CAPÍTULO 4

Introducción a la probabilidad

CONTENIDO

LA ESTADÍSTICA *en* LA PRÁCTICA

LA EMPRESA ROHM AND HASS*
Filadelfia, Pensilvania

Rohm and Hass es el principal productor de materiales especiales, entre los que se encuentran materiales electrónicos, polímeros para pinturas y artículos para el cuidado personal. Los productos de esta empresa permiten la creación de bienes de consumo de vanguardia en mercados como el farmacéutico, el de alimentos, el de suministros para la construcción, equipos de comunicación y productos para el hogar. La fuerza de trabajo de la empresa es de más de 17 000 personas y sus ventas anuales son de $8 mil millones. Una red de más de 100 puntos de fabricación, investigación técnica y servicio al cliente proporciona los productos y servicios de Rohm and Hass en 27 países.

En el área de productos químicos especiales, la empresa ofrece diversos productos químicos destinados a satisfacer las especificaciones únicas de sus clientes. Para un cliente determinado, la empresa produce un catalizador caro que el cliente emplea en sus procesos químicos. Algunos, pero no todos los lotes que produce la empresa satisfacen las especificaciones del producto. El contrato estipula que el cliente debe probar cada lote después de recibirlo y determinar si el catalizador podrá realizar la función esperada. Los lotes que no pasen la prueba del cliente serán regresados. Con el tiempo, la experiencia ha mostrado que el cliente acepta 60% de los lotes y regresa 40%. Ni el cliente ni la empresa estaban satisfechos con este servicio.

La empresa examinó la posibilidad de, antes de enviar el lote, replicar la prueba que hacía el cliente. Sin embargo, los elevados costos del equipo especial que se necesitaba para la prueba hicieron que esta posibilidad no fuera factible. Los químicos de la empresa encargados del problema propusieron una prueba diferente de costo bajo que se podía practicar antes de enviar el lote al cliente. La empresa creyó que la nueva prueba podría indicar si el catalizador pasaría la compleja prueba que practicaba el cliente.

*Los autores agradecen a Michael Haskell, de la subsidiaria Morton International de Rohm and Hass por haberles proporcionado este artículo para *La estadística en la práctica.*

Una nueva prueba antes de enviar el lote al cliente mejora el servicio al cliente. © Keith Word/Stone.

La pregunta era: ¿cuál es la probabilidad de que el catalizador pase la prueba del cliente dado que pasó la nueva prueba antes de enviar el lote?

La empresa produjo una muestra del catalizador y la sometió a la nueva prueba. Entonces sólo los lotes de catalizador que pasaban la prueba se enviaban al cliente. Mediante el análisis de probabilidad de los datos se supo que si el catalizador pasaba la nueva prueba antes de ser enviado al cliente, la probabilidad de que el catalizador pasara la prueba del cliente era 0.909. O que si el catalizador pasaba la prueba de la empresa, la probabilidad de que no pasara la prueba del cliente y fuera rechazado era 0.091. El análisis de probabilidad aportó evidencias para poner en uso el procedimiento de la prueba antes de enviar el lote. Esta nueva prueba tuvo una mejora inmediata en el servicio al cliente y redujo tanto los costos como los gastos de envío y el manejo de los lotes regresados.

A la probabilidad de que un lote sea aceptado por el cliente, dado que pasó la nueva prueba, se le llama probabilidad condicional. En este capítulo aprenderá cómo calcular la probabilidad condicional y otras probabilidades útiles en la toma de decisiones.

Los administradores sustentan sus decisiones en un análisis de incertidumbres como las siguientes:

1. ¿Qué posibilidades hay de que disminuyan las ventas si aumentamos los precios?
2. ¿Qué posibilidad hay de que un método nuevo de ensamblado aumente la productividad?
3. ¿Cuáles son las posibilidades de que el producto se tenga listo a tiempo?
4. ¿Qué oportunidad existe de que una nueva invención sea rentable?

Algunos de los primeros trabajos sobre probabilidad se dieron en una serie de cartas entre Pierre de Fermat y Blaise Pascal durante el año de 1650.

La **probabilidad** es una medida numérica de la posibilidad de que ocurra un evento. Por tanto, las probabilidades son una medida del grado de incertidumbre asociado con cada uno de los eventos previamente enunciados. Si cuenta con las probabilidades, tiene la capacidad de determinar la posibilidad de ocurrencia que tiene cada evento.

Los valores de probabilidad se encuentran en una escala de 0 a 1. Los valores cercanos a 0 indican que las posibilidades de que ocurra un evento son muy pocas. Los cercanos a 1 indican que es casi seguro que ocurra un evento. Otras probabilidades entre cero y uno representan distintos grados de posibilidad de que ocurra un evento. Por ejemplo, si considera el evento "que llueva mañana", se entiende que si el pronóstico del tiempo dice "la probabilidad de que llueva es cercana a cero", implica que casi no hay posibilidades de que llueva. En cambio, si informan que la probabilidad de que llueva es 0.90, sabe que es muy posible que llueva. La probabilidad de 0.50 indica que es igual de posible que llueva como que no llueva. En la figura 4.1 se presenta la probabilidad como una medida numérica de la posibilidad de que ocurra un evento.

4.1 Experimentos, reglas de conteo y asignación de probabilidades

En el contexto de la probabilidad, un **experimento** es definido como un proceso que genera resultados definidos. Y en cada una de las repeticiones del experimento, habrá uno y sólo uno de los posibles resultados experimentales. A continuación se dan varios ejemplos de experimentos con sus correspondientes resultados.

Experimento	**Resultado experimental**
Lanzar una moneda	Cara, cruz
Tomar una pieza para inspeccionarla	Con defecto, sin defecto
Realizar una llamada de ventas	Hay compra, no hay compra
Lanzar un dado	1, 2, 3, 4, 5, 6
Jugar un partido de futbol	Ganar, perder, empatar

Al especificar todos los resultados experimentales posibles, está definiendo el **espacio muestral** de un experimento.

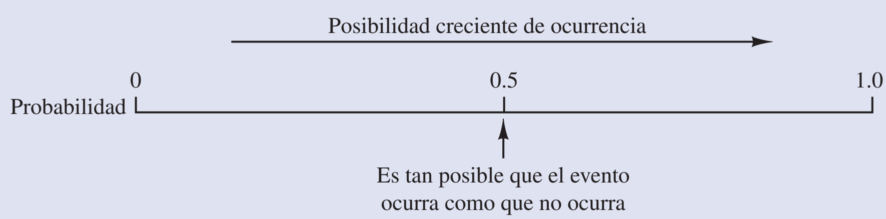
ESPACIO MUESTRAL

El espacio muestral de un experimento es el conjunto de todos los resultados experimentales.

A los resultados experimentales también se les llama puntos muestrales.

A un resultado experimental también se le llama **punto muestral** para identificarlo como un elemento del espacio muestral.

FIGURA 4.1 PROBABILIDAD COMO MEDIDA NUMÉRICA DE LA POSIBILIDAD DE QUE UN EVENTO OCURRA



Considere el primer experimento presentado en la tabla anterior, lanzar una moneda. La cara de la moneda que caiga hacia arriba —cara o cruz— determina el resultado experimental (puntos muestrales). Si denota con S el espacio muestral, puede emplear la notación siguiente para describir el espacio muestral.

$$S = \{\text{Cara, cruz}\}$$

En el segundo experimento de la tabla –tomar una pieza para revisarla– puede describir el espacio muestral como sigue:

$$S = \{\text{Defectuosa, no defectuosa}\}$$

Los dos experimentos descritos tienen dos resultados experimentales (puntos muestrales). Pero, observe ahora el cuarto experimento enumerado en la tabla, lanzar un dado. Los resultados experimentales, definidos por el número de puntos del dado en la cara que cae hacia arriba, son los seis puntos del espacio muestral de este experimento.

$$S = \{1, 2, 3, 4, 5, 6\}$$

Reglas de conteo, combinaciones y permutaciones

Al asignar probabilidades es necesario saber identificar y contar los resultados experimentales. A continuación tres reglas de conteo que son muy utilizadas.

Experimentos de pasos múltiples La primera regla de conteo sirve para experimentos de pasos múltiples. Considere un experimento que consiste en lanzar dos monedas. Defina los resultados experimentales en términos de las caras y cruces que se observan en las dos monedas. ¿Cuántos resultados experimentales tiene este experimento? El experimento de lanzar dos monedas es un experimento de dos pasos: el paso 1 es lanzar la primera moneda y el paso 2 es lanzar la segunda moneda. Si se emplea H para denotar cara y T para denotar cruz, (H, H) será el resultado experimental en el que se tiene cara en la primera moneda y cara en la segunda moneda. Si continúa con esta notación, el espacio muestral (S) en este experimento del lanzamiento de monedas será el siguiente:

$$S = \{(H, H), (H, T), (T, H), (T, T)\}$$

Por tanto, hay cuatro resultados experimentales. En este caso es fácil enumerar todos los resultados experimentales.

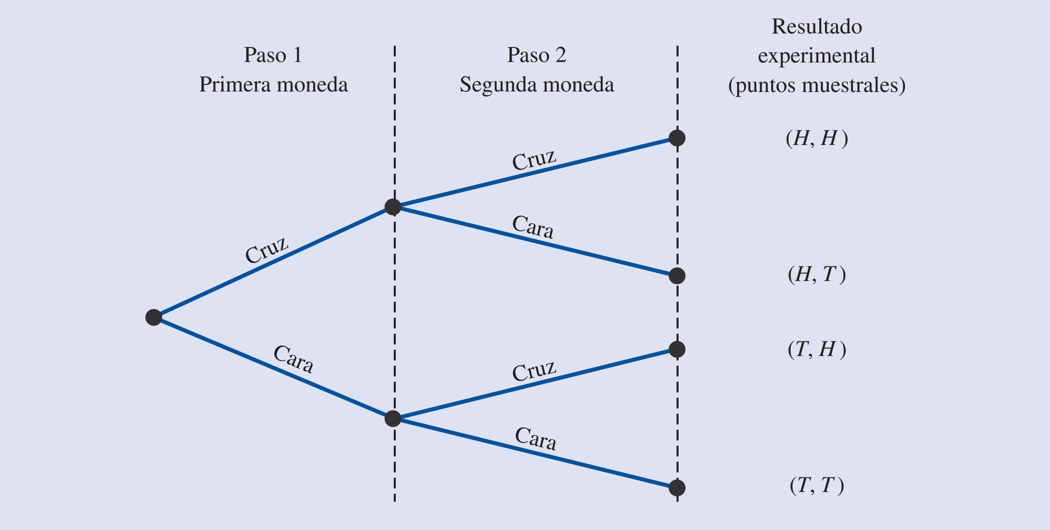
La regla de conteo para experimentos de pasos múltiples permite determinar el número de resultados experimentales sin tener que enumerarlos.

REGLA DE CONTEO PARA EXPERIMENTOS DE PASOS MÚLTIPLES

Un experimento se describe como una sucesión de k pasos en los que hay n_1 resultados posibles en el primer paso, n_2 resultados posibles en el segundo paso y así en lo sucesivo, entonces el número total de resultados experimentales es $(n_1)\,(n_2)\,.\,.\,.\,(n_k)$.

Si considera el experimento del lanzamiento de dos monedas como la sucesión de lanzar primero una moneda ($n_1 = 2$) y después lanzar la otra ($n_2 = 2$), siguiendo la regla de conteo $(2)(2) = 4$, entonces hay cuatro resultados distintos. Como ya se mostró, estos resultados son $S = \{(H, H), (H, T), (T, H), (T, T)\}$. El número de resultados experimentales de seis monedas es $(2)(2)(2)(2)(2)(2) = 64$.

FIGURA 4.2 DIAGRAMA DE ÁRBOL PARA EL LANZAMIENTO DE DOS MONEDAS



Sin el diagrama de árbol podría pensarse que sólo se pueden tener tres resultados experimentales en dos lanzamientos de una moneda: 0 caras, 1 cara y 2 caras.

Un **diagrama de árbol** es una representación gráfica que permite visualizar un experimento de pasos múltiples. En la figura 4.2 aparece un diagrama de árbol para el experimento del lanzamiento de dos monedas. La secuencia de los pasos en el diagrama va de izquierda a derecha. El paso 1 corresponde al lanzamiento de la primera moneda, el paso 2 al de la segunda moneda. En cada paso, los dos resultados posibles son cruz o cara. Observe que a cada uno de los resultados posibles en el paso 1 pertenecen dos ramas por los dos posibles resultados en el paso 2. Cada uno de los puntos en el extremo derecho del árbol representa un resultado experimental. Cada trayectoria a través del árbol, desde el nodo más a la izquierda hasta uno de los nodos en el extremo derecho del árbol, muestra una secuencia única de resultados.

Ahora una aplicación de la regla de conteo para experimentos de pasos múltiples en el análisis de un proyecto de expansión de la empresa Kentucky Power & Light (KP&L). Kentucky Power & Light ha empezado un proyecto que tiene como objetivo incrementar la capacidad de generación de una de sus plantas en el norte de Kentucky. El proyecto fue dividido en dos etapas o pasos sucesivos: etapa 1 (diseño) y etapa 2 (construcción). A pesar de que cada etapa se planeará y controlará con todo el cuidado posible, a los administrativos no les es posible pronosticar el tiempo exacto requerido en cada una de las etapas del proyecto. En un análisis de proyectos de construcción similares encuentran que la posible duración de la etapa de diseño es de 2, 3, o 4 meses y que la duración de la construcción es de 6, 7 u 8 meses. Además, debido a la necesidad urgente de más energía eléctrica, los administrativos han establecido como meta 10 meses para la terminación de todo el proyecto.

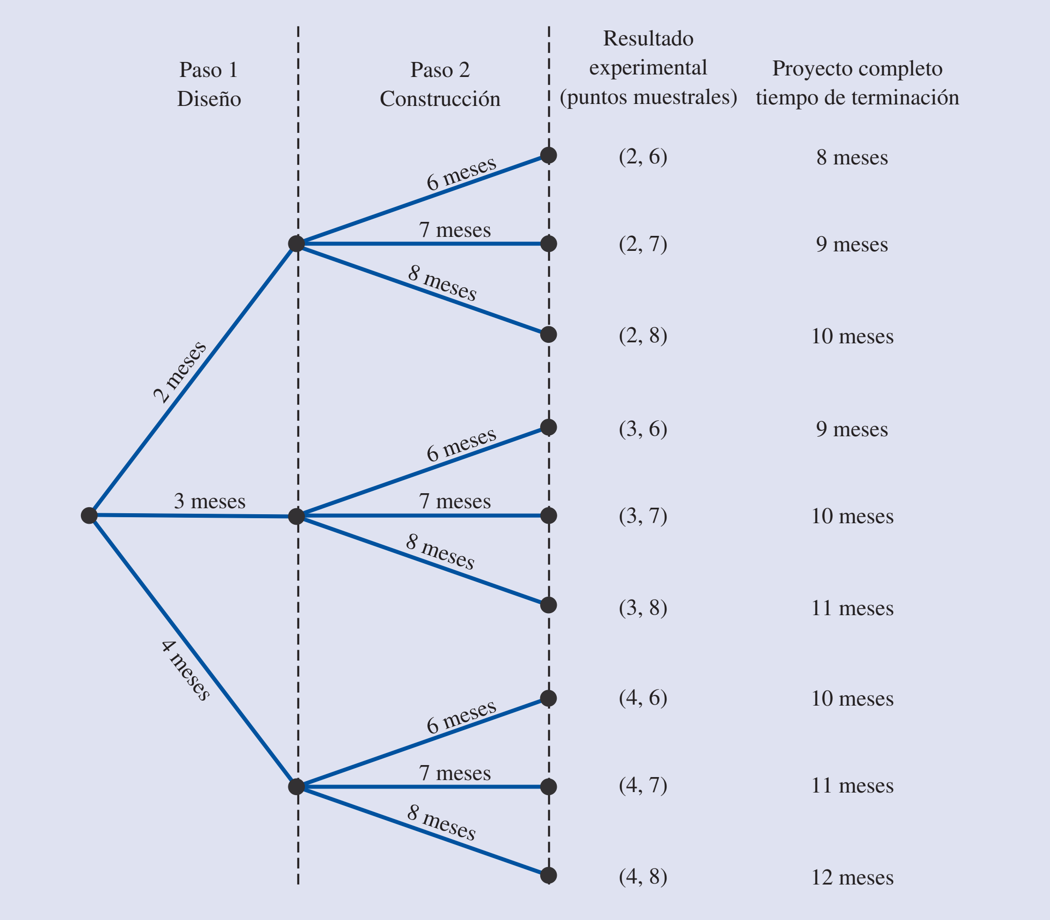
Como hay tres posibles periodos para la etapa del diseño (paso 1) y tres para la etapa de la construcción (paso 2) cabe aplicar la regla de conteo para experimentos de pasos múltiples, entonces el total de resultados posibles es $(3)(3) = 9$. Para describir los resultados experimentales emplean una notación de dos números; por ejemplo, (2, 6) significa que la etapa del diseño durará 2 meses y la etapa de la construcción 6. Esto da como resultado una duración de $2 + 6 = 8$ meses para todo el proyecto. En la tabla 4.1 aparecen los nueve resultados experimentales que hay para el problema de KP&L. El diagrama de árbol de la figura 4.3 muestra como se presentan los nueve resultados (puntos muestrales).

La regla de conteo y el diagrama de árbol ayudan al administrador del proyecto a identificar los resultados experimentales y a determinar la posible duración del proyecto. De acuerdo con la

TABLA 4.1 RESULTADOS EXPERIMENTALES (PUNTOS MUESTRALES) PARA EL PROYECTO KP&L

Duración (meses)			
Etapa 1 Diseño	**Etapa 2 Construcción**	**Notación para los resultados experimentales**	**Proyecto completo: duración (meses)**
2	6	(2, 6)	8
2	7	(2, 7)	9
2	8	(2, 8)	10
3	6	(3, 6)	9
3	7	(3, 7)	10
3	8	(3, 8)	11
4	6	(4, 6)	10
4	7	(4, 7)	11
4	8	(4, 8)	12

FIGURA 4.3 DIAGRAMA DE ÁRBOL PARA EL PROYECTO KP&L



información de la figura 4.3, la duración del proyecto es de 8 a 12 meses, y seis de los nueve resultados experimentales tienen la duración deseada de 10 meses o menos. Aun cuando identificar los resultados experimentales ayuda, es necesario considerar cómo asignar los valores de probabilidad a los resultados experimentales antes de evaluar la probabilidad de que el proyecto dure los 10 meses deseados.

Combinaciones Otra regla de conteo útil le permite contar el número de resultados experimentales cuando el experimento consiste en seleccionar n objetos de un conjunto (usualmente mayor) de N objetos. Ésta es la regla de conteo para combinaciones.

REGLA DE CONTEO PARA COMBINACIONES

El número de combinaciones de N objetos tomados de n en n es

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} \qquad \textbf{(4.1)}$$

donde

$$N! = N(N-1)(N-2)\cdots(2)(1)$$
$$n! = n(n-1)(n-2)\cdots(2)(1)$$

y por definición, $0! = 1$

Cuando se hace un muestreo de una población finita de tamaño N, la regla de conteo para combinaciones sirve para hallar el número de muestras de tamaño n que pueden seleccionarse.

La notación ! significa *factorial;* por ejemplo, 5 factorial es $5! = (5)(4)(3)(2)(1) = 120$.

Como ejemplo del uso de la regla de conteo para combinaciones, considere un procedimiento de control de calidad en el que un inspector selecciona al azar dos de cinco piezas para probar que no tengan defectos. En un conjunto de cinco partes, ¿cuántas combinaciones de dos partes pueden seleccionarse? De acuerdo con la regla de conteo de la ecuación (4.1) es claro que con $N = 5$ y $n = 2$ se tiene

$$C_2^5 = \binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)(3)(2)(1)} = \frac{120}{12} = 10$$

De manera que hay 10 resultados posibles en este experimento de la selección aleatoria de dos partes de un conjunto de cinco. Si etiqueta dichas partes como A, B, C, D y E, las 10 combinaciones o resultados experimentales serán AB, AC, AD, AE, BC, BD, BE, CD, CE y DE.

Para ver otro ejemplo, considere la lotería de Florida en la que se seleccionan seis números de un conjunto de 53 números para determinar al ganador de la semana. Para establecer las distintas variables en la selección de seis enteros de un conjunto de 53, se usa la regla de conteo para combinaciones.

$$\binom{53}{6} = \frac{53!}{6!(53-6)!} = \frac{53!}{6!47!} = \frac{(53)(52)(51)(50)(49)(48)}{(6)(5)(4)(3)(2)(1)} = 22\,957\,480$$

La regla de conteo para combinaciones muestra que la probabilidad de ganar en esta lotería es muy pequeña.

La regla de conteo para combinaciones arroja casi 23 millones de resultados experimentales en esta lotería. Si una persona compra un billete de lotería, tiene una en 22 957 480 posibilidades de ganar la lotería.

Permutaciones La tercera regla de conteo que suele ser útil, es para permutaciones. Dicha regla permite calcular el número de resultados experimentales cuando se seleccionan n objetos de un conjunto de N objetos y el orden de selección es relevante. Los mismos n objetos seleccionados en orden diferente se consideran un resultado experimental diferente.

REGLA DE CONTEO PARA PERMUTACIONES

El número de permutaciones de N objetos tomados de n en n está dado por

$$P_n^N = n!\binom{N}{n} = \frac{N!}{(N-n)!} \quad \textbf{(4.2)}$$

La regla de conteo para permutaciones tiene relación estrecha con la de combinaciones; sin embargo, con el mismo número de objetos, el número de permutaciones que se obtiene en un experimento es mayor que el número de combinaciones, ya que cada selección de n objetos se ordena de n! maneras diferentes.

Para ver un ejemplo, reconsidere el proceso de control de calidad en el que un inspector selecciona dos de cinco piezas para probar que no tienen defectos. ¿Cuántas permutaciones puede seleccionar? La ecuación (4.2) indica que si $N = 5$ y $n = 2$, se tiene

$$P_2^5 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{(5)(4)(3)(2)(1)}{(3)(2)(1)} = \frac{120}{6} = 20$$

De manera que el experimento de seleccionar aleatoriamente dos piezas de un conjunto de cinco piezas, teniendo en cuenta el orden en que se seleccionen, tiene 20 resultados. Si las piezas se etiquetan A, B, C, D y E, las 20 permutaciones son AB, BA, AC, CA, AD, DA, AE, EA, BC, CB, BD, DB, BE, EB, CD, DC, CE, EC, DE y ED.

Asignación de probabilidades

Ahora verá cómo asignar probabilidades a los resultados experimentales. Los tres métodos comúnmente usados son el método clásico, el método de la frecuencia relativa y el método subjetivo. Sin importar el método que se use, es necesario satisfacer los **requerimientos básicos para la asignación de probabilidades**.

REQUERIMIENTOS BÁSICOS PARA LA ASIGNACIÓN DE PROBABILIDADES

1. La probabilidad asignada a cada resultado experimental debe estar entre 0 y 1, inclusive. Si denota con E_i el i-ésimo resultado experimental y con $P(E_i)$ su probabilidad, entonces exprese este requerimiento como

$$0 \leq P(E_i) \leq 1 \text{ para toda } i \quad \textbf{(4.3)}$$

2. La suma de las probabilidades de los resultados experimentales debe ser igual a 1.0. Para resultados experimentales n escriba este requerimiento como

$$P(E_1) + P(E_2) + \cdots + P(E_n) = 1 \quad \textbf{(4.4)}$$

El **método clásico** de asignación de probabilidades es apropiado cuando todos los resultados experimentales tienen la misma posibilidad. Si existen n resultados experimentales, la probabilidad asignada a cada resultado experimental es $1/n$. Cuando emplee este método, satisfará en automático los dos requerimientos básicos de la asignación de probabilidades.

Por ejemplo, considere el experimento del lanzamiento de una moneda, los dos resultados experimentales —cruz o cara— tienen la misma posibilidad. Como uno de los dos resultados igualmente posibles es cara, la probabilidad de que caiga cara es 1/2 o 0.50. Asimismo, la probabilidad de que caiga cruz también es 1/2 o 0.50.

Otro ejemplo, considere el experimento de lanzar un dado. Es razonable pensar que los seis resultados que pueden presentarse son igualmente posibles y, por tanto, la probabilidad asignada a cada resultado es 1/6. Si $P(1)$ denota la probabilidad de que la cara del dado que caiga hacia arriba sea la que tiene un punto, entonces $P(1) = 1/6$. De manera similar $P(2) = 1/6$, $P(3) = 1/6$, $P(4) = 1/6$, $P(5) = 1/6$ y $P(6) = 1/6$. Observe que dichas probabilidades satisfacen los dos requerimientos básicos de las ecuaciones (4.3) y (4.4), porque cada una es mayor o igual que cero y juntas suman 1.0.

El **método de frecuencia relativa** para la asignación de probabilidades es el más conveniente cuando existen datos para estimar la proporción de veces que se presentarán los resultados si el experimento se repite muchas veces. Considere, por ejemplo un estudio sobre los tiempos de espera en el departamento de rayos *x* de un hospital pequeño. Durante 20 días sucesivos un empleado registra el número de personas que están esperando el servicio a las 9:00 a.m.; los resultados son los siguientes.

Número de personas que esperan	Número de días: resultados de ocurrencia
0	2
1	5
2	6
3	4
4	3
Total	20

En estos datos aparece que 2 de los 20 días, había cero pacientes esperando el servicio, 5 días había un paciente en espera y así sucesivamente. Con el método de la frecuencia relativa, la probabilidad que se le asignará al resultado experimental cero pacientes esperan el servicio, será $2/20 = 0.10$; al resultado experimental un paciente espera el servicio, $5/20 = 0.25$; $6/20 = 0.30$ a dos pacientes esperan el servicio; $4/20 = 0.20$ a tres pacientes esperan el servicio y $3/20 = 0.15$ a cuatro pacientes esperan el servicio. Como sucede con el método clásico, al usar el método de frecuencia relativa se satisfacen en automático los dos requerimientos básicos correspondientes a las ecuaciones (4.3) y (4.4).

El **método subjetivo** de asignación de probabilidades es el más indicado cuando no es factible suponer que todos los resultados de un experimento sean igualmente posibles y, además, cuenta con pocos datos relevantes. El método subjetivo de asignación de probabilidades a los resultados de un experimento, usa toda la información disponible, por ejemplo, la propia experiencia o la intuición. Después de considerar dicha información se asigna un valor de probabilidad que expresa el *grado de confianza* (en una escala de 0 a 1) que tiene acerca de que un resultado experimental ocurra. Como la probabilidad subjetiva expresa el grado de confianza que tiene un individuo, es personal. Cuando se usa el método de probabilidad subjetiva, es de esperarse que personas distintas asignen probabilidades diferentes a los mismos resultados de un experimento.

En el método subjetivo hay que tener cuidado de que se satisfagan los dos requerimientos básicos expresados en las ecuaciones (4.3) y (4.4). Sea cual sea el grado de confianza que tenga la persona, el valor de probabilidad asignado a cada resultado experimental debe estar entre 0 y 1, inclusive, y la suma de las probabilidades de todos los resultados experimentales debe ser 1.0.

Considere el caso en el que Tom y Judy Elsbernd hacen una oferta para la compra de una casa. Hay dos resultados posibles:

$$E_1 = \text{su oferta será aceptada}$$
$$E_2 = \text{su oferta no será aceptada}$$

El teorema de Bayes (véase sección 4.5) proporciona un medio para combinar la probabilidad a priori determinada subjetivamente con probabilidades obtenidas por otros medios para obtener probabilidades a posteriori o revisadas.

Judy cree que la probabilidad de que su oferta sea aceptada es 0.8; por tanto, Judy establece que $P(E_1) = 0.8$ y $P(E_2) = 0.2$; Tom, por su parte, cree que la probabilidad de que su oferta sea aceptada es 0.6; por tanto, Tom establecerá $P(E_1) = 0.6$ y $P(E_2) = 0.4$. Observe que la estimación de probabilidad de E_1 que hace Tom refleja bastante pesimismo de que su oferta sea aceptada.

Tanto Judy como Tom asignaron probabilidades que satisfacen los dos requerimientos básicos. El hecho de que sus probabilidades sean diferentes subraya la naturaleza personal del método subjetivo.

Incluso en situaciones de negocios en que es posible emplear el método clásico o el de las probabilidades relativas, los administradores suelen proporcionar estimaciones subjetivas de una probabilidad. En tales casos, la mejor estimación de una probabilidad suele obtenerse combinando las estimaciones del método clásico o del método de las frecuencias relativas con las estimaciones subjetivas de una probabilidad.

Probabilidades para el proyecto KP&L

Para continuar con el análisis del proyecto KP&L hay que hallar las probabilidades de los nueve resultados experimentales enumerados en la tabla 4.1. De acuerdo con la experiencia, los administrativos concluyen que los resultados experimentales no son todos igualmente posibles. Por tanto, no emplean el método clásico de asignación de probabilidades. Entonces deciden hacer un estudio sobre la duración de los proyectos similares realizados por KP&L en los últimos tres años. En la tabla 4.2 se resume el resultado de este estudio considerando 40 proyectos similares.

Después de analizar los resultados de este estudio, los administrativos deciden emplear el método de frecuencia relativa para asignar las probabilidades. Los administrativos podrían haber aportado probabilidades subjetivas, pero se dieron cuenta de que el proyecto actual era muy similar a los 40 proyectos anteriores. Así, consideraron que el método de frecuencia relativa sería el mejor.

Si emplea la tabla 4.2 para calcular las probabilidades, observará que el resultado (2, 6) —duración de la etapa 1, 2 meses, y duración de la etapa 2, 6 meses— se encuentra seis veces en los 40 proyectos. Con el método de las frecuencias relativas, la probabilidad signada a este resultado es $6/40 = 0.15$. También el resultado (2, 7) se encuentra seis veces en los 40 proyectos $6/40 = 0.15$. Continuando de esta manera, se obtienen, para los puntos muestrales del proyecto de KP&L, las asignaciones de probabilidad que se muestran en la tabla 4.3. Observe que $P(2, 6)$ representa la probabilidad del punto muestral (2, 6), $P(2, 7)$ representa la probabilidad del punto muestral (2, 7) y así sucesivamente.

TABLA 4.2 DURACIÓN DE 40 PROYECTOS DE KP&L

Duración (meses) Etapa 1 Diseño	Etapa 2 Construcción	Punto muestral	Número de proyectos que tuvieron esta duración
2	6	(2, 6)	6
2	7	(2, 7)	6
2	8	(2, 8)	2
3	6	(3, 6)	4
3	7	(3, 7)	8
3	8	(3, 8)	2
4	6	(4, 6)	2
4	7	(4, 7)	4
4	8	(4, 8)	6
		Total	40

TABLA 4.3 ASIGNACIÓN DE PROBABILIDADES PARA EL PROYECTO KP&L, EMPLEANDO EL MÉTODO DE LAS FRECUENCIAS RELATIVAS

Punto muestral	Tiempo de terminación del proyecto	Probabilidad del punto muestral
(2, 6)	8 meses	$P(2, 6) = 6/40 = 0.15$
(2, 7)	9 meses	$P(2, 7) = 6/40 = 0.15$
(2, 8)	10 meses	$P(2, 8) = 2/40 = 0.05$
(3, 6)	9 meses	$P(3, 6) = 4/40 = 0.10$
(3, 7)	10 meses	$P(3, 7) = 8/40 = 0.20$
(3, 8)	11 meses	$P(3, 8) = 2/40 = 0.05$
(4, 6)	10 meses	$P(4, 6) = 2/40 = 0.05$
(4, 7)	11 meses	$P(4, 7) = 4/40 = 0.10$
(4, 8)	12 meses	$P(4, 8) = 6/40 = 0.15$
		Total 1.00

NOTAS Y COMENTARIOS

1. En estadística la noción de experimento difiere un poco del concepto de experimento de las ciencias físicas. En las ciencias físicas, los investigadores suelen realizar los experimentos en laboratorios o en ambientes controlados, con objeto de investigar causas y efectos. En los experimentos estadísticos, la probabilidad determina los resultados. Aun cuando un experimento se repita con exactitud, el resultado puede ser completamente diferente. Debido a esta influencia que tiene la probabilidad sobre los resultados, a los experimentos en estadística también se les conoce como *experimentos aleatorios.*

2. Cuando de una población de tamaño N se extrae una muestra aleatoria sin reemplazarla, se emplea la regla de conteo para combinaciones para calcular la cantidad de muestras de tamaño n que pueden seleccionarse.

Ejercicios

Métodos

1. Un experimento consta de tres pasos; para el primer paso hay tres resultados posibles, para el segundo hay dos resultados posibles y para el tercer paso hay cuatro resultados posibles. ¿Cuántos resultados distintos hay para el experimento completo?

Auto examen

2. ¿De cuántas maneras es posible seleccionar tres objetos de un conjunto de seis objetos? Use las letras A, B, C, D, E y F para identificar a los objetos y enumere todas las combinaciones diferentes de tres objetos.

3. ¿Cuántas permutaciones de tres objetos se pueden seleccionar de un grupo de seis objetos? Use las letras A, B, C, D, E y F para identificar a los objetos y enumere cada una de las permutaciones factibles para los objetos B, D y F.

4. Considere el experimento de lanzar una moneda tres veces.
   a. Elabore un diagrama de árbol de este experimento.
   b. Enumere los resultados del experimento.
   c. ¿Cuál es la probabilidad que le corresponde a cada uno de los resultados?

5. Suponga que un experimento tiene cinco resultados igualmente posibles: E_1, E_2, E_3, E_4 y E_5. Asigne probabilidades a los resultados y muestre que satisfacen los requerimientos expresados por las ecuaciones (4.3) y (4.4). ¿Qué método empleó?

Auto examen

6. Un experimento que tiene tres resultados es repetido 50 veces y se ve que E_1 aparece 20 veces, E_2 13 veces y E_3 17 veces. Asigne probabilidades a los resultados. ¿Qué método empleó?

7. La persona que toma las decisiones asigna las probabilidades siguientes a los cuatro resultados de un experimento: $P(E_1) = 0.10$, $P(E_2) = 0.15$, $P(E_3) = 0.40$ y $P(E_4) = 0.20$. ¿Son válidas estas asignaciones de probabilidades? Argumente.

Aplicaciones

8. En una ciudad las solicitudes de cambio de uso de suelo pasan por un proceso de dos pasos: una revisión por la comisión de planeación y la decisión final tomada por el consejo de la ciudad. En el paso 1 la comisión de planeación revisa la solicitud de cambio de uso de suelo y hace una recomendación positiva o negativa respecto al cambio. En el paso 2 el consejo de la ciudad revisa la recomendación hecha por la comisión de planeación y vota para aprobar o desaprobar el cambio de suelo. Suponga que una empresa dedicada a la construcción de complejos departamentales presenta una solicitud de cambio de uso de suelo. Considere el proceso de la solicitud como un experimento. ¿Cuántos puntos muestrales tiene este experimento? Enumérelos. Construya el diagrama de árbol del experimento.

Autoexamen

9. El muestreo aleatorio simple usa una muestra de tamaño n tomada de una población de tamaño N para obtener datos para hacer inferencias acerca de las características de la población. Suponga que, de una población de 50 cuentas bancarias, desea tomar una muestra de cuatro cuentas con objeto de tener información acerca de la población. ¿Cuantas muestras diferentes de cuatro cuentas pueden obtener?

Autoexamen

10. El capital de riesgo es una fuerte ayuda para los fondos disponibles de las empresas. De acuerdo con Venture Economics (*Investor's Business Daily*, 28 de abril de 2000) de 2374 desembolsos en capital de riesgo, 1434 son de empresas en California, 390 de empresas en Massachussets, 217 de empresas en Nueva York y 112 de empresas en Colorado. Veintidós por ciento de las empresas que reciben fondos se encuentran en las etapas iniciales de desarrollo y 55% en la etapa de expansión. Suponga que desea tomar en forma aleatoria una de estas empresas para saber cómo son usados los fondos de capital de riesgo.
    a. ¿Cuál es la probabilidad de que la empresa que seleccione sea de California?
    b. ¿De que la empresa no sea de ninguno de los estados citados?
    c. ¿De que la empresa elegida no se encuentre en las etapas iniciales de desarrollo?
    d. Si admite que las empresas en las etapas iniciales de desarrollo tuvieran una distribución homogénea en todo el país, ¿cuántas empresas de Massachussets que reciben fondos de capital de riesgo se encuentran en las etapas iniciales de desarrollo?
    e. La cantidad total de fondos invertidos es \$32.4 mil millones. Estime la cantidad destinada a Colorado.

11. La National Highway Traffic Safety Administration (NHTSA) realizó una investigación para saber si los conductores de Estados Unidos están usando sus cinturones de seguridad (Associated Press, 25 de agosto de 2003). Los datos muestrales fueron los siguientes.

	Conductores que emplean el cinturón	
Región	**Sí**	**No**
Noreste	148	52
Oeste medio	162	54
Sur	296	74
Oeste	252	48
Total	858	228

    a. ¿Cuál es la probabilidad de que en Estados Unidos un conductor lleve puesto el cinturón?
    b. Un año antes, la probabilidad en Estados Unidos de que un conductor llevara puesto el cinturón era 0.75. El director de NHTSA, doctor Jeffrey Runge esperaba que en 2003 la probabilidad llegara a 0.78. ¿Estará satisfecho con los resultados del estudio del 2003?

c. ¿Cuál es la probabilidad de que se use el cinturón en las distintas regiones del país? ¿En qué región se usa más el cinturón?
d. En la muestra, ¿qué proporción de los conductores provenía de cada región del país? ¿En qué región se seleccionaron más conductores? ¿Qué región viene en segundo lugar?
e. Si admite que en todas las regiones la cantidad de conductores es la misma, ¿ve usted alguna razón para que la probabilidad estimada en el inciso a sea tan alta? Explique.

12. En Estados Unidos hay una lotería que se juega dos veces por semana en 28 estados, en las Islas Vírgenes y en el Distrito de Columbia. Para jugar, debe comprar un billete y seleccionar cinco números del 1 al 55 y un número del 1 al 42. Para determinar al ganador se sacan 5 bolas blancas entre 55 bolas blancas y una bola roja entre 42 bolas rojas. Quien atine a los cinco números de bolas blancas y al número de la bola roja es el ganador. Ocho trabajadores de una empresa tienen el récord del mayor premio, ganaron \$365 millones al atinarle a los números 15-17-43-44-49 de las bolas blancas y al 29 de las bolas rojas. En cada juego hay también otros premios. Por ejemplo, quien atina a los cinco números de las bolas blancas se lleva un premio de \$200 000 (www.powerball.com, 19 de marzo de 2006).
   a. ¿De cuántas maneras se pueden seleccionar los primeros cinco números?
   b. ¿Cuál es la probabilidad de ganar los \$200 000 atinándole a los cinco números de bolas blancas?
   c. ¿Cuál es la probabilidad de atinarle a todos los números y ganar el premio mayor?

13. Una empresa que produce pasta de dientes está analizando el diseño de cinco empaques diferentes. Suponiendo que existe la misma posibilidad de que los clientes elijan cualquiera de los empaques, ¿cuál es la probabilidad de selección que se le asignaría a cada diseño de empaque? En un estudio, se pidió a 100 consumidores que escogieran el diseño que más les gustara. Los resultados se muestran en la tabla siguiente. ¿Confirman estos datos la creencia de que existe la misma posibilidad de que los clientes elijan cualquiera de los empaques? Explique

Diseño	**Número de veces que fue elegido**
1	5
2	15
3	30
4	40
5	10

4.2 Eventos y sus probabilidades

En la introducción de este capítulo el término *evento* fue aplicado tal como se usa en el lenguaje cotidiano. Después, en la sección 4.1 se presentó el concepto de experimento y de los correspondientes resultados experimentales o puntos muestrales. Puntos muestrales y eventos son la base para el estudio de la probabilidad. Por tanto, ahora se le presenta la definición formal de **evento** como se emplea en relación con los puntos muestrales. Con esto se tiene la base para poder dar probabilidades de eventos.

EVENTO

Un evento es una colección de puntos muestrales.

Para dar un ejemplo recuerde el proyecto de KP&L. Considere que al encargado del proyecto le interesa conocer la probabilidad de terminar el proyecto en 10 meses o menos. En la tabla 4.3 aparecen los puntos muestrales (2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6) correspondientes a una duración del proyecto de 10 meses o menos. C denota el evento de que el proyecto dura 10 meses o menos:

$$C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$$

Si cualquiera de estos puntos muestrales es el resultado experimental, entonces ocurre el evento C.

Otros eventos de posible interés para el administrador del proyecto KP&L son los siguientes:

L = El evento de que el proyecto esté acabado en *menos* de 10 meses

M = El evento de que el proyecto esté acabado en *más* de 10 meses

De acuerdo con la tabla 4.3 dichos eventos consisten de los siguientes puntos muestrales

$$L = \{(2, 6), (2, 7), (3, 6)\}$$
$$M = \{(3, 8), (4, 7), (4, 8)\}$$

Para el proyecto KP&L existen otros muchos eventos, pero todos serán una colección de puntos muestrales del experimento.

Dadas las probabilidades de los puntos muestrales que se presentan en la tabla 4.3, para calcular la probabilidad de cualquier evento que interese al administrador del proyecto KP&L, se emplea la definición siguiente.

PROBABILIDAD DE UN EVENTO

La probabilidad de cualquier evento es igual a la suma de las probabilidades de los puntos muestrales que forman el evento.

De acuerdo con esta definición, la probabilidad de un determinado evento se calcula sumando las probabilidades de los puntos muestrales (resultados experimentales) que forman el evento. Ahora es posible calcular la probabilidad de que el proyecto dure 10 meses o menos. Como este evento está dado por $C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$, la probabilidad del evento C denotada por $P(C)$ está dada por

$$P(C) = P(2, 6) + P(2, 7) + P(2, 8) + P(3, 6) + P(3, 7) + P(4, 6)$$

Al consultar las probabilidades de los puntos muestrales de la tabla 4.3, se tiene

$$P(C) = 0.15 + 0.15 + 0.05 + 0.10 + 0.20 + 0.05 = 0.70$$

Así, como el evento de que el proyecto dure menos de 10 meses está dado por $L = \{(2, 6), (2, 7), (3, 6)\}$, la probabilidad de este evento será

$$\begin{aligned} P(L) &= P(2, 6) + P(2, 7) + P(3, 6) \\ &= 0.15 + 0.15 + 0.10 = 0.40 \end{aligned}$$

Por último, el evento de que el proyecto dure más de 10 meses está dado por $M = \{(3, 8), (4, 7), (4, 8)\}$ y por tanto

$$\begin{aligned} P(M) &= P(3, 8) + P(4, 7) + P(4, 8) \\ &= 0.05 + 0.10 + 0.15 = 0.30 \end{aligned}$$

Con estas probabilidades, ahora puede informarle al administrador del proyecto KP&L las probabilidades siguientes: que el proyecto dure 10 meses o menos es 0.70; que dure menos de 10 meses es 0.40 y que dure más de 10 meses es 0.30. Este procedimiento para calcular las probabilidades de los eventos aplica para cualquier evento que interese al administrador del proyecto KP&L.

Siempre que se puedan identificar todos los puntos muestrales de un experimento y asignar a cada uno su probabilidad, es factible calcular la probabilidad de un evento usando la definición. Sin embargo, en muchos experimentos la gran cantidad de puntos muestrales hace en extremo difícil, si no imposible, la determinación de los puntos muestrales, así como la asignación de sus probabilidades correspondientes. En las secciones restantes de este capítulo se presentan algunas relaciones básicas de probabilidad útiles para calcular la probabilidad de un evento, sin necesidad de conocer las probabilidades de todos los puntos muestrales.

NOTAS Y COMENTARIOS

1. El espacio muestral S es un evento. Puesto que contiene todos los resultados experimentales, su probabilidad es 1; es decir $P(S) = 1$.
2. Cuando se usa el método clásico para asignar probabilidades, se parte de que todos los resultados experimentales son igualmente posibles. En tales casos la probabilidad de un evento es calculable contando el número de resultados experimentales que hay en el evento y dividiendo el resultado entre el número total de resultados experimentales.

Ejercicios

Métodos

14. Para un experimento hay cuatro resultados que son igualmente posibles: E_1, E_2, E_3 y E_4.
    a. ¿Cuál es la probabilidad de que ocurra E_2?
    b. ¿De que ocurra cualquiera de dos resultados (por ejemplo, E_1 o E_2)?
    c. ¿De que ocurran tres de estos resultados (E_1 o E_2 o E_4)?

Autoexamen

15. Considere el experimento de seleccionar un naipe de una baraja con 52 naipes. Cada naipe es un punto muestral y su probabilidad es 1/52.
    a. Enumere los puntos muestrales del evento si selecciona un as.
    b. Enumere los puntos muestrales del evento si selecciona un trébol.
    c. Enumere los puntos muestrales del evento si selecciona una figura (sota, rey o reina).
    d. Halle la probabilidad correspondiente a cada uno de los eventos de los incisos a, b y c.

16. Considere el experimento que consiste en lanzar un par de dados. Suponga que lo relevante es la suma de los puntos en las dos caras que caen hacia arriba.
    a. ¿Cuántos puntos muestrales habrá? (*Sugerencia:* Use la regla de conteo para experimentos de pasos múltiples.)
    b. Enumere los puntos muestrales.
    c. ¿Cuál es la probabilidad de obtener un 7?
    d. ¿De obtener un 9 o un número mayor?
    e. Como en cada lanzamiento son factibles seis valores pares (2, 4, 6, 8, 10, y 12) y sólo cinco impares (3, 5, 7, 9 y 11), se tendrán más veces resultados pares que impares. ¿Está de acuerdo? Explique
    f. ¿Qué método usó para calcular las probabilidades pedidas?

Aplicaciones

Auto examen

17. Consulte las tablas 4.2 y 4.3 que muestran los puntos muestrales del proyecto KP&L y sus probabilidades.
    a. La etapa del diseño (etapa 1) saldrá del presupuesto si su duración es mayor a 4 meses. Liste los puntos muestrales del evento si la etapa del diseño sale del presupuesto.
    b. ¿Cuál es la probabilidad de que la etapa del diseño salga del presupuesto?
    c. La etapa de la construcción (etapa 2) saldrá del presupuesto si su duración es mayor a 8 meses. Enumere los puntos muestrales del evento si la etapa de construcción sale del presupuesto.
    d. ¿Cuál es la probabilidad de que la etapa de construcción salga del presupuesto?
    e. ¿Cuál es la probabilidad de que las dos etapas salgan del presupuesto?
18. Suponga que el administrador de un complejo grande de departamentos proporciona la siguiente estimación de probabilidades subjetivas acerca del número de departamentos libres que habrá el mes próximo.

Departamentos libres	Probabilidad
0	0.05
1	0.15
2	0.35
3	0.25
4	0.10
5	0.10

    Dé la probabilidad de cada uno de los eventos siguientes.
    a. No haya departamentos libres.
    b. Haya por lo menos 4 departamentos libres.
    c. Haya 2 o menos departamentos libres.
19. Una asociación deportiva realiza un sondeo entre las personas mayores a 6 años respecto de su participación en actividades deportivas. (*Statistical Abstract of the United States: 2002*). El total de la población de estas edades fue 248.5 millones, de los cuales 120.9 millones eran hombres y 127.6 millones mujeres. A continuación se presenta el número de participantes en los cinco deportes principales.

	Participantes (en millones)	
Actividad	**Hombres**	**Mujeres**
Andar en bicicleta	22.2	21.0
Acampar	25.6	24.3
Caminar	28.7	57.7
Hacer ejercicio con aparatos	20.4	24.4
Nadar	26.4	34.4

    a. Estime la probabilidad de que una mujer, elegida al azar, participe en cada una de estas actividades deportivas.
    b. Estime la probabilidad de que un hombre, elegido en forma aleatoria, participe en cada una de estas actividades deportivas.
    c. Estime la probabilidad de que una persona, elegida en forma aleatoria, haga ejercicio caminando.
    d. Suponga que acaba de ver una persona que pasa caminando para hacer ejercicio. ¿Cuál es la probabilidad de que sea mujer?, ¿de que sea hombre?

20. La revista *Fortune* publica anualmente una lista de las 500 empresas más grandes de Estados Unidos. A continuación se presentan los cinco estados en los que hay más de estas 500 empresas de *Fortune*.

Estado	Número de empresas
Nueva York	54
California	52
Texas	48
Illinois	33
Ohio	30

Suponga que se elige una de las 500 empresas de *Fortune*. ¿Cuál es la probabilidad de cada uno de los eventos siguientes?

a. Sea N el evento: la empresa se encuentra en Nueva York. Halle $P(N)$.
b. Sea T el evento: la empresa se encuentra en Texas. Halle $P(T)$.
c. Sea B el evento: la empresa se encuentra en uno de estos cinco estados. Halle $P(B)$.

21. En la tabla siguiente se dan las edades de la población de Estados Unidos (*The World Almanac 2004*). Los datos aparecen en millones de personas.

Edad	Cantidad
19 y menos	80.5
20 a 24	19.0
25 a 34	39.9
35 a 44	45.2
45 a 54	37.7
55 a 64	24.3
65 y más	35.0

Suponga una selección aleatoria de una persona de esta población.
a. ¿Cuál es la probabilidad de que la persona tenga entre 20 y 24 años?
b. ¿De que la persona tenga entre 20 y 34 años?
c. ¿De que tenga 45 años o más?

4.3 Algunas relaciones básicas de probabilidad

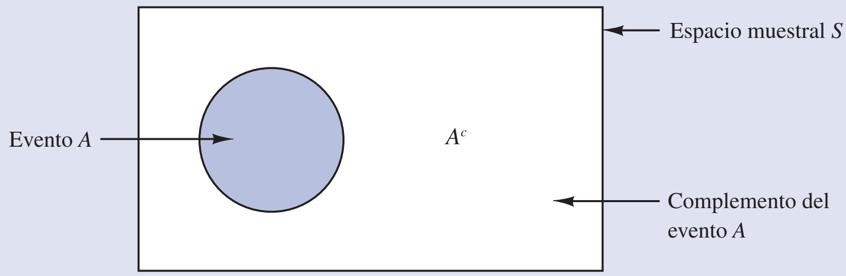
Complemento de un evento

Dado un evento A, el **complemento de A** se define como el evento que consta de todos los puntos muestrales que *no* están en A. El complemento de A se denota A^c. Al diagrama de la figura 4.4 se le llama **diagrama de Venn** e ilustra el concepto del complemento. El área rectangular representa el espacio muestral del experimento y, por tanto, contiene todos los puntos muestrales. El círculo representa el evento A y encierra sólo los puntos muestrales que pertenecen a A. La región del rectángulo que aparece sombreada incluye todos los puntos muestrales que no están en el evento A y es, por definición, el complemento de A.

En cualquier aplicación de la probabilidad ocurre un evento A o su complemento A^c. Por tanto,

$$P(A) + P(A^c) = 1$$

FIGURA 4.4 EL COMPLEMENTO DEL EVENTO A ES EL ÁREA QUE APARECE SOMBREADA



Despejando $P(A)$, obtiene lo siguiente.

CÁLCULO DE UNA PROBABILIDAD USANDO EL COMPLEMENTO

$$P(A) = 1 - P(A^c) \tag{4.5}$$

La ecuación (4.5) indica que la probabilidad de un evento A se puede calcular si se conoce la probabilidad de su complemento, $P(A^c)$.

Por ejemplo, considere el caso de un administrador de ventas que, después de revisar los informes de ventas, encuentra que 80% de los contactos con clientes nuevos no producen ninguna venta. Si A denota el evento hubo venta y A^c el evento no hubo venta, el administrador tiene que $P(A^c) = 0.80$. Mediante la ecuación (4.5) se ve que

$$P(A) = 1 - P(A^c) = 1 - 0.80 = 0.20$$

La conclusión es que la probabilidad de una venta en el contacto con un cliente nuevo es 0.20.

Otro ejemplo, un gerente de compras encuentra que la probabilidad de que el proveedor surta un pedido sin piezas defectuosas es 0.90, empleando el complemento podemos concluir que la probabilidad de que el pedido contenga piezas defectuosas es de $1 - 0.90 = 0.10$.

Ley de la adición

La ley de la adición sirve para determinar la probabilidad de que ocurra por lo menos uno de dos eventos. Es decir, si A y B son eventos, nos interesa hallar la probabilidad de que ocurra el evento A o el B o ambos.

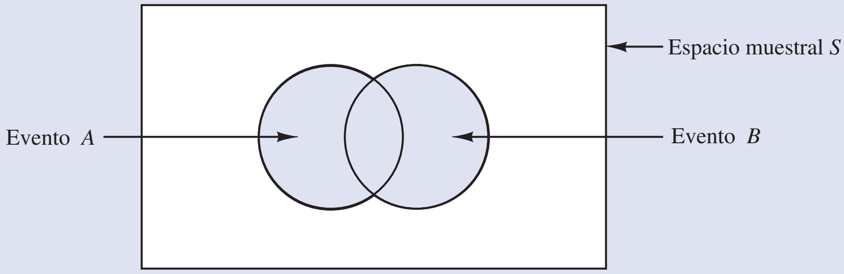
Antes de presentar la ley de la adición es necesario ver dos conceptos relacionados con la combinación de eventos: la *unión* y la *intersección* de eventos. Dados dos eventos, A y B, la **unión de A y B** se define.

UNIÓN DE DOS EVENTOS

La unión de A y B es el evento que contiene todos los puntos muestrales que pertenecen a A o a B o a ambos. La unión se denota $A \cup B$.

El diagrama de Venn de la figura 4.5 representa la unión de los eventos A y B. Observe que en los dos círculos están contenidos todos los puntos muestrales del evento A y todos los puntos

FIGURA 4.5 LA UNIÓN DE LOS EVENTOS *A* Y *B* APARECE SOMBREADA



muestrales del evento *B*. El que los círculos se traslapen indica que algunos puntos muestrales están contenidos tanto en *A* como en *B*.

A continuación la definición de la **intersección de *A* y *B***:

INTERSECCIÓN DE DOS EVENTOS

Dados dos eventos *A* y *B*, la intersección de *A* y *B* es el evento que contiene los puntos muestrales que pertenecen tanto a *A* como a *B*.

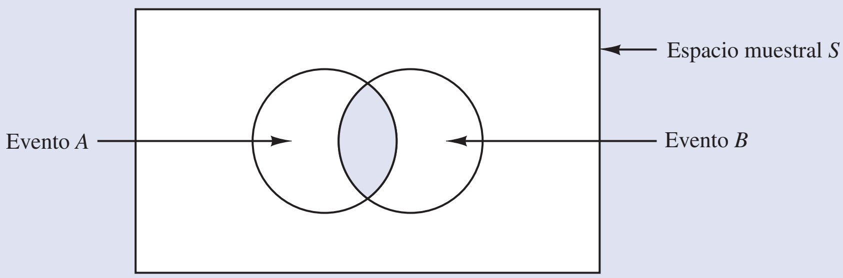
El diagrama de Venn ilustra la intersección de los eventos *A* y *B* mostrados en la figura 4.6. El área donde los círculos se sobreponen es la intersección que contiene una muestra de los puntos que están tanto en *A* como en *B*.

Ahora ya puede continuar con la ley de la adición. La **ley de la adición** proporciona una manera de calcular la probabilidad de que ocurra el evento *A* o el evento *B* o ambos. En otras palabras, la ley de la adición se emplea para calcular la probabilidad de la unión de los dos eventos. La ley de la adición se expresa.

LEY DE LA ADICIÓN

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{4.6}$$

FIGURA 4.6 LA INTERSECCIÓN DE LOS EVENTOS *A* Y *B* APARECE SOMBREADA



Para que logre un entendimiento intuitivo de la ley de la adición, observe que en la ley de la adición, los dos primeros términos $P(A) + P(B)$, corresponden a los puntos muestrales en $A \cup B$. Pero, como los puntos muestrales que se encuentran en la intersección $A \cap B$ están tanto en A como en B, cuando se calcula $P(A) + P(B)$, los puntos que se encuentran en $A \cap B$ cuentan dos veces. Esto se corrige restando $P(A \cap B)$.

Para ver un ejemplo de la aplicación de la ley de la adición, considere el caso de una pequeña empresa de ensamble en la que hay 50 empleados. Se espera que todos los trabajadores terminen su trabajo a tiempo y que pase la inspección final. A veces, alguno de los empleados no satisface el estándar de desempeño, ya sea porque no termina a tiempo su trabajo o porque no ensambla bien una pieza. Al final del periodo de evaluación del desempeño, el jefe de producción encuentra que 5 de los 50 trabajadores no terminarón su trabajo a tiempo, 6 de los 50 trabajadores ensamblaron mal una pieza y 2 de los 50 trabajadores no terminarón su trabajo a tiempo *y* armaron mal una pieza.

Sea

$$
\begin{aligned}
L &= \text{el evento no se terminó el trabajo a tiempo} \\
D &= \text{el evento se armó mal la pieza}
\end{aligned}
$$

La información de las frecuencias relativas lleva a las probabilidades siguientes.

$$
\begin{aligned}
P(L) &= \frac{5}{50} = 0.10 \\
P(D) &= \frac{6}{50} = 0.12 \\
P(L \cap D) &= \frac{2}{50} = 0.04
\end{aligned}
$$

Después de analizar los datos del desempeño, el jefe de producción decide dar una calificación baja al desempeño de los trabajadores que no terminaron a tiempo su trabajo o que armaron mal alguna pieza; por tanto, el evento de interés es $L \cup D$. ¿Cuál es la probabilidad de que el jefe de producción dé a un trabajador una calificación baja de desempeño?

Observe que esta pregunta sobre probabilidad se refiere a la unión de dos eventos. En concreto, se desea hallar $P(L \cup D)$, usando la ecuación (4.6) se tiene

$$P(L \cup D) = P(L) + P(D) - P(L \cap D)$$

Como conoce las tres probabilidades del lado derecho de esta expresión, se tiene

$$P(L \cup D) = 0.10 + 0.12 - 0.04 = 0.18$$

Estos cálculos indican que la probabilidad de que un empleado elegido al azar obtenga una calificación baja por su desempeño es 0.18

Para ver otro ejemplo de la ley de la adición, considere un estudio reciente efectuado por el director de personal de una empresa importante de software. En el estudio encontró que 30% de los empleados que se van de la empresa antes de dos años, lo hacen por estar insatisfechos con el salario, 20% se van de la empresa por estar descontentos con el trabajo y 12% por estar insatisfechos con las *dos* cosas, el salario y el trabajo. ¿Cuál es la probabilidad de que un empleado que se vaya de la empresa en menos de dos años lo haga por estar insatisfecho con el salario, con el trabajo o con las dos cosas?

Sea

S = el evento el empleado se va de la empresa por insatisfacción con el salario
W = el evento el empleado se va de la empresa por insatisfacción con el trabajo

Se tiene $P(S) = 0.30$, $P(W) = 0.20$ y $P(S \cap W) = 0.12$. Al aplicar la ecuación (4.6), de la ley de la adición, se tiene

$$P(S \cup W) = P(S) + P(W) - P(S \cap W) = 0.30 + 0.20 - 0.12 = 0.38.$$

Así, la probabilidad de que un empleado se vaya de la empresa por el salario o por el trabajo es 0.38.

Antes de concluir el estudio de la ley de la adición se considerará un caso especial que surge cuando los **eventos son mutuamente excluyentes**.

EVENTOS MUTUAMENTE EXCLUYENTES

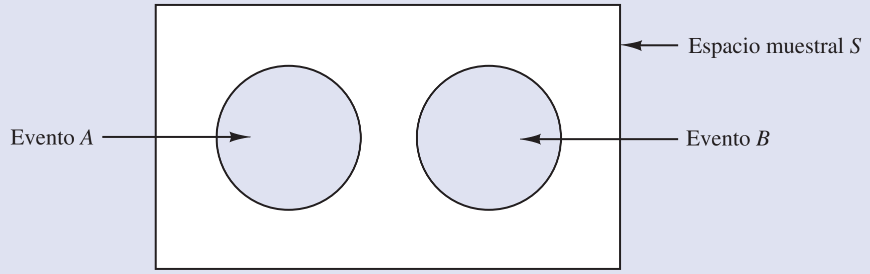
Se dice que dos eventos son mutuamente excluyentes si no tienen puntos muestrales en común.

Los eventos A y B son mutuamente excluyentes si, cuando un evento ocurre, el otro no puede ocurrir. Por tanto, para que A y B sean mutuamente excluyentes, se requiere que su intersección no contenga ningún punto muestral. En la figura 4.7 aparece el diagrama de Venn que representa dos eventos, A y B, mutuamente excluyentes. En este caso $P(A \cap B) = 0$ y la ley de la adición se expresa como sigue:

LEY DE LA ADICIÓN PARA EVENTOS MUTUAMENTE EXCLUYENTES

$$P(A \cup B) = P(A) + P(B)$$

FIGURA 4.7 EVENTOS MUTUAMENTE EXCLUYENTES



Ejercicios

Métodos

22. Suponga que tiene un espacio muestral con cinco resultados experimentales que son igualmente posibles: E_1, E_2, E_3, E_4 y E_5. Sean

$$A = \{E_1, E_2\}$$
$$B = \{E_3, E_4\}$$
$$C = \{E_2, E_3, E_5\}$$

 a. Halle $P(A)$, $P(B)$ y $P(C)$.
 b. Calcule $P(A \cup B)$. ¿A y B son mutuamente excluyentes?
 c. Estime A^c, C^c, $P(A^c)$ y $P(C^c)$.
 d. Halle $A \cup B^c$ y $P(A \cup B^c)$.
 e. Halle $P(B \cup C)$.

Auto examen

23. Suponga que se tiene el espacio muestral $S = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7\}$, donde $E_1, E_2, ..., E_7$ denotan puntos muestrales. La asignación de probabilidades es la siguiente: $P(E_1) = 0.05$, $P(E_2) = 0.20$, $P(E_3) = 0.20$, $P(E_4) = 0.25$, $P(E_5) = 0.15$, $P(E_6) = 0.10$ y $P(E_7) = 0.05$. Sea

$$A = \{E_1, E_4, E_6\}$$
$$B = \{E_2, E_4, E_7\}$$
$$C = \{E_2, E_3, E_5, E_7\}$$

 a. Halle $P(A)$, $P(B)$ y $P(C)$.
 b. Encuentre $A \cup B$ y $P(A \cup B)$.
 c. Halle $A \cap B$ y $P(A \cap B)$.
 d. ¿Los eventos A y B son mutuamente excluyentes?
 e. Halle B^c y $P(B^c)$

Aplicaciones

24. Las autoridades de Clarkson University realizaron un sondeo entre sus alumnos para conocer su opinión acerca de su universidad. Una pregunta fue si la universidad no satisface sus expectativas, si las satisface o si supera sus expectativas. Encontraron que 4% de los interrogados no dieron una respuesta, 26% respondieron que la universidad no llenaba sus expectativas y 56% indicó que la universidad superaba sus expectativas.
 a. Si toma un alumno al azar, ¿cuál es la probabilidad de que diga que la universidad supera sus expectativas?
 b. Si toma un alumno al azar, ¿cuál es la probabilidad de que diga que la universidad satisface o supera sus expectativas?

25. La Oficina de Censos de Estados Unidos cuenta con datos sobre la cantidad de adultos jóvenes, entre 18 y 24 años, que viven en casa de sus padres.* Sea

M = el evento adulto joven que vive en casa de sus padres
F = el evento adulta joven que vive en casa de sus padres

Si toma al azar un adulto joven y una adulta joven, los datos de dicha oficina permiten concluir que $P(M) = 0.56$ y $P(F) = 0.42$ (*The World Almanac,* 2006). La probabilidad de que ambos vivan en casa de sus padres es 0.24.
 a. ¿Cuál es la probabilidad de que al menos uno de dos adultos jóvenes seleccionados viva en casa de sus padres?
 b. ¿Cuál es la probabilidad de que los dos adultos jóvenes seleccionados vivan en casa de sus padres?

*En estos datos se incluye a los adultos jóvenes solteros que viven en los internados de las universidades, porque es de suponer que estos adultos jóvenes vuelven a las casas de sus padres en las vacaciones.

26. Datos sobre las 30 principales acciones y fondos balanceados proporcionan los rendimientos porcentuales anuales y a 5 años para el periodo que termina el 31 de marzo de 2000 (*The Wall Street Journal*, 10 de abril de 2000). Suponga que considera altos un rendimiento anual arriba de 50% y un rendimiento a cinco años arriba de 300%. Nueve de los fondos tienen un rendimiento anual arriba de 50%, siete de los fondos a cinco años lo tienen arriba de 300% y cinco de los fondos tienen tanto un rendimiento anual arriba de 50% como un rendimiento a cinco años arriba de 300%.
    a. ¿Cuál es la probabilidad de un rendimiento anual alto y cuál es la probabilidad de un rendimiento a cinco años alto?
    b. ¿Cuál es la probabilidad de ambos, un rendimiento anual alto y un rendimiento a cinco años alto?
    c. ¿Cuál es la probabilidad de que no haya un rendimiento anual alto ni un rendimiento a cinco años alto?

27. En una encuesta en la pretemporada de futbol americano de la NCAA 2001 se preguntó: "¿Este año habrá un equipo del Big Ten o del Pac-10 en el juego del Rose Bowl?" De los 13 429 interrogados, 2961 dijeron que habría uno del Big Ten, 4494 señalaron que habría uno del Pac-10 y 6823 expresaron que ni el Big Ten ni el Pac-10 tendría un equipo en el Rose Bowl (www.yahoo.com, 30 de agosto de 2001).
    a. ¿Cuál es la probabilidad de que el interrogado responda que ni el Big Ten ni el Pac-10 tendrán un equipo en el Rose Bowl?
    b. ¿De que afirme que el Big Ten o el Pac-10 tendrán un equipo en el campeonato Rose Bowl?
    c. Halle la probabilidad de que la respuesta sea que tanto el Big Ten como el Pac-10 tendrán un equipo en el Rose Bowl.

Autoexamen

28. En una encuesta aplicada a los suscriptores de una revista se encontró que en los últimos 12 meses 45.8% habían rentado un automóvil por razones de trabajo, 54% por razones personales y 30% por razones de trabajo y personales.
    a. ¿Cuál es la probabilidad de que un suscriptor haya rentado un automóvil en los últimos 12 meses por razones de trabajo o por razones personales?
    b. ¿Cuál es la probabilidad de que un suscriptor no haya rentado un automóvil en los últimos 12 meses ni por razones de trabajo ni por razones personales?

29. En Estados Unidos cada año hay más estudiantes con buenas calificaciones que desean inscribirse a las mejores universidades del país. Como el número de lugares permanece relativamente estable, algunas universidades rechazan solicitudes de admisión anticipadas. La universidad de Pensilvania recibió 2851 solicitudes para admisión anticipada. De éstas admitió a 1033 estudiantes, rechazó definitivamente a 854 estudiantes y dejó a 964 para el plazo de admisión normal. Esta universidad admitió a cerca de 18% de los solicitantes en el plazo normal para hacer un total (número de admisiones anticipadas más número de admisiones normales) de 2375 estudiantes (*USA Today* 24 de enero de 2001). Sean los eventos: E, un estudiante que solicita admisión anticipada es admitido; R rechazado definitivamente y D dejado para el plazo normal de admisión, sea A el evento de que un estudiante es admitido en el plazo normal.
    a. Use los datos para estimar $P(E)$, $P(R)$ y $P(D)$.
    b. ¿Son mutuamente excluyentes los eventos E y D? Halle $P(E \cap D)$.
    c. De los 2375 estudiantes admitidos en esta universidad, ¿cuál es la probabilidad de que un estudiante tomado en forma aleatoria haya tenido una admisión anticipada.
    d. Suponga que un estudiante solicita admisión anticipada en esta universidad. ¿Cuál es la probabilidad de que el estudiante tenga una admisión anticipada o en el periodo normal de admisión?

4.4 Probabilidad condicional

Con frecuencia, en la probabilidad de un evento influye el hecho de que un evento relacionado con él ya haya ocurrido. Suponga que tiene un evento A cuya probabilidad es $P(A)$**.** Si obtiene información nueva y sabe que un evento relacionado con él, denotado por B, ya ha ocurrido, deseará aprovechar esta información y volver a calcular la probabilidad del evento A. A esta nueva probabilidad del evento A se le conoce como **probabilidad condicional** y se expresa $P(A \mid B)$. La notación | indica que se está considerando la probabilidad del evento A *dada* la condición de que el evento B ha ocurrido. Por tanto, la notación $P(A \mid B)$ se lee "la probabilidad de A dado B".

Como ejemplo de la probabilidad condicional, considere el caso de las promociones de los agentes de policía de una determinada ciudad. La fuerza policiaca consta de 1200 agentes, 960 hombres y 240 mujeres. De éstos, en los últimos dos años, fueron promovidos 340. En la tabla 4.4 se muestra cómo quedaron repartidas estas promociones entre los hombres y mujeres.

Después de analizar el registro de las promociones, un comité femenil protestó, ya que habían sido promovidos 288 agentes hombres, frente a sólo 36 mujeres. Los directivos de la fuerza policiaca argumentaron que el número de mujeres promovidas no se debía a una discriminación, sino a que el número de mujeres que son agentes de policía es una cantidad pequeña. Ahora verá cómo emplear la probabilidad condicional para analizar esta acusación de discriminación.

Sean

$$
\begin{aligned}
M &= \text{el evento que un agente de policía sea hombre} \\
W &= \text{el evento que un agente de policía sea mujer} \\
A &= \text{el evento que un agente de policía sea promovido} \\
A^c &= \text{el evento que un agente de policía no sea promovido}
\end{aligned}
$$

Dividir los valores de los datos de la tabla 4.4 entre el total de agentes de policía, 1200, permite concretar la información que se tiene en las probabilidades siguientes.

$P(M \cap A) = 288/1200 = 0.24 =$ probabilidad de que un agente de policía, escogido en forma aleatoria, sea hombre *y* haya sido promovido

$P(M \cap A^c) = 672/1200 = 0.56 =$ probabilidad de que un agente de policía, escogido en forma aleatoria, sea hombre *y* no haya sido promovido

$P(W \cap A) = 36/1200 = 0.03 =$ probabilidad de que un agente de policía, escogido en forma aleatoria, sea mujer *y* haya sido promovido

$P(W \cap A^c) = 204/1200 = 0.17 =$ probabilidad de que un agente de policía, escogido en forma aleatoria, sea mujer *y* no haya sido promovido

Como cada uno de estos valores da la probabilidad de la intersección de dos eventos, se les llama **probabilidades conjuntas**. A la tabla 4.5, que proporciona la información de las probabilidades de promoción de los agentes de policía, se le conoce como *tabla de probabilidades conjuntas*.

Las cantidades que aparecen en los márgenes de una tabla de las probabilidades conjuntas son las probabilidades de cada uno de los eventos por separado. Es decir, $P(M) = 0.80$, $P(W) =$

TABLA 4.4 PROMOCIONES, EN LOS ÚLTIMOS DOS AÑOS, DE LOS AGENTES DE POLICÍA

	Hombre	Mujer	Total
Promovido	288	36	324
No promovido	672	204	876
Total	960	240	1200

TABLA 4.5 TABLA DE PROBABILIDAD CONJUNTA PARA LAS PROMOCIONES

Las probabilidades conjuntas aparecen en el cuerpo de la tabla.

	Hombre (M)	Mujer (W)	Total
Promovido (A)	0.24	0.03	0.27
No promovido (A^c)	0.56	0.17	0.73
Total	0.80	0.20	1.00

Las probabilidades marginales aparecen en los márgenes de la tabla.

0.20, $P(A) = 0.27$, $P(A^c) = 0.73$. A estas probabilidades se les conoce como **probabilidades marginales** por encontrarse en los márgenes de una tabla de probabilidad conjunta.

Observe que las probabilidades marginales se obtienen al sumar las probabilidades conjuntas del renglón o columna correspondiente de la tabla de probabilidades conjuntas. Por ejemplo, la probabilidad marginal de ser promovido es $P(A) = P(M \cap A) + P(W \cap A) = 0.24 + 0.03 = 0.27$. En las probabilidades marginales se observa que 80% de la fuerza policiaca está formada por hombres y 20% por mujeres, que 27% de los agentes de policía fueron promovidos y 73% no fueron promovidos.

Ahora empiece con el análisis de la probabilidad condicional calculando la probabilidad de que un agente de policía sea promovido dado que ese agente sea hombre. Emplee la notación para probabilidad condicional para determinar $P(A \mid M)$. Para calcular $P(A \mid M)$ se observa, primero, que esta notación sólo significa que se considera la probabilidad del evento A (promoción) ya que la condición designada como evento M (que el agente de policía sea hombre) está dada. Así que $P(A \mid M)$ indica que sólo interesan los promovidos dentro de los 960 agentes de policía que son hombres. Como 288 de los 960 agentes de policía que son hombres fueron promovidos, la probabilidad de ser promovido dado que se es un agente hombre es $288/960 = 0.30$. En otras palabras, puesto que un agente de policía es hombre, ese agente tuvo 30% de probabilidades de ser promovido en los dos últimos años.

Resultó fácil aplicar este procedimiento, ya que en la tabla 4.4 se muestra el número de agentes de policía en cada categoría. Ahora es interesante mostrar cómo calcular probabilidades condicionales, como $P(A \mid M)$, a partir de las probabilidades de eventos relacionados y no a partir de los datos de frecuencias de la tabla 4.4.

Entonces, $P(A \mid M) = 288/960 = 0.30$. Ahora, tanto el numerador como el denominador de esta fracción se dividen entre 1200, cantidad total de agentes de policía en el estudio.

$$P(A \mid M) = \frac{288}{960} = \frac{288/1200}{960/1200} = \frac{0.24}{0.80} = 0.30$$

Observe que la probabilidad condicional se obtiene de 0.24/0.80. Regrese a la tabla de probabilidad conjunta (tabla 4.5) y observe que 0.24 es la probabilidad conjunta de A y M; es decir, $P(A \cap M) = 0.24$; también que 0.80 es la probabilidad marginal de que un agente de la policía seleccionado aleatoriamente sea hombre. Es decir, $P(M) = 0.80$. Por tanto, la probabilidad condicional $P(A \mid M)$ se calcula como la razón entre $P(A \cap M)$ y la probabilidad marginal $P(M)$.

$$P(A \mid M) = \frac{P(A \cap M)}{P(M)} = \frac{0.24}{0.80} = 0.30$$

El hecho de que la probabilidad condicional se pueda calcular como la razón entre una probabilidad conjunta respecto a una probabilidad marginal proporciona la siguiente fórmula para el cálculo de la probabilidad condicional de dos eventos A y B.

PROBABILIDAD CONDICIONAL

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{4.7}$$

o

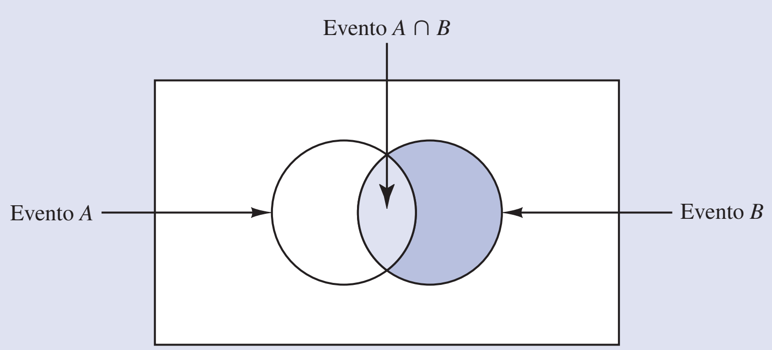
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{4.8}$$

El diagrama de Venn de la figura 4.8 ayuda a lograr una comprensión intuitiva de la probabilidad condicional. El círculo de la derecha muestra que el evento B ha ocurrido, la parte del círculo que se superpone con el evento A se denota $(A \cap B)$. Una vez que el evento B ha ocurrido, la única manera de que también sea observable el evento A es que ocurra el evento $(A \cap B)$. De manera que la razón $P(A \cap B)/P(B)$ aporta la probabilidad condicional de que se observe el evento A dado que el evento B ya ha ocurrido.

Ahora, considere de nuevo el asunto de la discriminación contra las mujeres agentes de policía. La probabilidad marginal del renglón 1 de la tabla 4.5 indica que la probabilidad de que un agente de la policía sea promovido (ya sea hombre o mujer) es $P(A) = 0.27$. Sin embargo, la cuestión relevante en el caso de la discriminación tiene que ver con las probabilidades condicionales $P(A \mid M)$ y $P(A \mid W)$. Es decir, ¿cuál es la probabilidad de que un agente de la policía sea promovido *dado que* es hombre y cuál es la probabilidad que un agente de la policía sea promovido *dado que* es mujer? Si estas dos probabilidades son iguales, no hay fundamentos para un argumento de discriminación ya que las oportunidades de ser promovidos son las mismas para agentes de la policía hombres o mujeres. Pero, si hay diferencia entre estas dos probabilidades condicionales se confirmará que los hombres y mujeres agentes de policía son considerados de manera distinta cuando se trata de las decisiones para promoverlos.

Ya se determinó que $P(A \mid M) = 0.30$. Ahora use los valores de probabilidad de la tabla 4.5 y la ecuación (4.7) de probabilidad condicional para calcular la probabilidad de que un agente de

FIGURA 4.8 PROBABILIDAD CONDICIONAL $P(A \mid B) = P(A \cap B)|P(B)$



la policía sea promovido dado que es mujer; es decir, $P(A \mid W)$. Use la ecuación (4.7) con W en lugar de B

$$P(A \mid W) = \frac{P(A \cap W)}{P(W)} = \frac{0.03}{0.20} = 0.15$$

¿Qué conclusión obtiene? La probabilidad de que un agente de policía sea promovido dado que es hombre es 0.30, el doble de 0.15, que es la probabilidad de que un agente de policía sea promovido dado que es mujer. Aunque el uso de la probabilidad condicional no demuestra por sí misma que haya discriminación en este caso, los valores de probabilidad condicional confirman el argumento presentado por las mujeres agentes de policía.

Eventos independientes

En el ejemplo anterior, $P(A) = 0.27$, $P(A \mid M) = 0.30$ y $P(A \mid W) = 0.15$. Es claro que a la probabilidad de ser promovido (evento A) le afecta o le influye el que el oficial sea un hombre o una mujer. En concreto, como $P(A \mid M) \neq P(A)$ los eventos A y M son eventos dependientes. Es decir, a la probabilidad del evento A (ser promovido) la altera o le afecta saber que se da el evento M (que el agente sea hombre). De manera similar, como $P(A \mid W) \neq P(A)$, los eventos A y W son *eventos dependientes.* Pero, si la probabilidad de un evento A no cambia por la existencia del evento M *—es decir, si* $P(A \mid M) = P(A)$—, entonces los eventos A y M son **eventos independientes**. Esto lleva a la definición de la independencia de dos eventos.

EVENTOS INDEPENDIENTES

Dos eventos A y B son independientes si

$$P(A \mid B) = P(A) \tag{4.9}$$

o

$$P(B \mid A) = P(B) \tag{4.10}$$

Si no es así, los eventos son dependientes.

Ley de la multiplicación

Mientras que la ley de las suma de probabilidades sirve para calcular la probabilidad de la unión de dos eventos, la ley de la multiplicación es útil para calcular la probabilidad de la intersección de dos eventos. La ley de la multiplicación se basa en la definición de probabilidad condicional. Al despejar en las ecuaciones (4.7) y (4.8) $P(A \cap B)$, se obtiene la **ley de la multiplicación**.

LEY DE LA MULTIPLICACIÓN

$$P(A \cap B) = P(B)P(A \mid B) \tag{4.11}$$

o

$$P(A \cap B) = P(A)P(B \mid A) \tag{4.12}$$

Para ilustrar el uso de la ley de la multiplicación, considere el caso del departamento de circulación de un periódico al que 84% de los hogares de cierta región están suscritos a la edición diaria del periódico. Si D denota el evento un hogar suscrito a la edición diaria, $P(D) = 0.84$. Además, sabe que la probabilidad de que un hogar ya suscrito a la edición diaria se suscriba también a la edición dominical (evento S) es 0.75; esto es, $P(S \mid D) = 0.75$.

¿Cuál es la probabilidad de que un hogar se subscriba a ambas, a la edición diaria y a la dominical? Emplee la ley de la multiplicación y calcule la probabilidad deseada, $P(S \cap D)$.

$$P(S \cap D) = P(D)P(S \mid D) = 0.84(0.75) = 0.63$$

Así, sabe que 63% de los hogares se suscriben a ambas ediciones, a la diaria y a la dominical.

Antes de terminar esta sección hay que considerar el caso especial de la ley de la multiplicación cuando los eventos involucrados son independientes. Recuerde que los eventos A y B son independientes si $P(A \mid B) = P(A)$ o $P(B \mid A) = P(B)$. Por tanto, con las ecuaciones (4.11) y (4.12) obtiene, para el caso especial de eventos independientes, la siguiente ley de la multiplicación.

LEY DE LA MULTIPLICACIÓN PARA EVENTOS INDEPENDIENTES

$$P(A \cap B) = P(A)P(B) \tag{4.13}$$

Para calcular la probabilidad de la intersección de dos eventos independientes, simplemente se multiplican las probabilidades correspondientes. Observe que la ley de la multiplicación para eventos independientes proporciona otra manera de determinar si dos eventos son independientes. Es decir, si $P(A \cap B) = P(A)P(B)$, entonces A y B son independientes; si $P(A \cap B) \neq P(A)P(B)$, entonces A y B son dependientes.

Como una aplicación de la ley de la multiplicación para eventos independientes considere el caso del jefe de una gasolinería que por experiencia sabe que 80% de los clientes usan tarjeta de crédito al pagar la gasolina. ¿Cuál es la probabilidad de que los dos siguientes clientes paguen la gasolina con tarjeta de crédito? Sean

A = el evento el primer cliente paga con tarjeta de crédito

B = el evento el segundo cliente paga con tarjeta de crédito

entonces el evento que interesa es $A \cap B$. Si no hay ninguna otra información, será razonable suponer que A y B son eventos independientes. Por tanto,

$$P(A \cap B) = P(A)P(B) = (0.80)(0.80) = 0.64$$

Para concluir esta sección, observe que el interés por la probabilidad condicional surgió porque los eventos suelen estar relacionados. En esos casos, los eventos son dependientes y para calcular la probabilidad de estos eventos se usan las fórmulas para probabilidad condicional de las ecuaciones (4.7) y (4.8). Si dos eventos no están relacionados, son independientes; en este caso a las probabilidades de ninguno de los eventos les afecta el hecho de que el otro evento ocurra o no.

NOTAS Y COMENTARIOS

No hay que confundir la noción de eventos mutuamente excluyentes con la de eventos independientes. Dos eventos cuyas probabilidades no son cero, no pueden ser mutuamente excluyentes e independientes. Si uno de los eventos mutuamente excluyentes ocurre, el otro evento no puede ocurrir; por tanto, la probabilidad de que ocurra el otro evento se reduce a cero.

Ejercicios

Métodos

Autoexamen

30. Suponga dos eventos, A y B, y que $P(A) = 0.50$, $P(B) = 0.60$ y $P(A \cap B) = 0.40$.
    a. Halle $P(A \mid B)$.
    b. Halle $P(B \mid A)$.
    c. ¿A y B son independientes? ¿Por qué sí o por qué no?

31. Suponga dos eventos, A y B, que son mutuamente excluyentes. Admita, además, que $P(A) = 0.30$ y $P(B) = 0.40$.
    a. Obtenga $P(A \cap B)$.
    b. Calcule $P(A \mid B)$.
    c. Un estudiante de estadística argumenta que los conceptos de eventos mutuamente excluyentes y eventos independientes son en realidad lo mismo y que si los eventos son mutuamente excluyentes deben ser también independientes. ¿Está usted de acuerdo? Use la información sobre las probabilidades para justificar su respuesta.
    d. Dados los resultados obtenidos, ¿qué conclusión sacaría usted acerca de los eventos mutuamente excluyentes e independientes?

Aplicaciones

32. Debido al aumento de los costos de los seguros, en Estados Unidos 43 millones de personas no cuentan con un seguro médico (*Time*, 1 de diciembre de 2003). En la tabla siguiente se muestran datos muestrales representativos de la cantidad de personas que cuentan con seguro médico.

		Seguro médico	
		Sí	**No**
Edad	**18 a 34**	750	170
	35 o mayor	950	130

    a. Con estos datos elabore una tabla de probabilidad conjunta y úsela para responder las preguntas restantes.
    b. ¿Qué indican las probabilidades marginales acerca de la edad de la población de Estados Unidos?
    c. ¿Cuál es la probabilidad de que una persona tomada en forma aleatoria no tenga seguro médico?
    d. Si la persona tiene entre 18 y 34 años, ¿cuál es la probabilidad de que no tenga seguro médico?
    e. Si la persona tiene 34 años o más ¿cuál es la probabilidad de que no tenga seguro médico?
    f. Si la persona no tiene seguro médico, ¿cuál es la probabilidad de que tenga entre 18 y 34 años?
    g. ¿Qué indica esta información acerca del seguro médico en Estados Unidos?

Autoexamen

33. Una muestra de estudiantes de la maestría en administración de negocios, arrojó la siguiente información sobre la principal razón que tuvieron los estudiantes para elegir la escuela en donde hacen sus estudios.

		Razones de su elección			
		Calidad de la escuela	**Costo de la escuela**	**Otras**	**Totales**
Tipo de estudiante	**Tiempo completo**	421	393	76	890
	Medio tiempo	400	593	46	1039
	Totales	821	986	122	1929

    a. Con estos datos elabore una tabla de probabilidad conjunta.
    b. Use las probabilidades marginales: calidad de la escuela, costo de la escuela y otras para comentar cuál es la principal razón por la que eligen una escuela.

c. Si es un estudiante de tiempo completo, ¿cuál es la probabilidad de que la principal razón para su elección de la escuela haya sido la calidad de la escuela?
d. Si es un estudiante de medio tiempo, ¿cuál es la probabilidad de que la principal razón para su elección de la escuela haya sido la calidad de la escuela?
e. Si *A* denota el evento es estudiante de tiempo completo y B denota el evento la calidad de la escuela fue la primera razón para su elección, ¿son independientes los eventos *A* y *B*? Justifique su respuesta.

34. La tabla siguiente muestra las probabilidades de los distintos tipos sanguíneo en la población.

	A	**B**	**AB**	**O**
Rh+	0.34	0.09	0.04	0.38
Rh−	0.06	0.02	0.01	0.06

a. ¿Cuál es la probabilidad de que una persona tenga sangre tipo O?
b. ¿De que tenga sangre Rh−?
c. ¿Cuál es la probabilidad de que una persona sea Rh− dado que la persona tiene sangre tipo O?
d. ¿Cuál es la probabilidad de que una persona tenga sangre tipo B dado que es Rh+?
e. ¿Cuál es la probabilidad de que en un matrimonio, los dos sean Rh−?
f. ¿Cuál es la probabilidad de que en un matrimonio, los dos tengan sangre AB?

35. El Departamento de Estadística Laboral de Estados Unidos reúne datos sobre las ocupaciones de las personas entre 25 y 64 años. La tabla siguiente presenta el número de hombres y mujeres (en millones) en cada una de las categorías ocupacionales.

Ocupación	**Hombres**	**Mujeres**
Directivo/Profesional	19 079	19 021
Enseñanza/Ventas/ Administrativo	11 079	19 315
Servicio	4 977	7 947
Producción con precisión	11 682	1 138
Operadores/Obrero	10 576	3 482
Agricultura/Ganadería/Silvicultura/Pesca	1 838	514

a. Desarrolle una tabla de probabilidad conjunta.
b. ¿Cuál es la probabilidad de que un trabajador mujer sea directivo o profesional?
c. ¿Cuál es la probabilidad de que un trabajador hombre esté en producción con precisión?
d. ¿Es la ocupación independiente del género? Justifique su respuesta con el cálculo de la probabilidad.

36. Reggie Miller de los Indiana Pacers tiene el record de la National Basketball Association de más canastas de 3 puntos anotadas en toda una carrera, acertando en 85% de sus tiros (*USA Today,* 22 de enero de 2004). Suponga que ya casi al final de un juego cometen una falta contra él y le conceden dos tiros.
a. ¿Cuál es la probabilidad de que acierte en los dos tiros?
b. ¿De que acierte en por lo menos uno de los dos tiros?
c. ¿De que no acierte en ninguno de los dos tiros?
d. Al final de un juego de básquetbol suele ocurrir que cometan faltas contra un jugador del equipo opuesto para detener el reloj del juego. La estrategia usual es cometer una falta contra el peor tirador del otro equipo. Suponga que el centro de los Indiana Pacers acierta 58% de sus tiros. Calcule para él las probabilidades calculadas en los incisos a, b y c y muestre que hacer una falta intencional contra el centro de los Indiana Pacers es mejor que hacerlo contra Reggie Miller.

37. Visa Card de Estados Unidos estudia con qué frecuencia usan sus tarjetas (de débito y de crédito) los consumidores jóvenes, entre 18 y 24 años. Los resultados del estudio proporcionan las probabilidades siguientes.

- La probabilidad de que un consumidor use su tarjeta al hacer una compra es 0.37.
- Dado que un consumidor usa su tarjeta, la probabilidad de que tenga entre 18 y 24 años es 0.19.
- Puesto que un consumidor usa su tarjeta, la probabilidad de que sea mayor de 24 años es 0.81.

Datos de la Oficina de Censos de Estados Unidos indican que 14% de los consumidores tienen entre 18 y 24 años.

a. Ya que un consumidor tiene entre 18 y 24 años, ¿cuál es la probabilidad de que use su tarjeta?
b. Dado que un consumidor tiene más de 24 años, ¿cuál es la probabilidad de que use su tarjeta?
c. ¿Qué interpretación se le da a las probabilidades de los incisos a y b?
d. ¿Empresas como Visa, Master Card y Discover deben proporcionar tarjetas a los consumidores entre 18 y 24 años, antes de que tengan una historia crediticia? Si no, explique. Si sí, ¿qué restricciones deben poner las empresas a estos consumidores?

38. En un estudio de Morgan Stanley Consumer Research se muestrearon hombres y mujeres y se les preguntó qué preferían tomar: agua de botella o una bebida deportiva como Gatorade o Propel Fitness (*The Atlanta Journal-Constitution*, 28 de diciembre de 2005). Suponga que en el estudio hayan participado 200 hombres y 200 mujeres y que de todos 280 hayan preferido el agua de botella. En el grupo de los que preferían bebidas deportivas, 80 eran hombres y 40 eran mujeres.

    Sea

    M = el evento el consumidor es hombre
    W = el evento el consumidor es mujer
    B = el evento el consumidor prefiere agua de botella
    S = el evento el consumidor prefiere una bebida deportiva

a. ¿Cuál es la probabilidad de que en este estudio una persona prefiera agua de botella?
b. ¿De que en este estudio una persona prefiera una bebida deportiva?
c. ¿Cuáles son las probabilidades condicionales $P(M \mid S)$ y $P(W \mid S)$?
d. ¿Cuáles son las probabilidades conjuntas $P(M \cap S)$ y $P(W \cap S)$?
e. Dado que un consumidor es hombre, ¿cuál es la probabilidad de que prefiera una bebida deportiva?
f. Ya que un consumidor es mujer, ¿cuál es la probabilidad de que prefiera una bebida deportiva?
g. ¿Depende la preferencia por una bebida deportiva de que el consumidor sea hombre o mujer? Explique usando la información sobre las probabilidades.

4.5 Teorema de Bayes

En el estudio de la probabilidad condicional vio que revisar las probabilidades cuando se obtiene más información es parte importante del análisis de probabilidades. Por lo general, se suele iniciar el análisis con una estimación de probabilidad inicial o **probabilidad previa** de los eventos que interesan. Después, de fuentes como una muestra, una información especial o una prueba del producto, se obtiene más información sobre estos eventos. Dada esta nueva información, se modifican o revisan los valores de probabilidad mediante el cálculo de probabilidades revisadas a las que se les conoce como **probabilidades posteriores**. El **teorema de Bayes** es un medio para calcular estas probabilidades. En la figura 4.9 se presentan los pasos de este proceso de revisión de la probabilidad.

FIGURA 4.9 REVISIÓN DE LA PROBABILIDAD USANDO EL TEOREMA DE BAYES

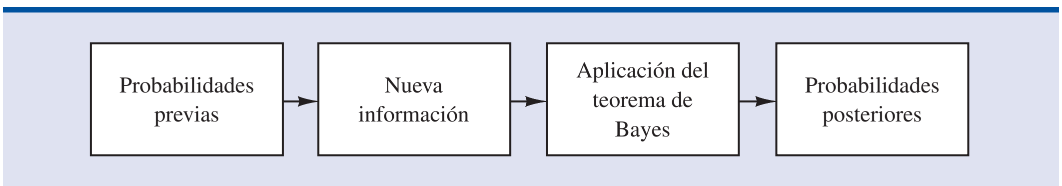


TABLA 4.6 CALIDAD DE DOS PROVEEDORES

	Porcentaje de piezas buenas	Porcentaje de piezas malas
Proveedor 1	98	2
Proveedor 2	95	5

Como aplicación del teorema de Bayes, considere una fábrica que compra piezas de dos proveedores. Sea A_1 el evento la pieza proviene del proveedor 1 y A_2 el evento la pieza proviene del proveedor 2. De las piezas que compra la fábrica, 65% proviene del proveedor 1 y 35% restante proviene del proveedor 2. Por tanto, si toma una pieza aleatoriamente, le asignará las probabilidades previas $P(A_1) = 0.65$ y $P(A_2) = 0.35$.

La calidad de las piezas compradas varía de acuerdo con el proveedor. Por experiencia, sabe que la calidad de los dos proveedores es como muestra la tabla 4.6. Si G denota el evento la pieza está buena y B denota el evento la pieza está mala, la información de la tabla 4.6 proporciona los siguientes valores de probabilidad condicional.

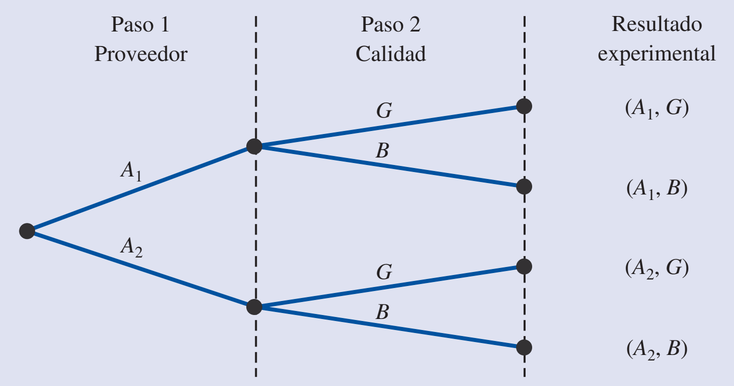
$$P(G \mid A_1) = 0.98 \qquad P(B \mid A_1) = 0.02$$
$$P(G \mid A_2) = 0.95 \qquad P(B \mid A_2) = 0.05$$

El diagrama de árbol de la figura 4.10 representa el proceso de recibir una pieza, de uno de los dos proveedores, y después determinar si la pieza es buena o mala como experimento de dos pasos. Se observa que existen cuatro resultados experimentales: dos corresponden a que la pieza esté buena y dos corresponden a que la pieza esté mala.

Cada uno de los resultados experimentales es la intersección de dos eventos, de manera que para calcular estas probabilidades puede usar la ley de la multiplicación. Por ejemplo,

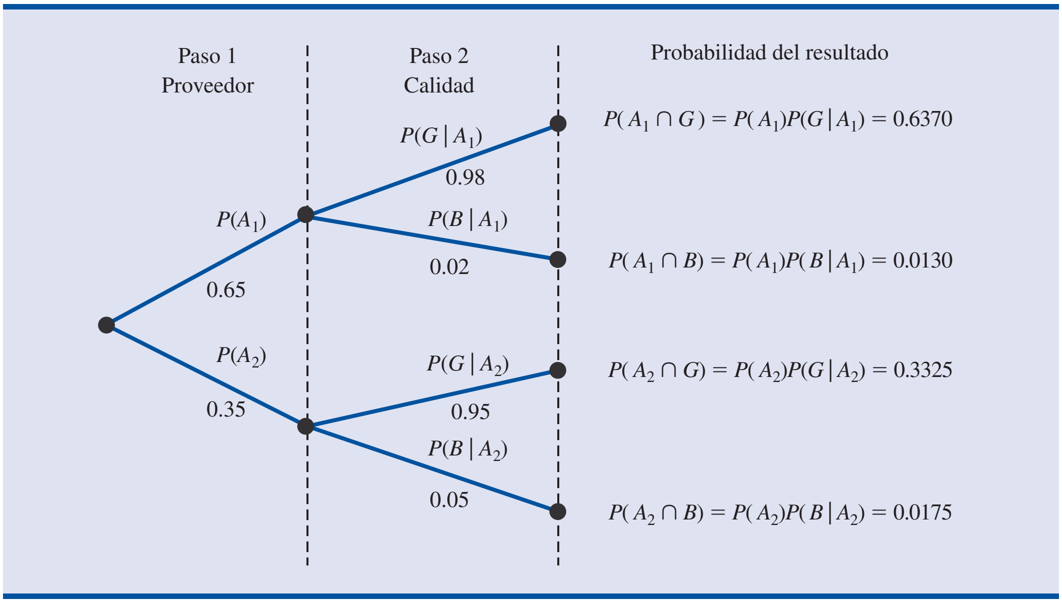
$$P(A_1, G) = P(A_1 \cap G) = P(A_1)P(G \mid A_1)$$

FIGURA 4.10 DIAGRAMA DE ÁRBOL PARA EL EJEMPLO DE LOS DOS PROVEEDORES



Nota: el paso 1 muestra que la pieza proviene de uno de los dos proveedores y el paso 2 muestra si la pieza está buena o mala.

FIGURA 4.11 ÁRBOL DE PROBABILIDAD PARA EL EJEMPLO DE LOS DOS PROVEEDORES



El proceso del cálculo de estas probabilidades conjuntas se representa mediante un árbol de probabilidad (figura 4.11). De izquierda a derecha por el árbol, las probabilidades de cada una de las ramas del paso 1 son probabilidades previas y las probabilidades de cada una de las ramas del paso 2 son probabilidades condicionales. Para hallar la probabilidad de cada uno de los resultados experimentales, simplemente se multiplican las probabilidades de las ramas que llevan a ese resultado. En la figura 4.11 se muestra cada una de estas probabilidades conjuntas junto con las probabilidades en cada rama.

Suponga ahora que las piezas de los dos proveedores se emplean en el proceso de fabricación de esta empresa y que una máquina se descompone al tratar de procesar una pieza mala. Dada la información de que la pieza está mala, ¿cuál es la probabilidad de que sea del proveedor 1 y cuál es la probabilidad de que sea del proveedor 2? Para responder estas preguntas aplique el teorema de Bayes usando la información del árbol de probabilidad (figura 4.11).

Como B es el evento la parte está mala, lo que busca son las probabilidades posteriores $P(A_1 \mid B)$ y $P(A_2 \mid B)$. De acuerdo con la ley para la probabilidad condicional

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)} \tag{4.14}$$

Del árbol de probabilidad

$$P(A_1 \cap B) = P(A_1)P(B \mid A_1) \tag{4.15}$$

Para hallar $P(B)$, se observa que B sólo puede presentarse de dos maneras: $(A_1 \cap B)$ y $(A_2 \cap B)$. Por tanto,

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) \\ &= P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) \end{aligned} \tag{4.16}$$

Sustituyendo las ecuaciones (4.15) y (4.16) en la ecuación (4.14) y expresando de manera similar $P(A_2 \mid B)$ se obtiene el teorema de Bayes para el caso de dos eventos.

Al reverendo Thomas Bayes, un ministro presbiteriano, se le atribuye la idea inicial que llevó a la versión del teorema de Bayes que se usa en la actualidad.

TEOREMA DE BAYES (CASO DE DOS EVENTOS)

$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \quad \textbf{(4.17)}$$

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \quad \textbf{(4.18)}$$

A partir de la ecuación (4.17) y los valores de probabilidad del ejemplo, se tiene

$$\begin{aligned} P(A_1 \mid B) &= \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \\ &= \frac{(0.65)(0.02)}{(0.65)(0.02) + (0.35)(0.05)} = \frac{0.0130}{0.0130 + 0.0175} \\ &= \frac{0.0130}{0.0305} = 0.4262 \end{aligned}$$

Y usando la ecuación (4.18) se encuentra $P(A_2 \mid B)$.

$$\begin{aligned} P(A_2 \mid B) &= \frac{(0.35)(0.05)}{(0.65)(0.02) + (0.35)(0.05)} \\ &= \frac{0.0175}{0.0130 + 0.0175} = \frac{0.0175}{0.0305} = 0.5738 \end{aligned}$$

Observe que al principio de este ejemplo, la probabilidad de seleccionar una pieza y que fuera del proveedor 1 era 0.65. Sin embargo, dada la información de que la pieza está mala, la probabilidad de que la pieza provenga del proveedor 1 bajó a 0.4262. En efecto, si la pieza está mala, la posibilidad de que sea del proveedor 2 es mayor que 50-50; es decir, $P(A_2 \mid B) = 0.5738$.

El teorema de Bayes es aplicable cuando los eventos para los que se quiere calcular la probabilidad revisada son mutuamente excluyentes y su unión es todo el espacio muestral.* En el caso de n eventos mutuamente excluyentes $A_1, A_2, \ldots, A_n$, cuya unión sea todo el espacio muestral, el teorema de Bayes aplica para calcular cualquiera de las probabilidades posteriores $P(A_i \mid B)$ como se muestra a continuación

TEOREMA DE BAYES

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n)} \quad \textbf{(4.19)}$$

*Si la unión de los eventos es todo el espacio muestral, los eventos son *colectivamente exhaustivos*.

Con las probabilidades previas $P(A_1)$, $P(A_2)$, …, $P(A_n)$ y las probabilidades condicionales adecuadas $P(B \mid A_1)$, $P(B \mid A_2)$, …, $P(B \mid A_n)$, se usa la ecuación (4.19) para calcular la probabilidad posterior de los eventos A_1, A_2, …, A_n

Método tabular

Para realizar los cálculos del teorema de Bayes es útil emplear un método tabular. En la tabla 4.7 se muestra este método aplicado al problema de las piezas de los proveedores. Los cálculos que se muestran ahí se realizan mediante los pasos siguientes.

Paso 1. Se harán las columnas siguientes:
Columna 1: Para los eventos mutuamente excluyentes A_i de los que quiere tener la probabilidad posterior
Columna 2: Para las probabilidades previas $P(A_i)$ de los eventos
Columna 3: Para las probabilidades condicionales $P(B \mid A_i)$ de la nueva información B dado cada evento

Paso 2. En la columna 4 se calculan las probabilidades conjuntas $P(A_i \cap B)$, de cada evento y la nueva información, empleando la ley de la multiplicación. Estas probabilidades conjuntas se encuentran multiplicando las probabilidades previas de la columna 2 por las correspondientes probabilidades condicionales de la columna 3; es decir, $P(A_i \cap B) = P(A_i)P(B \mid A_i)$.

Paso 3. Sume las probabilidades de la columna 4. Esta suma es la probabilidad de la nueva información, $P(B)$. Así, en la tabla 4.7 se ve que la probabilidad de que una pieza sea del proveedor 1 y esté mala es 0.0130 y que la probabilidad de que la pieza sea del proveedor 2 y esté mala es 0.0175. Como éstas son las únicas dos maneras de tener una pieza mala, la suma 0.0130 + 0.0175, que es 0.0305, da la probabilidad de hallar una pieza mala en las piezas recibidas de los dos proveedores.

Paso 4. En la columna 5 se calculan las probabilidades posteriores usando la relación básica de la probabilidad condicional.

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)}$$

Observe que las probabilidades conjuntas $P(A_i \cap B)$ están en la columna 4 y que la probabilidad $P(B)$ es la suma de la columna 4.

TABLA 4.7 MÉTODO TABULAR PARA LOS CÁLCULOS DEL TEOREMA DE BAYES APLICADO AL EJEMPLO DE LOS DOS PROVEEDORES

(1) Eventos A_i	(2) Probabilidades previas $P(A_i)$	(3) Probabilidades condicionales $P(B \mid A_i)$	(4) Probabilidades conjuntas $P(A_i \cap B)$	(5) Probabilidades posteriores $P(A_i \mid B)$
A_1	0.65	0.02	0.0130	0.0130/0.0305 = 0.4262
A_2	0.35	0.05	0.0175	0.0175/0.0305 = 0.5738
	1.00		$P(B)$ = 0.0305	1.0000

NOTAS Y COMENTARIOS

1. El teorema de Bayes se usa mucho en la toma de decisiones. Las probabilidades previas suelen ser estimaciones subjetivas dadas por la persona que toma las decisiones. Se obtiene información muestral y se usan las probabilidades posteriores para emplearlas en la toma de decisiones.
2. Un evento y su complemento son mutuamente excluyentes y su unión es todo el espacio muestral. Por tanto, el teorema de Bayes siempre se emplea para calcular la probabilidad posterior de un evento y su complemento.

Ejercicios

Métodos

Auto examen

39. Las probabilidades previas de los eventos A_1 y A_2 son $P(A_1) = 0.40$ y $P(A_2) = 0.60$. Sabe también que $P(A_1 \cap A_2) = 0$. Suponga que $P(B \mid A_1) = 0.20$ y $P(B \mid A_2) = 0.05$.
    a. ¿A_1 y A_2 son eventos mutuamente excluyentes? Explique.
    b. Calcule $P(A_1 \cap B)$ y $P(A_2 \cap B)$.
    c. Calcule $P(B)$.
    d. Emplee el teorema de Bayes para calcular $P(A_1 \mid B)$ y $P(A_2 \mid B)$.
40. Las probabilidades previas de los eventos A_1, A_2 y A_3 son $P(A_1) = 0.20$, $P(A_2) = 0.50$ y $P(A_3) = 0.30$. Las probabilidades condicionales del evento B dados los eventos A_1, A_2 y A_3 son $P(B \mid A_1) = 0.50$, $P(B \mid A_2) = 0.40$ y $P(B \mid A_3) = 0.30$.
    a. Calcule $P(B \cap A_1)$, $P(B \cap A_2)$ y $P(B \cap A_3)$.
    b. Emplee el teorema de Bayes, ecuación (4.19), para calcular la probabilidad posterior $P(A_2 \mid B)$.
    c. Use el método tabular para emplear el teorema de Bayes en el cálculo de $P(A_1 \mid B)$, $P(A_2 \mid B)$ y $P(A_3 \mid B)$.

Aplicaciones

41. Una empresa de consultoría presenta una oferta para un gran proyecto de investigación. El director de la firma piensa inicialmente que tiene 50% de posibilidades de obtener el proyecto. Sin embargo, mas tarde, el organismo al que se le hizo la oferta pide más información sobre la oferta. Por experiencia se sabe que en 75% de las ofertas aceptadas y en 40% de las ofertas no aceptadas, este organismo solicita más información.
    a. ¿Cuál es la probabilidad previa de que la oferta sea aceptada (es decir, antes de la solicitud dé más información)?
    b. ¿Cuál es la probabilidad condicional de que se solicite más información dado que la oferta será finalmente aceptada?
    c. Calcule la probabilidad posterior de que la oferta sea aceptada dado que se solicitó más información.

Auto examen

42. Un banco local revisa su política de tarjetas de crédito con objeto de retirar algunas de ellas. En el pasado aproximadamente 5% de los tarjetahabientes incumplieron, dejando al banco sin posibilidad de cobrar el saldo pendiente. De manera que el director estableció una probabilidad previa de 0.05 de que un tarjetahabiente no cumpla. El banco encontró también que la probabilidad de que un cliente que es cumplido no haga un pago mensual es 0.20. Por supuesto la probabilidad de no hacer un pago mensual entre los que incumplen es 1.
    a. Dado que un cliente no hizo el pago de uno o más meses, calcule la probabilidad posterior de que el cliente no cumpla.
    b. El banco deseará retirar sus tarjetas si la probabilidad de que un cliente no cumpla es mayor que 0.20. ¿Debe retirar el banco una tarjeta si el cliente no hace un pago mensual?

43. En los automóviles pequeños el rendimiento de la gasolina es mayor, pero no son tan seguros como los coches grandes. Los automóviles pequeños constituyen 18% de los vehículos en circulación, pero en accidentes con automóviles pequeños se registraron 11 898 victimas mortales en uno de los últimos años (*Reader´s Digest,* mayo de 2000). Suponga que la probabilidad de que un automóvil pequeño tenga un accidente es 0.18. La probabilidad de que en un accidente con un automóvil pequeño haya una víctima mortal es 0.128 y la probabilidad de que haya una víctima mortal si el automóvil no es pequeño es 0.05. Usted se entera de un accidente en el que hubo una víctima mortal. ¿Cuál es la probabilidad de que el accidente lo haya tenido un automóvil pequeño?

44. La American Council of Education informa que en Estados Unidos 47% de los estudiantes que ingresan en la universidad terminan sus estudios en un lapso de cinco años (Associated Press, 6 de mayo de 2002). Suponga que en los registros de terminación de estudios encuentra que 50% de los estudiantes que terminan sus estudios en cinco años son mujeres y 45% de quienes no terminan sus estudios en cinco años son mujeres. Los estudiantes que no terminan sus estudios en cinco años son estudiantes que han abandonado sus estudios o que están por terminarlos.
    a. Sea A_1 =el estudiante termina sus estudios en cinco años
       A_2 = el estudiante no termina sus estudios en cinco
       W = el estudiante es mujer
       Empleando la información dada, dé las probabilidades siguientes: $P(A_1)$, $P(A_2)$, $P(W \mid A_1)$ y $P(W \mid A_2)$.
    b. ¿Cuál es la probabilidad de que una estudiante termine sus estudios en cinco años?
    c. ¿Cuál es la probabilidad de que un estudiante termine sus estudios en cinco años?
    d. Dados los resultados anteriores, ¿cuál es el porcentaje de mujeres y cuál es el porcentaje de hombres que entran en la universidad?

45. En un artículo acerca del crecimiento de las inversiones, la revista *Money* informa que las acciones en medicamentos muestran una poderosa tendencia de largo plazo y ofrecen a los inversionistas potenciales inigualables y duraderas ganancias. La Health Care Financing Administration confirma estas conclusiones con su pronóstico de que para 2010 el consumo de medicamentos llegará a $366 mil millones, cuando en 2000 era de $117 mil millones. Muchas de las personas de 65 años o más necesitan medicamentos. Entre estas personas, 82% necesita medicamentos de manera regular, 55% usa tres o más medicamentos de manera regular y 40% necesita cinco o más medicamentos regularmente. En cambio entre las personas menores de 65 años, 49% usa medicamentos de manera regular, 37% necesita tres o más medicamentos de manera regular y 28% usa cinco o más medicamentos regularmente (*Money*, septiembre de 2001). La Oficina de Censos de Estados Unidos informa que de los 281 421 906 habitantes de Estados Unidos, 34 991 753 son personas de 65 años o mayores (U.S. Census Bureau, *Census 2000*).
    a. Calcule la probabilidad de que en Estados Unidos una persona tenga 65 años o más.
    b. Calcule la probabilidad de que una persona necesite medicamentos de manera regular.
    c. Calcule la probabilidad de que una persona tenga 65 años o más y necesite cinco o más medicamentos.
    d. Dado que una persona usa cinco o más medicamentos, calcule la probabilidad de que tenga 65 años o más.

Resumen

En este capítulo se introdujeron conceptos básicos de probabilidad y se ilustró cómo usar el análisis de probabilidad para obtener información útil para la toma de decisiones. Se describió cómo interpretar la probabilidad como una medida numérica de la posibilidad de que ocurra un evento. Además, se vio que la probabilidad de un evento se puede calcular, ya sea sumando las probabilidades de los resultados experimentales (puntos muestrales) que comprende el evento o usando las relaciones que establecen las leyes de probabilidad de la adición, de la probabilidad condicional y de la multiplicación. En el caso de que se obtenga información adicional, se mostró cómo usar el teorema de Bayes para obtener probabilidades revisadas o posteriores.

Glosario

Probabilidad Medida numérica de la posibilidad de que ocurra un evento.
Experimento Proceso para generar resultados bien definidos.

Espacio muestral Conjunto de todos los resultados experimentales.
Punto muestral Un elemento del espacio muestral. Un punto muestral que representa un resultado experimental.
Diagrama de árbol Representación gráfica que ayuda a visualizar un experimento de pasos múltiples.
Requerimientos básicos en la asignación de probabilidades Dos requerimientos que restringen la manera en que se asignan probabilidades son: 1) Para cada resultado experimental E_i se debe tener $0 \leq P(E_i) \leq 1$; 2) si $E_1, E_2, \ldots, E_n$ son todos los resultados experimentales, se debe tener que $P(E_1) + P(E_2) + \cdots + P(E_n) = 1.0$.
Método clásico Sirve para la asignación de probabilidades, es apropiado cuando todos los resultados experimentales son igualmente posibles.
Método de las frecuencias relativas Útil para la asignación de probabilidades, es conveniente cuando se tienen datos para estimar la proporción de veces que se presentará un resultado experimental si se repite un gran número de veces.
Método subjetivo Método para la asignación de probabilidades basado en un juicio.
Evento Colección de puntos muestrales
Complemento de *A* El evento que consta de todos los puntos muestrales que no están en A.
Diagrama de Venn Una representación gráfica para mostrar de manera simbólica el espacio muestral y las operaciones con eventos en la cual el espacio muestral se representa como un rectángulo y los eventos se representan como círculos dentro del espacio muestral.
Unión de *A* y *B* Evento que contiene todos los puntos muestrales que pertenecen a A o a B o a ambos. La unión se denota $A \cup B$.
Intersección de *A* y *B* Evento que contiene todos los puntos muestrales que pertenecen tanto a A como a B. La intersección se denota $A \cap B$.
Ley de la adición Ley de probabilidad que se usa para calcular la unión de dos eventos. Es $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Si los eventos son mutuamente excluyentes, $P(A \cap B) = 0$; en este caso la ley de la adición se reduce a $P(A \cup B) = P(A) + P(B)$.
Eventos mutuamente excluyentes Eventos que no tienen puntos muestrales en común; es decir, $A \cap B$ es vacío y $P(A \cap B) = 0$.
Probabilidad condicional Probabilidad de un evento dado que otro evento ya ocurrió. La probabilidad condicional de A dado B es $P(A \mid B) = P(A \cap B)/P(B)$.
Probabilidad conjunta La probabilidad de que dos eventos ocurran al mismo tiempo; es decir, la probabilidad de la intersección de dos eventos.
Probabilidad marginal Los valores en los márgenes de una tabla de probabilidad conjunta que dan las probabilidades de cada evento por separado.
Eventos independientes Son dos eventos, A y B, para los que $P(A \mid B) = P(A)$ o $P(B \mid A) = P(B)$; es decir, los eventos no tienen ninguna influencia uno en otro.
Ley de la multiplicación Una ley de probabilidad que se usa para calcular la probabilidad de la intersección de dos eventos. Esto es $P(A \cap B) = P(B)P(A \mid B)$ o $P(A \cap B) = P(\mathrm{A})P(B \mid A)$. Para eventos independientes se reduce a $P(A \cap B) = P(\mathrm{A})P(B)$
Probabilidades previas Estimaciones iniciales de las probabilidades de eventos.
Probabilidades posteriores Probabilidades revisadas de eventos basadas en informaciones adicionales.
Teorema de Bayes Método usado para calcular las probabilidades posteriores.

Fórmulas clave

Regla de conteo para combinaciones

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} \tag{4.1}$$

Regla de conteo para permutaciones

$$P_n^N = n!\binom{N}{n} = \frac{N!}{(N-n)!} \tag{4.2}$$

Cálculo de la probabilidad usando el complemento

$$P(A) = 1 - P(A^c) \tag{4.5}$$

Ley de la adición

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{4.6}$$

Probabilidad condicional

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{4.7}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{4.8}$$

Ley de la multiplicación

$$P(A \cap B) = P(B)P(A \mid B) \tag{4.11}$$

$$P(A \cap B) = P(A)P(B \mid A) \tag{4.12}$$

Ley de la multiplicación para eventos independientes

$$P(A \cap B) = P(A)P(B) \tag{4.13}$$

Teorema de Bayes

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n)} \tag{4.19}$$

Ejercicios complementarios

46. En un sondeo se les pidió a 1035 adultos su opinión respecto a los negocios (*BusinessWeek*, 11 de septiembre de 2000). Una de las preguntas era: "¿Cómo califica usted a las empresas estadounidenses respecto a la calidad de los productos y competitividad a nivel mundial?" Las respuestas fueron: excelentes, 18%; bastante buenas, 50%; regulares, 26%; malas, 5% y no saben o no contestaron 1%.
    a. ¿Cuál es la probabilidad de que un interrogado considere a las empresas estadounidenses bastante buenas o excelentes?
    b. ¿Cuántos de los interrogados consideraron malas a las empresas estadounidenses?
    c. ¿Cuántos de los interrogados dijo no saber o no contestó?
47. Un administrador financiero realiza dos nuevas inversiones, una en la industria del petróleo y otra en bonos municipales. Después de un año cada una de las inversiones se clasificará como buena o no. Considere como un experimento el resultado que se obtiene con estas dos acciones.
    a. ¿Cuántos puntos muestrales hay en este experimento?
    b. Presente un diagrama de árbol y enumere los puntos muestrales.
    c. Sea O = el evento la inversión en la industria del petróleo es buena y M = el evento la inversión en los fondos municipales es buena. Dé los puntos muestrales de O y de M.
    d. Enumere los puntos muestrales de la unión de los eventos $(O \cup M)$.
    e. Cuente los puntos muestrales de la intersección de los eventos $(O \cap M)$.
    f. ¿Son mutuamente excluyentes los eventos O y M? Explique.

48. A principios de 2003, el presidente de Estados Unidos propuso eliminar los impuestos a los dividendos de los accionistas con el argumento de que era un doble impuesto. Las corporaciones pagan impuestos sobre las ganancias que luego son repartidas como dividendos. En un sondeo realizado a 671 estadounidenses, Techno Metrica Market Intelligence halló que 47% estaban a favor de la propuesta, 44% se oponían a ella y 9% no estaban seguros (*Investor's Business Daily,* 13 de enero de 2003). Al analizar las respuestas de acuerdo con la pertenencia a los partidos políticos, se encontró en el sondeo que 29% de los demócratas estaban a favor, 64% de los republicanos estaban a favor y 48% de los independientes estaban a favor.
    a. ¿Cuántos de los encuestados estuvieron a favor de la eliminación de los impuestos a los dividendos?
    b. ¿Cuál es la probabilidad condicional de que una persona esté a favor de la propuesta dado que es demócrata?
    c. ¿Es la afiliación partidaria independiente de que una persona esté a favor de la propuesta?
    d. Si se supone que las respuestas de las personas estuvieron de acuerdo con sus propios intereses, ¿qué grupo se beneficiará más con la aceptación de la propuesta?
49. En un estudio realizado con 31 000 ingresos a hospitales en el estado de Nueva York se encontró que 4% de los ingresados sufrieron daños a causa del tratamiento. Un séptimo de estos daños condujeron a la muerte y un cuarto se debió a negligencia médica. En uno de cada 7.5 casos de negligencia médica se levantó una demanda y en una de cada dos demandas se tuvo que pagar una indemnización.
    a. ¿Cuál es la probabilidad de que una persona que ingresa en un hospital sufra un daño a causa del tratamiento debido a negligencia médica?
    b. ¿Cuál es la probabilidad de que una persona que ingresa en un hospital muera a causa de daños producidos por el tratamiento?
    c. En el caso de daños causado por negligencia médica, ¿cuál es la probabilidad de que la demanda ocasione una indemnización?
50. En una encuesta por teléfono para determinar la opinión de los televidentes respecto a un nuevo programa de televisión se obtuvieron las opiniones siguientes:

Opinión	**Frecuencia**
Malo	4
Regular	8
Bueno	11
Muy bueno	14
Excelente	13

    a. ¿Cuál es la probabilidad de que un televidente tomado aleatoriamente opine que el nuevo programa es bueno o le dé un calificativo mejor.
    b. ¿Cuál es la probabilidad de que un televidente tomado aleatoriamente opine que el nuevo programa es regular o le dé un calificativo inferior?
51. En la siguiente tabulación cruzada se muestra el ingreso familiar de acuerdo con el nivel de estudios del cabeza de familia (*Statistical Abstract of the United States: 2002*).

	Ingreso familiar (en miles de $)					
Nivel de estudios	**Menos de 25**	**25.0–49.9**	**50.0–74.9**	**75.0–99.9**	**100 o más**	**Total**
Preparatoria sin terminar	9 285	4 093	1 589	541	354	15 862
Preparatoria terminada	10 150	9 821	6 050	2 737	2 028	30 786
Estudios universitarios sin terminar	6 011	8 221	5 813	3 215	3 120	26 380
Estudios universitarios terminados	2 138	3 985	3 952	2 698	4 748	17 521
Estudios de posgrado	813	1 497	1 815	1 589	3 765	9 479
Total	28 397	27 617	19 219	10 780	14 015	100 028

a. Elabore una tabla de probabilidad conjunta.
b. ¿Cuál es la probabilidad de que el cabeza de familia no haya terminado la preparatoria?
c. ¿Cuál es la probabilidad de que el cabeza de familia haya terminado la universidad o tenga estudios de posgrado?
d. ¿Cuál es la probabilidad de que si el cabeza de familia terminó la universidad, el ingreso familiar sea $100 000 o más?
e. ¿Cuál es la probabilidad de que el ingreso familiar sea menor a $25 000?
f. ¿Cuál es la probabilidad de que una familia en la que el cabeza de familia terminó la universidad, tenga un ingreso familiar menor a $25 000?
g. ¿El ingreso familiar es independiente del nivel de educación?

52. En un estudio realizado entre los 2010 nuevos estudiantes inscritos a las maestrías de negocios se obtuvieron los datos siguientes.

		Hizo solicitudes en varias universidades	
		Sí	**No**
Grupos de edades	**23 o menos**	207	201
	24–26	299	379
	27–30	185	268
	31–35	66	193
	36 o más	51	169

a. Para un estudiante de maestría tomado en forma aleatoria elabore una tabla de probabilidad conjunta para el experimento que consiste en observar la edad del estudiante y si hizo solicitudes en varias universidades.
b. ¿Cuál es la probabilidad de que un estudiante tomado en forma aleatoria tenga 23 años o menos?
c. ¿Cuál es la probabilidad de que un estudiante tomado en forma aleatoria tenga más de 26 años?
d. ¿Cuál es la probabilidad de que un estudiante tomado en forma aleatoria haya hecho solicitud en varias universidades?

53. Vaya nuevamente a los datos de los nuevos estudiantes inscritos a las maestrías de negocios del ejercicio 52.
a. Dado que una persona hizo solicitudes en varias universidades, ¿cuál es la probabilidad de que tenga entre 24 y 26 años?
b. Ya que una persona tiene 36 años o más, ¿cuál es la probabilidad de que haya hecho solicitudes en varias universidades?
c. ¿Cuál es la probabilidad de que una persona entre 24 y 26 años haya hecho solicitudes en varias universidades?
d. Suponga que la persona sólo hizo solicitud para una universidad. ¿Cuál es la probabilidad de que la persona tenga 31 años o más?
e. ¿La edad y el hacer solicitudes en varias universidades son independientes? Explique.

54. En una encuesta realizada por IBD/TIPPP para obtener información sobre la opinión respecto a las inversiones para el retiro (*Investor´s Business Daily,* 5 de mayo de 2000) se les preguntó a los hombres y mujeres interrogados qué tan importante les parecía que era el nivel de riesgo al elegir una inversión para el retiro. Con los datos obtenidos se elaboró la siguiente tabla de probabilidades conjuntas. "Importante" significa que el interrogado respondió que el nivel de riesgo era importante o muy importante.

	Hombre	**Mujer**	**Total**
Importante	0.22	0.27	0.49
No importante	0.28	0.23	0.51
Total	0.50	0.50	1.00

a. ¿Cuál es la probabilidad de que uno de los interrogados diga que es importante?
b. ¿Cuál es la probabilidad de que una de las mujeres interrogadas diga que es importante?
c. ¿Cuál es la probabilidad de que uno de los hombres interrogados diga que es importante?
d. ¿El nivel de riesgo es independiente del género del interrogado?
e. ¿La opinión de hombres y mujeres difiere respecto al riesgo?

55. Una empresa grande de productos de consumo transmite por televisión publicidad para uno de sus jabones. De acuerdo con una encuesta realizada, se asignaron probabilidades a los eventos siguientes.

$$B = \text{una persona compra el producto}$$
$$S = \text{una persona recuerda haber visto la publicidad}$$
$$B \cap S = \text{una persona compra el producto y recuerda haber visto la publicidad.}$$

Las probabilidades fueron $P(B) = 0.20$, $P(B) = 0.40$ y $P(B \cap S) = 0.12$.

a. ¿Cuál es la probabilidad de que una persona compre el producto dado que recuerda haber visto la publicidad? ¿Ver la publicidad aumenta la probabilidad de que el individuo compre el producto? Si usted tuviera que tomar la decisión, ¿recomendaría que continuara la publicidad (suponiendo que los costos sean razonables)?
b. Si una persona que no compra el producto de la empresa compra el de la competencia. ¿Cuál sería su estimación de la participación de la empresa en el mercado? ¿Esperaría que continuando con la publicidad aumentara la participación de la empresa en el mercado? ¿Por qué sí o por qué no?
c. La empresa probó también otra publicidad y los valores de probabilidad asignados fueron $P(S) = 0.30$, $P(B \cap S) = 0.10$. Dé $P(B \mid S)$ en el caso de esta otra publicidad. ¿Qué publicidad parece tener mejor efecto en la compra de los clientes?

56. Cooper Realty es una empresa inmobiliaria pequeña que se encuentra en Albany, Nueva York y que se especializa en la venta de casas residenciales. Últimamente quiso saber cuál era la posibilidad de que una de las casas que tiene en venta se vendiera en menos de un determinado número de días. Mediante un análisis de 800 casas vendidas por la empresa en los años anteriores se obtuvieron los datos siguientes.

		Días en venta hasta la compra			
		Menos de 30	**31–90**	**Más de 90**	**Total**
Precio pedido inicialmente	**Menos de $150 000**	50	40	10	100
	$150 000–$199 999	20	150	80	250
	$200 000–$250 000	20	280	100	400
	Más de $250 000	10	30	10	50
	Total	100	500	200	800

a. Si A se define como el evento de que la casa esté en venta más de 90 días antes de ser vendida, estime la probabilidad de A.
b. Si B se define como el evento de que el precio inicial sea menor que $150 000, estime la probabilidad de B.
c. ¿Cuál es la probabilidad de $A \cap B$?
d. Suponga que se acaba de firmar un contrato para vender una casa en un precio inicial menor que $150 000, ¿cuál es la probabilidad de que a Cooper Realty le tome menos de 90 días venderla?
e. ¿Los eventos A y B son independientes?

57. Una empresa estudió el número de accidentes ocurridos en su planta de Brownsville, Texas. De acuerdo con información anterior, 6% de los empleados sufrieron accidentes el año pasado. Los directivos creen que un programa especial de seguridad reducirá este año los accidentes a 5%. Se estima además que 15% de los empleados que sufrieron un accidente el año pasado tendrán un accidente este año.
    a. ¿Qué porcentaje de los empleados sufrirá accidentes en los dos años?
    b. ¿Qué porcentaje de los empleados sufrirá por lo menos un accidente en este periodo de dos años?
58. El departamento de recolección de impuestos de Estados Unidos en Dallas, preocupado por las declaraciones de impuestos fraudulentas, cree que la probabilidad de hallar una declaración de impuestos fraudulenta, dado que la declaración contiene deducciones que exceden el estándar, es 0.20. Dado que las deducciones no exceden el estándar, la probabilidad de una declaración fraudulenta disminuye a 0.02. Si 8% de las declaraciones exceden el estándar de deducciones, ¿cuál es la mejor estimación del porcentaje de declaraciones fraudulentas?
59. Una empresa petrolera compra una opción de tierra en Alaska. Los estudios geológicos preliminares asignaron las probabilidades previas siguientes.

$$P(\text{petróleo de alta calidad}) = 0.50$$
$$P(\text{petróleo de calidad media}) = 0.20$$
$$P(\text{que no haya petróleo}) = 0.30$$

    a. ¿Cuál es la probabilidad de hallar petróleo?
    b. Después de 200 pies de perforación en el primer pozo, se toma una prueba de suelo. Las probabilidades de hallar el tipo de suelo identificado en la prueba son las siguientes.

$$P(\text{suelo} \mid \text{petróleo de alta calidad}) = 0.20$$
$$P(\text{suelo} \mid \text{petróleo de calidad media}) = 0.80$$
$$P(\text{suelo} \mid \text{que no haya petróleo}) = 0.20$$

    ¿Cómo debe interpretar la empresa la prueba de suelo? ¿Cuáles son las probabilidades revisadas y cuáles son las nuevas probabilidades de hallar petróleo?
60. Las empresas que hacen negocios por Internet suelen obtener información acerca del visitante de un sitio Web a partir de los sitios visitados previamente. El artículo "Internet Marketing" (*Interfaces,* marzo/abril de 2001) describe cómo los datos sobre el flujo de clics en los sitios Web visitados se usan junto a un modelo de actualización Bayesiano para determinar el género de una persona que visita la Web. ParFore creó un sitio Web para la venta de equipo y ropa para golf. A los directivos de la empresa les gustaría que apareciera una determinada oferta para los visitantes del sexo femenino y otra oferta determinada para los visitantes del sexo masculino. En una muestra de visitas anteriores al sitio Web se sabe que 60% de las personas que visitan el sitio son hombres y 40% mujeres.
    a. ¿Cuál es la probabilidad previa de que el siguiente visitante del sitio Web sea mujer?
    b. Suponga que el actual visitante de ParFore.com visitó previamente el sitio de la Web de Dillard, y que es tres veces más probable que ese sitio sea visitado por mujeres que por hombres. ¿Cuál es la probabilidad revisada de que el visitante actual de ParFore.com sea mujer? ¿Desplegaría la oferta que está dirigida más a hombres o a mujeres?

Caso problema Los jueces del condado de Hamilton

Los jueces del condado de Hamilton llevan miles de casos cada año. En su inmensa mayoría la sentencia queda dictada. Sin embargo, en algunos casos hay apelaciones y algunas apelaciones revocan la sentencia. Kristen DelGuzzi de *The Cincinnati Enquirer* realizó, durante tres años, un estudio sobre los casos llevados por los jueces del condado de Hamilton. En la tabla 4.8 se muestran los resultados de los 182 908 casos llevados por 38 jueces en tribunales de primera instan-

archivo en CD
Judge

TABLA 4.8 CASOS DESPACHADOS, APELADOS Y REVOCADOS EN LOS TRIBUNALES DEL CONDADO DE HAMILTON

Tribunal de primera instancia

Juez	Casos despachados	Casos apelados	Casos revocados
Fred Cartolano	3 037	137	12
Thomas Crush	3 372	119	10
Patrick Dinkelacker	1 258	44	8
Timothy Hogan	1 954	60	7
Robert Kraft	3 138	127	7
William Mathews	2 264	91	18
William Morrissey	3 032	121	22
Norbert Nadel	2 959	131	20
Arthur Ney Jr.	3 219	125	14
Richard Niehaus	3 353	137	16
Thomas Nurre	3 000	121	6
John O'Connor	2 969	129	12
Robert Ruehlman	3 205	145	18
J. Howard Sundermann	955	60	10
Ann Marie Tracey	3 141	127	13
Ralph Winkler	3 089	88	6
Total	43 945	1762	199

Tribunal de relaciones domésticas

Juez	Casos despachados	Casos apelados	Casos revocados
Penelope Cunningham	2 729	7	1
Patrick Dinkelacker	6 001	19	4
Deborah Gaines	8 799	48	9
Ronald Panioto	12 970	32	3
Total	30 499	106	17

Tribunal municipal

Juez	Casos despachados	Casos apelados	Casos revocados
Mike Allen	6 149	43	4
Nadine Allen	7 812	34	6
Timothy Black	7 954	41	6
David Davis	7 736	43	5
Leslie Isaiah Gaines	5 282	35	13
Karla Grady	5 253	6	0
Deidra Hair	2 532	5	0
Dennis Helmick	7 900	29	5
Timothy Hogan	2 308	13	2
James Patrick Kenney	2 798	6	1
Joseph Luebbers	4 698	25	8
William Mallory	8 277	38	9
Melba Marsh	8 219	34	7
Beth Mattingly	2 971	13	1
Albert Mestemaker	4 975	28	9
Mark Painter	2 239	7	3
Jack Rosen	7 790	41	13
Mark Schweikert	5 403	33	6
David Stockdale	5 371	22	4
John A. West	2 797	4	2
Total	108 464	500	104

cia, tribunales de relaciones domésticas y tribunales municipales. Dos de los jueces (Dinkelacker y Hogan) no prestaron sus servicios en el mismo tribunal durante los tres años completos.

El objetivo del estudio de este periódico fue evaluar el trabajo de los jueces. Las apelaciones suelen ser el resultado de errores cometidos por los jueces, y el periódico deseaba saber qué jueces realizan bien su trabajo y qué jueces cometían demasiados errores. Se le solicita su ayuda para realizar el análisis de datos. Emplee sus conocimientos de probabilidad y de probabilidad condicional para ayudar a la clasificación de los jueces. Podrá analizar también la posibilidad de apelación y de revocación en los casos tratados en los distintos tribunales.

Informe administrativo

Elabore un informe con su clasificación de los jueces. Incluya un análisis de la posibilidad de apelación y de revocación del caso en los tres tribunales. Como mínimo su informe debe contener lo siguiente:

1. La probabilidad de que los casos sean apelados y revocados en los distintos tribunales.
2. La probabilidad, para cada juez, de que un caso sea apelado.
3. La probabilidad, para cada juez, de que un caso sea revocado.
4. La probabilidad, para cada juez, de revocación dada una apelación.
5. Clasifique a los jueces de cada tribunal de mejor a peor. Dé el criterio que usa y proporcione el fundamento que justifique su elección.

CAPÍTULO 5

Distribuciones de probabilidad discreta

CONTENIDO

LA ESTADÍSTICA *en* LA PRÁCTICA

CITIBANK*
LONG ISLAND CITY, NUEVA YORK

Citibank, una división de Citigroup, proporciona una amplia gama de servicios financieros, que comprende cuentas de cheques y de ahorro, préstamos e hipotecas, seguros y servicios de inversión, todos dentro del marco de una estrategia única llamada Citibanking. Citibanking significa una identidad de marca consistente en todo el mundo, una oferta coherente de productos y servicios de calidad para el cliente. Citibanking permite al cliente disponer de dinero en cualquier momento, en cualquier parte y de la manera que lo desee. Ya sea que el cliente desee ahorrar para el futuro o solicitar un préstamo para hoy, lo puede hacer en Citibank.

Los cajeros automáticos de Citibank, localizados en los Citicard Banking Center (CBC), permiten al cliente hacer todas sus operaciones bancarias en un solo lugar con un simple toque de su dedo, 24 horas al día y 7 días a la semana. Más de 150 operaciones bancarias diferentes, desde depósitos hasta manejo de inversiones, pueden ser realizadas con facilidad. Los cajeros automáticos Citibanking son mucho más que un simple cajero automático y en la actualidad los clientes realizan en ellos 80% de sus transacciones.

Cada Citibank CBC opera como un sistema de espera en línea al que los clientes llegan en forma aleatoria a solicitar el servicio de uno de los cajeros automáticos. Si todos los cajeros automáticos están ocupados, debe esperar en la fila. Con periodicidad realizan estudios acerca de la capacidad de los CBC para determinar los tiempos de espera para el cliente y establecer si son necesarios más cajeros automáticos.

Los datos recolectados por Citibank muestran que la llegada aleatoria de los clientes sigue una distribución de probabilidad conocida como distribución de Poisson. Mediante la distribución de Poisson, Citibank calcula las probabilidades de que llegue un número determinado de clientes a un CBC durante un determinado periodo y decidir cuál es el número de cajeros que necesita. Por ejemplo, sea x la cantidad de clientes que llega en un periodo de un minuto. Suponga que la tasa media de llegadas de clientes a un determinado CBC es dos clientes por minuto, la tabla siguiente da las probabilidades de que llegue un determinado número de clientes por minuto.

Un vanguardista cajero automático de Citibank.

x	**Probabilidad**
0	0.1353
1	0.2707
2	0.2707
3	0.1804
4	0.0902
5 o más	0.0527

Las distribuciones de probabilidad discretas como la empleada por Citibank, son el tema de este capítulo. Además de la distribución de Poisson, verá las distribuciones binomial e hipergeométrica; conocerá también cómo emplear estas distribuciones de probabilidad para obtener información de utilidad.

*Los autores agradecen a Stacey Karter, Citibank, por proporcionarnos este artículo para *La estadística en práctica*.

En este capítulo se continúa con el estudio de la probabilidad introduciendo los conceptos de variable aleatoria y distribuciones de probabilidad. El punto sustancial de este capítulo son las distribuciones de probabilidad discreta de tres distribuciones de probabilidad discreta que serán estudiadas son: la binomial, la de Poisson y la hipergeométrica.

5.1 Variables aleatorias

En el capítulo 4 se definió el concepto de experimento con sus correspondientes resultados experimentales. Una variable aleatoria proporciona un medio para describir los resultados experimen-

VARIABLE ALEATORIA

Una **variable aleatoria** es una descripción numérica del resultado de un experimento.

Las variables aleatorias deben tomar valores numéricos.

tales empleando valores numéricos. Las variables aleatorias deben tomar valores numéricos. En efecto, una variable aleatoria asocia un valor numérico a cada uno de los resultados experimentales. El valor numérico de la variable aleatoria depende del resultado del experimento. Una variable aleatoria puede ser *discreta* o *continua*, depende del tipo de valores numéricos que asuma.

Variables aleatorias discretas

A una variable aleatoria que asuma ya sea un número finito de valores o una sucesión infinita de valores tales como 0, 1, 2, . . ., se le llama **variable aleatoria discreta**. Considere, por ejemplo, el siguiente experimento: un contador presenta el examen para certificarse como contador público. El examen tiene cuatro partes. Defina una variable aleatoria x como $x =$ número de partes del examen aprobadas. Ésta es una variable aleatoria discreta porque puede tomar el número finito de valores 0, 1, 2, 3 o 4.

Para tener otro ejemplo de una variable aleatoria discreta considere el experimento de observar los automóviles que llegan a una caseta de peaje. La variable aleatoria que interesa es $x =$ número de automóviles que llega a la caseta de peaje en un día. Los valores que puede tomar la variable aleatoria son los de la secuencia 0, 1, 2, etc. Así, x es una variable aleatoria discreta que toma uno de los valores de esta sucesión infinita.

Aunque los resultados de muchos experimentos se describen mediante valores numéricos, los de otros no. Por ejemplo, en una encuesta se le puede preguntar a una persona si recuerda el mensaje de un comercial de televisión. Este experimento tiene dos resultados: que la persona no recuerda el mensaje y que la persona recuerda el mensaje. Sin embargo, estos resultados se describen numéricamente definiendo una variable aleatoria x como sigue: sea $x = 0$ si la persona no recuerda el mensaje y sea $x = 1$ si la persona recuerda el mensaje. Los valores numéricos de esta variable son arbitrarios (podría haber usado 5 y 10), pero son aceptables de acuerdo con la definición de una variable aleatoria, es decir, x es una variable aleatoria porque proporciona una descripción numérica de los resultados del experimento.

En la tabla 5.1 aparecen algunos otros ejemplos de variables aleatorias discretas. Observe que en cada ejemplo la variable aleatoria discreta asume un número finito de valores o asume los valores de una secuencia infinita como 0, 1, 2, ... Este tipo de variables aleatorias discretas se estudia con detalle en este capítulo.

TABLA 5.1 EJEMPLOS DE VARIABLES ALEATORIAS DISCRETAS

Experimento	Variable aleatoria (x)	Valores posibles para la variable aleatoria
Llamar a cinco clientes	Número de clientes que hacen un pedido	0, 1, 2, 3, 4, 5
Inspeccionar un envío de 50 radios	Número de radios que tienen algún defecto	$0, 1, 2, \cdots, 49, 50$
Hacerse cargo de un restaurante durante un día	Número de clientes	$0, 1, 2, 3, \cdots$
Vender un automóvil	Sexo del cliente	0 si es hombre; 1 si es mujer

Variables aleatorias continuas

A una variable que puede tomar cualquier valor numérico dentro de un intervalo o colección de intervalos se le llama **variable aleatoria continua**. Los resultados experimentales basados en escalas de medición tales como tiempo, peso, distancia y temperatura pueden ser descritos por variables aleatorias continuas. Considere, por ejemplo, el experimento de observar las llamadas telefónicas que llegan a la oficina de atención de una importante empresa de seguros. La variable aleatoria que interesa es $x =$ tiempo en minutos entre dos llamadas consecutivas. Esta variable aleatoria puede tomar cualquier valor en el intervalo $x \geq 0$. En efecto, x puede tomar un número infinito de valores, entre los que se encuentran valores como 1.26 minutos, 2.751 minutos, 4.3333 minutos, etc. Otro ejemplo, considere el tramo de 90 millas de una carretera entre Atlanta y Georgia. Para el servicio de ambulancia de emergencia en Atlanta, la variable aleatoria x es $x =$ número de millas hasta el punto en que se localiza el siguiente accidente de tráfico en este tramo de la carretera. En este caso, x es una variable aleatoria continua que toma cualquier valor en el intervalo $0 \leq x \leq 90$. En la tabla 5.2 aparecen otros ejemplos de variables aleatorias continuas. Observe que cada ejemplo describe una variable aleatoria que toma cualquier valor dentro de un intervalo de valores. Las variables aleatorias continuas y sus distribuciones de probabilidad serán tema del capítulo 6.

TABLA 5.2 EJEMPLOS DE VARIABLES ALEATORIAS CONTINUAS

Experimento	Variable aleatoria (x)	Valores posibles para la variable aleatoria
Operar un banco	Tiempo en minutos entre la llegada de los clientes	$x \geq 0$
Llenar una lata de refresco (máx. 12.1 onzas)	Cantidad de onzas	$0 \leq x \leq 12.1$
Construir una biblioteca	Porcentaje del proyecto terminado en seis meses	$0 \leq x \leq 100$
Probar un proceso químico nuevo	Temperatura a la que tiene lugar la reacción deseada (min. 150°F; máx. 212°F)	$150 \leq x \leq 212$

NOTAS Y COMENTARIOS

Un modo de determinar si una variable aleatoria es discreta o continua es imaginar los valores de la variable aleatoria como puntos sobre un segmento de recta. Elegir dos puntos que representen valores de la variable aleatoria. Si todo el segmento de recta entre esos dos puntos representa también valores posibles para la variable aleatoria, entonces la variable aleatoria es continua.

Ejercicios

Métodos

Autoexamen

1. Considere el experimento que consiste en lanzar una moneda dos veces.
   a. Enumere los resultados experimentales.
   b. Defina una variable aleatoria que represente el número de caras en los dos lanzamientos.
   c. Dé el valor que la variable aleatoria tomará en cada uno de los resultados experimentales.
   d. ¿Es una variable aleatoria discreta o continua?

2. Considere el experimento que consiste en un empleado que arma un producto.
   a. Defina la variable aleatoria que represente el tiempo en minutos requerido para armar el producto.
   b. ¿Qué valores toma la variable aleatoria?
   c. ¿Es una variable aleatoria discreta o continua?

Aplicaciones

Auto examen

3. Tres estudiantes agendan entrevistas para un empleo de verano en el Brookwood Institute. En cada caso el resultado de la entrevista será una oferta de trabajo o ninguna oferta. Los resultados experimentales se definen en términos de los resultados de las tres entrevistas.
   a. Enumere los resultados experimentales.
   b. Defina una variable aleatoria que represente el número de ofertas de trabajo. ¿Es una variable aleatoria continua?
   c. Dé el valor de la variable aleatoria que corresponde a cada uno de los resultados experimentales.
4. Suponga que conoce la tasa hipotecaria de 12 instituciones de préstamo. La variable aleatoria que interesa es el número de las instituciones de préstamo en este grupo que ofrecen una tasa fija a 30 años de 8.5% o menos. ¿Qué valores toma esta variable aleatoria?
5. Para realizar cierto análisis de sangre, los técnicos laboratoristas tienen que llevar a cabo dos procedimientos. En el primero requieren uno o dos pasos y en el segundo requieren uno, dos o tres pasos.
   a. Enumere los resultados experimentales correspondientes a este análisis de sangre.
   b. Si la variable aleatoria que interesa es el número de pasos requeridos en todo el análisis (los dos procedimientos), dé los valores que toma la variable aleatoria en cada uno de los resultados experimentales.
6. A continuación se da una serie de experimentos y su variable aleatoria correspondiente. En cada caso determine qué valores toma la variable aleatoria y diga si se trata de una variable aleatoria discreta o continua.

Experimento	Variable aleatoria (x)
a. Hacer un examen con 20 preguntas	Número de preguntas contestadas correctamente
b. Observar los automóviles que llegan a una caseta de peaje en 1 hora	Número de automóviles que llegan a la caseta de peaje
c. Revisar 50 declaraciones de impuestos	Número de declaraciones que tienen algún error
d. Observar trabajar a un empleado	Número de horas no productivas en una jornada de 8 horas
e. Pesar un envío	Número de libras

5.2 Distribuciones de probabilidad discreta

La **disribución de probabilidad** de una variable aleatoria describe cómo se distribuyen las probabilidades entre los valores de la variable aleatoria. En el caso de una variable aleatoria discreta x, la distribución de probabilidad está definida por una **función de probabilidad**, denotada por $f(x)$. La función de probabilidad da la probabilidad de cada valor de la variable aleatoria.

Como ejemplo de una variable aleatoria discreta y de su distribución de probabilidad, considere las ventas de automóviles en DiCarlo Motors en Saratoga, Nueva York. Durante los últimos 300 días de operación, los datos de ventas muestran que hubo 57 días en los que no se vendió ningún automóvil, 117 días en los que se vendió 1 automóvil, 72 días en los que se vendieron 2 automóviles, 42 días en los que se vendieron 3 automóviles, 12 días en los que se vendieron 4 automóviles y 3 días en los que se vendieron 5 automóviles. Suponga que considera el experimento de seleccionar un día de operación en DiCarlo Motors y se define la variable aleatoria de interés como x = número de automóviles vendidos en un día. De acuerdo con datos del pasado, se sabe que x es una variable aleatoria discreta que puede tomar los valores 0, 1, 2, 3, 4 o 5. En la notación de funciones de probabilidad $f(0)$ da la probabilidad de vender 0 automóviles, $f(1)$ da la probabilidad de vender 1 automóvil, y así en lo sucesivo. Como los datos del pasado indican que en 54 de 300 días se vendieron 0 automóviles, a $f(0)$ se le asigna el valor $54/300 = 0.18$, lo que significa que la probabilidad de que se vendan 0 automóviles en un día es 0.18. De manera similar, como en 117 de los 300 días se vendió un automóvil, a $f(1)$ se le asigna el valor $117/300 = 0.39$, que significa que la probabilidad de que se venda exactamente 1 automóvil en un día es 0.39. Continuando de esta manera con los demás valores de la variable aleatoria, se obtienen los valores de $f(2)$, $f(3)$, $f(4)$ y $f(5)$, valores que se muestran en la tabla 5.3, que es la distribución de probabilidad para el número de automóviles vendidos en un día en DiCarlo Motors.

Una ventaja importante de definir una variable aleatoria y su correspondiente distribución de probabilidad es que una vez que se conoce la distribución de probabilidad, es relativamente fácil determinar la probabilidad de diversos eventos que pueden ser útiles para tomar decisiones. Por ejemplo, empleando la distribución de probabilidad de DiCarlo Motors, tabla 5.3, se observa que el número de automóviles que es más probable vender en un día es 1, ya que es $f(1) = 0.39$. Además se observa que la probabilidad de vender tres o más automóviles en un día es $f(3) + f(4) + f(5) = 0.14 + 0.04 + 0.01 = 0.19$. Estas probabilidades, junto con otras que pueden interesar para tomar decisiones, proporcionan información que sirve de ayuda al encargado de la toma de decisiones para entender la venta de automóviles en DiCarlo Motors.

Al elaborar una función de probabilidad para una variable aleatoria discreta, deben satisfacerse las dos condiciones siguientes.

Estas condiciones son análogas a los dos requerimientos básicos, presentados en el capítulo 4, para asignar probabilidades a los resultados experimentales.

CONDICIONES REQUERIDAS PARA UNA FUNCIÓN DE PROBABILIDAD DISCRETA

$$f(x) \geq 0 \tag{5.1}$$

$$\Sigma f(x) = 1 \tag{5.2}$$

En la tabla 5.3 se observa que las probabilidades de la variable aleatoria x satisfacen la ecuación (5.1); para todos los valores de x, $f(x)$ es mayor o igual que 0; además, como estas probabilidades suman 1, también se satisface la ecuación (5.2). Por tanto, la función de probabilidad de DiCarlo Motors es una función de probabilidad discreta válida.

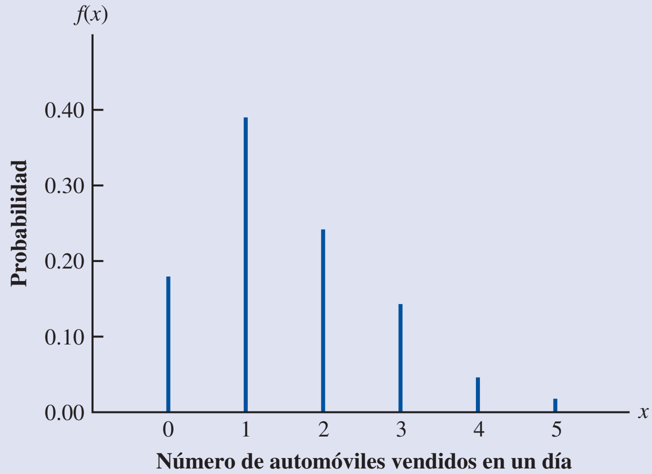
Las distribuciones de probabilidad también se representan gráficamente. En la figura 5.1, en el eje horizontal aparecen los valores de la variable aleatoria x para el caso de DiCarlo Motors y en el eje vertical aparecen las probabilidades correspondientes a estos valores.

Además de tablas y gráficas, para describir las funciones de probabilidad se suele usar una fórmula que da el valor de la función de probabilidad, $f(x)$, para cada valor x. El ejemplo más sencillo

TABLA 5.3 DISTRIBUCIÓN DE PROBABILIDAD PARA EL NÚMERO DE AUTOMÓVILES VENDIDOS EN UN DÍA EN DICARLO MOTORS

x	$f(x)$
0	0.18
1	0.39
2	0.24
3	0.14
4	0.04
5	0.01
Total	1.00

FIGURA 5.1 REPRESENTACIÓN GRÁFICA DE LA DISTRIBUCIÓN DE PROBABILIDAD DEL NÚMERO DE AUTOMÓVILES VENDIDOS EN UN DÍA EN DICARLO MOTORS



de una distribución de probabilidad discreta dada mediante una fórmula es la **distribución de probabilidad uniforme discreta**. Su función de probabilidad está definida por la ecuación (5.3).

FUNCIÓN DE PROBABILIDAD UNIFORME DISCRETA

$$f(x) = 1/n \tag{5.3}$$

donde

n = número de valores que puede tomar la variable aleatoria.

Por ejemplo, si en el experimento que consiste en lanzar un dado se define una variable aleatoria x como el número de puntos en la cara del dado que cae hacia arriba. En este experimento la variable aleatoria toma $n = 6$ valores; $x = 1, 2, 3, 4, 5, 6$. Por tanto, la función de probabilidad de esta variable aleatoria uniforme discreta es

$$f(x) = 1/6 \qquad x = 1, 2, 3, 4, 5, 6$$

Los valores de la variable aleatoria con sus probabilidades correspondientes se presentan a continuación.

x	$f(x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Otro ejemplo, la variable aleatoria x tiene la siguiente distribución de probabilidad discreta.

x	$f(x)$
1	1/10
2	2/10
3	3/10
4	4/10

Esta distribución de probabilidad se define mediante la fórmula

$$f(x) = \frac{x}{10} \quad \text{para } x = 1, 2, 3 \text{ o } 4$$

Si evalúa $f(x)$ para un valor determinado de la variable aleatoria obtiene la probabilidad correspondiente. Por ejemplo, con la función de probabilidad dada arriba se ve que $f(2) = 2/10$ da la probabilidad de que la variable aleatoria tome el valor 2.

Las funciones de probabilidad discreta más empleadas suelen especificarse mediante fórmulas. Tres casos importantes son las distribuciones binomial, de Poisson e hipergeométrica; estas distribuciones se estudian más adelante en este capítulo

Ejercicios

Métodos

Autoexamen

7. A continuación se presenta la distribución de probabilidad de una variable aleatoria x.

x	$f(x)$
20	0.20
25	0.15
30	0.25
35	0.40

   a. ¿Es válida esta distribución de probabilidad?
   b. ¿Cuál es la probabilidad de que $x = 30$?
   c. ¿Cuál es la probabilidad de que x sea menor o igual que 25?
   d. ¿Cuál es la probabilidad de que x sea mayor que 30?

Aplicaciones

Autoexamen

8. Los datos siguientes se obtuvieron contando el número de salas de operaciones de un hospital que fueron usadas en un periodo de 20 días. Tres de estos 20 días sólo se usó una sala de operaciones, cinco de estos 20 días se usaron dos, ocho de estos 20 días se usaron tres salas de operaciones y cuatro de estos 20 días se usaron las cuatro salas de operaciones del hospital.
   a. Use el método de las frecuencias relativas para elaborar una distribución de probabilidad para el número de salas de operaciones usadas en un día.
   b. Elabore una gráfica a partir de la distribución de probabilidad.
   c. Muestre que la distribución de probabilidad elaborada satisface las condiciones requeridas para una distribución de probabilidad.

9. En Estados Unidos 38% de los niños de cuarto grado no pueden leer un libro adecuado a su edad. La tabla siguiente muestra, de acuerdo con las edades, el número de niños que tienen problemas de lectura. La mayoría de estos niños tienen problemas de lectura que debieron ser detectados y corregidos antes del tercer grado.

Edad	**Número de niños**
6	37 369
7	87 436
8	160 840
9	239 719
10	286 719
11	306 533
12	310 787
13	302 604
14	289 168

Si desea tomar una muestra de niños que tienen problemas de lectura para que participen en un programa que mejora las habilidades de lectura. Sea x la variable aleatoria que indica la edad de un niño tomado en forma aleatoria.

a. Con estos datos elabore una distribución de probabilidad para x. Especifique los valores de la variable aleatoria y los correspondientes valores de la función de probabilidad $f(x)$.
b. Trace la gráfica de esta distribución de probabilidad.
c. Muestre que la distribución de probabilidad satisface las ecuaciones (5.1) y (5.2).

10. En la tabla 5.4 se muestra la distribución de frecuencias porcentuales para las puntuaciones dadas a la satisfacción con el trabajo por una muestra de directivos en sistemas de información de nivel alto y de nivel medio. Las puntuaciones van de 1 (muy insatisfecho) a 5 (muy satisfecho).

TABLA 5.4 DISTRIBUCIÓN DE FRECUENCIA PORCENTUAL DE LAS PUNTUACIONES DADAS POR DIRECTIVOS DE NIVEL ALTO Y DE NIVEL MEDIO A LA SATISFACCIÓN CON EL TRABAJO

Puntuación de la satisfacción con el trabajo	**Directivos de alto nivel**	**Directivos de nivel medio**
1	5	4
2	9	10
3	3	12
4	42	46
5	41	28

a. Elabore una distribución de probabilidad con las puntuaciones dadas a la satisfacción con el trabajo por los directivos de nivel alto.
b. Elabore una distribución de probabilidad con las puntuaciones dadas a la satisfacción con el trabajo por los directivos de nivel medio.
c. ¿Cuál es la probabilidad de que un ejecutivo de nivel alto dé una puntuación de 4 o 5 a su satisfacción con el trabajo?
d. ¿Cuál es la probabilidad de que un ejecutivo de nivel medio esté muy satisfecho?
e. Haga una comparación entre la satisfacción con el trabajo de los ejecutivos de nivel alto y la que tienen los ejecutivos de nivel medio.

11. Un técnico da servicio a máquinas franqueadoras de empresas en el área de Phoenix. El servicio puede durar 1, 2, 3 o 4 horas dependiendo del tipo de falla. Los distintos tipos de fallas se presentan aproximadamente con la misma frecuencia.

a. Elabore una distribución de probabilidad de las duraciones de los servicios.
b. Elabore una gráfica de la distribución de probabilidad.
c. Muestre que la distribución de probabilidad que ha elaborado satisface las condiciones requeridas para ser una distribución de probabilidad discreta.
d. ¿Cuál es la probabilidad de que un servicio dure tres horas?
e. Acaba de llegar una solicitud de servicio y no se sabe cuál es el tipo de falla. Son las 3:00 p.m. y los técnicos de servicio salen a las 5:00 de la tarde. ¿Cuál es la probabilidad de que el técnico de servicio tenga que trabajar horas extras para reparar la máquina hoy?

12. El jefe del departamento de admisión de una universidad calcula subjetivamente una distribución de probabilidad para x, el número de estudiantes que ingresarán en la universidad. A continuación se presenta esta distribución de probabilidad.

x	$f(x)$
1000	0.15
1100	0.20
1200	0.30
1300	0.25
1400	0.10

a. ¿Es válida esta distribución de probabilidad? Explique.
b. ¿Cuál es la probabilidad de que ingresen 1200 o menos estudiantes? Explique.

13. Un psicólogo encuentra que el número de sesiones necesarias para ganarse la confianza de un paciente es 1, 2 o 3. Sea x la variable aleatoria que representa el número de sesiones necesarias para ganarse la confianza de un paciente. Se ha propuesto la función de probabilidad siguiente.

$$f(x) = \frac{x}{6} \qquad \text{para } x = 1, 2 \text{ o } 3$$

a. ¿Es válida esta función de probabilidad? Explique.
b. ¿Cuál es la probabilidad de que se necesiten exactamente 2 sesiones para ganarse la confianza del paciente?
c. ¿De que se necesiten por lo menos 2 sesiones para ganarse la confianza del paciente?

14. La tabla siguiente es una distribución parcial de probabilidades para las ganancias proyectadas de MRA Company (x ganancias en miles de dólares) durante el primer año de operación (los valores negativos indican pérdida).

x	$f(x)$
−100	0.10
0	0.20
50	0.30
100	0.25
150	0.10
200	

a. ¿Cuál es el valor adecuado para $f(200)$? ¿Qué interpretación le da a este valor?
b. ¿Cuál es la probabilidad de que la empresa sea rentable?
c. ¿Cuál es la probabilidad de que la empresa gane por lo menos $100 000?

5.3 Valor esperado y varianzas

Valor esperado

El **valor esperado**, o media, de una variable aleatoria es una medida de la localización central de la variable aleatoria. A continuación se da la fórmula para obtener el valor esperado de una variable aleatoria x.

El valor esperado es un promedio ponderado de los valores que toma la variable aleatoria. Los pesos son las probabilidades.

VALOR ESPERADO DE UNA VARIABLE ALEATORIA DISCRETA

$$E(x) = \mu = \Sigma x f(x) \tag{5.4}$$

Las dos notaciones $E(x)$ y μ se usan para denotar el valor esperado de una variable aleatoria x.

La ecuación (5.4) indica que para calcular el valor esperado de una variable aleatoria discreta se multiplica cada valor de la variable aleatoria por su probabilidad correspondiente $f(x)$ y después se suman estos productos. Usando el ejemplo de la sección 5.2 sobre las ventas de automóviles en DiCarlo Motors, en la tabla 5.5 se muestra cómo se calcula el valor esperado del número de automóviles vendidos en un día. La suma de las entradas en la columna $xf(x)$ indica que el valor esperado es 1.50 automóviles por día. Por tanto, aunque se sabe que en un día las ventas pueden ser de 0, 1, 2, 3, 4 o 5 automóviles, DiCarlo prevé que a la larga se venderán 1.50 automóviles por día. Si en un mes hay 30 días de operación, el valor esperado, 1.50, se emplea para pronosticar que las ventas promedio mensuales serán de 30(1.5) = 45 automóviles.

El valor esperado no tiene que ser un valor que pueda tomar la variable aleatoria.

Varianza

Aunque el valor esperado proporciona el valor medio de una variable aleatoria, también suele ser necesaria una medida de la variabilidad o dispersión. Así como en el capítulo 3 se usó la varianza para resumir la variabilidad de los datos, ahora se usa la **varianza** para resumir la variabilidad en los valores de la variable aleatoria. A continuación se da la fórmula para calcular la

La varianza es un promedio ponderado de los cuadrados de las desviaciones de una variable aleatoria de su media. Los pesos son las probabilidades.

VARIANZA DE UNA VARIABLE ALEATORIA DISCRETA

$$\text{Var}(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x) \tag{5.5}$$

TABLA 5.5 CÁLCULO DEL VALOR ESPERADO PARA EL NÚMERO DE AUTOS QUE SE VENDEN EN UN DÍA EN DICARLO MOTORS

x	$f(x)$	$xf(x)$
0	0.18	0(.18) = 0.00
1	0.39	1(.39) = 0.39
2	0.24	2(.24) = 0.48
3	0.14	3(.14) = 0.42
4	0.04	4(.04) = 0.16
5	0.01	5(.01) = 0.05
		1.50

$E(x) = \mu = \Sigma x f(x)$

TABLA 5.6 CÁLCULO DE LA VARIANZA PARA EL NÚMERO DE AUTOS QUE SE VENDEN EN UN DÍA EN DICARLO MOTORS

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2 f(x)$
0	$0 - 1.50 = -1.50$	2.25	0.18	$2.25(0.18) = 0.4050$
1	$1 - 1.50 = -0.50$	0.25	0.39	$0.25(0.39) = 0.0975$
2	$2 - 1.50 = 0.50$	0.25	0.24	$0.25(0.24) = 0.0600$
3	$3 - 1.50 = 1.50$	2.25	0.14	$2.25(0.14) = 0.3150$
4	$4 - 1.50 = 2.50$	6.25	0.04	$6.25(0.04) = 0.2500$
5	$5 - 1.50 = 3.50$	12.25	0.01	$12.25(0.01) = 0.1225$
				1.2500

$\sigma^2 = \Sigma(x - \mu)^2 f(x)$

varianza de una variable aleatoria. Como indica la ecuación (5.5), un parte esencial de la fórmula de la varianza es la desviación $x - \mu$, la cual mide qué tan alejado del valor esperado, o media μ, se encuentra un valor determinado de la variable aleatoria. Para calcular la varianza de una variable aleatoria, estas desviaciones se elevan al cuadrado y después se ponderan con el correspondiente valor de la función de probabilidad. A la suma de estas desviaciones al cuadrado, ponderadas, se le conoce como *varianza*. Para denotar la varianza de una variable aleatoria se usan las notaciones Var(x) y σ^2.

En la tabla 5.6 aparece en forma resumida el cálculo de la varianza de la distribución de probabilidad del número de automóviles vendidos en un día en DiCarlo Motors. Como ve, la varianza es 1.25. La **desviación estándar**, σ, se define como la raíz cuadrada positiva de la varianza. Por tanto, la desviación estándar del número de automóviles vendidos en un día es

$$\sigma = \sqrt{1.25} = 1.118$$

La desviación estándar se mide en las mismas unidades que la variable aleatoria ($\sigma = 1.1180$ automóviles) y por tanto suele preferirse para describir la variabilidad de una variable aleatoria. La varianza σ^2 se mide en unidades al cuadrado por lo que es más difícil de interpretar.

Ejercicios

Métodos

15. La tabla siguiente muestra la distribución de probabilidad de una variable aleatoria x.

x	$f(x)$
3	0.25
6	0.50
9	0.25

a. Calcule $E(x)$, el valor esperado de x.
b. Calcule σ^2, la varianza de x.
c. Calcule σ, la desviación estándar de x.

Auto examen

16. La tabla siguiente muestra la distribución de probabilidad de una variable aleatoria y.

y	$f(y)$
2	0.20
4	0.30
7	0.40
8	0.10

a. Calcule $E(y)$.
b. Calcule $\mathrm{Var}(y)$ y σ.

Aplicaciones

17. Una ambulancia de voluntarios realiza de 0 a 5 servicios por día. A continuación se presenta la distribución de probabilidad de los servicios por día.

Número de servicios	Probabilidad	Número de servicios	Probabilidad
0	0.10	3	0.20
1	0.15	4	0.15
2	0.30	5	0.10

a. ¿Cuál es el valor esperado del número de servicios?
b. ¿Cuál es la varianza del número de servicios? ¿Cuál es la desviación estándar?

Auto examen

18. Los datos siguientes son el número de recámaras en casas rentadas y en casas propias en ciudades centrales de Estados Unidos (www.census.gov, 31 de marzo de 2003).

	Número de casas (en miles)	
Recámaras	**Rentadas**	**Propias**
0	547	23
1	5012	541
2	6100	3832
3	2644	8690
4 o más	557	3783

a. Defina una variable aleatoria x = número de recámaras en casas rentadas y elabore una distribución de probabilidad para esta variable. ($x = 4$ representará 4 recámaras o más.)
b. Calcule el valor esperado y la varianza del número de recámaras en casas rentadas.
c. Defina una variable aleatoria y = número de recámaras en casas propias y elabore una distribución de probabilidad para esta variable. ($y = 4$ representará 4 recámaras o más.)
d. Calcule el valor esperado y la varianza del número de recámaras en casas propias.
e. ¿Qué observaciones resultan al comparar el número de recámaras en casas rentadas y en casas propias?

19. La National Basketball Association (NBA) lleva diversas estadísticas de cada equipo. Dos se refieren al porcentaje de tiros de campo hechos por un equipo y el porcentaje de tiros de tres puntos hechos por un equipo. En parte de la temporada del 2004, el registro de tiros de los 29 equipos de la NBA indicaba que la probabilidad de anotar dos puntos en un tiro de campo era 0.44, y que la probabilidad de anotar tres puntos en un tiro de tres puntos era 0.34 (www.nba.com, 3 de enero de 2004).

a. ¿Cuál es el valor esperado para un tiro de dos puntos de estos equipos?
b. ¿Cuál es el valor esperado para un tiro de tres puntos de estos equipos?
c. Si la probabilidad de hacer un tiro de dos puntos es mayor que la probabilidad de hacer uno de tres puntos, ¿por qué los entrenadores permiten a algunos jugadores hacer un tiro de tres puntos si tienen oportunidad? Use el valor esperado para explicar su respuesta.

20. A continuación se presenta la distribución de probabilidad para los daños pagados por una empresa de seguros para automóviles, en seguros contra choques.

Pago	Probabilidad
0	0.85
500	0.04
1 000	0.04
3 000	0.03
5 000	0.02
8 000	0.01
10 000	0.01

a. Use el pago esperado para determinar la prima en el seguro de choques que le permitirá a la empresa cubrir los gastos.
b. La empresa de seguros cobra una tasa anual de $520 por la cobertura de choques. ¿Cuál es el valor esperado de un seguro de choques para un asegurado? (*Indicación*: son los pagos esperados de la empresa menos el costo de cobertura.) ¿Por qué compran los asegurados un seguro de choques con este valor esperado?

21. La siguiente distribución de probabilidad sobre puntuaciones dadas a la satisfacción con el trabajo por una muestra de directivos de alto nivel y de nivel medio en sistemas de la información va desde 1 (muy insatisfecho) hasta 5 (muy satisfecho).

	Probabilidad	
Puntuación de la satisfacción con el trabajo	Directivo de nivel alto	Directivo de nivel medio
1	0.05	0.04
2	0.09	0.10
3	0.03	0.12
4	0.42	0.46
5	0.41	0.28

a. ¿Cuál es el valor esperado en las puntuaciones dadas a la satisfacción con el trabajo por los ejecutivos de nivel alto?
b. ¿Cuál es el valor esperado en las puntuaciones dadas a la satisfacción con el trabajo por los directivos de nivel medio?
c. Calcule la varianza de las puntuaciones dadas a la satisfacción con el trabajo por los directivos de nivel medio.
d. Calcule la desviación estándar de las puntuaciones dadas a la satisfacción con el trabajo en las dos distribuciones de probabilidad.
e. Compare la satisfacción con el trabajo de los directivos de alto nivel con la que tienen los directivos de nivel medio.

22. La demanda de un producto de una empresa varía enormemente de mes a mes. La distribución de probabilidad que se presenta en la tabla siguiente, basada en los datos de los dos últimos años, muestra la demanda mensual de la empresa.

Demanda unitaria	Probabilidad
300	0.20
400	0.30
500	0.35
600	0.15

a. Si la empresa basa las órdenes mensuales en el valor esperado de la demanda mensual, ¿cuál será la cantidad ordenada mensualmente por la empresa para este producto?
b. Suponga que cada unidad demandada genera \$70 de ganancia y que cada unidad ordenada cuesta \$50. ¿Cuánto ganará o perderá la empresa en un mes si coloca una orden con base en su respuesta al inciso a y la demanda real de este artículo es de 300 unidades?

23. El estudio 2002 New York City Housing and Vacancy Survey indicó que había 59 324 viviendas con renta controlada y 236 263 unidades con renta estabilizada construidas en 1947 o después. A continuación se da la distribución de probabilidad para el número de personas que viven en estas unidades (www.census.gov, 12 de enero de 2004).

Número de personas	**Renta controlada**	**Renta estabilizada**
1	0.61	0.41
2	0.27	0.30
3	0.07	0.14
4	0.04	0.11
5	0.01	0.03
6	0.00	0.01

a. ¿Cuál es el valor esperado para el número de personas que viven en cada tipo de unidad?
b. ¿Cuál es la varianza para el número de personas que viven en cada tipo de unidad?
c. Haga comparaciones entre el número de personas que viven en una unidad de renta controlada y el número de personas que viven en una unidad de renta estabilizada.

24. J. R. Ryland Computer Company está considerando hacer una expansión a la fábrica para empezar a producir una nueva computadora. El presidente de la empresa debe determinar si hacer un proyecto de expansión a mediana gran escala. La demanda del producto nuevo es incierta, la cual, para los fines de planeación puede ser demanda pequeña, mediana o grande. Las probabilidades estimadas para la demanda son 0.20, 0.50 y 0.30, respectivamente. Con x y y representando ganancia anual en miles de dólares, los encargados de planeación en la empresa elaboraron el siguiente pronóstico de ganancias para los proyectos de expansión a mediana y gran escala.

		Ganancia con la expansión a mediana escala		**Ganancia con la expansión a gran escala**	
		x	$f(x)$	y	$f(y)$
	Baja	50	0.20	0	0.20
Demanda	**Mediana**	150	0.50	100	0.50
	Alta	200	0.30	300	0.30

a. Calcule el valor esperado de las ganancias correspondientes a las dos alternativas de expansión. ¿Cuál de las decisiones se prefiere para el objetivo de maximizar la ganancia esperada?
b. Calcule la varianza de las ganancias correspondientes a las dos alternativas de expansión. ¿Cuál de las decisiones se prefiere para el objetivo de minimizar el riesgo o la incertidumbre?

5.4 Distribución de probabilidad binomial

La distribución de probabilidad binomial es una distribución de probabilidad que tiene muchas aplicaciones. Está relacionada con un experimento de pasos múltiples al que se le llama experimento binomial.

Un experimento binomial

Un **experimento binomial** tiene las cuatro propiedades siguientes.

PROPIEDADES DE UN EXPERIMENTO BINOMIAL

1. El experimento consiste en una serie de n ensayos idénticos.
2. En cada ensayo hay dos resultados posibles. A uno de estos resultados se le llama *éxito* y al otro se le llama *fracaso*.
3. La probabilidad de éxito, que se denota p, no cambia de un ensayo a otro. Por ende, la probabilidad de fracaso, que se denota $1 - p$, tampoco cambia de un ensayo a otro.
4. Los ensayos son independientes.

Jacob Bernoulli (1654-1705), el primero de la familia Bernoulli de matemáticos suizos, publicó un tratado sobre probabilidad que contenía la teoría de las permutaciones y de las combinaciones, así como el teorema del binomio.

Si se presentan las propiedades 2, 3 y 4, se dice que los ensayos son generados por un proceso de Bernoulli. Si, además, se presenta la propiedad 1, se trata de un experimento binomial. En la figura 5.2 se presenta una sucesión de éxitos y fracasos de un experimento binomial con ocho ensayos.

En un experimento binomial lo que interesa es el *número de éxitos en* n *ensayos*. Si x denota el número de éxitos en n ensayos, es claro que x tomará los valores 0, 1, 2, 3, ..., n. Dado que el número de estos valores es finito, x es una variable aleatoria *discreta*. A la distribución de probabilidad correspondiente a esta variable aleatoria se le llama **distribución de probabilidad binomial**. Por ejemplo, considere el experimento que consiste en lanzar una moneda cinco veces y observar si la cara de la moneda que cae hacia arriba es cara o cruz. Suponga que se desea contar el número de caras que aparecen en los cinco lanzamientos. ¿Presenta este experimento las propiedades de un experimento binomial? ¿Cuál es la variable aleatoria que interesa? Observe que:

1. El experimento consiste en cinco ensayos idénticos; cada ensayo consiste en lanzar una moneda.
2. En cada ensayo hay dos resultados posibles: cara o cruz. Se puede considerar cara como éxito y cruz como fracaso.
3. La probabilidad de éxito y la probabilidad de fracaso son iguales en todos los ensayos, siendo $p = 0.5$ y $1 - p = 0.5$.
4. Los ensayos o lanzamientos son independientes porque al resultado de un ensayo no afecta a lo que pase en los otros ensayos o lanzamientos.

FIGURA 5.2 UNA POSIBLE SUCESIÓN DE ÉXITOS Y FRACASOS EN UN EXPERIMENTO BINOMIAL DE OCHO ENSAYOS

Propiedad 1: El experimento consiste en $n = 8$ ensayos idénticos.

Propiedad 2: En cada ensayo se obtiene como resultado un éxito o un fracaso.

Ensayos →	1	2	3	4	5	6	7	8
Resultados →	*S*	*F*	*F*	*S*	*S*	*F*	*S*	*S*

Por tanto, se satisfacen las propiedades de un experimento binomial. La variable aleatoria que interesa es x = número de caras que aparecen en cinco ensayos. En este caso, x puede tomar los valores 0, 1, 2, 3, 4 o 5.

Otro ejemplo, considere a un vendedor de seguros que visita a 10 familias elegidas en forma aleatoria. El resultado correspondiente de la visita a cada familia se clasifica como éxito si la familia compra un seguro y como fracaso si la familia no compra ningún seguro. Por experiencia, el vendedor sabe que la probabilidad de que una familia tomada aleatoriamente compre un seguro es 0.10. Al revisar las propiedades de un experimento binomial aparece que:

1. El experimento consiste en 10 ensayos idénticos; cada ensayo consiste en visitar a una familia.
2. En cada ensayo hay dos resultados posibles: la familia compra un seguro (éxito) o la familia no compra ningún seguro (fracaso).
3. Las probabilidades de que haya compra y de que no haya compra se supone que son iguales en todas las visitas, siendo $p = 0.10$ y $1 - p = 0.90$.
4. Los ensayos son independientes porque las familias se eligen en forma aleatoria.

Como estos cuatro puntos se satisfacen, este ejemplo es un experimento binomial. La variable aleatoria que interesa es el número de ventas al visitar a las 10 familias. En este caso los valores que puede tomar x son 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10.

La propiedad 3 de un experimento binomial se llama *suposición de estacionaridad* y algunas veces se confunde con la propiedad 4, independencia de los ensayos. Para ver la diferencia entre estas dos propiedades, reconsidere el caso del vendedor que visita a las familias para venderles un seguro. Si a medida que el día avanza, el vendedor se va cansando y va perdiendo entusiasmo, la probabilidad de éxito puede disminuir, por ejemplo, a 0.05 en la décima llamada. En tal caso la propiedad 3 (estacionaridad) no se satisface, y no se tiene un experimento binomial. Incluso si la propiedad 4 se satisface —en cada familia la decisión de comprar o no se hizo de manera independiente— si no se satisface la propiedad 3, no se trata de un experimento binomial.

En las aplicaciones de los experimentos binomiales se emplea una fórmula matemática llamada *función de probabilidad binomial* que sirve para calcular la probabilidad de x éxitos en n ensayos. Empleando los conceptos de probabilidad presentados en el capítulo 4, se mostrará, en el contexto de un ilustrativo problema, cómo se desarrolla la fórmula.

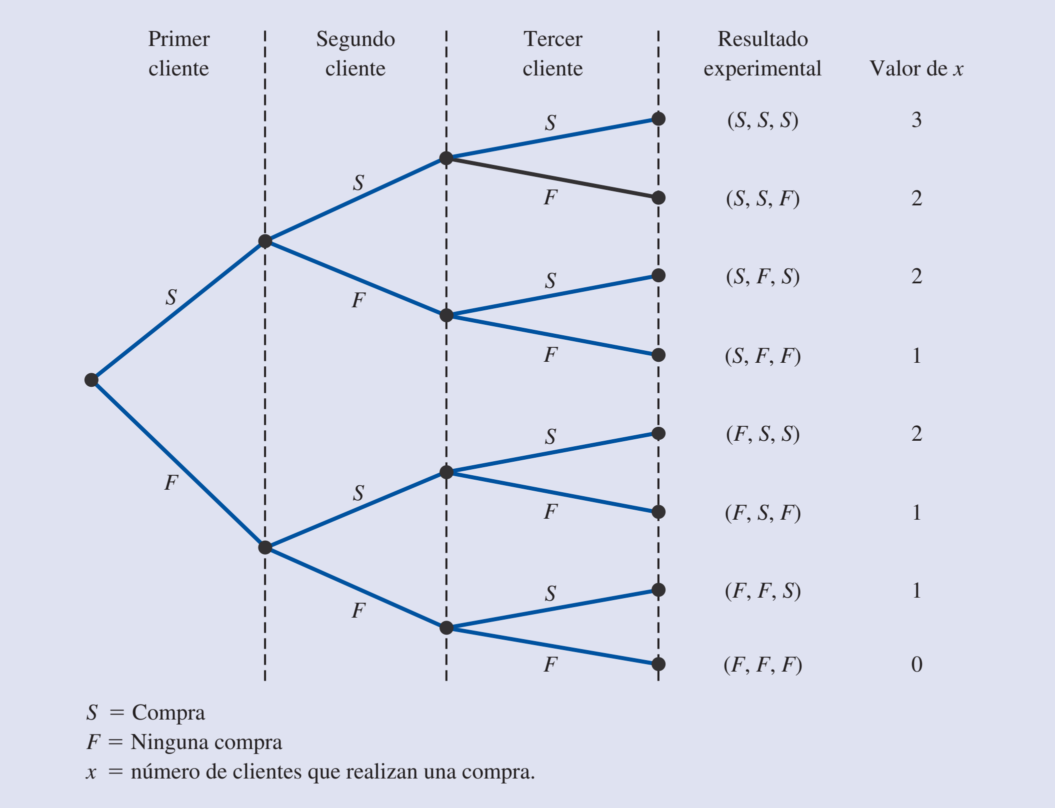
El problema de la tienda de ropa Martin Clothing Store

Considere las decisiones de compra de los próximos tres clientes que lleguen a la tienda de ropa Martin Clothing Store. De acuerdo con la experiencia, el gerente de la tienda estima que la probabilidad de que un cliente realice una compra es 0.30. ¿Cuál es la probabilidad de que dos de los próximos tres clientes realicen una compra?

Un diagrama de árbol (figura 5.3), permite advertir que el experimento de observar a los tres clientes para ver si cada uno de ellos decide realizar una compra tiene ocho posibles resultados. Entonces, si S denota éxito (una compra) y F fracaso (ninguna compra), lo que interesa son los resultados experimentales en los que haya dos éxitos (decisiones de compra) en los tres ensayos. A continuación verifique que el experimento de las tres decisiones de compra es un experimento binomial. Al verificar los cuatro requerimientos de un experimento binomial, se observa que:

1. Es posible describir el experimento como una serie de tres ensayos idénticos, un ensayo por cada uno de los tres clientes que llegan a la tienda.
2. Cada ensayo tiene dos posibles resultados: el cliente hace una compra (éxito) o el cliente no hace ninguna compra (fracaso).
3. La probabilidad de que el cliente haga una compra (0.30) o de que no haga una compra (0.70) se supone que es la misma para todos los clientes.
4. La decisión de comprar de cada cliente es independiente de la decisión de comprar de los otros clientes.

FIGURA 5.3 DIAGRAMA DE ÁRBOL PARA EL PROBLEMA DE LA TIENDA DE ROPA MARTIN CLOTHING STORE



En consecuencia, se satisfacen las propiedades de un experimento binomial.

Con la fórmula siguiente* se calcula el número de resultados experimentales en los que hay exactamente x éxitos en n ensayos.

NÚMERO DE RESULTADOS EXPERIMENTALES EN LOS QUE HAY EXACTAMENTE x ÉXITOS EN n ENSAYOS

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{5.6}$$

donde

$$n! = n(n-1)(n-2)\cdots(2)(1)$$

y por definición,

$$0! = 1$$

Ahora regrese al experimento de las decisiones de compra de tres clientes de la tienda Martin Clothing Store. La ecuación (5.6) sirve para determinar el número de resultados experimentales

* Esta fórmula presentada en el capítulo 4, determina el número de combinaciones de n objetos tomados de x a la vez. En el experimento binomial esta fórmula combinatoria da el número de resultados experimentales (series de n ensayos) en los que hay x éxitos.

en los que hay dos compras; el número de maneras en que son posibles $x = 2$ éxitos en $n = 3$ ensayos. De acuerdo con la ecuación (5.6)

$$\binom{n}{x} = \binom{3}{2} = \frac{3!}{2!(3-2)!} = \frac{(3)(2)(1)}{(2)(1)(1)} = \frac{6}{2} = 3$$

La ecuación (5.6) indica que en tres de los resultados experimentales hay dos éxitos. En la figura 5.3 aparecen denotados por (*S, S, F*), (*S, F, S*) y (*F, S, S*).

Empleando la ecuación (5.6) para determinar en cuántos resultados experimentales hay tres éxitos (compras) en tres ensayos, se obtiene

$$\binom{n}{x} = \binom{3}{3} = \frac{3!}{3!(3-3)!} = \frac{3!}{3!0!} = \frac{(3)(2)(1)}{3(2)(1)(1)} = \frac{6}{6} = 1$$

El único resultado experimental con tres éxitos es el identificado por (*S, S, S*) mostrado en la figura 5.3.

Ya sabe que usando la ecuación (5.6) es posible determinar el número de resultados experimentales en los que hay x éxitos. Sin embargo, si va a determinar la probabilidad de x éxitos en n ensayos, es necesario conocer también la probabilidad correspondiente a cada uno de estos resultados experimentales. Como en un experimento binomial, los ensayos son independientes, para hallar la probabilidad de una determinada sucesión de éxitos y fracasos simplemente se multiplican las probabilidades correspondientes al resultado de cada ensayo.

La probabilidad de que los dos primeros clientes compren y el tercero no compre, denotada por (*S, S, F*) está dada por

$$pp(1-p)$$

Puesto que la probabilidad de compra en cualquier ensayo es 0.30, la probabilidad de que haya una compra en los dos primeros ensayos y que no haya compra en el tercer ensayo es

$$(0.30)(0.30)(0.70) = (0.30)^2(0.70) = 0.063$$

Hay otros dos resultados experimentales en los que también se obtienen dos éxitos y un fracaso. A continuación se presentan las probabilidades de los tres resultados experimentales en los que hay dos éxitos.

Resultados de los ensayos 1er. cliente	 2o. cliente	 3er. cliente	Resultado experimental	Probabilidad de este resultado experimental
Compra	Compra	No hay compra	(*S, S, F*)	$pp(1-p) = p^2(1-p)$ $= (0.30)^2(0.70) = 0.063$
Compra	Compra	Compra	(*S, F, S*)	$p(1-p)p = p^2(1-p)$ $= (0.30)^2(0.70) = 0.063$
No hay compra	Compra	Compra	(*F, S, S*)	$(1-p)pp = p^2(1-p)$ $= (0.30)^2(0.70) = 0.063$

Observe que los tres resultados experimentales en los que hay dos éxitos tienen la misma probabilidad. Esto se cumple en general. En cualquier experimento binomial todas las series de resultados de ensayos en las que hay x éxitos en n ensayos tienen la *misma probabilidad* de ocurrencia. A continuación se presenta la probabilidad de cada una de las series de ensayos en las que hay x éxitos en n ensayos.

$$\begin{array}{l}\text{Probabilidad de una} \\ \text{determinada serie de} \\ \text{resultados de ensayos}\end{array} = p^x(1 - p)^{(n-x)}$$

En el caso de la tienda de ropa Martin Clothing Store, esta fórmula indica que la probabilidad de cualquier resultado experimental con dos éxitos es $p^2(1 - p)^{(3 - 2)} = p^2(1 - p)^1 = (0.30)^2(0.70^1) = 0.63$.

Como la ecuación (5.6) da el número de resultados de un experimento binomial en el que hay x éxitos, y la ecuación (5.7) da la probabilidad de cada serie en la que hay x éxitos, combinando las ecuaciones (5.6) y (5.7) se obtiene la **función de probabilidad binomial** siguiente.

FUNCIÓN DE PROBABILIDAD BINOMIAL

$$f(x) = \binom{n}{x} p^x(1 - p)^{(n-x)} \tag{5.8}$$

donde

$$\begin{aligned} f(x) &= \text{probabilidad de } x \text{ éxitos en } n \text{ ensayos} \\ n &= \text{número de ensayos} \\ \binom{n}{x} &= \frac{n!}{x!(n - x)!} \\ p &= \text{probabilidad de un éxito en cualquiera de los ensayos} \\ 1 - p &= \text{probabilidad de un fracaso en cualquiera de los ensayos} \end{aligned}$$

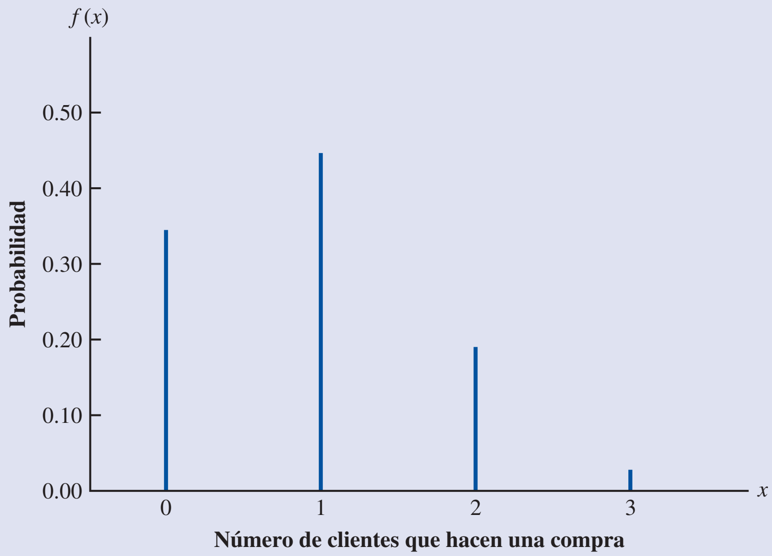
En el ejemplo de la tienda de ropa Martin Clothing Store se calculará ahora la probabilidad de que ningún cliente realice una compra, de que exactamente un cliente realice una compra, de que exactamente dos clientes realicen una compra y de que los tres clientes realicen una compra. Los cálculos se presentan en forma resumida en la tabla 5.7, que da la distribución de probabilidad para el número de clientes que hacen una compra. La figura 5.4 es una gráfica de esta distribución de probabilidad.

La función de probabilidad binomial es aplicable a *cualquier* experimento binomial. Si encuentra que una situación presenta las propiedades de un experimento binomial y conoce los valores de n y p, use la ecuación (5.8) para calcular la probabilidad de x éxitos en n ensayos.

TABLA 5.7 DISTRIBUCIÓN DE PROBABILIDAD BINOMIAL PARA EL NÚMERO DE CLIENTES QUE HACEN UNA COMPRA

x	$f(x)$
0	$\frac{3!}{0!3!}(0.30)^0(0.70)^3 = 0.343$
1	$\frac{3!}{1!2!}(0.30)^1(0.70)^2 = 0.441$
2	$\frac{3!}{2!1!}(0.30)^2(0.70)^1 = 0.189$
3	$\frac{3!}{3!0!}(0.30)^3(0.70)^0 = 0.027$
	1.000

FIGURA 5.4 REPRESENTACIÓN GRÁFICA DE LA DISTRIBUCIÓN DE PROBABILIDAD BINOMIAL PARA EL NÚMERO DE CLIENTES QUE HACEN UNA COMPRA



Si considera variaciones del experimento de la tienda de ropa, por ejemplo, que lleguen a la tienda 10 clientes en lugar de tres clientes, también se emplea la función de probabilidad binomial dada por la ecuación (5.8). Suponga que tiene un experimento binomial con $n = 10$, $x = 4$ y $p = 0.30$. La probabilidad de que cuatro de los 10 clientes que entran en la tienda de ropa realicen una compra es

$$f(4) = \frac{10!}{4!6!}(0.30)^4(0.70)^6 = 0.2001$$

Uso de las tablas de probabilidades binomiales

Existen tablas que dan la probabilidad de x éxitos en n ensayos de un experimento binomial. Estas tablas son fáciles de usar y los resultados se obtienen más rápidamente que con la ecuación (5.8). La tabla 5 del apéndice B es una de estas tablas de probabilidades binomiales. Una parte de esta tabla se presenta en la tabla 5.8. Para usarla es necesario especificar los valores de n, p y x en el experimento binomial de que se trate. En el ejemplo que se presenta en la parte superior de la tabla 5.8 se ve que la probabilidad de $x = 3$ éxitos en un experimento binomial con $n = 10$ y $p = 0.40$ es 0.2150. Use la ecuación (5.8) para verificar que este mismo resultado se obtiene si usa la función de probabilidad binomial directamente.

Ahora se usará la tabla 5.8 para corroborar la probabilidad de 4 éxitos en 10 ensayos en el problema de la tienda de ropa Martin Clothing Store. Observe que el valor de $f(4) = 0.2001$ se lee directamente de la tabla de probabilidades binomiales, eligiendo $n = 10$, $x = 4$ y $p = 0.30$.

Con las calculadoras modernas estas tablas son casi innecesarias. Es muy fácil evaluar la ecuación (5.8) directamente.

Aun cuando las tablas de probabilidades binomiales son relativamente fáciles de utilizar, es imposible contar con tablas que tengan todos los valores de n y p de un experimento binomial. Sin embargo, con las calculadoras de hoy en día, usar la ecuación (5.8) para calcular la probabilidad deseada no es difícil, en especial si el número de ensayos no es grande. En los ejercicios tendrá la oportunidad de usar la ecuación (5.8) para calcular probabilidades binomiales, a menos que el problema pida que use la tabla de probabilidad binomial.

TABLA 5.8 ALGUNOS VALORES DE LA TABLA DE PROBABILIDAD BINOMIAL
EJEMPLO: $n = 10, x = 3, p = 0.40; f(3) = 0.2150$

						p					
n	x	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020
	1	0.2985	0.3874	0.3679	0.3020	0.2253	0.1556	0.1004	0.0605	0.0339	0.0176
	2	0.0629	0.1722	0.2597	0.3020	0.3003	0.2668	0.2162	0.1612	0.1110	0.0703
	3	0.0077	0.0446	0.1069	0.1762	0.2336	0.2668	0.2716	0.2508	0.2119	0.1641
	4	0.0006	0.0074	0.0283	0.0661	0.1168	0.1715	0.2194	0.2508	0.2600	0.2461
	5	0.0000	0.0008	0.0050	0.0165	0.0389	0.0735	0.1181	0.1672	0.2128	0.2461
	6	0.0000	0.0001	0.0006	0.0028	0.0087	0.0210	0.0424	0.0743	0.1160	0.1641
	7	0.0000	0.0000	0.0000	0.0003	0.0012	0.0039	0.0098	0.0212	0.0407	0.0703
	8	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0013	0.0035	0.0083	0.0176
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0008	0.0020
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.3151	0.3874	0.3474	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098
	2	0.0746	0.1937	0.2759	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439
	3	0.0105	0.0574	0.1298	0.2013	0.2503	0.2668	0.2522	0.**2150**	0.1665	0.1172
	4	0.0010	0.0112	0.0401	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051
	5	0.0001	0.0015	0.0085	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461
	6	0.0000	0.0001	0.0012	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051
	7	0.0000	0.0000	0.0001	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172
	8	0.0000	0.0000	0.0000	0.0001	0.0004	0.0014	0.0043	0.0106	0.0229	0.0439
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010

Los paquetes de software para estadística como Minitab y los paquetes de hojas de cálculo como Excel también están habilitadas para calcular probabilidades binomiales. Considere el ejemplo de la tienda de ropa Martin Clothing Store con $n = 10$ y $p = 0.30$. En la figura 5.5 se muestran las probabilidades binomiales para todos los valores posibles de x, generadas por Minitab. Observe que estos valores son los mismos que se encuentran en la columna $p = 0.30$ de la tabla 5.8. En el apéndice 5.1 se da paso por paso el procedimiento en Minitab para generar el resultado que se muestra en la figura 5.5. En el apéndice 5.2 se describe cómo usar Excel para calcular probabilidades binomiales.

Valor esperado y varianza en la distribución binomial

En la sección 5.3 se dieron las fórmulas para calcular el valor esperado y la varianza de una variable aleatoria discreta. En el caso especial de que la variable aleatoria tenga una distribución binomial para la que se conoce el número de ensayos n y la probabilidad de éxito p, las fórmulas generales para el valor esperado y la varianza se simplifican. El resultado se muestra a continuación.

VALOR ESPERADO Y VARIANZA EN LA DISTRIBUCIÓN BINOMIAL

$$E(x) = \mu = np \qquad \textbf{(5.9)}$$

$$\text{Var}(x) = \sigma^2 = np(1 - p) \qquad \textbf{(5.10)}$$

FIGURA 5.5 RESULTADOS DE MINITAB QUE MUESTRAN LAS PROBABILIDADES BINOMIALES PARA EL PROBLEMA DE LA TIENDA DE ROPA MARTIN CLOTHING STORE

x	P(X = x)
0.00	0.0282
1.00	0.1211
2.00	0.2335
3.00	0.2668
4.00	0.2001
5.00	0.1029
6.00	0.0368
7.00	0.0090
8.00	0.0014
9.00	0.0001
10.00	0.0000

Para el problema de los tres clientes de la tienda de ropa Martin Clothing Store, use la ecuación (5.9) para calcular el número esperado de clientes que harán una compra.

$$E(x) = np = 3(0.30) = 0.9$$

Suponga que Martin Clothing Store pronostica que el mes próximo 1000 clientes visitarán la tienda. ¿Cuál es el número esperado de clientes que harán una compra? La respuesta es $\mu = np = (1000)(0.30) = 300$. Así, para aumentar el número esperado de compras, Martin debe hacer que más clientes visiten su tienda o de alguna manera aumentar la probabilidad de que una persona que visite la tienda haga una compra.

En el caso de los tres clientes de la tienda de ropa Martin Clothing Store, la varianza y la desviación estándar del número de clientes que harán una compra son

$$\sigma^2 = np(1 - p) = 3(0.3)(0.7) = 0.63$$
$$\sigma = \sqrt{0.63} = 0.79$$

Para los próximos 1000 clientes que visiten la tienda, la varianza y la desviación estándar del número de clientes que harán una compra son

$$\sigma^2 = np(1 - p) = 1000(0.3)(0.7) = 210$$
$$\sigma = \sqrt{210} = 14.49$$

NOTAS Y COMENTARIOS

1. En las tablas binomiales del apéndice B los valores de p llegan sólo hasta 0.50. Es posible pensar que estas tablas no son útiles cuando la probabilidad de éxito es mayor a 0.50. Sin embargo, puede usarlas observando que la probabilidad de $n - x$ fracasos es también la probabilidad de x éxitos. Cuando la probabilidad de éxito es mayor que $p = 0.50$, en lugar de la probabilidad de éxito calcule la probabilidad de $n - x$ fracasos. Cuando $p > 0.50$, la probabilidad de fracaso, $1 - p$, será menor que 0.50.

2. En algunas fuentes se presentan tablas binomiales en forma acumulada. Al usar estas tablas para hallar la probabilidad de x éxitos en n ensayos hay que hacer una resta. Por ejemplo, $f(2) = P(x \leq 2) - P(x \leq 1)$. Las tablas que se presentan en este libro dan estas probabilidades. Para calcular probabilidades acumuladas usando las tablas de este libro, sume las probabilidades individuales. Por ejemplo, para calcular $P(x \leq 2)$ usando las tablas del libro, sume $f(0) + f(1) + f(2)$.

Ejercicios

Métodos

Auto examen

25. Considere un experimento binomial con dos ensayos y $p = 0.4$.
   a. Dibuje un diagrama de árbol para este experimento (véase figura 5.3).
   b. Calcule la probabilidad de un éxito, $f(1)$.
   c. Calcule $f(0)$.
   d. Calcule $f(2)$.
   e. Calcule la probabilidad de por lo menos un éxito.
   f. Calcule el valor esperado, la varianza y la desviación estándar.

26. Considere un experimento binomial con $n = 10$ y $p = 0.10$.
   a. Calcule $f(0)$.
   b. Calcule $f(2)$.
   c. Calcule $P(x \leq 2)$.
   d. Calcule $P(x \geq 1)$.
   e. Calcule $E(x)$.
   f. Calcule $\text{Var}(x)$ y σ.

27. Considere un experimento binomial con $n = 20$ y $p = 0.70$.
   a. Calcule $f(12)$.
   b. Calcule $f(16)$.
   c. Calcule $P(x \geq 16)$.
   d. Calcule $P(x \leq 15)$.
   e. Calcule $E(x)$.
   f. Calcule $\text{Var}(x)$ y σ.

Aplicaciones

28. Una encuesta de Harris Interactive para InterContinental Hoteld and Resorts preguntó: "Cuando viaja al extranjero, ¿suele aventurarse usted solo para conocer la cultura o prefiere permanecer con el grupo de su *tour* y apegarse al itinerario?" Se encontró que 23% prefiere permanecer con el grupo de su *tour* (*USA Today,* 21 de enero de 2004).
   a. ¿Cuál es la probabilidad de que en una muestra de seis viajeros, dos prefieran permanecer con su grupo?
   b. ¿De que en una muestra de seis viajeros, por lo menos dos prefieran permanecer con su grupo?
   c. ¿De que en una muestra de 10 viajeros, ninguno prefiera permanecer con su grupo?

29. En San Francisco, 30% de los trabajadores emplean el transporte público (*USA Today,* 21 de diciembre de 2005).
   a. ¿Cuál es la probabilidad de que en una muestra de 10 trabajadores exactamente tres empleen el transporte público?
   b. ¿De que en una muestra de 10 trabajadores por lo menos tres empleen el transporte público?

Auto examen

30. Cuando una máquina nueva funciona adecuadamente, sólo 3% de los artículos producidos presentan algún defecto. Suponga que selecciona aleatoriamente dos piezas producidas con la nueva máquina y que busca el número de piezas defectuosas.
   a. Describa las condiciones en las que éste será un experimento binomial.
   b. Elabore un diagrama de árbol como el de la figura 5.3 en el que se muestre este problema como un experimento de dos ensayos.
   c. ¿En cuántos resultados experimentales hay exactamente una pieza defectuosa?
   d. Calcule las probabilidades de hallar ninguna pieza defectuosa, exactamente una pieza defectuosa y dos piezas defectuosas.

31. Nueve por ciento de los estudiantes tienen un balance en su tarjeta de crédito mayor a \$7000 (*Reader's Digest*, julio de 2002). Suponga que selecciona aleatoriamente 10 estudiantes para entrevistarlos respecto del uso de su tarjeta de crédito

a. ¿Es la selección de 10 estudiantes un experimento binomial? Explique.
b. ¿Cuál es la probabilidad de que dos de los estudiantes tengan un balance en su tarjeta de crédito superior a $7000?
c. ¿De que ninguno tenga un balance en su tarjeta de crédito superior a $7000?
d. ¿De que por lo menos tres tengan un balance en su tarjeta de crédito superior a $7000?

32. Los radares militares y los sistemas para detección de misiles tienen por objeto advertir a un país de un ataque enemigo. Una cuestión de confiabilidad es si el sistema de detección será capaz de detectar un ataque y emitir un aviso. Suponga que la probabilidad de que un determinado sistema de detección detecte un ataque con misiles es 0.90. Use la distribución de probabilidad binomial para responder las preguntas siguientes.
a. ¿Cuál es la probabilidad de que un solo sistema de detección detecte un ataque?
b. Si se instalan dos sistemas de detección en una misma área y los dos operan independientemente, ¿cuál es la probabilidad de que por lo menos uno de los sistemas detecte el ataque?
c. Si se instalan tres sistemas, ¿cuál es la probabilidad de que por lo menos uno de los sistemas detecte el ataque?
d. ¿Recomendaría que se usaran varios sistemas de detección? Explique.

33. Cincuenta por ciento de los estadounidenses creyeron que el país se encontraba en una recesión aun cuando en la economía no se habían observado dos trimestres seguidos con crecimiento negativo. (*BusinessWeek,* 30 de julio de 2001). Dada una muestra de 20 estadounidenses, calcule lo siguiente.
a. Calcule la probabilidad de que exactamente 12 personas hayan creído que el país estaba en recesión.
b. De que no más de cinco personas hayan creído que el país estaba en recesión
c. ¿Cuántas personas esperaría usted que dijeran que el país estuvo en recesión?
d. Calcule la varianza y la desviación estándar del número de personas que creyeron que el país estuvo en recesión.

34. En una encuesta realizada por la Oficina de Censos de Estados Unidos se encontró que 25% de las personas de 25 años o más habían estudiado cuatro años en la universidad (*The New York Times Almanac,* 2006). Dada una muestra de 15 individuos de 25 años o más, conteste las preguntas siguientes.
a. ¿Cuál es la probabilidad de que cuatro hayan estudiado cuatro años en la universidad?
b. ¿De que tres o más hayan estudiado cuatro años en la universidad?

35. En una universidad se encontró que 20% de los estudiantes no terminan el primer curso de estadística, al curso se inscriben 20 estudiantes.
a. Calcule la probabilidad de que dos o menos no terminen.
b. De que cuatro, exactamente, no terminen.
c. De que más de tres no terminen.
d. ¿Cuál es el número esperado de estudiantes que no terminan?

36. En el caso particular de una variable aleatoria binomial, es factible calcular la varianza empleando la fórmula $\sigma^2 = np(1 - p)$. En el caso del problema de la tienda de ropa Martin Clothing Store, en donde $n = 3$ y $p = 0.3$, se encontró que $\sigma^2 = np(1 - p) = 3(0.3)(0.7) = 0.63$. Aplique la definición general de varianza para una variable aleatoria discreta, ecuación (5.5), y las probabilidades de la tabla 5.7 para comprobar que la varianza es 0.63

37. Veintitrés por ciento de los automóviles no cuenta con un seguro (CNN, 23 de febrero de 2006). En un fin de semana determinado hay 35 automóviles que sufren un accidente.
a. ¿Cuál es el número esperado de estos automóviles que no cuentan con un seguro?
b. ¿Cuál es la varianza y la desviación estándar?

5.5 Distribución de probabilidad de Poisson

En esta sección estudiará una variable aleatoria discreta que se suele usar para estimar el número de veces que sucede un hecho determinado (ocurrencias) en un intervalo de tiempo o de espacio. Por ejemplo, la variable de interés va desde el número de automóviles que llegan (llegadas) a un lavado de coches en una hora o el número de reparaciones necesarias en 10 millas de una autopista hasta el número de fugas en 100 millas de tubería. Si se satisfacen las condiciones si-

La distribución de probabilidad de Poisson suele emplearse para modelar las llegadas aleatorias a una línea de espera (fila).

guientes, el número de ocurrencias es una variable aleatoria discreta, descrita por la **distribución de probabilidad de Poisson**.

PROPIEDADES DE UN EXPERIMENTO DE POISSON

1. La probabilidad de ocurrencia es la misma para cualesquiera dos intervalos de la misma magnitud.
2. La ocurrencia o no-ocurrencia en cualquier intervalo es independiente de la ocurrencia o no-ocurrencia en cualquier otro intervalo.

La **función de probabilidad de Poisson** se define mediante la ecuación (5.11).

Simeon Poisson dio clases de matemáticas en la Ecole Polytechnique de París de 1802 a 1808. En 1837 publicó un trabajo titulado "Investigación sobre la probabilidad de veredictos en materia criminal y civil" en el que presenta un estudio sobre lo que después se conoció como distribución de Poisson.

FUNCIÓN DE PROBABILIDAD DE POISSON

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \tag{5.11}$$

en donde

$f(x)$ = probabilidad de x ocurrencias en un intervalo
μ = valor esperado o número medio de ocurrencias en un intervalo
e = 2.71828

Antes de considerar un ejemplo para ver cómo se usa la distribución de Poisson, observe que el número de ocurrencias x, no tiene límite superior. Ésta es una variable aleatoria discreta que toma los valores de una sucesión infinita de números ($x = 0, 1, 2, \ldots$).

Un ejemplo considerando intervalos de tiempo

Los laboratorios Bell usaron la distribución de Poisson para modelar las llegadas de llamadas telefónicas.

Suponga que desea saber el número de llegadas, en un lapso de 15 minutos, a la rampa del cajero automático de un banco. Si se puede suponer que la probabilidad de llegada de los automóviles es la misma en cualesquiera dos lapsos de la misma duración y si la llegada o no–llegada de un automóvil en cualquier lapso es independiente de la llegada o no–llegada de un automóvil en cualquier otro lapso, se puede aplicar la función de probabilidad de Poisson. Dichas condiciones se satisfacen y en un análisis de datos pasados encuentra que el número promedio de automóviles que llegan en un lapso de 15 minutos es 10; en este caso use la función de probabilidad siguiente.

$$f(x) = \frac{10^x e^{-10}}{x!}$$

Aquí la variable aleatoria es x = número de automóviles que llegan en un lapso de 15 minutos.

Si la administración desea saber la probabilidad de que lleguen exactamente cinco automóviles en 15 minutos, $x = 5$, y se obtiene

$$\begin{array}{l}\text{Probabilidad de que lleguen}\\ \text{exactamente 5 automóviles}\\ \text{en 15 minutos}\end{array} = f(5) = \frac{10^5 e^{-10}}{5!} = 0.0378$$

Aunque esta probabilidad se obtuvo evaluando la función de probabilidad con $\mu = 10$ y $x = 5$, suele ser más fácil consultar una tabla de probabilidad de Poisson. Dichas tablas proporcionan las probabilidades para valores específicos de x y μ. La tabla 7 del apéndice B es una tabla de probabilidad de Poisson. Para mayor comodidad, en la tabla 5.9 se reproduce parte de la tabla 7 del apéndice B. Observe que para usar una tabla de probabilidades de Poisson se necesitan sólo

TABLA 5.9 ALGUNOS VALORES DE LAS TABLAS DE PROBABILIDAD DE POISSON
EJEMPLO: $\mu = 10$, $x = 5$; $f(5) = .0378$

	μ									
x	**9.1**	**9.2**	**9.3**	**9.4**	**9.5**	**9.6**	**9.7**	**9.8**	**9.9**	**10**
0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
1	0.0010	0.0009	0.0009	0.0008	0.0007	0.0007	0.0006	0.0005	0.0005	0.0005
2	0.0046	0.0043	0.0040	0.0037	0.0034	0.0031	0.0029	0.0027	0.0025	0.0023
3	0.0140	0.0131	0.0123	0.0115	0.0107	0.0100	0.0093	0.0087	0.0081	0.0076
4	0.0319	0.0302	0.0285	0.0269	0.0254	0.0240	0.0226	0.0213	0.0201	0.0189
5	0.0581	0.0555	0.0530	0.0506	0.0483	0.0460	0.0439	0.0418	0.0398	0.**0378**
6	0.0881	0.0851	0.0822	0.0793	0.0764	0.0736	0.0709	0.0682	0.0656	0.0631
7	0.1145	0.1118	0.1091	0.1064	0.1037	0.1010	0.0982	0.0955	0.0928	0.0901
8	0.1302	0.1286	0.1269	0.1251	0.1232	0.1212	0.1191	0.1170	0.1148	0.1126
9	0.1317	0.1315	0.1311	0.1306	0.1300	0.1293	0.1284	0.1274	0.1263	0.1251
10	0.1198	0.1210	0.1219	0.1228	0.1235	0.1241	0.1245	0.1249	0.1250	0.1251
11	0.0991	0.1012	0.1031	0.1049	0.1067	0.1083	0.1098	0.1112	0.1125	0.1137
12	0.0752	0.0776	0.0799	0.0822	0.0844	0.0866	0.0888	0.0908	0.0928	0.0948
13	0.0526	0.0549	0.0572	0.0594	0.0617	0.0640	0.0662	0.0685	0.0707	0.0729
14	0.0342	0.0361	0.0380	0.0399	0.0419	0.0439	0.0459	0.0479	0.0500	0.0521
15	0.0208	0.0221	0.0235	0.0250	0.0265	0.0281	0.0297	0.0313	0.0330	0.0347
16	0.0118	0.0127	0.0137	0.0147	0.0157	0.0168	0.0180	0.0192	0.0204	0.0217
17	0.0063	0.0069	0.0075	0.0081	0.0088	0.0095	0.0103	0.0111	0.0119	0.0128
18	0.0032	0.0035	0.0039	0.0042	0.0046	0.0051	0.0055	0.0060	0.0065	0.0071
19	0.0015	0.0017	0.0019	0.0021	0.0023	0.0026	0.0028	0.0031	0.0034	0.0037
20	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012	0.0014	0.0015	0.0017	0.0019
21	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009
22	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004
23	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

dos valores, x y μ. En la tabla 5.9 la probabilidad de cinco llegadas en un lapso de 15 minutos se obtiene localizando el valor que se encuentra en el renglón correspondiente a $x = 5$ y la columna correspondiente a $\mu = 10$. Así obtiene $f(5) = 0.0378$

Una propiedad de la distribución de Poisson es que la media y la varianza son iguales.

La media de la distribución de Poisson en el ejemplo anterior fue $\mu = 10$ llegadas en un lapso de 15 minutos. Una propiedad de la distribución de Poisson es que la media y la varianza de la distribución son *iguales*. Por tanto, la varianza del número de llegadas en un lapso de 15 minutos es $\sigma^2 = 10$. La desviación estándar es $\sigma = \sqrt{10} = 3.16$.

En el ejemplo anterior se usó un lapso de 15 minutos, pero también se usan otros lapsos. Suponga que desea calcular la probabilidad de una llegada en un lapso de 3 minutos. Como 10 es el número esperado de llegadas en un lapso de 15 minutos: $10/15 = 2/3$ es el número esperado de llegadas en un lapso de un minuto y que $(2/3)(3 \text{ minutos}) = 2$ es el número esperado de llegadas en un lapso de 3 minutos. Entonces, la probabilidad de x llegadas en un lapso de 3 minutos con $\mu = 2$ está dada por la siguiente función de probabilidad de Poisson.

$$f(x) = \frac{2^x e^{-2}}{x!}$$

La probabilidad de una llegada en un lapso de 3 minutos se obtiene como sigue:

$$\begin{array}{l}\text{Probabilidad de exactamente} \\ \text{una llegada en 3 minutos}\end{array} = f(1) = \frac{2^1 e^{-2}}{1!} = 0.2707$$

Antes se calculó la probabilidad de cinco llegadas en un lapso de 15 minutos; se obtuvo 0.0378. Observe que la probabilidad de una llegada en un lapso de tres minutos (0.2707) no es la misma. Para calcular la probabilidad de Poisson en un lapso diferente, primero hay que convertir la llegada media al lapso que interesa y después calcular la probabilidad.

Un ejemplo considerando intervalos de longitud o de distancia

Ahora se da un ejemplo en el que no aparecen intervalos de tiempo y en el que se usa la distribución de Poisson. Asuma que le interesa la ocurrencia de una avería importante en una autopista un mes después de que ha sido repavimentada. Supondrá que la probabilidad de que haya una avería es la misma en cualesquiera dos tramos, de una misma longitud, de la autopista y que la ocurrencia o no–ocurrencia de una avería en un tramo es independiente de la ocurrencia o no-ocurrencia de una avería en cualquier otro tramo. Por tanto, emplea la distribución de Poisson.

También sabe que el promedio de averías importantes, un mes después de la repavimentación, son dos averías por milla. Desea determinar la probabilidad de que no haya ninguna avería en un determinado tramo de tres millas de autopista. Como lo que interesa es un intervalo cuya longitud es de tres millas, μ = (2 averías/milla)(3 millas) = 6 representa el número esperado de averías importantes en un tramo de tres millas de autopista. Mediante la ecuación (5.11), la probabilidad de que no haya ninguna avería importante es $f(0) = 6^0 e^{-6}/0! = 0.0025$. Por tanto, es poco probable que no haya ninguna avería importante en este tramo de tres millas. En efecto, este ejemplo indica que hay una probabilidad de $1 - 0.0025 = 0.9975$ de que haya por lo menos una avería importante en este tramo de tres millas de autopista.

Ejercicios

Métodos

38. Considere una distribución de Poisson con $\mu = 3$.
    a. Dé la adecuada función de probabilidad de Poisson.
    b. Calcule $f(2)$.
    c. Calcule $f(1)$.
    d. Calcule $P(x \geq 2)$.

Autoexamen

39. Considere una distribución de Poisson en que la media es de dos ocurrencias por un periodo de tiempo.
    a. Dé la adecuada función de probabilidad de Poisson.
    b. ¿Cuál es el número esperado de ocurrencias en tres periodos de tiempo?
    c. Dé la adecuada función de probabilidad de Poisson para determinar la probabilidad de x ocurrencias en tres lapsos.
    d. Calcule la probabilidad de dos ocurrencias en un periodo de tiempo.
    e. Calcule la probabilidad de seis ocurrencias en tres periodos de tiempo.
    f. Calcule la probabilidad de cinco ocurrencias en dos periodos de tiempo.

Aplicaciones

40. A la oficina de reservaciones de una aerolínea regional llegan 48 llamadas por hora.
    a. Calcule la probabilidad de recibir cinco llamadas en un lapso de 5 minutos.
    b. Estime la probabilidad de recibir exactamente 10 llamadas en un lapso de 15 minutos.
    c. Suponga que no hay ninguna llamada en espera. Si el agente de viajes necesitará 5 minutos para la llamada que está atendiendo, ¿cuántas llamadas habrá en espera para cuando él termine? ¿Cuál es la probabilidad de que no haya ninguna llamada en espera?
    d. Si en este momento no hay ninguna llamada, ¿cuál es la probabilidad de que el agente de viajes pueda tomar 3 minutos de descanso sin ser interrumpido por una llamada?

41. Durante el periodo en que una universidad recibe inscripciones por teléfono, llegan llamadas a una velocidad de una cada dos minutos.
    a. ¿Cuál es el número esperado de llamadas en una hora?
    b. ¿Cuál es la probabilidad de que haya tres llamadas en cinco minutos?
    c. ¿De que no haya llamadas en un lapso de cinco minutos?

Auto examen

42. En Estados Unidos, cada año, más de 50 millones de huéspedes se alojan en un "Bread and breakfast" (B&B). El sitio Web dedicado a los alojamientos tipo Bread and Breakfast en Estados Unidos (www.bestinns.net), que tiene un promedio aproximado de siete visitantes por minuto, permite a muchos B&B obtener huéspedes (*Time*, septiembre de 2001).
    a. Calcule la probabilidad de que no haya ningún visitante al sitio Web en un lapso de un minuto.
    b. De que haya dos o más visitantes al sitio Web en un lapso de un minuto.
    c. De que haya uno o más visitantes al sitio Web en un lapso de 30 segundos.
    d. De que haya cinco o más visitantes al sitio Web en un lapso de un minuto.
43. Los pasajeros de las aerolíneas llegan en forma aleatoria e independiente al mostrador de revisión de pasajeros. La tasa media de llegada es 10 pasajeros por minuto.
    a. Calcule la probabilidad de que no llegue ningún pasajero en un lapso de un minuto.
    b. Calcule la probabilidad de que lleguen tres o menos pasajeros en un lapso de un minuto.
    c. De que no llegue ningún pasajero en un lapso de 15 segundos.
    d. De que llegue por lo menos un pasajero en un lapso de 15 segundos.
44. Cada año ocurren en promedio 15 accidentes aéreos (*The World Almanac and Book of Facts*, 2004).
    a. Calcule el número medio de accidentes aéreos por mes.
    b. Calcule la probabilidad de que no haya ningún accidente en un mes.
    c. De que haya exactamente un accidente en un mes.
    d. De que haya más de un accidente en un mes.
45. El National Safety Council de Estados Unidos estima que los accidentes fuera del trabajo tienen para las empresas un costo de casi \$200 mil millones anuales en pérdida de productividad. Con base en estos datos, las empresas que tienen 50 empleados esperan tener por lo menos tres accidentes fuera del trabajo por año. Para estas empresas con 50 empleados, conteste las preguntas siguientes.
    a. ¿Cuál es la probabilidad de que no haya ningún accidente fuera del trabajo en un año?
    b. ¿De que haya por lo menos dos accidentes fuera del trabajo en un año?
    c. ¿Cuál es el número esperado de accidentes fuera del trabajo en un lapso de seis meses?
    d. ¿Cuál es la probabilidad de que no haya ningún accidente fuera del trabajo en los próximos seis meses?

5.6 Distribución de probabilidad hipergeométrica

La **distribución de probabilidad hipergeométrica** está estrechamente relacionada con la distribución binomial. Pero difieren en dos puntos: en la distribución hipergeométrica los ensayos no son independientes y la probabilidad de éxito varía de ensayo a ensayo.

En la notación usual en la distribución hipergeométrica, r denota el número de elementos considerados como éxitos que hay en una población de tamaño N, y $N - r$ denota el número de elementos considerados como fracasos que hay en dicha población. La **función de probabilidad hipergeométrica** se usa para calcular la probabilidad de que en una muestra aleatoria de n elementos, seleccionados sin reemplazo, se tengan x éxitos y $n - x$ fracasos. Para que se presente este resultado, debe tener x éxitos de los r éxitos que hay en la población y $n - x$ fracasos de los $N - r$ fracasos. La siguiente función de probabilidad hipergeométrica proporciona $f(x)$, la probabilidad de tener x éxitos en una muestra de tamaño n.

FUNCIÓN DE PROBABILIDAD HIPERGEOMÉTRICA

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad \text{para } 0 \leq x \leq r \tag{5.12}$$

donde

$f(x) =$ probabilidad de x éxitos en n ensayos
$n =$ número de ensayos
$N =$ número de elementos en la población
$r =$ número de elementos en la población considerados como éxitos

Observe que $\binom{N}{n}$ representa el número de maneras en que es posible tomar una muestra de tamaño n de una población de tamaño N; $\binom{r}{x}$ representa el número de formas en que se toman x éxitos de un total de r éxitos que hay en la población, y $\binom{N-r}{n-x}$ representa el número de maneras en que se puede tomar $n - x$ fracasos de un total de $N - r$ que hay en la población.

Para ilustrar los cálculos que se emplean al usar la ecuación (5.12), considere la siguiente aplicación al control de calidad. Una empresa fabrica fusibles que empaca en cajas de 12 unidades cada una. Asuma que un inspector selecciona al azar tres de los 12 fusibles de una caja para inspeccionarlos. Si la caja contiene exactamente cinco fusibles defectuosos, ¿cuál es la probabilidad de que el inspector encuentre que uno de los tres fusibles está defectuoso? En esta aplicación $n = 3$ y $N = 12$. Si $r = 5$ fusibles defectuosos en la caja, la probabilidad de hallar $x = 1$ defectuoso es

$$f(1) = \frac{\binom{5}{1}\binom{7}{2}}{\binom{12}{3}} = \frac{\left(\frac{5!}{1!4!}\right)\left(\frac{7!}{2!5!}\right)}{\left(\frac{12!}{3!9!}\right)} = \frac{(5)(21)}{220} = 0.4773$$

Ahora suponga que desea conocer la probabilidad de hallar *por lo menos* un fusible defectuoso. La manera más sencilla de contestar es calcular primero la probabilidad de que el inspector no encuentre ningún fusible defectuoso. La probabilidad de $x = 0$ es

$$f(0) = \frac{\binom{5}{0}\binom{7}{3}}{\binom{12}{3}} = \frac{\left(\frac{5!}{0!5!}\right)\left(\frac{7!}{3!4!}\right)}{\left(\frac{12!}{3!9!}\right)} = \frac{(1)(35)}{220} = 0.1591$$

Si la probabilidad de cero fusibles defectuosos es $f(0) = 0.1591$, se concluye que la probabilidad de hallar por lo menos un fusible defectuoso debe ser $1 - 0.1591 = 0.8409$. Así, existe una probabilidad razonablemente alta de que el inspector encuentre por lo menos un fusible defectuoso.

La media y la varianza de una distribución hipergeométrica son las siguientes.

$$E(x) = \mu = n\left(\frac{r}{N}\right) \tag{5.13}$$

$$\text{Var}(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right) \tag{5.14}$$

En el ejemplo anterior $n = 3$, $r = 5$ y $N = 12$. Por tanto, la media y la varianza del número de fusibles defectuosos es

$$\mu = n\left(\frac{r}{N}\right) = 3\left(\frac{5}{12}\right) = 1.25$$

$$\sigma^2 = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right) = 3\left(\frac{5}{12}\right)\left(1 - \frac{5}{12}\right)\left(\frac{12-3}{12-1}\right) = 0.60$$

La desviación estándar es $\sigma = \sqrt{0.60} = 0.77$.

NOTAS Y COMENTARIOS

Considere una distribución hipergeométrica con n ensayos. Sea $p = (r/N)$ la probabilidad de un éxito en el primer ensayo. Si el tamaño de la población es grande, el término $(N - n)/(N - 1)$ de la ecuación (5.14) se aproxima a 1. Entonces, el valor esperado y la varianza se expresan como $E(x) = np$ y $\text{Var}(x) = np(1 - p)$. Preste atención a que estas expresiones son las mismas que se usan para calcular el valor esperado y la varianza en una distribución binomial, ecuaciones (5.9) y (5.10). Cuando el tamaño de la población es grande, se aproxima una distribución hipergeométrica mediante una distribución binomial con n ensayos y probabilidad de éxito $p = (r/N)$.

Ejercicios

Métodos

Auto examen

46. Suponga que $N = 10$ y $r = 3$. Calcule las probabilidades hipergeométricas correspondientes a los valores siguiente de n y x.
    a. $n = 4, x = 1$.
    b. $n = 2, x = 2$.
    c. $n = 2, x = 0$.
    d. $n = 4, x = 2$.
47. Suponga que $N = 15$ y $r = 4$. ¿Cuál es la probabilidad de $x = 3$ para $n = 10$?

Aplicaciones

48. En una encuesta realizada por Gallup Organization, se les preguntó a los interrogados, “Cuál es el deporte que prefieres ver”. Futbol y básquetbol ocuparon el primero y segundo lugar de preferencia (www.gallup.com, 3 de enero de 2004). Si en un grupo de 10 individuos, siete prefieren futbol y tres prefieren básquetbol. Se toma una muestra aleatoria de tres de estas personas.
    a. ¿Cuál es la probabilidad de que exactamente dos prefieren el futbol?
    b. ¿De que la mayoría (ya sean dos o tres) prefiere el futbol?
49. Blackjack, o veintiuno, como se le suele llamar, es un popular juego de apuestas en los casinos de Las Vegas. A un jugador se le reparten dos cartas. Las figuras (sotas, reinas y reyes) y los 10 valen 10 puntos. Los ases valen 1 u 11. Una baraja de 52 cartas tiene 16 cartas que valen 10 (sotas, reinas, reyes y dieces) y cuatro ases.

a. ¿Cuál es la probabilidad de que las dos cartas repartidas sean ases o cartas que valgan 10 puntos?
b. ¿De que las dos cartas sean ases?
c. ¿De que las dos cartas valgan 10?
d. Un blackjack es una carta de 10 puntos y un as que suman 21. Use sus respuestas a los incisos a, b y c para determinar la probabilidad de que a un jugador se le reparta blackjack. (*Indicación:* El inciso c no es un problema hipergeométrico. Desarrolle su propio razonamiento lógico para combinar las probabilidades hipergeométricas de los incisos a, b y c para responder esta pregunta.)

Auto examen

50. Una empresa fabrica computadoras personales en dos fábricas, una en Texas y la otra en Hawai. La fábrica de Texas tiene 40 empleados; la fábrica de Hawai tiene 20 empleados. A una muestra aleatoria de 20 empleados se le pide que llene un cuestionario sobre prestaciones.
    a. ¿Cuál es la probabilidad de que ninguno de los empleados de la muestra trabaje en la fábrica de Hawai?
    b. ¿De que uno de los empleados de la muestra trabaje en la fábrica de Hawai?
    c. ¿De que dos o más de los empleados de la muestra trabajen en la fábrica de Hawai?
    d. ¿De que nueve de los empleados de la muestra trabajen en la fábrica de Texas?
51. En una revista de encuestas se da información sobre la evaluación a los platillos, la decoración y el servicio de varios de los principales restaurantes de Estados Unidos. En 15 de los mejor evaluados restaurantes de Boston, el costo promedio de una cena, que incluye una bebida y la propina, es $48.60. Usted va a ir en viaje de negocios a Boston y le gustaría cenar en tres de estos restaurantes. Su empresa le pagará máximo $50 por cena. Sus conocidos en Boston le han informado que en una tercera parte de estos restaurantes una cena cuesta más de $50. Suponga que escoge al azar tres de estos restaurantes para ir a cenar.
    a. ¿Cuál es la probabilidad de que el costo de ninguna de las cenas sea mayor a la cantidad que paga su empresa?
    b. ¿De que el costo de una de las cenas sea mayor a la cantidad que paga su empresa?
    c. ¿De que el costo de dos de las cenas sea mayor a la cantidad que paga su empresa?
    d. ¿De que el costo de las tres cenas sea mayor a la cantidad que paga su empresa?
52. En un pedido de 10 artículos hay dos defectuosos y ocho no defectuosos. Para la inspección del pedido se tomará una muestra y se inspeccionará. Si se encuentra un artículo defectuoso todo el pedido de 10 artículos será devuelto.
    a. Si toma una muestra de tres artículos, ¿cuál es la probabilidad de que devuelva el pedido?
    b. Si toma una muestra de cuatro artículos, ¿cuál es la probabilidad de que devuelva el pedido?
    c. Si toma una muestra de cinco artículos, ¿cuál es la probabilidad de que devuelva el pedido?
    d. Si la administración desea que la probabilidad de rechazar un pedido en el que haya dos artículos defectuosos y ocho no defectuosos sea 0.90, ¿de qué tamaño recomienda que sea la muestra?

Resumen

Una variable aleatoria da una descripción numérica de los resultados de un experimento. La distribución de probabilidad de una variable aleatoria describe cómo se reparten las probabilidades entre los valores que toma dicha variable. En toda variable aleatoria discreta, x, su distribución de probabilidad se define mediante una función de probabilidad, que se denota $f(x)$ y la cual da la probabilidad que corresponde a cada valor de la variable aleatoria. Una vez que se ha definido la función de probabilidad, es posible calcular el valor esperado, la varianza y la desviación estándar de la variable aleatoria.

La distribución binomial se usa para determinar la probabilidad de x éxitos en n ensayos, siempre que el experimento satisfaga las propiedades siguientes:

1. El experimento consista en una serie de n ensayos idénticos.
2. En cada ensayo haya dos resultados posibles, uno llamado éxito y el otro fracaso.
3. La probabilidad de un éxito no varíe de un ensayo a otro. Por tanto, la probabilidad de fracaso, $(1 - p)$, tampoco variará de un resultado a otro.
4. Los ensayos sean independientes.

Si se satisfacen estas cuatro propiedades, la probabilidad de x éxitos en n ensayos se determina usando la función de probabilidad binomial. También se presentaron las fórmulas para hallar la media y la varianza de una distribución binomial.

La distribución de Poisson se usa cuando se quiere obtener la probabilidad de x ocurrencias de un evento en un determinado intervalo de tiempo o de espacio. Para que se emplee la distribución de Poisson deben satisfacerse las condiciones siguientes:

1. La probabilidad de una ocurrencia del evento es la misma para cualesquier dos intervalos de la misma longitud.
2. La ocurrencia o no–ocurrencia del evento en un determinado intervalo es independiente de la ocurrencia o no–ocurrencia del evento en cualquier otro intervalo.

En la sección 5.6 se presentó la tercera distribución discreta de probabilidad presentada, la distribución hipergeométrica. Es como la binomial, que se usa para calcular la probabilidad de x éxitos en n ensayos, pero, a diferencia de ésta, la probabilidad de éxito si varía de un ensayo a otro.

Glosario

Variable aleatoria Una descripción numérica del resultado de un experimento.
Variable aleatoria discreta Una variable aleatoria que puede asumir un número finito de valores o un número infinito de valores de una sucesión.
Variable aleatoria continua Ésta toma cualquier valor de un intervalo o de una colección de intervalos.
Distribución de probabilidad Descripción de cómo se distribuyen las probabilidades entre los valores de una variable aleatoria.
Función de probabilidad Se denota $f(x)$ y da la probabilidad de que x tome un determinado valor de una variable aleatoria.
Distribución de probabilidad uniforme discreta Distribución de probabilidad para la cual cada posible valor de la variable aleatoria tienen la misma probabilidad.
Valor esperado Medida de localización central de una variable aleatoria.
Varianza Medida de la variabilidad o dispersión de una variable aleatoria.
Desviación estándar Raíz cuadrada positiva de la varianza.
Experimento binomial Un experimento que tiene cuatro propiedades que se dan al principio de la sección 5.4.
Distribución de probabilidad binomial Distribución de probabilidad da la probabilidad de x éxitos en n ensayos de un experimento binomial.
Función de probabilidad binomial La función usada para calcular las probabilidades binomiales.
Distribución de probabilidad de Poisson Distribución de probabilidad da la probabilidad de x ocurrencias de un evento en un determinado intervalo de tiempo o de espacio.
Función de probabilidad de Poisson La función usada para calcular las probabilidades de Poisson.
Distribución de probabilidad hipergeométrica Distribución de probabilidad da la probabilidad de x éxitos en n ensayos a partir de una población en la que hay r éxitos y $N - r$ fracasos.
Función de probabilidad hipergeométrica La función usada para calcular probabilidades hipergeométricas

Fórmulas clave

Función de probabilidad uniforme discreta

$$f(x) = 1/n \tag{5.3}$$

Valor esperado en una variable aleatoria discreta

$$E(x) = \mu = \Sigma x f(x) \tag{5.4}$$

Varianza en una variable aleatoria discreta

$$\text{Var}(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x) \tag{5.5}$$

Número de resultados experimentales en los que se encuentran exactamente *x* éxitos en *n* ensayos

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{5.6}$$

Función de probabilidad binomial

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \tag{5.8}$$

Valor esperado en una distribución binomial

$$E(x) = \mu = np \tag{5.9}$$

Varianza en una distribución binomial

$$\text{Var}(x) = \sigma^2 = np(1-p) \tag{5.10}$$

Función de probabilidad de Poisson

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \tag{5.11}$$

Función de probabilidad hipergeométrica

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad \text{para } 0 \leq x \leq r \tag{5.12}$$

Valor esperado en la distribución hipergeométrica

$$E(x) = \mu = n\left(\frac{r}{N}\right) \tag{5.13}$$

Varianza en la distribución hipergeométrica

$$\text{Var}(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right) \tag{5.14}$$

Ejercicios complementarios

53. El *Barron´s* Big Money Poll preguntó a 131 gerentes de inversiones de Estados Unidos acerca de sus puntos de vista sobre las inversiones a corto plazo (*Barron´s*, 28 de octubre de 2002). De acuerdo con las respuestas 4% se encontraban muy optimistas, 39 % se encontraban optimistas, 29% se encontraban neutrales, 21% se encontraban pesimistas y 7% se encontraban muy pesimistas. Sea x la variable aleatoria que refleje el grado de optimismo y que vaya desde $x = 1$ para muy pesimista hasta $x = 5$ para muy optimista.
    a. Elabore una distribución de probabilidad para el grado de optimismo de los gerentes de inversiones.
    b. Calcule el valor esperado del grado de optimismo.
    c. Calcule la varianza y la desviación estándar del grado de optimismo.
    d. Haga un comentario sobre lo que le dicen sus resultados acerca del grado de optimismo y su variabilidad.

54. La American Association of Individual Investors publica una guía anual con los principales fondos mutualistas (*The Individual Investor´s Guide to the Top Mutual Founds*, 22ª ed., American Association of Individual Investors, 2003). En la tabla 5.10 se presenta la clasificación de 29 fondos mutualistas de acuerdo con el riesgo.
    a. Sea x una variable que va desde $x = 1$ con el menor riesgo hasta el mayor riesgo con $x = 5$. Elabore una distribución de probabilidad para el nivel de riesgo.
    b. ¿Cuál es el valor esperado y la varianza del nivel de riesgo?
    c. Se encontró que 11 de éstos eran fondos de renta fija. De ellos siete se clasificaron como bajos y cuatro como abajo del promedio. Compare el riesgo de los fondos de renta fija con los 18 fondos de acciones.

TABLA 5.10 DE 29 FONDOS MUTUALISTAS

Número de fondos	Nivel de riesgo: categorías
Bajo	7
Bajo el promedio	6
Promedio	3
Sobre el promedio	6
Alto	7

55. Al hacer el presupuesto de gastos para el próximo año en una universidad, se obtuvieron los siguientes pronósticos de gastos (dados en millones de dólares) \$9, \$10, \$11, \$12 y \$13. Como no se sabe cuáles son los gastos actuales, a los gastos calculados se les asignaron las probabilidades 0.3, 0.2, 0.25, 0.05 y 0.2.
    a. Dé la distribución de probabilidad para estos pronósticos de gastos.
    b. ¿Cuál es el valor esperado en estos pronósticos de gastos?
    c. ¿Cuál es la varianza en el pronóstico de gastos para el año próximo?
    d. Si las proyecciones de ingreso estiman que éste será de \$12 millones, ¿cómo será la situación financiera de la universidad?

56. En un estudio realizado por la Bureau of Transportation Statistics se encontró que, en promedio, la duración del recorrido de la casa al trabajo de una persona es de 26 minutos. También que 5% de las personas necesitan más de una hora para transportarse de su casa al trabajo.
    a. Si interroga a 20 de estas personas, ¿cuál es la probabilidad de que informen que necesitan más de una hora para ir de su casa el trabajo?
    b. Si interroga a 20 de estas personas, ¿cuál es la probabilidad de que ninguna de ellas informe que necesita más de una hora para ir de su casa al trabajo?

c. Si en una empresa hay 2000 empleados, ¿cuál es el número esperado de empleados que necesita más de una hora para trasladarse de su casa al trabajo?
d. Si en una empresa hay 2000 empleados, ¿cuál es la varianza y la desviación estándar del número de empleados que necesitan más de una hora para trasladarse de su casa al trabajo.

57. Una empresa piensa entrevistar a los usuarios de Internet para saber cómo será recibida su página por los grupos de las distintas edades. De acuerdo con la Census Bureau, 40% de las personas entre 18 y 54 años y 12% de las personas de 55 años o más usan Internet.
    a. ¿Cuántas personas entre 18 y 54 años hay que contactar para hallar un número esperado de por lo menos 10 usuarios de Internet?
    b. ¿Cuántas personas de 55 años o más hay que contactar para hallar un número esperado de por lo menos 10 usuarios de Internet?
    c. Si se contacta el número de personas entre 18 y 54 años sugerido por el inciso a, ¿cuál es la desviación estándar del número que será usuario de Internet?
    d. Si se contacta el número de personas de entre 55 años o más sugerido por el inciso b, ¿cuál es la desviación estándar del número de quienes serán usuarios de Internet?

58. Muchas empresas usan una técnica de control de calidad conocida como muestreo de aceptación para vigilar los pedidos que reciben de piezas, materia prima, etc. En la industria electrónica, los componentes se suelen recibir por lotes grandes. La inspección de una muestra de *n* componentes se considera como *n* ensayos de un experimento binomial. El resultado de la revisión de cada componente (ensayo) es que el componente sea clasificado como bueno o como defectuoso. Reynolds Electronics acepta el lote de un determinado distribuidor si los componentes defectuosos encontrados en el lote no son más de 1%. Suponga que se prueba una muestra aleatoria de cinco artículos del último lote recibido.
    a. Asuma que 1% del lote recibido está defectuoso. Calcule la probabilidad de que ningún elemento de la muestra esté defectuoso.
    b. Admita que 1% del lote recibido está defectuoso. Calcule la probabilidad de que exactamente un elemento de la muestra esté defectuoso.
    c. ¿Cuál es la probabilidad de encontrar uno o más artículos defectuosos si 1% del lote está defectuoso?
    d. ¿Estaría usted tranquilo al aceptar el lote si se encuentra un artículo defectuoso? ¿Por qué sí o por qué no?

59. La tasa de desempleo es 4.1% (*Barron's*, 4 de septiembre de 2000). Suponga que selecciona aleatoriamente 100 personas empleables.
    a. ¿Cuál es el número esperado de personas que están desempleadas?
    b. ¿Cuál es la varianza y la desviación estándar del número de personas que están desempleadas?

60. Un sondeo de Zogby encontró que de los estadounidenses para quienes la música es "muy importante" en su vida, 30% dice que su estación de radio "siempre" toca la clase de música que le gusta. Suponga que toma una muestra de 800 personas para quienes la música es muy importante en su vida.
    a. ¿Cuántas afirmarían que su estación de radio siempre toca la música que les gusta?
    b. ¿Cuál es la desviación estándar del número de interrogados para quienes su estación de radio siempre toca la música que les gusta?
    c. ¿Cuál es la desviación estándar del número de interrogados para quienes su estación de radio no siempre toca la música que les gusta?

61. A un lavado de coches los automóviles llegan en forma aleatoria e independiente; la probabilidad de una llegada es la misma en cualesquiera dos intervalos de la misma duración. La tasa de llegada media es 15 automóviles por hora. ¿Cuál es la probabilidad de que en una hora cualquiera de operación lleguen 20 o más automóviles?

62. En un proceso nuevo de producción automática hay en promedio 1.5 interrupciones por día. Debido al elevado costo de las interrupciones, los directivos están preocupados por la posibilidad de que en un día haya tres o más interrupciones. Suponga que las interrupciones se presentan en forma aleatoria, que la probabilidad de una interrupción es la misma en cualesquiera dos intervalos de una misma duración y que las interrupciones en un intervalo de tiempo son independientes de

las interrupciones en otro intervalo de tiempo. ¿Cuál es la probabilidad de que haya tres o más interrupciones en un día?

63. Un director regional responsable del desarrollo de los negocios en una determinada área está preocupado por el número de fracasos de pequeños negocios. Si en promedio fracasan 10 pequeños negocios por mes, ¿Cuál es la probabilidad de que exactamente cuatro pequeños negocios fracasen en un mes determinado? Suponga que la probabilidad de fracasos es la misma en cada dos meses que se tomen y que la ocurrencia o no–ocurrencia de fracasos en un determinado mes es independiente de la ocurrencia o no–ocurrencia de fracasos en cualquier otro mes

64. Las llegadas de los clientes a un banco son aleatorias e independientes; la probabilidad de una llegada en un lapso cualquiera de un minuto es la misma que la probabilidad de una llegada en otro lapso cualquiera de un minuto. Conteste las preguntas siguientes suponiendo que la tasa media de llegadas en un lapso de un minuto es tres clientes.
   a. ¿Cuál es la probabilidad de exactamente tres llegadas en un minuto?
   b. ¿Cuál es la probabilidad de por lo menos tres llegadas en un minuto?

65. Una baraja contiene 52 cartas, de las cuales cuatro son ases. ¿Cuál es la probabilidad de que en una repartición de cinco cartas haya:
   a. Un par de ases?
   b. Exactamente un as?
   c. Ningún as?
   d. Por lo menos un as?

66. En la semana que terminó el 16 de septiembre de 2001, Tiger Woods estuvo a la cabeza en ganancia de dinero en el PGA Tour, con una ganancia total de $5 517 777. De los 10 principales jugadores en ganancias de dinero siete usaron pelotas de golf de la marca Titleist (www.pgatour.com). Suponga que toma al azar a dos de estos principales ganadores.
   a. ¿Cuál es la probabilidad de que exactamente uno use una pelota de golf de la marca Titleist?
   b. ¿De que los dos usen una pelota de golf de la marca Titleist?
   c. ¿De que ninguno use una pelota de golf de la marca Titleist?

Apéndice 5.1 Distribuciones de probabilidad con Minitab

Los paquetes para estadística como Minitab ofrecen procedimientos relativamente fáciles y eficientes para calcular probabilidades binomiales. En este apéndice se muestra paso a paso el procedimiento para hallar las probabilidades binomiales del problema de la tienda de ropa Martin Clothing Store de la sección 5.4. Recuerde que las probabilidades binomiales deseadas son para $n = 10$ y $p = 0.30$. Antes de empezar con la rutina de Minitab, el usuario debe ingresar los valores deseados de la variable aleatoria en una columna de la hoja de cálculo. Aquí se han ingresado los valores 0, 1, 2, . . ., 10 en la columna 1 (véase la figura 5.5) para generar la distribución de probabilidad binomial completa. Los pasos para obtener las probabilidades binomiales deseadas usando Minitab son los siguientes.

Paso 1. Seleccionar el menú **Calc**
Paso 2. Elegir **Probability distributions**
Paso 3. Elegir **Binomial**
Paso 4. Cuando aparezca el cuadro de diálogo Binomial Distribution:
Seleccionar **Probability**
Ingresar 10 en el cuadro **Number of trials**
Ingresar 0.3 en el cuadro **Probability of succes**
Ingresar C1 en el cuadro **Input column**
Clic en **OK**

El resultado que da Minitab con las probabilidades binomiales aparecerá como se muestra en la figura 5.5.

De manera similar, Minitab proporciona probabilidades de Poisson e hipergeométricas. Por ejemplo, para calcular probabilidades de Poisson, las únicas diferencias están en el paso 3, en el que se deberá seleccionar la opción **Poisson** y en el paso 4, en el que se deberá ingresar **Mean** en lugar del número de ensayos y la probabilidad de éxito

Apéndice 5.2 Distribuciones de probabilidad discreta con Excel

Excel proporciona funciones para calcular las probabilidades de las distribuciones binomial, de Poisson e hipergeométrica tratadas en este capítulo. La función de Excel para calcular probabilidades binomiales es DISTR.BINOM. Esta función tiene cuatro argumentos: x (el número de éxitos), n (el número de ensayos), p (la probabilidad de éxito) y acumulado. Se usa FALSO como cuarto argumento (acumulado) si se quiere la probabilidad de x éxitos y VERDADERO se usa como cuarto argumento si se desea la probabilidad acumulada de x o menos éxitos. A continuación se muestra cómo calcular la probabilidad de 0 a 10 éxitos en el caso del problema de la tienda de ropa Martin Clothing Store de la sección 5.4 (véase figura 5.5).

A medida que se describe la elaboración de la hoja de cálculo consulte la figura 5.6; la hoja de cálculo con las fórmulas aparece en segundo plano y la hoja de cálculo con los valores en primer plano. En la celda B1 ingrese el número de ensayos (10), en la celda B2 la probabilidad de

FIGURA 5.6 HOJA DE CÁLCULO DE EXCEL PARA CALCULAR PROBABILIDADES BINOMIALES

	A	B	C	D
1	**Number of Trials (*n*)**	10		
2	**Probability of Success (*p*)**	0.3		
3				
4		***x***	***f*(*x*)**	
5		0	=BINOMDIST(B5,B1,B2,FALSE)	
6		1	=BINOMDIST(B6,B1,B2,FALSE)	
7		2	=BINOMDIST(B7,B1,B2,FALSE)	
8		3	=BINOMDIST(B8,B1,B2,FALSE)	
9		4	=BINOMDIST(B9,B1,B2,FALSE)	
10		5	=BINOMDIST(B10,B1,B2,FALSE)	
11		6	=BINOMDIST(B11,B1,B2,FALSE)	
12		7	=BINOMDIST(B12,B1,B2,FALSE)	
13		8	=BINOMDIST(B13,B1,B2,FALSE)	
14		9	=BINOMDIST(B14,B1,B2,FALSE)	
15		10	=BINOMDIST(B15,B1,B2,FALSE)	
16				

	A	B	C	D
1	**Number of Trials (*n*)**	10		
2	**Probability of Success (*p*)**	0.3		
3				
4		***x***	***f*(*x*)**	
5		0	0.0282	
6		1	0.1211	
7		2	0.2335	
8		3	0.2668	
9		4	0.2001	
10		5	0.1029	
11		6	0.0368	
12		7	0.0090	
13		8	0.0014	
14		9	0.0001	
15		10	0.0000	
16				

éxito y en las celdas B5:B15 los valores de la variable aleatoria. Con los pasos siguientes generará las probabilidades deseadas.

Paso 1. Usar la función DISTR.BINOM para calcular la probabilidad de $x = 0$ ingresando la fórmula siguiente en la celda C5:

=BINOMDIST(B5,B1,B2,FALSO)

Paso 2. Copiar la fórmula de la celda C5 en las celdas C6:C15.

La hoja de cálculo con los valores en la figura 5.6 muestra que las probabilidades obtenidas son las mismas que aparecen en la figura 5.5. Las probabilidades de Poisson e hipergeométrica se calculan de manera similar. Se emplean las funciones POISSON y DISTR.HIPERGEOM. La herramienta de Excel Insertar función puede ayudar al usuario a ingresar los argumentos adecuados para estas funciones (véase apéndice 2.2).

CAPÍTULO 6

Distribuciones de probabilidad continua

CONTENIDO

LA ESTADÍSTICA *en* LA PRÁCTICA

PROCTER & GAMBLE*
CINCINNATI, OHIO

Procter & Gamble (P&G) produce y comercializa productos como detergentes, pañales desechables, productos farmacéuticos que no requieren receta, dentífricos, jabones de tocador y toallas de papel. En todo el mundo P&G tiene la marca líder en más categorías que cualquiera otra empresa de productos de consumo. Desde su fusión con Gillette, P&G también comercializa rasuradoras, navajas para afeitar y muchos otros productos para el cuidado personal.

Al ser uno de los líderes en aplicación de los métodos estadísticos para la toma de decisiones, P&G emplea personas con diversas formaciones académicas: ingenieros, especialistas en estadística, en investigación de operaciones y en negocios. Las principales tecnologías cuantitativas en las que estos profesionistas aplican sus conocimientos son decisiones probabilísticas y análisis de riesgos, simulación avanzada, mejoramiento de la calidad y métodos cuantitativos (por ejemplo, programación lineal, análisis de regresión, análisis de probabilidad).

La División de Productos Químicos para la Industria de P&G es una de las principales proveedoras de alcoholes grasos obtenidos de sustancias naturales, como el aceite de coco, y de derivados del petróleo. La división deseaba saber qué riesgos económicos y cuáles oportunidades existen para la expansión de sus instalaciones dedicadas a la producción de alcoholes grasos; por tanto, solicitó la ayuda de los expertos de P&G en decisiones probabilísticas y en análisis de riesgos. Después de estructurar y modelar el problema, los expertos determinaron que la clave para la rentabilidad era la diferencia entre los costos de las materias primas provenientes del petróleo y del coco. Los costos futuros no se podían saber, pero los analistas los calcularon mediante las siguientes variables aleatorias continuas.

x = precio del aceite de coco por libra de alcoholes grasos

y

y = precio de la materia prima proveniente del petróleo por libra de alcoholes grasos

Algunos de los muchos productos de Procter & Gamble son bien conocidos. © AFP/Getty Images.

Como la clave de la rentabilidad era la diferencia entre estas dos variables aleatorias, se empleó una tercera variable aleatoria para el análisis $d = x - y$. Para determinar las distribuciones de probabilidad de x y y entrevistaron a varios expertos. Después, esta información se empleó para elaborar una distribución de probabilidad de la diferencia entre los precios d. En esta distribución de probabilidad continua se encontró que la probabilidad de que la diferencia entre los precios fuera \$0.0655 o menos, era 0.90 y que la probabilidad de que la diferencia entre los precios fuera \$0.035 o menos era 0.50. Además, la probabilidad de que la diferencia fuera \$0.0045 o menos era sólo 0.10.[†]

La dirección de esta división pensó que la clave para alcanzar un consenso estaba en poder cuantificar el impacto de las diferencias entre los precios de las materias primas. Las probabilidades obtenidas se usaron en un análisis sensible a la diferencia entre los precios de las materias primas. Este análisis arrojó suficiente información como para sustentar una recomendación para los directivos.

Usar variables aleatorias continuas y sus distribuciones de probabilidad ayudó a P&G a analizar los riesgos económicos relacionados con su producción de alcoholes grasos. En este capítulo el lector conocerá las variables aleatorias continuas y sus distribuciones de probabilidad, entre ellas una de las distribuciones de probabilidad más importantes en la estadística, la distribución normal.

*Los autores agradecen a Joel Kahn de P&G por proporcionar este artículo para *La estadística en la práctica*.

[†]Las diferencias de precios dadas aquí están modificadas para proteger los datos.

En el capítulo anterior se estudiaron las variables aleatorias discretas y sus distribuciones de probabilidad. En este capítulo se tratan las variables aleatorias continuas. En específico verá tres distribuciones de probabilidad continua: la uniforme, la normal y la exponencial.

Una diferencia fundamental entre las variables aleatorias discretas y las variables aleatorias continuas es cómo se calculan las probabilidades. En las variables aleatorias discretas la función de probabilidad $f(x)$ da la probabilidad de que la variable aleatoria tome un valor determinado. En las variables aleatorias continuas, la contraparte de la función de probabilidad es la **función de densidad de probabilidad**, que también se denota $f(x)$. La diferencia está en que la función de densidad de probabilidad no da probabilidades directamente. Si no que el área bajo la curva de $f(x)$ que corresponde a un intervalo determinado proporciona la probabilidad de que la variable aleatoria tome uno de los valores de ese intervalo. De manera que cuando se calculan probabilidades de variables aleatorias continuas se calcula la probabilidad de que la variable aleatoria tome alguno de los valores dentro de un intervalo.

Como en cualquier punto determinado el área bajo la gráfica de $f(x)$ es cero, una de las consecuencias de la definición de la probabilidad de una variable aleatoria continua es que la probabilidad de cualquier valor determinado de la variable aleatoria es cero. Estos conceptos se demuestran en la sección 6.1 con una variable que tiene una distribución uniforme.

Gran parte del capítulo se dedica a describir y mostrar aplicaciones de la distribución normal. La distribución normal es muy importante por tener muchas aplicaciones y un amplio uso en la inferencia estadística. El capítulo concluye con el estudio de la distribución exponencial. La distribución exponencial es útil en aplicaciones en las que intervienen factores como tiempos de espera y tiempos de servicios.

6.1 Distribución de probabilidad uniforme

Siempre que una probabilidad sea proporcional a la longitud del intervalo, la variable aleatoria estará distribuida uniformemente.

Considere una variable aleatoria x que representa el tiempo de vuelo de un avión que viaja de Chicago a Nueva York. Suponga que el tiempo de vuelo es cualquier valor en el intervalo de 120 minutos a 140 minutos. Dado que la variable aleatoria x toma cualquier valor en este intervalo, x es una variable aleatoria continua y no una variable aleatoria discreta. Admita que cuenta con datos suficientes como para concluir que la probabilidad de que el tiempo de vuelo esté en cualquier intervalo de 1 minuto es el mismo que la probabilidad de que el tiempo de vuelo esté en cualquier otro intervalo de 1 minuto dentro del intervalo que va de 120 a 140 minutos. Como cualquier intervalo de 1 minuto es igual de probable, se dice que la variable aleatoria x tiene una **distribución de probabilidad uniforme**. La función de densidad de probabilidad que define la distribución uniforme de la variable aleatoria tiempo de vuelo, es

$$f(x) = \begin{cases} 1/20 & \text{para } 120 \leq x \leq 140 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

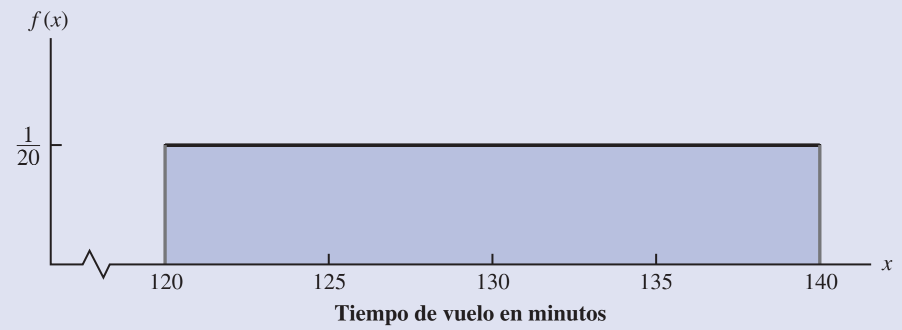
La figura 6.1 es una gráfica de esta función de densidad de probabilidad. En general, la función de densidad de probabilidad uniforme de una variable aleatoria x se define mediante la fórmula siguiente.

FUNCIÓN DE DENSIDAD DE PROBABILIDAD UNIFORME

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{para } a \leq x \leq b \\ 0 & \text{en cualquier otro caso} \end{cases} \tag{6.1}$$

En el caso de la variable aleatoria tiempo de vuelo, $a = 120$ y $b = 140$.

FIGURA 6.1 DISTRIBUCIÓN DE PROBABILIDAD UNIFORME PARA EL TIEMPO DE VUELO

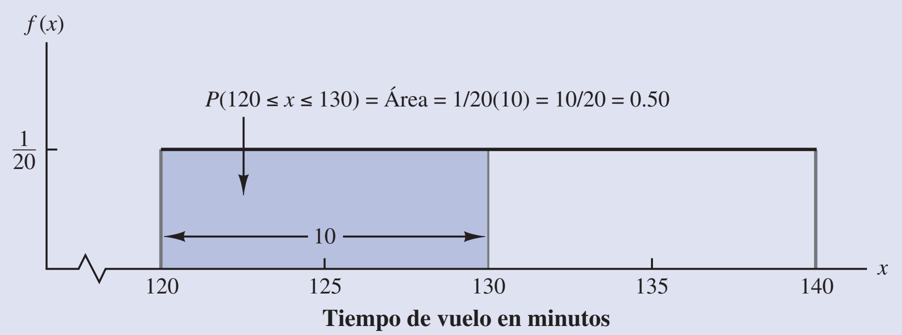


Como se hizo notar en la introducción, en el caso de una variable aleatoria continua, sólo se considera la probabilidad en términos de la posibilidad de que la variable aleatoria tome un valor dentro de un determinado intervalo. En el ejemplo del tiempo de vuelo, una pregunta aceptable acerca de una probabilidad es: ¿Cuál es la probabilidad de que el tiempo de vuelo se encuentre entre 120 y 130 minutos? Es decir, ¿cuál es $P(120 \leq x \leq 130)$? Como el tiempo de vuelo debe estar entre 120 y 140 minutos y como se ha dicho que la probabilidad es uniforme en este intervalo, es factible decir que $P(120 \leq x \leq 130) = 0.50$. En la sección siguiente se muestra que esta probabilidad se calcula como el área bajo la gráfica de $f(x)$ desde 120 hasta 130 (véase figura 6.2)

Áreas como medida de probabilidad

Ahora una observación acerca de la gráfica de la figura 6.2. Considere el área bajo la gráfica de $f(x)$ en el intervalo que va de 120 a 130. Esta área es rectangular y el área de un rectángulo es simplemente el ancho multiplicado por la altura. Si el ancho del intervalo es igual a $130 - 120 = 10$ y la altura es igual al valor de la función de densidad de probabilidad $f(x) = 1/20$, se tiene, área = ancho $\times$ alto = $10(1/20) = 10(1/20) = 10/20 = 0.50$.

FIGURA 6.2 EL ÁREA PROPORCIONA LA PROBABILIDAD DE QUE EL TIEMPO DE VUELO ESTÉ ENTRE 120 Y 130 MINUTOS.



¿Qué observación se puede hacer acerca del área bajo la curva de $f(x)$ y la probabilidad? ¡Son idénticas! En efecto, esta observación es correcta y válida para todas las variables aleatorias continuas. Una vez que se ha dado la función de densidad de probabilidad $f(x)$, la probabilidad de que x tome un valor entre algún valor menor x_1 y otro valor mayor x_2 se encuentra calculando el área bajo la gráfica de $f(x)$ y sobre el intervalo de va de x_1 a x_2.

Dada la distribución uniforme del tiempo de vuelo y usando la interpretación de área como probabilidad es posible contestar cualquier pregunta acerca de la probabilidad de los tiempos de vuelo. Por ejemplo, ¿cuál es la probabilidad de un tiempo de vuelo entre 128 y 136 minutos? El ancho del intervalo es $136 - 128 = 8$. Como la altura uniforme de $f(x) = 1/120$, se ve que $P(128 \leq x \leq 136) = 8(1/20) = 0.40$.

Observe que $P(120 \leq x \leq 140) = 20(1/20) = 1$; es decir, el área total bajo la gráfica de $f(x)$ es igual a 1. Esta propiedad es válida para todas las distribuciones de probabilidad continua y es el análogo de la condición de que la suma de las probabilidades debe ser igual a 1 en el caso de una función de probabilidad discreta.

Dos diferencias importantes sobresalen entre el tratamiento de una variable aleatoria continua y el tratamiento de una variable aleatoria discreta.

Para ver que la probabilidad de un solo punto es 0, consulte la figura 6.2 y calcule la probabilidad de un solo punto, por ejemplo, $x = 125$. $P(x = 125) = P(125 \leq x \leq 125) = 0(1/20) = 0$.

1. Ya no se habla de la probabilidad de que una variable aleatoria tome un determinado valor. Se habla de la probabilidad de que una variable aleatoria tome un valor dentro de un intervalo dado.
2. La probabilidad de que una variable aleatoria continua tome un valor dentro de un determinado intervalo que va de x_1 a x_2 se define como el área bajo la gráfica de la función de densidad de probabilidad entre x_1 y x_2. Como un solo punto es un intervalo cuyo ancho es cero, esto implica que la probabilidad de que una variable aleatoria continua tome un valor exacto, cualquiera, es cero. Esto también significa que en cualquier intervalo la probabilidad de que una variable aleatoria continua tome un valor es la misma, ya sea que se incluyan o no los extremos del intervalo.

El cálculo del valor esperado y de la varianza de una variable aleatoria continua es análogo al de una variable aleatoria discreta. Sin embargo, como en este caso interviene el cálculo integral la deducción de estas fórmulas queda para cursos más avanzados.

En el caso de la distribución de probabilidad continua uniforme presentada en esta sección, las fórmulas para el valor esperado y para la varianza son

$$E(x) = \frac{a + b}{2}$$

$$\text{Var}(x) = \frac{(b - a)^2}{12}$$

En estas fórmulas a es el menor valor y b es el mayor valor que toma la variable aleatoria.

Al aplicar estas fórmulas a la distribución uniforme de los tiempos de vuelo de Chicago a Nueva York, se obtiene,

$$E(x) = \frac{(120 + 140)}{2} = 130$$

$$\text{Var}(x) = \frac{(140 - 120)^2}{12} = 33.33$$

La desviación estándar de los tiempos de vuelo se encuentra sacando la raíz cuadrada de la varianza. Por tanto, $\sigma = 5.77$ minutos.

NOTAS Y COMENTARIOS

Para ver más claramente por qué la altura de una función de densidad de probabilidad no es una probabilidad, considere la variable aleatoria cuya distribución de probabilidad uniforme es la siguiente.

$$f(x) = \begin{cases} 2 & \text{para } 0 \leq x \leq 0.5 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

La altura de la función de densidad de probabilidad, $f(x)$ es 2 para todos los valores de x entre 0 y 0.5. Pero se sabe que las probabilidades nunca pueden ser mayores a 1. Por tanto, $f(x)$ no se interpreta como la probabilidad de x.

Ejercicios

Métodos

Auto examen

1. La variable aleatoria x está distribuida uniformemente entre 1.0 y 1.5.
   a. Dé la gráfica de la función de densidad de probabilidad.
   b. Calcule $P(x = 1.25)$.
   c. Calcule $P(1.0 \leq x \leq 1.25)$.
   d. Calcule $P(1.20 < x < 1.5)$.
2. La variable aleatoria x está distribuida uniformemente entre 10 y 20.
   a. Dé la gráfica de la función de densidad de probabilidad.
   b. Calcule $P(x < 15)$.
   c. Calcule $P(12 \leq x \leq 18)$.
   d. Calcule $E(x)$.
   e. Calcule $\text{Var}(x)$.

Aplicaciones

3. En su vuelo de Cincinati a Tampa, Delta Airlines da como tiempo de vuelo 2 horas, 5 minutos. En realidad los tiempos de vuelo están distribuidos uniformemente entre 2 horas y 2 horas, 20 minutos.
   a. Dé la gráfica de la función de densidad de probabilidad del tiempo de vuelo.
   b. ¿Cuál es la probabilidad de que un vuelo no se retrase más de 5 minutos?
   c. ¿De que un vuelo no se retrase más de 10 minutos?
   d. ¿Cuál es el tiempo de vuelo esperado?

Auto examen

4. La mayoría de los lenguajes de computadora tienen una función para generar números aleatorios. En Excel, la función ALEATORIO se usa para generar números aleatorios entre 0 y 1. Si x denota un número aleatorio generado mediante ALEATORIO, entonces x es una variable aleatoria continua, cuya función de densidad de probabilidad es la siguiente.

$$f(x) = \begin{cases} 1 & \text{para } 0 \leq x \leq 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

   a. Haga la gráfica de la función de densidad de probabilidad.
   b. ¿Cuál es la probabilidad de generar un número aleatorio entre 0.25 y 0.75?
   c. ¿De generar un número aleatorio menor o igual que 0.30?
   d. ¿De generar un número aleatorio mayor o igual que 0.60?
   e. Genere 50 números aleatorios ingresando = ALEATORIO() en 50 celdas de una hoja de cálculo de Excel.
   f. Calcule la media y la desviación estándar de los números del inciso e.

5. La *driving distance* de los 100 mejores golfistas del Tour PGA está entre 284.7 y 310.6 yardas (*Golfweek*, 29 de marzo de 2003). Suponga que las *driving distance* de estos golfistas se encuentran uniformemente distribuidas en este intervalo.
   a. Dé una expresión matemática de la función de densidad de probabilidad correspondiente a estas *driving distance*
   b. ¿Cuál es la probabilidad de que la *driving distance* de uno de estos golfistas sea menor que 290 yardas?
   c. ¿De que la *driving distance* de uno de estos golfistas sea por lo menos de 300 yardas?
   d. ¿De que la *driving distance* de uno de estos golfistas esté entre 290 y 305 yardas?
   e. ¿Cuántos de estos jugadores lanzan la pelota por lo menos a 290 yardas?
6. En las botellas de un detergente líquido se indica que el contenido es de 12 onzas por botella. En la operación de producción se llenan las botellas uniformemente de acuerdo con la siguiente función de densidad de probabilidad.

$$f(x) = \begin{cases} 8 & \text{para } 11.975 \leq x \leq 12.100 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

   a. ¿Cuál es la probabilidad de que el contenido de una botella esté entre 12 y 12.05 onzas?
   b. ¿De que el contenido de una botella sea 12.02 onzas o más?
   c. En el control de calidad se acepta que una botella sea llenada con más o menos 0.02 onzas de lo indicado en la etiqueta. ¿Cuál es la probabilidad de que una de las botellas de detergente no satisfaga estos estándares?
7. Suponga que quiere comprar un terreno y sabe que también hay otros compradores interesados.* El vendedor revela que aceptará la oferta mayor que sea superior a $10 000. Si la oferta del competidor x es una variable aleatoria que está uniformemente distribuida entre $10 000 y $15 000.
   a. Asuma que usted ofrece $12 000. ¿Cuál es la probabilidad de que su oferta sea aceptada?
   b. Si usted ofrece $14 000. ¿Cuál es la probabilidad de que su oferta sea aceptada?
   c. ¿Cuál es la cantidad que deberá ofrecer para maximizar la probabilidad de obtener la propiedad?
   d. Suponga que conoce a quien está dispuesto a pagar $16 000 por la propiedad. ¿Consideraría la posibilidad de ofrecer una cantidad menor que la del inciso c?

6.2 Distribución de probabilidad normal

Abraham de Moivre, un matemático francés, publicó en 1733 Doctrina de las posibilidades*. De Moivre dedujo la distribución normal.*

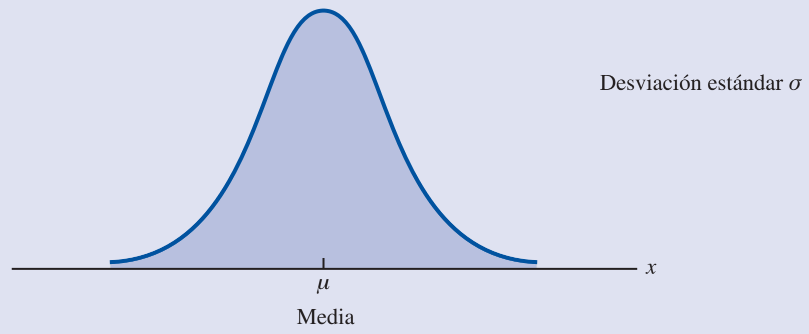
La distribución de probabilidad más usada para describir variables aleatorias continuas es la **distribución de probabilidad normal**. La distribución normal tiene gran cantidad de aplicaciones prácticas, en las cuales la variable aleatoria puede ser el peso o la estatura de las personas, puntuaciones de exámenes, resultados de mediciones científicas, precipitación pluvial u otras cantidades similares. La distribución normal también tiene una importante aplicación en inferencia estadística, tema principal del resto de este libro. En estas aplicaciones, la distribución normal describe qué tan probables son los resultados obtenidos de un muestreo

Curva normal

En la figura 6.3 aparece la forma de la distribución normal, una curva normal en forma de campana. A continuación se presenta la función de densidad de probabilidad que define la curva en forma de campana de la distribución normal.

* Este ejercicio está basado en un problema sugerido por el profesor Roger Myerson de la Universidad de Northwestern.

FIGURA 6.3 CURVA EN FORMA DE CAMPANA DE UNA DISTRIBUCIÓN NORMAL



FUNCIÓN DE DENSIDAD DE PROBABILIDAD NORMAL

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad \textbf{(6.2)}$$

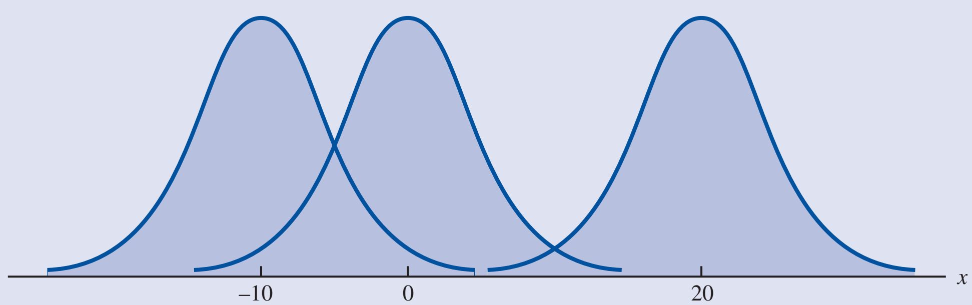
donde

$$\begin{aligned} \mu &= \text{media} \\ \sigma &= \text{desviación estándar} \\ \pi &= 3.14159 \\ e &= 2.71828 \end{aligned}$$

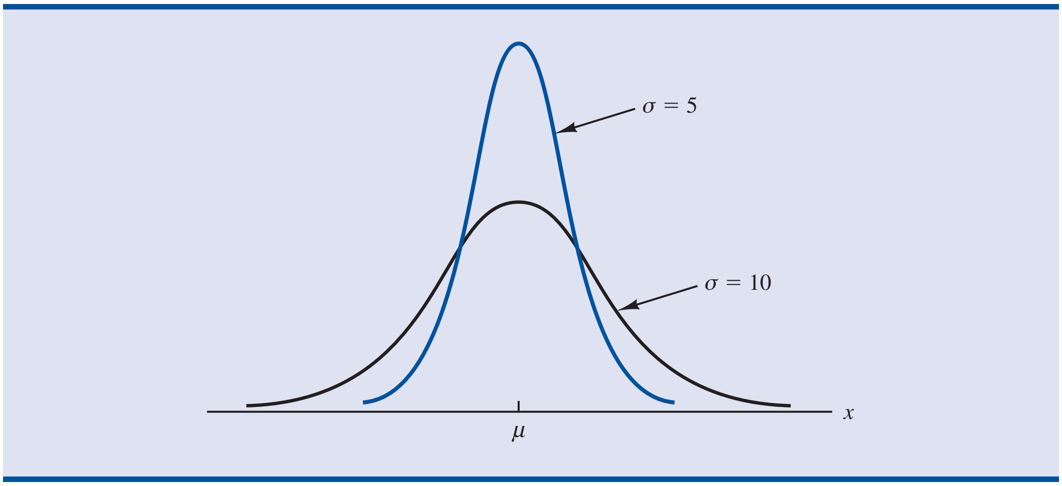
Las siguientes son observaciones importantes acerca de las características de las distribuciones normales.

La curva normal tiene dos parámetros, μ y σ. Estos parámetros determinan la localización y la forma de la distribución normal.

1. Toda la familia de distribuciones normales se diferencia por medio de dos parámetros: la media μ y la desviación estándar σ.
2. El punto más alto de una curva normal se encuentra sobre la media, la cual coincide con la mediana y la moda.
3. La media de una distribución normal puede tener cualquier valor: negativo, positivo o cero. A continuación se muestran tres distribuciones normales que tienen la misma desviación estándar, pero diferentes medias. (−10, 0 y 20).



4. La distribución normal es simétrica, siendo la forma de la curva normal al lado izquierdo de la media, la imagen especular de la forma al lado derecho de la media. Las colas de la curva normal se extienden al infinito en ambas direcciones y en teoría jamás tocan el eje horizontal. Dado que es simétrica, la distribución normal no es sesgada; su sesgo es cero.
5. La desviación estándar determina qué tan plana y ancha es la curva normal. Desviaciones estándar grandes corresponden a curvas más planas y más anchas, lo cual indica mayor variabilidad en los datos. A continuación se muestran dos curvas normales que tienen la misma media pero distintas desviaciones estándar.



6. Las probabilidades correspondientes a la variable aleatoria normal se dan mediante áreas bajo la curva normal. Toda el área bajo la curva de una distribución normal es 1. Como esta distribución es simétrica, el área bajo la curva y a la izquierda de la media es 0.50 y el área bajo la curva y a la derecha de la media es 0.50.
7. Los porcentajes de los valores que se encuentran en algunos intervalos comúnmente usados son:
   a. 68.3% de los valores de una variable aleatoria normal se encuentran más o menos una desviación estándar de la media.
   b. 95.4% de los valores de una variable aleatoria normal se encuentran más o menos dos desviaciones estándar de la media.
   c. 99.7% de los valores de una variable aleatoria normal se encuentran más o menos tres desviaciones estándar de la media.

Estos porcentajes son la base de la regla empírica que se presentó en la sección 3.3.

En la figura 6.4 aparece una gráfica de las propiedades a, b y c.

Distribución de probabilidad normal estándar

Una variable aleatoria que tiene una distribución normal con una media cero y desviación estándar de uno tiene una **distribución normal estándar**. Para designar esta variable aleatoria normal se suele usar la letra z. La figura 6.5 es la gráfica de la distribución normal estándar. Esta distribución tiene el mismo aspecto general que cualquier otra distribución normal, pero tiene las propiedades especiales, $\mu = 0$ y $\sigma = 1$.

FIGURA 6.4 ÁREAS BAJO LA CURVA DE CUALQUIER DISTRIBUCIÓN NORMAL

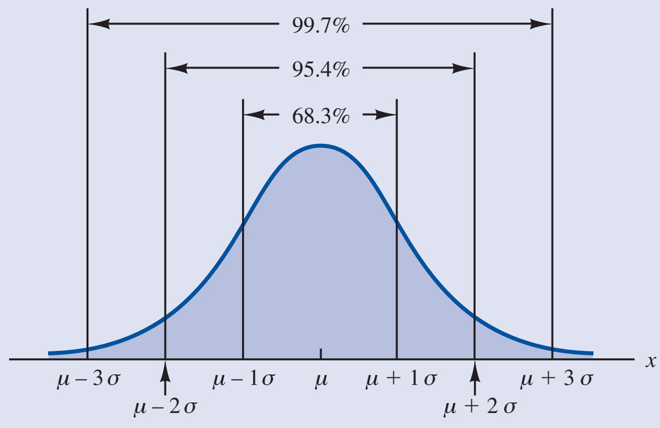
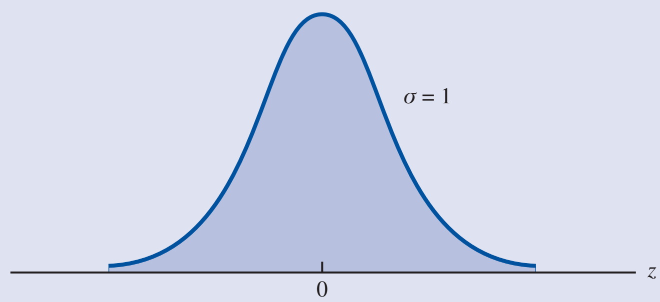


FIGURA 6.5 DISTRIBUCIÓN NORMAL ESTÁNDAR



Dado que $\mu = 0$ y $\sigma = 1$, la fórmula de la función de densidad de probabilidad normal estándar es una versión más simple de la ecuación (6.2).

FUNCIÓN DE DENSIDAD NORMAL ESTÁNDAR

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

Como ocurre con otras variables aleatorias continuas, los cálculos de la probabilidad en cualquier distribución normal se hacen calculando el área bajo la gráfica de la función de densidad de probabilidad. Por tanto, para hallar la probabilidad de que una variable aleatoria normal esté dentro de un determinado intervalo, se tiene que calcular el área que se encuentra bajo la curva normal y sobre ese intervalo.

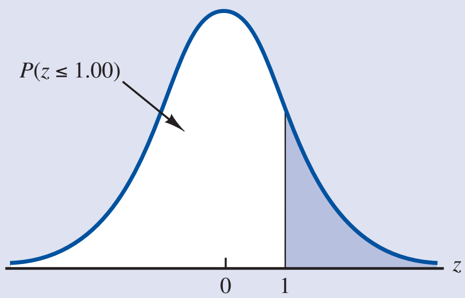
En la función de densidad de probabilidad normal, la altura de la curva varía y para calcular las áreas que representan las probabilidades se requiere de matemáticas más avanzadas.

Para la distribución normal estándar ya se encuentran calculadas las áreas bajo la curva normal y se cuenta con tablas que dan estas áreas y que se usan para calcular las probabilidades. Estas tablas se encuentran en los forros interiores al inicio del libro. La tabla del forro izquierdo contiene áreas, o probabilidades acumuladas, correspondientes a valores de z menores o iguales a la media, cero. La tabla siguiente contiene áreas, o probabilidades acumuladas, correspondientes a valores de z mayores o iguales a la media de cero.

Los tres tipos de probabilidades que se necesitan calcular son: (1) la probabilidad de que la variable aleatoria normal estándar z sea menor o igual que un valor dado; (2) la probabilidad de que z esté entre dos valores dados, y (3) la probabilidad de que z sea mayor o igual que un valor dado. Para mostrar el uso de las tablas de probabilidad acumulada de la distribución normal estándar en el cálculo de estos tres tipos de probabilidades, se consideran algunos ejemplos.

Debido a que la variable aleatoria normal estándar es continua, $P(z \leq 1.00) = P(z < 1.00)$.

Se empieza por mostrar cómo se calcula la probabilidad de que z sea menor o igual a 1.00; es decir $P(z \leq 1.00)$. Esta probabilidad acumulada es el área bajo la curva normal a la izquierda de $z = 1.00$ como se muestra en la gráfica siguiente.

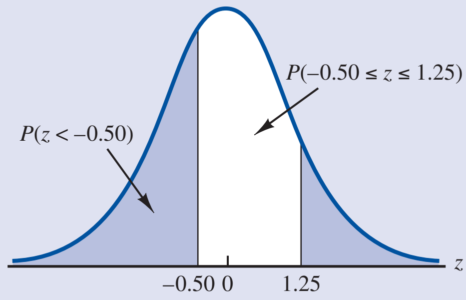


Consulte la página del lado derecho de la tabla de la distribución de probabilidad normal estandar que se encuentra dentro de la cubierta frontal del libro. Esta probabilidad acumulada correspondiente a $z = 1.00$ es el valor que en la tabla se localiza en la intersección del renglón cuyo encabezado es 1.0 y la columna cuyo encabezado es 0.00. Primero localice 1.0 en la columna del extremo izquierdo de la tabla y después localice 0.00 en el renglón en la parte superior de la tabla. Observe que en el interior de la tabla, el renglón 1.0 y la columna 0.00 se cruzan en el valor 0.8413; por tanto, $P(z \leq 1.00) = 0.8413$. Estos pasos se muestran en el extracto siguiente de las tablas de probabilidad.

z	**0.00**	**0.01**	**0.02**
.			
.			
.			
0.9	0.8159	0.8186	0.8212
1.0	$\boxed{\mathbf{0.8413}}$	0.8438	0.8461
1.1	0.8643	0.8665	0.8686
1.2	0.8849	0.8869	0.8888
.			
.			
.			

$P(z \leq 1.00)$

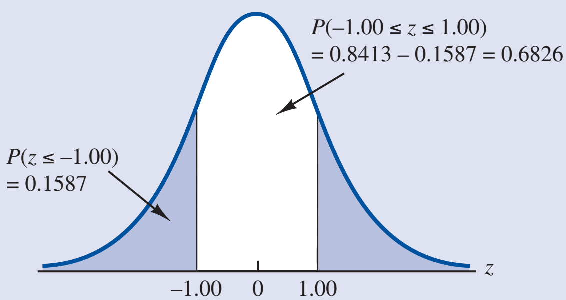
Para ilustrar el segundo tipo de cálculo de una probabilidad se muestra cómo calcular la probabilidad de que z esté en el intervalo entre -0.50 y 1.25; esto es, $P(-.50 \leq z \leq 1.25)$. En la gráfica siguiente se muestra esta área o probabilidad.



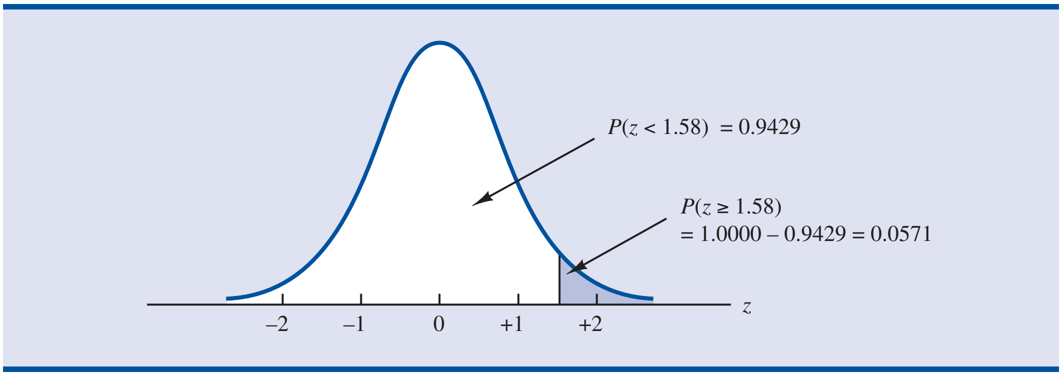
Para calcular esta probabilidad son necesarios tres pasos. Primero, se encuentra el área bajo la curva normal a la izquierda de $z = 1.25$. Segundo, se encuentra el área bajo la curva normal a la izquierda de $z = -0.50$. Por último, se resta el área a la izquierda de $z = -0.50$ del área a la izquierda de $z = 1.25$ y se encuentra, $P(-0.50 \le z \le 1.25)$.

Para encontrar el área bajo la curva normal a la izquierda de $z = 1.25$, primero se localiza en la tabla de probabilidad normal estándar el renglón 1.2 y después se avanza por ese renglón hasta la columna 0.05. Como el valor que aparece en el renglón 1.2 columna 0.05 es 0.8944, $P(z \le 1.25) = 0.8944$. De manera similar, para encontrar el área bajo la curva a la izquierda de $z = -0.50$ se usa el forro izquierdo de la tabla para localizar el valor en el renglón -0.5 columna 0.00; como el valor que se encuentra es 0.3085, $P(z \le -0.50) = 0.3085$. Por tanto, $P(-0.50 \le z \le 1.25) = P(z \le 1.25) - P(z \le -0.50) = 0.8944 - 0.3085 = 0.5859$.

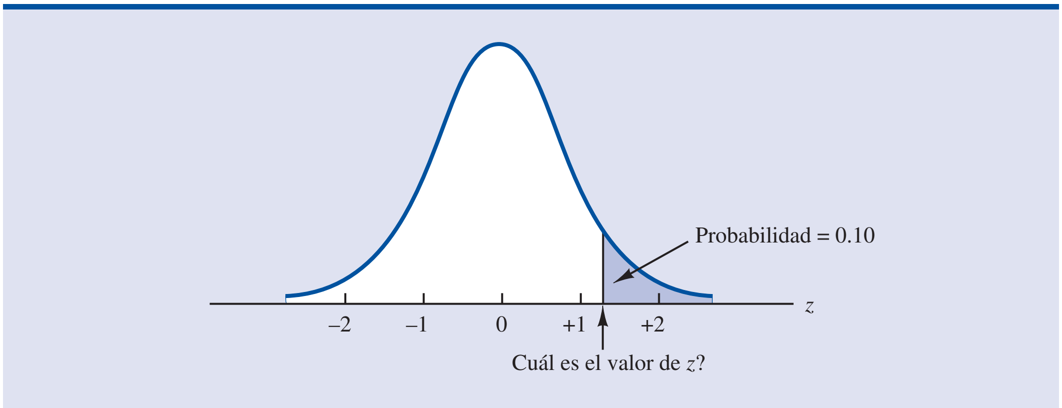
A continuación se presenta otro ejemplo para calcular la probabilidad de que z esté en el intervalo entre dos valores dados. Con frecuencia se desea calcular la probabilidad de que una variable aleatoria normal tome un valor dentro de cierto número de desviaciones estándar respecto a la media. Suponga que desea calcular la probabilidad de que la variable aleatoria normal estándar se encuentre a no más de una desviación estándar de la media; es decir, $P(-1.00 \le z \le 1.00)$. Para calcular esta probabilidad tiene que hallar el área bajo la curva entre -1.00 y 1.00. Antes encontró que $P(z \le 1.00) = 0.8413$. Si va al forro izquierdo, encontrará que el área bajo la curva a la izquierda de $z = -1.00$ es 0.1587, de manera que $P(z \le -1.00) = 0.1587$. Por tanto, $P(-1.00 \le z \le 1.00) = P(z \le 1.00) - P(z \le -1.00) = 0.8413 - 0.1587 = 0.6826$. Esta probabilidad se muestra en forma gráfica en la figura siguiente.



Para ilustrar cómo se calcula el tercer tipo de probabilidad, suponga que desea calcular la probabilidad de tener un valor z por lo menos igual a 1.58; es decir, $P(z \geq 1.58)$. El valor en el renglón $z = 1.5$, columna 0.08 de la tabla normal acumulada es 0.9429; por tanto, $P(z < 1.58) = 0.9429$. Pero, como toda el área bajo la curva normal es 1, $P(z \geq 1.58) = 1 - 0.9429 = 0.0571$. En la figura siguiente se muestra esta probabilidad.



En los ejemplos anteriores se muestra cómo calcular probabilidades dados determinados valores de z. En algunas situaciones se da una probabilidad y se trata de hacer lo contrario, encontrar el correspondiente valor de z. Suponga que desea hallar un valor z tal que la probabilidad de obtener un valor z mayor sea 0.10. En la figura siguiente se muestra en forma gráfica esta situación.



Dada una probabilidad, se puede usar la tabla normal estándar para encontrar el correspondiente valor de z.

Este problema es la situación contraria a la presentada en los ejemplos anteriores, en ellos se dio el valor z y se halló la probabilidad o área correspondiente. En este ejemplo se da la probabilidad, o el área, y se pide hallar el valor correspondiente de z. Para esto se usa la tabla de probabilidad normal estándar de una manera un poco diferente.

Recuerde que la tabla de probabilidad normal estándar da el área bajo la curva a la izquierda de un determinado valor z. Se ha recibido la información de que el área en la cola superior (derecha) de la curva es 0.10. Por tanto, el área bajo la curva a la izquierda del valor desconocido de z debe ser 0.9000. Al recorrer el cuerpo de la tabla, se encuentra que 0.8997 es la probabilidad acumulada más cercana a 0.9000. A continuación se reproduce la sección de la tabla en la que se encuentra este resultado.

z	**0.06**	**0.07**	**0.08**	**0.09**
.				
.				
.				
1.0	0.8554	0.8577	0.8599	0.8621
1.1	0.8770	0.8790	0.8810	0.8830
1.2	0.8962	0.8980	$\boxed{0.8997}$	0.9015
1.3	0.9131	0.9147	0.9162	0.9177
1.4	0.9279	0.9292	0.9306	0.9319
.				
.				
.				

Probabilidad acumulada más cercana a 0.9000

Al leer el valor de z en la columna del extremo izquierdo y en el renglón superior de la tabla, se encuentra que el valor de z es 1.28. De manera que un área de aproximadamente 0.9000 (en realidad de 0.8997) es la que se encuentra a la izquierda de $z = 1.28$.* En términos de la pregunta originalmente planteada, 0.10 es la probabilidad aproximada de que z sea mayor que 1.28.

Estos ejemplos ilustran que la tabla de probabilidades acumuladas para la distribución de probabilidad normal estándar se puede usar para hallar probabilidades correspondientes a valores de la variable aleatoria normal estándar z. Es posible hacer dos tipos de preguntas. En el primer tipo de pregunta se dan valores, o un valor de z, y se pide usar la tabla para determinar el área o probabilidad correspondiente. En el segundo tipo de pregunta se da un área, o probabilidad, y se pide usar la tabla para encontrar el correspondiente valor de z. Por tanto, se necesita tener flexibilidad para usar la tabla de probabilidad normal estándar para responder la pregunta deseada. En la mayoría de los casos, hacer un bosquejo de la gráfica de la distribución de probabilidad normal estándar y sombrear el área deseada será una ayuda para visualizar la situación y encontrar la respuesta correcta.

Cálculo de probabilidades en cualquier distribución de probabilidad normal

La razón por la cual la distribución normal estándar se ha visto de manera tan amplia es que todas las distribuciones normales son calculadas mediante la distribución normal estándar. Esto es, cuando distribución normal con una media μ cualquiera y una desviación estándar σ cualquiera, las preguntas sobre las probabilidades en esta distribución se responden pasando primero a la distribución normal estándar. Use las tablas de probabilidad normal estándar y los valores apropiados de z para hallar las probabilidades deseadas. A continuación se da la fórmula que se emplea para convertir cualquier variable aleatoria x con media μ y desviación estándar σ en la variable aleatoria normal estándar z.

La fórmula para la variable aleatoria normal estándar es semejante a la fórmula que se dio en el capítulo 3 para los puntos z de un conjunto de datos.

CONVERSIÓN A LA VARIABLE ALEATORIA NORMAL ESTÁNDAR

$$z = \frac{x - \mu}{\sigma} \tag{6.3}$$

* Se podía haber hecho una interpolación en el cuerpo de la tabla para obtener una aproximación más exacta al valor z que corresponde al área 0.9000. Al hacerlo en busca de un lugar decimal más preciso se obtiene 1.282. Sin embargo, en la mayor parte de las situaciones prácticas, es suficiente con la precisión obtenida usando el valor más cercano al valor deseado que da la tabla.

Un valor x igual a su media μ da como resultado $z = (\mu - \mu)/\sigma = 0$. De manera que un valor x igual a su media corresponde a $z = 0$. Ahora suponga que x se encuentra una desviación estándar arriba de su media. Es decir, $x = \mu + \sigma$. Aplicando la ecuación (6.3) el valor correspondiente es $z = [(\mu + \sigma) - \mu]/\sigma = \sigma/\sigma = 1$. Así que un valor de x que es una desviación estándar mayor que su media corresponde a $z = 1$. En otras palabras, *z se interpreta como el número de desviaciones estándar a las que está una variable aleatoria* x *de su media μ.*

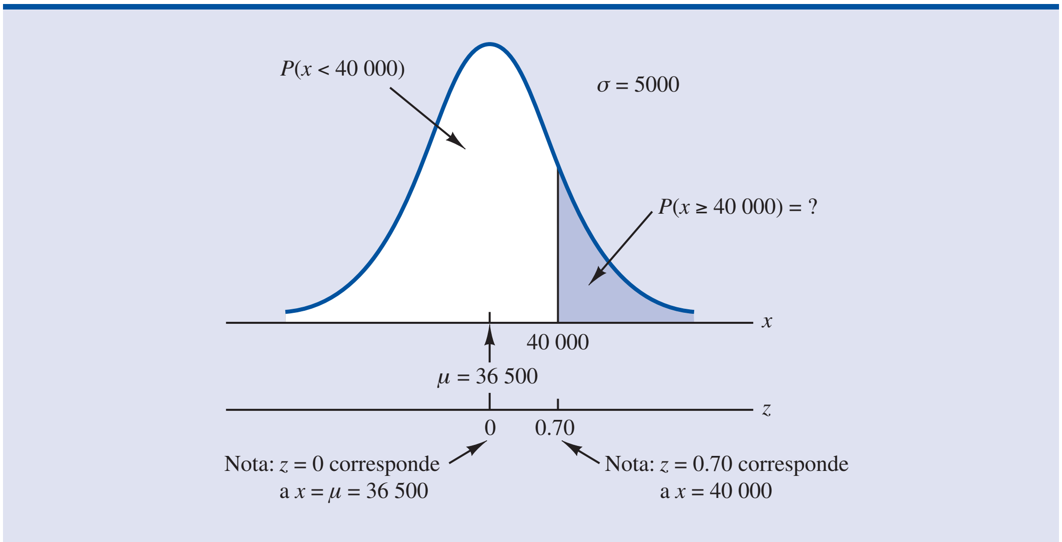
Para ver cómo esta distribución permite calcular probabilidades en cualquier distribución normal, admita que tiene una distribución en la que $\mu = 10$ y $\sigma = 2$. ¿Cuál es la probabilidad de que la variable aleatoria x esté entre 10 y 14? Empleando la ecuación (6.3) se ve que para $z = (x - \mu)/\sigma = (10 - 10)/2 = 0$ y que para $x = 14$, $z = (14 - 10)/2 = 4/2 = 2$. Así, la respuesta a la pregunta acerca de la probabilidad de que x esté entre 10 y 14 está dada por la probabilidad equivalente de que z esté entre 0 y 2 en la distribución normal estándar. En otras palabras, la probabilidad que se está buscando es que la variable aleatoria x esté entre su media y dos desviaciones estándar arriba de la media. Usando $z = 2$ y la tabla de probabilidad normal estándar del forro interior, se ve que $P(z \leq 2) = 0.9772$. Como $P(z \leq 0) = 0.5000$, se tiene que $P(0.00 \leq z \leq 2.00) = P(z \leq 2) - P(z \leq 0) = 0.9772 - 0.5000 = 0.4772$. Por tanto, la probabilidad de que x esté entre 10 y 14 es 0.4772.

El problema de la empresa Grear Tire

Para una aplicación de la distribución de probabilidad normal, suponga que Grear Tire Company ha fabricado un nuevo neumático que será vendido por una cadena nacional de tiendas de descuento. Como este neumático es un producto nuevo, los directivos de Grear piensan que la garantía de duración será un factor importante en la aceptación del neumático. Antes de finalizar la póliza de garantía, los directivos necesitan información probabilística acerca de x = duración del neumático en número de millas.

De acuerdo con las pruebas realizadas al neumático, los ingenieros de Grear estiman que la duración media en millas es $\mu = 36\,500$ millas y que la desviación estándar es $\sigma = 5000$. Además, los datos recogidos indican que es razonable suponer una distribución normal. ¿Qué porcentaje de los neumáticos se espera que duren más de 40 000 millas? En otras palabras, ¿cuál es la probabilidad de que la duración de los neumáticos sea superior a 40 000? Esta pregunta se responde hallando el área de la región sombreada que se observa en la gráfica de la figura 6.6.

FIGURA 6.6 DISTRIBUCIÓN DE DURACIÓN EN MILLAS PARA GREAR TIRE COMPANY



Para $x = 40\,000$, se tiene

$$z = \frac{x - \mu}{\sigma} = \frac{40\,000 - 36\,500}{5000} = \frac{3500}{5000} = 0.70$$

Observe que en la parte inferior de la figura 6.6 el valor $x = 40\,000$ en la distribución normal de Grear Tire corresponde a $z = 0.70$ en la distribución normal estándar. Mediante la tabla de probabilidad normal estándar se encuentra que el área bajo la curva normal estándar a la izquierda de $z = 0.70$ es 0.7580. De manera que $1.000 - 0.7580 = 0.2420$ es la probabilidad de que z sea mayor a 0.70 y por tanto de que x sea mayor a 40 000. Entonces 24.2% de los neumáticos durará más de 40 000 millas.

Ahora suponga que Grear está considerando una garantía que dé un descuento en la sustitución del neumático original si éste no dura lo que asegura la garantía. ¿Cuál deberá ser la duración en millas especificada en la garantía si Grear desea que no más de 10% de los neumáticos alcancen la garantía? En la figura 6.7 se plantea esta pregunta en forma gráfica.

De acuerdo con la figura 6.7, el área bajo la curva a la izquierda de la cantidad desconocida de millas para la garantía debe ser 0.10. De manera que primero se debe encontrar el valor de z que deja un área de 0.10 en el extremo de la cola izquierda de la distribución normal estándar. Según la tabla de probabilidad normal estándar $z = -1.28$ deja un área de 0.10 en el extremo de la cola izquierda. Por tanto, $z = -1.28$ es el valor de la variable aleatoria normal estándar que corresponde a las millas de duración deseadas para la garantía en la distribución normal de Grear Tire. Para hallar el valor de x que corresponde a $z = -1.28$, se tiene

Las millas de garantía que se desean encontrar están a 1.28 desviaciones estándar abajo de la media. Por tanto, $x = \mu - 1.28\sigma$.

$$z = \frac{x - \mu}{\sigma} = -1.28$$
$$x - \mu = -1.28\sigma$$
$$x = \mu - 1.28\sigma$$

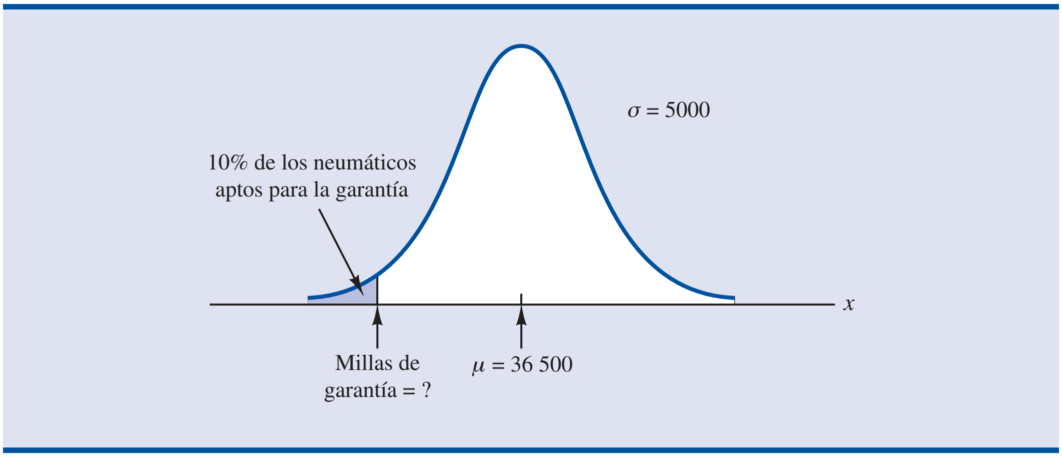
Como $\mu = 36\,500$ y $s = 5000$,

$$x = 36\,500 - 1.28(5000) = 30\,100$$

Al establecer una garantía a partir de 30 000 millas, el porcentaje real apto para la garantía será 9.68%.

Por tanto, una garantía de 30 100 millas cumplirá con el requerimiento de que aproximadamente 10% de los neumáticos sean aptos para la garantía. Con esta información, quizá la empresa establezca una garantía de 30 000 millas.

FIGURA 6.7 GARANTÍA DE GREAR



Nuevamente, se observa la importancia de las distribuciones de probabilidad en el suministro de información para la toma de decisiones. A saber, una vez que la distribución de probabilidad es establecida para una aplicación en particular, puede ser usada para obtener información probabilística acerca del problema. La probabilidad no recomienda directamente una decisión, pero suministra información que ayuda a tomarla entendiendo mejor los riesgos y la incertidumbre asociados al problema. Finalmente, esta información ayuda a enriquecer una buena decisión.

Ejercicios

Métodos

8. Usando como guía la figura 6.4, dibuje la curva normal de la variable aleatoria x cuya media es $\mu = 100$ con desviación estándar de $\sigma = 10$. Indique en el eje horizontal los valores 70, 80, 90, 100, 110, 120 y 130.
9. Una variable aleatoria es normalmente distribuida con media $\mu = 50$ y desviación estándar $\sigma = 5$.
   a. Dibuje la curva normal de la función de densidad de probabilidad. En el eje horizontal dé los valores 35, 40, 45, 50, 55, 60 y 65. En la figura 6.4 se observa que la curva normal casi toca el eje horizontal en los puntos que se encuentran tres desviaciones estándar arriba de la media y tres desviaciones estándar debajo de la media (en este caso en 35 y 65).
   b. ¿Cuál es la probabilidad de que la variable aleatoria tome un valor entre 45 y 55?
   c. ¿De que la variable aleatoria tome un valor entre 40 y 60?
10. Dibuje la gráfica de la distribución normal estándar. Etiquete el eje horizontal con los valores -3, -2, -1, 0, 1, 2 y 3. Después use la tabla de probabilidades de la distribución normal estándar que se encuentra en el forro interior del libro para calcular las probabilidades siguientes.
    a. $P(z \leq 1.5)$
    b. $P(z \leq 1)$
    c. $P(1 \leq z \leq 1.5)$
    d. $P(0 < z < 2.5)$
11. Dado que z es la variable normal estándar, calcule las probabilidades siguientes.
    a. $P(z \leq -1.0)$
    b. $P(z \geq -1)$
    c. $P(z \geq -1.5)$
    d. $P(-2.5 \leq z)$
    e. $P(-3 < z \leq 0)$
12. Dado que z es la variable normal estándar, calcule las probabilidades siguientes.
    a. $P(0 \leq z \leq 0.83)$
    b. $P(-1.57 \leq z \leq 0)$
    c. $P(z > 0.44)$
    d. $P(z \geq -0.23)$
    e. $P(z < 1.20)$
    f. $P(z \leq -0.71)$

Autoexamen

13. Dado que z es la variable normal estándar, calcule las probabilidades siguientes.
    a. $P(-1.98 \leq z \leq 0.49)$
    b. $P(0.52 \leq z \leq 1.22)$
    c. $P(-1.75 \leq z \leq -1.04)$
14. Dado que z es la variable normal estándar, encuentre z en cada una de las situaciones siguientes.
    a. El área a la izquierda de z es 0.9750.
    b. El área entre 0 y z es 0.4750.
    c. El área a la izquierda de z es 0.7291.
    d. El área a la derecha de z es 0.1314.
    e. El área a la izquierda de z es 0.6700.
    f. El área a la derecha de z es 0.3300.

Autoexamen

15. Dado que z es la variable normal estándar, halle z en cada una de las situaciones siguientes.
    a. El área a la izquierda de z es 0.2119
    b. El área entre $-z$ y z es 0.9030.
    c. El área entre $-z$ y z es 0.2052.
    d. El área a la izquierda de z es 0.9948.
    e. El área a la derecha de z es 0.6915.
16. Dado que z es la variable normal estándar, encuentre z en cada una de las situaciones siguientes.
    a. El área a la derecha de z es 0.01
    b. El área a la derecha de z es 0.025.
    c. El área a la derecha de z es 0.05.
    d. El área a la derecha de z es 0.10.

Aplicaciones

17. Una persona con una buena historia crediticia tiene una deuda promedio de $15 015 (*BusinessWeek*, 20 de marzo de 2006). Suponga que la desviación estándar es de $3 540 y que los montos de las deudas están distribuidos normalmente.
    a. ¿Cuál es la probabilidad de que la deuda de una persona con buena historia crediticia sea mayor a $18 000?
    b. ¿De que la deuda de una persona con buena historia crediticia sea de menos de $10 000?
    c. ¿De que la deuda de una persona con buena historia crediticia esté entre $12 000 y $18 000?
    d. ¿De que la deuda de una persona con buena historia crediticia sea mayor a $14 000?

Autoexamen

18. El precio promedio de las acciones que pertenecen a S&P500 es de $30 y la desviación estándar es $8.20 (*BusinessWeek*, Special Annual Issue, primavera de 2003). Suponga que los precios de las acciones están distribuidos normalmente.
    a. ¿Cuál es la probabilidad de que el precio de las acciones de una empresa sea por lo menos de $40?
    b. ¿De que el precio de las acciones de una empresa no sea mayor a $20?
    c. ¿De cuánto deben ser los precios de las acciones de una empresa para que esté entre las 10% mejores?
19. La cantidad promedio de precipitación pluvial en Dallas, Texas, durante el mes de abril es 3.5 pulgadas (*The World Almanac*, 2000). Suponga que se puede usar una distribución normal y que la desviación estándar es 0.8 pulgadas.
    a. ¿Qué porcentaje del tiempo la precipitación pluvial en abril es mayor que 5 pulgadas?
    b. ¿Qué porcentaje del tiempo la precipitación pluvial en abril es menor que 3 pulgadas?
    c. Un mes se considera como extremadamente húmedo si la precipitación pluvial es 10% superior para ese mes. ¿Cuánta debe ser la precipitación pluvial en abril para que sea considerado un mes extremadamente húmedo?
20. En enero de 2003 un empleado estadounidense pasaba, en promedio, 77 horas conectado a Internet durante las horas de trabajo (CNBC, 15 de marzo de 2003). Suponga que la media poblacional es 77 horas, tiempos que están distribuidos normalmente y que la desviación estándar es 20 horas.
    a. ¿Cuál es la probabilidad de que en enero de 2003 un empleado seleccionado aleatoriamente haya pasado menos de 50 horas conectado a Internet?
    b. ¿Qué porcentaje de los empleados pasó en enero de 2003 más de 100 horas conectado a Internet?
    c. Un usuario es clasificado como intensivo si se encuentra en el 20% superior de uso. ¿Cuántas horas tiene un empleado que haber estado conectado a Internet en enero de 2003 para que se le considerara un usuario intensivo?
21. La puntuación de una persona en una prueba de IQ debe estar en el 2% superior para que sea clasificado como miembro del grupo Mensa, la sociedad internacional de IQ elevado (*U.S. Airways Attaché*, septiembre de 2000). Si las puntuaciones de IQ tienen una distribución normal con una media de 100 y desviación estándar de 15, ¿cuál debe ser la puntuación de una persona para que se le considere miembro del grupo Mensa?

22. La tasa de remuneración media por hora para administrativos financieros en una determinada región es \$32.62 y la desviación estándar es \$2.32 (Bureau of Labor Statistics, septiembre de 2005). Suponga que estas tasas de remuneración están distribuidas normalmente.
    a. ¿Cuál es la probabilidad de que un directivo financiero tenga una remuneración entre \$30 y \$35 por hora?
    b. ¿Qué tan alta debe ser la remuneración por hora para que un directivo financiero tenga un pago 10% superior?
    c. ¿Cuál es la probabilidad de que la remuneración por hora de un directivo financiero sea menos de \$28 por hora?
23. El tiempo necesario para hacer un examen final en un determinado curso de una universidad tiene una distribución normal cuya media es 80 minutos con desviación estándar de 10 minutos. Conteste las preguntas siguientes
    a. ¿Cuál es la probabilidad de terminar el examen en una hora o menos?
    b. ¿Cuál es la probabilidad de que un estudiante termine el examen en más de 60 minutos pero en menos de 75 minutos?
    c. Suponga que en la clase hay 60 estudiantes y que el tiempo para resolver el examen es de 90 minutos. ¿Cuántos estudiantes piensa usted que no podrán terminar el examen en este tiempo?

archivo en CD
Volume

24. El volumen de negociaciones en la Bolsa de Nueva York es más intenso en la primera media hora (en la mañana temprano) y la última media hora (al final de la tarde) de un día de trabajo. A continuación se presentan los volúmenes (en millones de acciones) de 13 días de enero y febrero.

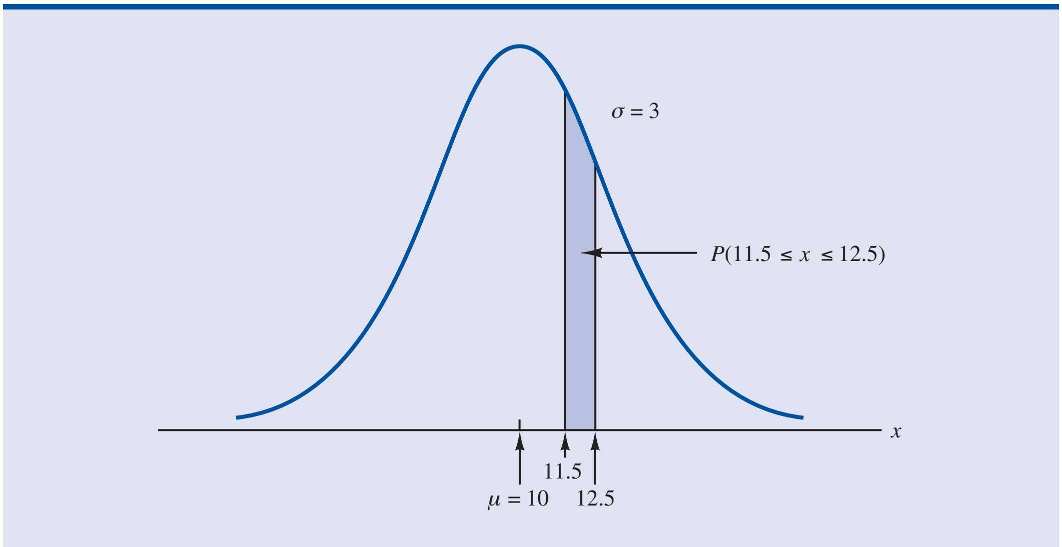
214	163	265	194	180
202	198	212	201	
174	171	211	211	

    La distribución de probabilidad de los volúmenes de negociaciones es aproximadamente normal.
    a. Calcule la media y la desviación estándar a usar como estimaciones de la media y de la desviación estándar de la población.
    b. ¿Cuál es la probabilidad de que, en un día elegido al azar, el volumen de negociaciones en la mañana temprano sea superior a 180 millones de acciones?
    c. ¿Cuál es la probabilidad de que, en un día elegido al azar, el volumen de negociaciones en la mañana temprano sea superior a 230 millones de acciones?
    d. ¿Cuántas acciones deberán ser negociadas para que el volumen de negociaciones en la mañana temprano de un día determinado pertenezca al 5% de los días de mayor movimiento?
25. De acuerdo con la Sleep Foundation, en promedio se duermen 6 .8 horas por noche. Suponga que la desviación estándar es 0.6 horas y que la distribución de probabilidad es normal.
    a. ¿Cuál es la probabilidad de que una persona seleccionada al azar duerma más de ocho horas?
    b. ¿De que una persona tomada aleatoriamente duerma seis horas o menos?
    c. Los médicos aconsejan dormir entre siete y nueve horas por noche. ¿Qué porcentaje de la población duerme esta cantidad?

6.3 Aproximación normal de las probabilidades binomiales

En la sección 5.4 se presentó la distribución binomial discreta. Recuerde que un experimento binomial consiste en una serie de n ensayos idénticos e independientes, habiendo para cada ensayo dos resultados posibles, éxito o fracaso. La probabilidad de éxito en un ensayo es la misma que en cualquier otro de los ensayos y se denota p. La variable aleatoria binomial es el número de éxitos en n ensayos y lo que se quiere saber es la probabilidad de x éxitos en n ensayos.

FIGURA 6.8 APROXIMACIÓN NORMAL A UNA PROBABILIDAD BINOMIAL DISTRIBUCIÓN EN LA QUE $n = 100$ Y $p = 0.10$ MOSTRANDO LA PROBABILIDAD DE 12 ERRORES



La evaluación de una función de probabilidad binomial, a mano o con una calculadora, se dificulta cuando el número de ensayos es muy grande. En los casos en que $np \geq 5$ y $n(1 - p) \geq 5$, la distribución normal proporciona una aproximación a las probabilidades binomiales que es fácil de usar. Cuando se usa la aproximación normal a la binomial, en la definición de la curva normal $\mu = np$ y $\sigma = \sqrt{np(1 - p)}$.

Para ilustrar la aproximación normal a la binomial, suponga que una empresa sabe por experiencia que 10% de sus facturas tienen algún error. Toma una muestra de 100 facturas y desea calcular la probabilidad de que 12 de estas facturas contengan algún error. Es decir, quiere hallar la probabilidad binomial de 12 éxitos en 100 ensayos. Aplicando la aproximación normal a este caso se tiene, $\mu = np = (100)(0.1) = 10$ y $\sigma = \sqrt{np(1 - p)} = \sqrt{(100)(0.1)(0.9)} = 3$. En la figura 6.8 se muestra la distribución normal con $\mu = 10$ y $\sigma = 3$.

Recuerde que en una distribución de probabilidad continua las probabilidades se calculan como áreas bajo la curva de la función de densidad de probabilidad. En consecuencia, la probabilidad que tiene un solo valor de la variable aleatoria es cero. Por tanto, para aproximar la probabilidad binomial de 12 éxitos se calcula el área correspondiente bajo la curva normal entre 11.5 y 12.5. Al 0.5 que se suma y se resta al 12 se le conoce como **factor de corrección por continuidad.** Este factor se introduce debido a que se está usando una distribución continua para aproximar una distribución discreta. Así, $P(x = 12)$ de una distribución binomial *discreta* se aproxima mediante $P(11.5 \leq x \leq 12.5)$ en la distribución normal *continua*.

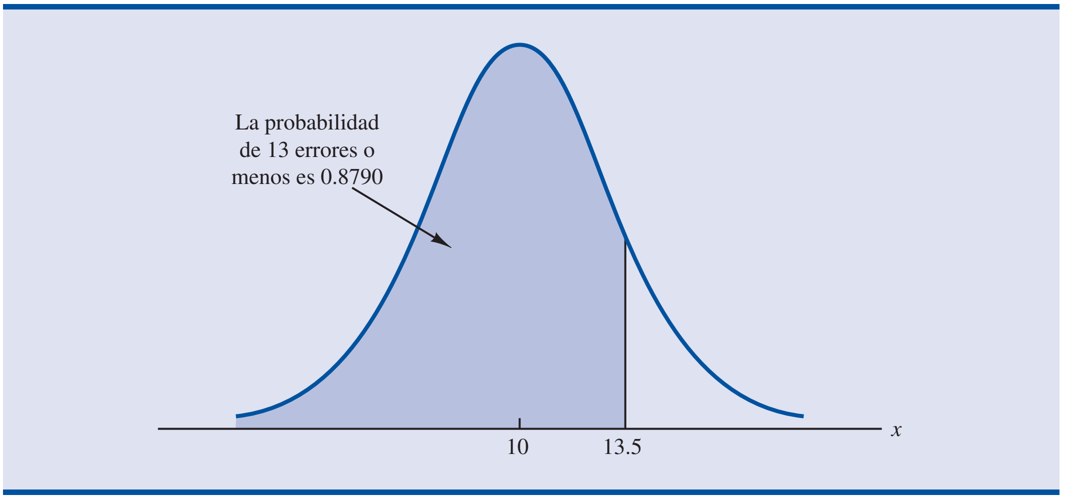
Convirtiendo la distribución normal estándar para calcular $P(11.5 \leq x \leq 12.5)$, se tiene

$$z = \frac{x - \mu}{\sigma} = \frac{12.5 - 10.0}{3} = 0.83 \quad \text{para } x = 12.5$$

y

$$z = \frac{x - \mu}{\sigma} = \frac{11.5 - 10.0}{3} = 0.50 \quad \text{para } x = 11.5$$

FIGURA 6.9 APROXIMACIÓN NORMAL A UNA PROBABILIDAD BINOMIAL: DISTRIBUCIÓN EN LA QUE $n = 100$ Y $p = 0.10$ MUESTRAN LA PROBABILIDAD DE 13 ERRORES O MENOS



En la tabla de la probabilidad normal estándar aparece que el área bajo la curva (figura 6.8) a la izquierda de 12.5 es 0.7967. De manera similar, el área bajo la curva a la izquierda de 11.5 es 0.6915. Por tanto, el área entre 11.5 y 12.5 es $0.7967 - 0.6915 = 0.1052$. El cálculo normal de la probabilidad de 12 éxitos en 100 ensayos es 0.1052.

Para tener un ejemplo más, suponga que se quiere calcular la probabilidad de 13 o menos facturas con errores en una muestra de 100 facturas. En la figura 6.9 se muestra el área bajo la curva que aproxima esta probabilidad. Observe que debido al uso del factor de continuidad el valor que se usa para calcular esta probabilidad es 13.5. El valor z que corresponde a 13.5 es

$$z = \frac{13.5 - 10.0}{3.0} = 1.17$$

En la tabla de probabilidad normal estándar se observa que el área bajo la curva normal estándar y a la izquierda de $z = 1.17$ es 0.8790. El área bajo la curva normal que aproxima la probabilidad de 13 o menos facturas con errores es la porción sombreada que se observa en la gráfica de la figura 6.9.

Ejercicios

Métodos

Auto examen

26. En una distribución de probabilidad binomial con $p = 0.20$ y $n = 100$.
    a. ¿Cuál es la media y la desviación estándar?
    b. ¿En esta situación las probabilidades binomiales pueden ser aproximadas por la distribución de probabilidad normal? Explique.
    c. ¿Cuál es la probabilidad de exactamente 24 éxitos?
    d. ¿Cuál es la probabilidad de 18 a 22 éxitos?
    e. ¿Cuál es la probabilidad de 15 o menos éxitos?
27. Suponga que se tiene una distribución de probabilidad binomial en la que $p = 0.60$ y $n = 200$.
    a. ¿Cuál es la media y la desviación estándar?
    b. ¿En esta situación las probabilidades binomiales puedan ser aproximadas por la distribución de probabilidad normal? Explique.
    c. ¿Cuál es la probabilidad de 100 a 110 éxitos?

d. ¿Cuál es la probabilidad de 130 o más éxitos?
e. ¿Cuál es la ventaja de usar la distribución de probabilidad normal para aproximar las probabilidades binomiales? Use el inciso d para explicar las ventajas.

Aplicaciones

Auto examen

28. El presidente Bush propuso eliminar los impuestos sobre los dividendos que pagan los accionistas debido a que esto resulta en un doble pago de impuestos. Las ganancias que se usan para pagar los dividendos ya han sido grabadas. En un sondeo sobre este tema se encontró que 47% de los estadounidenses estaban a favor de esta propuesta. La posición de los partidos políticos era 64% de los republicanos y 29% de los demócratas a favor de la propuesta (*Investor´s Business Daily*, 13 de enero de 2003). Suponga que 250 estadounidenses se reúnen para una conferencia acerca de la propuesta.
    a. ¿Cuál es la probabilidad de que por lo menos la mitad del grupo esté a favor de la propuesta?
    b. Más tarde se entera de que en el grupo hay 150 republicanos y 100 demócratas. Ahora, ¿cuál es su estimación del número esperado a favor de la propuesta?
    c. Ahora que conoce la composición del grupo, ¿cree que un conferencista a favor de la propuesta sea mejor recibido que uno que esté en contra de la propuesta?
29. La tasa de desempleo es de 5.8% (Bureau of Labor Statistics, www.bls.gov, 3 de abril de 2003). Suponga que se seleccionan aleatoriamente 100 personas que se pueden emplear.
    a. ¿Cuál es el número esperado de quienes están desempleados?
    b. ¿Cuál es la varianza y la desviación estándar del número de los que están desempleados?
    c. ¿Cuál es la probabilidad de que exactamente seis estén desempleados?
    d. ¿Cuál es la probabilidad de que por lo menos cuatro estén desempleados?
30. Cuando usted firma un contrato para una tarjeta de crédito, ¿lee cuidadosamente el contrato? En un sondeo FindLaw.com le preguntó a las personas "¿Qué tan cuidadosamente lee usted un contrato para una tarjeta de crédito?" Los hallazgos fueron que 44% leen cada palabra, 33% leen lo suficiente para entender el contrato, 11% sólo le echa una mirada y 4% no lo leen en absoluto.
    a. En una muestra de 500 personas ¿cuántas esperaría usted que respondan que leen cada palabra de un contrato para una tarjeta de crédito?
    b. En una muestra de 500 personas ¿cuál es la probabilidad de que 200 o menos digan que leen cada palabra de un contrato para una tarjeta de crédito?
    c. En una muestra de 500 personas ¿cuál es la probabilidad de que por lo menos 15 digan que no leen en absoluto un contrato para una tarjeta de crédito?
31. El Myrtle Beach hotel tiene 120 habitaciones. En los meses de primavera su ocupación es de 75%.
    a. ¿Cuál es la probabilidad de que por lo menos la mitad de las habitaciones estén ocupadas en un día dado?
    b. ¿De que 100 o más de las habitaciones estén ocupadas en un día dado?
    c. ¿De que 80 o menos de las habitaciones estén ocupadas en un día dado?

6.4 Distribución de probabilidad exponencial

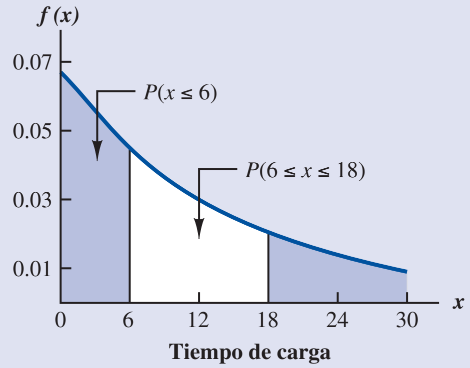
La **distribución de probabilidad exponencial** se aplica a variables como las llegadas de automóviles a un lavado de coches, los tiempos requeridos para cargar un camión, las distancias entre dos averías en una carretera, etc. A continuación se da la función de densidad de probabilidad exponencial.

FUNCIÓN DE DENSIDAD DE PROBABILIDAD EXPONENCIAL

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \qquad \text{para } x \geq 0, \mu > 0 \tag{6.4}$$

donde μ = valor esperado o media

FIGURA 6.10 DISTRIBUCIÓN EXPONENCIAL PARA EL EJEMPLO DEL ÁREA DE CARGA



Como ejemplo de la distribución exponencial, suponga que x representa el tiempo que se necesita para cargar un camión en un área de carga, y que este tiempo de carga sigue una distribución exponencial. Si el tiempo de carga medio o promedio es 15 minutos ($\mu = 15$), la función de densidad de probabilidad apropiada para x es

$$f(x) = \frac{1}{15} e^{-x/15}$$

La figura 6.10 es la gráfica de esta función de densidad de probabilidad.

Cálculo de probabilidades en la distribución exponencial

En aplicaciones de colas de espera, la distribución exponencial suele emplearse para tiempos de servicio.

Como ocurre con cualquier distribución de probabilidad continua, el área bajo la curva correspondiendo a un intervalo da la probabilidad de que la variable aleatoria tome algún valor en ese intervalo. En el ejemplo del área de carga, la probabilidad de que cargar un camión necesite 6 minutos o menos $P\ (x \leq 6)$ está definida como el área bajo la curva de la figura 10.6 que va desde $x = 0$ hasta $x = 6$. De manera similar, la probabilidad de que el tiempo de carga sean 18 minutos o menos $P\ (x \leq 18)$ es el área bajo la curva desde $x = 0$ hasta $x = 18$. Observe también que la probabilidad de que el tiempo de carga esté entre 6 y 18 minutos $P\ (6 \leq x \leq 18)$ corresponde al área bajo la curva desde $x = 6$ hasta $x = 18$.

Para calcular probabilidades exponenciales como las que se acaban de describir, se usa la fórmula siguiente. Esta fórmula aporta la probabilidad acumulada de obtener un valor de la variable aleatoria exponencial que sea menor o igual que algún valor específico denotado por x_0.

DISTRIBUCIÓN EXPONENCIAL: PROBABILIDADES ACUMULADAS

$$P(x \leq x_0) = 1 - e^{-x_0/\mu} \quad \textbf{(6.5)}$$

En el ejemplo del área de carga, x = tiempo de carga en minutos y $\mu = 15$ minutos. A partir de la ecuación (6.5)

$$P(x \leq x_0) = 1 - e^{-x_0/15}$$

Por tanto, la probabilidad de que cargar un camión requiera 6 minutos o menos es

$$P(x \leq 6) = 1 - e^{-6/15} = 0.3297$$

Con la ecuación (6.5) se calcula la probabilidad de que cargar un camión requiera 18 minutos o menos.

$$P(x \leq 18) = 1 - e^{-18/15} = 0.6988$$

De manera que la probabilidad de que para cargar un camión se necesiten entre 6 y 18 minutos es igual a $0.6988 - 0.3297 = 0.3691$. Probabilidades para cualquier otro intervalo se calculan de manera semejante.

La distribución exponencial tiene la propiedad de que la media y la desviación estándar son iguales.

En el ejemplo anterior el tiempo medio para cargar un camión fue $\mu = 15$ minutos. La distribución exponencial tiene la propiedad de que la media de la distribución y la desviación estándar de la distribución son iguales. Por tanto, la desviación estándar del tiempo que se necesita para cargar un camión es $\sigma = 15$ minutos y la varianza es $\sigma^2 = (15)^2 = 225$.

Relación entre la distribución de Poisson y la exponencial

En la sección 5.5 se presentó la distribución de probabilidad de Poisson como una distribución de probabilidad discreta que se usa para examinar el número de ocurrencias de un evento en un determinado intervalo de tiempo o de espacio. Recuerde que la función de probabilidad de Poisson es

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

donde

$$\mu = \text{valor esperado o número medio de ocurrencias en un determinado intervalo}$$

Si las llegadas siguen una distribución de Poisson, el tiempo entre las llegadas debe seguir una distribución exponencial.

La distribución de probabilidad exponencial continua está relacionada con la distribución discreta de Poisson. Si la distribución de Poisson da una descripción del número de ocurrencias por intervalo, la distribución exponencial aporta una descripción de la longitud de los intervalos entre las ocurrencias.

Para ilustrar esta relación, suponga que el número de automóviles que llegan a un lavado de coches durante una hora se describe mediante la distribución de probabilidad de Poisson, con una media de 10 automóviles por hora. La función de probabilidad de Poisson que da la probabilidad de x llegadas por hora es

$$f(x) = \frac{10^x e^{-10}}{x!}$$

Dado que el número promedio de llegadas es 10 automóviles por hora, el tiempo promedio entre las llegadas de los automóviles es

$$\frac{1 \text{ hora}}{10 \text{ automóviles}} = 0.1 \text{ hora/automóvil}$$

Entonces, la distribución exponencial que describe el tiempo entre las llegadas tiene una media de $\mu = 0.1$ por automóvil; la función de densidad de probabilidad exponencial es

$$f(x) = \frac{1}{0.1} e^{-x/0.1} = 10e^{-10x}$$

NOTAS Y COMENTARIOS

Como se observa en la figura 6.10, la distribución exponencial es sesgada a la derecha. En efecto, la medida del sesgo en la distribución exponencial es 2. La distribución exponencial da una idea clara de cómo es una distribución sesgada.

Ejercicios

Métodos

32. Considere la siguiente función de densidad de probabilidad exponencial.

$$f(x) = \frac{1}{8} e^{-x/8} \quad \text{para } x \geq 0$$

a. Halle $P(x \leq 6)$.
b. Encuentre $P(x \leq 4)$.
c. Halle $P(x \geq 6)$.
d. Encuentre $P(4 \leq x \leq 6)$.

Auto examen

33. Considere la siguiente función de densidad de probabilidad exponencial.

$$f(x) = \frac{1}{3} e^{-x/3} \quad \text{para } x \geq 0$$

a. Dé la fórmula para hallar $P(x \leq x_0)$..
b. Halle $P(x \leq 2)$.
c. Encuentre $P(x \geq 3)$.
d. Halle $P(x \leq 5)$.
e. Halle $P(2 \leq x \leq 5)$.

Aplicaciones

34. El tiempo requerido para pasar por la inspección en los aeropuertos puede ser molesto para los pasajeros. El tiempo medio de espera en los periodos pico en el Cincinnnati/Northern Kentucky Internacional Airport es de 12.1 minutos (*The Cincinnati Enquirer*, 2 de febrero de 2006). Suponga que los tiempos para pasar por la inspección de seguridad tienen una distribución exponencial.
   a. ¿Cuál es la probabilidad de que durante los periodos pico se requieran 10 minutos para pasar la inspección de seguridad?
   b. ¿De que durante los periodos pico se requieran más de 20 minutos para pasar la inspección de seguridad?
   c. ¿De que durante los periodos pico se requieran entre 10 y 20 minutos para pasar la inspección de seguridad?
   d. Son las 8 de la mañana (periodo pico) y usted se acaba de formar en la fila para la inspección de seguridad. Para alcanzar su avión, tiene que estar en la puerta de arribo en no más de 30 minutos. Si necesitara 12 minutos una vez pasada la inspección de seguridad para llegar a la puerta de arribo, ¿cuál es la probabilidad de que pierda el avión?

Auto examen

35. Los tiempos entre las llegadas de vehículos a un determinado entronque siguen una distribución de probabilidad exponencial cuya media es 12 segundos.
   a. Dibuje esta distribución de probabilidad exponencial.
   b. ¿Cuál es la probabilidad de que los tiempos de llegada entre vehículos sean 12 segundos o menos?

c. ¿Cuál es la probabilidad de que los tiempos de llegada entre vehículos sean 6 segundos o menos?
d. ¿Cuál es la probabilidad de 30 o más segundos entre los tiempos de llegada?

36. El tiempo de vida (en hora) de un dispositivo electrónico es una variable aleatoria que tiene la siguiente función de densidad de probabilidad exponencial.

$$f(x) = \frac{1}{50} e^{-x/50} \qquad \text{para } x \geq 0$$

a. ¿Cuál es la media del tiempo de vida de este dispositivo?
b. ¿Cuál es la probabilidad de que el dispositivo tenga una falla en las primeras 25 horas de funcionamiento?
c. ¿Cuál es la probabilidad de que el dispositivo funcione 100 o más horas sin fallar?

37. Sparagowsky & Associates hace un estudio sobre los tiempos necesarios para atender a un cliente en la ventanilla de su automóvil en los restaurantes de comida rápida. En McDonald´s el tiempo medio para atender a un cliente fue 2.78 minutos (*The Cincinnati Enquirer*, 9 de julio de 2000). Tiempos de servicio como los de estos restaurantes suelen seguir una distribución exponencial.
a. ¿Cuál es la probabilidad de que el tiempo para atender a un cliente sea menor que 2 minutos?
b. ¿De que el tiempo para atender a un cliente sean más de 5 minutos?
c. ¿De que el tiempo para atender a un cliente sean más de 2.78 minutos?

38. ¿Las interrupciones durante su trabajo reducen su productividad? De acuerdo con un estudio realizado por la University of California–Irvine, las personas de negocios son interrumpidas aproximadamente 51/2 veces por hora (*Fortune*, 20 de marzo de 2006). Suponga que el número de interrupciones sigue una distribución de probabilidad de Poisson.
a. Dé la distribución de probabilidad para el tiempo entre las interrupciones.
b. ¿Cuál es la probabilidad de que una persona de negocios no tenga ninguna interrupción en 15 minutos?
c. ¿Cuál es la probabilidad de que la siguiente interrupción a una determinada persona de negocios ocurra en no más de 10 minutos?

Resumen

En este capítulo se amplía el estudio de las distribuciones de probabilidad al caso de las variables aleatorias continuas. La principal diferencia conceptual entre distribuciones de probabilidades discretas y continuas está en el método para calcular las probabilidades. En el caso de distribuciones discretas la función de probabilidad $f(x)$ da la probabilidad de que la variable aleatoria x tome diversos valores. En el caso de las distribuciones continuas, la función de densidad de probabilidad $f(x)$ no da directamente valores de probabilidad. Aquí, las probabilidades están dadas por áreas bajo la curva o gráfica de la función de densidad de probabilidad $f(x)$. Como el área bajo la curva sobre un solo punto es 0, se observa que en una variable aleatoria continua la probabilidad de cualquier valor particular es 0.

Se vieron a detalle tres distribuciones de probabilidad continua: la uniforme, la normal y la exponencial. La distribución normal es muy empleada en la inferencia estadística y será muy empleada en lo que resta del libro.

Glosario

Función de densidad de probabilidad Función que se usa para calcular probabilidades de una variable aleatoria continua. El área bajo la gráfica de una función de densidad de probabilidad y sobre un intervalo representa probabilidad.

Distribución de probabilidad uniforme Distribución de probabilidad continua en la cual la probabilidad de que una variable aleatoria tome un valor en cualquier intervalo es igual para intervalos de igual longitud.

Distribución de probabilidad normal Una distribución de probabilidad continua. Su función de densidad de probabilidad tiene forma de campana y está determinada por la media μ y la desviación estándar σ.
Distribución de probabilidad normal estándar Distribución normal en la cual la media es cero y la desviación estándar es uno.
Factor de corrección por continuidad Valor de 0.5 que se suma o resta al valor de x cuando se usa una distribución normal continua para aproximar una distribución binomial discreta.
Distribución de probabilidad exponencial Una distribución de probabilidad continua útil para calcular probabilidades acerca del tiempo que se necesita para realizar una tarea.

Fórmulas clave

Función de densidad de probabilidad uniforme

$$f(x) = \begin{cases} \dfrac{1}{b - a} & \text{para } a \leq x \leq b \\ 0 & \text{si no es así} \end{cases} \tag{6.1}$$

Función de densidad de probabilidad normal

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{6.2}$$

Conversión a la variable aleatoria normal estándar

$$z = \frac{x - \mu}{\sigma} \tag{6.3}$$

Función de densidad de probabilidad exponencial

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \qquad \text{para } x \geq 0, \mu > 0 \tag{6.4}$$

Distribución exponencial: probabilidades acumuladas

$$P(x \leq x_0) = 1 - e^{-x_0/\mu} \tag{6.5}$$

Ejercicios complementarios

39. Una ejecutiva de negocios se va a mudar de Chicago a Atlanta y necesita vender rápidamente su casa en Chicago. Un empleado le ofrece comprársela en \$210 000, pero la oferta expira al final de esa semana. En este momento la ejecutiva no tiene otra oferta mejor, pero piensa que puede dejar la casa en el mercado un mes más. De acuerdo con las pláticas que ha tenido con su agente inmobiliario la ejecutiva cree que los precios que pueden ofrecerle dejando la casa un mes más en el mercado están distribuidos uniformemente entre \$200 000 y \$225 000.
    a. Si deja la casa en el mercado un mes más, ¿cuál es la expresión matemática para la función de densidad de probabilidad de los precios de venta que le sean ofrecidos?
    b. Si la deja en el mercado un mes más, ¿cuál es la probabilidad de que venda la casa en por lo menos \$215 000?
    c. Si la deja en el mercado un mes más, ¿cuál es la probabilidad de que venda la casa en menos de \$210 000?
    d. ¿Deberá dejar la ejecutiva su casa en el mercado un mes más? ¿Por qué sí o por qué no?

40. La U.S. Bureau of Labor Statistics informa que el gasto promedio anual en alimentos y bebidas de una familia es \$5700 (*Money*, diciembre de 2003). Suponga que los gastos anuales en alimentos y bebidas están distribuidos en forma normal y que la desviación estándar es \$1500.
    a. ¿Cuál es el intervalo en que se encuentran los gastos de 10% de las familias que tienen los menores gastos anuales en alimentos y bebidas?
    b. ¿Qué porcentaje de las familias gasta más de \$7000 anualmente en alimentos y bebidas?
    c. ¿Cuál es el intervalo en el que se encuentran los gastos de 5% de las familias que tienen los gastos más altos en alimentos y bebidas?
41. Motorola usa la distribución normal para determinar la probabilidad de defectos y el número de defectos esperados en un proceso de producción. Suponga que en un proceso de producción el peso promedio de los artículos producidos es 10 onzas. Calcule la probabilidad de un defecto y el número esperado de defectos en 100 unidades producidas en las situaciones siguientes.
    a. La desviación estándar del proceso es 0.15 y los límites para el proceso se han fijado en más o menos una desviación estándar. Las unidades que pesen más de 9.85 o menos de 10.15 onzas se clasifican como defectuosas.
    b. Después de hacer mejoras al proceso de producción, la desviación estándar se reduce a 0.05. Asuma que se siguen usando los mismos límites para el proceso; artículos que pesen menos de 9.85 o más de 10.15 onzas se clasifican como defectuosos.
    c. ¿Cuál es la ventaja de haber reducido la variación en el proceso de producción, haciendo que los límites se encuentren a un número mayor de desviaciones estándar de la media?
42. El promedio anual de gastos de una familia estadounidense en transporte diario es \$6312 (*Money*, agosto de 2001). Suponga que dicha cantidad está distribuida normalmente.
    a. Si sabe que 5% de las familias estadounidenses gastan menos de \$1000 en el transporte diario. ¿Cuál es la desviación estándar en esta cantidad de gasto?
    b. ¿Cuál es la probabilidad de que un hogar gaste entre \$4000 y \$6000?
    c. ¿En que intervalo se encuentran los gastos de las familias que constituyen 3% de las familias con los gastos más altos en transporte?
43. *Condé Nast Traveler* publica la lista de oro de los mejores hoteles en todo el mundo. Broadmoor Hotel en Colorado Springs tiene 700 habitaciones y estuvo en la lista de oro en 2004 (*Condé Nast Traveler*, enero de 2004). El grupo encargado del marketing de este hotel pronostica una demanda media de 670 habitaciones para el próximo fin de semana. Suponga que la demanda para el próximo fin de semana está distribuida normalmente y que la desviación estándar es 30.
    a. ¿Cuál la probabilidad de que se ocupen todas las habitaciones del hotel?
    b. ¿Cuál la probabilidad de que se ocupen 50 o más habitaciones del hotel?
    c. ¿Recomendaría al hotel hacer una promoción para aumentar la demanda? ¿Qué consideraciones serían importantes?
44. Ward Doering Auto Sales está pensando en ofrecer un contrato especial de servicio que cubra todos los costos de servicio de los automóviles vendidos. De acuerdo con la experiencia, el director de la empresa estima que los costos anuales de servicio están distribuidos casi normalmente con una media de \$150 y una desviación estándar de \$25.
    a. Si la empresa ofrece a los clientes el contrato de servicio por una cantidad anual de \$200, ¿cuál es la probabilidad de que el costo de un servicio sea mayor a los \$200 del precio del contrato?
    b. ¿Cuál es la ganancia esperada por la empresa en estos contratos de servicio?
45. ¿La falta de sueño es causa de accidentes de tráfico de consecuencias fatales? En un estudio se encontró que el número promedio por año de accidentes de tráfico con consecuencias fatales ocasionados por conductores somnolientos es 1550 (*BusinessWeek*, 26 de enero de 2004). Suponga que el número promedio anual de accidentes de tráfico de consecuencias fatales está distribuido normalmente con una desviación estándar de 300.
    a. ¿Cuál es la probabilidad de que haya menos de 1000 accidentes fatales en un año?
    b. ¿De que el número anual de accidentes fatales esté entre 1000 y 2000?
    c. Para que un año se encuentre en el 5% superior en número de accidentes fatales, cuántos de éstos tendrán que ocurrir?

46. Suponga que las puntuaciones obtenidas en el examen de admisión a una universidad están distribuidas en forma normal con una media de 450 y una desviación estándar de 100.
    a. ¿Qué porcentaje de las personas que hacen el examen tendrá una puntuación entre 400 y 500?
    b. Si la puntuación que obtiene un estudiante es 630. ¿Qué porcentaje de los estudiantes que hacen el examen tendrá una puntuación mayor? ¿Qué porcentaje tendrá una puntuación menor?
    c. Si la universidad no admite estudiantes que obtengan una puntuación menor a 480, ¿qué porcentaje de los estudiantes que hacen el examen podrá ser aceptado?
47. De acuerdo con *Adversiting Age*, el salario base promedio de las mujeres que trabajan como publicistas es superior al salario base promedio de los hombres. El salario base promedio de las mujeres es $67 000 y el salario base promedio de los hombres es $65 500 (*Working Woman*, julio/agosto de 2000). Suponga que los salarios están distribuidos normalmente con una desviación estándar de $7000 tanto para hombres como para mujeres.
    a. ¿Cuál es la probabilidad de que una mujer tenga un salario mayor que $75 000?
    b. ¿De que un hombre tenga un salario mayor que $75 000?
    c. ¿De que una mujer tenga un salario mayor que $50 000?
    d. ¿Cuánto tendrá que ganar una mujer para tener un salario mayor que 99% de los hombres?
48. Una máquina llena recipientes con un determinado producto. De acuerdo con datos anteriores se sabe que la desviación estándar en los pesos rellenados es 0.6 onzas. Si sólo 2% de los recipientes llenados tienen menos de 18 onzas, ¿cuál es el peso medio de llenado de la máquina? Es decir, a cuánto es igual μ? Suponga que los pesos llenados tienen una distribución normal.
49. Considere un examen de opción múltiple con 50 preguntas. Para cada pregunta hay cuatro respuestas posibles. Suponga que un estudiante que ha hecho las tareas y asistido a clase tiene 75% de probabilidad de contestar correctamente las preguntas.
    a. Para obtener A de calificación, un estudiante tiene que contestar correctamente 43 o más preguntas. ¿Qué porcentaje de los estudiantes que hicieron las tareas y asistieron a clase obtendrá A de calificación?
    b. Para obtener C de calificación, un estudiante tiene que contestar correctamente de 35 a 39 preguntas. ¿Qué porcentaje de los estudiantes que hicieron las tareas y asistieron a clases obtendrá C de calificación?
    c. Para aprobar el examen, un estudiante tiene que contestar correctamente 30 preguntas o más. ¿Qué porcentaje de los estudiantes que hicieron las tareas y asistieron a clases pasará el examen?
    d. Suponga que un estudiante no asistió a clases ni hizo las tareas. Además, dicho estudiante sólo tratará de adivinar las respuestas a las preguntas. ¿Cuál es la probabilidad de que el estudiante conteste correctamente 30 o más preguntas y pase el examen?
50. En Las Vegas un jugador de blackjack se entera de que la casa proporcionará una habitación gratis a quien juegue cuatro horas con un promedio de apuesta de $50. La estrategia del jugador tiene una probabilidad de ganar en cualquier mano de 0.49 y el jugador sabe que hay 60 manos por hora. Suponga que el jugador juega durante cuatro horas con una apuesta de $50 por mano.
    a. ¿Cuál es la ganancia esperada del jugador?
    b. ¿Cuál es la probabilidad de que el jugador pierda $1000 o más?
    c. ¿Cuál es la probabilidad de que el jugador gane?
    d. Si el jugador empieza con $1500. ¿Cuál es la probabilidad de que el jugador se vaya a la bancarrota?
51. El tiempo, en minutos, que un estudiante usa una terminal de computadora en el centro de cálculo de una universidad sigue una distribución de probabilidad exponencial con una media de 36 minutos. Suponga que un estudiante llega a una terminal precisamente en el momento en que otro estudiante quería usar la terminal.
    a. ¿Cuál es la probabilidad de que el segundo estudiante tenga que esperar 15 minutos o menos?
    b. ¿De que el segundo estudiante tenga que esperar entre 15 y 45 minutos?
    c. ¿Cuál la probabilidad de que el segundo estudiante tenga que esperar una hora o más?
52. El sitio Web de Bed and Breakfast Inns of North America (www.cimarron.net) recibe aproximadamente siete visitas por minuto (*Time*, septiembre de 2001). Suponga que el número de visitantes por minuto sigue una distribución de probabilidad de Poisson.

a. ¿Cuál es el tiempo medio entre las visitas a este sitio de la Web?
b. Muestre la función de densidad de probabilidad exponencial para los tiempos entre las visitas a este sitio.
c. ¿Cuál es la probabilidad de que nadie visite este sitio en un lapso de 1 minuto?
d. ¿Cuál es la probabilidad de que nadie visite este sitio en un lapso de 12 minutos?

53. En la ciudad de Nueva York el tiempo de recorrido promedio al trabajo es de 36.5 minutos.
 a. Suponga que la distribución de probabilidad exponencial es aplicable y muestre la función de densidad de probabilidad para las duraciones de los recorridos al trabajo en Nueva York.
 b. ¿Cuál es la probabilidad de que un neoyorquino típico necesite entre 20 y 40 minutos para transportarse a su trabajo?
 c. ¿De que un neoyorquino típico necesite más de 40 minutos para transportarse a su trabajo?

54. El tiempo (en minutos) entre dos llamadas telefónicas en la oficina de solicitud de servicios de una aseguradora tiene la siguiente distribución de probabilidad exponencial.

$$f(x) = 0.50e^{-0.50x} \quad \text{para } x \geq 0$$

 a. ¿Cuál es el tiempo medio entre las llamadas telefónicas?
 b. ¿Cuál es la probabilidad de que pasen 30 segundos o menos entre llamadas telefónicas?
 c. ¿De que pase 1 minuto o menos entre las llamadas telefónicas?
 d. ¿Cuál es la probabilidad de que pasen 5 minutos o más sin que haya llamadas telefónicas?

Caso problema Specialty Toys

Specialty Toys, Inc. vende una gran variedad de nuevos e innovadores juguetes para niños. Los directivos saben que la época prenavideña es la mejor oportunidad para la introducción de un nuevo juguete, en esta época muchas personas buscan cosas novedosas para los regalos navideños. Cuando Specialty descubre un nuevo juguete con un buen potencial de mercado, elige alguna fecha en octubre para su lanzamiento.

Para contar con los juguetes en octubre Specialty hace los pedidos a sus proveedores en junio o julio de cada año. La demanda de juguetes para niños puede ser muy volátil. Si un nuevo juguete se pone de moda, la posibilidad de que se agote suele incrementar la demanda hasta niveles altos y se pueden obtener grandes ganancias. Sin embargo, un nuevo juguete también puede fracasar dejando a Specialty con un gran inventario que debe vender a precios reducidos. La interrogante más importante que enfrenta la empresa es decidir cuántas unidades comprar de un juguete nuevo para satisfacer la demanda. Si compra muy pocos, perderá ventas; si compra demasiados, las ganancias se reducirán por los precios bajos que tendrá que ofrecer en una liquidación.

En la próxima temporada Specialty desea introducir un juguete nuevo que se llama *El osito pronosticador del clima*. Esta variación de un osito de peluche que habla es fabricada por una empresa en Taiwan. Cuando un niño oprime la mano del osito, éste empieza a hablar. El osito tiene un barómetro que le ayuda, de acuerdo con el estado del tiempo, a elegir una de cinco frases que pronostican el estado del tiempo. Las frases van desde "¡Parece que es un bonito día! Que te diviertas" hasta "Parece que va a llover. No se te olvide llevar tu paraguas". Pruebas realizadas con el producto indican que, aunque no es preciso, sus pronósticos del tiempo son sorprendentemente buenos. Varios de los directivos de Specialty opinan que los pronósticos del tiempo del osito son tan buenos como muchos de los pronósticos del tiempo que se dan en televisión.

Como ocurre con todos los productos, Specialty se enfrenta a la pregunta de cuántos ositos ordenar para la temporada siguiente. Las cantidades que sugieren los directivos son 15 000, 18 000, 24 000 o 28 000 unidades. El intervalo tan amplio en que se encuentran estas cantidades indica una considerable discrepancia en lo que se refiere al potencial de mercado. El equipo de directivos le solicita a usted un análisis de las probabilidades de terminar el inventario de acuerdo con diversas cantidades a comprar, así como una estimación del potencial de ganancias y su ayuda para hacer una recomendación de la cantidad que se debe comprar. Specialty espera vender *El osito pronosticador del clima* a $24 con base en un costo de $16 por unidad. Si hay inventario sobrante después de la temporada de las fiestas decembrinas, Specialty venderá las unidades restantes a $5 cada una. Después de revisar las ventas anteriores de productos semejantes, los expertos de Specialty pronostican una demanda esperada de 20 000 unidades y 0.95 de probabilidad de que la demanda esté entre 10 000 y 30 000.

Informe administrativo

Elabore un informe sobre los puntos siguientes y recomiende la cantidad a comprar de *El osito pronosticador del clima*.

1. Use los pronósticos de ventas para describir una distribución de probabilidad normal que pueda servir para aproximar la distribución de la demanda. Dibuje la distribución y dé su media y su desviación estándar.
2. Calcule la probabilidad de terminar el inventario de acuerdo con las cantidades a comprar sugeridas por los miembros del equipo de directivos.
3. Calcule las ganancias proyectadas de acuerdo con las cantidades a comprar sugeridas por los miembros del equipo de directivos bajo tres escenarios: el peor de los casos, en el cual se venderán 10 000 unidades, en el caso más probable, en el cual se venderán 20 000 unidades y en el mejor de los casos en el cual se venderán 30 000 unidades.
4. Uno de los directivos de Specialty encuentra que el potencial de ganancia es tan bueno que la cantidad a comprar debe tener 70% de posibilidades de satisfacer la demanda y 30% de posibilidades de quedarse sin mercancía. De acuerdo con esto, ¿qué cantidad debe comprarse y cuál es la ganancia proyectada bajo cada uno de los tres escenarios?
5. Dé su propia recomendación sobre la cantidad que debe comprarse y muestre las proyecciones de ganancia correspondientes. Fundamente su recomendación.

Apéndice 6.1 Distribuciones de probabilidad continua con Minitab

Para demostrar el procedimiento de Minitab para el cálculo de probabilidades continuas se retomará el problema de la empresa Grear Tire, en el que la duración de los neumáticos en millas se describió mediante una distribución normal en la que $\mu = 36\,500$ y $\sigma = 5000$. Una de las preguntas que se plantearon fue: ¿cuál es la probabilidad de que los neumáticos duren más de 40 000 millas?

Para distribuciones de probabilidad continua, Minitab proporciona probabilidades acumuladas; es decir, Minitab da la probabilidad de que la variable aleatoria tome un valor menor o igual que una constante específica. En el caso de la pregunta sobre la duración de los neumáticos, Minitab se puede usar para determinar la probabilidad acumulada de que un neumático dure 40 000 millas o menos. (En este caso la constante específica es 40 000.) Una vez que se tiene la probabilidad acumulada que proporciona Minitab, es necesario restar esta probabilidad de 1 para determinar la probabilidad de que el neumático dure más de 40 000 millas.

Para que Minitab calcule una probabilidad, es necesario ingresar la constante específica en una de las columnas de la hoja de cálculo. En este caso se introduce la constante específica 40 000 en la columna C1 de la hoja de cálculo de Minitab. A continuación se presentan los pasos necesarios para que Minitab calcule la probabilidad acumulada de que la variable aleatoria normal tome valores menores o iguales que 40 000.

Paso 1. Seleccionar el menú **Calc**
Paso 2. Elegir **Probability Distributions**
Paso 3. Elegir **Normal**
Paso 4. Cuando aparezca el cuadro de diálogo Normal Distribution:
- Seleccionar **Cumulative probability**
- Ingresar 36 500 en el cuadro **Mean**
- Ingresar 5 000 en el cuadro **Standard deviation**
- Ingresar C1 en el cuadro **Input column** (la columna que contiene 40 000)
- Clic en **OK**

Después de que el usuario hace clic en **OK**, Minitab da la probabilidad acumulada de que la variable aleatoria normal tome un valor menor o igual que 40 000. Minitab indica que esta probabilidad es 0.7580. Como lo que interesa es la probabilidad de que el neumático dure más de 40 000, la probabilidad buscada es $1 - 0.7580 = 0.2420$.

Otra pregunta en el problema de la empresa Grear Tire fue: ¿cuál es la duración en millas que la empresa debe establecer en la garantía de manera que en no más de 10% de los neumáticos se tenga que pagar la garantía? En este caso se da una probabilidad y se quiere hallar el valor correspondiente de la variable aleatoria. Minitab usa una rutina de cálculo inverso para hallar el valor de la variable aleatoria que corresponde a la probabilidad acumulada dada. Primero, se ingresa la probabilidad acumulada en la hoja de cálculo de Minitab (por ejemplo en C1). En este caso la probabilidad acumulada es 0.10. Después, los tres primeros pasos del procedimiento de Minitab son los dados antes. En el paso 4 se selecciona **Inverse cumulative probability** en lugar de **Cumulative probability** y se realiza la parte restante de este paso. Minitab da entonces 30 092 millas para la duración en la garantía.

Minitab también calcula las probabilidades de otras distribuciones de probabilidad continua, entre las que se encuentra la distribución de probabilidad exponencial. Para calcular probabilidades exponenciales se sigue el procedimiento antes dado para la distribución de probabilidad normal eligiendo la opción **Exponential** en el paso 3. El paso 4 es igual, salvo que no es necesario ingresar la desviación estándar. Minitab da los resultados de probabilidades acumuladas o probabilidades acumuladas inversas en la misma forma que se describió para la distribución de probabilidad normal.

Apéndice 6.2 Distribuciones de probabilidad continua con Excel

Excel permite calcular las probabilidades de varias distribuciones de probabilidad continuas. Entre las que se encuentran las distribuciones de probabilidad normal y exponencial. En este apéndice, se describe cómo usar Excel para calcular probabilidades en cualquier distribución normal. El procedimiento para la exponencial y para las otras distribuciones continuas es semejante al descrito aquí para la distribución normal.

Recuerde el problema de la empresa Grear Tire, la duración de los neumáticos en millas se describe mediante una distribución normal con media $\mu = 36\,500$ y $\sigma = 5000$. Suponga que se desea conocer la probabilidad de que un neumático dure más de 40 000 millas.

La función de Excel DISTR.NORM. suministra probabilidades acumuladas de una distribución normal. La forma general de la función es DISTR.NORM (x,media,desv_estándar,acum). En el cuarto argumento se especifica VERDADERO si se desea una probabilidad acumulada. De esta manera, para calcular la probabilidad acumulada de que la duración de un neumático sea menor o igual que 40 000 millas se ingresará la fórmula siguiente en cualquier celda de la hoja de cálculo Excel:

=DISTR.NORM(40000,36500,5000,VERDADERO)

En este momento, en la celda en que se ingresó la fórmula aparecerá 0.7580, indicando que la probabilidad de que la duración del neumático sea 40 000 millas es 0.7580. Por tanto, la probabilidad de que un neumático dure más de 40 000 millas es $1 - 0.7580 = 0.2420$.

La función de Excel DISTR.NORM.INV. usa un cálculo inverso para hallar el valor de x que corresponde a una probabilidad acumulada dada. Por ejemplo, si se desea hallar la duración que Grear debe ofrecer en su garantía de manera que no más de 10% de los neumáticos sean aptos para solicitar la garantía. Se ingresará en cualquier celda de la hoja de cálculo de Excel la fórmula siguiente:

=DISTR.NORM.INV(.1,36500,5000)

En este momento, en la celda en la que se ingresó la fórmula aparecerá 30092, indicando que la probabilidad de que un neumático dure 30 092 millas es 0.10.

La función de Excel para calcular probabilidades exponenciales es DISTR.EXP. Usar esta función es muy sencillo. Pero si se necesita ayuda para especificar los argumentos adecuados, se puede usar la herramienta Insertar Función de Excel (véase apéndice E).