***4.35** Si Y es una variable aleatoria continua tal que $E[(Y-a)^2] < \infty$ para toda a, demuestre que $E[(Y-a)^2]$ se minimiza cuando $a = E(Y)$. [*Sugerencia:* $E[(Y-a)^2] = E(\{[Y-E(Y)] + [E(Y)-a]\}^2)$.]

***4.36** ¿El resultado obtenido en el Ejercicio 4.35 también es válido para variables aleatorias discretas? ¿Por qué?

***4.37** Si Y es una variable aleatoria continua con función de densidad $f(y)$ que es simétrica alrededor de 0 (es decir, $f(y) = f(-y)$ para toda y) y $E(Y)$ existe, demuestre que $E(Y) = 0$. [*Sugerencia:* $E(Y) = \int_{-\infty}^{0} yf(y)\,dy + \int_{0}^{\infty} yf(y)\,dy$. Haga el cambio de variable $w = -y$ en la primera integral.]

4.4 La distribución de probabilidad uniforme

Suponga que un autobús llega siempre a una parada particular entre las 8:00 y las 8:10 a.m. y que la probabilidad de que llegue en cualquier subintervalo dado es proporcional sólo a la duración del subintervalo. Esto es, es igual de probable que llegue entre las 8:00 y 8:02 a que llegue entre las 8:06 y las 8:08. Denote con Y el tiempo que una persona deba esperar para que llegue el autobús si llegó a la parada exactamente a las 8:00. Si con cuidado medimos en minutos cuánto tiempo después de las 8:00 llegó el autobús en varias mañanas, podríamos desarrollar un histograma de frecuencia relativa para los datos.

A partir de la descripción que acabamos de dar, debe ser evidente que la frecuencia relativa con la cual observamos un valor de Y entre 0 y 2 sería aproximadamente la misma que la frecuencia relativa con la cual observamos un valor de Y entre 6 y 8. Un modelo razonable para la función de densidad de Y se muestra en la Figura 4.9. Como las áreas bajo las curvas representan probabilidades para variables aleatorias continuas y $A_1 = A_2$ (por inspección), se deduce que $P(0 \le Y \le 2) = P(6 \le Y \le 8)$, como se desea.

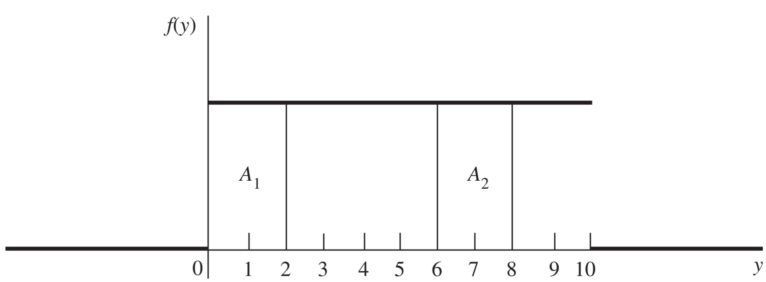
La variable aleatoria Y que acabamos de examinar es un ejemplo de una variable aleatoria que tiene una distribución uniforme. La forma general para la función de densidad de una variable aleatoria con una distribución uniforme es como sigue.

DEFINICIÓN **4.6**

Si $\theta_1 < \theta_2$, se dice que una variable aleatoria Y tiene *distribución de probabilidad uniforme* en el intervalo (θ_1, θ_2) si y sólo si la función de densidad de Y es

$$f(y) = \begin{cases} \dfrac{1}{\theta_2 - \theta_1}, & \theta_1 \le y \le \theta_2, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

FIGURA **4.9** Función de densidad para Y



En el problema del autobús podemos tomar $\theta_1 = 0$ y $\theta_2 = 10$ porque estamos interesados sólo en un intervalo particular de diez minutos. La función de densidad que se estudia en el Ejemplo 4.2 es una distribución uniforme con $\theta_1 = 0$ y $\theta_2 = 1$. Las gráficas de la función de distribución y función de densidad para la variable aleatoria del Ejemplo 4.2 se dan en las Figuras 4.4 y 4.5, respectivamente.

DEFINICIÓN **4.7**

Las constantes que determinan la forma específica de una función de densidad se denominan *parámetros* de la función de densidad.

Las cantidades θ_1 y θ_2 son parámetros de la función de densidad uniforme y son valores numéricos claramente significativos asociados con la función de densidad teórica. Tanto la amplitud como la probabilidad de que Y caiga en cualquier intervalo determinado dependen de los valores de θ_1 y θ_2.

Algunas variables aleatorias continuas en física, administración y ciencias biológicas tienen distribuciones de probabilidad aproximadamente uniformes. Por ejemplo, suponga que el número de eventos, como las llamadas que entran en un conmutador, que se presentan en el intervalo $(0, t)$ tienen una distribución de Poisson. Si se sabe que exactamente uno de estos eventos ha ocurrido en el intervalo $(0, t)$, entonces el tiempo real del suceso está distribuido de manera uniforme en este intervalo.

EJEMPLO **4.7** La llegada de clientes a una caja en un establecimiento sigue una distribución de Poisson. Se sabe que durante un periodo determinado de 30 minutos, un cliente llega a la caja. Encuentre la probabilidad de que el cliente llegue durante los últimos 5 minutos del periodo de 30 minutos.

Solución Como acabamos de citar, el tiempo real de llegada sigue una distribución uniforme en el intervalo de (0, 30). Si Y denota el tiempo de llegada, entonces

$$P(25 \le Y \le 30) = \int_{25}^{30} \frac{1}{30}\, dy = \frac{30 - 25}{30} = \frac{5}{30} = \frac{1}{6}.$$

La probabilidad de que la llegada ocurra en cualquier otro intervalo de 5 minutos también es $1/6$. ■

Como veremos, la distribución uniforme es muy importante por razones teóricas. Los estudios de simulación son técnicas valiosas para validar modelos en estadística. Si deseamos un conjunto de observaciones de una variable aleatoria Y con función de distribución $F(y)$, a menudo podemos obtener los resultados deseados si transformamos un conjunto de observaciones en una variable aleatoria uniforme. Por esta razón, casi todos los sistemas de cómputo contienen un generador de números aleatorios que produce valores observados para una variable aleatoria que tiene una distribución uniforme continua.

TEOREMA **4.6** Si $\theta_1 < \theta_2$ y Y es una variable aleatoria uniformemente distribuida en el intervalo (θ_1, θ_2), entonces

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \quad \text{y} \quad \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}.$$

Prueba Por la Definición 4.5,

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y f(y)\, dy \\ &= \int_{\theta_1}^{\theta_2} y \left(\frac{1}{\theta_2 - \theta_1}\right) dy \\ &= \left(\frac{1}{\theta_2 - \theta_1}\right) \frac{y^2}{2}\bigg]_{\theta_1}^{\theta_2} = \frac{\theta_2^2 - \theta_1^2}{2(\theta_2 - \theta_1)} \\ &= \frac{\theta_2 + \theta_1}{2}. \end{aligned}$$

Observe que la media de una variable aleatoria uniforme es simplemente el valor que está a la mitad entre los valores de los dos parámetros, θ_1 y θ_2. La obtención de la varianza se deja como ejercicio.

Ejercicios

4.38 Suponga que Y tiene una distribución uniforme en el intervalo (0, 1).

a Encuentre $F(y)$.

b Demuestre que $P(a \leq Y \leq a + b)$, para $a \geq 0$, $b \geq 0$, y $a + b \leq 1$ depende sólo del valor de b.

4.39 Si una paracaidista aterriza en un punto aleatorio en una recta entre los marcadores A y B, encuentre la probabilidad de que ella esté más cerca de A que de B. Encuentre la probabilidad de que su distancia hasta A sea más de tres veces su distancia a B.

4.40 Suponga que tres paracaidistas operan de manera independiente como se describe en el Ejercicio 4.39. ¿Cuál es la probabilidad de que exactamente uno de los tres aterrice en el punto medio entre A y B?

4.41 Una variable aleatoria Y tiene una distribución uniforme en el intervalo (θ_1, θ_2). Obtenga la varianza de Y.

4.42 La *mediana* de la distribución de una variable aleatoria continua Y es el valor $\phi_{.5}$ de manera que $P(Y \leq \phi_{.5}) = 0.5$. ¿Cuál es la mediana de la distribución uniforme en el intervalo (θ_1, θ_2)?

4.43 Un círculo de radio r tiene área $A = \pi r^2$. Si un círculo aleatorio tiene un radio que está uniformemente distribuido en el intervalo (0, 1), ¿cuáles son la media y la varianza del área del círculo?

4.44 El cambio en profundidad de un río de un día al siguiente, medida (en pies) en un lugar específico, es una variable aleatoria Y con la siguiente función de densidad:

$$f(y) = \begin{cases} k, & -2 \leq y \leq 2 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

a Determine el valor de k.

b Obtenga la función de distribución para Y.

4.45 Al estudiar bajas cotizaciones para contratos de embarques, una empresa fabricante de microcomputadoras encuentra que los contratos interestatales tienen bajas cotizaciones que están uniformemente distribuidas entre 20 y 25, en unidades de miles de dólares. Encuentre la probabilidad de que la baja cotización en el siguiente contrato interestatal

a esté por debajo de \$22,000.

b sea de más de \$24,000.

4.46 Consulte el Ejercicio 4.45. Encuentre el valor esperado de bajas cotizaciones en contratos del tipo descrito ahí.

4.47 La falla de una tarjeta de circuito que utiliza un sistema de cómputo interrumpe el trabajo hasta que se instala una nueva. El tiempo de entrega, Y, está uniformemente distribuido en el intervalo de uno a cinco días. El costo de la falla de una tarjeta y la interrupción incluye el costo fijo c_0 de una nueva tarjeta y un costo que aumenta proporcionalmente con Y^2. Si C es el costo en que se incurre, $C = c_0 + c_1 Y^2$.

a Encuentre la probabilidad de que el tiempo de entrega exceda de dos días.

b En términos de c_0 y c_1, encuentre el costo esperado asociado con una sola tarjeta de circuito que falle.

4.48 Si un punto se localiza *al azar* en un intervalo (a, b) y si Y denota la ubicación del punto, entonces se supone que Y tiene una distribución uniforme en (a, b). Una experta en eficiencia de la planta selecciona al azar un lugar, a lo largo de una línea de ensamble de 500 pies, desde el cual observa hábitos de los trabajadores de la línea. ¿Cuál es la probabilidad de que el punto que ella seleccione se encuentre

a a no más de 25 pies del final de la línea?

b a no más de 25 pies del principio de la línea?

c más cerca del principio de la línea que al final de la línea?

4.49 Una llamada telefónica llega a un conmutador al azar en un intervalo de no más de un minuto. El conmutador estuvo totalmente ocupado durante 15 segundos en este periodo de un minuto. ¿Cuál es la probabilidad de que la llamada llegara cuando el conmutador no hubiera estado totalmente ocupado?

4.50 Empezando a las 12:00 de la noche, un centro de computadoras funciona durante una hora y deja de operar dos horas en un ciclo regular. Una persona que desconoce este horario marca al centro en una hora al azar entre las 12:00 de la noche y las 5:00 a.m. ¿Cuál es la probabilidad de que el centro esté funcionando cuando entre la llamada de la persona?

4.51 El tiempo de ciclo para camiones que transportan concreto al lugar de construcción de una carretera está uniformemente distribuido en el intervalo de 50 a 70 minutos. ¿Cuál es la probabilidad de que el tiempo de ciclo exceda de 65 minutos si se sabe que el tiempo de ciclo excede de 55 minutos?

4.52 Consulte el Ejercicio 4.51. Encuentre la media y la varianza de los tiempos de ciclo para los camiones.

4.53 El número de tarjetas de circuito defectuosas que salen de una máquina soldadora sigue una distribución de Poisson. Durante un día específico de ocho horas, se encontró una tarjeta defectuosa.

a Encuentre la probabilidad de que haya sido producida durante la primera hora de operación durante ese día.

b Encuentre la probabilidad de que haya sido producida durante la última hora de operación durante ese día.

c Dado que no se produjeron tarjetas defectuosas durante las primeras cuatro horas de operación, encuentre la probabilidad de que la tarjeta defectuosa se fabricara durante la quinta hora.

4.54 Al usar el método de triangulación para determinar el alcance de una sonda acústica, el equipo de prueba debe medir con precisión el tiempo que tarda en llegar el frente de onda esférica a un sensor de recepción.

De acuerdo con Perruzzi y Hilliard (1984), los errores de medición se pueden modelar como si tuvieran una distribución uniforme de –0.05 a $+0.05$ μs (microsegundos).

a ¿Cuál es la probabilidad de que una medición de tiempo de llegada sea precisa con tolerancia de 0.01 μs?

b Encuentre la media y varianza de los errores de medición.

4.55 Consulte el Ejercicio 4.54. Suponga que los errores de medición están uniformemente distribuidos entre –0.02 a $+0.05$ μs.

a ¿Cuál es la probabilidad de que una medición particular de tiempo de llegada sea precisa con tolerancia de no más de 0.01 μs?

b Encuentre la media y varianza de los errores de medición.

4.56 Consulte el Ejemplo 4.7. Encuentre la probabilidad condicional de que un cliente llegue durante los últimos 5 minutos del periodo de 30 minutos, si se sabe que ninguno llega durante los primeros 10 minutos del periodo.

4.57 De acuerdo con Zimmels (1983), los tamaños de partículas empleadas en experimentos de sedimentación a menudo tienen una distribución uniforme. En sedimentación que comprenda mezclas de partículas de varios tamaños, las más grandes impiden los movimientos de las más pequeñas. Entonces, es importante estudiar la media y la varianza de los tamaños de partículas. Suponga que las partículas esféricas tienen diámetros que están uniformemente distribuidos entre .01 y .05 centímetros. Encuentre la media y la varianza de los *volúmenes* de estas partículas. (Recuerde que el volumen de una esfera es $(4/3)\pi r^3$.)

4.5 La distribución de probabilidad normal

La distribución de probabilidad continua que más se utiliza es la distribución normal, con la conocida forma de campana que estudiamos en relación con la regla empírica. Los ejemplos y ejercicios de esta sección ilustran algunas de las numerosas variables aleatorias que tienen distribuciones que se calculan en forma muy cercana por medio de una distribución de probabilidad normal. En el Capítulo 7 presentaremos un argumento que explica, al menos parcialmente, el suceso común de distribuciones normales de datos en la naturaleza. La función de densidad normal es como sigue:

DEFINICIÓN **4.8**

Se dice que una variable Y tiene una *distribución normal de probabilidad* si y sólo si, para $\sigma > 0$ y $-\infty < \mu < \infty$, la función de densidad de Y es

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}, \qquad -\infty < y < \infty.$$

Observe que la función de densidad normal contiene dos parámetros, μ y σ.

TEOREMA **4.7**

Si Y es una variable aleatoria normalmente distribuida con parámetros μ y σ, entonces

$$E(Y) = \mu \quad \text{y} \quad V(Y) = \sigma^2.$$

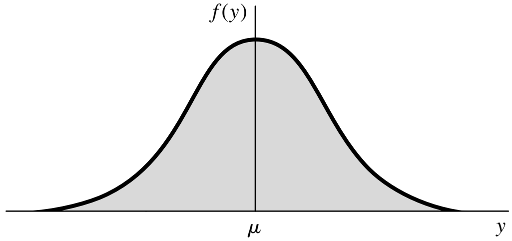


FIGURA **4.10** La función de densidad de probabilidad normal

La demostración de este teorema se difiere a la Sección 4.9, donde obtendremos la función generadora de momento de una variable aleatoria normalmente distribuida. Los resultados contenidos en el Teorema 4.7 implican que el parámetro μ localiza el centro de la distribución y que σ mide su dispersión. Una gráfica de una función de densidad normal se muestra en la Figura 4.10.

Las áreas bajo la función de densidad normal correspondientes a $P(a \leq Y \leq b)$ requieren la evaluación de la integral

$$\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}\, dy.$$

Desafortunadamente, no existe una expresión de forma cerrada para esta integral; en consecuencia, su evaluación requiere el uso de técnicas de integración numérica. Las probabilidades y cuantiles para variables aleatorias con distribuciones normales se encuentran fácilmente usando *R* y *S*–Plus. Si Y tiene una distribución normal con media μ y desviación estándar σ, el comando `pnorm(`y_0, μ, σ`)` de *R* (o *S*–Plus) genera $P(Y \leq y_0)$ mientras que `qnorm(`p, μ, σ`)` da el p–ésimo cuantil, el valor de ϕ_p tal que $P(Y \leq \phi_p) = p$. Aun cuando hay un número infinito de distribuciones normales (μ puede tomar cualquier valor finito, en tanto que σ puede tomar cualquier valor finito positivo), sólo necesitamos una tabla —la Tabla 4, Apéndice 3— para calcular áreas bajo densidades normales. Las probabilidades y cuantiles asociados con variables aleatorias normalmente distribuidas también se pueden hallar usando la aplicación breve (applet) *Normal Tail Areas and Quantiles* accesibles en www.thomsonedu.com/statistics/wackerly. El único beneficio real obtenido al usar software para obtener probabilidades y cuantiles asociados con variables aleatorias normalmente distribuidas, es que el software da respuestas que son correctas hasta un gran número de lugares decimales.

La función de densidad normal es simétrica alrededor del valor μ, de modo que las áreas tienen que ser tabuladas en sólo un lado de la media. Las áreas tabuladas están a la derecha de los puntos z, donde z es la distancia desde la media, medida en desviaciones estándar. Esta área está sombreada en la Figura 4.11.

EJEMPLO **4.8** Denote con Z una variable aleatoria normal con media 0 y desviación estándar 1.

a Encuentre $P(Z > 2)$.
b Encuentre $P(-2 \leq Z \leq 2)$.
c Encuentre $P(0 \leq Z \leq 1.73)$.

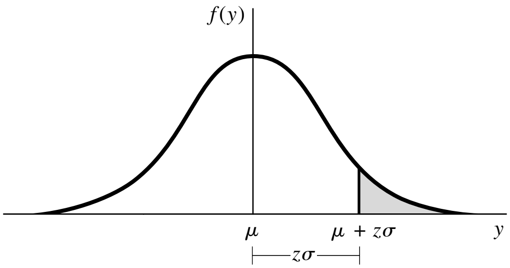


FIGURA **4.11** Área tabulada para la función de densidad normal

Solución

a Como $\mu = 0$ y $\sigma = 1$, el valor 2 está en realidad a $z = 2$ desviaciones estándar arriba de la media. Avance hacia abajo en la primera columna (z) de la Tabla 4, Apéndice 3, y lea el área opuesta a $z = 2.0$. Esta área, denotada por el símbolo $A(z)$, es $A(2.0) = .0228$. Entonces, $P(Z > 2) = .0228$.

b Consulte la Figura 4.12, donde hemos sombreado el área de interés. En el inciso a determinamos que $A_1 = A(2.0) = 0.228$. Como la función de densidad es simétrica alrededor de la media $\mu = 0$, se deduce que $A_2 = A_1 = .0228$ y por tanto que

$$P(-2 \leq Z \leq 2) = 1 - A_1 - A_2 = 1 - 2(.0228) = .9544.$$

c Como $P(Z > 0) = A(0) = .5$, obtenemos que $P(0 \leq Z \leq 1.73) = .5 - A(1.73)$, donde $A(1.73)$ se obtiene al bajar por la columna z de la Tabla 4, Apéndice 3, a la entrada 1.7 y luego en sentido horizontal por la parte superior de la tabla a la columna marcada .03 para leer $A(1.73) = .0418$. De esta manera,

$$P(0 \leq Z \leq 1.73) = .5 - .0418 = .4582.$$

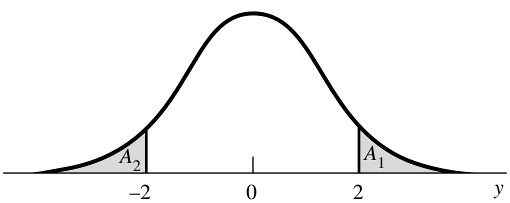


FIGURA **4.12** Área deseada para el Ejemplo 4.8(b)

■

EJEMPLO **4.9** Las calificaciones para un examen de admisión a una universidad están normalmente distribuidas con media de 75 y desviación estándar 10. ¿Qué fracción de las calificaciones se encuentra entre 80 y 90?

Solución Recuerde que z es la distancia desde la media de una distribución normal expresada en unidades de desviación estándar. Entonces,

$$z = \frac{y - \mu}{\sigma}.$$

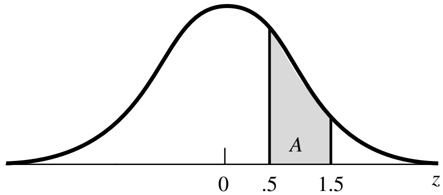


FIGURA **4.13**
Área requerida para el Ejemplo 4.9

Entonces la fracción deseada de la población está dada por el área entre

$$z_1 = \frac{80 - 75}{10} = .5 \quad \text{y} \quad z_2 = \frac{90 - 75}{10} = 1.5.$$

Esta área está sombreada en la Figura 4.13.

Usted puede ver en la Figura 4.13 que $A = A(.5) - A(1.5) = .3085 - .0668 = .2417$. ■

Siempre podemos transformar una variable aleatoria normal Y en una variable aleatoria *normal estándar* Z si usamos la relación

$$Z = \frac{Y - \mu}{\sigma}.$$

La Tabla 4, Apéndice 3 se puede usar para calcular probabilidades, como se muestra aquí. Z localiza un punto medido desde la media de una variable aleatoria normal, con la distancia *expresada en unidades de la desviación estándar* de la variable aleatoria normal original. Entonces, el valor medio de Z debe ser 0 y su desviación estándar debe ser igual a 1. La prueba de que la *variable aleatoria normal estándar,* Z, está normalmente distribuida con media 0 y desviación estándar 1 se proporciona en el Capítulo 6.

En la aplicación breve *Normal Probabilities*, accesible en www.thomsonedu.com/statistics/wackerly, se ilustra la correspondencia entre probabilidades normales en las escalas originales y transformadas (z). Para contestar la pregunta planteada en el Ejemplo 4.9, localice el intervalo de interés, (80, 90), en el eje horizontal inferior marcado como Y. Las calificaciones z correspondientes se dan en el eje horizontal superior y es evidente que el área sombreada da $P(80 < Y < 90) = P(0.5 < Z < 1.5) = 0.2417$ (vea la Figura 4.14). Algunos de los ejercicios del final de esta sección sugieren que el estudiante use esta aplicación para reforzar los cálculos de probabilidades asociadas con variables aleatorias normalmente distribuidas.

Ejercicios

4.58 Use la Tabla 4, Apéndice 3 para hallar las siguientes probabilidades para una variable Z aleatoria normal estándar:

a $P(0 \le Z \le 1.2)$.
b $P(-.9 \le Z \le 0)$.
c $P(.3 \le Z \le 1.56)$.

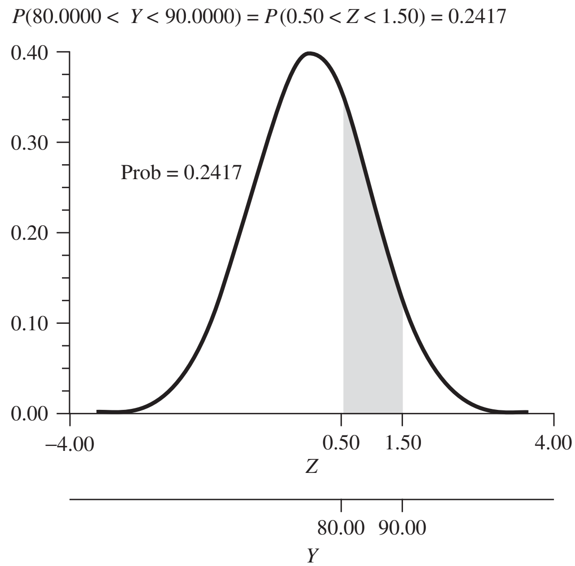


FIGURA **4.14** Área requerida para el Ejemplo 4.9, que usa escalas original y transformada (z).

d $P(-.2 \leq Z \leq .2)$.

e $P(-1.56 \leq Z \leq -.2)$.

f **Ejercicio Applet** Use la aplicación breve *Normal Probabilities* para obtener $P(0 \leq Z \leq 1.2)$. ¿Por qué son idénticos los valores dados en los dos ejes horizontales?

4.59 Si Z es una variable aleatoria normal estándar, encuentre el valor z_0 tal que

a $P(Z > z_0) = .5$.

b $P(Z < z_0) = .8643$.

c $P(-z_0 < Z < z_0) = .90$.

d $P(-z_0 < Z < z_0) = .99$.

4.60 Una variable aleatoria normalmente distribuida tiene función de densidad

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}, \qquad -\infty < y < \infty.$$

Usando las propiedades fundamentales asociadas con cualquier función de densidad, demuestre que el parámetro σ debe ser tal que $\sigma > 0$.

4.61 ¿Cuál es la mediana de una variable aleatoria normalmente distribuida con media μ y desviación estándar σ?

4.62 Si Z es una variable aleatoria normal estándar, ¿cuál es

a $P(Z^2 < 1)$?

b $P(Z^2 < 3.84146)$?

4.63 Una compañía que manufactura y embotella jugo de manzana usa una máquina que automáticamente llena botellas de 16 onzas. Hay alguna variación, no obstante, en las cantidades de líquido que se ponen en las botellas que se llenan. Se ha observado que la cantidad de líquido está normalmente distribuida en forma aproximada con media de 16 onzas y desviación estándar de 1 onza.

a Use la Tabla 4, Apéndice 3 para determinar la proporción de botellas que tendrán más de 17 onzas.

b **Ejercicio Applet** Use la aplicación breve *Normal Probabilities* para obtener la respuesta al inciso a.

4.64 Se observó que la cantidad semanal de dinero gastado por una compañía durante largo tiempo en mantenimiento y reparaciones, está normalmente distribuida en forma aproximada con media de \$400 y desviación estándar de \$20. Si están presupuestados \$450 para la próxima semana, ¿cuál es la probabilidad de que los costos reales rebasen la cantidad presupuestada?

a Conteste la pregunta usando la Tabla 4, Apéndice 3.

b **Ejercicio Applet** Use la aplicación breve *Normal Probabilities* para obtener la respuesta.

c ¿Por qué son diferentes los valores marcados en los dos ejes horizontales?

4.65 En el Ejercicio 4.64, ¿cuánto debe presupuestarse para reparaciones y mantenimiento semanal para lograr que la probabilidad de que la cantidad presupuestada en una semana determinada sea excedida sólo .1?

4.66 Una operación de maquinado produce cojinetes con diámetros que están normalmente distribuidos con media de 3.0005 pulgadas y desviación estándar de .0010 pulgadas. Las especificaciones requieren que los diámetros de los cojinetes se encuentren en el intervalo $3.000 \pm .0020$ pulgadas. Los cojinetes que estén fuera de este intervalo son considerados de desecho y deben volver a maquinarse. Con el ajuste de la máquina existente, ¿qué fracción de la producción total se desechará?

a Conteste la pregunta usando la Tabla 4, Apéndice 3.

b **Ejercicio Applet** Obtenga la respuesta usando la aplicación breve *Normal Probabilities.*

4.67 En el Ejercicio 4.66, ¿cuál debe ser el diámetro medio para que la fracción de cojinetes desechados sea mínima?

4.68 Los promedios de calificaciones (GPA, por sus siglas en inglés) de una gran población de estudiantes universitarios están normalmente distribuidos en forma aproximada, con media de 2.4 y desviación estándar .8. ¿Qué fracción de los estudiantes alcanzarán un GPA de más de 3.0?

a Conteste la pregunta usando la Tabla 4, Apéndice 3.

b **Ejercicio Applet** Use la aplicación breve *Normal Tail Areas and Quantiles* para obtener la respuesta.

4.69 Consulte el Ejercicio 4.68. Si los estudiantes que alcancen un GPA menor que 1.9 serán suspendidos de la universidad, ¿qué porcentaje de los estudiantes será suspendido?

4.70 Consulte el Ejercicio 4.68. Suponga que se seleccionan al azar tres estudiantes del alumnado. ¿Cuál es la probabilidad de que los tres alcancen un GPA de más de 3.0?

4.71 Se especifica que los cables manufacturados para usarse en un sistema de computadora deben tener resistencias entre .12 y .14 ohms. Las resistencias medidas reales de los cables producidos por la compañía A tienen una distribución de probabilidad normal con media de .13 ohms y desviación estándar .005 ohm.

a ¿Cuál es la probabilidad de que un cable seleccionado al azar de la producción de la compañía A satisfaga las especificaciones?

b Si cuatro de estos cables se usan en el sistema de cada computadora y todos son seleccionados de la compañía A, ¿cuál es la probabilidad de que los cuatro en un sistema seleccionado al azar satisfagan las especificaciones?

4.72 Un método para llegar a pronósticos económicos es usar un método de consenso. Un pronóstico se obtiene de todos y cada uno de un gran número de analistas; el promedio de los pronósticos de estas personas es el pronóstico de consenso. Suponga que los pronósticos de tasa de interés preferencial de enero de 1996 de todos los analistas económicos están distribuidos normalmente en forma aproximada con media de 7% y desviación estándar de 2.6%. Si un solo analista se selecciona al azar entre este grupo, ¿cuál es la probabilidad de que el pronóstico del analista de la tasa de interés preferencial

a exceda de 11%?

b sea menor que 9%?

4.73 El ancho de rollos de tela está normalmente distribuido con media de 950 mm (milímetros) y desviación estándar de 10 mm.

a ¿Cuál es la probabilidad de que un rollo seleccionado al azar tenga un ancho de entre 947 y 958 mm?

b ¿Cuál es el valor apropiado para C de manera que un rollo seleccionado al azar tenga un ancho menor que C con probabilidad .8531?

4.74 Se supone que las calificaciones de un examen están normalmente distribuidas con media de 78 y varianza de 36.

a ¿Cuál es la probabilidad de que una persona que haga el examen alcance calificaciones mayores de 72?

b Suponga que los estudiantes que alcancen el 10% más alto de esta distribución reciben una calificación de A. ¿Cuál es la calificación mínima que un estudiante debe recibir para ganar una calificación de A?

c ¿Cuál debe ser el punto límite para pasar el examen si el examinador desea pasar sólo a 28.1% más alto de todas las calificaciones?

d ¿Aproximadamente qué proporción de estudiantes tienen calificaciones de 5 o más puntos arriba de la calificación que corta al 25% más bajo?

e **Ejercicio Applet** Conteste los incisos a–d usando la aplicación breve *Normal Tail Areas and Quantiles.*

f Si se sabe que la calificación de un estudiante excede de 72, ¿cuál es la probabilidad de que su calificación exceda de 84?

4.75 Una máquina expendedora de bebidas gaseosas puede ser regulada para descargar un promedio de μ onzas por vaso. Si las onzas están normalmente distribuidas con desviación estándar de 0.3 onzas, determine los valores para μ de modo que vasos de 8 onzas se sirvan sólo 1% del tiempo.

4.76 La máquina descrita en el Ejercicio 4.75 tiene desviación estándar σ que se puede fijar en ciertos niveles al ajustar la máquina con todo cuidado. ¿Cuál es el máximo valor de σ que permitirá que la cantidad real servida esté a no más de 1 onza de la media con probabilidad de al menos .95?

4.77 Los exámenes de admisión SAT y ACT (de aptitud y universitario) se aplican a miles de estudiantes cada año. Las secciones de matemáticas de cada uno de estos exámenes producen calificaciones que están normalmente distribuidas, en forma aproximada. En años recientes las calificaciones de exámenes SAT de matemáticas han promediado 480 con desviación estándar de 100. El promedio y desviación estándar para calificaciones ACT de matemáticas son 18 y 6, respectivamente.

a Una escuela de ingeniería establece 550 como calificación mínima SAT de matemáticas para estudiantes de nuevo ingreso. ¿Qué porcentaje de estudiantes obtendrá una calificación por debajo de 550 en un año típico?

b ¿Qué calificación debe establecer la escuela de ingeniería como estándar comparable en el examen ACT de matemáticas?

4.78 Demuestre que el máximo valor de la densidad normal con parámetros μ y σ es $1/(\sigma\sqrt{2\pi})$ y sucede cuando $y = \mu$.

4.79 Demuestre que la densidad normal con parámetros μ y σ tiene puntos de inflexión en los valores $\mu - \sigma$ y $\mu + \sigma$. (Recuerde que un punto de inflexión es aquel donde la curva cambia de dirección de cóncava hacia arriba a cóncava hacia abajo o viceversa y ocurre cuando la segunda derivada cambia de signo. Este cambio en signo puede presentarse cuando la segunda derivada es igual a cero.)

4.80 Suponga que Y está normalmente distribuida con media μ y desviación estándar σ. Después de observar el valor de Y, un matemático construye un rectángulo con longitud $L = |Y|$ y ancho $W = 3|Y|$. Denote con A el área del triángulo resultante. ¿Cuál es $E(A)$?

4.6 La distribución de probabilidad gamma

Algunas variables aleatorias son siempre no negativas y por varias razones dan distribuciones de datos que está sesgadas (no simétricas) a la derecha. Esto es, casi toda el área bajo la función de densidad está ubicada cerca del origen y la función de densidad cae gradualmente conforme y aumenta. En la Figura 4.15 se muestra una función de densidad de probabilidad sesgada.

Los intervalos de tiempo entre mal funcionamiento de motores de aviones poseen una distribución de frecuencia sesgada, al igual que los intervalos de llegada en una fila de espera en las cajas de un supermercado (esto es, la fila de espera para llegar a la caja a pagar). Del mismo modo, los intervalos de tiempo para completar una revisión de mantenimiento para un motor de automóvil o de avión poseen una distribución de frecuencia sesgada. La población asociada con estas variables aleatorias posee con frecuencia funciones de densidad que son modeladas de manera adecuada por una función de densidad gamma.

DEFINICIÓN **4.9**

Se dice que una variable aleatoria Y tiene una *distribución gamma con parámetros* $\alpha > 0$ y $\beta > 0$ si y sólo si la función de densidad de Y es

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & 0 \le y < \infty, \\ 0, & \text{en cualquier otro punto,} \end{cases}$$

donde

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\, dy.$$

La cantidad $\Gamma(\alpha)$ se conoce como *función gamma.* La integración directa verificará que $\Gamma(1) = 1$. La integración por partes verifica que $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ para cualquier $\alpha > 1$ y que $\Gamma(n) = (n - 1)!$, siempre que n sea un entero.

En la Figura 4.16 se dan gráficas de funciones de densidad gamma para $\alpha = 1$, 2 y 4 y $\beta = 1$. Observe en la Figura 4.16 que la forma de la densidad gamma difiere para los diferentes valores de α. Por esta razón, α recibe a veces el nombre de *parámetro de forma* asociado con una distribución gamma. El parámetro β generalmente se llama *parámetro de escala* porque multiplicar una variable aleatoria con distribución gamma por una constante positiva (y por tanto cambiando la escala en la que se hace la medición) produce una variable aleatoria

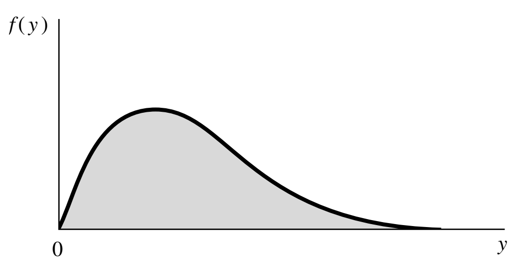


FIGURA **4.15** Función de densidad de probabilidad sesgada

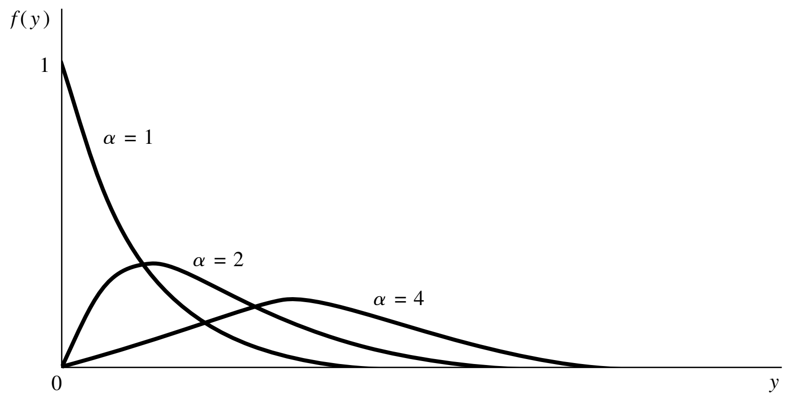


FIGURA **4.16** Funciones de densidad gamma, $\beta = 1$

que también tiene una distribución gamma con el mismo valor de α (parámetro de forma) pero con un valor alterado de β.

En el caso especial cuando α es un entero, la función de distribución de una variable aleatoria con distribución gamma puede expresarse como una suma de ciertas probabilidades de Poisson. Encontrará esta representación en el Ejercicio 4.99. Si α no es un entero y $0 < c < d < \infty$, es imposible dar una expresión de forma cerrada para

$$\int_c^d \frac{y^{\alpha-1}e^{-y/\beta}}{\beta^\alpha \Gamma(\alpha)}\,dy.$$

Por esta razón, excepto cuando $\alpha = 1$ (una distribución exponencial), es imposible obtener áreas bajo la función de densidad gamma por integración directa. Valores tabulados para integrales como ésta se dan en *Tables of the Incomplete Gamma Function* (Pearson 1965). Por mucho, la forma más fácil de calcular probabilidades asociadas con variables aleatorias de distribución gamma es usar un software de estadística. Si Y es una variable aleatoria con distribución gamma y parámetros α y β, el comando `pgamma(y0,α,1/β)` de R (o S–Plus) genera $P(Y \le y_0)$, mientras que `qgamma(q,α,1/β)` da el p–ésimo cuantil, el valor de ϕ_p tal que $P(Y \le \phi_p) = p$. Además, una de las aplicaciones breves, *Gamma Probabilities and Quantiles*, accesible en www.thomsonedu.com/statistics/wackerly, se puede usar para determinar probabilidades y cuantiles asociados con variables aleatorias de distribución gamma. Otra aplicación breve en la página web de Thomson, *Comparison of Gamma Density Functions,* permitirá visualizar y comparar funciones de densidad gamma con diferentes valores para α y/o β. Estas aplicaciones breves se usarán para contestar algunos de los ejercicios del final de esta sección.

Como se indica en el siguiente teorema, la media y la varianza de variables aleatorias de distribución gamma son fáciles de calcular.

TEOREMA **4.8** Si Y tiene una distribución gamma con parámetros α y β, entonces

$$\mu = E(Y) = \alpha\beta \quad \text{y} \quad \sigma^2 = V(Y) = \alpha\beta^2.$$

Demostración

$$E(Y) = \int_{-\infty}^{\infty} yf(y)\,dy = \int_{0}^{\infty} y\left(\frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}\right) dy.$$

Por definición, la función de densidad gamma es tal que

$$\int_{0}^{\infty} \frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}\,dy = 1.$$

Por tanto,

$$\int_{0}^{\infty} y^{\alpha-1}e^{-y\ \beta}\,dy = \beta^{\alpha}\Gamma(\alpha),$$

y

$$E(Y) = \int_{0}^{\infty} \frac{y^{\alpha}e^{-y/\beta}\,dy}{\beta^{\alpha}\Gamma(\alpha)} = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_{0}^{\infty} y^{\alpha}e^{-y/\beta}\,dy$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}[\beta^{\alpha+1}\Gamma(\alpha+1)] = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta.$$

Del Ejercicio 4.24, $V(Y) = E[Y^2] - [E(Y)]^2$. Además,

$$E(Y^2) = \int_{0}^{\infty} y^2\left(\frac{y^{\alpha-1}e^{-y\ \beta}}{\beta^{\alpha}\Gamma(\alpha)}\right) dy = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_{0}^{\infty} y^{\alpha+1}e^{-y\ \beta}\,dy$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}[\beta^{\alpha+2}\Gamma(\alpha+2)] = \frac{\beta^2(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha(\alpha+1)\beta^2.$$

Entonces $V(Y) = E[Y^2] - [E(Y)]^2$, donde, desde la primera parte de la derivación, $E(Y) = \alpha\beta$. Sustituyendo $E[Y^2]$ y $E(Y)$ en la fórmula para $V(Y)$, obtenemos

$$V(Y) = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha^2\beta^2 + \alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$$

Dos casos especiales de variables aleatorias con distribución gamma ameritan consideración particular.

DEFINICIÓN 4.10 Sea ν un entero positivo. Se dice que una variable aleatoria Y tiene *distribución ji cuadrada con* ν *grados de libertad* si y sólo si Y es una variable aleatoria con distribución gamma y parámetros $\alpha = \nu/2$ y $\beta = 2$.

Una variable aleatoria con distribución ji cuadrada se denomina *variable aleatoria* (χ^2) *ji cuadrada*. Estas variables aleatorias se presentan con frecuencia en teoría estadística. La motivación que hay detrás de llamar al parámetro ν como grados de libertad de la distribución χ^2 se apoya en una de las principales formas de generar una variable aleatoria con esta distribución y se da en el Teorema 6.4. La media y la varianza de una variable aleatoria χ^2 provienen directamente del Teorema 4.8.

TEOREMA **4.9** Si Y es una variable aleatoria ji cuadrada con ν grados de libertad, entonces

$$\mu = E(Y) = \nu \quad \text{y} \quad \sigma^2 = V(Y) = 2\nu.$$

Demostración Aplique el Teorema 4.8 con $\alpha = \nu/2$ y $\beta = 2$.

En casi todos los textos de estadística se pueden ver tablas que dan probabilidades asociadas con distribuciones χ^2. La Tabla 6, Apéndice 3, da puntos porcentuales asociados con distribuciones χ^2 para numerosas opciones de ν. No se dispone fácilmente de tablas de la distribución gamma general, pero demostraremos en el Ejercicio 6.46 que si Y tiene una distribución gamma con $\alpha = n/2$ para algún entero n, entonces $2Y/\beta$ tiene una distribución χ^2 con n grados de libertad. De ahí que, por ejemplo, si Y tiene una distribución gamma con $\alpha = 1.5 = 3/2$ y $\beta = 4$, entonces $2Y/\beta = 2Y/4 = Y/2$ tiene una distribución χ^2 con 3 grados de libertad. Entonces, $P(Y < 3.5) = P([Y/2] < 1.75)$ se puede hallar usando tablas de la distribución χ^2 de las que se puede disponer fácilmente.

La función de densidad gamma en la que $\alpha = 1$, se llama *función de densidad exponencial.*

DEFINICIÓN **4.11** Se dice que una variable aleatoria Y tiene una *distribución exponencial con parámetro* $\beta > 0$ si y sólo si la función de densidad de Y es

$$f(y) = \begin{cases} \dfrac{1}{\beta} e^{-y/\beta}, & 0 \le y < \infty, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

La función de densidad exponencial a menudo es de ayuda para modelar la vida útil de componentes electrónicos. Suponga que el tiempo que ya ha operado un componente no afecta su probabilidad de operar durante al menos b unidades de tiempo adicionales. Esto es, la probabilidad de que el componente opere durante más de $a + b$ unidades de tiempo, dado que ya ha operado durante al menos a unidades de tiempo, es la misma que la probabilidad de que un componente nuevo opere al menos b unidades de tiempo si el componente nuevo se pone en servicio en el tiempo 0. Un fusible es un ejemplo de un componente para el cual a veces esta suposición es razonable. Veremos en el siguiente ejemplo que la distribución exponencial proporciona un modelo para la distribución de la vida útil de ese componente.

TEOREMA **4.10** Si Y es una variable aleatoria exponencial con parámetro β, entonces

$$\mu = E(Y) = \beta \quad \text{y} \quad \sigma^2 = V(Y) = \beta^2.$$

Demostración La demostración se sigue directamente del Teorema 4.8 con $\alpha = 1$.

EJEMPLO **4.10** Suponga que Y tiene una función de densidad de probabilidad exponencial. Demuestre que, si $a > 0$ y $b > 0$,

$$P(Y > a + b | Y > a) = P(Y > b).$$

Solución De la definición de probabilidad condicional, tenemos que

$$P(Y > a + b|Y > a) = \frac{P(Y > a + b)}{P(Y > a)}$$

porque la intersección de los eventos $(Y > a + b)$ y $(Y > a)$ es el evento $(Y > a + b)$. Ahora

$$P(Y > a + b) = \int_{a+b}^{\infty} \frac{1}{\beta} e^{-y/\beta}\, dy = -e^{-y/\beta}\Big]_{a+b}^{\infty} = e^{-(a+b)/\beta}.$$

De manera similar,

$$P(Y > a) = \int_{a}^{\infty} \frac{1}{\beta} e^{-y/\beta}\, dy = e^{-a/\beta},$$

y

$$P(Y > a + b|Y > a) = \frac{e^{-(a+b)/\beta}}{e^{-a/\beta}} = e^{-b/\beta} = P(Y > b).$$

Esta propiedad de la distribución exponencial en ocasiones recibe el nombre de *propiedad sin memoria* de la distribución. ■

Como recordará del Capítulo 3, la distribución geométrica, que es una distribución discreta, también tenía propiedad *sin memoria.* Una relación interesante entre las distribuciones exponencial y geométrica se da en el Ejercicio 4.95.

Ejercicios

4.81 **a** Si $\alpha > 0$, $\Gamma(\alpha)$ está definida por $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\,dy$, demuestre que $\Gamma(1) = 1$.

*__b__ Si $\alpha > 1$, integre por partes para demostrar que $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.

4.82 Use los resultados obtenidos en el Ejercicio 4.81 para demostrar que si n es un entero positivo, entonces $\Gamma(n) = (n - 1)!$. ¿Cuáles son los valores numéricos de $\Gamma(2)$, $\Gamma(4)$ y $\Gamma(7)$?

4.83 **Ejercicio Applet** Use la aplicación breve *Comparison of Gamma Density Functions* para obtener los resultados dados en la Figura 4.16.

4.84 **Ejercicio Applet** Consulte el Ejercicio 4.83. Use la aplicación breve *Comparison of Gamma Density Functions* para comparar funciones con densidad gamma con $(\alpha = 4,\ \beta = 1)$, $(\alpha = 40,\ \beta = 1)$ y $(\alpha = 80,\ \beta = 1)$.

a ¿Qué observa usted acerca de las formas de estas tres funciones de densidad? ¿Cuáles son menos sesgadas y más simétricas?

b ¿Qué diferencias observa acerca de la ubicación de los centros de estas funciones de densidad?

c Dé una explicación de lo que observó en el inciso b.

4.85 **Ejercicio Applet** Use la aplicación breve *Comparison of Gamma Density Functions* para comparar funciones de densidad gamma con $(\alpha = 1, \ \beta = 1)$, $(\alpha = 1, \ \beta = 2)$ y $(\alpha = 1, \ \beta = 4)$.

a ¿Qué otro nombre se da a las funciones de densidad que observó?

b ¿Estas densidades tienen la misma forma general?

c El parámetro β es un parámetro "de escala". ¿Qué observa usted acerca de la "dispersión" de estas tres funciones de densidad?

4.86 **Ejercicio Applet** Cuando examinamos la distribución χ^2 en esta sección, presentamos (con justificación en el Capítulo 6) el hecho de que si Y tiene distribución gamma con $\alpha = n/2$ para algún entero n, entonces $2Y/\beta$ tiene una distribución χ^2. En particular, se dijo que cuando $\alpha = 1.5$ y $\beta = 4$, $W = Y/2$ tiene una distribución χ^2 con tres grados de libertad.

a Use la aplicación *Gamma Probabilities and Quantiles* para hallar $P(Y < 3.5)$.

b Use la aplicación *Gamma Probabilities and Quantiles* para hallar $P(W < 1.75)$. [*Sugerencia:* Recuerde que la distribución χ^2 con ν grados de libertad es simplemente una distribución gamma con $\alpha = \nu/2$ y $\beta = 2$.]

c Compare sus respuestas a los incisos a y b.

4.87 **Ejercicio Applet** Hagamos que Y y W tengan las distribuciones dadas en el Ejercicio 4.86.

a Use la aplicación *Gamma Probabilities and Quantiles* para hallar el cuantil .05 de la distribución de Y.

b Use la aplicación *Gamma Probabilities and Quantiles* para hallar el cuantil .05 de la distribución χ^2 con tres grados de libertad.

c ¿Cuál es la relación entre el cuantil .05 de la distribución gamma $(\alpha = 1.5, \beta = 4)$ y el cuantil .05 de la distribución χ^2 con tres grados de libertad? Explique esta relación.

4.88 La magnitud de temblores registrados en una región de América del Norte puede modelarse como si tuviera una distribución exponencial con media 2.4, según se mide en la escala de Richter. Encuentre la probabilidad de que un temblor que ocurra en esta región

a sea mayor que 3.0 en la escala de Richter.

b caiga entre 2.0 y 3.0 en la escala de Richter.

4.89 Si Y tiene una distribución exponencial y $P(Y > 2) = .0821$, ¿cuál es

a $\beta = E(Y)$?

b $P(Y \leq 1.7)$?

4.90 Consulte el Ejercicio 4.88. De los siguientes diez temblores que afecten esta región, ¿cuál es la probabilidad de que al menos uno de ellos sea mayor que 5.0 en la escala de Richter?

4.91 El operador de una estación de bombeo ha observado que la demanda de agua durante las primeras horas de la tarde tiene una distribución aproximadamente exponencial con media de 100 pes (pcs cúbicos por segundo).

a Encuentre la probabilidad de que la demanda sea mayor que 200 pcs durante las primeras horas de la tarde en un día seleccionado al azar.

b ¿Qué capacidad de bombeo de agua debe mantener la estación durante las primeras horas de la tarde para que la probabilidad de que la demanda sea mayor que la capacidad en un día seleccionado al azar sea de sólo .01?

4.92 El tiempo Y necesario para completar una operación clave en la construcción de casas tiene una distribución exponencial con media de 10 horas. La fórmula $C = 100 + 40Y + 3Y^2$ relaciona el costo C de completar esta operación con el cuadrado del tiempo para completarla. Encuentre la media y la varianza de C.

4.93 Una evidencia histórica indica que los tiempos entre accidentes mortales en vuelos nacionales de horario programado en aviones de pasajeros en Estados Unidos tienen una distribución aproximadamente exponencial. Suponga que el tiempo medio entre accidentes es de 44 días.

a Si uno de los accidentes ocurrió el 1 de julio de un año seleccionado al azar en el periodo de estudio, ¿cuál es la probabilidad de que ocurra otro accidente ese mismo mes?

b ¿Cuál es la varianza de los tiempos entre accidentes?

4.94 Concentraciones de monóxido de carbono de una hora en muestras de aire de una gran ciudad tienen una distribución aproximadamente exponencial, con media de 3.6 ppm (partes por millón).

a Encuentre la probabilidad de que la concentración de monóxido de carbono exceda de 9 ppm durante un periodo de una hora seleccionado al azar.

b Una estrategia de control de tránsito redujo la media a 2.5 ppm. Ahora encuentre la probabilidad de que la concentración exceda de 9 ppm.

4.95 Sea Y una variable aleatoria distribuida exponencialmente con media β. Defina una variable aleatoria X en la siguiente forma: $X = k$ si $k - 1 \leq Y < k$ para $k = 1, 2, \ldots$

a Encuentre $P(X = k)$ para cada $k = 1, 2, \ldots$

b Demuestre que su respuesta al inciso a se puede escribir como

$$P(X = k) = \left(e^{-1/\beta}\right)^{k-1}\left(1 - e^{-1/\beta}\right), \qquad k = 1,\ 2, \ldots$$

y que X tiene una distribución geométrica con $p = (1 - e^{-1/\beta})$.

4.96 Suponga que una variable aleatoria Y tiene una función de densidad de probabilidad dada por

$$f(y) = \begin{cases} ky^3 e^{-y/2}, & y > 0. \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

a Encuentre el valor de k que haga de $f(y)$ una función de densidad.

b ¿Tiene Y una distribución χ^2? Si es así, ¿de cuántos grados de libertad?

c ¿Cuáles son la media y la desviación estándar de Y?

d **Ejercicio Applet** ¿Cuál es la probabilidad de que Y se encuentre a no más de 2 desviaciones estándar de su media?

4.97 Una planta de manufactura utiliza un producto específico a granel. La cantidad de producto empleada en un día puede ser modelada por una distribución exponencial con $\beta = 4$ (medida en toneladas). Encuentre la probabilidad de que la planta utilice más de 4 toneladas en un día determinado.

4.98 Considere la planta del Ejercicio 4.97. ¿Cuánto producto a granel debe tener en existencia para que la probabilidad de que se agote el producto en la planta sea de sólo .05?

4.99 Si $\lambda > 0$ y α es un entero positivo, la relación entre integrales gamma incompletas y sumas de probabilidades de Poisson está dada por

$$\frac{1}{\Gamma(\alpha)} \int_{\lambda}^{\infty} y^{\alpha-1} e^{-y}\, dy = \sum_{x=0}^{\alpha-1} \frac{\lambda^x e^{-\lambda}}{x!}.$$

a Si Y tiene una distribución gamma con $\alpha = 2$ y $\beta = 1$, encuentre $P(Y > 1)$ usando la igualdad anterior y la Tabla 3 del Apéndice 3.

b **Ejercicio Applet** Si Y tiene una distribución gamma con $\alpha = 2$ y $\beta = 1$, encuentre $P(Y > 1)$ con el uso de la aplicación breve *Gamma Probabilities*.

***4.100** Sea Y una variable aleatoria con distribución gamma donde α es un entero positivo y $\beta = 1$. El resultado dado en el Ejercicio 4.99 implica que si $y > 0$,

$$\sum_{x=0}^{\alpha-1} \frac{y^x e^{-y}}{x!} = P(Y > y).$$

Suponga que X_1 tiene distribución de Poisson con media λ_1 y X_2 tiene distribución de Poisson con media λ_2, donde $\lambda_2 > \lambda_1$.

a Demuestre que $P(X_1 = 0) > P(X_2 = 0)$.

b Sea k cualquier entero positivo fijo. Demuestre que $P(X_1 \leq k) = P(Y > \lambda_1)$ y $P(X_2 \leq k) = P(Y > \lambda_2)$, donde Y es una distribución gamma con $\alpha = k + 1$ y $\beta = 1$.

c Sea k cualquier entero fijo positivo. Use el resultado que obtuvo en el inciso b y el hecho de que $\lambda_2 > \lambda_1$ para demostrar que $P(X_1 \leq k) > P(X_2 \leq k)$.

d Como el resultado del inciso c es válido para cualquier $k = 1, 2, 3, \ldots$ y el inciso a también es válido, hemos establecido que $P(X_1 \leq k) > P(X_2 \leq k)$ para toda $k = 0, 1, 2\ldots$ Interprete este resultado.

4.101 **Ejercicio Applet** Consulte el Ejercicio 4.88. Suponga que la magnitud de los terremotos que afectan la región tiene una distribución gamma con $\alpha = .8$ y $\beta = 2.4$.

a ¿Cuál es la magnitud media de los terremotos que afectan la región?

b ¿Cuál es la probabilidad de que la magnitud de un terremoto que afecte la región exceda de 3.0 en la escala Richter?

c Compare sus respuestas con el Ejercicio 4.88(a). ¿Cuál probabilidad es mayor? Explique.

d ¿Cuál es la probabilidad de que un terremoto que afecte las regiones caiga entre 2.0 y 3.0 en la escala Richter?

4.102 **Ejercicio Applet** Consulte el Ejercicio 4.97. Suponga que la cantidad de producto usado en un día tiene una distribución gamma con $\alpha = 1.5$ y $\beta = 3$.

a Encuentre la probabilidad de que la planta use más de 4 toneladas en un día determinado.

b ¿Cuánto del producto a granel debe haber en existencia para que la probabilidad de que la planta agote el producto sea de sólo .05?

4.103 Materiales explosivos que se usan en operaciones de minería producen cráteres casi circulares cuando se hacen detonar. Los radios de estos cráteres están distribuidos exponencialmente con media de 10 pies. Encuentre la media y la varianza de las áreas producidas por estos materiales explosivos.

4.104 La vida útil (en horas) Y de un componente electrónico es una variable aleatoria con función de densidad dada por

$$f(y) = \begin{cases} \frac{1}{100} e^{-y/100}, & y > 0, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Tres componentes operan de manera independiente en una pieza de equipo. El equipo falla si fallan al menos dos de los componentes. Encuentre la probabilidad de que el equipo opere durante al menos 200 horas sin fallar.

4.105 La cantidad total de lluvia de cuatro semanas en verano en una parte del medio oeste de Estados Unidos tiene aproximadamente una distribución gamma con $\alpha = 1.6$ y $\beta = 2.0$.

a Encuentre la media y varianza del total de lluvia de cuatro semanas.

b **Ejercicio Applet** ¿Cuál es la probabilidad de que la cantidad total de lluvia de cuatro semanas sea mayor que 4 pulgadas?

4.106 Los tiempos de respuesta en una terminal de computadora en línea tienen aproximadamente una distribución gamma con media de cuatro segundos y varianza de ocho segundos^2.

a Escriba una función de densidad de probabilidad para los tiempos de respuesta.

b **Ejercicio Applet** ¿Cuál es la probabilidad de que el tiempo de respuesta en la terminal sea menor que cinco segundos?

4.107 Consulte el Ejercicio 4.106.

a Use el teorema de Tchebysheff para dar un intervalo que contenga al menos 75% de los tiempos de respuesta.

b **Ejercicio Applet** ¿Cuál es la probabilidad real de observar un tiempo de respuesta en el intervalo que obtuvo en el inciso a?

4.108 Los ingresos anuales de jefes de familia en una parte de una ciudad tienen aproximadamente una distribución gamma con $\alpha = 20$ y $\beta = 1000$.

a Encuentre la media y la varianza de estos ingresos.

b ¿Esperaría hallar muchos ingresos de más de \$30,000 en esta sección de la ciudad?

c **Ejercicio Applet** ¿Qué proporción de jefes de familia de esta sección de la ciudad tienen ingresos de más de \$30,000?

4.109 El tiempo improductivo por semana Y (en horas) de una máquina industrial tiene aproximadamente una distribución gamma con $\alpha = 3$ y $\beta = 2$. La pérdida L (en dólares) para la operación industrial como resultado de este tiempo improductivo está dada por $L = 30Y + 2Y^2$. Encuentre la varianza y el valor esperados de L.

4.110 Si Y tiene una función de densidad de probabilidad dada por

$$f(y) = \begin{cases} 4y^2e^{-2y}, & y > 0, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

obtenga $E(Y)$ y $V(Y)$ por inspección.

4.111 Suponga que Y tiene una distribución gamma con parámetros α y β.

a Si a es cualquier valor positivo o negativo tal que $\alpha + a > 0$, demuestre que

$$E(Y^a) = \frac{\beta^a \Gamma(\alpha + a)}{\Gamma(\alpha)}.$$

b ¿Por qué la respuesta en el inciso a requirió que $\alpha + a > 0$?

c Demuestre que, con $a = 1$, el resultado del inciso a da $E(Y) = \alpha\beta$.

d Use el resultado del inciso a para dar una expresión para $E(\sqrt{Y})$. ¿Qué es necesario suponer acerca de α?

e Utilice el resultado del inciso a para obtener una expresión para $E(1/Y)$, $E(/\sqrt{y})$ y $E(1/Y^2)$. ¿Qué es necesario suponer sobre α en cada caso?

4.112 Suponga que Y tiene una distribución χ^2 con ν grados de libertad. Use los resultados del Ejercicio 4.111 en sus respuestas a lo siguiente. Estos resultados serán útiles cuando estudiemos las distribuciones t y F en el Capítulo 7.

a Proporcione una expresión para $E(Y^a)$ si $\nu > -2a$.

b ¿Por qué la respuesta en el inciso a requirió que $\nu > -2a$?

c Use el resultado del inciso a para dar una expresión para $E(\sqrt{Y})$. ¿Qué es necesario suponer acerca de ν?

d Use el resultado del inciso a para dar una expresión para $E(1/Y)$, $E(1/\sqrt{Y})$ y $E(1/Y^2)$. ¿Qué es necesario suponer acerca de ν en cada caso?

4.7 La distribución de probabilidad beta

La función de densidad beta es una función de densidad de dos parámetros definida sobre el intervalo cerrado $0 \le y \le 1$. Frecuentemente se usa como modelo para proporciones, por ejemplo como la proporción de impurezas en un producto químico o la proporción de tiempo que una máquina está en reparación.

DEFINICIÓN **4.12**

Se dice que una variable aleatoria Y tiene una *distribución de probabilidad beta con parámetros* $\alpha > 0$ *y* $\beta > 0$ si y sólo si la función de densidad de Y es

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}, & 0 \le y \le 1, \\ 0, & \text{en cualquier otro punto,} \end{cases}$$

donde

$$B(\alpha,\beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\,dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Las gráficas de funciones de densidad beta toman formas muy diferentes para diversos valores de los dos parámetros α y β. Algunos de éstos se muestran en la Figura 4.17. Ciertos ejercicios del final de esta sección piden al lector usar la applet *Comparison of Beta Density Functions* accesibles en www.thomsonedu.com/statistics/wackerly para explorar y comparar las formas de más densidades beta.

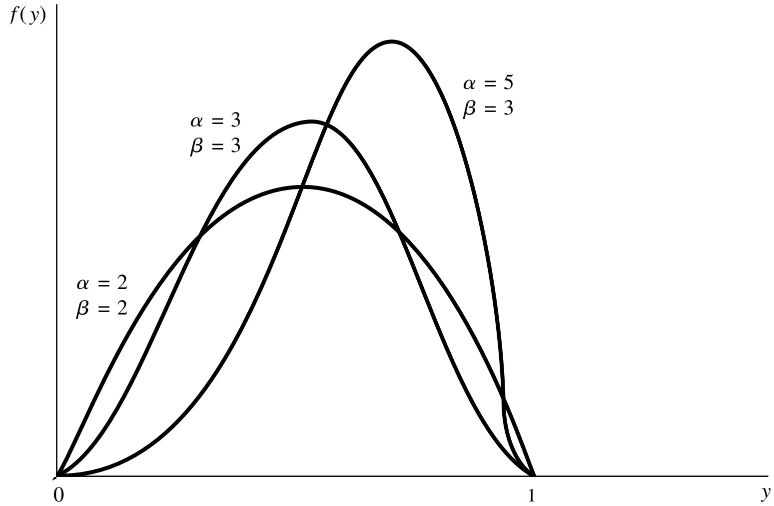
Observe que definir y sobre el intervalo $0 \le y \le 1$ no restringe el uso de la distribución beta. Si $c \le y \le d$, entonces $y^* = (y - c)/(d - c)$ define una nueva variable tal que $0 \le y^* \le 1$. Entonces, la función de densidad beta se puede aplicar a una variable aleatoria definida en el intervalo $c \le y \le d$ por traducción y un cambio de escala.

La función de distribución acumulativa para la variable aleatoria beta comúnmente se denomina *función beta incompleta* y está denotada por

$$F(y) = \int_0^y \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}\,dt = I_y(\alpha,\beta)$$

Una tabulación de $I_y(\alpha, \beta)$ se da en la obra *Tables of the Incomplete Beta Function* (Pearson, 1968). Cuando α y β son enteros positivos $I_y(\alpha, \beta)$ está relacionada con la función de proba-

FIGURA **4.17** Funciones de densidad beta



bilidad binomial. Es posible usar integración por partes para demostrar que $0 < y < 1$, y α y β ambos enteros,

$$F(y) = \int_0^y \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha, \beta)}\, dt = \sum_{i=\alpha}^{n} \binom{n}{i} y^i (1-y)^{n-i},$$

donde $n = \alpha + \beta - 1$. Observe que la suma del lado derecho de esta expresión es precisamente la suma de probabilidades asociadas con una variable aleatoria binomial $n = \alpha + \beta - 1$ y $p = y$. La función de distribución acumulativa binomial se presenta en la Tabla 1, Apéndice 3, para $n = 5$, 10, 15, 20 y 25 y $p = .01$, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95 y .99. El modo más eficiente de obtener probabilidades binomiales es usar un software de estadística como el *R* o *S*–Plus (vea el Capítulo 3). Una forma incluso más fácil para hallar probabilidades y cuantiles asociados con variables aleatorias de distribución beta es usar directamente software apropiado. La página web de Thomson contiene una aplicación breve, *Beta Probabilities,* que proporciona probabilidades de "cola superior" [es decir, $P(Y > y_0)$] y cuantiles asociados con variables aleatorias con distribución beta. Además, si Y es una variable aleatoria con distribución beta y parámetros α y β, el comando `pbeta(y0, α, 1/β)` de *R* (o *S*–Plus) genera $P(Y \leq y_0)$, mientras que `qbeta(p, α, 1/β)` da el p–ésimo cuantil, el valor de ϕ_p de manera que $P(Y \leq \phi_p) = p$.

TEOREMA **4.11**

Si Y es una variable aleatoria con distribución beta $\alpha > 0$ y $\beta > 0$, entonces

$$\mu = E(Y) \quad \frac{\alpha}{\alpha + \beta} \quad \text{y} \quad \sigma^2 = V(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Demostración Por definición,

$$
\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty} yf(y)\,dy \\
&= \int_{0}^{1} y\left[\frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}\right] dy \\
&= \frac{1 \quad \beta}{B(\alpha,\beta)}\int_{0}^{1} y^{\alpha}(1-y)^{\beta-1}\,dy \\
&= \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} \qquad \text{(porque } \alpha > 0 \text{ implica que } \alpha+1>0\text{)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\alpha}{(\alpha+\beta)}.
\end{aligned}
$$

La obtención de la varianza se deja al lector (vea Ejercicio 4.130).

Veremos en el siguiente ejemplo que la función de densidad beta se puede integrar directamente cuando α y β son enteros.

EJEMPLO **4.11** Una distribuidora mayorista de gasolina tiene tanques de almacenamiento a granel que contienen suministros fijos y se llenan cada lunes. De interés para la mayorista es la proporción de este suministro que se vende durante la semana. Durante varias semanas de observación, la distribuidora encontró que esta proporción podría ser modelada por una distribución beta con $\alpha = 4$ y $\beta = 2$. Encuentre la probabilidad de que la mayorista venda al menos 90% de su existencia en una semana determinada.

Solución Si Y denota la proporción vendida durante la semana, entonces

$$
f(y) = \begin{cases} \dfrac{\Gamma(4+2)}{\Gamma(4)\Gamma(2)} y^3(1-y), & 0 \le y \le 1, \\[2ex] 0, & \text{en otro lugar,} \end{cases}
$$

y

$$
\begin{aligned}
P(Y > .9) &= \int_{.9}^{\infty} f(y)\,dy = \int_{.9}^{1} 20(y^3 - y^4)\,dy \\
&= 20\left\{ \left.\frac{y^4}{4}\right]_{.9}^{1} - \left.\frac{y^5}{5}\right]_{.9}^{1} \right\} = 20(.004) = .08.
\end{aligned}
$$

No es muy probable que 90% de la existencia se venda en una semana determinada. ■

Ejercicios

4.113 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para obtener los resultados dados en la Figura 4.17.

4.114 **Ejercicio Applet** Consulte el Ejercicio 4.113. Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = 1,\ \beta = 1)$, $(\alpha = 1,\ \beta = 2)$ y $(\alpha = 2,\ \beta = 1)$.

a ¿Cómo hemos llamado previamente a la distribución beta con $(\alpha = 1, \beta = 1)$?

b ¿Cuál de estas densidades beta es simétrica?

c ¿Cuál de estas densidades beta está sesgada a la derecha?

d ¿Cuál de estas densidades beta está sesgada a la izquierda?

***e** En el Capítulo 6 veremos que si Y tiene distribución beta con parámetros α y β, entonces $Y^* = 1 - Y$ tiene una distribución beta con parámetros $\alpha^* = \beta$ y $\beta^* = \alpha$. ¿Esto explica las diferencias en las gráficas de las densidades beta?

4.115 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = 2,\ \beta = 2)$, $(\alpha = 3,\ \beta = 3)$ y $(\alpha = 9,\ \beta = 9)$.

a ¿Cuáles son las medias asociadas con las variables aleatorias con cada una de estas distribuciones beta?

b ¿Qué es semejante acerca de estas densidades?

c ¿Cómo difieren estas densidades? En particular, ¿qué observa usted acerca de la “dispersión” de estas tres funciones de densidad?

d Calcule las desviaciones estándar asociadas con las variables aleatorias con cada una de estas densidades beta. ¿Los valores de estas desviaciones estándar explican lo que se observó en el inciso c? Explique.

e Grafique algunas densidades beta adicionales con $\alpha = \beta$. ¿Qué se puede conjeturar acerca de la forma de las densidades beta con $\alpha = \beta$?

4.116 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = 1.5,\ \beta = 7)$, $(\alpha = 2.5,\ \beta = 7)$ y $(\alpha = 3.5,\ \beta = 7)$.

a ¿Son simétricas estas densidades? ¿Sesgadas a la izquierda? ¿Sesgadas a la derecha?

b ¿Qué observa usted a medida que el valor de α se acerca a 7?

c Grafique algunas densidades beta adicionales con $\alpha > 1$, $\beta > 1$, y $\alpha < \beta$. ¿Qué puede conjeturar acerca de la forma de las densidades beta cuando $\alpha > 1$, $\beta > 1$ y $\alpha < \beta$?

4.117 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = 9, \beta = 7)$, $(\alpha = 10, \beta = 7)$ y $(\alpha = 12, \beta = 7)$.

a ¿Son simétricas estas densidades? ¿Sesgadas a la izquierda? ¿Sesgadas a la derecha?

b ¿Qué observa usted a medida que el valor de α se acerca a 12?

c Grafique algunas densidades beta adicionales con $\alpha > 1$, $\beta > 1$ y $\alpha > \beta$. ¿Qué se puede conjeturar acerca de la forma de las densidades beta con $\alpha > \beta$ y $\alpha > 1$ y $\beta > 1$?

4.118 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = .3, \beta = 4)$, $(\alpha = .3, \beta = 7)$ y $(\alpha = .3, \beta = 12)$.

a ¿Son simétricas estas densidades? ¿Sesgadas a la izquierda? ¿Sesgadas a la derecha?

b ¿Qué observa usted a medida que el valor de β se acerca a 12?

c ¿Cuál de estas distribuciones beta da la más alta probabilidad de observar un valor mayor que .2?

d Grafique algunas densidades beta adicionales con $\alpha < 1$ y $\beta > 1$. ¿Qué se puede conjeturar acerca de la forma de las densidades beta con $\alpha < 1$ y $\beta > 1$?

4.119 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = 4,\ \beta = .3)$, $(\alpha = 7,\ \beta = .3)$, y $(\alpha = 12,\ \beta = .3)$.

a ¿Son simétricas estas densidades? ¿Sesgadas a la izquierda? ¿Sesgadas a la derecha?

b ¿Qué observa usted cuando el valor de α se acerca a 12?

c ¿Cuál de estas distribuciones beta da la más alta probabilidad de observar un valor menor que .8?

d Grafique algunas densidades beta adicionales con $\alpha > 1$ y $\beta < 1$. ¿Qué se puede conjeturar acerca de la forma de las densidades beta con $\alpha > 1$ y $\beta < 1$?

***4.120** En el Capítulo 6 veremos que si Y tiene distribución beta con parámetros α y β, entonces $Y^* = 1 - Y$ tiene una distribución beta con parámetros $\alpha^* = \beta$ y $\beta^* = \alpha$. ¿Explica esto las diferencias y similitudes en las gráficas de las densidades beta en los Ejercicios 4.118 y 4.119?

4.121 **Ejercicio Applet** Use la aplicación breve *Comparison of Beta Density Functions* para comparar funciones de densidad beta con $(\alpha = .5,\ \beta = .7)$, $(\alpha = .7,\ \beta = .7)$ y $(\alpha = .9,\ \beta = .7)$.

a ¿Cuál es la forma general de estas densidades?

b ¿Qué observa cuando el valor de α aumenta?

4.122 **Ejercicio Applet** Las densidades beta con $\alpha < 1$ y $\beta < 1$ son difíciles de exhibir debido a problemas de escala o resolución.

a Use la aplicación breve *Beta Probabilities and Quantiles* para calcular $P(Y > .1)$ si Y tiene una distribución beta con $(\alpha = .1, \beta = 2)$.

b Use la aplicación breve *Beta Probabilities and Quantiles* para calcular $P(Y < .1)$ si Y tiene una distribución beta con $(\alpha = .1, \beta = 2)$.

c Con base en su respuesta el inciso b, ¿a cuáles valores de Y se les asignan altas probabilidades si Y tiene una distribución beta con $(\alpha = .1, \beta = 2)$?

d Use la aplicación breve *Beta Probabilities and Quantiles* para calcular $P(Y < .1)$ si Y tiene una distribución beta con $(\alpha = .1, \beta = .2)$.

e Use la aplicación breve *Beta Probabilities and Quantiles* para calcular $P(Y > .9)$ si Y tiene una distribución beta con $(\alpha = .1, \beta = .2)$.

f Use la aplicación breve *Beta Probabilities and Quantiles* para calcular $P(.1 < Y < .9)$ si Y tiene una distribución beta con $(\alpha = 1, \beta = .2)$.

g Con base en sus respuestas a los incisos d, e y f, ¿a cuáles valores de Y se les asignan altas probabilidades si Y tiene una distribución beta con $(\alpha = .1, \beta = .2)$?

4.123 La humedad relativa Y, cuando se mide en una localidad, tiene una función de densidad de probabilidad dada por

$$f(y) = \begin{cases} ky^3(1-y)^2, & 0 \le y \le 1, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

a Encuentre el valor de k que haga de $f(y)$ una función de densidad.

b **Ejercicio Applet** Use la aplicación breve *Beta Probabilities and Quantiles* para hallar un valor de humedad que se exceda sólo 5% del tiempo.

4.124 El porcentaje de impurezas por lote en un producto químico es una variable aleatoria Y con función de densidad

$$f(y) = \begin{cases} 12y^2(1-y), & 0 \le y \le 1, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Un lote con más de 40% de impurezas no se puede vender.

a Integre la densidad directamente para determinar la probabilidad de que un lote seleccionado al azar no se pueda vender por su exceso de impurezas.

b **Ejercicio Applet** Use la aplicación breve *Beta Probabilities and Quantiles* para hallar la respuesta al inciso a.

4.125 Consulte el Ejercicio 4.124. Encuentre la media y la varianza del porcentaje de impurezas en un lote seleccionado al azar del producto químico.

4.126 Suponga que una variable aleatoria Y tiene una función de densidad de probabilidad dada por

$$f(y) = \begin{cases} 6y(1-y), & 0 \le y \le 1, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

a Encuentre $F(y)$.

b Grafique $F(y)$ y $f(y)$.

c Encuentre $P(.5 \le Y \le .8)$.

4.127 Verifique que si Y tiene una distribución beta con $\alpha = \beta = 1$, entonces Y tiene una distribución uniforme en el intervalo (0, 1). Esto es, la distribución uniforme en el intervalo (0, 1) es un caso especial de una distribución beta.

4.128 El costo Y de reparaciones semanales para una máquina tiene una función de densidad de probabilidad dada por

$$f(y) = \begin{cases} 3(1-y)^2, & 0 < y < 1, \\ 0. & \text{en cualquier otro punto,} \end{cases}$$

con mediciones en cientos de dólares. ¿Cuánto dinero debe presupuestarse a la semana para costos de reparación, de modo que el costo real rebase la cantidad presupuestada sólo 10% del tiempo?

4.129 Durante un turno de ocho horas la proporción de tiempo Y que una máquina troqueladora de láminas metálicas está sin operar por mantenimiento o reparaciones tiene una distribución beta con $\alpha = 1$ y $\beta = 2$. Esto es,

$$f(y) = \begin{cases} 2(1-y), & 0 \le y \le 1, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

El costo (en cientos de dólares) de este tiempo improductivo, debido a producción perdida y costo de mantenimiento y reparación, está dado por $C = 10 + 20Y + 4Y^2$. Encuentre la media y la varianza de C.

4.130 Demuestre que la varianza de una variable aleatoria con distribución beta y parámetros α y β es

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

4.131 Errores en la medición del tiempo de llegada de un frente de onda, desde una fuente acústica, en ocasiones tienen una distribución beta aproximada. Suponga que estos errores, medidos en microsegundos, tienen aproximadamente una distribución beta con $\alpha = 1$ y $\beta = 2$.

a ¿Cuál es la probabilidad de que el error de medición en un ejemplo seleccionado al azar sea menor que .5 μs?

b Obtenga la media y la desviación estándar de los errores de medición.

4.132 La mezcla apropiada de polvos finos y gruesos antes de la sinterización (aglutinación) del cobre es esencial para lograr uniformidad en el producto acabado. Un modo de verificar la homogeneidad de la mezcla es seleccionar muchas muestras pequeñas de los polvos mezclados y medir la proporción del peso total aportado por los polvos finos en cada una. Estas mediciones deben ser relativamente estables si se ha obtenido una mezcla homogénea.

a Suponga que la proporción del peso total aportada por los polvos finos tiene una distribución beta con $\alpha = \beta = 3$. Encuentre la media y la varianza de la proporción de peso aportada por los polvos finos.

b Repita el inciso a si $\alpha = \beta = 2$.

c Repita el inciso a si $\alpha = \beta = 1$.

d ¿Cuál de los casos —incisos a, b o c— dan la mezcla más homogénea?

4.133 La proporción de tiempo por día que todas las cajas en un supermercado están ocupadas es una variable aleatoria Y con una función de densidad dada por

$$f(y) = \begin{cases} cy^2(1-y)^4, & 0 \le y \le 1, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

a Encuentre el valor de c que haga de $f(y)$ una función de densidad de probabilidad.

b Encuentre $E(Y)$. (Use lo que haya aprendido acerca de la distribución tipo beta. Compare sus respuestas con las obtenidas en el Ejercicio 4.28.)

c Calcule la desviación estándar de Y.

d **Ejercicio Applet** Utilice la aplicación breve *Beta Probabilities and Quantiles* para hallar $P(Y > \mu + 2\sigma)$.

4.134 En el texto de esta sección observamos la relación entre la función de distribución de una variable aleatoria con distribución beta y las sumas de probabilidades binomiales. Específicamente, si Y tiene una distribución beta con valores enteros positivos para α y β y $0 < y < 1$,

$$F(y) = \int_0^y \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha, \beta)}\, dt = \sum_{i=\alpha}^{n} \binom{n}{i} y^i (1-y)^{n-i},$$

donde $n = \alpha + \beta - 1$.

a Si Y tiene una distribución beta con $\alpha = 4$ y $\beta = 7$, use las tablas binomiales apropiadas para hallar $P(Y \le .7) = F(.7)$.

b Si Y tiene una distribución beta con $\alpha = 12$ y $\beta = 14$, use las tablas binomiales apropiadas para hallar $P(Y \le .6) = F(.6)$.

c **Ejercicio Applet** Use la aplicación breve *Beta Probabilities and Quantiles* para hallar las probabilidades en los incisos a y b.

***4.135** Suponga que Y_1 y Y_2 son variables aleatorias binomiales con parámetros (n, p_1) y (n, p_2), respectivamente, donde $p_1 < p_2$. (Observe que el parámetro n es el mismo para las dos variables.)

a Use la fórmula binomial para deducir que $P(Y_1 = 0) > P(Y_2 = 0)$.

b Use la relación entre la función de distribución beta y las sumas de probabilidades binomiales dadas en el Ejercicio 4.134 para deducir que, si k es un entero entre 1 y $n - 1$,

$$P(Y_1 \le k) = \sum_{i=0}^{k} \binom{n}{i} (p_1)^i (1-p_1)^{n-i} = \int_{p_1}^{1} \frac{t^k(1-t)^{n-k-1}}{B(k+1, n-k)}\, dt.$$

c Si k es un entero entre 1 y n – 1, el mismo argumento empleado en el inciso b dice que

$$P(Y_2 \leq k) = \sum_{i=0}^{k} \binom{n}{i} (p_2)^i (1 - p_2)^{n-i} = \int_{p_2}^{1} \frac{t^k (1-t)^{n-k-1}}{B(k+1, n-k)} \, dt.$$

Demuestre que si k es cualquier entero entre 1 y n – 1, $P(Y_1 \leq k) > P(Y_2 \leq k)$. Interprete este resultado.

4.8 Algunos comentarios generales

Recuerde que las funciones de densidad son modelos teóricos para poblaciones de datos reales que se presentan en fenómenos aleatorios. ¿Cómo sabemos cuál modelo usar? ¿Cuánto importa si usamos la densidad errónea como modelo para la realidad?

Para contestar primero esta última pregunta, es muy poco probable que alguna vez seleccionemos una función de densidad que proporcione una representación perfecta de la naturaleza; pero la bondad de ajuste no es el criterio para evaluar lo adecuado de nuestro modelo. La finalidad de un modelo probabilístico es proveer el mecanismo para hacer inferencias acerca de una población con base en información contenida en una muestra. La probabilidad de la muestra observada (o una cantidad proporcional a ella) es útil para hacer una inferencia acerca de la población. Se deduce que una función de densidad que proporcione un mal ajuste a la distribución de frecuencia poblacional (pero no necesariamente) da como resultado enunciados incorrectos de probabilidad y lleva a inferencias erróneas acerca de la población. Un buen modelo es aquel que produzca buenas inferencias acerca de la población de interés.

Seleccionar un modelo razonable es a veces cuestión de actuar con base en consideraciones teóricas. Con frecuencia, por ejemplo, una situación en la que la variable aleatoria discreta de Poisson es apropiada, queda indicada por el comportamiento aleatorio de eventos en el tiempo. Sabiendo esto, podemos demostrar que el tiempo entre cualquier par adyacente de eventos sigue una distribución exponencial. Del mismo modo, si a y b son enteros, $a < b$, entonces el tiempo entre los sucesos de los eventos a-ésimo y b-ésimo posee una distribución gamma con $\alpha = b - a$. Más adelante encontraremos un teorema (llamado *teorema central del límite*) que compendia algunas condiciones que implican que una distribución normal sería una aproximación apropiada para la distribución de datos.

Una segunda forma de seleccionar un modelo es formar un histograma de frecuencia (Capítulo 1) para datos sacados de la población y escoger una función de densidad que visualmente parecería dar una curva de frecuencia similar. Por ejemplo, si un conjunto de $n = 100$ mediciones muestrales diera una distribución de frecuencia en forma de campana, podríamos concluir que la función de densidad normal representaría en forma adecuada la distribución de frecuencia poblacional.

No toda la selección de un modelo es completamente subjetiva. Hay procedimientos estadísticos para probar la hipótesis de que una distribución de frecuencia poblacional es de un tipo particular. También podemos calcular una medida de la bondad de ajuste para varias distribuciones y seleccionar la mejor. Se han hecho estudios de numerosos métodos inferenciales comunes para determinar la magnitud de los errores de inferencia introducidos por modelos incorrectos de población. Es alentador saber que muchos métodos estadísticos de inferencia son insensibles a suposiciones acerca de la forma de la distribución de frecuencia poblacional subyacente.