada por la distribución de probabilidad de Poisson con $\lambda = np = 5$.

```
MTB > PDF;
SUBC > BINOMINAL 100 0.05.

BINOMINAL WITH N = 100 P = 0.050000
   K          P( X = K)
   0          0.0059
   1          0.0312
   2          0.0812
   3          0.1396
   4          0.1781
   5          0.1800
   6          0.1500
   7          0.1060
   8          0.1649
   9          0.0349
  10          0.0167
  11          0.0072
  12          0.0028
  13          0.0010
  14          0.0003
  15          0.0001
  16          0.0000
```

```
MTB> PDF;
SUBC > POISSON 5.

POISSON WITH MEAN = 5.000
   K      P( X = K)
   0      0.0067
   1      0.0337
   2      0.0842
   3      0.1404
   4      0.1755
   5      0.1755
   6      0.1462
   7      0.1044
   8      0.0653
   9      0.0363
  10      0.0181
  11      0.0082
  12      0.0034
  13      0.0013
  14      0.0005
  15      0.0002
  16      0.0000
```

La fórmula de probabilidad binomial se usa en experimentos binomiales para determinar la probabilidad de obtener un número específico de éxitos en un número fijo de intentos, cuando se conoce la probabilidad de un éxito aislado. Por otro lado, la fórmula de probabilidad de Poisson (6.13) se usa para determinar la probabilidad del número de éxitos que tienen lugar por unidad de tiempo o de espacio, más que por un número específico de intentos; además de utilizarla para aproximar probabilidades binomiales cuando $n \geq 100$ y $\mu = np < 10$; pudiendo usarla también para determinar probabilidades asociadas con experimentos referentes a fenómenos aleatorios tales como el número de

- llamadas telefónicas (0, 1, 2, etc.,) que entran a un conmutador durante cierto periodo;
- errores en una página típica impresa por cierto editor de libros;

- hamburguesas gigantes vendidas a la hora del almuerzo;
- pacientes que llegan a la sala de emergencias de un hospital particular durante un cierto día;
- peces capturados en un gran lago durante un día;
- defectos de un rollo de tela;
- accidentes por hora en cierta parte de una carretera;
- componentes defectuosos encontrados en una automóvil recién fabricado;
- pedazos de chocolate contenidos en un tipo de galleta con chocolate;
- clientes que llegan a la caja registradora en determinado horario;
- solicitudes procesadas en un día por una empresa aseguradora.

El número de pedazos de chocolate observados en una galleta seleccionada al azar, es un ejemplo de fenómeno relacionado con el espacio, otro ejemplo es la cantidad de errores en una página impresa seleccionada al azar.

Si un experimento referente a procesos relacionados con el espacio satisface las tres condiciones siguientes, se denomina **experimento de Poisson:**

1. El número promedio de veces (μ) que ocurre un éxito por cada unidad de tiempo o de espacio es constante.
2. La probabilidad de más de un suceso es una unidad de tiempo o de espacio muy pequeña.
3. El número de acontecimientos en intervalos ajenos de tiempo o de espacio son independientes unos de otros.

Si la variable aleatoria X representa el número de sucesos resultantes de un experimento de Poisson, se le llama **variable aleatoria de Poisson**. La fórmula de probabilidad de Poisson se usa para calcular probabilidades asociadas con una variable aleatoria de Poisson.

APLICACIÓN 6.19

Representemos por la variable aleatoria X el número de trozos de chocolate hallados en una galleta con chocolate sacada aleatoriamente en la fábrica de galletas. Explique por qué es razonable suponer que X es una variable aleatoria de Poisson.

Solución: Es razonable suponer que una galleta puede subdividirse en n regiones con áreas iguales, en forma tal que cada una contenga a lo más un trozo de chocolate; un intento es igual a inspeccionar una región para determinar si contiene trozos de chocolate, observar una región que contenga un trozo de chocolate equivale a un éxito. Si p denota la probabilidad de que una región contenga un trozo de chocolate, es razonable suponer que p es constante de un intento a otro y que los intentos son independientes; asumimos que el tamaño de una galleta es grande con respecto al de un trozo de chocolate, y también que hay pocos trozos de chocolate en cada galleta; esto es, creemos que n es grande y p pequeño. Bajo las hipótesis anteriores, tenemos un experimento binomial, por tanto, podemos aproximar las probabilidades binomiales usando la fórmula de probabilidad de Poisson. En consecuencia, es posible suponer que X es una variable aleatoria de Poisson. ■

Se puede demostrar que una variable aleatoria de Poisson tiene su media igual a su varianza. Esto es,

Propiedades de una variable aleatoria de Poisson X con un parámetro $\lambda > 0$.

1. $\mu_x = \lambda$ 2. $\sigma_x^2 = \lambda$

APLICACIÓN 6.20

Un banco recibe un promedio de tres cheques sin fondos por día. ¿Cuál es la probabilidad de que en cierto día reciba cuatro o cinco cheques sin fondos?

Solución: Notemos que $\lambda = \mu = 3$, y usemos la fórmula (6.11). La solución es $P(4) + P(5)$, donde

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Con la ayuda de una calculadora o de la tabla 2 del apéndice B, encontramos que:

$$P(4) = \frac{3^4 e^{-3}}{4!} = 0.168$$

y

$$P(5) = \frac{3^5 e^{-3}}{5!} = 0.101$$

En consecuencia, $P(4) + P(5) = 0.168 + 0.101 = 0.269$ ■

APLICACIÓN 6.21

Las llamadas recibidas por un conmutador siguen una distribución de probabilidad de Poisson y, en promedio, entran seis por hora. ¿Cuál es la probabilidad de que ocurran dos o más llamadas en una hora?

Solución: Como la suma de probabilidades de Poisson es 1, tenemos

$$P(X = 0) + P(X = 1) + P(X \geq 2) = 1$$

Por tanto, la probabilidad de que haya dos o más llamadas de una hora es igual a:

$$P(X \geq 2) = 1 - [P(0) + P(1)]$$

Podemos usar la fórmula (6.11) y una calculadora o la tabla 2 del apéndice B para calcular valores de $P(0)$ y $P(1)$.

$$P(0) = \frac{6^0 e^{-6}}{0!} = 0.0025$$

$$P(1) = \frac{6^1 e^{-6}}{1!} = 0.0149$$

Por tanto, la probabilidad de dos o más llamadas en una hora es:

$$1 - (0.0025 + 0.0149) = 1 - 0.0174 = 0.9826$$ ■

GRUPO DE EJERCICIOS 6.6

Habilidades básicas

1. Para cada caso, calcule la probabilidad de Poisson asociada con los valores mostrados:
   a) $x = 4, \lambda = 3$;
   b) $x = 7, \lambda = 0.6$;
   c) $x = 10, \lambda = 10$;
   d) $x = 5, \lambda = 1$
2. Haga igual que el anterior:
   a) $x = 5, \lambda = 4$;
   b) $x = 2, \lambda = 9$;
   c) $x = 1, \lambda = 1$;
   e) $x = 0, \lambda = 1.5$
3. Sea X una variable aleatoria binomial con $n = 500$ y $p = 0.01$, determine una aproximación de Poisson para:
   a) $P(X = 4)$;
   b) $P(X < 1)$;
   c) $P(X = 10)$;
   d) $P(X > 1)$.
4. Si X es una variable aleatoria binomial con $n = 1000$ y $p = 0.002$, calcule una aproximación de Poisson para:
   a) $P(X = 1)$;
   b) $P(X < 2)$;
   c) $P(X = 3)$;
   d) $P(X > 1)$.

Más aplicaciones

5. Ciento cincuenta estudiantes principiantes de leyes presentan el examen para ingresar a la Barra de Abogados de un estado, la probabilidad de que un estudiante de primer año pase ese examen es de 0.05. Aproxime la probabilidad de que:
   a) ocho estudiantes de leyes pasen el examen;
   b) al menos tres aspirantes pasen el examen;
   c) a lo más dos estudiantes aprueben el examen;
   d) al menos dos estudiantes califiquen positivo el examen.
6. Doscientos estudiantes de medicina presentan el examen de admisión al Colegio de Medicina (MCAT por sus siglas en inglés); la probabilidad de que alguien obtenga un puntaje promedio de 16 en el MCAT es 0.001. Aproxime la probabilidad de que:
   a) dos estudiantes obtengan un promedio de 16;
   b) al menos dos estudiantes consigan un promedio de 16;
   c) a lo más dos aspirantes logren el promedio 16;
   d) ninguno obtenga el promedio mencionado.
7. Los registros de policía muestran que hay un promedio de tres accidentes por semana en la Ruta 40; si suponemos que esos percances siguen una distribución de Poisson, determine la probabilidad de que durante cierta semana seleccionada al azar haya:
   a) cuatro accidentes;
   b) cuatro o cinco accidentes;
   c) a lo más tres accidentes;
   d) al menos tres percances.
8. El número de errores por página para un periódico local se distribuye como una variable aleatoria de Poisson con parámetro $\lambda = 2.5$. Estime la probabilidad de que una página elegida al azar tenga:
   a) cinco errores;
   b) dos o cuatro errores;
   c) al menos dos fallas;
   d) a lo más dos equívocos.
9. Suponga que el número de cartas perdidas en el correo en un día es una variable aleatoria de Poisson con parámetro $\lambda = 4$. ¿Cuál es la probabilidad de que en un día equis:
   a) se pierdan tres cartas en el correo?;
   b) se extravíen cuatro o cinco?;
   c) al menos una carta desaparezca en el correo?;
   d) se pierdan a lo más dos cartas en la oficina postal?
10. El número de coches que pasan por cierto puente es una variable aleatoria con $\lambda = 10$. Diga la probabilidad de que, para un día en particular,

a) cinco coches pasen por el puente.
b) ningún coche pase.
c) al menos cinco coches crucen por el puente.
d) a lo más dos coches utilicen el puente.

Un paso más allá

11. El número de llamadas telefónicas que entran a un conmutador local durante cualquier intervalo de tiempo de longitud t medido en minutos, es una variable aleatoria de Poisson con parámetro $\lambda = 5$.
    a) Estime la probabilidad de que entren 12 llamadas al conmutador durante un periodo específico de dos minutos.
    b) Encuentre la probabilidad de que se reciba al menos una llamada por minuto durante un periodo de dos minutos.

12. Si X es una variable aleatoria de Poisson con parámetro λ, y si $P(X = 0) = 0.2$, encuentre $P(X > 2)$.

13. Si X es una variable aleatoria de Poisson con parámetro λ, encuentre el valor de a tal que $P(X = a)$ es lo más grande posible.

14. Si X es una variable aleatoria de Poisson tal que $3P(X = 2) = 2P(X = 1)$, encuentre:
    a) $P(X = 0)$.
    b) $P(X = 3)$.

RESUMEN DEL CAPÍTULO

En este capítulo estudiamos experimentos binomiales y sus propiedades; aprendimos cómo asociar los resultados de un experimento binomial con una variable aleatoria discreta binomial; cómo calcular probabilidades asociadas a una variable aleatoria binomial y cómo encontrar la media, la varianza y la desviación estándar de las distribuciones de probabilidad binomiales. Estudiamos también experimentos multinomiales, hipergeométricos y de Poisson como variaciones de experimentos binomiales, con cada uno usamos la función de probabilidad para calcular probabilidades.

REVISIÓN DEL CAPÍTULO

▪ *TÉRMINOS GENERALES* ▪

Los términos siguientes usados en el capítulo se han mezclado para proporcionarle una mejor práctica de revisión, para cada uno dé una definición con sus propias palabras, luego compare sus respuestas con las definiciones del capítulo.

coeficiente binomial

distribución binomial

experimento binomial

fórmula de probabilidad binomial

media de una distribución binomial

coeficiente multinomial

experimento multinomial

fórmula de probabilidad multinomial

experimento hipergeométrico

distribución hipergeométrica

distribución binomial empírica

desviación estándar de una distribución binomial

intento

fórmula de probabilidad de Poisson

varianza de una distribución binomial

variable aleatoria de Poisson

media de una variable aleatoria de Poisson

fórmula de probabilidad hipergeométrica

experimento de Poisson

distribución de probabilidad de Poisson

experimentos trinomiales

coeficiente trinomial

fórmula de probabilidad trinomial

fórmula del coeficiente trinomial

SÍMBOLOS IMPORTANTES

n, el número de intentos

E, éxito

F, fracaso

p, la probabilidad de éxito

$1-p$, la probabilidad de fracaso

x, el número de éxitos.

$E(x)$, número promedio de éxitos

$\binom{n}{x}$, coeficiente binomial

$\binom{n}{x_1\ x_2\ x_3}$, coeficiente trinomial

$\binom{n}{x_1\ x_2\ x_3 \cdots x_k}$, coeficiente multinomial

n_1, número de éxitos para un experimento hipergeométrico

n_2, número de fracasos para un experimento hipergeométrico

λ, media de una variable aleatoria de Poisson

HECHOS Y FÓRMULAS IMPORTANTES

Coeficiente binomial:
$\binom{n}{x} = \frac{n!}{x!(n-x)!}$. (6.1)
$\binom{n}{x} = \binom{n}{n-x}$.

Fórmula de probabilidad binomial:
$P(x) = \binom{n}{x}p^x(1-p)^{n-x}$. (6.2)

Media de una distribución binomial: $\mu = np$. (6.4)

Varianza de una distribución binomial: $\sigma^2 = np(1-p)$. (6.6)

Desviación estándar de una distribución binomial: $\sigma = \sqrt{np(1-p)}$. (6.7)

Fórmula del coeficiente trinomial:
$\binom{n}{x_1\ x_2\ x_3} = \frac{n!}{x_1!\ x_2!\ x_3!}$. (6.8)

Fórmula de probabilidad trinomial:
$P(x_1, x_2, x_3) = \binom{n}{x_1\ x_2\ x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$,
donde $x_1 + x_2 + x_3 = n$ y $p_1 + p_2 + p_3 = 1$. (6.9)

Coeficiente multinomial:
$\binom{n}{x_1\ x_2\ x_3 \ldots x_k} = \frac{n!}{x_1!\ x_2!\ x_3! \ldots x_k!}$,
donde $x_1 + x_2 + x_3 \ldots + x_k = n$. (6.10)

Fórmula de probabilidad multinomial:
$P(x_1, x_2, \ldots x_k) = \binom{n}{x_1\ x_2\ x_3 \cdots x_k} p_1^{x_1} p_2^{x_2} p_3^{x_3} \ldots p_k^{x_k}$ donde, $n = x_1 + x_2 + \ldots + x_k$ y $p_1 + p_2 + \ldots + p_k = 1$. (6.11)

Fórmula de probabilidad hipergeométrica: $P(x) = \frac{\binom{n_1}{x}\binom{n_2}{n-x}}{\binom{n_1+n_2}{n}}$,
donde, $n \leq n_1 + n_2$. (6.12)

Fórmula de probabilidad de Poisson: $P(x) = \frac{\lambda^x e^{-\lambda}}{x!}, 1 > 0$. (6.13)

EJERCICIOS DE REPASO

1. Treinta y cinco por ciento de todos los automóviles vendidos en Estados Unidos son de fabricación extranjera, se elige aleatoriamente cinco coches nuevos.
   a) Construya una tabla de distribución de probabilidad binominal para el número de coches de fabricación extranjera.
   b) Dibuje una gráfica de rectas verticales para la distribución binominal.
   c) Encuentre la media del número de coches de fabricación extranjera.
   d) Encuentre la desviación estándar de la cantidad de automóviles que son manufacturados fuera de Estados Unidos.
   e) Estime la probabilidad de que a lo más cuatro coches provengan del extranjero.
2. La probabilidad de que un recién nacido sea varón es de 0.5; si en un hospital local nacen ocho niños, encuentre:
   a) la probabilidad de que tres sean varones.
   b) que ninguno sea niño.

c) la probabilidad de que dos sean niñas.

d) el número esperado de niños.

3. Si un jugador de beisbol con un porcentaje de bateo de 0.250 va diez veces al bat en una serie, ¿cuál es la probabilidad de que consiga:
   a) tres hits?;
   b) al menos uno?;
   c) a lo más tres hits?
4. Si se supone que un jugador de gol mandará la pelota a una trampa de arena el 17% del tiempo, ¿cuál es la probabilidad de que mande la pelota a la arena exactamente en cuatro de los primeros nueve hoyos?
5. Refiérase al ejercicio 4. Encuentre el número esperado de veces que la pelota caerá en una trampa de arena.
6. Un cierto tipo de píldoras anticonceptivas tienen una eficacia del 90%, 500 mujeres las usan, ¿cuántos nacimientos no planeados espera usted?
7. Se estima que 85% de todas las plantas caseras están sobrehumidificadas; para diez plantas caseras, encuentre la media y la desviación estándar de la cantidad que está sobrehumidificada, e interprete sus resultados.
8. Si en un saco hay cinco pelotas rojas y 15 negras, y se eligen tres, una por vez y con remplazo, ¿cuál es la probabilidad de que las tres sean rojas?
9. Si una estudiante no preparada presenta un examen de opción múltiple donde cada pregunta tiene cinco opciones y responde adivinando cada una de las 25 preguntas, ¿cuál es la probabilidad de que reciba una calificación de 28%?
10. En promedio, una tienda ha logrado ventas al por mayor de más de 5000 dólares en uno de siete días a lo largo de los diez meses pasados, asumimos que este fenómeno continuará, ¿cuál es la probabilidad de que tenga ventas al por mayor de más de 5,000 dólares al menos cuatro veces en los siguientes siete días?
11. Si X es una variable aleatoria de Poisson con $\lambda = 4$, encuentre:
    a) $P(X = 1)$;
    b) $P(X = 3)$;
    c) $P(X \leq 3)$;
    d) $P(X \geq 4)$.
12. El número de las llamadas telefónicas que entran a un conmutador es, en promedio, una cada tres minutos, si ese total sigue un proceso de Poisson, ¿cuál es la probabilidad de que lleguen dos o más llamadas en un periodo de tres minutos?
13. Si un fabricante de circuitos integrados sabe que el 2% de sus productos están defectuosos, aproxime la probabilidad de que una producción de 150 circuitos contenga a lo más uno defectuoso.
14. En 1927, Babe Ruth logró la estadística siguiente:

JJ	VB	H	2B	3B	HR	CA	CJ	BB	P	BR	PB
151	540	192	29	8	60	158	164	138	89	7	0.356

Suponga que en cierto juego fue cuatro veces al bat, oficialmente, estime con base en los datos de arriba la probabilidad de que obtuviera dos jonrones y dos "outs" no necesariamente en ese orden.

15. Un juego de formar palabras utiliza 44 vocales y 54 consonantes. Si usted elige siete letras al azar, ¿cuál es la probabilidad de que haya:
    a) únicamente vocales?;
    b) 4 vocales?;
    c) 3 consonantes?

Aplicaciones de computación

1. Use un programa de computadora para encontrar la probabilidad que entre 200 recién nacidos, el número total de niñas sea 95 o menos, si la probabilidad de ese suceso es 0.485.
2. Los datos adjuntos representan los resultados de realizar un experimento binomial; si 1 significa un éxito y 0 un fracaso, use un programa de computadora para:
   a) estimar la probabilidad de un éxito;
   b) determinar la media y la desviación estándar de la distribución.

1 0 0 1 0 0 0 0 0 1 0 1 1 1 0
0 0 0 1 0 0 0 0 0 0 0 1 0 1 0
1 0 0 0 0 1 1 1 0 1 0 0 1 1 0
0 0 0 0 0 0 0 0 1 0 0 1 0 0 0
1 0 1 1 0 1 1 1 0 1 0 0 1 1 0
1 1 1 1 1 1 1 1 0 1 0 0 1 0 1
1 1 1 0 0 0 1 0 0 0

3. a) Use un programa de computadora para simular 100 experimentos binomiales con parámetros $n = 4$ y $p = 0.4$. Calcule la media y la desviación estándar de los 100 resultados y use un histograma para resumirlos.
   b) Ordene los resultados del inciso a y construya una tabla de frecuencias relativas para ellos.
   c) Determine una tabla de probabilidad para la distribución binomial que tiene parámetros $n = 4$ y $p = 0.4$, compárela con la construida en el inciso b.

4. Simule lanzar 100 veces un dado común.
   a) Encuentre la media y la desviación estándar de los 100 resultados.
   b) Ordene los 100 resultados y construya una tabla de frecuencias relativas para ellos.

■ EXAMEN DE CONOCIMIENTOS DEL CAPÍTULO ■

1. Se estima que el 5% de los transistores fabricados por la AJAX Company son deficientes; si se examina un lote de 20, encuentre la probabilidad de que:
   a) ninguno sea deficiente;
   b) a lo más uno lo sea;
   c) 19 estén defectuosos;
   d) al menos 5 no sirvan.
2. Un examen de opción múltiple tiene diez preguntas, cada una con cinco opciones; si un estudiante sin preparación responde cada pregunta adivinando, encuentre la probabilidad de que tenga:
   a) cinco aciertos;
   b) a lo más dos respuestas correctas;
   c) cuando mucho nueve aciertos;
   d) dos o más aciertos.
3. Para el ejercicio 2, encuentre lo que se pide e interprete sus respuestas.
   a) El número esperado en respuestas correctas.
   b) La varianza del número de respuestas correctas.
4. Las caras de un tetraedro se numeran del 1 al 4; suponga que el tetraedro se lanza cinco veces y que cada lado tiene la misma probabilidad de caer hacia abajo.
   a) Construya una tabla de probabilidad para el número de unos obtenido.
   b) Trace una gráfica que muestre las probabilidades para el número de unos obtenido.
   c) Encuentre μ y σ para los unos obtenidos.
5. El número X de nacimientos en el Memorial Hospital en un día cualquiera sigue una distribución de Poisson con media igual a dos por día. ¿Cuál es la probabilidad de que sean tres en cierto día?
6. Suponga que una alacena grande tiene tres tipos de compartimientos: pequeños, medianos y grandes, 25% de los cuales son pequeños, 30% medianos y 50% grandes. Si se eligen seis compartimentos al azar, ¿cuál es la probabilidad de que dos sean pequeños, dos medianos y dos grandes?
7. El promedio de llamadas telefónicas que entran en un conmutador es de dos cada tres minutos, si la cantidad total de llamadas sigue un proceso de Poisson, ¿cuál es la probabilidad de que entren dos en un periodo de tres minutos?
8. Suponga que una urna contiene tres pelotas blancas y cuatro negras, si se sacan dos pelotas sin remplazo, ¿cuál es la probabilidad de sacar una pelota blanca y una negra?

7 Distribuciones continuas

DESCRIPCIÓN

OBJETIVOS DEL CAPÍTULO

En este capítulo estudiaremos:

- *Distribuciones uniformes.*
- *Cómo calcular probabilidades asociadas a distribuciones uniformes.*
- *Distribuciones normales.*
- *Distribuciones normales estándar.*
- *Cómo asociar áreas bajo la curva normal estándar con probabilidades.*
- *Cómo calcular probabilidades asociadas con la distribución normal estándar.*
- *Cómo calcular probabilidades asociadas con cualquier distribución normal.*
- *Aplicaciones prácticas de distribuciones normales.*
- *Cómo usar las distribuciones normales para aproximar distribuciones binomiales.*
- *Distribuciones exponenciales.*
- *Cómo calcular probabilidades asociadas con distribuciones exponenciales.*

MOTIVADOR 7

Las distribuciones normales son de gran importancia en estadística. Fueron descubiertas inicialmente por Abraham de Moivre (1667-1754), un matemático francés radicado en Londres, al tratar de resolver problemas de los apostadores. Después fueron estudiadas por Pierre Laplace y Karl Gauss. Gauss (1777-1855), matemático alemán, fue el primero en investigar las propiedades de las distribuciones normales que llegaron a ser conocidas como la ley gaussiana del error, en referencia a su uso para describir la distribución de los errores de estimación en el método de los mínimos cuadrados, un método que el inventó en 1795, a la edad de 18 años. Además de su utilidad en la solución de problemas en los juegos de apuestas, se encontró que eran aplicables a errores de observación en astronomía y, de hecho, en toda clase de medidas físicas.

El término "distribución normal" surge de la observación de que los errores en medidas repetidas de la misma cantidad física tienden a exhibir una gran regularidad en la distribución de sus frecuencias; era un lugar común esperar que los errores en las medidas repetidas de una cantidad como la altura, el peso o la inteligencia, tuvieran una distribución acampanada o normal. Las distribuciones normales se llaman también distribuciones gaussianas, especialmente en áreas de ciencias naturales, en Francia se denominan también distribuciones laplacianas.

Panorama del capítulo

Un experimento que requiere medidas cuyos valores pueden suponerse en un continuo puede identificarse con una variable aleatoria continua; los valores de esta variable consisten en medidas que son números reales pertenecientes a un continuo. Al conjunto infinito de medidas que contiene los resultados posibles de un experimento que utiliza alguna forma de medida se le llama: **distribución continua**. Recuerde de la sección 5.7 que las probabilidades asociadas con una distribución continua se identifican con áreas bajo una curva. El área total comprendida entre la curva y el eje x debe ser igual a 1.

Como ejemplo de una distribución continua, considere el tiempo que tarda un individuo en reaccionar a un estímulo dado, y suponga que el rango máximo del tiempo de reacción es 2 segundos; si los tiempos de reacción se midieran hasta los milésimos de segundo, habría dos mil resultados posibles solamente, a causa de las limitantes en los instrumentos de medición. Teóricamente, los tiempos de reacción podrían ser cualquiera de los números infinitos correspondientes a valores que van desde el tiempo más corto de reacción hasta el más grande; así que los tiempos teóricos de reacción forman una distribución continua.

En este capítulo estudiaremos tres tipos muy importantes de distribuciones continuas de probabilidad uniformes, normales y exponenciales.

SECCIÓN 7.1 Distribuciones uniformes

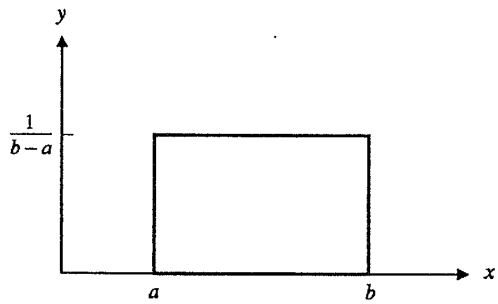
En la sección 5.2 construimos histogramas de probabilidad para datos continuos seleccionando intervalos de clase y considerando un rectángulo para cada uno; el área de un rectángulo correspondía a la proporción de las medidas que caían en el intervalo, y la suma de las áreas de todos los rectángulos era 1. Si tenemos un gran número de medidas y permitimos un gran número de intervalos de clase, entonces cada intervalo tendría un ancho reducido, donde el ancho de cada rectángulo delgado correspondería a la probabilidad de que una medida cayera en él. Así, mientras más medidas tengamos mayor será el parecido del histograma con una curva continua (véase la figura 2.12 del capítulo 2); esto sugiere que si tenemos una curva suave para la cual el área bajo ella es 1, podemos identificarla con una distribución continua de probabilidad. Si una variable aleatoria tiene como función de probabilidad una curva suave, necesariamente debe tener un número infinito de valores; la probabilidad de cualquier valor aislado debe ser cero; de otra forma, la suma de un número infinito de valores no cero excedería de 1. Cuando se trate de una variable aleatoria continua, consideraremos la probabilidad como la posibilidad de que un valor particular de dicha variable caiga en un cierto intervalo. Las funciones de probabilidad de variables aleatorias continuas se denominan **funciones de densidad de probabilidad**, porque las probabilidades no cero se identifican con áreas bajo sus gráficas y para distinguirlas de las funciones de probabilidad de variables aleatorias discretas.

La variable aleatoria más simple es la **variable aleatoria uniforme**, para la que podemos identificar una **distribución de probabilidad uniforme.** La función de probabilidad para una variable aleatoria uniforme tiene como gráfica un segmento de recta horizontal situado sobre el eje x. El área comprendida entre el segmento y el eje x es igual a 1.

Sean a y b cualesquiera dos números reales con $a < b$.

Función de densidad de probabilidad uniforme

$$f(x) = \frac{1}{b-a} \quad \text{para } a < x < b \tag{7.1}$$



Si x no está entre a y b, entonces definimos $f(x) = 0$. Note que el área del rectángulo formado por el eje x, el segmento de recta horizontal y los segmentos verticales $x = a$ y $x = b$ es 1.

$$\text{Área} = (\text{longitud})\ (\text{altura}) = (b-a) \text{ o } \frac{1}{b-a} = 1$$

La inclusión o exclusión de un punto extremo de un intervalo no afecta la probabilidad de que una variable aleatoria continua caiga en el intervalo, porque la probabilidad de un punto extremo es cero. Así, si X es una variable aleatoria continua:

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$

EJEMPLO 7.1

Suponga que una rueda de un tranvía tiene un radio de r pulgadas y corre sobre rieles de acero, que se elige un punto de referencia 0 en la circunferencia de la rueda y que la variable aleatoria X denota la distancia de un punto de la circunferencia al punto de referencia. Cuando el tranvía se frena, algún punto de la rueda hace contacto con el riel y se desliza momentáneamente a lo largo de éste, este punto puede etiquetarse por la distancia x a la que está del punto de referencia de la rueda; al repetir el frenado el valor de x representa una variable aleatoria que se distribuye uniformemente en el intervalo de 0 a $2\pi r$ (la longitud de la circunferencia de la rueda); si algunos puntos de la rueda hacen contacto más frecuentemente que otros, entonces se generarán áreas lisas a lo largo de la rueda y esta mostrará picos al final de estos puntos perdiendo su redondez. La función de probabilidad para la variable aleatoria X es:

$$f(x) = \frac{1}{2\pi r}$$

para x entre 0 y $2\pi r$.

APLICACIÓN 7.1

Una máquina electrónica hace pernos de 3/8 de pulgada que deben tener una longitud de 3 pulgadas. Si en realidad las longitudes de los pernos de 3/8 de pulgada se distribuyen uniformemente en el intervalo que va de las 2.5 a las 3.5 pulgadas, ¿cuál es la probabilidad de que uno de esos pernos elegido al azar de un lote terminado tenga una longitud que:

a) esté entre 2.75 y 3.25 pulgadas?;
b) sea mayor de 3.25 pulgadas?;
c) sea exactamente igual a 3 pulgadas?

Solución: Denotemos por la variable aleatoria X la longitud de los pernos de 3/8 de pulgada. Sean $a = 2.5$ y $b = 3.5$. Si usamos la fórmula (7.1), la función de densidad de probabilidad f está definida por:

$$f(x) = \frac{1}{b-a} = \frac{1}{3.5-2.5} = 1$$

La gráfica de la función de probabilidad para X se muestra en la figura 7.1.

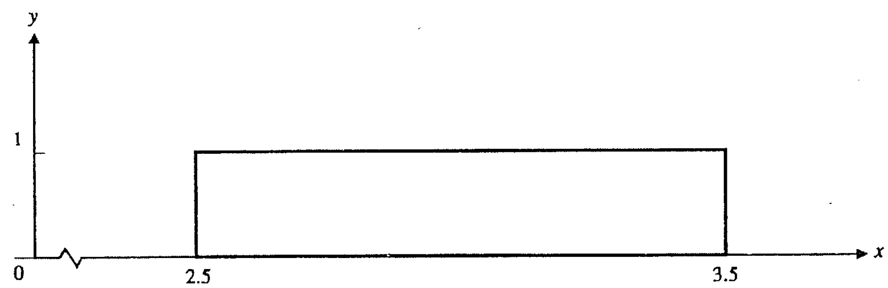


FIGURA 7.1
Gráfica de la función de probabilidad para la aplicación 7.1

a) Para determinar la probabilidad de que el valor de la variable aleatoria caiga entre 2.75 y 3.25, encontramos el área del rectángulo determinado por la gráfica del eje x y las rectas entre $x = 2.75$ y $x = 3.25$. Esta área se muestra sombreada en la figura 7.2

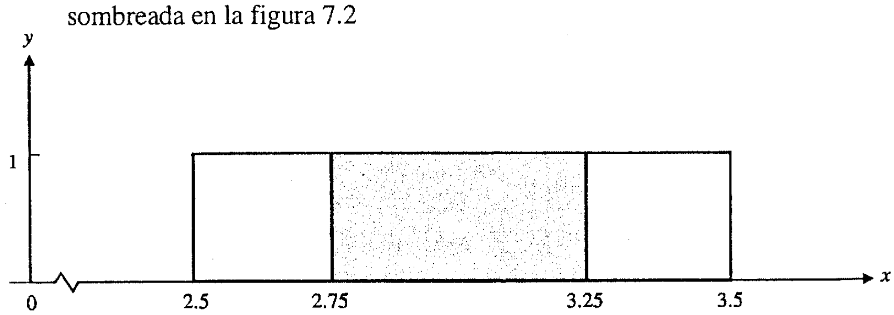


FIGURA 7.2
Gráfica de la probabilidad de que la variable aleatoria quede entre 2.75 y 3.25

y representa la probabilidad $P(2.75 < X < 3.25)$. La longitud del rectángulo sombreado es $3.25 - 2.75 = 0.5$, y la altura es 1. En consecuencia, el área es

$$\text{Área} = (\text{longitud})\ (\text{altura}) = (0.5)(1) = 0.5$$

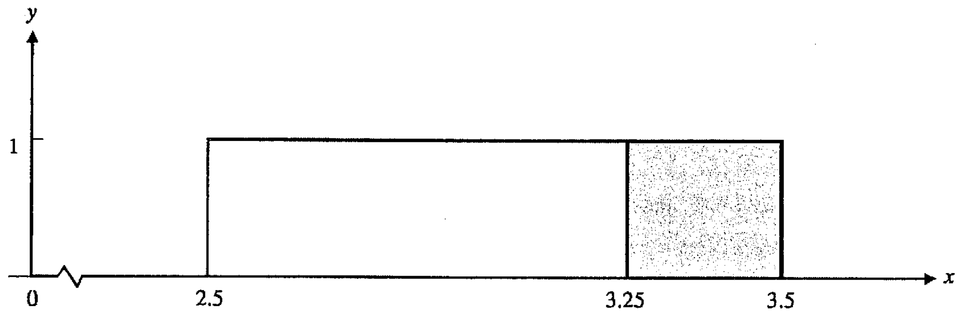
En consecuencia, la probabilidad de que el perno elegido al azar tenga una longitud entre 2.75 y 3.25 pulgadas es 0.5.

b) La probabilidad de que un perno tenga una longitud mayor que 3.25 pulgadas es el área del rectángulo formado por la gráfica del eje x y las rectas $x = 3.25$ y $x = 3.5$. El área sombreada en la figura 7.3 representa la probabilidad $P(X > 3.25) = P(3.25 < X < 3.5)$. El área de este rectángulo es:

$$\text{Área} = (\text{longitud})(\text{altura}) = (3.5 - 3.25)(1) = 0.25$$

FIGURA 7.3

Gráfica de la probabilidad de que la variable aleatoria sea mayor que 3.25



Por tanto, la probabilidad de que un perno tenga una longitud mayor que 3.25 pulgadas es 0.25.

c) La probabilidad de que un perno elegido al azar tenga una longitud de exactamente 3 pulgadas es 0. ■

APLICACIÓN 7.2

Un servicio de llamadas telefónicas se ha diseñado de forma tal que el tiempo mínimo de espera de quien llame sea de 20 segundos y el máximo de 50; si los tiempos de respuesta se distribuyen uniformemente, encuentre la probabilidad de que, al llamar, una persona tenga un tiempo de respuesta:

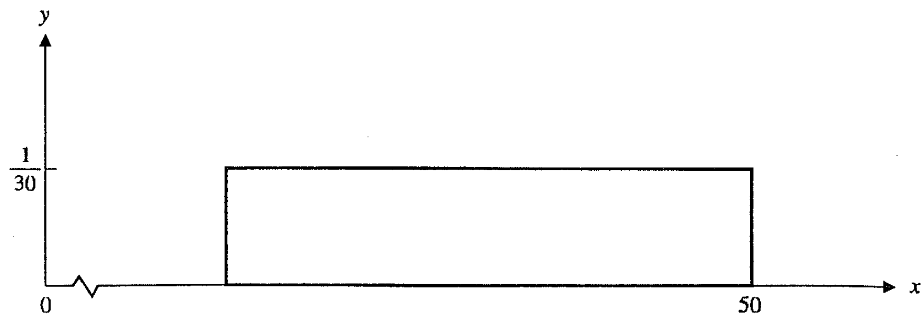
a) Entre 25 y 45 segundos.
b) Menor que 30 segundos o mayor que 40.

Solución: Note que $a = 20$ y $b = 50$. La función de densidad de probabilidad es:

$$f(x) = \frac{1}{b - a} = \frac{1}{50 - 20} = \frac{1}{30}$$

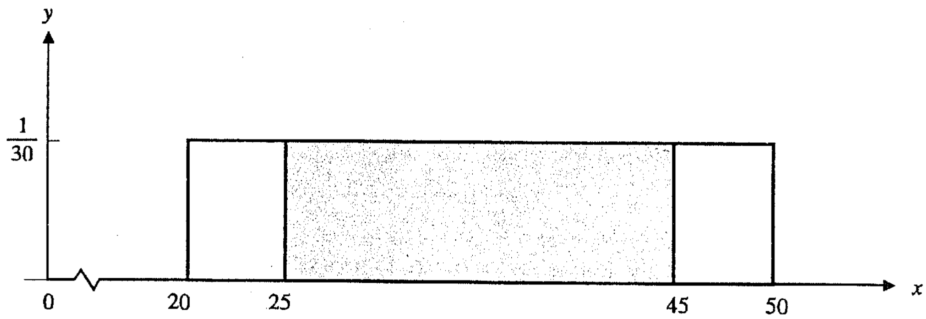
donde x está entre 20 y 50. La gráfica de la función de densidad de probabilidad está dada en la figura 7.4.

FIGURA 7.4
Gráfica de la función de densidad de probabilidad para la aplicación 7.2



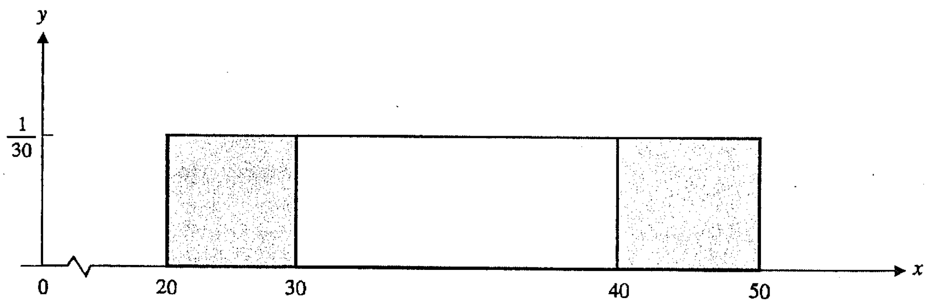
a) Necesitamos encontrar el área bajo la recta $f(x) = 1/30$, entre $x = 25$ y $x = 45$ (vea el área sombreada en la figura 7.5). El área sombreada del rectángulo es $(45 - 25)(1/30) = 15/30 = 0.667$. Por lo tanto, la probabilidad de que una persona al llamar deba esperar entre 25 y 45 segundos para que le contesten es de 0.667.

FIGURA 7.5
Gráfica de la probabilidad de que *X* esté entre 25 y 45



b) La probabilidad, como lo indican las regiones sombreadas de la figura 7.6, es:

FIGURA 7.6
Gráfica de probabilidad de que *X* sea menor que 30 o mayor que 40



$$
\begin{aligned}
P[(X < 30) \text{ o } (P(X > 40))] &= P(20 < X < 30) + P(40 < X < 50) \\
&= \left(\frac{1}{30}\right)(30 - 20) + \left(\frac{1}{30}\right)(50 - 40) \\
&= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}
\end{aligned}
$$

En consecuencia, la probabilidad de que una persona que llama, elegida al azar, deba esperar menos de 30 segundos o más de 40 para que le respondan es de 2/3. ■

GRUPO DE EJERCICIOS 7.1

Habilidades básicas

1. Si la variable aleatoria X se distribuye uniformemente para valores mayores que 20 pero menores que 80, encuentre:
   a) $P(X = 25)$. d) $P(X > 45)$.
   b) $P(30 < X < 50)$. e) $P(10 < X < 20)$.
   c) $P(X < 70)$.

2. Suponga que la variable aleatoria X se distribuye uniformemente sobre el intervalo que se extiende de 25 a 75. Encuentre:
   a) $P(30 < X < 70)$. d) $P(X > 70)$.
   b) $P(X < 37)$. e) $P(80 < X < 90)$.
   c) $P(X = 50)$.

3. Igual que en el anterior, para el intervalo que se extiende de –20 a 50.
   a) $P(-5 < Y < 5)$. c) $P(Y < 40)$.
   b) $P(Y > -10)$. d) $P(Y = 0)$.

4. Suponga que la variable aleatoria Y se distribuye uniformemente sobre el intervalo que va de –30 a 30. Localice:
   a) $P(-10 < Y < 5)$. c) $P(Y < -5)$.
   b) $P(Y > 20)$. d) $P(Y = -30)$.

Más aplicaciones

5. Suponga que una despachadora automática de café nunca da menos de 6 onzas ni más de 10; cualquier cantidad de café entre 6 y 10 onzas tiene la misma probabilidad de ocurrir. Al despachar cierta cantidad, determine la probabilidad de que:
   a) sea menor de 7 onzas;
   b) sea mayor de 6 onzas;
   c) esté entre 7 y 9 onzas.

6. Suponga que la precipitación pluvial anual en una región se distribuye uniformemente con valores de entre 10 y 14 pulgadas. Si se elige un año al azar, determine la probabilidad de que la precipitación pluvial sea de:
   a) menos de 10 pulgadas;
   b) más de 11 pulgadas;
   c) entre 11 y 14 pulgadas.

7. Considere que la temperatura invernal promedio, en grados Fahrenheit, para una región del norte de Estados Unidos se distribuye de igual manera entre niveles que van desde –20° hasta 20°. Si se escoge al azar una temporada invernal, calcule la probabilidad de que la temperatura promedio:
   a) sea menor de 10°;
   b) esté entre –5° y 15°;
   c) sea menor que 10° y mayor que 5°;
   d) esté entre –10° y 10°.

8. Si una profesora nunca llega justo a tiempo a la clase, ni con más de 5 minutos de adelanto ni de retraso, y arriba con la misma frecuencia para los valores en ese rango de tiempo; encuentre la probabilidad de que la profesora llegue:
   a) al menos 3 minutos tarde;
   b) a lo más 4 minutos antes;
   c) justo a tiempo;
   d) no más de 2 minutos antes ni menos de 2 minutos después.

Un paso más allá

9. Si X se distribuye uniformemente sobre el intervalo que se extiende entre a y b, entonces su media o valor esperado está dado por $\mu_x = E(X) = (a + b)/2$, y su varianza por $\sigma_x^2 = (b - a)^2/12$; encuentre la media y la varianza para la variable aleatoria del ejercicio 5.

10. Encuentre la media y la desviación estándar de la variable aleatoria dada en el ejercicio 6.

11. Si X es una variable aleatoria continua que se distribuye de manera uniforme en el intervalo que va de a a b, determine $E(X^2)$.

12. Suponga que una variable aleatoria continua X tiene una función de densidad de probabilidad g dada por $g(x) = (1/2)x$ para valores de x que van de 0 a 2; encuentre la probabilidad de que X:
    a) sea menor que 1;
    b) esté entre 1 y 2;
    c) sea mayor que 0.5;
    d) sea menor que 0.5 o mayor que 1.5.

13. Dada la función f definida por $f(x) = (5/4)x$ para valores de x de 3 hasta b; encuentre el valor de b tal que f es una función de densidad de probabilidad.

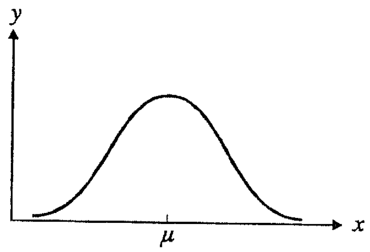
SECCIÓN 7.2 Distribuciones normales

Una de las clases más importantes de distribuciones continuas es la **distribución normal;** desde su descubrimiento hace ya más de 350 años, se ha desenvuelto como una herramienta indispensable en cualquier rama de la ciencia, la industria y el comercio. Muchos eventos reales y naturales tienen una distribución de frecuencias cuya forma es muy parecida a la distribución normal.

EJEMPLO 7.2

La distribución de frecuencias del contenido de nitrógeno de las hojas de un árbol tiende a ser normal. Las medidas físicas suelen distribuirse normalmente; las presiones sanguíneas sistólica y diastólica, las pulsaciones del corazón, los niveles de colesterol en la sangre, las estaturas de los hombres adultos y los pesos de las niñas de dos años son todos ejemplos de distribuciones de datos que tienden a seguir la distribución normal; además, las calificaciones de muchos exámenes de aplicación amplia, como el SAT y el American College Test (ACT), poseen distribuciones cuya forma es acampanada.

FIGURA 7.7
Gráfica de una distribución normal



Una distribución normal tiene la forma de una montaña o la apariencia de una campana, como lo ilustra la figura 7.7. La ecuación de una curva con forma de campana está dada por:

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x - \mu)^2/2\sigma^2}$$

donde μ representa la media de la población, σ es la desviación estándar de la población, $e \approx 2.718$ y x es cualquier número real.

Los dos parámetros μ y σ especifican por completo la posición y la forma respectivamente, de una distribución normal; un valor pequeño de σ significa que la curva normal es una campana delgada, picuda; mientras que un valor grande de σ significa que la curva normal es ancha, aplanada (véase la figura 7.8). La figura 7.9 ilustra tres distribuciones normales que tienen una media de 20 pero cuyas desviaciones estándar difieren. Y la figura 7.10 nos deja ver tres distribuciones normales con medias distintas pero con la misma desviación estándar. En muchas aplicaciones prácticas tenemos colecciones de medidas cuyas gráficas aproximan una curva acampanada.

FIGURA 7.8
Efecto de μ y σ en la posición y la forma de la gráfica de una distribución normal

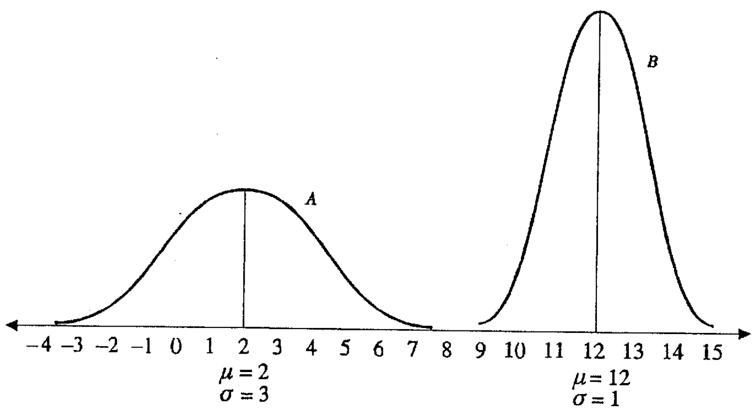


FIGURA 7.9
Gráficas de distribuciones normales con la misma media, pero con distintas desviaciones estándar

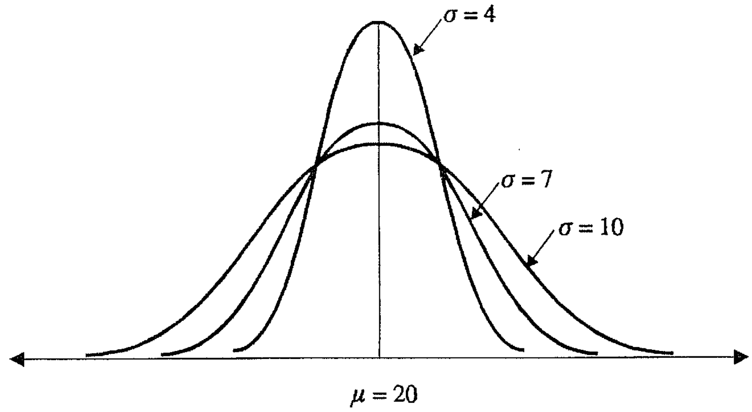
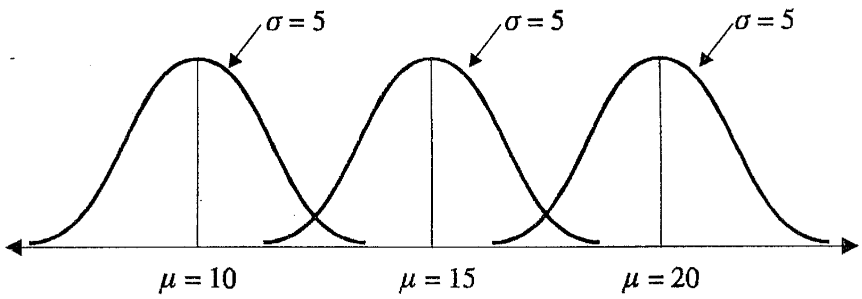


FIGURA 7.10
Gráficas de distribuciones normales con medias distintas, pero con la misma desviación estándar



EJEMPLO 7.3

Considere la colección siguiente de 100 medidas de presión sanguínea sistólica, pertenecientes a un grupo de individuos de entre 20 y 40 años de edad

98	141	127	112	107	111	116	104	87	130
131	124	126	129	65	126	139	136	137	102
128	152	115	124	96	114	116	119	108	113
126	116	107	134	82	141	85	126	131	141
128	85	145	150	97	125	109	119	159	114
117	131	122	104	145	119	123	138	109	102
115	109	143	128	136	118	125	124	109	133
131	127	100	110	109	112	125	166	93	92
113	119	116	136	93	131	141	153	85	119
92	112	140	121	108	92	98	159	91	125

El diagrama de tallo y hojas para este conjunto de datos sugiere que las medidas se distribuye normalmente.

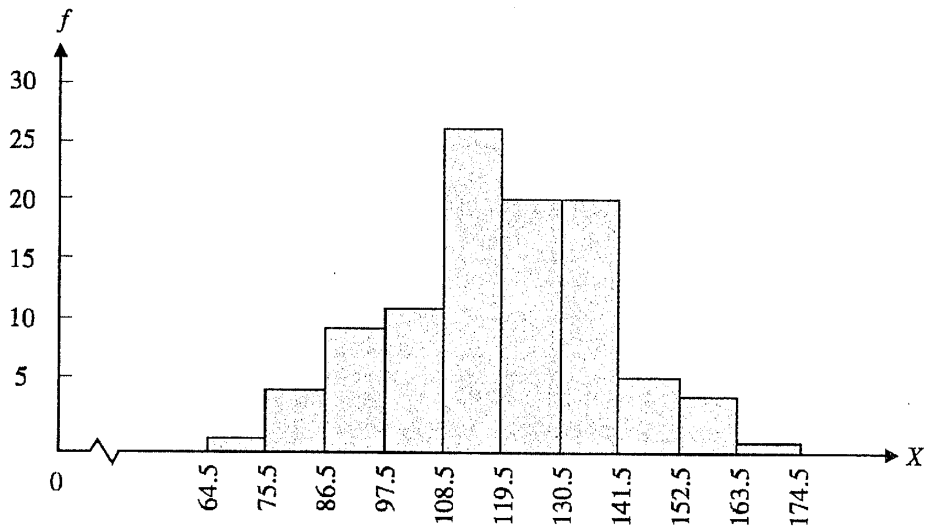
6	5
7	
8	2 5 5 5 7
9	1 2 2 2 3 6 6 7 8 8
10	0 2 2 4 4 7 7 8 8 9 9 9 9 9
11	0 1 2 2 2 3 3 4 4 5 5 6 6 6 6 7 8 8 9 9 9 9
12	1 2 4 4 4 5 5 5 5 6 6 6 6 7 7 8 8 8 9
13	0 1 1 1 1 1 2 3 4 5 5 5 7 8 9
14	0 1 1 1 1 3 5 5
15	0 2 3 9 9
16	6

Una tabla de frecuencias agrupadas para los datos anteriores es como sigue:

Límites de clase	f
67-75	1
76-86	4
87-97	9
98-108	11
109-119	27
120-130	20
131-141	19
142-152	5
153-163	3
164-174	1

En la figura 7.11 vemos un histograma de frecuencias que, aunque no es simétrico, tiene una forma aproximadamente acampanada y sugiere que las medidas de la presión sanguínea pueden distribuirse normalmente.

FIGURA 7.11
Histograma para las medidas de la presión sanguínea



Propiedades de las distribuciones normales

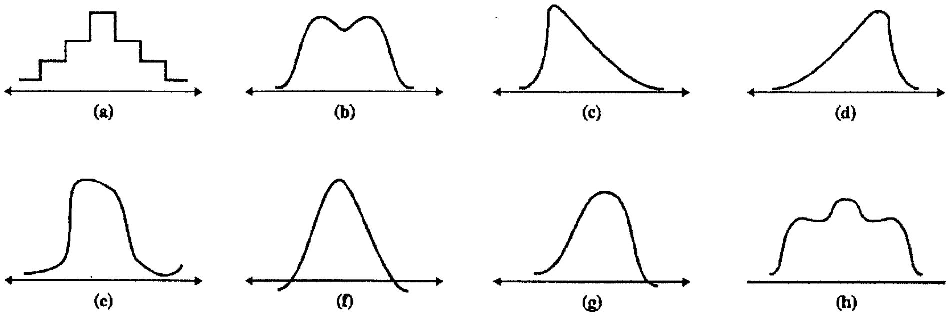
Algunas de las propiedades de las distribuciones normales se listan aquí.

Propiedades de las distribuciones normales

1. Una distribución normal tiene forma de montaña o de campana.
2. El área bajo una curva normal y sobre el eje x es siempre igual a 1.
3. La media se localiza en el centro de la distribución y la curva normal es simétrica con respecto a la línea perpendicular, al eje horizontal en el valor de la media.
4. La media, la mediana y la moda coinciden.
5. Una curva para una distribución normal se extiende indefinidamente a la izquierda y a la derecha de la media y tiende hacia el eje horizontal.
6. Una curva para una distribución normal nunca toca al eje horizontal.
7. La forma y la posición de una distribución normal dependen de los parámetros μ y σ, en consecuencia hay un número infinito de distribuciones normales.

APLICACIÓN 7.3

Explique si cada una de las curvas siguientes puede representar una distribución normal

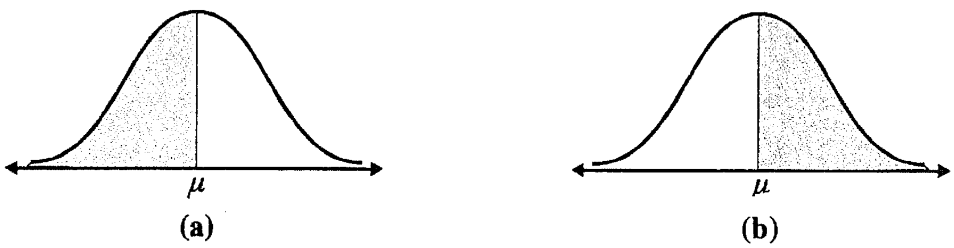


Solución:

a) No, porque no tiene forma de montaña.
b) No, tampoco tiene forma de montaña.
c) No, no es simétrica.
d) No, tampoco es simétrica.
e) No, porque no tiende hacia el eje horizontal.
f) No, la curva corta el eje horizontal.
g) No, también corta al eje horizontal.
h) No, la curva no presenta forma de montaña.

APLICACIÓN 7.4

Encuentre el área de las regiones sombreadas bajo las curvas normales que se muestran aquí.



Solución: En ambos casos el área sombreada es 0.5, porque el área total bajo una curva normal es 1 y la curva es simétrica respecto a su medida central μ. ■

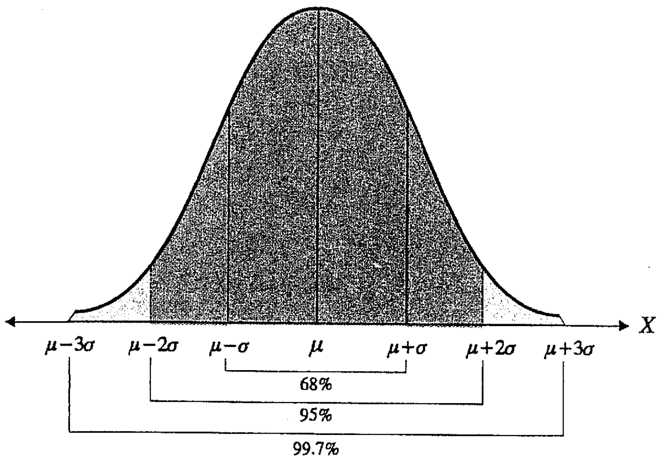
Regla empírica

La **regla empírica** se aplica a cualquier distribución normal. La figura 7.12 ilustra la regla empírica.

Regla empírica

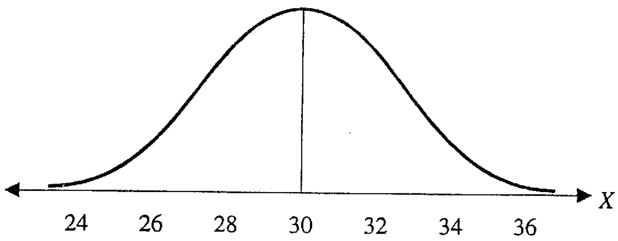
a) Aproximadamente el 68% de las medidas distan menos de una desviación estándar de la media, es decir, caen en el intervalo $(\mu \pm \sigma)$.
b) Casi un 95% de las medidas distan menos de dos desviaciones estándar de la media; caen en el intervalo $(\mu \pm 2\sigma)$.
c) Alrededor del 99.7% de las medidas distan menos de tres desviaciones estándar de la media, esto es, pertenecen al intervalo $(\mu \pm 3\sigma)$.

FIGURA 7.12
Ilustración gráfica de la regla empírica



APLICACIÓN 7.5

Suponga que $\sigma = 2$ y $\mu = 30$ para la distribución normal dibujada aquí.

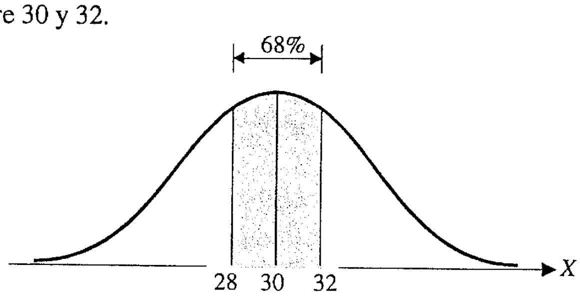


Encuentre el porcentaje de las medidas que caen:

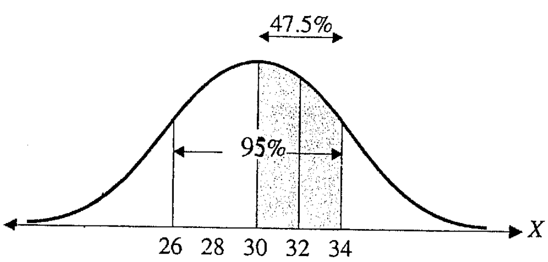
a) entre 30 y 32;
b) entre 32 y 34;
c) arriba de 32.

Solución:

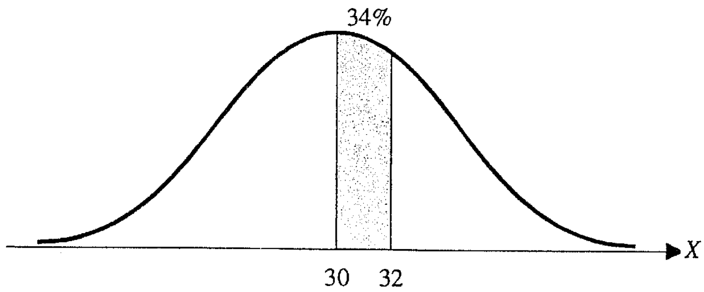
a) Como la curva es simétrica con respecto a $\mu = 30$ y aproximadamente el 68% de los puntajes caen entre 28 y 32, casi un (1/2) (68%) = 34% de las medidas, caen entre 30 y 32.



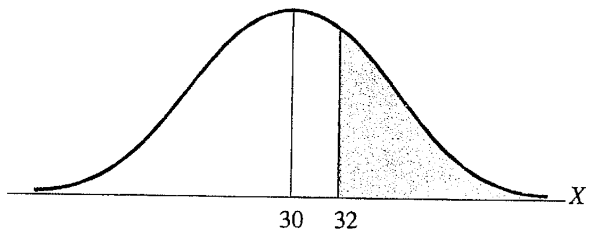
b) Este problema requiere que encontremos las áreas sombreadas ilustradas en los dibujos siguientes:



Aproximadamente 95% de los puntajes caen entre 26 y 34; como la curva es simétrica con respecto a $\mu = 30$, $(1/2)(0.95) = 0.475 = 47.5\%$ de los puntajes caen entre 30 y 34, en consecuencia, el porcentaje ilustrado en el primer dibujo es 47.5%; como resultado del inciso a, el porcentaje ilustrado en el segundo dibujo es 34%, por tanto, restamos 34% de 47.5% y obtenemos 13.5%, que es el porcentaje de medidas que caen entre 32 y 34.



c) Por el inciso a tenemos que un 34% de los puntajes caen entre 30 y 32, como el 50% se sitúa arriba de 30, se sigue que $0.50 - 0.34 = 0.16 = 16\%$ de los puntajes caen arriba de 32.



Recuerde de la sección 3.4 que un puntaje z para una población está definido por

$$z = \frac{x - \mu}{\sigma}$$

e indica el número de desviaciones estándar que un puntaje dista de la media. Una medida igual a $(\mu + \sigma)$ está una desviación estándar arriba de la media y tiene un puntaje z igual a 1:

$$\begin{aligned} z &= \frac{x - \mu}{\sigma} \\ &= \frac{(\mu + \sigma) - \mu}{\sigma} \\ &= \frac{\sigma}{\sigma} = 1 \end{aligned}$$

Por el mismo razonamiento, la medida $(\mu + 2\sigma)$ está dos desviaciones estándar arriba de la media y tiene un puntaje z igual a 2:

$$\begin{aligned} z &= \frac{x - \mu}{\sigma} \\ &= \frac{(\mu + 2\sigma) - \mu}{\sigma} = 2 \end{aligned}$$

Y la medida ($\mu + 3\sigma$) está tres desviaciones estándar arriba de la media con un puntaje de z de 3.

$$z = \frac{x - \mu}{\sigma}$$
$$= \frac{(\mu + 3\sigma) - \mu}{\sigma} = 3$$

Como una consecuencia de estos resultados, la regla empírica puede reformularse como sigue:

Reformulación de la regla empírica

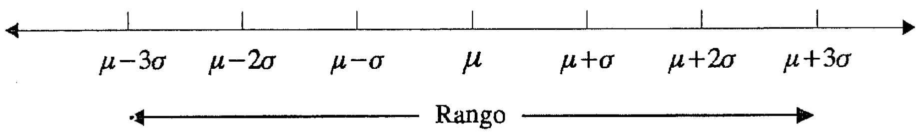
a) En cualquier distribución normal, aproximadamente 68% de los puntajes z caen entre -1 y $+1$.
b) En cualquier distribución normal, casi un 95% de los puntajes z caen entre -2 y $+2$.
c) En cualquier distribución normal, alrededor del 99.7% de los puntajes z caen entre -3 y $+3$.

Aproximación de σ y s

En la sección 3.2 se estableció que $R/4$ da una buena aproximación de s, donde R es el rango y s la desviación estándar muestral. El rango de una distribución normal es infinito y aproximadamente 99.7% de una distribución muestral cae dentro del intervalo $\mu \pm 3\sigma$. Si restringimos la distribución a $\mu \pm 3\sigma$, entonces el rango de la distribución restringida es aproximadamente igual a 6σ; por tanto, el rango R de cualquier distribución acampanada es aproximadamente igual a 6σ, como se ilustra en la figura 7.13. Por eso, tenemos:

$$\sigma \approx \frac{R}{6}$$

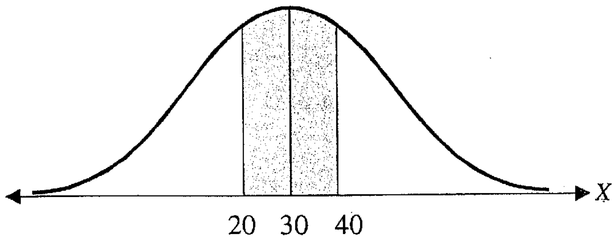
FIGURA 7.13
El rango $R \approx 6\sigma$ para una distribución acampanada



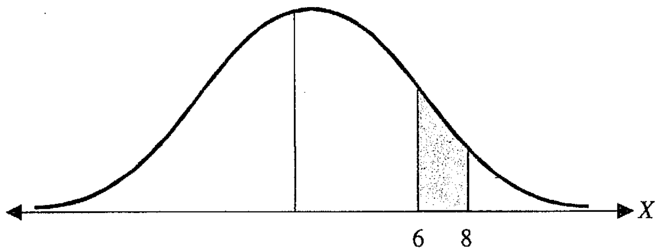
En consecuencia, $R/6$ proporciona una aproximación para σ y también para s. Además, como 95% de las medidas en una distribución normal caen en el intervalo $\mu \pm 2\sigma$, $R/4$ también da una aproximación para s basada en la misma línea de razonamiento, como $R/4$ siempre da una aproximación más holgada, y por tanto más conservadora, de s que $R/6$, la estimación $R/4$ comúnmente se usa más. En muchos casos donde una aproximación es simétrica, $R/4$ proporciona una aproximación razonable de s; pero para ***distribuciones asimétricas***, en que las medida aparecen con mayor frecuencia arriba o abajo de la media, $R/4$ proporciona usualmente una aproximación pobre de s.

Probabilidad y área

Las probabilidades relativas a una distribución normal se identifican con áreas. Por ejemplo, la probabilidad de que una medida x de una distribución normal caiga entre 20 y 40 se puede identificar con el área sombreada de esta figura:

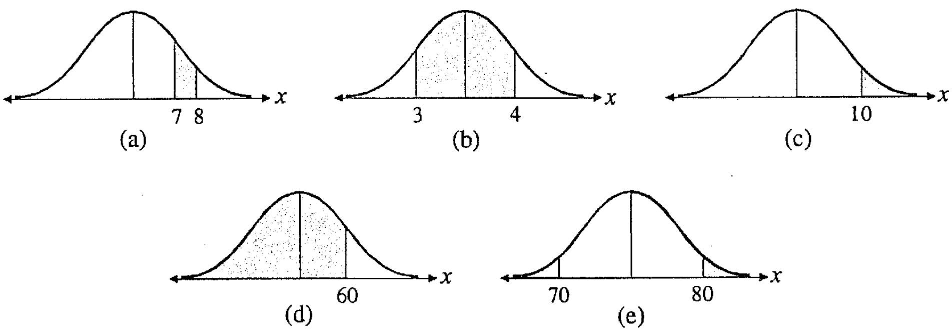


En igual forma, el área sombreada bajo la curva normal siguiente se puede interpretar como la probabilidad de que una medida x esté entre 6 y 8 y se escribe como $P(6 < X < 8)$.



APLICACIÓN 7.6

Escriba una afirmación probabilística para cada una de las áreas siguientes bajo curvas normales:



Solución: Las siguientes son las afirmaciones probabilísticas y sus correspondientes interpretaciones:

Mediante desigualdades	**Interpretación**
a) $P(7 < X < 8)$	la probabilidad de que x caiga entre 7 y 8;
b) $P(3 < X < 4)$	la probabilidad de que x caiga entre 3 y 4;
c) $P(X > 10)$	la probabilidad de que x sea mayor que 10;
d) $P(X < 60)$	la probabilidad de que x sea menor que 60;
e) $P(X < 70 \text{ o } X > 80)$	la probabilidad de que x menor que 70 o mayor que 80. ■

Distribución normal estándar

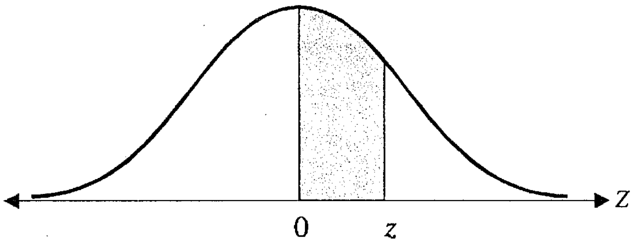
Recuerde de la sección 3.4 que un puntaje z tiene media cero y desviación estándar 1. Si X es una variable aleatoria con media μ y desviación estándar σ, entonces la variable $Z = (X - \mu)/\sigma$ es una variable estandarizada cuyas unidades son desviaciones estándar. La variable normal estándar Z tiene una media de 0 y desviación estándar de 1. La distribución de probabilidad de Z

se denomina **distribución normal estándar**, y el símbolo Z se usa para denotarla.

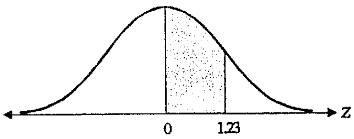
Determinación de probabilidades usando la tabla normal estándar

Queremos ser capaces de determinar la probabilidad de que la variable aleatoria Z tenga uno de sus valores en un intervalo específico. Se puede usar la regla empírica para encontrar ciertas probabilidades o áreas asociadas con la distribución normal estándar, sin embargo, muchas aplicaciones prácticas no requerirán intervalos determinados por exactamente un radio de una, dos o tres desviaciones estándar de la media; por ejemplo, la regla empírica no puede usarse para encontrar $P(0 < Z < 1.23)$, en un caso así, la probabilidad se debe encontrar usando una **tabla normal estándar**, (véase tabla z en la primera de forros). Esta tabla proporciona las áreas de regiones comprendidas bajo la curva normal estándar y sobre el eje horizontal, a la derecha de 0 con un valor específico z, como se ilustra en la figura 7.14. Para encontrar un área, o probabilidad, en la tabla z localizamos los dígitos de las unidades y las décimas de z bajo la columna de la izquierda, y los dígitos de las centésimas a lo largo del renglón superior; el número en el que se intersecan el renglón y la columna es el área deseada.

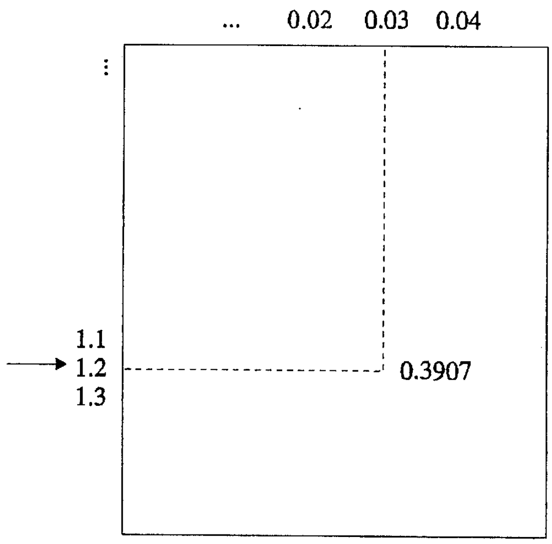
FIGURA 7.14
Área bajo la curva normal estándar dada en la tabla z



EJEMPLO 7.4



Considere el área sombreada de la figura. Para encontrar el área entre 0 y $z = 1.23$ bajo una curva normal estándar, localizamos el renglón 1.2 a la izquierda de la tabla z y la columna de 0.03 a lo largo del renglón superior; la entrada en el cuerpo de la tabla correspondiente a la intersección del renglón 1.2 y la columna 0.03 es 0.3907, como se muestra aquí:



Así, 0.3907 es el área de la región sombreada, que es también la probabilidad de que z esté entre 0 y 1.23.

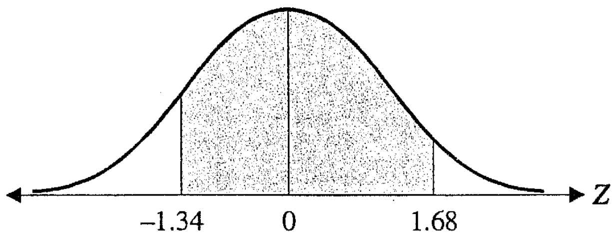
APLICACIÓN 7.7

Encuentre las probabilidades siguientes asociadas con la distribución normal estándar, usando la tabla z:

a) $P(Z$ está entre -1.34 y $1.68)$
b) $P(Z$ está entre 1.34 y $2.38)$
c) $P(Z$ es menor que $-1.25)$

Solución:

a) Un primer paso importante para encontrar cualquier valor de probabilidad normal, es dibujar un diagrama que ilustre el área correspondiente a la probabilidad que se quiere encontrar.



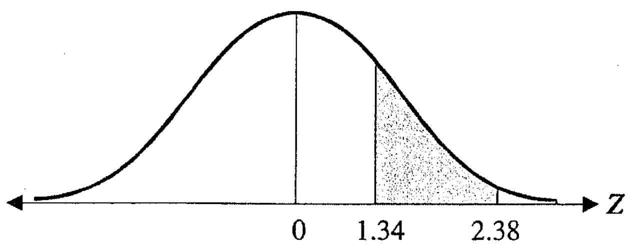
Los pasos son:

1. Encuentre el área de $z = 0$ a $z = 1.68$.
2. Encuentre el área desde $z = -1.34$ a $z = 0$.
3. Sume dichas áreas.

El área desde $z = 0$ a $z = 1.68$ es 0.4535 y, debido a la simetría, el área desde $z = -1.34$ a $z = 0$ es la misma que el área desde $z = 0$ hasta $z = 1.34$, que es 0.4099, al sumar estas dos áreas obtenemos 0.8634. En consecuencia,

$$P(-1.34 < Z < 1.68) = 0.4535 + 0.4099 = 0.8634$$

b)

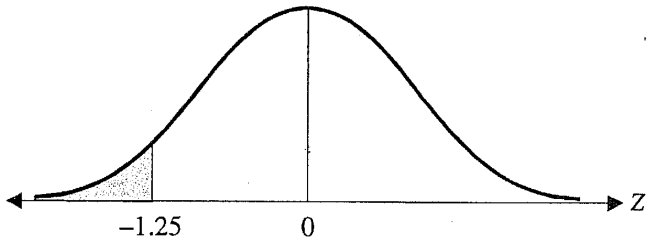


Los pasos son como sigue:

1. Encuentre el área desde $z = 0$ hasta $z = 2.38$.
2. Localice el área desde $z = 0$ hasta $z= 1.34$.
3. Reste la segunda área de la primera.

$$\text{Así, } P(1.34 < Z < 2.38) = 0.4913 - 0.4099 = 0.0814.$$

c)



Los pasos son como sigue:

1. Encontrar el área entre $z = -1.25$ y $z = 0$.
2. Restar esta área de 0.5.

Así, $P(Z < -1.25) = 0.5 - 0.3944 = 0.1056$. ■

EJEMPLO 7.5

En algunas ocasiones necesitamos calcular probabilidades asociadas con valores z que caen en las regiones extremas de la distribución normal estándar. Por ejemplo, podemos necesitar determinar $P(Z < -5)$ o $P(Z > 6)$; en ambos casos, las probabilidades son pequeñas, pero no iguales a cero, porque las curvas se extienden indefinidamente en ambas direcciones. Para probabilidades muy pequeñas, cuyo valor es 0.0000 hasta la cuarta cifra decimal, usaremos el símbolo 0^+, además, indicaremos probabilidades acumuladas asociadas con valores grandes de z como $P(Z < 7)$, usando el símbolo 1^-.

Verificación de la regla empírica

Las probabilidades listadas en la tabla normal (tabla z, véase la primera de forros), pueden usarse para verificar la regla empírica. Advierte que:

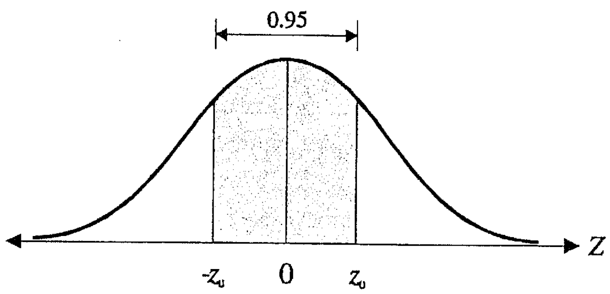
$$P(-1 < Z < 1) = 2P(0 < Z < 1) = 2(0.3413) = 0.6826 \approx 0.68$$
$$P(-2 < Z < 2) = 2P(0 < Z < 2) = 2(0.4772) = 0.9544 \approx 0.95$$
$$P(-3 < Z < 3) = 2P(0 < Z < 3) = 2(0.4987) = 0.9974 \approx 0.997$$

Cómo encontrar puntajes z dadas las áreas

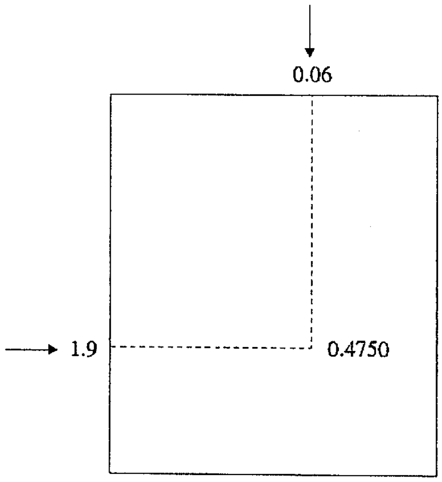
Con frecuencia debemos encontrar el valor z dada el área bajo la curva normal estándar. Éste es el proceso inverso de encontrar probabilidades o áreas que corresponden a valores de z; considere las aplicaciones 7.8-7.11; si el área o probabilidad que buscamos en la tabla normal estándar es igualmente parecida a dos valores de la tabla, escogemos el valor más grande de z, como lo ilustramos en la aplicación 7.9.

APLICACIÓN 7.8

Encuentre el valor normal estándar z_0 tal que $P(Z$ esté entre $-z_0$ y $z_0) = 0.95$.



Solución: El área sombreada en la figura adjunta es 0.95. Como la tabla z lista áreas bajo la curva normal desde $z = 0$ hasta $z = z_0$, dividimos 0.95 entre 2 para obtener una área de 0.475 entre $z = 0$ y $z = z_0$, luego localizamos el área 0.475 en el cuerpo de la tabla z, como se muestra aquí.



Así, el valor de z_0 es 1.96 y $P(-1.96 < Z < 1.96) = 0.95$ ■

APLICACIÓN 7.9

Encuentre el valor normal estándar z_0 tal que $P(Z > z_0) = 0.05$.

Solución: $P(0 < Z < z_0) = 0.45$, como $P(0 < Z < z_0) + P(Z > z_0) = 0.5$. De la tabla z encontramos:

Área	Valores de z
0.4495	1.64
0.4500	
0.4505	1.65

Como 0.45 equidista de 0.4495 y 0.4505 y 1.65 es el valor mayor de z elegimos $z_0 = 1.65$. En consecuencia, $P(Z > 1.65) \approx 0.05$. ■

APLICACIÓN 7.10

Encuentre un valor normal estándar z_0 tal que $P(Z < z_0) = 0.75$.

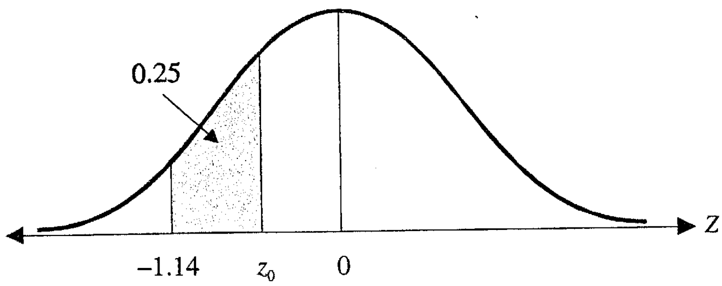
Solución: Como $P(Z < 0) = 0.5$, $P(0 < Z < z_0) = 0.25$. Si buscamos en la tabla z encontramos $z_0 = 0.67$, porque 0.25 es más cercano a 0.2486 que 0.2517.

	Área	Valores de z
	0.2486	0.67
0.0014		
	0.2500	
0.0017		
	0.2517	0.68

Así, $P(Z < 0.67) \approx 0.75$. ■

APLICACIÓN 7.11 Encuentre un valor normal estándar z_0 tal que $P(-1.14 < Z < z_0) = 0.25$.

Solución: Por simetría, $P(-1.14 < Z < 0) = P(0 < Z < 1.14) = 0.3729$. En consecuencia, z_0 debe ser menor que 0 y tenemos la ilustración siguiente:



Así, $P(0 < Z < -z_0) = 0.3729 - 0.25 = 0.1229$. Si leemos la tabla z, vemos que el área más cercana a 0.1229 es 0.1217.

Área	Valor de z
0.1217	0.31
0.1229	
0.1255	0.32

En consecuencia, $-z_0 = 0.31$ y $z_0 = -0.31$. ■

GRUPO DE EJERCICIOS 7.2

Habilidades básicas

1. Suponga que una distribución normal dada tiene una media de 20 y una desviación estándar de 4. Aproxime las probabilidades siguientes usando la regla empírica:
   a) $P(X > 24)$ c) $P(12 < X < 24)$
   b) $P(16 < X < 24)$ d) $P(X > 28)$
2. Haga lo mismo que en el ejercicio anterior:
   a) $P(8 < X < 12)$ c) $P(X < 12)$
   b) $P(24 < X < 32)$ d) $P(20 < X < 28)$
3. Encuentre las probabilidades siguientes usando la tabla z de la primera de forros:
   a) $P(Z$ es mayor que 0)
   b) $P(Z$ está entre -2.5 y 1.5)
   c) $P(Z$ está entre -2.7 y -1.3)
   d) $P(Z$ está entre 2 y 3)
4. Igual que en el anterior:
   a) $P(Z$ es mayor que 2.8)
   b) $P(Z$ es mayor que 3.7)
   c) $P(Z$ está entre -4 y 4)
   d) $P(Z$ está entre 2.1 y 2.9)
5. Encuentre las probabilidades siguientes usando la tabla z de la primera de forros:
   a) $P(-1.23 < Z > 0)$
   b) $P(0 < Z < 2.34)$
   c) $P(-1.78 < Z < 2.38)$
   d) $P(-1.12 < Z < -1.01)$
6. Localice las probabilidades como en el ejercicio tres.
   a) $P(1.23 < Z < 2.75)$ c) $P(Z > 2.97)$
   b) $P(Z < 2.34)$ d) $P(Z > 4.38)$
7. Encuentre el área bajo la curva normal estándar
   a) entre $z = 1.8$ y $z = 2.4$;
   b) arriba de $z = -1.78$;
   c) abajo de $z = 2.5$.
8. Encuentre el área bajo la curva normal estándar.
   a) abajo de $z = -2.17$;
   b) arriba de $z = 1.98$;
   c) entre $z = -1.28$ y $z = 2.35$.
9. Para cada caso, encuentre el valor normal estándar z_0 tal que:
   a) $P(Z$ está entre $-z_0$ y $z_0) = 0.90$.

b) $P(Z$ es mayor que $z_0) = 0.025$.
c) $P(Z > z_0) = 0.10$.
d) $P(-z_0 < Z < z_0) = 0.99$.

10. Para cada caso, encuentre el valor normal estándar z_0 tal que
a) $P(Z > z_0) = 0.85$
b) $P(Z$ está entre $-z_0$ y $2) = 0.94$
c) $P(Z < z_0$ o $z > 1.1) = 0.90$
d) $P(Z < z_0$ o $z > 1.3) = 0.24$

Un paso más allá

11. Una aproximación manual para áreas bajo la curva normal estándar desde 0 hasta z ($z > 0$), está dada por la regla siguiente:

Valor z	Área aproximada
$0 \le z \le 2.22$	$\frac{z(4.44 - z)}{10}$
$2.2 < z < 2.6$	0.49
$z \ge 2.6$	0.50

El error absoluto máximo para la aproximación es 0.0052.[35] Use esta aproximación para encontrar las probabilidades en el ejercicio 3 y compare los resultados.

12. Si u_1 y u_2 son dos variables aleatorias definidas en el intervalo (0, 1), es posible usar las fórmulas siguientes para generar dos variables aleatorias normales estándar:

$$z_1 = \sqrt{-\ln(u_2)}\cos(2\pi u_1)$$
$$z_2 = \sqrt{-\ln(u_2)}\operatorname{sen}(2\pi u_1)$$

donde el ángulo se expresa en radianes. Los valores de u_1 y u_2 pueden obtenerse usando una tabla de números aleatorios y seleccionado tres dígitos al mismo tiempo. Use las fórmulas dadas para simular dos observaciones de la distribución normal estándar.

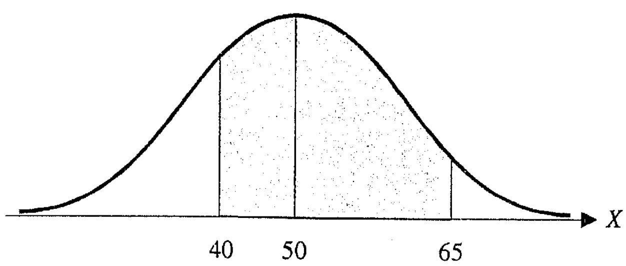
13. a) Escriba un programa computacional para simular 25 observaciones de la distribución normal estándar.
b) Calcule la media y la desviación estándar para las 25 medidas del inciso a, y construya un histograma para verificar que la muestra es de la distribución normal estándar.

SECCIÓN 7.3 *Aplicaciones de las distribuciones normales*

No es posible usar la tabla normal estándar para calcular directamente probabilidades asociadas con una distribución normal distinta de la distribución normal estándar.

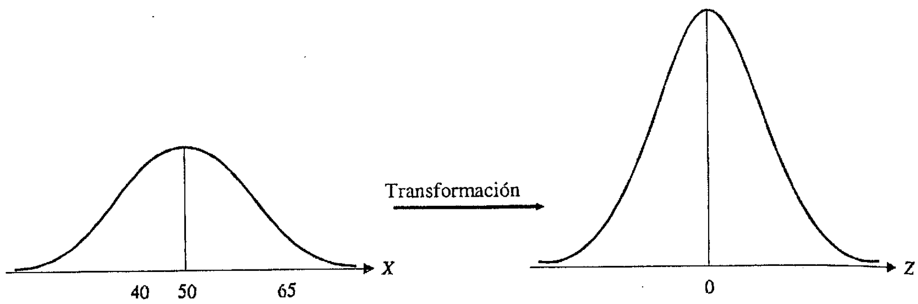
EJEMPLO 7.6

Suponga que estamos interesados en encontrar el área siguiente para la distribución normal estándar con una media de 50 y una desviación estándar de 5.



No podemos usar directamente la tabla normal estándar, tabla z de la primera de forros, porque la variable aleatoria de interés no es la normal estándar; para determinar probabilidades asociadas con una distribución normal no estándar, necesitamos una regla que nos permita transformar cualquier distribución normal en la distribución normal estándar, como se indica en la figura 7.15. Entonces sí podríamos determinar probabilidades usando la tabla z.

FIGURA 7.15
Transformación a la distribución normal



La regla de transformación que usamos es la fórmula del puntaje estándar:

$$z = \frac{x - \mu}{\sigma}$$

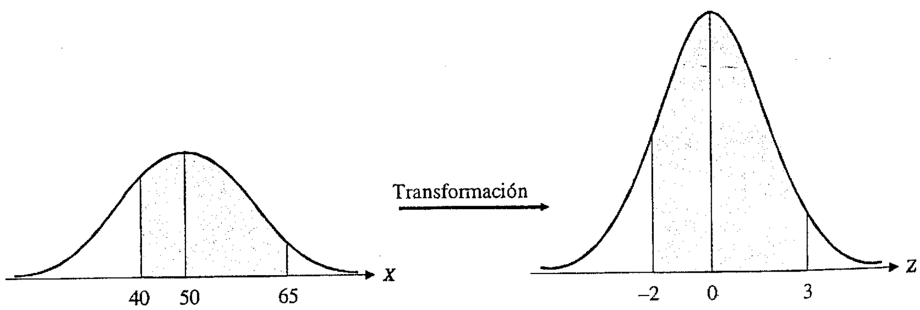
Se usa para transformar cualquier puntaje normal x en un puntaje normal estándar z, así, para la ilustración anterior, tenemos

$$z = \frac{40 - 50}{5} = -2$$

y

$$z = \frac{65 - 50}{5} = 3$$

En consecuencia, obtenemos el diagrama siguiente donde las dos áreas sombreadas son iguales.



Esto es:

$$P(40 < X < 65) = P(-2 < Z < 3)$$

La probabilidad de que z esté entre −2 y 3 es

$$P(-2 < Z < 3) = 0.4772 + 0.4987 = 0.9759.$$

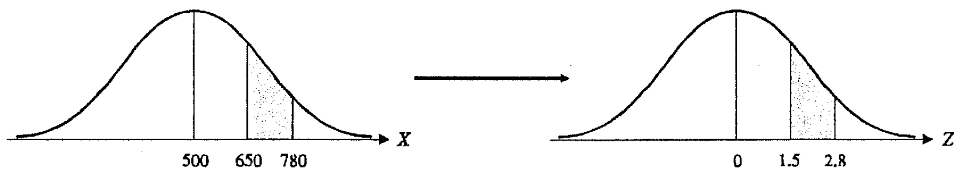
Así, $P(40 < X < 65) = 0.9759$.

Siempre es una buena idea dibujar ilustraciones de las curvas normales cuando intentamos calcular probabilidades asociadas con una distribución normal, pero no es necesario que se dibujen con gran precisión, porque su función principal es identificar las regiones apropiadas para calcular las áreas. Las ilustraciones de las curvas normales usadas en este texto pueden no ser técnicamente precisas porque se usan únicamente con intención conceptual.

La fórmula del puntaje z nos permite pensar en la tabla de probabilidad de la normal estándar, como en una lista de probabilidades asociadas con intervalos que comienzan en la media y se extienden a z desviaciones estándar, a la derecha de la media de cualquier curva normal. Considere las aplicaciones 7.12 y 7.13.

APLICACIÓN 7.12

Una compañía fabrica focos con vida media de 500 horas y desviación estándar de 100; si se supone que los tiempos de vida útil de los focos se distribuyen normalmente, esto es que los tiempos de vida forman una distribución normal, encuentre el número de focos de un total de 10 000 que puede esperarse duren entre 650 y 780 horas.



Solución: Primero calculamos los puntajes z para $x = 650$ y $x = 780$. El puntaje z para $x = 650$ es

$$\frac{650 - 500}{100} = 1.5$$

El puntaje z para $x = 780$ es

$$\frac{780 - 500}{100} = 2.8$$

Así,

$$\begin{aligned} P(650 < X < 780) &= P(1.5 < Z < 2.8) \\ &= 0.4974 - 0.4332 = 0.0642 \end{aligned}$$

Por tanto, podemos esperar que (10,000) (0.0642) = 642 focos duren entre 650 y 780 horas. ■

En aplicaciones prácticas de estadística tratamos, en primera instancia, con distribuciones finitas de medidas correspondientes a datos en bruto, que solo pueden ser aproximaciones de distribuciones normales, porque éstas son continuas y encierran un número infinito de medidas. En la literatura sobre estadística es un lugar común referirse a tales distribuciones como "normales" aun cuando en el contexto resulte evidente que tales distribuciones solo pueden aproximarse a una normal. Éste es el caso del planteamiento de las aplicaciones 7.12 y 7.13.

APLICACIÓN 7.13

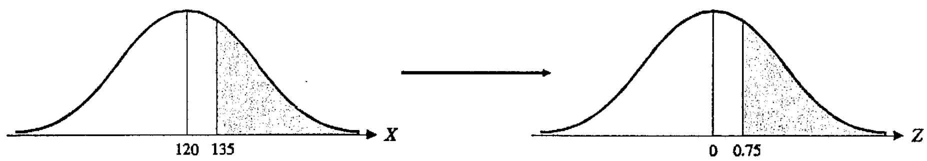
Si las medidas X de la presión sanguínea sistólica, de un grupo cuyas edades oscilan entre 20 y 24 años, se distribuyen normalmente con una media de 120 y una desviación estándar de 20, encuentre:

a) $P(X > 135)$.
b) $P(X < 146)$.
a) $P(105 < X < 110)$.

Solución:

a) El puntaje z para $x = 135$ es:

$$z = \frac{x - \mu}{\sigma}$$
$$= \frac{135 - 120}{20} = 0.75$$

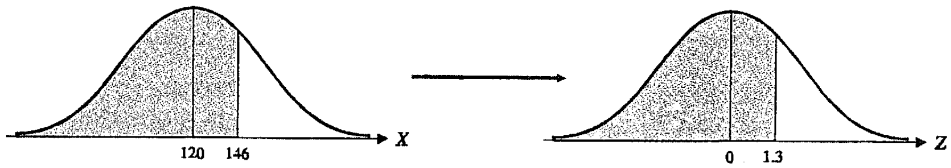


Si usamos la tabla z en Apéndices, encontraremos:

$$P(X > 135) = P(Z > 0.75)$$
$$= 0.5 - P(0 < Z < 0.75)$$
$$= 0.5 - 0.2734 = 0.2266$$

b) El puntaje z para $x = 146$ es:

$$z = \frac{x - \mu}{\sigma}$$
$$= \frac{146 - 120}{20} = 1.3$$



De la tabla z:

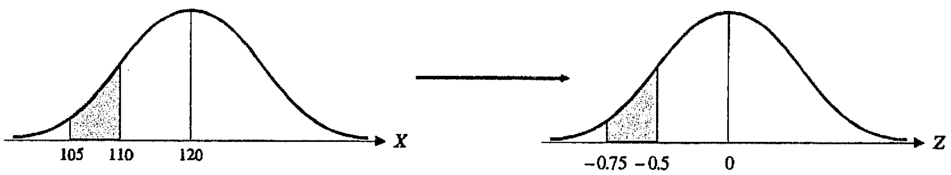
$$P(X < 146) = P(Z < 1.3)$$
$$= 0.5 + P(0 < Z < 1.3)$$
$$= 0.5 + 0.4032 = 0.9032$$

c) El puntaje z para $x = 105$ es

$$\frac{105 - 120}{20} = -0.75$$

y el puntaje z para $x = 110$ es

$$\frac{110 - 120}{20} = -0.5$$



Si usamos la tabla z, hallamos:

$$P(105 < X < 110) = P(-0.75 < Z < -0.5)$$
$$= P(0 < Z < 0.75) - P(0 < Z < 0.5)$$
$$= 0.2734 - 0.1915 = 0.0819 \quad ■$$

Percentiles, cuartiles y deciles asociados con distribuciones estándar

Recordemos que los puntos de posición se discutieron en el capítulo 3, tres medidas de posición usadas comúnmente son percentiles, cuartiles y deciles. Los **percentiles** son los números que dividen a un intervalo de medidas en 100 partes en forma tal, que cada parte contiene el 1% de dichas medidas. Hay 99 puntos percentiles, identificados por $P_1, P_2, \ldots, P_{99}$. En otras palabras, etiquetamos el n-ésimo percentil con P_n; siendo el septuagésimo quinto percentil, P_{75}, aquella medida bajo la cual queda el 75% de las medidas.

Los **cuartiles** son números que dividen un intervalo de medidas, que van de la medida mínima a la máxima, en cuatro partes iguales que contienen el 25% de las medidas. Hay tres puntos cuartiles que se denotan por Q_1, Q_2 y Q_3 e identificamos el n-nésimo cuartil por Q_n; siendo el segundo cuartil el quincuagésimo percentil o la mediana; el primer cuartil Q_1 es aquel número en que cae el 25% de las medidas debajo de él, es también igual a P_{25} y al percentil vigésimo quinto.

Los **deciles** son números que dividen un intervalo de medidas, desde la mínima hasta la máxima, en 10 partes tales que cada parte contiene al 10% de esas medidas. Escribimos como D_n el n-ésimo decil y denotamos los nueve puntos deciles por D_1, D_2, D_3, ..., D_9. Las aplicaciones 7.14 y 7.15 muestran cómo se encuentran estos puntos de posición para distribuciones normales.

APLICACIÓN 7.14

Las calificaciones de matemáticas del SAT se distribuyen normalmente con una media de 500 y una desviación estándar de 100. Encuentre P_{80}, el octagésimo percentil.

Solución: P_{80} es aquel número o calificación bajo el cual cae el 80% de las calificaciones; entonces el área bajo la curva normal entre $x = 500$ y P_{80} es 0.30 (véase la Fig. 7.16); si localizamos el puntaje z en la tabla z correspondiente a un área de 0.30, podemos usar la fórmula del puntaje z para encontrar P_{80}, porque

$$z = \frac{P_{80} - 500}{100}$$

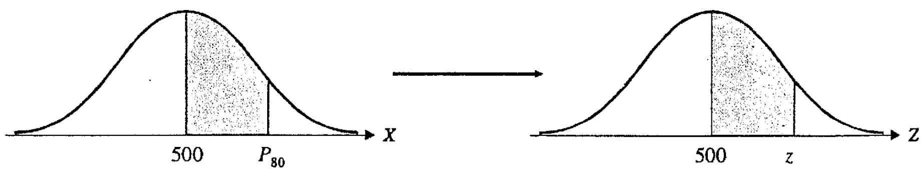


FIGURA 7.16
Determinación de P_{80} para las calificaciones de matemáticas del SAT

El valor aproximado de z que produce un área de 0.30 resulta ser, según la tabla z, 0.84. En consecuencia, tenemos:

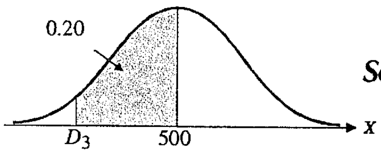
$$0.84 = \frac{P_{80} - 500}{100}$$

Multiplicamos ambos lados de esta ecuación por 100 y luego sumamos 500 también en ambos lados y obtenemos:

$$P_{80} = 584 \quad \blacksquare$$

APLICACIÓN 7.15 Para los puntajes de matemáticas del SAT con $\mu = 500$ y $\sigma = 100$, encuentre:

a) D_3, el tercer decil.
b) Q_1, el primer cuartil.



Solución:

a) Encontramos el puntaje z para D_3.

$$z = \frac{D_3 - 500}{100}$$

Si usamos la tabla z en Apéndices, encontraremos: $P(0 < Z < 0.52)$ ' 0.20. Así, el puntaje z para el tercer decil es:

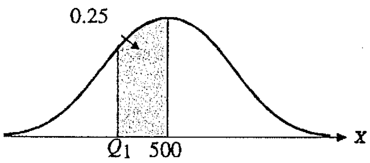
$$-0.52 = \frac{D_3 - 500}{100}$$

Si multiplicamos ambos lados por 100 y después sumamos 500 a ambos lados, obtenemos:

$$D_3 = 448$$

b) Encontremos el puntaje z para Q_1 como sigue:

$$z = \frac{Q_1 - 500}{100}$$



De la tabla z encontramos:

$$P(0 < Z < 0.67) = 0.25$$

Así,

$$-0.67 = \frac{Q_1 - 500}{100}$$

Si despejamos Q_1 tendremos:

$$Q_1 = 433 \quad \blacksquare$$

Una medida x tiene un **rango percentil** de n si x es igual, o aproximadamente igual, al n-ésimo percentil, esto es, si $x = P_n$. En la aplicación 7.16 vemos cómo se encuentra el rango percentil de una medida dada.

APLICACIÓN 7.16 Juan presenta el examen de matemáticas del SAT y obtiene una calificación de 572, encuentre su rango percentil.

Solución: Considere la figura 7.17; para encontrar el rango percentil de la calificación de Juan, ubicamos el área sombreada en la figura y la redondeamos al porcentaje más cercano. Localizaremos el área determinando primero el puntaje z para $x = 572$:

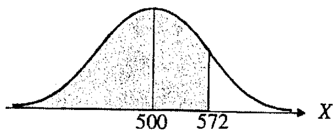
$$z = \frac{x - \mu}{\sigma}$$
$$= \frac{572 - 500}{100} = 0.72$$

El área sombreada en la figura 7.17 es:

$$\begin{aligned} P(X < 572) &= P(Z < 0.72) \\ &= 0.5 + P(0 < Z < 0.72) \\ &= 0.5 + 0.2642 = 0.7642 \end{aligned}$$

En consecuencia, $P_{76} \approx 572$ y el rango percentil de 572 es 76. ■

FIGURA 7.17
Determinación del rango percentil de la calificación 572 en matemáticas del SAT



Se puede usar MINITAB para determinar el rango percentil de la calificación de Juan; las órdenes y las respuestas están contenidas en la pantalla 7.1.

Pantalla 7.1

```
MTB > CDF 572;
SUBC > NORMAL 500, 100.
   572.0000 0.7642
MTB >
```

La orden para la función de distribución acumulativa, CDF por sus siglas en inglés, y la suborden asociada NORMAL determinan la probabilidad $P(X \leq 572)$ para la variable aleatoria normal X que tiene $\mu = 500$ y $\sigma = 100$.

Verificación de la posibilidad de que una muestra proceda de una distribución normal

Para determinar cuándo una distribución normal puede servir como un modelo razonable para la población que produjo la muestra, es útil recurrir a procedimientos gráficos que permitan detectar diferencias serias con la normalidad; los histogramas pueden exhibir falta de simetría, y la regla empírica se puede usar para verificar las proporciones de los datos que caen en los intervalos determinados por $\bar{x} \pm s$, $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$.

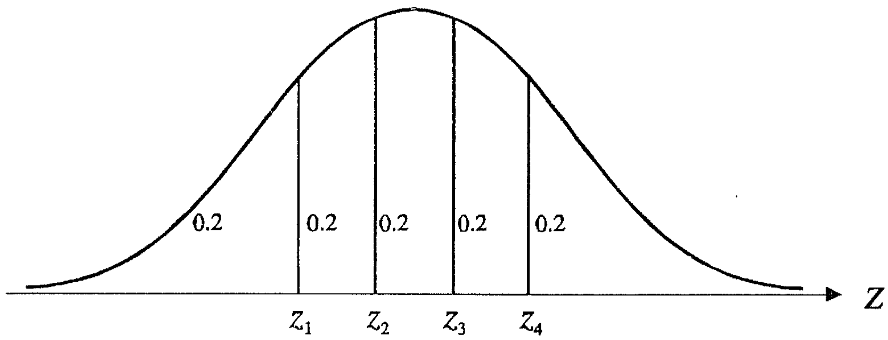
Una **gráfica de los puntajes** z de los datos y los correspondientes puntos percentiles de la distribución normal estándar, puede ser un procedimiento efectivo para verificar la posibilidad de un modelo normal cuando el tamaño de la muestra es menor que 20; si el tamaño es n, se pueden encontrar n valores normales estándar que dividan la distribución normal estándar en $(n + 1)$ intervalos, cada uno de los cuales tiene la misma probabilidad. Las medidas de la muestra se ordenan de menor a mayor y se aparean con los n puntos percentiles, se localizan entonces las parejas para obtener una gráfica de los puntajes z. Para facilitar el análisis, suponga que la muestra es de tamaño 4; primero quisiéramos encontrar los cuatro puntos percentiles P_{20}, P_{40}, P_{60} y P_{80} que dividen a la distribución normal estándar en cinco intervalos, cada uno con una probabilidad de 0.20. El valor mínimo de la muestra se aparea con P_{20}, el siguiente en tamaño con P_{40}, el siguiente con P_{60} y el máximo con P_{80}; estas cuatro parejas se localizan en una gráfica. Un patrón de línea recta apoyará la probabilidad de que la muestra se ha extraído de una población normal; una diferencia con la normalidad la constituye que la gráfica tenga una apariencia curva. El patrón de línea recta también se puede detectar mediante el coeficiente de correlación de Pearson; los valores de $|r|$ cercanos a 1 indican una relación lineal fuerte. La aplicación 7.17 ilustra únicamente las ideas; en las aplicaciones prácticas, el tamaño de la muestra debe ser al menos de 20.

APLICACIÓN 7.17

Construya una gráfica de puntajes z de la muestra siguiente de datos para determinar si pudo haber sido producida por una población normal.

y: 48 55 57 60

Solución: Los cuatro puntos percentiles que dividen a la distribución en cinco intervalos igualmente probables, se encuentran con ayuda de la tabla z en Apéndices (véase la figura adjunta): $P(Z \le -0.84) = 0.2$, $P(Z \le -0.25) = 0.4$, $P(Z \le 0.25) = 0.6$, $P(Z \le 0.84) = 0.8$.

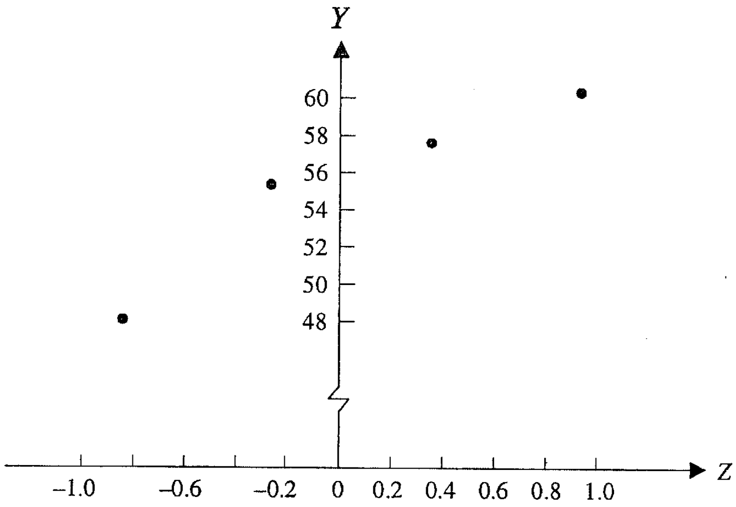


Estos puntajes z se aparean entonces con las cuatro observaciones:

Observaciones muestrales, y	Puntos porcentiles, puntajes z
48	-0.84 ($= z_1$)
55	-0.25 ($= z_2$)
57	0.25 ($= z_3$)
60	0.84 ($= z_4$)

La gráfica de puntajes z se muestra en la figura 7.18.

FIGURA 7.18
Gráfica de puntajes z para la aplicación 7:17



Como los cuatro puntos de los datos parecen seguir un patrón lineal, no tenemos razón para creer que la muestra no fue producida por una población normal; además, el coeficiente de correlación de Pearson para los cuatro pares de datos es $r = 0.965$, lo cual apoya el fuerte patrón lineal de la gráfica de los puntajes z. Si la distribución (X) que produjo la muestra tiene una media μ y una desviación estándar σ, cada punto percentil z se puede convertir en un puntaje x mediante la transformación $x = \mu + \sigma z$. Esperaríamos que cada valor muestral fuera cercano a su correspondiente valor x si la población es normal, es decir, una gráfica de los valores muestrales contra los puntajes normales x debería producir un patrón de línea recta ($y = \mu + \sigma z$), donde la pendiente de la recta es σ y la intersección con el eje vertical es μ.

Valor muestral (y)	Valor x
48	$\mu + \sigma_{z_1}$
55	$\mu + \sigma_{z_2}$
57	$\mu + \sigma_{z_3}$
60	$\mu + \sigma_{z_4}$

■

GRUPO DE EJERCICIOS 7.3

Habilidades básicas

1. Si una variable aleatoria X se distribuye normalmente con una media de 70 y una desviación estándar de 5, encuentre la probabilidad de que X:
   a) esté entre 60 y 80;
   b) sea menor que 65;
   c) sea menor que 81;
   d) sea mayor que 67.5;
   e) sea menor que 67 o mayor que 77.

2. Si una variable aleatoria X se distribuye normalmente con una media de 50 y una desviación estándar de 8, encuentre la probabilidad de que X:
   a) sea mayor que 60;
   b) sea menor que 65;
   c) esté entre 40 y 60;
   d) sea menor que 38;
   e) sea menor que 39 y mayor que 65.

Más aplicaciones

3. Una máquina produce pernos cuyos diámetros se distribuyen normalmente con un promedio de 0.25 pulgadas y una desviación estándar de 0.02 pulgadas. ¿Cuál es la probabilidad de que un perno tenga un diámetro:

a) mayor de 0.3 pulgadas?;
b) entre 0.2 y 0.3 pulgadas?;
c) menor que 0.19 pulgadas?;
d) mayor que 0.1 pulgadas?

4. Cierta marca de focos tiene una vida media de 750 horas y una desviación estándar de 75. Si sus tiempos de vida se distribuyen normalmente, ¿cuál es la probabilidad de que un foco dure:
a) entre 750 y 850 horas?;
b) entre 785 y 800 horas?;
c) más de 700 horas?;
d) menos de 800 horas?

5. Para los puntajes de matemáticas del SAT, que se distribuyen normalmente con una media de 500 y una desviación estándar de 100, encuentre:
a) P_{30}.
b) P_{50}.
c) P_{70}.
d) P_{90}.

6. Para los mismos puntajes del ejercicio anterior calcule:
a) Q_1.
b) D_7.
c) D_3.
d) Q_3.

7. Refiérase al ejercicio 3. Si se produjeron 100,000 pernos, encuentre el número esperado con un diámetro:
a) mayor de 0.3 pulgadas;
b) entre 0.2 y 0.3 pulgadas;
c) menor de 0.19 pulgadas;
d) mayor que 0.1 pulgadas.

8. Refiérase al ejercicio 4. Si se produjeron 100,000 focos, encuentre el número esperado con un tiempo de vida:
a) entre 750 y 850 horas;
b) entre 785 y 800 horas;
c) mayor de 700 horas;
d) mayor de 800 horas.

9. La Apex Taxi Company ha encontrado que el costo de los viajes se distribuye normalmente con una μ = 4.30 dólares y σ = 1.25 dólares. Si un chofer acude a una llamada escogida al azar, ¿cuál es la probabilidad de que su costo sea menor de 3.05 dólares?

10. La media GPA en una universidad es 2.2 y la desviación estándar 0.5. En una convocatoria honorífica especial, se dará un reconocimiento a los estudiantes con GPA dentro del 3% más alto; si los GPA se distribuyen normalmente, ¿cuál es el mínimo que debe tener un estudiante para recibir el reconocimiento?

11. Las estaturas de los hombres adultos se distribuyen normalmente con una media de 70 pulgadas y una desviación estándar de 2.6. ¿Qué tan alta debe ser una puerta que se va a construir si se quiere que el 90% de los hombres pasen por ella sin tener que agacharse?

12. Una cafetera se programa para llenar tazas con un promedio de 8 onzas de café. Si las cantidades servidas se distribuyen normalmente con una desviación estándar de 0.4 onzas, ¿qué porcentaje de tazas de 9 onzas llenará hasta derramar el café?

13. El número de horas que un universitario dedica por semana al estudio se distribuye normalmente con una media de 25 y una desviación estándar de 10.
a) ¿Qué porcentaje de estudiantes estudia menos de 30 horas?
b) ¿Qué porcentaje estudia menos de 20 horas?
c) ¿Qué porcentaje estudia más de 30 horas?
d) De una clase de 100 estudiantes, ¿aproximadamente cuántos estudian entre 12 y 38 horas por semana?

14. Si X se distribuye normalmente con $\mu = 30$ y $\sigma = 5$, encuentre el valor x_0 tal que:
a) $P(X > x_0) = 0.05$ b) $P(X < x_0) = 0.005$

Un paso más allá

15. La vida útil de una marca particular de tubos de rayos catódicos (CRT por sus siglas en inglés), usado en una televisión a color se distribuye normalmente con media de 7.7 años y una desviación estándar de 1.9 años.
a) ¿Cuál es la probabilidad de que un CRT dure más de 8 años?
b) Si el CRT está garantizado por 2 años, ¿qué porcentaje de tubos deberá ser remplazado?
c) Si el fabricante quiere remplazar solo el 2% de los CRTs, ¿qué periodo de garantía, aproximado al mes más cercano, deberá establecer en la póliza respectiva?

16. Un fabricante de motores eléctricos quiere determinar el tiempo que debe garantizar sus motores de forma tal que solo el 4% de ellos deban ser remplazados. Si las vidas de los motores se distribuyen normalmente con una media de 10 años y una desviación estándar de 0.75, encuentre el periodo de garantía aproximado al mes más cercano.

17. Si la variable aleatoria X se distribuye normalmente con media μ y varianza σ^2, encuentre el rango percentil de:
a) $\mu + \sigma$. c) $\mu + 2\sigma$.
b) $\mu - \sigma$. d) $\mu - 2\sigma$.

18. Construya una gráfica de puntajes z para determinar si una distribución normal puede servir como modelo de la muestra siguiente, es decir, estime si esta muestra pudo haberse extraído de una población normal.

36	58	42	37	24
52	49	56	34	32
44	53	51	47	37
51	30	31	47	50
43	31	53	44	27

19. Si u_1 y u_2 son dos valores aleatorios en el intervalo (0, 1), entonces las fórmulas siguientes se pueden usar para generar dos variables aleatorias normales estándar:

$$z_1 = \sqrt{-\ln(u_2)}\cos(2\pi u_1)$$
$$z_2 = \sqrt{-\ln(u_2)}\operatorname{sen}(2\pi u_1)$$

donde el ángulo se expresa en radianes. Los valores de u_1 y u_2 se pueden obtener usando un generador de números aleatorios o una tabla de números aleatorios y escogiendo tres dígitos al mismo tiempo. Use estas fórmulas para simular dos observaciones de la distribución normal que tiene una media de 30 y una desviación estándar de 5, mediante las relaciones $x_1 = \mu + \sigma z_1$ y $x_2 = \mu + \sigma z_2$

20. a) Escriba un programa computacional para simular 25 observaciones de la distribución normal que tiene una media de 30 y una desviación estándar de 5.
b) Calcule la media y la desviación estándar de las 25 medidas del inciso a y construya un histograma para verificar que la muestra es de la distribución allí mencionada.

SECCIÓN 7.4 *Uso de distribuciones normales para aproximar distribuciones binomiales*

Hay muchos experimentos binomiales para los cuales el cálculo de probabilidades binomiales requiere de un gran número de operaciones largas y tediosas.

EJEMPLO 7.7

Considere el experimento binomial siguiente.

> Durante un periodo largo, se ha determinado que el 70% de los abogados que presentan el examen de la barra de abogados lo aprueban. De 500 abogados que presentan el examen siguiente, encuentre la probabilidad de que al menos 370 de ellos pasen.

Para resolver este problema, necesitamos encontrar $P(370) + P(372) + \cdots + P(499) + P(500)$, esto requeriría usar 131 veces la fórmula de probabilidad binomial. Intentemos calcular $P(370)$ en primer lugar.
Si usamos la fórmula (6.2), obtenemos:

$$P(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$
$$P(370) = \binom{500}{370} (0.7)^{370} (0.3)^{130}$$

Ahora tratamos de encontrar el coeficiente binomial $\binom{500}{370}$.

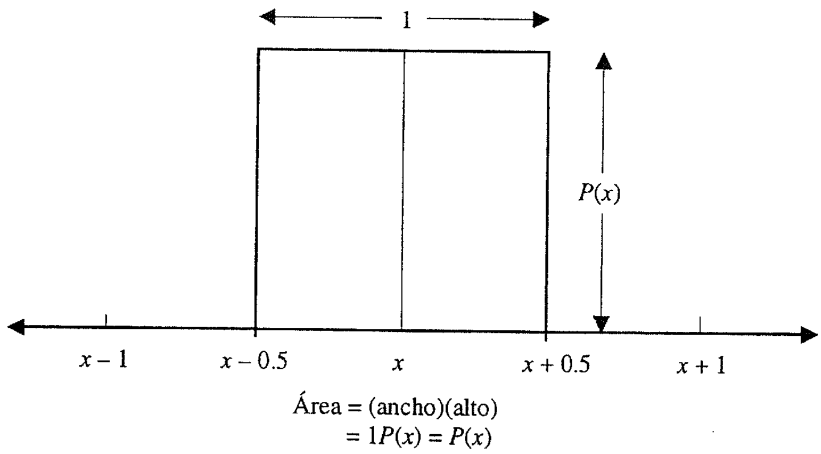
$$\binom{500}{370} = \frac{(500)!}{(370)!\,(130)!}$$

Una calculadora manual no nos ayudará en este cálculo, y no deberá sorprendernos que la computadora escolar tampoco lo haga. Más adelante, en la aplicación 7.18, usaremos una distribución normal para ayudarnos a aproximar la probabilidad de que al menos 370 abogados aprueben el examen de la barra.

Gráficas de barras para distribuciones binomiales

Recuerde que usamos gráficas de segmentos de recta verticales en el capítulo 6 para representar distribuciones binomiales de probabilidad binomial. Otro tipo de gráfica de probabilidad, llamada *gráfica de barras* o *histograma*, utiliza barras verticales en lugar de segmentos de recta verticales. Por cada valor de x en el eje horizontal se marcan dos puntos en $x - 0.5$ y $x + 0.5$; la distancia entre estos puntos, que es 1, será la longitud del lado correspondiente al ancho de una barra rectangular y la longitud del otro lado será $P(x)$. El área de la barra x-ésima es entonces $P(x)$, como lo indica la figura 7.19. Note que el área bajo una gráfica de barras de probabilidad es 1, porque $\Sigma P(x) = 1$; aunque las áreas se asocian generalmente con distribuciones continuas de probabilidad, deberemos tener cuidado de no clasificar una distribución binomial como continua; el uso de una gráfica de barras para una distribución binomial se debe únicamente a la conveniencia, y es especialmente útil cuando se aproximan probabilidades binomiales.

FIGURA 7.19
Gráfica de barras de probabilidad

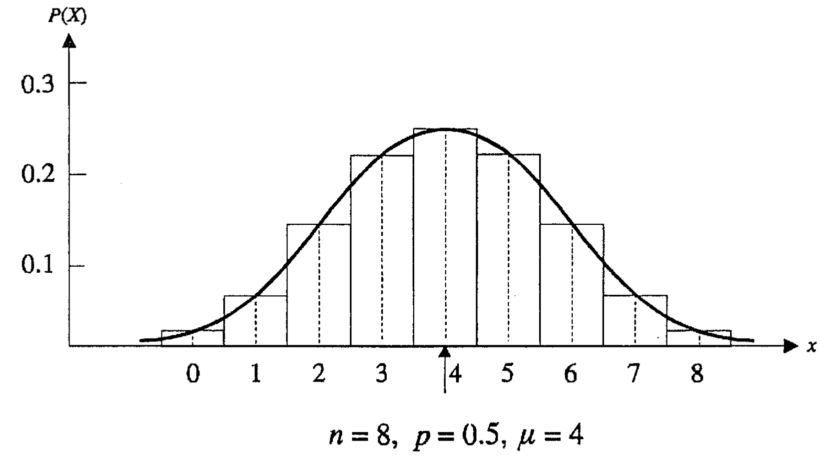


Para una distribución binomial que es simétrica, o aproximadamente simétrica respecto a su media, la gráfica de barras se parece mucho a una distribución normal, como resultado, una distribución normal se puede usar para aproximar una distribución binomial (véase la figura 7.20).

Recuerde que en cualquier distribución normal con media μ y desviación estándar σ, el 99.7%, para propósitos prácticos casi toda la distribución, está entre $\mu - 3\sigma$ y $\mu + 3\sigma$. Suponga que una distribución binomial tiene media $\mu = np$ y desviación estándar $\sigma = \sqrt{np(1-p)}$, si la vamos a aproximar por una distribución normal, debe tener media np y desviación estándar igual a $\sqrt{np(1-p)}$. Como los valores de la distribución normal son mayores o iguales a cero, y menores o iguales a n, debemos tener $0 \le \mu - 3\sigma$ y $\mu + 3\sigma \le n$,

FIGURA 7.20

Aproximación de una distribución binomial por una distribución normal



donde μ y σ representan la media y la desviación estándar, respectivamente, para la binomial. Si es así, deben cumplirse las dos desigualdades siguientes:

1. $0 \leq np - 3\sqrt{np(1-p)}$
2. $np + 3\sqrt{np(1-p)} \leq n$

Si $p \leq 0.5$, se puede demostrar que las desigualdades son equivalentes a $np \geq 4.5$, y si $p \geq 0.5$, serán equivalentes a $n(1-p) \geq 4.5$. Así, establecemos la regla:

> Una distribución binomial se puede aproximar por una distribución normal si $np \geq 5$ y $n(1-p) \geq 5$.

Observe que para valores grandes de n ambas desigualdades deben ser ciertas, especialmente para valores pequeños de p o de $1-p$.

EJEMPLO 7.8

Si $p = 0.01$ y $n = 600$, entonces esta distribución binomial se puede aproximar por una distribución normal, pues $np = (600)(0.01) = 6$, y $n(1-p) = (600)(0.99) = 594$.

Para usar una distribución normal como aproximación de una distribución binomial, seguimos los pasos:

1. Verifique la validez de las desigualdades $np \geq 5$ y $n(1-p) \geq 5$.
2. Si eso ocurre, calcule μ y σ para la distribución binomial.
3. Se usa la distribución normal con media μ y desviación estándar σ para la aproximación.
4. Construya una gráfica de barras para la distribución binomial y determine los límites para encontrar áreas bajo la curva normal.
5. Encuentre los puntajes z para estos límites y use la normal estándar para encontrar el(las) área(s) correspondiente(s) ; este número será la aproximación a la probabilidad binomial buscada.

APLICACIÓN 7.18

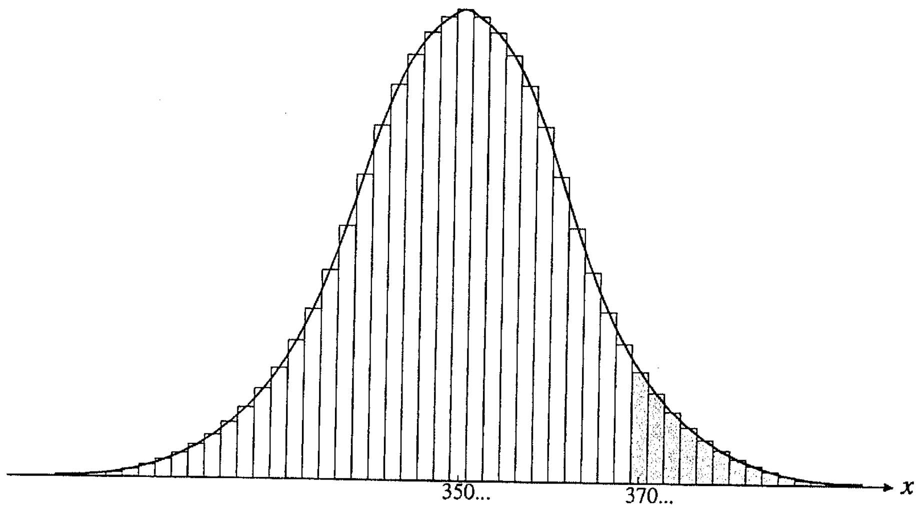
Encuentre $P(x \geq 370)$ para el experimento binomial analizado en el ejemplo 7.7, al principio de esta sección, donde $n = 500$ y $p = 0.70$.

Solución: Como $np = (500)(0.7) = 350 \geq 5$ y $n(1-p) = (500)(0.3) = 150 \geq 5$, es posible aproximar por una normal. Si usamos las fórmulas (6.4) y (6.7), tenemos:

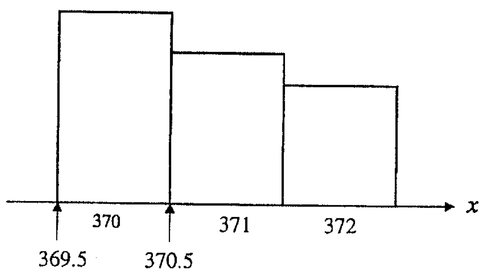
$$\begin{aligned} \mu &= np = 350 \\ \sigma &= \sqrt{np(1-p)} \\ &= \sqrt{(350)(0.3)} = 10.25 \end{aligned}$$

Así, la distribución normal con $\mu = 350$ y $\sigma = 10.25$, será utilizada con propósitos de aproximación.

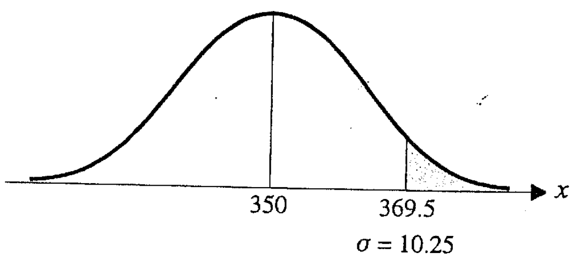
Se muestra aquí una gráfica de barras para la binomial junto con la distribución normal con $\mu = 350$ y $\sigma = 10.25$.



Queremos encontrar la suma de las áreas de las barras sombreadas, y como aproximación a esa suma encontraremos el área bajo la curva normal a la derecha de 369.5. Para ver esto, observe que la barra marcada con 370 tiene un ancho de uno y se extiende de 369.5 a 370.5; una distancia de 1, como lo ilustra la figura:



En consecuencia, encontramos el área sombreada bajo la curva normal siguiente:



Para obtener el área sombreada, encontramos el puntaje z para 369.5:

$$\begin{aligned} z &= \frac{x - \mu}{\sigma} \\ &= \frac{369.5 - 350}{10.25} = 1.90 \end{aligned}$$

De la tabla normal estándar, obtenemos:

$$\begin{aligned} P(Z \geq 1.90) &= 0.5 - 0.4713 \\ &= 0.0287 \end{aligned}$$

Así,

$$P(X \geq 370) \approx 0.0287 \quad ■$$

Se puede usar MINITAB para determinar $P(X < 370)$ para el problema del examen ante la barra de abogados. Determinamos $P(X \geq 370)$ restando. La pantalla 7.2 deja ver las órdenes de MINITAB usadas y la respuesta correspondiente.

Pantalla 7.2

```
MTB > CDF 369.5;
SUBC > NORMAL 350 10.25.
   369.5000 0.9714
MTB >
```

La tercera línea indica $P(X < 370) = 0.9714$. Al restar encontramos:

$$P(X \geq 370) = 1 - 0.9714 = 0.0286$$

La diferencia de 0.0001 entre la respuesta encontrada en la aplicación 7.18 y la obtenida con la ayuda de MINITAB, se debe a la mayor precisión numérica que se logra con MINITAB.

APLICACIÓN 7.19

Si se sabe que el 60% del ganado inoculado con un suero queda protegido de cierta enfermedad, encuentre la probabilidad de que 8 o 9, de 15 vacas inoculadas, no contraigan la enfermedad, utilice:

a) La fórmula de probabilidad binomial.
b) Una distribución normal que aproxime a la binomial.
c) Las tablas de probabilidad binomial.

Después compare los resultados.

Solución: Tenemos $n = 15$, $p = 0.6$ y $x = 8, 9$.

a) Mediante la fórmula de probabilidad binomial (6.2), tenemos:

$$\begin{aligned} P(8) &= \binom{15}{8}(0.6)^8(0.4)^7 \\ &= (6435)(0.6)^8(0.4)^7 = 0.177 \end{aligned}$$

y

$$P(9) = \binom{15}{9}(0.6)^9(0.4)^6$$
$$= (5005)(0.6)^9(0.4)^6 = 0.207$$

En consecuencia, la probabilidad de que 8 o 9 de las 15 vacas inoculadas no contraigan la enfermedad es:

$$P(8) + P(9) = 0.177 + 0.207 = 0.384$$

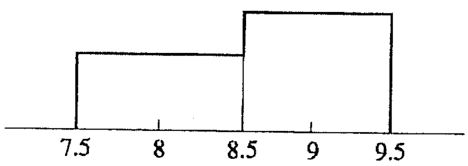
b) Primero, notemos que $np = (15)\ (0.6) = 9$ y $n(1-p) = (15)\ (0.4) = 6$, ambos son mayores que 5. Si usamos la fórmula (6.4), tenemos:

$$\mu = np = 9$$

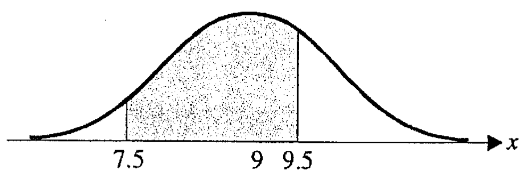
y usando la fórmula (6.7) resulta:

$$\sigma = \sqrt{np(1-p)}$$
$$= \sqrt{(9)(0.4)} = 1.897$$

A continuación dibujamos parte de la distribución binomial para determinar los límites para el área bajo la curva normal con $\mu = 9$ y $\sigma = 1.897$.



Nos interesa determinar en la figura el área bajo la curva normal 7.5 y 9.5.



Ahora, encontremos los puntajes z de 7.5 y 9.5. El puntaje z para $x = 7.5$ es:

$$z = \frac{7.5 - 9}{1.897} = -0.79$$

y para $x = 9.5$:

$$z = \frac{9.5 - 9}{1.897} = 0.26$$

En la tabla normal estándar encontramos que el área bajo la curva normal estándar y arriba del eje horizontal entre $z = -0.79$ y $z = 0.26$ es:

$$P(-0.79 < Z < 0.26) = 0.2852 + 0.1026 = 0.3878$$

Este valor, comparado con 0.3840, representa menos de 1% de error; la aproximación es buena aun para un n tan pequeño como 15.

c) Si usamos las tablas de probabilidad binomial, encontramos $P(8) = 0.177$ y $P(9) = 0.207$. Así,

$$\begin{aligned} P(8 \text{ o } 9) &= P(8) + P(9) \\ &= 0.177 + 0.207 \\ &= 0.384 \end{aligned}$$

lo que concuerda con el resultado del inciso a. ■

GRUPO DE EJERCICIOS 7.4

Habilidades básicas

1. Para un experimento binomial con $p = 0.5$ y $n = 10$, encuentre $P(4 \le X \le 6)$ usando:
   a) La fórmula de probabilidad binomial.
   b) Una aproximación normal.
   c) Las tablas de probabilidad binomial.
   Compare los resultados.

2. Para un experimento binomial con $p = 0.6$ y $n = 15$, calcule $P(8 \le X \le 10)$ con:
   a) La fórmula de probabilidad binomial.
   b) Una aproximación normal.
   c) Las tablas de probabilidad binomial.
   Compare los resultados.

Más aplicaciones

3. La probabilidad de que la pelusa provoque una reacción adversa es 0.02. Si 1000 individuos se escogen aleatoriamente para exponerlos a la pelusa, encuentre la probabilidad aproximada de que:
   a) a lo más 20 personas tengan una reacción adversa.
   b) al menos 24 personas tengan una reacción adversa.
   c) entre 21 y 29, inclusive, presenten reacción adversa.
   d) 210 personas reaccionen desfavorablemente.

4. Suponga que la probabilidad de que una perinola caiga apuntando hacia arriba es 0.7; si se baila 1000 veces encuentre la probabilidad de que:
   a) 715 o más veces caiga hacia arriba.
   b) entre 685 y 720 eventos inclusive, señale hacia arriba.
   c) 675 caiga hacia arriba.
   d) a lo más 730 apunte hacia arriba.

5. Si una moneda se lanza 1000 veces, encuentre la probabilidad aproximada de que:
   a) a lo más 530 veces caiga mostrando "sol".
   b) entre 485 y 520, inclusive, lanzamientos den lugar a "soles".
   c) 525 lanzamientos den "sol".
   d) el número de lanzamientos que muestren "sol" sea mayor de 490.

6. El 60% de las familias de cierta ciudad tienen casa propia, si se escogen al azar 500 familias, ¿cuál es la probabilidad aproximada de que:
   a) al menos 315 tengan casa propia?
   b) ¿entre 290 y 310, inclusive, posean su casa?
   c) ¿305 sean propietarios?
   d) a lo más sean 320 los que tengan casa propia?

7. En un programa de control de calidad para un proceso de producción se considera que este es satisfactorio solo si de una muestra de 100 piezas, menos de 10 resulten defectuosas; ¿cuál es la probabilidad aproximada de que en un lote de 100 artículos el proceso de producción se considere:
   a) Satisfactorio cuando, de hecho, el proceso produzca 15% de artículos defectuosos?
   b) No satisfactorio cuando se produzca 15% de artículos defectuosos?

8. Para determinar si un fármaco nuevo es efectivo, será administrado a 500 pacientes; si 440 o más registran resultados positivos, la droga se clasificará como efectiva, de otra forma, se señalará como no efectiva. Encuentre la probabilidad aproximada de que el fármaco se considere como no efectivo aunque, de hecho, resulte positivo en el 90% de los casos.

9. La probabilidad de que una marca de perinolas señale hacia arriba al caer es 0.60; si se bailan 500 perinolas, ¿cuál es la probabilidad de que 325 o menos caigan hacia arriba?

10. Si se lanzan 1000 monedas el resultado más probable es 500 "soles"; use la aproximación normal para encontrar la probabilidad de obtener 500 "soles".

Un paso más allá

11. La ecuación de la *curva normal estándar* es

$$f(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$$

El resultado del ejercicio 10 podrá aproximarse multiplicando $f(0)$ por $1/\sigma$, pues el ancho de la quingentésima barra en unidades x es 1 y en unidades z es $1/\sigma$. Realice este cálculo y compare su respuesta con la obtenida en el ejercicio 10.

12. Demuestre que si $0 \leq np - 3\sqrt{np(1-p)}$ y $np + 3\sqrt{np(1-p)} \leq n$, entonces $np \geq 4.5$ o $n(1-p) \geq 4.5$

SECCIÓN 7.5 *Distribuciones exponenciales*

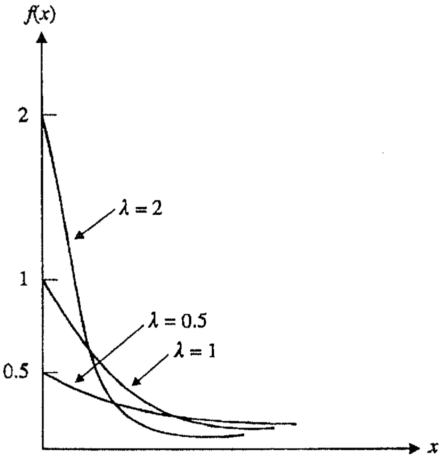
Una distribución estrechamente relacionada con la distribución de Poisson analizada en la sección 6.6, es la *distribución exponencial.* La distribución de Poisson se refiere al número de eventos que ocurren en un intervalo de tiempo, mientras que la distribución exponencial se puede usar para modelar los tiempos entre la ocurrencia de dos eventos sucesivos, cuya distribución a menudo resulta exponencial. Ejemplos:

- llegadas sucesivas de clientes a una estación de servicio.
- llamadas sucesivas que entran a un conmutador.
- caídas una tras otra de un sistema de cómputo.
- temblores sucesivos en cierta región.
- fallas continuas de un sistema equis.

Se dice que una variable aleatoria X tiene una distribución exponencial si la función de densidad de probabilidad (fdp) presenta la forma siguiente:

FIGURA 7.21

Gráfica de funciones de densidad exponencial



Función de densidad de probabilidad exponencial

$$f(x) = \lambda e^{-\lambda x}$$

donde $x > 0$ y $\lambda > 0$

Una **variable aleatoria exponencial** posee las propiedades:

1. Su media es igual a $1/\lambda$.
2. Su varianza es igual a $1/\lambda^2$.

Note que una variable aleatoria distribuida exponencialmente tiene su media igual a su desviación estándar; como la media y la desviación estándar de una variable así dependen solo de la constante λ, llamada **parámetro**, la forma de la gráfica de su fdp depende de λ. La figura 7.21 muestra gráficas de tres **funciones de densidad exponencial** correspondientes a $\lambda = 0.5$, $\lambda = 1$ y $\lambda = 2$. Observe que cada curva cruza el eje y en λ.

Como una variable aleatoria exponencial es continua, para encontrar probabilidades asociadas con ella debemos ser capaces de determinar áreas bajo la gráfica de su fdp y sobre el eje x. Se puede demostrar mediante cálculo que el área bajo la curva y sobre el eje x entre las rectas $x = 0$ y $x = c (c > 0)$ puede encontrarse evaluando la expresión $1 - e^{-\lambda c}$. La función F definida por $F(x) = 1 - e^{-\lambda x}$ $(x > 0)$ se llama **función de distribución acumulativa** (fda) para la variable aleatoria exponencial X que tiene parámetro λ.

Función de distribución acumulativa para una variable aleatoria exponencial

$$F(x) = 1 - e^{-\lambda x} \text{ donde } x > 0 \text{ y } \lambda > 0 \qquad (7.2)$$

EJEMPLO 7.9

Supongamos que X es una variable aleatoria distribuida exponencialmente con parámetro $\lambda = 0.5$. Calculemos la probabilidad de que el valor de la variable aleatoria sea menor que 3, es decir, determinemos $P(X < 3) = F(3)$; para determinar el valor $F(3)$ podemos usar la fórmula (7.2).

$$F(3) = 1 - e^{-(0.5)(3)} = 1 - e^{-1.5}$$

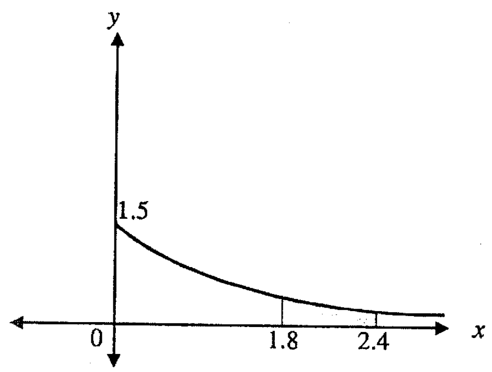
En este punto podemos elegir entre la tabla 3 del apéndice B o una calculadora con la capacidad de calcular el valor de la exponencial, e^{-6}. La tabla 3 proporciona los valores de e^{-x} para valores x de 0 a 10 con incrementos de 0.05, para usarla a fin de aproximar el valor de $e^{-1.5}$. Localizamos el valor de $x = 1.5$ en la primera columna y, junto a ella, en la segunda columna está el valor 0.2231. Así, el valor aproximado de $F(3)$ es

$$F(3) = 1 - e^{-1.5} \approx 1 - (0.2231) = 0.7769$$

APLICACIÓN 7.20

Si una variable aleatoria tiene la distribución exponencial con $\lambda = 1.5$, encuentre $P(1.8 \leq X \leq 2.4)$.

Solución: Note que $P(1.8 \leq X \leq 2.4) = P(X \leq 2.4) - P(X < 1.8) = F(2.4) - F(1.8)$, donde F es la función de distribución acumulativa.



Determinamos las probabilidades acumuladas $F(2.4)$ y $F(1.8)$ usando la fórmula (7.2).

$$F(2.4) = 1 - e^{-(1.5)(2.4)} = 1 - e^{-3.6}$$

y

$$F(1.8) = 1 - e^{-(1.5)(1.8)} = 1 - e^{-2.7}$$

Podemos usar la tabla 3 del apéndice B o una calculadora con capacidad para evaluar $e^{-3.6}$ y $e^{-2.7}$, determinamos las aproximaciones:

$$e^{-3.6} \approx 0.0273$$

y

$$e^{-2.7} \approx 0.0672$$

En consecuencia,

$$\begin{aligned} F(2.4) - F(1.8) &= (1 - e^{-3.6}) - (1 - e^{-2.7}) \\ &= (1 - 0.0273) - (1 - 0.0672) \\ &= 0.9727 - 0.9328 \approx 0.0399 \end{aligned}$$

Por tanto, $P(1.8 \leq X \leq 2.4)$ es aproximadamente igual a 0.04. ■

APLICACIÓN 7.21

Suponga que un usuario de computadoras está trabajando en una terminal de una red conectada a una central. El tiempo en segundos que le toma a la computadora central responder a una petición del usuario, tiene una distribución exponencial con un tiempo esperado de respuesta de 2 segundos. Determine la probabilidad de que dicha respuesta tarde a lo más 5 segundos.

Solución: Representemos por la variable aleatoria X el tiempo de respuesta del sistema. Como $E(X) = 2$ segundos y $E(X) = 1/\lambda$, sabiendo que $\lambda = 1/2$. Mediante la función de probabilidad acumulativa 7.2, tenemos:

$$\begin{aligned} P(X \leq 5) &= F(5) \\ &= 1 - e^{-(0.5)(5)} \\ &= 1 - e^{-2.5} \\ &= 1 - 0.0821 \quad \text{(de la tabla 3, apéndice)} \\ &= 0.9179 \end{aligned}$$

Por lo tanto, la probabilidad de que un usuario deba esperar a lo más 5 segundos una respuesta del sistema central de cómputo es 0.9179. ■

Se puede usar MINITAB para determinar $F(5) = P(X < 5)$ en la aplicación 7.21, con la orden CDF y la suborden EXPONENTIAL. La pantalla 7.3 deja ver las órdenes usadas y la respuesta correspondiente. Nota: con la suborden se proporciona el valor de la media, no el del parámetro.

Pantalla 7.3

```
MTB > CDF 5;
SUBC > EXPONENCIAL 2.
        5.0000    0.9179
```

APLICACIÓN 7.22

El tiempo en minutos entre las llamadas telefónicas que llegan a un conmutador se distribuye exponencialmente con parámetro $\lambda = 1.8$.

a) Encuentre la probabilidad de que haya una espera de al menos 1 minuto entre la primera y la segunda llamadas que llegan a ese conmutador cierto día.
b) Calcule la probabilidad de que haya una espera de al menos 1 minuto para cada uno de los tres primeros telefonemas.

Solución: Representemos por X el tiempo en minutos entre llamadas sucesivas. La función de distribución acumulativa es $F(X) = 1 - e^{-1.8x}$.

a) La probabilidad de que haya una espera de al menos un minuto entre la primera y la segunda llamadas es:

$$\begin{aligned} P(X \geq 1) &= 1 - F(1) \\ &= 1 - (1 - e^{-(1.8)(1)}) \\ &= e^{-1.8} \\ &= 0.1653 \quad \text{(de la tabla 3, apéndice B)} \end{aligned}$$

b) Como las llamadas llegan al conmutador de manera independiente, la probabilidad de que haya una espera de al menos 1 minuto para cada una de las tres primeras llamadas que llegan un cierto día es $(0.1653)^3 = 0.0045$. ■

La *confiabilidad de un componente, o de un sistema, en el tiempo t, $R(t)$,* se define por $R(t) = P(T > 1)$, donde T es el tiempo de vida del componente, o del sistema. Si el tiempo transcurrido entre dos caídas sucesivas de un componente o sistema, se distribuye como una variable aleatoria exponencial con parámetro λ, entonces la **función de confiabilidad** se define como:

Función de confiabilidad

$$R(t) = e^{-\lambda t}$$

APLICACIÓN 7.23

Suponga que el tiempo de vida, en horas, de un artefacto electrónico se distribuye exponencialmente con $\lambda = 0.002$. Si ese artefacto tiene una confiabilidad de 95%, determine el número de horas que funcionará antes de descomponerse.

Solución: Como $R(t) = e^{-\lambda t}$ y $\lambda = 0.002$, la función de confiabilidad para ese artefacto es $R(t) = e^{-0.002t}$. En consecuencia, tendremos $0.95 = e^{-0.002t}$, pudiendo usar logaritmos para despejar t de esta ecuación.

$$-0.002t = \ln(0.95)$$

$$t = \frac{\ln(0.95)}{-0.002} = 25.65 \text{ horas}$$

Por tanto, si uno de cada 100 aparatos opera 25.65 horas, aproximadamente 95 de ellos no se descompondrán durante ese periodo. ■

GRUPO DE EJERCICIOS 7.5

Habilidades básicas

1. Use una calculadora para determinar el valor de $e^{-c\lambda}$ para cada una de las parejas siguientes:
a) $\lambda = 2, c = 3$ b) $\lambda = 1.5, c = 2.3$
c) $\lambda = 0.75, c = 4.1$

2. Igual que el anterior.
a) $\lambda = 1.2, c = 2.3$ b) $\lambda = 1.8, c = 3.3$
c) $\lambda = 1.65, c = 10.4$

3. Suponga que X se distribuye exponencialmente con parámetro $\lambda = 6.4$ y determine su media y su varianza.

4. Si X se distribuye exponencialmente con parámetro $\lambda = 0.02$. Calcule su media y su varianza.

5. Asuma que X se distribuye exponencialmente con parámetro $\lambda = 0.12$, para estimar las probabilidades siguientes
a) $P(X > 8.3)$ b) $P(X > 9.3)$ c) $P(8.2 < X < 11.2)$

6. Si X se distribuye exponencialmente con parámetro $\lambda = 0.32$, encuentre las probabilidades:
a) $P(X > 2.3)$ b) $P(X < 3)$
c) $P(2.8 < X < 4.1)$

Más aplicaciones

7. Suponga que el tiempo de vida, en horas, de un artefacto electrónico se distribuye exponencialmente con $\lambda = 0.02$, y que posee una confiabilidad de 95%; determine el número de horas que funcionará sin descomponerse.

8. Como el supuesto anterior, con $\lambda = 0.012$ y una confiabilidad del 89%, ¿cuántas horas funcionará antes de fallar?

9. El tiempo de vida de un artefacto electrónico se distribuye exponencialmente con una media de 500 horas; se sabe que la confiabilidad del aparato es de 0.88. ¿Cuántas horas debe operarse para alcanzar una confiabilidad de 0.95?

10. El tiempo de vida de un artefacto electrónico se distribuye exponencialmente con una media de 450 horas, siendo su confiabilidad de 0.80 horas. ¿Cuánto tiempo ha de operarse para alcanzar una confiabilidad de 0.90?

11. Represente por X el tiempo en minutos entre los sucesivos ingresos a la sala de emergencias de un hospital, y suponga que X se distribuye exponencialmente con parámetro $\lambda = 0.10$. Encuentre:
a) El tiempo esperado entre dos ingresos sucesivos.
b) La desviación estándar entre dos ingresos sucesivos.
c) $P(X \leq 10.5)$.
d) $P(X > 9.6)$.

12. Suponga que el tiempo de funcionamiento, en años, de una batería especial se distribuye exponencialmente con un tiempo de vida esperado de dos años. Calcule la probabilidad de que una batería de este tipo tenga un tiempo de funcionamiento:
a) mayor de tres años;
b) de entre 1 y tres años;
c) menor de 2 años.

13. El tiempo en minutos entre los clientes sucesivos en una ventanilla bancaria se distribuye exponencialmente con parámetro $\lambda = 0.5$. Determine la probabilidad de que al menos uno de los cuatro clientes siguientes deba esperar más de 3 minutos.

14. El tiempo en minutos entre dos ingresos sucesivos a la sala de emergencias de un hospital se distribuye exponencialmente con parámetro $\lambda = 0.4$. ¿Qué probabilidad hay de que al menos dos de los cuatro ingresos subsecuentes deban esperar menos de 2 minutos?

15. Sea X una variable aleatoria exponencial con parámetro $\lambda = 0.125$. Encuentre:
a) el sexagésimo percentil;
b) el tercer cuartil;
c) la mediana;
d) el cuarto decil.

16. Sea X una variable aleatoria exponencial con parámetro $\lambda = 0.08$. Ubique:
a) el cuadragésimo percentil;
b) el primer cuartil;
c) la mediana;
d) el séptimo decil.

17. El tiempo de espera para que un cliente sea atendido en una cafetería se distribuye exponencialmente con una media de 3 minutos. ¿Cuál es la probabilidad de que una persona sea atendida en menos de 2 minutos en al menos tres de los seis días siguientes?

18. Un sistema electrónico contiene un circuito integrado cuyo tiempo de vida en años se distribuye exponencialmente con un valor esperado de siete años; si seis circuitos integrados similares se instalan en sistemas diferentes, ¿cuál es la probabilidad de que al menos dos estén en funcionamiento al término de diez años?

19. Un fabricante de abanicos eléctricos quiere determinar el tiempo de vida de sus productos para garantizarlos sin tener que reemplazar más del 5%. Si los tiempos

de vida de los abanicos se distribuyen exponencialmente con una media de siete años, encuentre el periodo de garantía al mes más cercano.

20. Un fabricante de baterías para automóvil quiere determinar el periodo de garantía para uno de los tipos de batería que produce. Si los tiempos de vida para ese tipo de productos se distribuyen exponencialmente con una media de dos años, calcule el periodo de garantía, aproximado al mes más cercano, si se espera reemplazar no más del 6% del producto.

RESUMEN DEL CAPÍTULO

En este capítulo estudiamos distribuciones de probabilidad uniformes, normales y exponenciales; aprendimos cómo asociar áreas bajo una curva con probabilidades, cómo encontrar probabilidades asociadas con distribuciones uniformes y la distribución normal estándar; también vimos la distribución normal estándar para calcular probabilidades asociadas con cualquier distribución normal; cómo usar distribuciones normales para aproximar distribuciones binomiales, mismas que se pueden usar cuando np y $n(1-p)$ son mayores o iguales que 5; finalmente aprendimos cómo encontrar probabilidades asociadas con distribuciones exponenciales, que son las distribuciones de los tiempos transcurridos entre la sucesión de dos eventos.

REVISIÓN DEL CAPÍTULO

■ TÉRMINOS GENERALES ■

Los términos siguientes utilizados en el capítulo se han mezclado para proporcionarle una mejor práctica de revisión; para cada uno dé una definición con sus propias palabras, después verifique sus respuestas contra las definiciones dadas en el capítulo.

variable aleatoria uniforme

distribución continua

rango percentil

deciles

regla empírica

distribución normal

percentiles

funciones de densidad de probabilidad

variable aleatoria exponencial

función de distribución acumulativa

función de densidad exponencial

cuartiles

gráfica de puntajes z

distribución asimétrica

distribución normal estándar

tabla de la normal estándar

distribución de probabilidad uniforme

parámetro

función de confiabilidad

■ SÍMBOLOS IMPORTANTES ■

Z, la variable aleatoria normal estándar

0^+, un número apenas mayor que 0

1^-, un número apenas menor que 1

P_n, el n-ésimo percentil

Q_n, el n-ésimo cuartil

D_n, el n-ésimo decil

λ, parámetro exponencial

▪ HECHOS Y FÓRMULAS IMPORTANTES ▪

Función de densidad de probabilidad uniforme: $f(x) = \frac{1}{b-a}$ para valores de x entre a y b donde $a < b$. (7.1)

El área bajo una curva normal es uno.

Regla empírica: en cualquier distribución normal, aproximadamente 68% de las medidas caen en $\mu \pm \sigma$, 95% de las medidas caen en $\mu \pm 2\sigma$, y 99% de las medidas cen en $\mu \pm 3\sigma$.

Regla de transformación: $z = \frac{x-\mu}{\sigma}$.

Una distribución normal con $\mu = np$ y $\sigma = \sqrt{np(1-p)}$ se puede usar para aproximar una distribución binomial con parámetros n y p si se cumple $np \geq 5$ y $n(1-p) \geq 5$.

La función de densidad de probabilidad para una variable aleatoria que se distribuye exponencialmente con parámetro λ: $f(x) = \lambda e^{-\lambda x}$ para $\lambda > 0$ y $x > 0$.

Media de una variable aleatoria exponencial con parámetro λ: $\mu = \frac{1}{\lambda}$.

Desviación estándar de una variable aleatoria exponencial con parámetro λ: $\sigma = \frac{1}{\lambda}$.

Función de distribución acumultiva para una variable aleatoria exponencial con parámetro λ: $F(x) = 1 - e^{-\lambda x}$ para $\lambda > 0$ y $x > 0$. (7.2)

Función de confiabilidad para una variable aleatoria exponencial con parámetro λ: $R(t) = e^{-\lambda t}$

EJERCICIOS DE REPASO

1. Suponga que X es una variable aleatoria con valores en el rango comprendido entre 10 y 90. Determine las probabilidades siguientes:
a) $P(20 < X < 50)$
b) $P(10 < X < 70)$
c) $P(X < 40)$
d) $P(X > 60)$
e) $P[(X < 50)$ o $(X > 79)]$

2. Encuentre las probabilidades:
a) $P(-1.2 < Z < 2.2)$
b) $P(Z > 1.24)$
c) $P(Z > -1.78)$
d) $P(Z < -2.14)$
e) $P(Z < 1.58)$
f) $P(Z > 1.58)$
g) $P(Z < -2.14$ o $Z > 2.18)$
h) $P(1.4 < Z < 2.76)$
i) $P(-2.11 < Z < -1.17)$

3. Para cada inciso de los dados a continuación, encuentre un valor de z_0 tal que:
a) $P(-z_0 < Z < z_0) = 0.70$.
b) $P(Z > z_0) = 0.60$.
c) $P(Z > z_0) = 0.40$.
d) $P(Z < z_0) = 0.80$.
e) $P(Z < z_0) = 0.30$.

4. Suponga que X se distribuye normalmente con $\mu = 40$ y $\sigma = 5$. Determine las probabilidades siguientes:
a) $P(X > 47)$
b) $P(X > 32)$
c) $P(X < 36)$
d) $P(X < 51)$
e) $P(36 < X < 48)$
f) $P(42 < X < 61)$

5. X se distribuye normalmente con $\mu = 50$ y $\sigma = 10$. En seguida calcule el valor de x_0 tal que:
a) $P(X > x_0) = 0.70$.
b) $P(X > x_0) = 0.30$.
c) $P(X < x_0) = 0.30$.
d) $P(X < x_0) = 0.60$.

6. Si en un examen las calificaciones de 400 estudiantes se distribuyen normalmente con una media de 100 y una desviación estándar de 10, ¿cuántos estudiantes, aproximadamente, tienen calificaciones entre 90 y 110?

7. Un profesor de estadística anuncia que calificará con A el 10% de los resultados de un examen final. Los resultados del examen final indican que, en puntos, la calificación promedio es 73 y la desviación estándar 6. ¿Qué calificación mínima debe obtener un estudiante para recibir A? Los grados se distribuyen normalmente.

8. La producción diaria de una línea de ensamblaje se distribuye normalmente con una media de 165 unidades y una desviación estándar de 5. Encuentre la probabilidad de que el número de unidades producidas por día:
a) sea menor que 162 unidades;
b) sea mayor que 173 unidades.

c) esté entre 152 y 174 unidades;
d) esté entre 171 y 177 unidades.

9. Refiérase al ejercicio 8. Encuentre el número de unidades x tal que en el 75% de los días el número de unidades producidas sea x.

10. Suponga que las velocidades típicas en las carreteras de Maryland se distribuyen normalmente con una media de 59 millas por hora y una desviación estándar de 4 mph. Si la policía estatal tiene órdenes de multar al 10% más veloz de los conductores, ¿cuál es la mayor velocidad a la que usted puede conducir ahí sin ser multado?

11. Un tipo de baterías tiene un tiempo de funcionamiento que se distribuye exponencialmente con una media de dos años. Encuentre la probabilidad de que una batería funcione:
a) más de tres años;
b) entre uno y tres años.

12. Un estudiante con un CI de 135 afirma que su puntaje en CI está en un 3% más alto de su grupo. ¿Es cierta su afirmación si los puntajes de sus compañeros se distribuyen normalmente con una media de 120 y una desviación estándar de 7?

13. Ciertas píldoras anticonceptivas tienen una efectividad del 90%, si 50 mujeres las usan ¿cuál es la probabilidad aproximada de que entre ellas se produzca un embarazo? Si hombres y mujeres se conciben con la misma probabilidad, ¿qué posibilidad existe en este caso, de que nazca un varón?

14. Se estima que el 85% de todas las plantas caseras están humedas en exceso. En un grupo de 500 plantas, ¿cuál es la probabilidad aproximada de que:
a) 433 estén sobrehumedecidas?;
b) entre 420 y 435, inclusive, tengan agua en exceso?;
c) al menos a 430 las saturen de agua?;
d) a lo más 440 estén excesivamente húmedas?

15. Si se estima que del 80% de todos los estudiantes que presentan el examen final de matemáticas 209 lo aprueban, ¿cuál es la probabilidad aproximada de que pase al menos la mitad de un grupo de 30 que presentan ese examen?

16. Suponga que X es una variable aleatoria de que se distribuye exponencialmente con un valor esperado de 5. Halle:
a) $P(X > 5.5)$ **c)** $P(3.5 < X < 5.5)$
b) $P(X < 4)$ **d)** $P(X = 5)$

Aplicaciones de computación

1. Los puntajes del SAT de matemáticas tienen una media de 500 y una desviación estándar de 100. Use un programa computacional para encontrar la probabilidad de obtener una calificación menor de 550. (*Sugerencia:* Use órdenes parecidas a:
CDF 500:
NORMAL MU = 500 SIGMA = 100.)

2. Genere una muestra aleatoria de 200 calificaciones de una población normal que tiene una media de 500 y una desviación estándar de 100. Ordene las medidas y exhíbalas, luego determine la proporción de medidas menores de 550. Compare este resultado con la respuesta obtenida en la aplicación de computación 1. (*Sugerencia:* Use órdenes similares a:
RANDON 200 C1;
NORMAL MU = 500 SIGMA = 100
SPORT C1 C2
PRINT C2.)

3. Repita la aplicación de computación 2 para 500 calificaciones en lugar de 200. (*Sugerencia:* Si se usa MINITAB, la orden CODE puede emplearse para contar las medidas menores de 550.)

4. Determine si la muestra siguiente de datos puede provenir de una población normal. ¿Cuáles son la media y la desviación estándar que pueden esperarse en la población?

68	80	68	89	82	61	67	80	76	74	63
74	90	74	78	73	64	71	93	60	74	87
74	83	73	88	90	80	87	81	68	100	79
70	91	83	68	79	74	79	65	84	71	74
62	69	95	76	97	62	71	71	80	85	67
78	57	66	76	75	84	84	62	71	82	82
67	66	86	75	60	60	85	76	80	95	79
79	74	71	79	91	74	62	75	80	75	77
70	71	72	74	73	61	64	67	91	70	79
77										

5. Si X es una variable aleatoria que se distribuye normalmente con $\mu = 30$ y $\sigma = 5$, determine un valor x_0 tal que $P(-x_0 < X < x_0) = 0.95$

■ EXAMEN DE CONOCIMIENTOS DEL CAPÍTULO ■

1. Suponga que X se distribuye uniformemente entre los valores 7 y 13. Encuentre las probabilidades:
a) $P(X > 8)$ **c)** $P(8 < X < 12)$
b) $P(X < 12)$

2. Igual que el anterior
a) $P(Z < 0.51)$ **d)** $P(Z > -1.14)$
b) $P(-1.2 < Z < 2.3)$ **e)** $P(Z < -0.24)$
c) $P(-2.4 < Z < -0.78)$

3. Para cada uno de los casos siguientes, encuentre el valor de z_0 tal que:
a) $P(Z > z_0) = 0.30$
b) $P(Z > z_0) = 0.75$
c) $P(Z < z_0) = 0.40$

4. Suponga que X se distribuye normalmente con $\mu = 20$ y $\sigma = 3$. Localice:
a) $P(X > 27)$
b) $P(X > 16)$
c) $P(13 < X < 15)$
d) $P(25 < X < 28)$

5. Para la pregunta 4, encuentre el valor z_0 que satisface a:
a) $P(X > z_0) = 0.90$ **c)** $P(X < z_0) = 0.78$
b) $P(X > z_0) = 0.40$

6. Si X se distribuye normalmente con $\mu = 40$ y $\sigma = 6$, encuentre:
a) Q_3 **b)** P_{40}

7. Los tiempos de vida útil en los focos de cierta marca se distribuyen normalmente con $\mu = 250$ horas y $\sigma = 50$. Si el fabricante garantiza que sus focos duran al menos 125 horas, ¿qué porcentaje de esos productos espera reemplazar bajo la garantía?

8. Un fabricante de partes de computadora afirma que solo el 1% de ellas resultan defectuosas; si suponemos cierta su afirmación, ¿cuál es la probabilidad de que no más de 25 partes salgan defectuosas en la siguiente producción de 2000 piezas?

9. Si X se distribuye exponencialmente con $\lambda = 3$, determine:
a) $P(X < 4)$ **c)** $P(2 < X < 5)$
b) $P(X > 2)$ **d)** la mediana

UNIDAD TRES

Estadística inferencial

8 Teoría del muestreo

DESCRIPCIÓN

8.1 *Tipos de errores y muestras aleatorias*

8.2 *Distribuciones muestrales*

8.3 *Muestreo de poblaciones normales*

8.4 *Muestro de poblaciones no normales*

8.5 *Distribución muestral de proporciones de muestras*

OBJETIVOS DEL CAPÍTULO

En este capítulo estudiaremos:

- *Sesgo.*
- *Error muestral.*
- *Cómo usar la tabla de números aleatorios.*
- *Varios métodos para obtener una muestra aleatoria.*
- *Distribuciones muestrales.*
- *La distribución muestral de la media.*
- *Significado de la media, de la varianza y de la desviación estándar de la distribución muestral de la media.*
- *Cómo encontrar probabilidades asociadas con la distribución muestral de la media cuando la población es normal.*
- *Distribuciones t.*
- *El teorema del límite central.*
- *Aplicaciones del teorema del límite central.*
- *Distribución muestral de la suma muestral.*
- *Media, varianza y desviación estándar de la distribución muestral de la suma muestral.*
- *Distribución muestral de la proporción muestral.*
- *Media, varianza y desviación estándar de la distribución muestral de la proporción muestral.*
- *Aplicaciones que usan la distribución muestral de la proporción muestral.*

MOTIVADOR 8

Considere el experimento de lanzar diez veces un dado para determinar la suma de los diez resultados. Por ejemplo, suponga que diez lanzamientos dan la secuencia: 2, 3, 3, 6, 5, 1, 4, 4, 1 y 5; la suma es 34. Más aún, suponga que el experimento se repite 1000 veces y que se construye un histograma para las 1000 sumas. ¿Esperaría usted que el histograma tenga una forma particular? ¿Cuál sería su expectativa de la media de las 1000 sumas? Las respuestas a estas dos preguntas son bastante sorprendentes. Usemos MINITAB para simular el experimento 1000 veces, dibujar un histograma con los resultados y calcular algunos estadísticos descriptivos. La pantalla 8.1 deja ver las órdenes usadas y las respuestas correspondientes.

Pantalla 8.1

```
MTB > RANDOM 1000 C1-C10;
SUBC > INTEGER 1 6.
MTB > RSUM C1 - C10 C11
MTB > HISTOGRAM C11

HISTOGRAM OF C11 N = 1000
EACH * REPRESENTS 10 OBS.

MIDPOINT
    16       1  *
    20       2  *
    24      28  ***
    28     124  ************
    32     221  **********************
    36     284  ****************************
    40     208  ********************
    44     102  **********
    48      28  ***
    52       2  *

MTB  DESCRIBE C11

          N      MEAN   MEDIAN  TRMEAN   STDEV  SEDMEAN
C11     1000    35.232  35.000  35.230   5.397   0.171

          MIN     MAX      Q1      Q3
C11     16.000  50.000  31.000  39.000
```

Note que el histograma tiene la forma de campana y se aproxima a una distribución normal; la media es muy cercana a 35, [el producto de 10 y la media de los seis resultados posibles cuando se lanza el dado]. Esto es, $35.232 \approx (10)[(1 + 2 + 3 + 4 + 5 + 6)/6] = (10)(3.5) = 35$. Si fuera posible repetir el experimento indefinidamente y anotar los resultados, la media de la distribución de las sumas sería exactamente igual a 35. Estos dos fenómenos no ocurren por accidente: están relacionados con un resultado importante, llamado teorema del límite central, que se estudiará en la sección 8.4. Aquí examinaremos este resultado y otros relacionados con él.

Panorama del capítulo

Uno de los propósitos de la estadística inferencial es estimar las características o parámetros poblacionales desconocidos, examinando la información obtenida de una subcolección, muestra, de una población. El punto de interés es la muestra, si es que se va a usar una, la cual debe ser representativa de la población objeto del estudio; seguiremos ciertos procedimientos de selección para asegurarnos de que las muestras reflejen fielmente a la población de la que proceden, ya que solo se pueden hacer afirmaciones probabilísticas sobre una población cuando se usan muestras representativas de la misma. Por ejemplo, suponga que estamos interesados en las actitudes hacia el estudio por parte de los estudiantes universitarios; en este caso, la población relevante en la colección de respuestas de los estudiantes universitarios sobre el estudio, una muestra es una subcolección de dicha población, y la usaremos para normar las actitudes de los estudiantes universitarios; si nos basamos únicamene en las respuestas de miembros de fraternidades sociales, obtendremos, sobre todo, respuestas sesgadas o resultados atípicos, o si usamos únicamente las respuestas de los miembros de fraternidades honoríficas, probablemente los resultados de desviarían a favor del estudio. Al usar una muestra sesgada, conseguimos una visión distorsionada de las actitudes de los estudiantes en conjunto.

SECCIÓN 8.1 Tipos de errores y muestras aleatorias

Cuando nos interesa estudiar las características de poblaciones grandes, utilizamos muestras por muchas razones; una enumeración completa de la población, llamada **censo**, puede ser económicamente imposible; o puede no haber tiempo suficiente para examinar a la población completa; en algunas situaciones, el censo puede ser imposible; por ejemplo, un censo de la población marina que vive en el Océano Atlántico es imposible.

EJEMPLO 8.1

A continuación veremos los usos del muestreo en diversos campos:

1. *Política.* Las muestras de las opiniones de los votantes se usan para que los candidatos midan la opinión pública y el apoyo en las elecciones.
2. *Educación.* Las muestras de las calificaciones de los exámenes de estudiantes se usan para determinar la eficiencia de una técnica o programa de enseñanza.
3. *Industria.* Muestras de los productos de una línea de ensamblaje sirve al propósito de controlar la calidad.
4. *Medicina.* Muestras de medidas de azúcar en la sangre de pacientes diabéticos prueban la eficacia de una técnica o de un fármaco nuevo.
5. *Agricultura.* Las muestras del maíz cosechado en una parcela proyectan en la producción los efectos de un fertilizante nuevo.
6. *Gobierno.* Una muestra de opiniones de los votantes se usaría para determinar los criterios del público sobre cuestiones relacionadas con el bienestar y la seguridad nacionales.

Cuando se usan valores muestrales, o estadísticos, para estimar valores poblacionales, o parámetros, pueden ocurrir dos tipos generales de errores: el error muestral y el error no muestral. El **error muestral** se refiere a la variación natural existente entre muestras tomadas de la misma población, cuando una muestra no es una copia exacta de la población; aún si hemos tenido gran cuidado para asegurar que dos muestras del mismo tamaño sean representativas de una cierta población, no esperaríamos que las dos sean idénticas en todos sus detalles. El error muestral es un concepto importante que nos ayudará a entender mejor la naturaleza de la estadística inferencial; más adelante, en este capítulo estudiaremos con profundidad el concepto de error muestral.

Los errores que surgen al tomar las muestras no pueden clasificarse como errores muestrales y se denominan **errores no muestrales**. El sesgo de las muestras es un tipo de error no muestral. El **sesgo muestral** se refiere a una tendencia sistemática inherente a un método de muestreo que da estimaciones de un parámetro que son, en promedio, menores, sesgo negativo, o mayores, sesgo positivo, que el parámetro real. Véanse los ejemplos 8.2 y 8.3 para errores que surgen de colecciones de datos, que caen en esta categoría. Los errores que resultan de la acumulación de datos o de su procesamiento se clasifican también como errores no muestrales, ejemplo 8.4.

EJEMPLO 8.2

Si queremos obtener información relativa a las actitudes hacia el aborto y obtenemos una muestra que consta preponderantemente de hombres, podríamos encontrar un sesgo muestral.

EJEMPLO 8.3

Un ejemplo clásico de sesgo muestral es el estudio de las actitudes de varios millones de personas, realizdo por el *Literary Digest*, un periódico popular en ese entonces, para predecir al ganador de la presidencia en 1936, cuando el republicando Alfred Landon competía contra el demócrata Franklin Roosevelt; los nombres de las personas que se incluyeron en la encuesta los obtuvo el *Digest* del directorio telefónico y de su lista de suscriptores, la mayoría de los entrevistados mostraron su prefrencia por Landon, y el periódico predijo que este candidato ganaría por un gran margen. Como sabemos, Landon perdió; mucha gente que prefería votar por Roosevelt no tenía teléfono y no se suscribía a periódicos, por eso la muestra estaba fuertemente sesgada a favor de Landon.

EJEMPLO 8.4

Al recabar datos pueden generarse errores no muestrales cuando los instrumentos usados para realizar las mediciones están fuera de ajuste o mal calibrados; pueden ocurrir errores de procesamiento si los datos están mal colocados, si se pierden al registrarlos o si las respuestas proporcionadas por las personas durante el estudio no son verdaderas, este último caso puede darse con preguntas relativas a la edad, en las que mucha gente miente por vanidad.

El sesgo muestral puede suprimirse, o minimizarse, usando la aleatorización. La **aleatorización** se refiere a cualquier proceso de selección de una muestra de la población en el que la selección es imparcial o no está sesgada; una muestra elegida con procedimientos aleatorios se llama **muestra aleatoria**; los tipos más comunes de técnicas de muestreo aleatorio son el muestreo aleatorio simple, el muestreo estratificado, el muestreo por conglomerados y el muestreo sistemático. Explicaremos ahora cada uno de ellos.

Muestras aleatorias

Si una muestra aleatoria se elige de forma tal que todas las muestras del mismo tamaño tengan igual probabilidad de ser elegidas, la llamamos **muestra aleatoria simple.**

EJEMPLO 8.5

Suponga que nos interesa elegir una muestra aleatoria de 5 estudiantes en un grupo de estadística de 20 alumnos; el coeficiente binomial $\binom{20}{5}$ da el número total de formas de elegir una muestra no ordenada de 5 estudiantes en un grupo de 20.

$$\binom{20}{5} = \frac{20!}{5!15!}$$