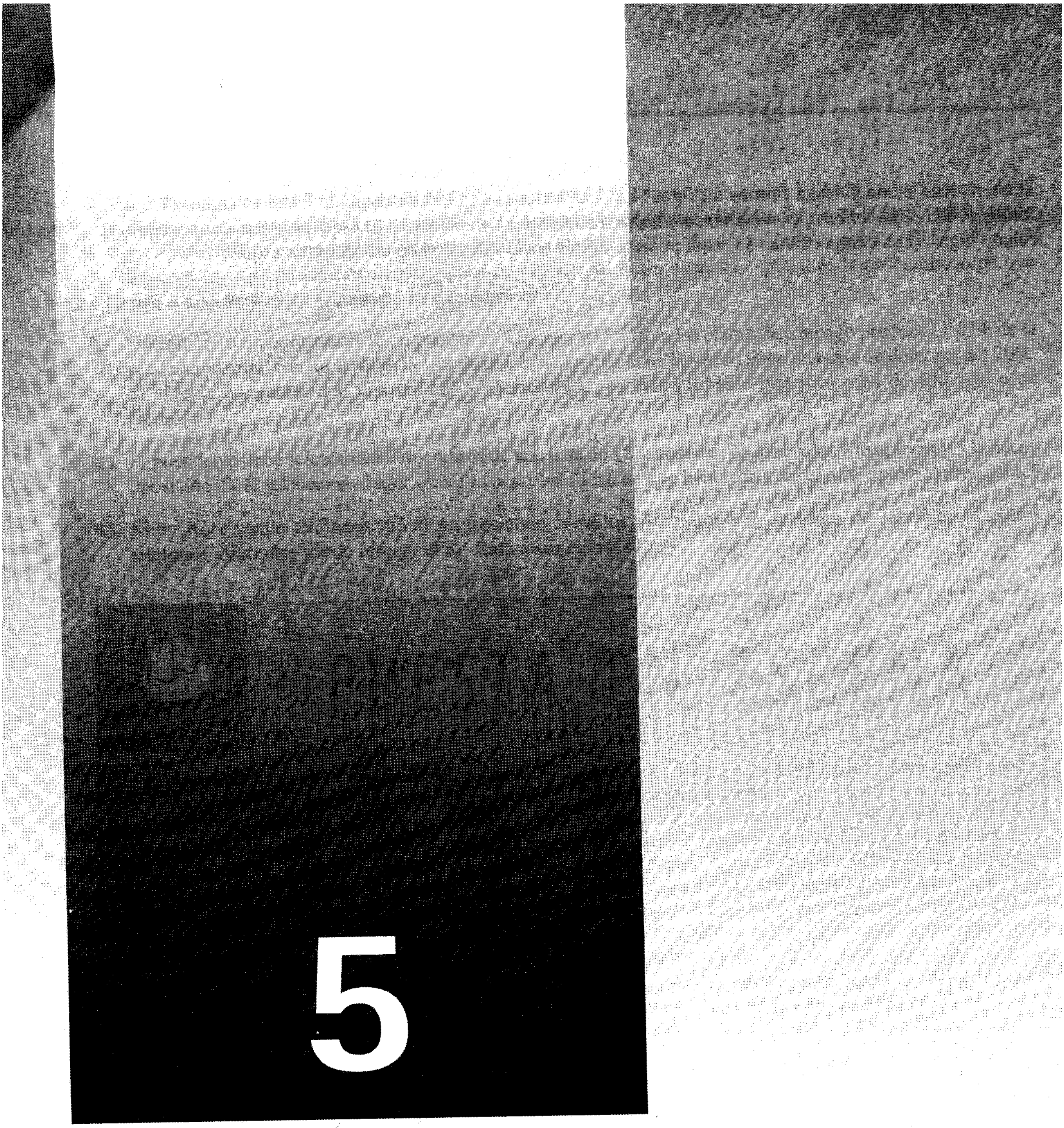
Estadística aplicada a los negocios y la economía

Tercera edición

- Incluye 850 ejercicios
- Disquete de datos
- Ejercicios y proyectos por internet

Allen L. Webster

Irwin
McGraw-Hill

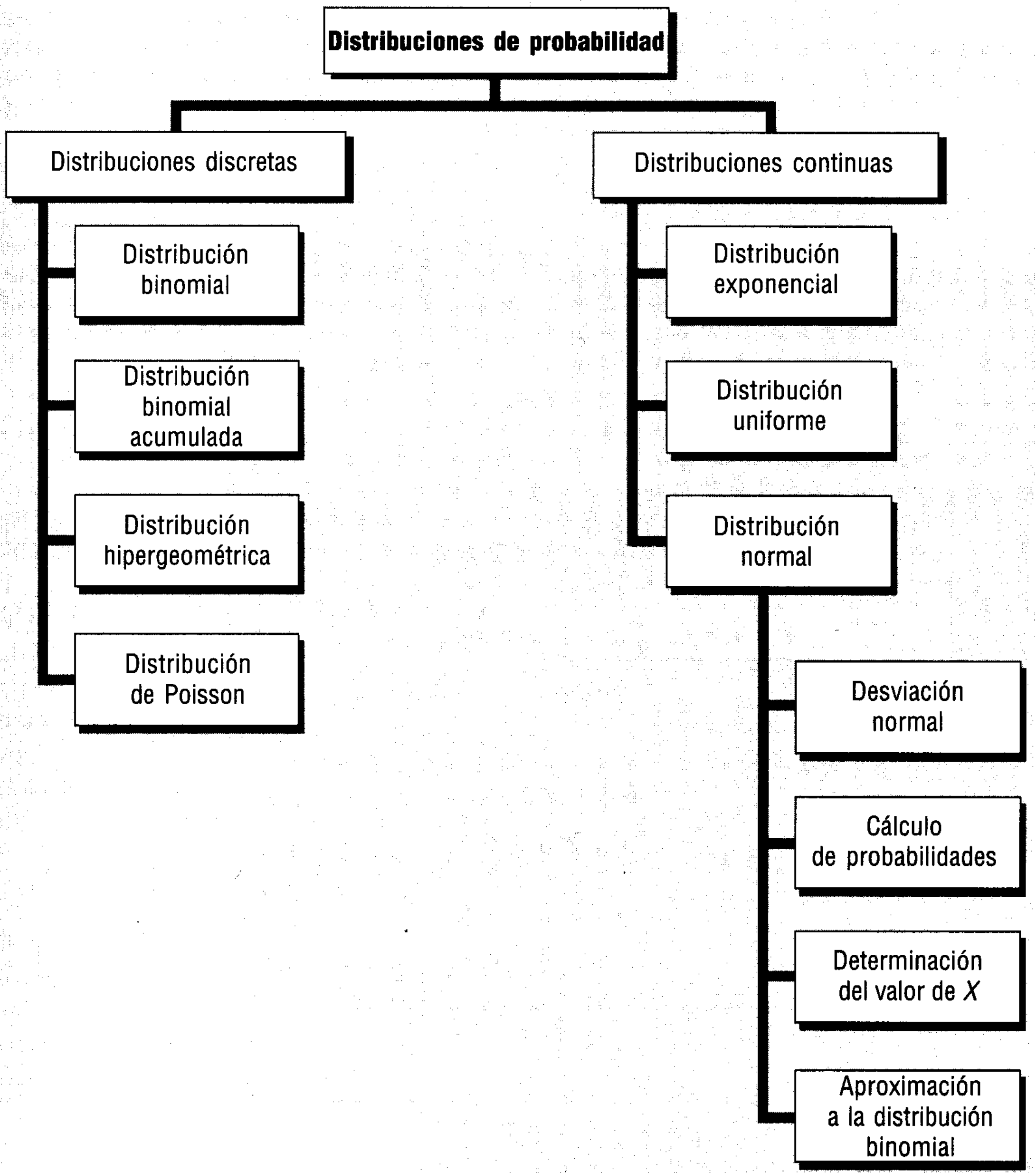


5

Distribuciones de probabilidad

Plan del capítulo

Este capítulo analiza cómo pueden utilizarse las distribuciones de probabilidad para solucionar muchos problemas de negocios. En las ilustraciones se utilizan variables tanto discretas como continuas.

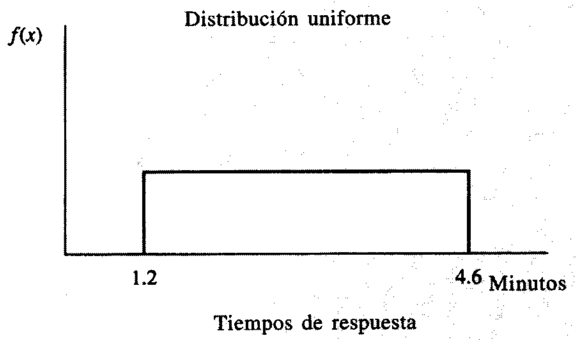


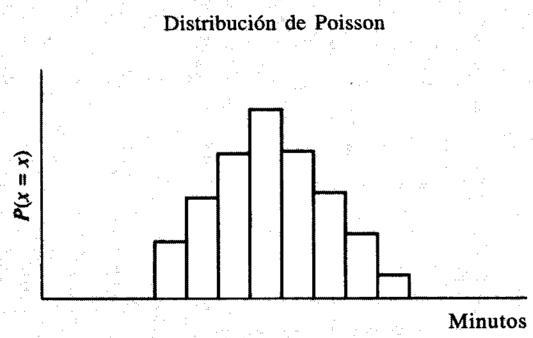
ESCENARIO

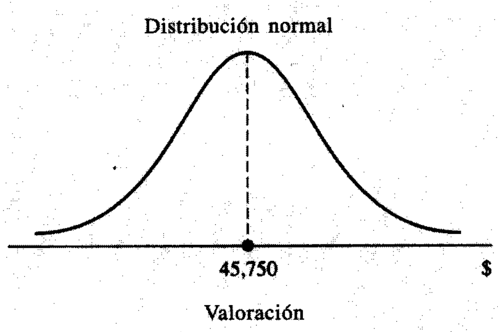
Un estudio de factibilidad realizado por profesores de la Escuela de Negocios de la Universidad de Bradley (*College of Business at Bradley University*) en Peoria, Illinois, reveló que los viernes y sábados en la noche el tiempo de respuesta a las llamadas al número 911 oscilaba entre 1.2 y 4.6 minutos y se comprobó que estaban distribuidas uniformemente. Las llamadas tenían la distribución de Poisson y alcanzaron una tasa promedio de 9 por hora. Si la policía de la ciudad tuviera que responder más de 3 llamadas, en cualquier momento podrían acudir a la policía estatal para solicitarles ayuda.

El alcalde de la ciudad deseaba reducir el tiempo promedio de respuesta a 2 minutos. Se estimaba que el costo de más patrullas, máquinas de bomberos y personal sería de US$575,000 por cada reducción de 30 segundos. El costo debía ser asumido por un impuesto predial a las casas cuyo valor en el avalúo catastral estuviera por encima de US$70,000. Las casas en Peoria tienen un avalúo catastral promedio de US$45,750 con una desviación estándar de US$15,110 y parecía estar distribuido normalmente. En el momento del estudio había 42,089 casas en los límites de la ciudad.

El estudio presentado a la alcaldía estaba diseñado para evaluar la respuesta de la ciudad a las emergencias, así como también la factibilidad de lograr la meta de reducción en el tiempo de respuesta que quería el alcalde. El reporte final necesitó de la aplicación de numerosas distribuciones de probabilidad así como de una evaluación del potencial de promulgar un sobrecargo al impuesto predial para financiar las mejoras del programa.







5.1 Introducción

En el capítulo anterior se analizó el concepto de probabilidad. El objetivo era determinar la probabilidad de un evento. En este capítulo definiremos las variables aleatorias y utilizaremos las leyes de probabilidad. Una **variable**

aleatoria es una variable cuyo valor es el resultado de un evento aleatorio. Se supone que se lanza una moneda tres veces y se anota el número de caras que se obtienen. Los posibles resultados son 0 caras, 1 cara, 2 caras, o 3 caras. La variable aleatoria es el número de caras que se obtienen, y los posibles resultados son los valores de la variable aleatoria. Como segundo ejemplo, los pesos de envío del agua mineral en contenedores oscilaban aleatoriamente entre 10 a 25 libras. Los pesos reales de los contenedores, en libras, son los valores de la variable aleatoria "peso".

Tal y como lo sugieren estos dos ejemplos, las variables aleatorias pueden ser discretas o continuas. Una **variable aleatoria discreta** puede asumir sólo ciertos valores, con frecuencia números enteros, y resulta principalmente del conteo. El número de caras en el experimento del lanzamiento de la moneda es un ejemplo de una variable aleatoria discreta. Los valores de la variable aleatoria se restringen sólo a ciertos números: 0, 1, 2, y 3. El resultado del lanzamiento de un dado, el número de camiones que llegan por hora al puerto de carga, y el número de clientes que están en fila para sacar sus libros favoritos, son otros ejemplos de variables aleatorias discretas.

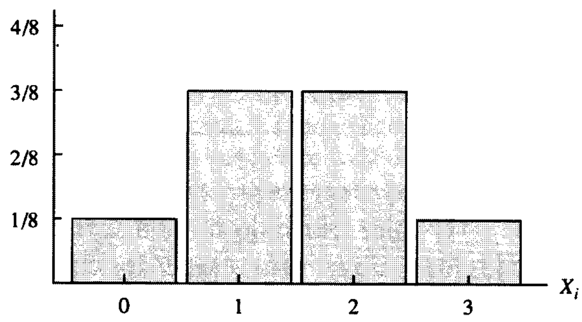
Una **variable aleatoria continua** resulta principalmente de la medición y puede tomar cualquier valor, al menos dentro de un rango dado. Los pesos del agua mineral es un ejemplo, debido a que los contenedores pueden tomar cualquier valor entre 10 y 25 libras. Otros ejemplos de variables aleatorias continuas incluyen la estatura de los clientes en una tienda de ropa, los ingresos de los empleados en un centro comercial local y el tiempo transcurrido entre la llegada de cada cliente a la biblioteca. En cada caso, la variable aleatoria puede medirse con cualquier valor, incluyendo fracciones de la unidad. Aunque las unidades monetarias no pueden dividirse en un número continuo o infinito de subdivisiones (el dólar puede subdividirse sólo 100 veces), comúnmente se tratan como distribuciones continuas de probabilidad.

Una **distribución de probabilidad** es un despliegue de todos los posibles resultados de un experimento junto con las probabilidades de cada resultado. Del trabajo realizado en el capítulo 4, se puede determinar que la probabilidad de lanzar una moneda tres veces y de obtener (1) ninguna cara es 1/8, (2) 1 cara es 3/8, (3) 2 caras es 3/8 y (4) 3 caras es 1/8. Esta distribución de probabilidad se presenta en la tabla 5.1 la cual muestra todos los resultados posibles y sus probabilidades. Vale la pena destacar que las probabilidades suman 1. La misma información también puede mostrarse gráficamente como en la figura 5.1

Tabla 5.1
Distribución discreta de probabilidad para el número de caras

Resultado (caras)	Probabilidad
0	1/8
1	3/8
2	3/8
3	1/8
	1

Figura 5.1
Distribución de probabilidad para el número de caras



Distribución de probabilidad Es una lista de todos los resultados posibles de algún experimento y de la probabilidad relacionada con cada resultado.

La probabilidad de que la variable aleatoria X tome algún valor específico, x_i, se escribe $P(X = x_i)$. Por tanto, la probabilidad de que los tres lanzamientos de una moneda resulten en dos caras es $P(X = 2) = 3/8$. Vale la pena notar que $0 \leq P(X = x_i) \leq 1$ y $\sum P(X = x_i) = 1$.

5.2 Media y la varianza de las distribuciones discretas

Así como en el capítulo 3 se calculó la media de un conjunto de datos, también se puede determinar la media de una distribución de probabilidad. La media aritmética de una distribución de probabilidad se llama el **valor esperado** *E(X)*, y se halla multiplicando cada resultado posible por su probabilidad y sumando los resultados, tal y como se muestra en la fórmula (5.1).

Media o valor esperado de una distribución de probabilidad discreta

$$\mu = E(X) = \Sigma[(x_i)P(x_i)] \qquad [5.1]$$

en donde x_i son los resultados individuales.

La distribución de probabilidad para el experimento de lanzar un dado se muestra en las primeras dos columnas de la tabla 5.2. La columna (3) ilustra el cálculo del valor esperado para el experimento utilizando la fórmula (5.1). Cada resultado se multiplica por su respectiva probabilidad, y los resultados se suman, produciendo $\mu = E(X) = 3.5$. Esto sugiere que si se lanza un dado ¿se puede esperar obtener 3.5? Difícilmente. Significa que si se promedian los resultados de los lanzamientos del dado (teóricamente, un número infinito), se obtendrá 3.5.

Tabla 5.2
Distribución discreta de probabilidad para el lanzamiento de un dado

(1) Solución (x_i)	(2) $P(x_i)$	(3) $(x_i)\cdot P(x_i)$	(4) $(x_i-\mu)^2\cdot P(x_i)$
1	1/6	1/6	$(1-3.5)^2\cdot 1/6 = 1.042$
2	1/6	2/6	$(2-3.5)^2\cdot 1/6 = 0.375$
3	1/6	3/6	$(3-3.5)^2\cdot 1/6 = 0.042$
4	1/6	4/6	$(4-3.5)^2\cdot 1/6 = 0.042$
5	1/6	5/6	$(5-3.5)^2\cdot 1/6 = 0.375$
6	1/6	6/6	$(6-3.5)^2\cdot 1/6 = 1.042$
	1.00	$3.5 = \mu = E(X)$	$2.92 = \sigma^2$

Valor esperado El valor esperado de una variable aleatoria discreta es la media ponderada de todos los posibles resultados en los cuales los pesos son las probabilidades respectivas de tales resultados.

La varianza de una distribución de probabilidad es conceptualmente la misma que la varianza que se calculó en el capítulo 3. Es el promedio de las desviaciones al cuadrado con respecto de la media. La varianza puede escribirse como:

Varianza de una distribución de probabilidad	$\sigma^2 = \Sigma[(x_i - \mu)^2 P(x_i)]$	[5.2]

La fórmula (5.2) mide la diferencia entre cada uno de los resultados y su media. Tales diferencias se elevan al cuadrado y se multiplican por sus respectivas probabilidades. Luego se suman los resultados. La columna (4) de la tabla 5.2 revela, $\sigma^2 = 2.92$.

La desviación estándar es $\sigma = \sqrt{\sigma^2} = \sqrt{2.92} = 1.71$. La varianza y la desviación estándar tienen la misma interpretación que se les dio en el capítulo 3. Miden la dispersión de los resultados alrededor de su media. La varianza se expresa en unidades al cuadrado, pero la desviación estándar se expresa en las mismas unidades que la variable aleatoria y por ende con frecuencia tiene una interpretación más racional.

Ejemplo 5.1

El número de casas que Ponder Real Estate vendió mensualmente varió de 5 a 20 junto con la frecuencia de cada nivel de ventas que aparece en las dos primeras columnas de la tabla que se muestra a continuación.

(1) Número de meses	(2) Casas (x_i)	(3) $P(x_i)$	(4) $(x_i)\,P(x_i)$	(5) $(x_i-\mu)^2\,P(x_i)$
3	5	3/24 = 0.125	0.625	$(5-10.912)^2\,(0.125) = 4.369$
7	8	7/24 = 0.292	2.336	$(8-10.912)^2\,(0.292) = 2.476$
4	10	4/24 = 0.167	1.670	$(10-10.912)^2\,(0.167) = 0.139$
5	12	5/24 = 0.208	2.496	$(12-10.912)^2\,(0.208) = 0.246$
3	17	3/24 = 0.125	2.125	$(17-10.912)^2\,(0.125) = 4.633$
2	20	2/24 = 0.083	1.660	$(20-10.912)^2\,(0.083) = 6.855$
24		1.000	$10.912 = \mu$	$18.718 = \sigma^2$

El Sr. Ponder espera que estas cifras reflejen un incremento en el número promedio de ventas, por encima del 7.3 que vendió en meses anteriores, y una reducción en la variabilidad de las ventas mensuales que habían sido de $\sigma = 5.7$. De lo contrario, él ha decidido vender el negocio y convertirse en un bufón de rodeo. ¿Qué consejo puede ofrecerle al Sr. Ponder?

Solución

Se debe determinar la probabilidad de cada nivel de ventas tal y como se muestra en la columna (3). Por ejemplo, en 3 de los 24 meses, se vendieron 5 casas: $P(x_i = 5) = 0.125$. El valor esperado o media se calcula multiplicando las probabilidades por sus respectivos niveles de venta. Esto se muestra en la columna (4) como $\mu = 10.912$ casas por mes. La variabilidad se mide por la varianza y se muestra en la última columna. La diferencia al cuadrado entre cada observación y la media de 10.912 se multiplica por las probabilidades apropiadas y se suman resultando $\sigma^2 = 18.718$ casas al cuadrado, con $\sigma = 4.236$ casas.

Interpretación

El Sr. Ponder puede tranquilizarse. Ha incrementado su promedio mensual de ventas y ha reducido su variabilidad. Debería quedarse en el negocio de finca raíz.

Ejercicios de la sección

1. Dé varios ejemplos tanto de distribuciones discretas de probabilidad como de distribuciones continuas de probabilidad que pueden aparecer comúnmente en un negocio. ¿Cuál es la diferencia entre una distribución discreta de probabilidad y una continua?

2. Las siguientes variables aleatorias ¿son discretas o continuas? En cada caso explique el porqué de su respuesta.
   a. Los carros vendidos por Harry el honesto.
   b. Los ingresos que gana Harry.
   c. Los tiempos de terminación de un trabajo en particular.
   d. Los empleados requeridos para completar dicho trabajo.

3. Calcule e interprete el valor esperado, la varianza, y la desviación estándar del experimento de lanzar una moneda tres veces y observe el número de caras.

4. El número de quejas de los empleados en Fidelity Services oscila entre 0 a 6 cada día como se muestra en la siguiente tabla. Calcule e interprete el valor esperado, la varianza y la desviación estándar.

Quejas	Número de días	Quejas	Número de días
0	3	4	2
1	4	5	1
2	3	6	4
3	6		

5. Para recolectar los datos de un proyecto de investigación, un estudiante de mercadeo en una universidad pequeña en el centro de Estados Unidos contó en 50 cursos de negocios el número de estudiantes que habían comprado recientemente discos compactos. En 12 clases no encontró estudiantes que hubieran hecho dicha compra, 3 estudiantes habían comprado en 8 clases, 4 habían comprado en 9 clases, 5 en 15 clases y 7 estudiantes, de las seis clases restantes habían aumentado sus colecciones de música. El estudiante deseaba comenzar su investigación resumiendo sus datos. ¿Cómo podría usted ayudarle?

Un gran número de decisiones empresariales depende de la distribución de probabilidad prevaleciente. Una de las más importantes es la distribución binomial.

5.3 La distribución binomial – una distribución discreta de probabilidad

El experimento de lanzar la moneda discutido anteriormente tiene sólo dos posibles resultados: (1) cara y (2) sello. La probabilidad de cada uno es conocida y constante de un intento (lanzamiento) al siguiente, y además el experimento puede repetirse muchas veces. Los experimentos de este tipo siguen una **distribución binomial**. Con base en el proceso de Bernoulli, llamado así por Jacob Bernoulli (1654-1705), miembro de una familia de matemáticos suizos, una distribución normal presenta cuatro propiedades:

- Sólo debe haber dos posibles resultados. Uno se identifica como éxito, y el otro como fracaso. Sin embargo, se advierte que estos términos no tienen ninguna connotación de "bueno" o "malo". Son completamente objetivos, y un "éxito" no implica necesariamente un resultado deseable.

- La probabilidad de un éxito, π, sigue siendo constante de un ensayo al siguiente, al igual que lo hace la probabilidad de fracaso, $1 - \pi$.
- La probabilidad de un éxito en un ensayo es totalmente independiente de cualquier otro ensayo.
- El experimento puede repetirse muchas veces.

Debe haber quedado claro por qué el lanzamiento de la moneda cumple con los requisitos de una distribución binomial.

Podrían citarse muchos ejemplos relacionados con los negocios. Los sindicatos laborales con frecuencia desean saber cuántos trabajadores: (1) están interesados en unirse al sindicato; (2) quienes no están interesados. Los banqueros pueden hacer encuestas a los expertos en economía sobre si las tasas de interés: (1) aumentarán o (2) no aumentarán. El personal de mercadeo desea saber si una persona: (1) prefiere o (2) no prefiere cierto producto. La aplicación de la distribución binomial al campo de los negocios es casi ilimitada.

Una distribución binomial Cada ensayo en una distribución binomial termina en sólo uno de dos resultados mutuamente excluyentes, uno de los cuales se identifica como un éxito y el otro como un fracaso. La probabilidad de cada resultado permanece constante de un ensayo al siguiente.

Si se conoce la probabilidad de que un ensayo determinado producirá un éxito, es posible estimar cuántos éxitos habrá en un número dado de ensayos. Por ejemplo, si se conoce la probabilidad de que un solo trabajador esté interesado en unirse al sindicato, entonces puede estimarse la probabilidad de que un número determinado de trabajadores de la fuerza laboral estaría interesado en unirse. La probabilidad de que de n número de trabajadores, un número x dado, esté interesado en unirse al sindicato es

La fórmula binomial

$$P(x) = \frac{n!}{x!(n-x)!}\pi^x(1-\pi)^{n-x} = {}_nC_x(\pi)^x(1-\pi)^{n-x} \qquad [5.3]$$

Aunque a simple vista la fórmula parece intimidante, no se desespere. Las probabilidades para diferentes valores de π, x y n ya se han calculado y se han tabulado en el apéndice III, tabla B, al final del libro.

Consideremos la siguiente situación: un gerente de crédito de *American Express* ha descubierto que $\pi = 10\%$ de los usuarios de tarjeta no paga el monto completo de la deuda durante un mes dado. Desea determinar la probabilidad que de $n = 20$ cuentas seleccionadas de manera aleatoria, $x = 5$ de las cuentas no sean pagadas. Esto puede expresarse como $P(X = 5 \mid n = 20, \pi = 0.10)$, lo cual se lee como "la probabilidad de cinco éxitos dado que hay 20 ensayos y la probabilidad de un éxito de cualquier ensayo es del 10%".

La probabilidad de que 5 cuentas de las 20 muestreadas sigan sin ser canceladas se puede calcular utilizando la fórmula (5.3). En donde $n = 20$, $X = 5$, y $\pi = 0.10$, entonces se tiene

$${}_{20}C_5(0.10)^5(0.90)^{20-5} = (15504)(0.00001)(0.2058911) = 0.0319$$

Si la probabilidad de que no se pague una cuenta cualquiera en su totalidad es $\pi = 0.10$, entonces existe un 3.19% de oportunidad de que exactamente 5 de 20 cuentas seleccionadas de manera aleatoria tengan un saldo a favor.

Esta información se obtiene más fácilmente utilizando la tabla B. Vale la pena destacar que las dos primeras columnas en la tabla muestran los posibles valores para n y x. Ubique el valor de 20 para n debido a que hay 20 ensayos (cuentas) en el experimento. Debido a que el gerente de crédito busca la probabilidad de que $x = 5$ éxitos (cuentas no pagadas), localice la fila que contenga los valores de probabilidad para $x = 5$. Proceda a lo largo de la fila hasta encontrar la columna encabezada con $\pi = 0.10$. Allí se encontrará el valor 0.0319, la respuesta a la pregunta del gerente de crédito.

Consideremos otro ejemplo de distribución binomial. Se tiene que el personal de ventas de Widgets, Inc., hace una venta al 15% de los clientes a los que visitan. Si un miembro del personal de ventas llama a 15 clientes hoy ¿cuál es la probabilidad de que venda exactamente dos aparatos? Dado $\pi = 0.15$, $n = 15$ y $x = 2$, ubique el valor para $n = 15$, luego la fila que pertenezca a $X = 2$. En la fila encabezada por la columna $\pi = 0.15$, encontrará $P(x = 2 \mid n = 15, \pi = 0.15) = 0.2856$. Existe un 28.56% de oportunidad de que se hagan exactamente dos ventas de las 15 llamadas.

Ejemplo 5.2

De acuerdo con el Periódico de Educación Superior (*Journal of Higher Education*), el 40% de todos los bachilleres trabajan durante el verano para ganar dinero para la educación universitaria correspondiente al siguiente período de otoño. Si 7 bachilleres se seleccionan de manera aleatoria, ¿cuál es la probabilidad de que (*a*) 5 tengan trabajos en el verano, (*b*) ninguno trabaje, (*c*) todos trabajen?

Solución

a. Ubique el valor de $n = 7$ y $\pi = 0.40$. La fila correspondiente a $x = 5$ da un valor de 0.0774. Existe un 7.74% de probabilidad de que 5 de 7 bachilleres hayan tomado trabajos de verano para ganar el dinero para su educación.

b. Dado $n = 7$ y $\pi = 0.40$, la probabilidad de que ninguno trabaje se muestra en la tabla como $P(x = 0) = 0.0280$.

c. La probabilidad de que todos los estudiantes trabajen parece ser $P(x = 7 \mid, \pi = 0.4) = 0.0016$.

Interpretación

Es poco probable que ninguno de los estudiantes trabaje.

La tabla binomial incluye los valores de π sólo hasta 0.5. ¿Qué debe hacerse si la probabilidad de un éxito es mayor? Supongamos que el 70% de todos los residentes de Flatbush tiene sus computadores enlazados con internet. ¿Cuál es la probabilidad de que de los 10 residentes seleccionados aleatoriamente, 6 estén "conectados"? ya que $\pi > 0.5$, no se puede utilizar la tabla B (Apéndice III) sin algún ajuste. Sin embargo, si la probabilidad de un éxito (un residente esté conectado a Internet) es $P(S) = 0.70$, la probabilidad de que no esté enlazado es $P(\bar{S}) = 0.30$. Además, si 6 de los 10 residentes son usuarios de internet, entonces 4 no lo son. Es decir, 6 éxitos a $\pi = 0.70$ es *lo mismo* que 4 fracasos a $\pi = 0.30$. En lugar de hallar x éxitos en π, se halla $n - x$ fracasos a $1.00 - \pi$.

Esta práctica puede ilustrarse construyendo dos arreglos ordenados como los que se ven aquí, uno de 0 a 10 a $\pi = 0.70$ y uno de 10 a $\pi = 0.30$.

0	1	2	3	4	5	**6**	7	8	9	10	($\pi = 0.70$)
10	9	8	7	6	5	**4**	3	2	1	0	($\pi = 0.30$)

Esto revela más claramente que $P(X = \mathbf{6} \mid n = 10, \pi = \mathbf{0.70}) = P(X = \mathbf{4} \mid n = 10, \pi = \mathbf{0.30})$. De la tabla B (Apéndice III) esto se ve como 0.2001.

A. La media y la varianza de una distribución binomial

Antes se mostró cómo determinar la media y la varianza de una distribución discreta utilizando las fórmulas (5.1) y (5.2). Sin embargo, si sólo hay dos resultados posibles, como en la distribución binomial, la media y la varianza pueden determinarse más fácilmente:

Media de una distribución binomial	$E(X) = \mu = n\pi$	[5.4]

y

Varianza de una distribución binomial	$\sigma^2 = n\pi(1-\pi)$	[5.5]

Para los residentes de Flatbush, si $n = 10$, $E(X) = (10)(0.70) = 7$. De las 10 personas seleccionadas aleatoriamente, se esperaría que 7 estuvieran inscritas en internet. La varianza es $\sigma^2 = (10)(0.70)(0.30) = 2.1$ y la desviación estándar es $\sigma = 1.45$.

B. Distribuciones binomiales acumuladas

Dados los datos del ejemplo 5.2 para los trabajos de verano de los estudiantes, se supone que se desea determinar la probabilidad de que 3 o menos estudiantes trabajaron. Este problema implica una distribución binomial *acumulada* debido a que se está interesado en un *rango* de valores (0 a 3) en lugar de un solo número específico. El siguiente arreglo ordenado ilustra este punto. La probabilidad del evento A (0 a 3 trabajan) es $P(A) = P(X) \leq 3)$.

$$\overbrace{0 \quad 1 \quad 2 \quad 3}^{\text{Evento A}} \quad 4 \quad 5 \quad 6 \quad 7 \qquad (\pi = 0.40)$$

En la tabla B (apéndice III), esto puede hallarse sumando $P(X=0) + P(X=1) + P(X=3) + P(X=3) = 0.7102$. Por motivos de conveniencia, estas sumas se compilan en la tabla C, la cual muestra la probabilidad del número de éxitos que es igual a *o menor que* cierta cantidad. En nuestro caso actual, se tiene que $P(X \leq 3 \mid n = 7, \pi = 0.40) = 0.7102$.

Vale la pena recordar que la tabla C proporciona la probabilidad de que el número de éxitos sea igual a o menor que cierta cantidad. Se supone que se desea conocer $P(A) = P(X \geq 5)$. La tabla C no dará directamente la probabilidad de que un número de éxitos sea igual a *o mayor que* alguna cantidad. Observando el arreglo ordenado se tiene que

$$\underbrace{0 \quad 1 \quad 2 \quad 3 \quad 4}_{\text{Evento } \bar{A}} \quad \overbrace{5 \quad 6 \quad 7}^{\text{Evento A}} \qquad (\pi = 0.40)$$

Si el evento A es $P(X \geq 5)$, entonces $\bar{A}$ es 4 o menos, lo cual puede hallarse en la tabla C. Se sabe que $P(A) = 1 - P(\bar{A})$. Entonces, $P(X \geq 5 \mid n = 7, \pi = 0.40) = 1 - P(X \leq 4 \mid n = 7, \pi = 0.40)$. De la tabla C, se observa que este es $1 - 0.9037 = 0.0963$. La probabilidad de que por lo menos 5 de 7 estudiantes tengan trabajo en verano es del 9.63%.

Se supone que se necesitaba determinar la probabilidad de que entre 3 y 5 estudiantes inclusive, trabajaron. De nuevo el arreglo prueba que es de utilidad.

$$\underbrace{0 \quad 1 \quad 2}_{P(X \leq 2) = 0.4199} \quad \overbrace{3 \quad 4 \quad 5}^{\text{Evento A}} \quad 6 \quad 7 \qquad (\pi = 0.40)$$

$P(X \leq 5) = 0.9812$

$P(3 \leq X \leq 5) = P(X \leq 5) - P(X \leq 2) = 0.9812 - 0.4199 = 0.5613$

$P(3 \leq X \leq 5 \mid n = 7, \pi = 0.40)$ debe determinarse en dos pasos. Primero se determina la probabilidad de que el número de estudiantes con trabajos sea de 0 a 5 (lo que incluye el intervalo de 3 a 5 que es el que se busca), y luego se resta la probabilidad de que el número de estudiantes emprendedores sea 2 o menos. Luego $P(3 \leq X \leq 5) = P(0 \leq X \leq 5) - P(0 \leq X \leq 2) = 0.9812 - 0.4199 = 0.5613$.

Si $\pi > 0.50$, se necesitan dos arreglos ordenados. Se asume que el 80% de los graduados tomaron trabajos de verano. Debe construirse un arreglo para $\pi = 0.80$ y uno para $\pi = 0.20$.

Evento A								
0	1	2	3	4	5	6	7	(π = 0.80)
7	6	5	4	3	2	1	0	(π = 0.20)
Evento A				Evento $\bar{A}$				

Si, al igual que antes, se desea la probabilidad de que 3 o menos estudiantes trabajen, se debe hallar $P(A) = P(X \leq 3 \mid n = 7, \pi = 0.80)$. Debido a que la tabla C no contiene valores para $\pi > 0.50$, se debe regresar al arreglo ordenado para $\pi = 1 - 0.80 = 0.20$. Se observa que la probabilidad 3 o menos trabajen con $\pi = 0.80$ es la *misma* que la probabilidad de que 4 o más no trabajen. Es decir, si 3 de 7 trabajan, 4 no; si 2 de 7 trabajan, 5 no, y así sucesivamente. Así, $P(A)$ también es igual a $P(X \geq 4 \mid n = 7, \pi = 0.20)$. Sin embargo, todavía se tiene un problema. La tabla C no dará directamente la probabilidad de que X sea igual a o mayor que algún valor, como 4 en este caso. La solución a este dilema persistente es la misma que la anterior: se halla la probabilidad de A, es decir, 3 o menos, y se resta de 1. $P(X \leq 3 \mid n = 7, \pi = 0.80) = 1 - P(X \leq 3 \mid n = 7, \pi = 0.20) = 1 - 0.9667. = 0.0333$.

C. Usando el computador

Tanto Minitab como Excel calculan fácilmente las probabilidades binomiales. Al utilizar Minitab para determinar las probabilidades de los trabajos de verano del ejemplo 5.2, ingrese los valores para X en las celdas de la columna (1) (o en cualquier otra columna). En este caso, ingrese 5, 0, y 7 en las primeras tres celdas de la columna (1). Luego escoja **Calc > Probability Distributions > Binomial > Probability**. Ingrese el número de ensayos, 7 en este caso, en la casilla de **Number of Trials**, 0.40 en la casilla de **Probability of a Success**, y *C1* (en donde se ingresó anteriormente los valores para X) en la casilla de **Input Column**. La impresión resultante se ve en la pantalla 5.1. Si se ha seleccionado **Cumulative Probability** en lugar de **Probability** antes mencionada, Minitab habrá devuelto las probabilidades acumuladas tal y como aparecen en la tabla C.

Pantalla de Minitab 5.1

Probability density function (Función de densidad de probabilidad)

```
Binomial con n = 7 y p = 0.400000
         X    P(X=x)
      5.00    0.0774
      0.00    0.0280
      7.00    0.0016
MTB >
```

Excel funciona de forma similar. Se coloca el cursor en la celda de la hoja de trabajo en donde se desea que aparezca la respuesta. Luego se selecciona **Insertar** > **Función** > **Estadísticas** (de la casilla de categoría de funciones) > **Distr.Binom** (de la casilla nombre de función). Se hace clic en **Aceptar**. Se ingresa 5 en la casilla de **Núm-éxito** (para 5 éxitos), 7 en la casilla de **Ensayos**, 0.4 en la casilla de **Prob-éxito**, y Falso en la casilla de **Acumulado**. Seleccione **Aceptar**. La respuesta aparecerá en la casilla de **Valor** en la esquina superior derecha y en la celda que se haya designado en la hoja de trabajo. Si se ha ingresado Verdadero en la casilla de **Acumulado**, la probabilidad acumulada para 5 éxitos se reportará tal y como aparece en la tabla C.

Ejercicios de la sección

6. ¿Cuáles son las cuatro características de una distribución binomial? Dé por lo menos tres ejemplos relacionados con negocios.
7. El 10% de los discos de computador producidos por un nuevo proceso salen defectuosos. Si hay 20 discos en una caja:
   a. ¿Cuántos esperaría usted que salieran defectuosos?
   b. ¿Cuál es la probabilidad de que el número de discos defectuosos sea igual al número esperado que usted determinó en su respuesta a la parte *a*?
8. Del problema anterior, ¿cuál variación se encontraría en los discos defectuosos de una caja a otra?
9. Sólo 20% de los empleados de la población civil que está en una base militar restringida porta su identificación personal. Si llegan 10 empleados, cuál es la probabilidad de que el guardia de seguridad encuentre:
   a. ¿Ocho empleados con identificación?
   b. ¿Cuatro empleados con identificación?
   c. ¿Por lo menos 4 empleados con identificación?
   d. ¿A lo sumo 5 empleados con identificación?
   e. ¿Entre 4 y 7 empleados inclusive con identificación?
10. Responda la pregunta anterior si 60% de todos los empleados portan identificación.
11. Usted ha contratado 8 recepcionistas telefónicas para que tomen los pedidos telefónicos para una línea de productos deportivos que su empresa está comercializando. Una recepcionista está ocupada el 30% del tiempo catalogando un pedido. Usted no desea que la probabilidad de que una llamada del cliente se reciba con una señal de ocupado exceda del 50%. ¿Debería usted contratar más recepcionistas si 3 clientes llaman?
12. Un estudiante debe obtener por lo menos el 60% en un examen de verdadero y falso con 18 preguntas por responder. Si el estudiante lanza una moneda para determinar la respuesta a cada pregunta, ¿cuál es la probabilidad de que el estudiante pase?

5.4 La distribución hipergeométrica

Como se acaba de explicar, la distribución binomial es apropiada sólo si la probabilidad de un éxito permanece constante para cada intento. Esto ocurre si el muestreo se realiza con reemplazo o de una población finita (o muy grande). Sin embargo, si la población es pequeña y ocurre el muestreo sin reemplazo, la probabilidad de un éxito variará. Si la probabilidad de un éxito no es constante, la **distribución hipergeométrica** es de especial utilidad. La función de probabilidad para la distribución hipergeométrica es:

Distribución hipergeométrica

$$P(x) = \frac{{}_{r}C_{x}\ {}_{N-r}C_{n-x}}{{}_{N}C_{n}} \qquad [5.6]$$

en donde N es el tamaño de la población
r es el número de éxitos en la población
n es el tamaño de la muestra
x es el número de éxitos en la muestra

La distribución hipergeométrica Si se selecciona una muestra sin reemplazo de una población finita conocida y contiene una proporción relativamente grande de la población, de manera que la probabilidad de éxito sea perceptiblemente alterada de una selección a la siguiente, debe utilizarse la distribución hipergeométrica.

Supongamos que en un establo de caballos de carrera hay $N = 10$ caballos, y $r = 4$ de ellos tienen una enfermedad contagiosa. ¿Cuál es la probabilidad de seleccionar una muestra de $n = 3$ en la cual $x = 2$ caballos enfermos?

$$\begin{aligned} P(X = 2) &= \frac{{}_4C_2 \; {}_{10-4}C_{3-2}}{{}_{10}C_3} \\ &= \frac{6 \times 6}{120} \\ &= 0.30 \end{aligned}$$

Existe un 30% de probabilidad de seleccionar tres caballos de carreras, dos de los cuales están enfermos.

Ejemplo 5.3 Uso de la distribución hipergeométrica para analizar la discriminación

En un caso reciente en el Distrito de Johnson en Kansas City, tres mujeres entablaron una demanda contra una empresa de servicios locales, por discriminación de sexos. De las nueve personas que eran elegibles para un ascenso, cuatro eran mujeres. Tres de las nueve personas recibieron en realidad el ascenso; pero sólo una de ellas era mujer. Las otras tres mujeres elegibles demandaron. Una consideración importante en el caso, unida con la probabilidad de que de las tres personas que recibieron ascenso sólo una mujer fuera seleccionada por casualidad. Es decir, si el género no era un factor, ¿cuál es la probabilidad de que no más que uno de los tres ascensos fuera asignado a una mujer?

Solución

Un consultor económico especializado en asuntos legales fue llamado por el abogado defensor para refutar los cargos. El economista calculó la probabilidad de que ante la ausencia de discriminación, sólo una de las mujeres sería ascendida en su cargo. Este cálculo se basó en que

$N = 9$; el número de personas elegibles para ser ascendidas

$r = 4$; el número en la población identificado como éxitos (mujeres)

$n = 3$; tamaño de la muestra (quienes fueron seleccionados para el ascenso)

$x \leq 1$; el número de éxitos (mujeres) en la muestra.

La probabilidad de que no más que una mujer fuera ascendida es $P(X = 0) + P(X = 1)$.

$$P(X = 1) = \frac{{}_4C_1 \; {}_5C_2}{{}_9C_3} = \frac{4 \times 10}{84} = 0.4762$$

$$P(X = 0) = \frac{{}_4C_0 \; {}_5C_3}{{}_9C_3} = \frac{1 \times 10}{84} = 0.1190$$

Por tanto, $P(X \leq 1) = 0.4762 + 0.1190 = 0.5952$.

Interpretación

Había casi un 60% de probabilidad, sin considerar el género que, no más de una mujer fuera ascendida. Con base en tales hallazgos, como con otras pruebas presentadas en el caso, la corte dictaminó que no había suficiente evidencia de discriminación.

A. Uso del computador

Excel funciona casi como magia en la solución de las distribuciones hipergeométricas. Simplemente haga clic en **INSERTAR > FUNCIÓN > ESTADÍSTICAS > DISTR. HIPERGEOM.** Luego ingrese los valores para x, n, r y N. La respuesta aparecerá en la casilla **Valor**.

Ejercicios de la sección

13. Como subgerente de su empresa de materias primas, usted debe contratar 10 personas entre 30 candidatos, 22 de los cuales tienen títulos universitarios. ¿Cuál es la probabilidad de que 5 de los que usted contrate tengan un título?

14. De los 15 altos ejecutivos de un negocio de importaciones y exportaciones, se seleccionan 12 para ser enviados al Japón a estudiar un nuevo proceso de producción. Ocho de los ejecutivos ya tienen algo de entrenamiento en el proceso. ¿Cuál es la probabilidad de que 5 de los enviados tengan algo de conocimiento sobre el proceso antes de partir para el lejano oriente?

15. Cuarenta trabajadores de su oficina han recibido nuevos computadores. Veintisiete tienen la nueva tecnología MMX. Si se seleccionan 10 aleatoriamente, ¿cuál es la probabilidad de que 3 estén equipados con MMX?

16. Una encuesta de la revista *Fortune* (marzo 17 de 1997) sirve como fuente para este problema, que su supervisor le solicita que resuelva. De los 10 empleados hombres, 7 tenían esposas que también trabajaban. ¿Cuál es la probabilidad de que a lo sumo un esposo tenga una esposa que esté empleada fuera de casa si se seleccionan 3 trabajadores al azar?

17. Del problema anterior, la encuesta reveló que 6 de los 10 empleados ganaban más de US$95,000 al año. De los 3 seleccionados, ¿cuál es la probabilidad de que todos tres ganen más de US$95,000?

5.5 La distribución de Poisson

Una variable aleatoria discreta de gran utilidad en la medición de la frecuencia relativa de un evento sobre alguna unidad de tiempo o espacio es la **distribución de Poisson**. Con frecuencia se utiliza para describir el número de llegadas de clientes por hora, el número de accidentes industriales cada mes, el número de conexiones eléctricas defectuosas por milla de cableado en un sistema eléctrico de una ciudad, o el número de máquinas que se dañan y esperan ser reparadas.

Distribución de Poisson Ideada por el matemático francés Simeon Poisson (1781 - 1840), la distribución de Poisson mide la probabilidad de un evento aleatorio sobre algún intervalo de tiempo o espacio.

Son necesarios dos supuestos para la aplicación de la distribución de Poisson:

- La probabilidad de ocurrencia del evento es constante para dos intervalos cualesquiera de tiempo o espacio.
- La ocurrencia del evento en un intervalo es independiente de la ocurrencia de otro intervalo cualquiera.

Dados estos supuestos, la función de probabilidad de Poisson puede expresarse como

Función de probabilidad de Poisson	$P(x) = \dfrac{\mu^x e^{-\mu}}{x!}$	[5.7]

en donde x es el número de veces que ocurre el evento

μ es el número promedio de ocurrencias por unidad de tiempo o de espacio

e = 2.71828, la base del logaritmo natural.

Supongamos que se está interesado en la probabilidad de que exactamente 5 clientes lleguen durante la siguiente hora (o en cualquier hora dada) laboral. La observación simple de las últimas 80 horas ha demostrado que 800 clientes han entrado al negocio. Por tanto, $\mu = 10$ por hora. Utilizando la fórmula (5.7),

$$P(5) = \frac{(10)^5 \times 2.71828^{-10}}{5!} = 0.0378$$

Debido a que esta fórmula es poco práctica, las probabilidades para los valores seleccionados se dan en la tabla D. Diríjase al extremo superior de la tabla hasta encontrar $\mu = 10$. Baje por esa columna hasta la fila en donde está x = 5. Allí usted encontrará 0.0378. Existe 3.78% de oportunidad de que exactamente 5 clientes ingresen a la tienda durante la siguiente hora.

Una compañía de pavimentación local obtuvo un contrato con el Ayuntamiento para hacer mantenimiento a las vías de un gran centro urbano. Las vías recientemente pavimentadas por esta compañía demostraron un promedio de dos defectos por milla, después de haber sido utilizadas durante un año. Si el condado sigue con esta compañía de pavimentación, ¿cuál es la probabilidad de que se presenten 3 defectos en cualquier milla de vía después de haber tenido tráfico durante un año?

$$P(1) = \frac{2^3 \times 2.71828^{-2}}{3!} = 0.1804$$

o 18.04%. Para utilizar la tabla D, halle la columna en donde $\mu = 2$ y la fila en donde $x = 3$. Allí encontrará el valor de 0.1804.

Se supone por el momento que se desea conocer la probabilidad de 3 defectos en 0.5 millas. Debido a que se da la media en ocurrencias por *una milla* (2 por milla) es necesario ajustar μ de acuerdo con la estipulación en el problema de 0.5 millas. Se debe determinar qué porcentaje es 0.5 millas de una milla = 0.5/1 = 0.5. Entonces la media en ocurrencias para este problema es μ = (0.5) (2 ocurrencias) = 1. Si el promedio es de 2 por milla, va a ser 1 por media milla. Por tanto, $P(X = 3 \mid \mu = 1) = 0.0613$. Vale la pena observar el ejemplo 5.4, especialmente la parte *c*.

También debe notarse que si en el problema los valores exceden los rangos limitados en la tabla D, es posible trabajar el problema con calculadoras manuales recordando la regla de exponentes : $e^{-\mu} = 1/e^{\mu}$. Entonces,

$$\begin{aligned} P(X = 3 \mid \mu = 1) &= \frac{\mu^x e^{-\mu}}{x!} \\ &= \frac{1^3\, 2.71828^{-1}}{3!} \\ &= \frac{(1)\left[\dfrac{1}{2.71828^1}\right]}{3!} \\ &= 0.0613 \end{aligned}$$

Ejemplo 5.4 Una distribución de Poisson para estudiantes prudentes

El profesor Bradley anima a sus estudiantes de estadística a "actuar de forma prudente" consultando al tutor si tienen alguna pregunta mientras se preparan para el examen final. Parece que la llegada de los estudiantes a la oficina del tutor se ajusta a una distribución de Poisson, con un promedio de 5.2 estudiantes cada 20 minutos. El profesor Bradley está preocupado porque si muchos estudiantes necesitan los servicios del tutor, puede resultar un problema de congestión.

a. El tutor debe determinar la probabilidad de que cuatro estudiantes lleguen durante cualquier intervalo de 20 minutos, lo cual podría causar el problema de congestión que teme el profesor Bradley. Si la probabilidad excede el 20%, se contratará un segundo tutor.

b. El tutor debe calcular la probabilidad de que más de cuatro estudiantes lleguen durante algún período de 20 minutos. Si es mayor que el 50%, las horas de oficina del tutor se aumentarán, permitiendo a los estudiantes extender el horario en las que vienen a ver al tutor.

c. Si la probabilidad de que más de siete estudiantes lleguen durante un período cualquiera de 30 minutos excede 50%, el mismo profesor Bradley ofrecerá tutoría adicional.

Solución

a. $P(X = 4 \mid \mu = 5.2) = 0.1681$

b. $P(X > 4 \mid \mu = 5.2)$

$$\begin{aligned} &= 1 - P(X \leq 4 \mid \mu = 5.2) \\ &= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)] \\ &= 1 - [0.0055 + 0.0287 + 0.0746 + 0.1293 + 0.1681] \\ &= 0.5938 \end{aligned}$$

c. Se tiene que $\mu = 5.2$ por cada 20 minutos. La estipulación del profesor cubre un período de 30 minutos. Se debe determinar qué porcentaje es 30 de 20:30/20 = 1.5.

Entonces, μ para cada 30 minutos es 5.2(1.5) = 7.8. Así, pues:

$$\begin{aligned} P(X > 7 \mid \mu = 7.8) &= 1 - [P(X \leq 7)] \\ &= 1 - [P(X = 0) + \cdots + P(X = 7)] \\ &= 0.5188 \end{aligned}$$

Interpretación

Debido a que $P(X = 4) = 0.1681 < 20\%$, un segundo tutor es innecesario. $P(X > 4) = 0.5938 > 50\%$; las horas de oficina del tutor se extenderán. Y $P(X > 7) = 0.5188 > 50\%$; el profesor Bradley ayudará en la tarea de tutoría.

A. Uso del computador

Las probabilidades de Poisson pueden obtenerse también utilizando Minitab y Excel. Se supone que se desea determinar la probabilidad de que 4 estudiantes lleguen a la oficina del tutor de la parte *a* del ejemplo 5.4. Para obtener las probabilidades de Poisson con Minitab, ingrese 4 en la columna (1) de la hoja de trabajo. Seleccione **Calc >Probability Distribution> Poisson > Probability**. Ingrese 5.2 en la casilla de **Mean** y C1 en la casilla de **Input Column**.

Para utilizar Excel, seleccione **Insertar > Función > Estadísticas> Poisson**. Haga clic en **Aceptar**. Ingrese 4 en la casilla **x**, 5.2 en la casilla **Media** y Falso en la casilla **Acumulado**. La respuesta aparecerá en la casilla de **Valor** en la esquina superior derecha.

Ejercicios de la sección

18. A un conmutador de la oficina principal de la compañía llegan llamadas a un promedio de dos por minuto y se sabe que tienen distribución de Poisson. Si el operador está distraído por un minuto, cuál es la probabilidad de que el número de llamadas no respondidas sea:
a. ¿Cero?
b. ¿Por lo menos una?
c. ¿Entre 3 y 5, inclusive?

19. ¿Cuáles serían las probabilidades en el ejercicio 18 si el operador se distrae por 4 minutos?

20. Un proceso de fabricación utilizado para hacer artefactos plásticos Incas presenta una tasa de defectos de 5 por cada 100 unidades. Las unidades se envían a los distribuidores en lotes de 200. Si la probabilidad de que más de 3 salgan defectuosos supera el 30%, usted planea vender en su lugar, camisetas Grateful Dead. ¿Cuál artículo agregará usted al inventario?

21. Usted compra partes para bicicleta de un proveedor en Toledo que tiene 3 defectos por cada 100 partes. Usted está en el mercado para comprar 150 partes pero no aceptará una probabilidad de más del 50% de que más de dos partes sean defectuosas. ¿Usted le compraría a dicho proveedor?

5.6 La distribución exponencial

Como se acaba de observar, la distribución de Poisson es una distribución discreta que mide el número de ocurrencias sobre algún intervalo de tiempo o espacio. Describe por ejemplo, el número de clientes que pueden llegar durante algún período determinado. Por el contrario, la **distribución exponencial** es una distribución continua. Mide el paso del tiempo entre tales ocurrencias. Mientras que la distribución de Poisson describe las tasas de llegada (de personas, camiones, llamadas telefónicas, etc.) dentro de algún período dado, la distribución exponencial estima el lapso entre tales arribos. Si el número de ocurrencias tiene distribución de Poisson, el lapso entre las ocurrencias estará distribuido exponencialmente.

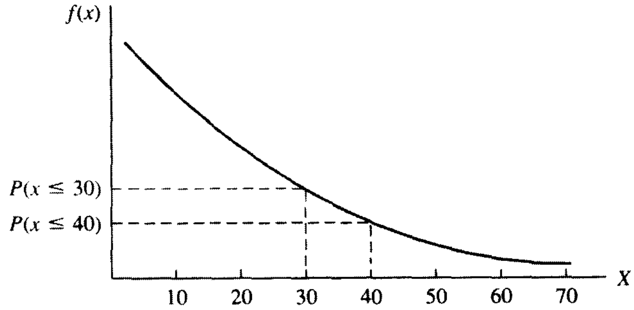
La probabilidad de que el lapso sea menor que o igual a cierta cantidad x es

Distribución exponencial $$P(X \leq x) = 1 - e^{-\mu t} \qquad [5.8]$$

en donde t es el lapso de tiempo
e es la base del logaritmo natural 2.71828
μ es la tasa promedio de ocurrencia

La distribución de una variable aleatoria exponencial se muestra en la figura 5.2. La curva en continuo descenso muestra que con el paso del tiempo X aumenta, y la probabilidad disminuye.

Figura 5.2
Distribución exponencial



La probabilidad de que pasen 30 minutos entre ocurrencias excede la probabilidad de que pasen 40 minutos: $P(X \leq 30) > P(X \leq 40)$. Esto se debe a que *siempre* deben pasar 30 minutos antes que pasen 40.

Así como se tiene que ajustar la tasa promedio de llegada para ajustar la estipulación de Poisson, es necesaria una corrección para la distribución exponencial. Sin embargo, aquí es más fácil ajustar el valor para el tiempo t en la fórmula (5.8) para ajustar el marco de tiempo estipulado en el problema. Se asume que la tasa promedio de llegada de los clientes es $\mu = 1.5$ por hora y se desea saber la probabilidad de que no más de dos horas transcurran entre llegadas. Usando la fórmula (5.7), t es 2. Entonces $P(X \leq 2) = 1 - e^{-(1.5)(2)} = 1 - e^{-3}$. La solución puede hallarse en la mayoría de las calculadoras manuales como $e^{-3} = 1/e^{3} = 0.0498$. Se puede querer hacer uso de la tabla de Poisson (tabla D) así como se hizo con la solución de los problemas de Poisson. El secreto es establecer que $x = 0$, la tabla D revela que $e^{3} = 0.0498$. La probabilidad de que no más de dos horas transcurran entre la llegada de los clientes es $1 - 0.0498 = 0.9592$. Existe 95.02% de probabilidad de que el segundo cliente ingrese a las dos horas o menos del primero si la tasa promedio de llegadas es de 1.5 por hora.

Los camiones llegan al puerto de carga a una tasa de $\mu = 2$ por hora. ¿Cuál es la probabilidad de que no más de 30 minutos transcurran entre llegadas? La tasa promedio de llegada está dada por hora, o 60 minutos, y el problema se plantea en minutos (30 de ellos). Para evitar "el asunto de las manzanas y las naranjas" se debe hallar *qué porcentaje de 60 minutos son 30.* Así $t = 30/60 = ½$. Entonces, $P(X \leq 30 \text{ minutos}) = 1 - e^{-(2)(1/2)} = 1 - e^{-2}$ utilizando una calculadora o confiando en la tabla D estableciendo que $x = 0$, se tiene que $1 - 0.1353 = 0.8647$.

Ejemplo 5.5

Cross City Cab Company programa sus taxis para que lleguen al aeropuerto local en una distribución de Poisson con una tasa promedio de llegada de 12 por hora. Usted acaba de aterrizar en el aeropuerto y debe llegar al centro a cerrar un gran negocio. Cuál es la probabilidad de que usted tenga que esperar máximo 5 minutos para conseguir un taxi? Su jefe es un tirano que no tolerará la falla, de manera que si la probabilidad de que pase otro taxi dentro de 5 minutos es menor al 50%, usted alquilará un carro para el viaje a la oficina.

Solución

Asumiendo lo peor, que el último taxi acaba de irse, usted debe determina $P(X \leq 5 \text{ minutos})$. Debido a que $\mu = 12$ por 60 minutos, usted debe determinar a qué porcentaje son 5 minutos de 60:5/60 =1/12. Por lo tanto, $t = 1/12$ y $P(X \leq 5) = 1 - e^{-(12)(1/12)} = 1 - e^{-1}$. Con una calculadora o utilizando la tabla D se determina $P(X \leq 5) = 1 - 0.3679 = 63.21\%$.

Interpretación

Usted puede relajarse y esperar el taxi, hay una probabilidad de 63.21% (>50%) de que llegue uno dentro de 5 minutos. Mientras espera el taxi, debe considerar que la probabilidad de que llegue uno entre 5 y 10 minutos es igual a $P(X \leq 10) - P(X \leq 5)$. Usted también puede desear matar el tiempo revisando sus reglas algebraicas de exponentes en caso de que el exponente para e no funcione de manera conveniente para un agradable y respetable número entero como ocurrió anteriormente. Vale la pena recordar que $e^{-3} = \frac{1}{e^3}$ y $e^{-0.4} = e^{-4/10} = 1/\sqrt[10]{e^4}$.

A. Usando el computador

Excel nuevamente prueba su valor al calcular probabilidades exponenciales. Simplemente haga clic en **Insertar > Función >Estadísticas> Distr. exp**. Una vez ingresado el valor *ajustado* de t en la casilla **x**, la media en la casilla de **Lambda** y Verdadero en la casilla de **Acum**, entonces de inmediato la respuesta aparece en la casilla de **Valor**.

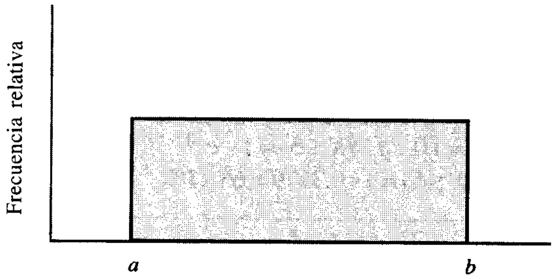
Ejercicios de la sección

22. Los aviones llegan al pequeño aeropuerto en Puerto Vallarta, México, a una proporción de dos por hora. Tomará una hora reparar una rampa utilizada para desembarcar pasajeros. ¿Cuál es la probabilidad de que un avión llegue mientras que la rampa está en reparación?
23. El computador principal de la universidad queda fuera de línea tres veces por semana. El profesor Mundane debe completar un proyecto esta semana que requiere del computador. ¿Cuál es la probabilidad de que el computador esté fuera de línea toda la semana?
24. En el ejercicio 23, ¿cuál es la probabilidad de que el computador esté fuera de línea por cualquier período de dos semanas?
25. Durante un día de trabajo típico de 8 horas, los computadores utilizados para vigilar la etapa de enfriamiento en la producción de neumáticos para autos señala que la temperatura no se mantiene de forma apropiada en 30 oportunidades. El Sr. Radial, director ejecutivo de la compañía, está por hacer una inspección de la planta durante 30 minutos. ¿Cuál es la probabilidad de que esté allí cuando se active la señal del computador?
26. En el ejercicio 25, ¿cuál es la probabilidad de que la visita del Sr. Radial sea interrumpida por la señal del computador?

5.7 La distribución uniforme

La **distribución de probabilidad uniforme** es una distribución en la cual las probabilidades de todos los resultados son las mismas. El experimento de lanzar un dado ilustrado en la tabla 5.2 es uno de los ejemplos. Todos los seis resultados tenían 1/6 de probabilidad de ocurrencia. La figura 5.3 muestra una distribución uniforme en la cual todos los resultados sobre el rango total de posibilidades de distribución son igualmente posibles, desde el mínimo de *a* hasta el máximo de *b*.

Figura 5.3
Distribución uniforme



Distribución uniforme En una distribución uniforme las probabilidades son las mismas para todos los posibles resultados.

La media o valor esperado de una distribución uniforme está a mitad de camino entre sus dos puntos extremos. Así:

Media de una distribución uniforme
$$E(x) = \mu = \frac{a+b}{2} \qquad [5.9]$$

en donde a y b son los valores más bajo y más alto, respectivamente.

La varianza es:

Varianza de una distribución uniforme de probabilidad	$\sigma^2 = \dfrac{(b-a)^2}{12}$	[5.10]

El área total bajo la curva, como en el caso de todas las distribuciones de probabilidad, debe ser igual a 1 o 100%. Debido a que el área es la altura por el ancho, la altura es

$$\text{Altura} = \frac{\text{Área}}{\text{Ancho}}$$

y por tanto

$$\text{Altura} = \frac{1}{b-a} \qquad [5.11]$$

en donde b - a es el ancho o rango de la distribución.

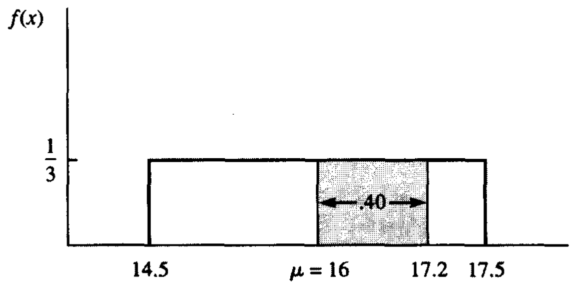
Suponga que los contenidos de las latas de 16 onzas de fruta enlatada producida por Del Monte oscila entre 14.5 y 17.5 onzas y se ajusta a una distribución uniforme. Esto se muestra en la figura 5.4. La media es

$$\mu = \frac{14.5 + 17.5}{2} = 16 \text{ onzas}$$

y la altura es

$$\text{Altura} = \frac{1}{17.5 - 14.5} = 1/3$$

Figura 5.4
Distribución uniforme de los productos enlatados



Asuma que Del Monte desea saber la probabilidad de que una sola lata pese entre 16 y 17.2 onzas. Este valor está dado por el área dentro de ese rango tal y como se muestra en la figura 5.4. La probabilidad de que una observación única esté comprendida dentro de dos valores X_1 y X_2 es

Probabilidad de que una observación caiga entre dos valores	$P(X_1 \leq X \leq X_2) = \dfrac{X_2 - X_1}{\text{rango}}$	[5.12]

Para la lata de Del Monte, se tiene

$$P(16 < X < 17.2) = \frac{17.2 - 16}{17.5 - 14.5}$$
$$= 0.40$$

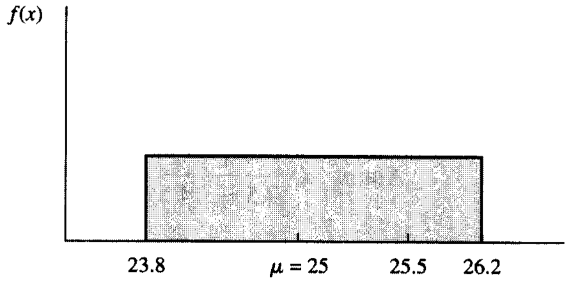
Ejemplo 5.6

Dow Chemical produce un fertilizante inorgánico para césped, para aquellos quienes fertilizan su pasto en casa, de manera que lo puedan podar con más frecuencia. Este tipo de fertilizante se vende en bolsas con un peso uniformemente distribuido, con una media de 25 libras y un rango de 2.4 libras. Harry Homeowner necesita 23 libras para fertilizar su césped, pero duda si comprar sólo una bolsa ya que se desvía de las 25 libras en un rango de 2.4 libras. También tiene curiosidad sobre la probabilidad de comprar una bolsa con más de 25.5 libras.

Solución

Si las bolsas tienen un promedio de 25 libras sobre un rango de 2.4 libras, entonces la mitad de ese rango, o 1.2 libras, debe estar por debajo de 25, y la otra mitad, por encima de 25 libras. Por consiguiente, el peso mínimo será 25 - 1.2 = 23.8 libras y el peso máximo es 25 + 1.2 = 26.2 libras, como se observa en la figura. La probabilidad de seleccionar una sola bolsa que contenga entre 25.5 y 26.2 libras es

$$P(25.5 < X < 26.2) = \frac{26.2 - 25.5}{2.4} = 0.2917$$



Interpretación

Harry no tiene que preocuparse. La bolsa más liviana que podría comprar pesa 23.8 libras. Definitivamente comprará por lo menos las 23 libras que necesita para su césped. Además, la probabilidad de seleccionar una bolsa con más de 25.5 libras es 29.17%.

A. Usando el computador

Minitab puede utilizarse de manera efectiva para determinar probabilidades uniformes. Ingrese los valores para los cuales desea hallar las probabilidades en la Columna (1). Haga clic en **Calc.** > **Probability Distribution>Uniform** > **Cumulative Probability**. Ingrese los valores más bajo y más alto y C1 en la casilla de **Input Column**.

Ejercicios de la sección

27. Generalmente le toma entre 1.2 y 1.7 horas aproximadamente hacer su tarea de estadística. Los tiempos están distribuidos de manera uniforme. ¿Qué tan probable es que usted termine a tiempo para reunirse con sus amigos dentro de 1.4 horas?

28. Las latas de alimento para perros Happy-Tale tienen un promedio de 16 onzas, con un rango de 4.2 onzas.
    a. ¿Cuál es la lata más pequeña en onzas que usted puede comprar para Weiner, su perro de raza poodle toy? ¿Cuál es la lata más grande que usted puede comprar para su perro lobo llamado Killer?
    b. ¿Si usted selecciona una lata al azar, cuál es la probabilidad de que pese entre 15.8 y 16.5 onzas?

29. El agua utilizada por Auto-Brite para lavar los carros es de 30 galones por carro. Lo menos que se utiliza son 27 galones, y su uso está distribuido uniformemente. Una encuesta muestra que los carros no quedan limpios a menos que se utilicen 32 galones de agua en la lavada. ¿Qué porcentaje de carros que salen de Auto-Brite queda limpios?

30. El tiempo requerido para conseguir una pista en una bolera local oscila entre 23.5 y 40.5 minutos. Asumiendo una distribución uniforme, si la probabilidad de que usted tenga que esperar más de 30 minutos excede del 60%, usted piensa jugar golf. ¿Cuál bolsa debería colocar en su baúl, la bolsa de golf o la de bolos?

31. Debido a que usted decidió jugar golf, dada su respuesta a la pregunta anterior, usted aprende que el tiempo promedio para jugar 18 hoyos en esta cancha es de 4.2 horas. La persona que completó este trayecto más rápidamente fue Rapid Roy Parr, quien tomó 2.9 horas. Si los tiempos están distribuidos uniformemente, ¿cuál es la probabilidad de que usted termine a tiempo para llegar a casa para ver el juego de fútbol entre Pittsburgh Steelers y Denver Broncos que comienza en 4 horas?

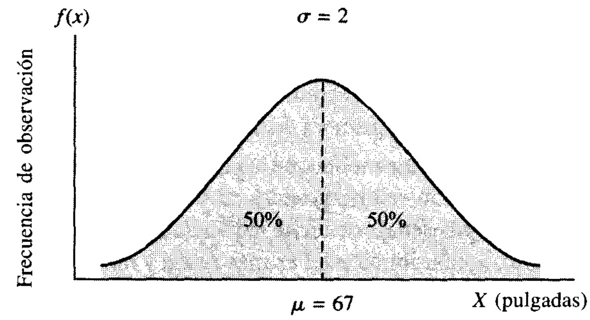
5.8 La distribución normal

De todas las distribuciones de probabilidad que se analizarán, la distribución normal es la más importante. En el capítulo 3 se hizo una introducción a la naturaleza básica de la distribución normal, su característica simétrica en forma de campana y la forma como se relacionaba con la regla empírica. En este momento debe recordarse que la distribución normal es una distribución continua (no discreta). Se utiliza para reflejar la distribución de variables tales como estaturas, pesos, distancias y otras medidas que son divisibles infinitamente. Tales variables continuas generalmente son el resultado de la medida.

Consideremos un caso en el cual ToppsWear, un gran fabricante de ropas, desea estudiar la distribución en la estatura de las personas. Topps Wear reconoció que el público estaba en constante cambio en su tamaño físico y en sus proporciones. En un esfuerzo por producir la ropa de mejor ajuste, la gerencia sintió que se necesitaba un análisis completo de las tendencias actuales en los tamaños de moda. Se supone que si Topps Wear fuera a medir las estaturas de todos sus clientes potenciales, encontrarían que las estaturas están distribuidas normalmente alrededor de una media de 67 pulgadas. Es decir, que mientras que la estatura promedio es de 67 pulgadas, algunas personas son más altas y algunas más bajas. Esta dispersión por encima y por debajo de la media podría medirse mediante la desviación estándar que se calculó en el capítulo 3. Se asume que la desviación estándar en las estaturas de los clientes es de 2 pulgadas.

Una gráfica de estas estaturas produciría la habitual forma de campana. La figura 5.5 muestra esta gráfica, colocando las observaciones individuales en el eje horizontal, y la frecuencia con la cual cada una de estas

Figura 5.5
Distribución normal de las estaturas para Topps Wear

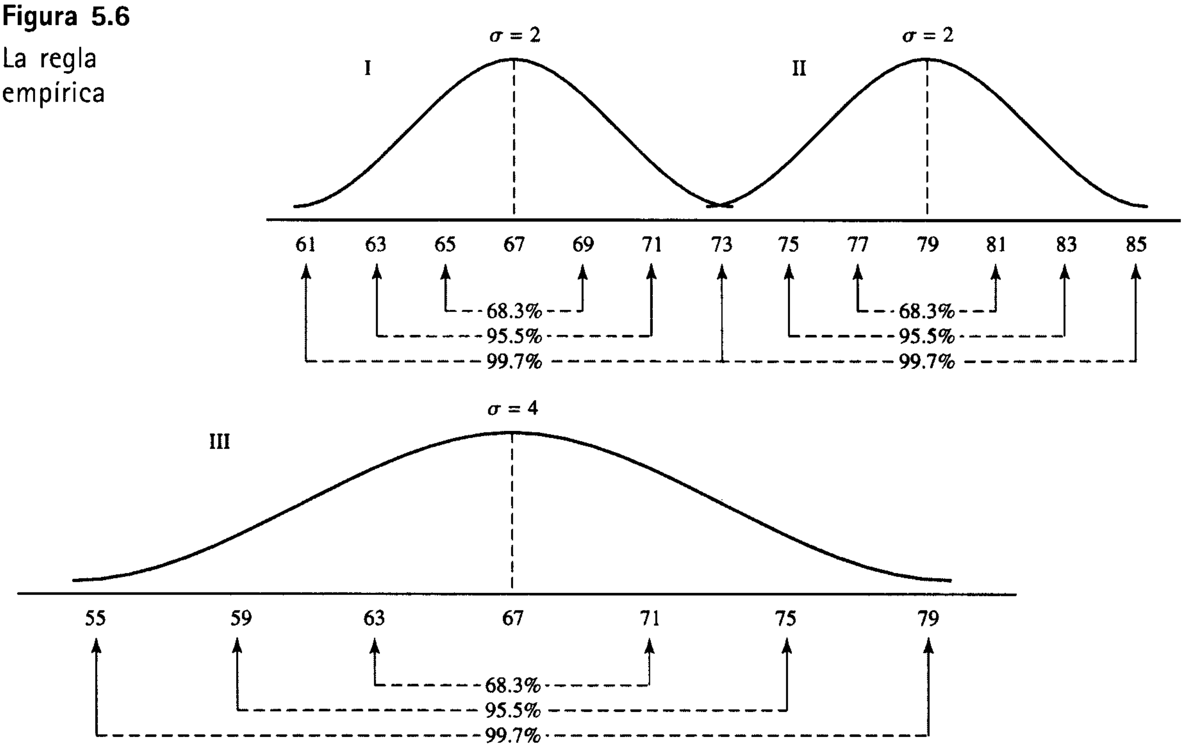


observaciones ocurrieron en el eje vertical. Si los valores son todavía normales (ej: distribuidos normalmente), entonces aparecerá la curva en forma de campana. Vale la pena recordar que más del 50% de las observaciones (estaturas) están por encima de la media y el 50% de éstas están debajo de la media. Similarmente el 50% de toda el área bajo la curva normal está a la derecha de la media y el 50% de esta área está a la izquierda de la media. Esto se observa también en la figura 5.5.

A. Comparación entre distribuciones normales

La forma y posición de una distribución normal están determinadas por dos parámetros: su media μ y su desviación estándar σ. La figura 5.6 muestra tres distribuciones normales diferentes de las tallas que ToppsWear puede encontrar en su estudio sobre las tendencias de la moda. La primera (I) corresponde a las distribuciones descritas anteriormente, las cuales tienen una media de $\mu = 67$ y una desviación estándar de $\sigma = 2$. Está centrada en 67 con la mitad de las observaciones por encima de 67 y la mitad por debajo. La desviación estándar de 2 indica el grado en el cual las observaciones están dispersas por encima y por debajo de 67.

Figura 5.6
La regla empírica



La segunda distribución (II) tiene una media más alta, de $\mu = 79$, pero la misma desviación estándar de $\sigma = 2$. Por tanto, está centrada más hacia la derecha, directamente encima de 79. Pero como tiene el mismo grado de dispersión ($\sigma = 2$), toma la misma forma que la primera distribución.

Una tercera distribución (III) tiene la misma media que la primera ($\mu = 67$) y por tanto está centrada en el mismo sitio. Sin embargo, su medida de dispersión es más grande, tal y como lo indicó la desviación estándar de $\sigma = 4$. Las observaciones varían por encima y por debajo de dicha media de 67 hasta un grado mayor que las observaciones de la primera distribución. Por tanto, la distribución III es más plana y más dispersa por encima y por debajo de la media de 67.

A pesar de sus diferencias, las tres son distribuciones normales. Son simétricas y en forma de campana. Además, como conjuntos de datos normalmente distribuidos, la regla empírica que se analizó en el capítulo 3 se aplica a cada distribución. La **regla empírica** especifica que, sin considerar el valor de la media o la desviación estándar,

El 68.3% de todas las observaciones está a una desviación estándar de la media.

El 95.5% de todas las observaciones está a dos desviaciones estándar de la media.

El 99.7% de todas las observaciones está a tres desviaciones estándar de la media.

La figura 5.6 ilustra la regla empírica. Vale la pena destacar que para los tres conjuntos de datos, sin considerar el valor de μ o σ, el 68.3% de todas las observaciones está a un σ de μ. Comparando la primera distribución (I) con la tercera (III). Debido a que la tercera distribución está altamente dispersa, es necesario tomar un intervalo más amplio para incluir la misma proporción de observaciones. Mientras que la primera distribución incluye el 68.3% de todas las observaciones dentro del intervalo 65 a 69, la tercera distribución comprende este mismo porcentaje sólo dentro de un intervalo más amplio de 63 a 71.

Incluir un cierto porcentaje de todas las observaciones dentro de un intervalo, significa también abarcar el mismo porcentaje de toda el área que está debajo de la curva dentro de dicho intervalo. Por tanto, mientras que el intervalo de 65 a 69 contiene el 68.3% de todas las observaciones en la primera distribución, ese mismo intervalo también contiene el 68.3% de toda el área que está debajo de la curva normal.

B. La desviación normal

Puede existir un número infinito de distribuciones normales posibles, cada una con su propia media y su desviación estándar. Ya que obviamente no se puede analizar un número tan grande de posibilidades, es necesario convertir todas estas distribuciones normales a una forma estándar. Esta conversión a la **distribución normal estándar** se efectúa con la **fórmula de conversión** (o fórmula- Z)

La desviación normal o fórmula-Z

$$Z = \frac{X - \mu}{\sigma} \qquad [5.13]$$

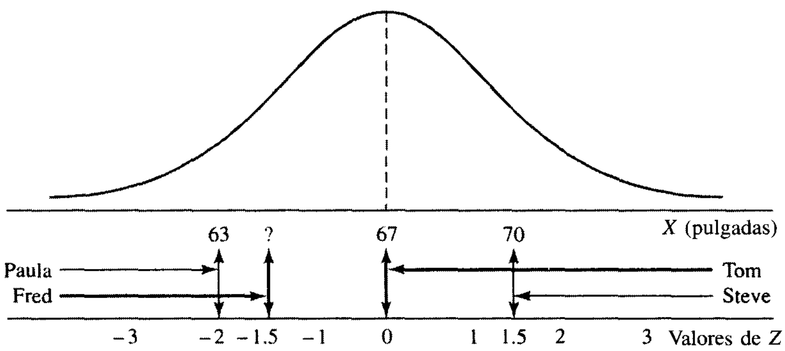
en donde Z es la **desviación normal** y X es algún valor específico de la variable aleatoria. Después de este proceso de conversión, la media de la distribución es 0 y la desviación estándar es 1. Es decir, sin considerar lo que valen la media y la desviación estándar, se miden en las unidades originales en la distribución, después de que se ha aplicado la fórmula de conversión la media es 0 y la desviación estándar es 1.

La figura 5.7 ilustra el uso de los datos de ToppsWear. El eje superior mide las observaciones de estatura X en pulgadas. La media es $\mu = 67$ pulgadas, y la desviación estándar es $\sigma = 2$ pulgadas. El eje inferior refleja estas estaturas en términos de sus valores de Z.

Tom Typical mide 67 pulgadas, la estatura promedio de todos los consumidores en el mercado de ropas de ToppsWear. Utilizando la fórmula (5.13), el valor de Z relacionado con una estatura de $X = 67$ es

$$\begin{aligned} Z &= \frac{X - \mu}{\sigma} \\ &= \frac{67 - 67}{2} \\ &= 0 \end{aligned}$$

Figura 5.7
Conversión de estaturas de los clientes



Esto se puede observar en el eje inferior, o eje Z, en la figura 5.7. Después de aplicar la fórmula de conversión, se encuentra que la estatura promedio de 67 pulgadas tiene un valor de Z de 0. Si se quisiera convertir la estatura de todos los clientes en todo el mercado, se encontraría que todos los valores resultantes de Z tendrían una media de cero y una desviación estándar de 1.

Z, la *desviación normal*, se define como "el número de desviaciones estándar a las que una observación está de la media". Paula Petite mide 63 pulgadas. Su valor de Z es

$$Z = \frac{63 - 67}{2}$$
$$= -2$$

Como lo muestra la figura 5.7, 63 pulgadas están a 4 pulgadas, o 2 desviaciones estándar por debajo de la media. Convirtiendo $X = 63$ pulgadas a su valor de Z resulta -2.00.

> **Valor de Z** Es el número de desviaciones estándar a las que una observación está por encima o por debajo de la media.

Steve Stretch mide 70 pulgadas. La figura 5.7 revela que convertir 70 pulgadas a un valor de Z resulta

$$Z = \frac{70 - 67}{2}$$
$$= 1.5$$

El valor de Z de Steve es 1.5. El está a 3 pulgadas o 1.5 desviaciones estándar por encima de la estatura promedio de 67 pulgadas. ¿Cuánto mide Fred si su valor de Z es –1.5?

C. Cálculo de probabilidades con la desviación normal

Estandarizar una distribución normal permite determinar más fácilmente la probabilidad de que ocurra cierto evento. El personal de ToppsWear puede hallar la probabilidad de que un solo cliente tenga entre 67 y 69 pulgadas de estatura, $P(67 \leq X \leq 69)$, simplemente hallando el área que está bajo la curva normal entre 67 y 69. Es decir, si se conoce el área se conocerá la probabilidad.

Considerándolo en este sentido, se supone que se está disparando a un objetivo o blanco, el cual tiene dos tercios pintado de verde y un tercio pintado de rojo. Se tiene la misma oportunidad de darle a un punto del objetivo como a cualquier otro punto. No necesariamente debe dar en el centro, sólo al objetivo en general. La probabilidad de que pegue en la parte verde es dos tercios. ¿Por qué? Debido a que dos tercios del área están pintados de verde. Si se conoce el área se conoce la probabilidad. Lo mismo puede decirse para el área que está bajo la curva normal.

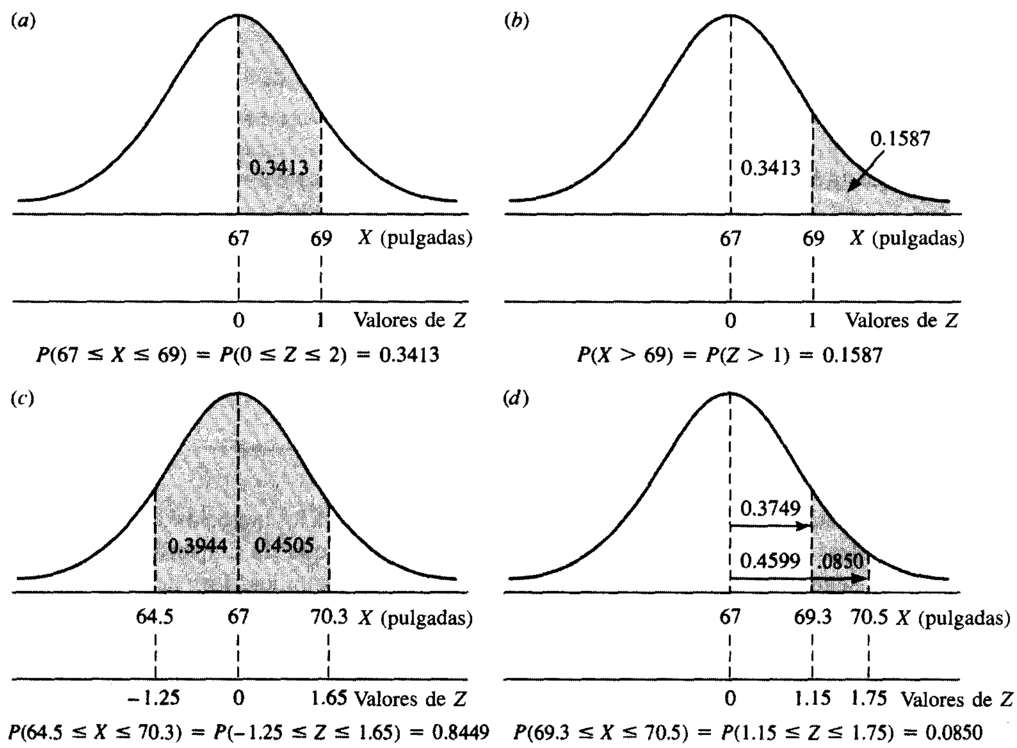
El área relacionada con un valor de Z dado puede hallarse en la tabla E del apéndice III. La figura 5.8(*a*) ilustra este proceso. Se desea saber el área que está entre 67 y 69. El valor de Z para 69 es

$$Z = \frac{69 - 67}{2}$$
$$= 1.00$$

La tabla E proporciona el área bajo la curva *desde la media hasta algún valor por encima o por debajo de ésta.* Ésta es justamente el área que se quiere. En la tabla E, se halla el valor de Z para 1.0 Se va hacia la derecha a la siguiente columna que comienza con 0.00 para obtener $Z = 1.00$. Allí se encontrará la entrada 0.3413. Es decir, el 34.13% del área que está bajo la curva está entre 67 y 69. Hay 34.13% de probabilidad de que un cliente seleccionado aleatoriamente mida entre 67 y 69 pulgadas.

Aunque la tabla E muestra solamente el área desde la media hasta algún valor por encima o por debajo de ella, otras probabilidades pueden hallarse fácilmente. Suponiendo que ToppsWear debe determinar la probabilidad de que un cliente mida más de 69 pulgadas. Como lo muestra la figura 5.8(*b*), ya se ha establecido que el 34.13% de todos los clientes miden entre 67 y 69 pulgadas. Además, también se sabe que el 50% de todos los clientes está por encima de la media de 67. Esto deja 0.5000 - 0.3413 = 0.1587 en el área de la cola que va más allá de 1.00. Hay un 15.87% de probabilidad que un cliente escogido aleatoriamente mida más de 69 pulgadas.

Figura 5.8
Áreas debajo de la curva normal



La figura 5.8(*c*), la cual busca el área comprendida entre 64.5 y 70.3, requiere que se calculen los valores de Z. Como la tabla dará solamente el área de la media hasta algún valor por encima o por debajo de ella, deben determinarse las áreas (1) entre 64.5 y 67 y (2) entre 67 y 70.3 y adicionarlas.

$$Z = \frac{64.5 - 67}{2}$$
$$= -1.25$$

Un valor de Z de 1.25 (se puede ignorar el signo negativo ya que la curva es simétrica y la mitad de la derecha es igual a la mitad de la izquierda) da un área de 0.3944. Para el área comprendida entre 67 y 70.3, se halla

$$Z = \frac{70.3 - 67}{2}$$
$$= 1.65$$

La tabla E revela que el área es 0.4505. Por tanto, $P(64.5 \leq X \leq 70.3) = 0.3944 + 0.4505 = 0.8449$. La probabilidad de que un cliente tenga entre 64.5 y 70.3 pulgadas de estatura es del 84.49%.

Determinar $P(69.3 \leq X \leq 70.5)$ también requiere dos cálculos de Z, como se muestra en la figura 5.8 (*d*). Se debe determinar el área de 67 a 70.5, la cual incluye el área que se desea y una que no se desea. Luego se calcula el área comprendida entre 67 y 69.3 y se resta:

$$Z = \frac{70.5 - 67}{2}$$
$$= 1.75$$

Un valor de Z de 1.75 da un área de 0.4599. Entonces,

$$Z = \frac{69.3 - 67}{2}$$
$$= 1.15$$

lo que produce un área de 0.3749. Entonces $P(69.3 \leq X \leq 70.5) = 0.4599 - 0.3749 = 0.0850$.

Vale la pena notar que entre mayor sea el valor de Z, menor será el área en la cola de la distribución. La tabla E muestra que a medida que Z se aproxima a 3.99, el área abarcada es de virtualmente el 50% por encima de la media, dejando muy poco en la cola más allá de $Z = 3.99$. Por tanto, $P(Z > 3.99) \approx 0$.

Incidentalmente, $P(X < x) = P(X \leq x)$, en donde x es cualquier valor dado. Esto se debe a que la distribución normal es una distribución continua. Existe un número infinito de posibles valores que puede tomar X. Por tanto, *incluir* el valor de x no incrementa la probabilidad de que el evento ocurra.

Ejemplo 5.7

TelCom Satellite presta servicios de comunicación a los negocios del área metropolitana de Chicago. Los funcionarios de la compañía han aprendido que la transmisión satélite promedio es de 150 segundos, con una desviación estándar de 15 segundos. Los tiempos parecen estar distribuidos normalmente.

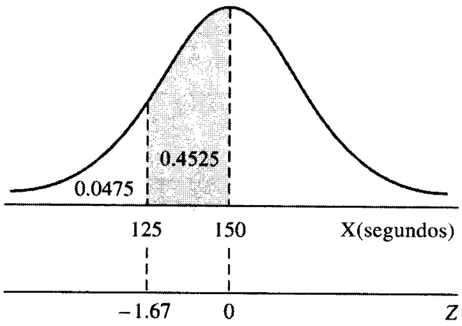
Para estimar de manera apropiada la demanda del cliente por sus servicios y establecer una estructura de tarifas que maximice las utilidades corporativas, TelCom debe determinar qué tan probable es que algunas llamadas se presenten. El director de servicios desea que usted proporcione estimados de la probabilidad de que una llamada dure:

a. Entre 125 y 150 segundos.
b. Menos de 125 segundos.
c. Entre 145 y 155 segundos.
d. Entre 160 y 165 segundos.

Solución

a.

$$Z = \frac{125 - 150}{15} = -1.67$$

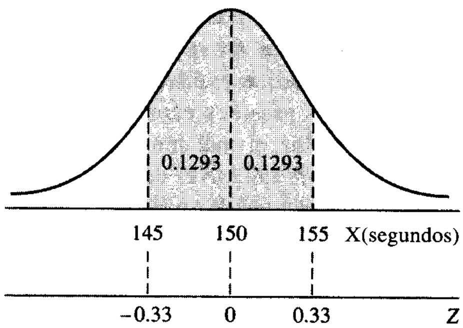


De la tabla E, un valor Z de 1.67 da un área de 0.4525. Por tanto, la probabilidad de que una transmisión dure entre 125 y 150 segundos es del 45.25%.

b. Si el 45.25% del área está entre 125 y 150, entonces 0.5000 – 0.4525 = 0.0475, o el 4.75% de todas las transmisiones requieren menos de 125 segundos. La probabilidad de que cualquier transmisión seleccionada aleatoriamente requiera 125 segundos o menos es del 4.75%.

c.

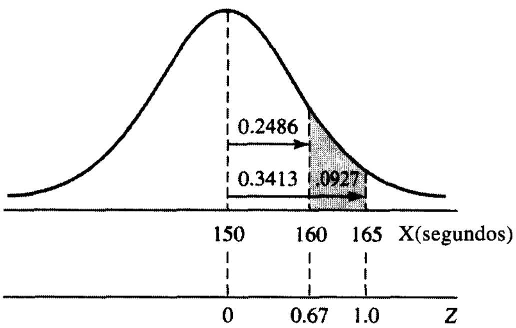
$$Z = \frac{145 - 150}{15} = -0.33$$



Dado $Z = -0.33$, el área que está entre 145 y 150 es 0.1293. Debido a que 155 está a una distancia por encima de la media de 150 igual a la que 145 está por debajo de la media, el área entre 150 y 155 también es 0.1293. Por tanto, $P(145 \leq X \leq 155) = 0.1293 + 0.1293 = 0.2586$.

d.

$$Z = \frac{165 - 150}{15}$$
$$= 1$$



Con $Z = 1$, el área es 0.3413. Para hallar el área entre 150 y 160,

$$Z = \frac{160 - 150}{15}$$
$$= 0.67$$

para un área de 0.2486. Por tanto, $P(160 \leq X \leq 165) = 0.3413 - 0.2486 = 0.0927$.

Interpretación

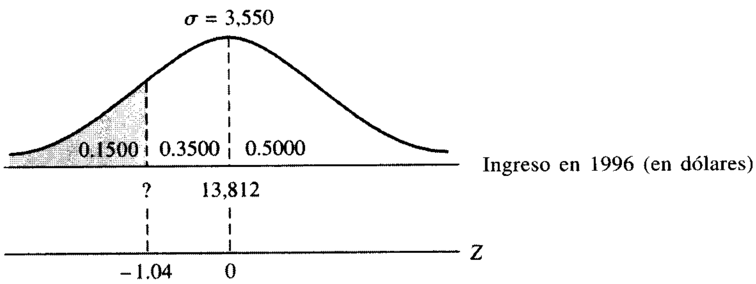
Con base en estas probabilidades, es posible para TelCom desarrollar un sentido de la demanda por sus servicios que le ayudará a establecer políticas respecto al uso de los servicios por parte de los clientes, así como también una estructura de tarifas óptima que TelCom pueda cobrar.

D. Cálculo de un valor X a partir de una probabilidad conocida

En la sección anterior se pidió calcular una probabilidad dado un valor de X. Es decir, que se proporcionaba el valor X para la variable aleatoria, y se debía hallar el área comprendida entre dicho valor y la media. Sin embargo, algunas veces se puede saber cuál probabilidad se requiere, y debe determinarse qué valor de X dará dicha probabilidad. Por ejemplo, se asume que los asesores económicos del presidente proponen un programa de bienestar social para ayudar a los desfavorecidos, el cual consta de un pago monetario al 15% de los más pobres de la nación. Entonces surge la pregunta sobre qué nivel de ingresos separa el 15% más pobre del resto de la gente. En 1996, el ingreso promedio por persona, medido en dólares, en el año de 1982 era de US$13,812. Se asume una desviación estándar de US$3,550. Esto se muestra en la figura 5.9. ¿Existe algún nivel de ingreso que aparezca como "?" que separe el 15% más pobre del 85% restante? Se asume que los ingresos están distribuidos normalmente.

Como se muestra en la figura 5.9, se conoce el área y se busca el valor correspondiente para X que está representado por un signo de interrogación. En problemas anteriores se calculó un valor de Z y se utilizó para

Figura 5.9
Ingresos del 15% más pobre



buscar el área en la tabla. En esta oportunidad se tiene un área y se puede utilizar la tabla E para buscar el valor correspondiente de Z. Aunque se está interesado en el valor de 0.15, se busca 0.3500 (0.5 – 0.15), ya que sólo el área de la media a algún valor por encima o por debajo de ella está dado en la tabla. Se busca en la estructura interna de la tabla E el área de 0.3500. Lo más próximo que se obtiene es 0.3508, lo que corresponde a un valor de Z de 1.04. (La extrapolación puede utilizarse cuando se requiere un grado mayor de exactitud). Debido a que

$$Z = \frac{X - \mu}{\sigma}$$

y a que se halló un valor Z de 1.04, se tiene que

$$-1.04 = \frac{X - 13{,}812}{3{,}550}$$

Cuando se despeja X y se halla X = US$10,120. Cualquiera con un ingreso de US$10,120 o menos recibirá el subsidio del gobierno.

Vale la pena destacar el signo negativo para el valor de Z. El signo algebraico Z no fue importante en problemas anteriores simplemente porque se utilizó el valor de Z para buscar un área en la tabla E. Sin embargo, éste no es nuestro caso. En esta ocasión, el valor de Z se utiliza para cálculos matemáticos adicionales para despejar X. Por consiguiente, su signo sí es de importancia. La regla general es que si se trabaja con el área a la izquierda de la media, el signo siempre es negativo.

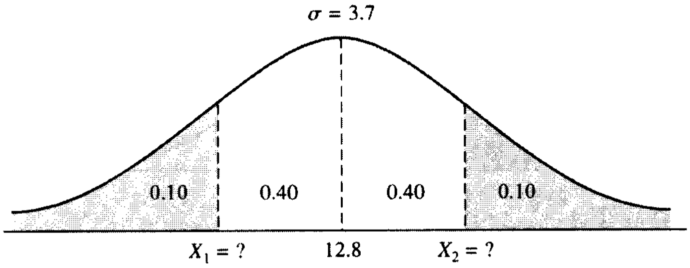
Ejemplo 5.8 Mejorando la prevención urbana contra incendios

Se ha creado una comisión estatal para reducir los tiempos de respuesta de tres estaciones de bomberos. Un grupo de expertos intenta identificar los departamentos de bomberos cuyos tiempos de respuesta estén en el 10% más bajo, o quienes toman más del 90% de todas las estaciones de bomberos en el estudio. Los del primer grupo sirven como modelos para las unidades menos eficientes del segundo grupo.

Los datos muestran que los tiempos promedio de respuesta para una cierta clase de estaciones de bomberos es de 12.8 minutos, con una desviación estándar de 3.7 minutos.

Solución

Se asume que los tiempos de respuesta están distribuidos normalmente; la figura adjunta ilustra el problema. Deben determinarse dos tiempos de respuesta. El primero es tan corto que sólo el 10% de todas las unidades contra incendios llegan al sitio del incendio dentro de dicho lapso. El segundo, es tan largo que sólo el 10% de las unidades toman el mayor tiempo. La fórmula Z se utiliza para determinar cada valor de X. Para establecer el tiempo de respuesta más rápido, se observa 0.4000 en la estructura interna de la tabla E. Aunque se está interesado en el 10% más bajo, se busca 0.4000,



ya que la tabla así está diseñada. La entrada 0.3997 es el valor más próximo y da un valor de Z de 1.28.

Debido a que se está buscando el valor para X en la cola izquierda, el valor de Z se da en el signo negativo apropiado.

$$Z_1 = \frac{X_1 - \mu}{\sigma}$$

$$-1.28 = \frac{X_1 - 12.8}{3.7}$$

$$X_1 = 8.06$$

y

$$Z_2 = \frac{X_2 - \mu}{\sigma}$$

$$1.28 = \frac{X_2 - 12.8}{3.7}$$

$$X_2 = 17.54$$

El valor de Z para X_2 está dado en un signo positivo debido a que se busca un valor en la cola derecha que es mayor que la media.

Interpretación

En esta clasificación sólo el 10% de las estaciones de bomberos respondió a las llamadas en menos de 8.06 minutos. Estas unidades de bomberos servirán como programas modelo para el 10% de las estaciones de bomberos cuyos recorridos superan los 17.54 minutos.

E. Aproximación normal a la distribución binomial

La distribución binomial involucra una serie de n ensayos que pueden producir (1) un éxito o (2) un fracaso. La probabilidad de un éxito se indica como π. Las respuestas pueden hallarse a menudo en la tabla binomial o utilizando la fórmula binomial, fórmula (5.3). Sin embargo, si n es demasiado grande, puede exceder los confines de cualquier tabla y la fórmula puede ser excesivamente engorrosa. Debe diseñarse un método alternativo. La solución puede hallarse con el uso de la distribución normal para aproximar la distribución binomial. Esta aproximación se considera lo suficientemente precisa si $n\pi \geq 5$ y $n(1 - \pi) \geq 5$ y si π está próximo a 0.50.

Se considera un sindicato laboral en el cual el 40% de los miembros está a favor de una huelga. Si se seleccionan 15 miembros de manera aleatoria, ¿cuál es la probabilidad de que 10 apoyen un paro? Para la tabla binomial se halla

$$P(X = 10 \mid n = 15,\ \pi = 0.40) = 0.0245$$

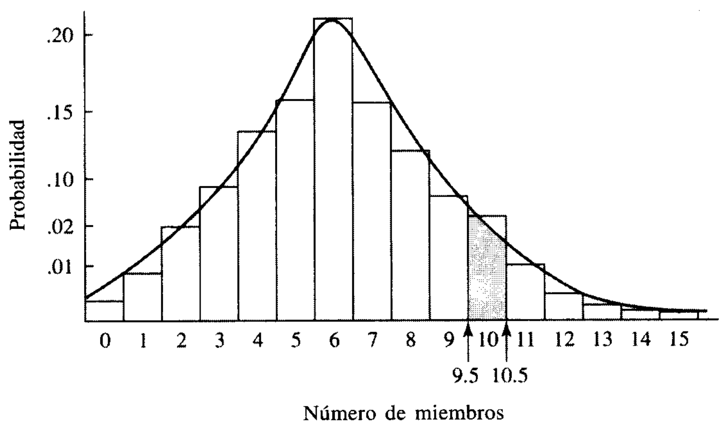
Si no se pudiera utilizar la tabla, podría aproximarse a la respuesta utilizando la distribución normal. Primero se debe hallar la media μ y la desviación estándar σ de la distribución normal así

$$\mu = n\pi \qquad \text{y} \qquad \sigma = \sqrt{n(\pi)(1 - \pi)}$$

En este caso, $\mu = (15)(0.40) = 6$ y $\sigma = \sqrt{15(0.40)(0.60)} = 1.897$.

Debido a que existe un número infinito de valores posibles en una distribución normal (o en cualquier distribución continua), la probabilidad de que la variable aleatoria sea exactamente igual a algún valor específico como 10, es cero. Cuando se utiliza una distribución continua para estimar una variable aleatoria discreta, es necesario un leve ajuste. Este ajuste, llamado *factor de corrección de continuidad*, requiere que se trate la probabilidad de exactamente 10 miembros como el intervalo entre 9.5 miembros y 10.5 miembros. Esto se ilustra en la figura 5.10, la cual muestra las probabilidades para cada valor de la variable aleatoria (número de miembros) tomada de la tabla B.

Figura 5.10
Aproximación normal a la binomial



La probabilidad de que exactamente 10 miembros estén a favor de una huelga está representado por el área del rectángulo centrado en 10. Vale la pena destacar que el rectángulo se extiende de 9.5 a 10.5. La curva normal está superpuesta sobre los rectángulos.

Utilizando la distribución normal para hallar $P(9.5 \leq X \leq 10.5)$, se tiene

$$Z = \frac{9.5 - 6}{1.897} = 1.85$$

para un área de 0.4678, y

$$Z = \frac{10.5 - 6}{1.897} = 2.37$$

para un área de 0.4911. Entonces, $P(9.5 \leq X \leq 10.5)$, = 0.4911 – 0.4678 = 0.0233, lo cual es una aproximación muy cercana al 0.0245 que se encuentra en la tabla B.

Ejercicios de la sección

32. Los paquetes de cereal Cheerios de General Mills vienen en cajas de 36 onzas que tienen una desviación estándar de 1.9 onzas. Se piensa que los pesos están distribuidos normalmente. Si se selecciona una caja aleatoriamente, cuál es la probabilidad de que la caja pese:

a. ¿Menos de 34.8 onzas?

b. ¿Más de 34.8 onzas?

c. ¿Entre 34.3 onzas y 38.9 onzas?

d. ¿Entre 39.5 onzas y 41.1 onzas?

33. Como ingeniero constructor usted compra bolsas de cemento de un promedio de 50 libras, con una desviación estándar de 5.2 libras. Desde que usted tuvo el accidente escalando una montaña, el médico le dijo que no levantara nada que pesara más de 60 libras. ¿Debería usted cargar una bolsa?

34. Se publica que los frenos de los nuevos autos de la marca Lambourginis duran un promedio de 35,000 millas con una desviación estándar de 1,114 millas. Cuál es la probabilidad de que los frenos del auto que usted acaba de comprar le duren:

a. ¿Más de 35,000 millas?

b. ¿Menos de 33.900 millas?

c. ¿Menos de 37.500 millas?

d. ¿Entre 35,200 y 36,900 millas?

35. Los sobrecostos por actualización de computadores en su empresa tienen un promedio de US$23,500, con una desviación estándar deUS$9,400. Como director ejecutivo de la División de Investigación, usted no desea arriesgarse a más de 34% de probabilidad que el sobrecosto en una actualización propuesta recientemente exceda de US$25,000. ¿Debería ejecutar la actualización?

36. El promedio de los salarios en los bancos comerciales en Illinois es de US$22,87 por hora, con una desviación estándar de US$5.87. Cuál debe ser su salario por hora si desea ganar:

a. ¿Más que el 80% de todos los empleados?

b. ¿Más que el 30% de todos los empleados?

c. ¿Menos que el 20% de todos los empleados?

d. ¿Más que el 50% de todos los empleados?

37. Los empleados en Coopers-Price and Lybrand trabajan un promedio de 55.8 horas por semana, con una desviación estándar de 9.8 horas. Los ascensos son más probables para los empleados que están dentro del 10% de los que pasan más tiempo trabajando. ¿Cuánto debe trabajar usted para mejorar sus oportunidades de ascenso?

38. Los registros muestran que 45% de todos los automóviles producidos por Ford Motor Company contiene partes importadas de Japón. ¿Cuál es la probabilidad de que los próximos 200 carros, 115 contengan partes japonesas?

Problemas resueltos

1. Distribución binomial. Un fabricante en California le suministra un diseño prototipo para una pieza de aeronave que requiere su negocio. Este nuevo producto, que es enviado en lotes de $n = 12$, sufre de una tasa de defectos de 40%.

a. Si usted no desea un riesgo mayor del 10% en la probabilidad de que 5 de los 12 sean defectuosos ¿debería comprarle a ese distribuidor?

De la tabla B, $P(X = 5 \mid n = 12, \pi = 0.40) = 0.2270 > 10\%$. No compre.

b. Si usted no desea enfrentar un riesgo mayor del 20% de probabilidad de que más de 5 salgan defectuosos, debería comprarle a este proveedor?