Se puede usar MINITAB para determinar la ecuación de regresión para los datos de la marca mundial. La pantalla 4.3 muestra la información de las órdenes y las respuestas.

Pantalla 4.3

```
MTB > READ C1 C2
DATA > 1 4.023
DATA > 10 3.990
DATA > 10 3.967
DATA > 13 3.953
DATA > 14 3.908
DATA > 18 3.907
DATA > 20 3.902
DATA > 21 3.893
DATA > 22 3.855
DATA > 23 3.852
DATA > 31 3.833
DATA > 31 3.823
DATA > 37 3.788
DATA > 41 3.772
DATA > END
    14 ROWS READ
MTB > NAME C1 'CODYEAR'
MTB > NAME C2 'CODETIME'
MTB > REGRESSION C2 1 C1
THE REGRESSION EQUATION IS CODETIME = 4.03
    - 0.00655 CODEYEAR
```

La relación entre r y m

Tanto el coeficiente de correlación r como la pendiente de regresión m utilizan las cantidades SS_{xy} y SS_x; en consecuencia, es posible obtener uno en términos del otro. Usando un poco de álgebra elemental, se puede demostrar la validez de la relación siguiente.

Relación entre r y m

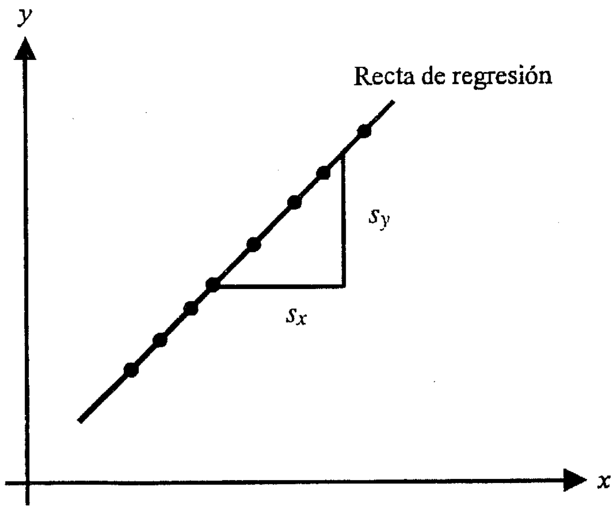
$$r = m\,\frac{s_x}{s_y} \tag{4.11}$$

donde s_x es la desviación estándar muestral de x y s_y es la desviación estándar muestral de y. Como s_x y s_y son mayores que cero, el coeficiente de correlación r concuerda en signo con la pendiente de la recta de regresión; por lo tanto, la fórmula (4.11) ofrece otra explicación de por qué la correlación es positiva si los puntos del diagrama de dispersión se acumulan de abajo a la izquierda a arriba a la derecha, y negativa si los puntos se acumulan de arriba a la izquierda y de abajo a la derecha.

Si despejamos m de la fórmula (4.11), tenemos:

$$m = \frac{s_y}{s_x}\, r$$

Si $r = 1$, entonces $m = \frac{s_y}{s_x}$ como se ilustra en el diagrama.



APLICACIÓN 4.8

Se puede usar un programa computacional para realizar un análisis de regresión que proporcione valores para b, m y r. Considere el ritmo cardiaco máximo y las edades que se registraron para diez individuos en un programa intensivo de ejercicios. Los datos son:

Edad	10	20	20	25	30	30	30	40	45	50
Ritmo cardiaco	210	200	195	195	190	180	185	180	170	165

La respuesta aquí ilustrada contiene el análisis de regresión apropiado; advierta que se usa una notación diferente: m representa la pendiente de la recta de regresión, n el número de pares de datos, B es la intersección con el eje y, R representa el coeficiente de correlación y S.D. significa la desviación estándar poblacional. El concepto de *grados de libertad* se estudiará posteriormente en el texto. Los paquetes de programación proporcionan con frecuencia más respuestas de las que se necesitan en una cierta etapa y usan notaciones distintas.

CORRELATION AND LINEAR REGRESION
VARIABLE X: AGE **VARIABLE Y: HEART RATE**
MEAN OF X = 30 **MEAN OF Y: 187**
S.D. OF X = 11.61895 **S.D. OF Y: 13.07670**
NUMBER OF PAIRS (N) = 10
CORRELATION COEFFICIENT (R) = –0.971
DEGREES OF FREEDOM (DF) = 8
SLOPE (M) OF REGRESSION LINE = –1.09259
Y INTERCEPT (B) FOR THE LINE = 219.778

Las respuestas computacionales de los programas comerciales con frecuencia no contienen alguna información deseada; pero muchas veces, la información faltante puede calcularse de la proporcionada en la respuesta.

Por ejemplo, de la respuesta anterior podemos determinar la ecuación de regresión y la suma de cuadrados de los errores, SSE. ¿Qué ritmo cardiaco máximo debería predecirse para una edad de 28 años?

Solución: La ecuación de regresión es $\hat{y} = 219.778 - 1.09259x$. Para la edad de 28 años, deberíamos esperar una pulsación máxima de:

$$\hat{y} = 219.778 - 1.09259(28) = 189.185$$

Para encontrar SSE, usaremos la fórmula (4.10); primero necesitamos encontrar SS_x, SS_y y SS_{xy}. Como la varianza poblacional está definida por $\sigma^2 = SS/N$ y la varianza muestral está definida por $s^2 = SS/(n-1)$, para determinar el valor de SS dado σ^2, multiplicamos σ^2 por N. Esto es, $s^2 = N/(n-1)\sigma^2$ y $SS = N\sigma^2$. En consecuencia,

$$\begin{aligned} SS_x &= N\sigma_x^2 = 10(11.61895)^2 = 1350 \\ SS_x &= N\sigma_y^2 = 10(13.07670)^2 = 1710 \end{aligned}$$

Como la pendiente de la recta de regresión está definida por:

$$m = \frac{SS_{xy}}{SS_x}$$

podemos despejar SS_{xy} de esta ecuación para obtener:

$$\begin{aligned} SS_{xy} &= mSS_x \\ &= (-1.09259)(1350) = -1474.9965 \end{aligned}$$

Por lo tanto, la suma de cuadrados de los errores es:

$$\begin{aligned} SSE &= SS_y - mSS_x \\ &= 1710 - (-1.09259)(-1474.9965) = 98.4336 \end{aligned}$$ ■

GRUPO DE EJERCICIOS 4.3

Habilidades básicas

1. Para cada una de las ecuaciones siguientes, encuentre la pendiente y la intersección de la recta con el eje y, y dibuje la gráfica.
   a) $y = 2x - 3$
   b) $y = x + 2$
   c) $2x + 3y = 6$
   d) $y = -2x$

2. Para cada una de las ecuaciones siguientes, encuentre la pendiente y la intercepción de la recta con el eje y y dibuje la gráfica:
   a) $y = -2x + 3$
   b) $y = 3x - 4$
   c) $4x - 3y = 12$
   d) $y = (2/3)x$

3. Encuentre la ecuación de regresión y SSE para los datos del ejercicio 1 del grupo de ejercicios 4.2. Los datos se repiten aquí por conveniencia.

x	1	5	2	4	8	9
y	3	7	2	6	7	4

4. Para el ejercicio 2 del grupo de ejercicios 4.2, encuentre la ecuación de regresión y SSE. Los datos se repiten aquí:

x	2	7	8	1	5	9	3
y	8	7	1	4	7	4	5

5. Para el ejercicio 7 del grupo de ejercicios 4.2, encuentre la ecuación de regresión y SSE. Los datos se repiten aquí:

x	0	1	4	–5	2	6
y	–2	–1	2	–7	0	4

6. Los datos del ejercicio 8 del grupo de ejercicios 4.2 se dan en la tabla. Encuentre la ecuación de regresión y SSE.

x	0	4	8	1	3	−1
y	2	6	−14	0	−4	4

7. Los datos del ejercicio 10 del grupo de ejercicios 4.2 se reproducen aquí. Encuentre la ecuación de regresión y SSE.

x	77	81	94	50	72	63	88	95
y	82	47	85	66	65	72	89	95

8. Los valores de los estadísticos siguientes se obtuvieron al analizar nueve pares de datos bivariados:

$\bar{x} = 2.8,\quad \bar{y} = 3.1\quad r = 0.049$
$s_x = 1.476,\quad s_y = 1.853,\quad n = 10$

Encuentre la ecuación de regresión $\hat{y} = b + mx$ y SSE.

9. Los valores siguientes se obtuvieron de nueve pares de datos bivariados:

$\bar{x} = 7.27667,\quad \bar{y} = 11.2722,\quad r = 0.622$
$s_x = 2.60702,\quad s_y = 5.24589,\quad n = 9$

Encuentre la ecuación de regresión $\hat{y} = b + mx$ y SSE.

Más aplicaciones

10. Considere los datos bivariados:

x	3	6	4	7
y	1	5	6	8

Sea $x' = x - \bar{x}$ y $y' = y - \bar{y}$ (recuerde que x' y y' son las desviaciones de los valores).

a) Determine la ecuación de regresión $\hat{y}' = b + mx'$.
b) Defina $\hat{y}'$ si $x' = 5$.

11. Considere los datos bivariados:

x	3	9	5	11
y	5	1	8	6

Sea $x' = x - \bar{x}$ y $y' = y - \bar{y}$ (recuerde que x' y y' son las desviaciones de los valores).

a) Determine la ecuación de regresión $\hat{y}' = b + mx'$.
b) Calcule $\hat{y}'$ si $x' = 11$.

12. El clima parece tener efecto en la ofensiva en el beisbol. Los datos adjuntos indican la relación entre temperatura y ofensiva desde 1987 hasta 1989.[30] (Esta aplicación se describió en el motivador 4.)

Temperatura	Porcentaje de bateo	Carreras por juego	Jonrones por juego
0° –59°	0.248	8.0	1.40
60° –69°	0.253	8.5	1.65
70° –79°	0.259	8.6	1.69
80° –89°	0.263	9.1	1.85
90° y más	0.263	9.1	1.83

Use las marcas de clase para las temperaturas 60°–89° para determinar la ecuación de regresión que puede usarse para predecir las carreras por juego para una temperatura dada.

13. Refiérase al ejercicio 12. Use las marcas de clase para las temperaturas 60°–89° para determinar la ecuación de regresión que puede usarse para predecir cuadrangulares por juego para una temperatura dada.

14. La tabla adjunta enlista las marcas olímpicas de natación, en segundos, en los 400 metros de estilo libre para mujeres desde 1924.

Año	Tiempo	Año	Tiempo
1924	362.20	1964	283.3
1928	342.8	1968	271.8
1932	328.5	1972	259.04
1936	326.4	1976	249.89
1948	317.8	1980	248.76
1952	312.1	1984	247.10
1956	294.6	1988	243.85
1960	290.6		

Encuentre la ecuación de regresión y úsela para predecir la marca mundial en 1962. (*Sugerencia*: codifique los años usando $x =$ año − 1923.)

15. Las marcas olímpicas de natación, en segundos, en los 400 metros de estilo libre para hombres desde 1924 son:

Año	Tiempo	Año	Tiempo
1924	304.02	1964	252.2
1928	301.6	1968	249.0
1932	288.4	1972	240.27
1936	284.5	1976	231.93
1948	281.0	1980	231.31
1952	270.7	1984	231.23
1956	267.3	1988	226.95
1960	258.3		

Encuentre la ecuación de regresión y úsela para predecir el tiempo de los hombres para 1992. (*Sugerencia*: codifique los años usando $x = \text{año} - 1923$.)

16. Se realizó un estudio para probar la efectividad de un nuevo fármaco para reducir el ritmo cardiaco en pacientes adultos que padecen del corazón, donde participaron mil enfermos; en la tabla adjunta se muestra la reducción promedio del ritmo cardiaco, medido en pulsaciones por minuto, para cada una de diez dosis, en miligramos, del fármaco.

Dosis (x)	50	75	100	125	150	175	200	225	250	275
Reducción promedio del ritmo cardiaco	8	5	13	11	13	12	18	18	16	19

a) Determine la ecuación de regresión para predecir la reducción promedio del ritmo cardiaco dada una dosis fija del fármaco.

b) Use la ecuación para predecir la reducción promedio del ritmo cardiaco de un paciente que toma 300 mg del medicamento.

Un paso más allá

17. Use las ecuaciones de regresión encontradas en los ejercicios 14 y 15, para calcular el año en el cual el tiempo predicho para las mujeres será igual al anunciado para los hombres en la competencia de 400 metros estilo libre. Analice sus resultados.

18. Para los datos en el ejercicio 10, encuentre los puntajes z, z_x para x, y los puntajes z, z_y para y. Determine entonces la ecuación de regresión $\hat{z}_y = b + mz_x$ y encuentre r.

19. Refiérase a la pantalla adjunta.

a) Encuentre la ecuación de la recta de regresión.

b) Encuentre SSE.

```
CORRELATION AND LINEAR REGRESSION
VARIABLE X:X              VARIABLE Y:Y
MEAN OF X = 20.13         MEAN OF Y = 10.31
S.D. OF X = 4.5927        S.D. OF Y = 8.33385
NUMBER OF PAIRS (N) = 10
CORRELATION COEFFICIENT (R) = 0.984
DEGREES OF FREEDOM (DF) = 8
SLOPE (M) OF REGRESSION LINE = 0.333427
Y INTERCEPT (B) FOR THE LINE = 3.59811
```

20. Demuestre la fórmula (4.11).

21. Compruebe que $\text{SSE} = \text{SS}_y - m^2\,\text{SS}_x$.

22. Una medida de la forma en que los puntos de un diagrama de dispersión se distribuyen alrededor de la recta de regresión es el **error estándar de estimación** s_e, lo cual está definido por $s_e = \sqrt{\text{SSE}/(n-2)}$. Encuentre s_e para los datos del ejercicio 10.

23. Demuestre que $\Sigma(y - \hat{y}) = 0$, donde $\hat{y} = b + mx$.

24. Compruebe que $(\bar{x}, \bar{y})$ es un punto de la recta de regresión.

25. Verifique que la ecuación de regresión puede escribirse como
$\hat{y} = \bar{y} + m(x - \bar{x})$.

26. Para los datos de la marca de la milla de la aplicación 4.4, encuentre la ecuación de regresión para predecir años, dados los tiempos para la marca; use esa ecuación para predecir el año en el cual la marca mundial para el recorrido de la milla será de 3:40. ¿En qué año (1999 o el año obtenido aquí) diría usted que dicha marca será de 3:40?

RESUMEN DEL CAPÍTULO

En este capítulo se introdujeron los conceptos de ecuaciones lineales, regresión lineal y correlación lineal; para determinar si existe una relación lineal entre dos variables, a menudo se usa un diagrama de dispersión; vimos que la fuerza de la relación lineal puede medirse por el coeficiente de correlación r y la covarianza muestral s_{xy}; los valores del coeficiente de correlación r pueden caer en cualquier lugar del intervalo entre -1 y 1, inclusive; si los puntos del diagrama de dispersión caen todos en una recta, el valor de r es 1 o -1, dependiendo de que la recta tenga pendiente positiva o negativa; un valor de $r = 0$ indica la falta de una relación lineal y un valor de r cercano a 1 o -1 no necesariamente implica una relación de causalidad. Aprendimos también cómo determinar la ecuación de regresión usando el método de mínimos cuadrados; la suma de los cuadrados de los errores, SSE, se minimiza cuando se usa el método de los mínimos cuadrados para calcular la ecuación de regresión.

REVISIÓN DEL CAPÍTULO

■ TÉRMINOS IMPORTANTES ■

Los términos del capítulo se han mezclado para proporcionar una mejor práctica de revisión; defina cada uno con sus propias palabras; después verifique sus respuestas contra las dadas en el capítulo.

error estándar de estimación	método de los mínimos cuadrados	línea recta
recta de regresión	ecuación lineal	suma de productos cruzados
codificación	ecuación de regresión	suma de cuadrados de los errores
análisis de regresión	diagrama de dispersión	covarianza muestral
variable dependiente	variable independiente	centroide
pendiente	dependencia lineal	análisis de correlación
datos bivariados	intercepción con el eje y	línea de mejor ajuste
coeficiente de correlación de Pearson	correlación espuria	

■ SÍMBOLOS IMPORTANTES ■

$\text{cov}(x, y)$, covarianza central	s_x, desviación estándar de x	$\hat{y}$, valor predicho de y
s_{xy}, covarianza muestral	s_y, desviación estándar de y	e_i, error de predicción
r, coeficiente de correlación	m, pendiente	SSE, suma de cuadrados de los errores
SS_{xy}, suma de productos cruzados	b, intercepción con el eje y	

■ HECHOS Y FÓRMULAS IMPORTANTES ■

Covarianza muestral: $\text{cov}(x, y) = s_{xy}$

$$= \frac{\Sigma (x - \bar{x})\ (y - \bar{y})}{n - 1} \qquad \textbf{(4.1)}$$

Coeficiente de correlación de Pearson:

$$r = \frac{\Sigma z_x z_y}{n - 1} \qquad \textbf{(4.2)}$$

Ecuación de una línea recta: $y = b + mx$

Suma de cuadrados de los errores:

$$SSE = \Sigma (y - \hat{y})^2 \qquad \textbf{(4.7)}$$

Suma de productos cruzados:

$$SS_{xy} = \Sigma xy - \frac{(\Sigma x)\ (\Sigma y)}{n} \qquad \textbf{(4.5)}$$

Fórmula para calcular el coeficiente de correlación muestral:

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}} \qquad \textbf{(4.6)}$$

Ecuación de regresión: $\hat{y} = b + mx$

Error de predicción: $e_i = y_i - \hat{y}_i$

Pendiente de la ecuación de regresión:

$$m = \frac{SS_{xy}}{SS_x} \quad (4.8)$$

Intersección con el eje y de la ecuación de regresión:

$$b = \bar{y} - m\bar{x} \quad (4.9)$$

Fórmula para calcular la suma de los cuadrados de los errores:

$$SSE = SS_y - mSS_{xy} \quad (4.10)$$

Relación entre r y m:

$$r = m\frac{s_x}{s_y} \quad (4.11)$$

EJERCICIOS DE REPASO

1. En un intento de determinar la relación entre el monto gastado en una campaña y el número de votos recibidos durante una elección, se recabaron los datos siguientes:

Monto gastado en miles de dólares: x	3	4	2	5	1
Votos recibidos (en miles): y	14	12	5	20	4

a) Dibuje un diagrama de dispersión.
b) Calcule el valor de s_{xy}
c) Determine el valor de r.
d) Encuentre la ecuación de regresión de SSE.
e) Prediga el número de votos recibidos si se gastaran 3,500 dólares en la campaña.
f) ¿Cuántos votos más pueden esperarse por cada 1,000 dólares adicionales gastados?
g) Dibuje una gráfica de la recta de regresión en el diagrama de dispersión.

2. Para estudiar la relación entre el número de veces que los estudiantes faltan a clases y sus calificaciones al final del curso, un instructor del grupo 209 de matemáticas obtuvo los datos mostrados aquí:

Número de faltas	1	2	3	3	4	4	5	6	2	0
Calificación	98	98	88	81	83	76	71	71	85	98

a) Dibuje un diagrama de dispersión.
b) Calcule el valor de s_{xy}.
c) Determine el valor de r.
d) Encuentre la ecuación de regresión y SSE.
e) Prediga la calificación final si un estudiante ha faltado a tres clases.
f) ¿En cuánto se supone que se afectará la calificación final por cada falta adicional?
g) Dibuje una gráfica de la recta de regresión en el diagrama de dispersión.

3. Estudiantes que presentan examen de admisión (x), número de inscritos por primera vez (y), en los pasados siete años en una universidad. Resuelva los incisos con estos datos:

x	3300	4100	5600	5200	5900	5500	5100
y	3000	3500	4200	4800	5000	5100	4700

a) Dibuje un diagrama de dispersión.
b) Calcule el valor de s_{xy}.
c) Determine el valor de r.
d) Encuentre la ecuación de regresión y SSE.
e) Localice $\hat{y}$ si $x = 5000$.
f) ¿Cuántas inscripciones más pueden esperarse por cada 1000 solicitudes adicionales?
g) Dibuje una gráfica de la recta de regresión en el diagrama de dispersión.

4. Un biomédico estudió el efecto de dosis diferentes (x) de un nuevo fármaco en el ritmo cardiaco (y) de los seres humanos. Los resultados para cinco individuos se indican en la tabla siguiente:

Dosis (x)	2.5	3	3.5	4	4.5
Descenso en el ritmo cardiaco (y)	8	11	9	16	19

a) Calcule el valor de r e interprete el resultado.
b) Encuentre la ecuación de regresión.
c) Encuentre SSE.
d) Determine $\hat{y}$ si $x = 3.75$.
e) Por cada unidad de incremento en la dosis, ¿cuál es el descenso predicho en el ritmo cardiaco?
f) Dibuje una gráfica de la recta de regresión en el diagrama de dispersión.

5. En un análisis de regresión se determinó la información siguiente:

$\hat{y} = 25.187x - 878.8583$

$s_y = 278.5247$ $\quad s_x = 9.3956$

$\bar{x} = 51.5$ $\quad \bar{y} = 418.3$ $\quad n = 10.$

a) Encuentre el valor de r.

b) Encuentre SSE.

c) Si $x = 45$, determine $\hat{y}$.

d) ¿En cuánto cambiará $\hat{y}$ por cada unidad de incremento en x?

6. Los datos adjuntos representan las ventas anuales de armas, en billones de dólares, de Estados Unidos a las naciones del Tercer Mundo.

Año	1976	1977	1978	1979	1980	1981	1982	1983
Ventas	8.2	9.8	10.1	9.2	6.4	6.8	7.9	9.7

a) Encuentre la ecuación de regresión para predecir las ventas:

b) Use la ecuación encontrada de la parte a, para estimar las ventas de Estados Unidos a las naciones del Tercer Mundo para el año de 1984. (*Sugerencia*: codifique el año usando: x = año – 1975.)

7. Nueve peces dorados fueron aclimatados a una temperatura del agua de 3° C; luego fueron sometidos a un incremento gradual de la temperatura del agua para averiguar si el metabolismo está relacionado con la temperatura; el metabolismo se midió contando los parpadeos por minuto. Los datos resultantes se enlistan en la tabla adjunta.

Temperatura °C	Número promedio de parpadeos por minuto
5.0	33.0
7.5	44.8
10.0	54.0
12.5	52.5
15.0	70.2
17.5	99.8
20.0	110.5
22.5	117.0
25.0	129.1

a) Dibuje un diagrama de dispersión.

b) Calcule r.

c) Encuentre la ecuación de regresión.

d) Determine SSE.

e) Si la temperatura fuera 0° C, ¿cuántos parpadeos por minuto deberían esperarse?

Los datos siguientes son para los ejercicios de revisión 8 y 9.

Distancia al árbol más cercano (en pies)	Altura (en pies)	DAP (en pulgadas)
1	39	5.7
19.5	72	9.2
4.0	69	9.3
5.0	67	9.5
10.5	73	9.5
9.0	78	9.7
8.0	79	9.8
15.0	81	10.4
12.5	65	10.7
18.0	60	11.7

8. Un guardabosques quiere saber la correlación entre la altura total y el diámetro a la altura de su pecho DAP, de una muestra de álamos temblones. Con los datos de la tabla determine:

a) el valor de r,

b) la ecuación de regresión,

c) SSE.

9. Un guardabosques quiere conocer la correlación entre el tamaño, medido como el diámetro a la altura de su pecho, y la distancia al árbol más cercano en una muestra de álamos temblones. Refiérase a los datos de la tabla.

a) Dibuje un diagrama de dispersión.

b) Calcule el valor de r.

c) Encuentre la ecuación de regresión para predecir DAP a partir de la distancia al árbol más cercano.

d) Encuentre SSE.

10. Para determinar si están relacionados el flujo del tránsito, medido en el número de vehículos por hora, y el contenido de plomo en la vegetación que crece cerca de las carreteras, se realizó un estudio en seis localidades, y se obtuvieron los siguientes datos:

Número de vehículos	103	216	294	402	416	573
Contenido de plomo	4.6	7.4	26.1	37.2	24.8	38.7

Encuentre la ecuación de regresión y úsela para predecir el contenido de plomo de la vegetación que experimenta un flujo de tránsito de 300 vehículos por hora.

Aplicaciones de computación

1. La siguiente tabla lista los precios de venta x, en miles de dólares, y el área y en cientos de pies cuadrados, para una muestra de diez casas en Somerset County el año pasado.

x	98	118	94.6	92.5	68.5	77	62.5	89.1	120.5	33.6
y	16	39	24	17	17	13	18	15	20	13

Utilice un programa computacional para:

a) Encontrar la ecuación de la recta de regresión para el área en términos del precio de venta.

b) Determinar la ecuación de la recta de regresión para el precio de venta en términos del área.

c) Saber el coeficiente de correlación.

d) Conocer la covarianza muestral.

e) Dibujar un diagrama de dispersión.

f) Estimar el precio de venta el año pasado de una casa que tiene un área de 1500 pies cuadrados.

g) Calcular el área de una casa que tuvo un precio de venta de 100,000 dólares el año pasado.

2. La tabla siguiente lista el precio de venta (x) en miles de dólares y el total anual de los impuestos estatales reales (y), en miles de dólares, para una muestra de diez casas vendidas en Somerset County el año pasado.

x	98	118	94.6	92.5	68.5	77	62.5	89.1	120.5	33.6
y	3.2	2.9	2.4	2.7	2.1	2.6	1.6	2.7	4.7	0.73

Use un programa computacional y resuelva los incisos:

a) encuentre la ecuación de la recta de regresión para el área en términos del monto del impuesto;

b) defina la ecuación de la recta de regresión para el monto del impuesto en términos del precio de venta;

c) ¿cuál es el coeficiente de correlación?;

d) calcule la covarianza muestral;

e) dibuje un diagrama de dispersión;

f) estime el precio de venta de una casa cuyo impuesto fue de 2,500 dólares el año pasado;

g) estime el impuesto estatal de una casa que tuvo un precio de venta de 100,000 dólares el año pasado.

3. Según un estudio realizado por la Northwestern National Life Insurance Company, los ciudadanos de Utah, el estado con la población más saludable, viven dos años y medio más que los ciudadanos de Delaware, el estado con la población menos saludable. Dicho trabajo ordena los estados del 1 al 50; el estado 1 tiene la población más saludable y el estado 50 la población menos saludable, usando seis factores:

Esperanza de vida – 1 = esperanza de vida en la fecha del nacimiento.

Enfermedad – 1 = baja frecuencia de enfermedades graves.

Estilo de vida – 1 = mayor participación en hábitos saludables y consumo mínimo de alcohol y cigarros.

Disponibilidad – 1 = máxima disponibilidad de médicos y máxima parte de la población con servicios de salud.

Tiempo perdido – 1 = mínimo de faltas a la escuela y al trabajo y mínimo de enfermedades agudas.

Mortalidad – 1 = mínimo de proporción de muertes totales de infantes y muertes prematuras.

La tabla adjunta lista los resultados del estudio:[31]

Lugar	Estado	Esperanza de vida	Enfermedad	Estilo de vida	Disponibilidad	Tiempo perdido	Mortalidad
1.	Utah	3	1	1	38	10	1
2.	Dakota del Norte	5	4	6	10	18	4
3.	Idaho	10	2	2	25	10	11
4.	Minnesota	2	5	18	3	31	6
5.	Hawaii	1	16	35	21	1	1
6.	Vermont	17	6	42	3	37	16
7.	Nebraska	6	7	3	17	37	22
8.	Colorado	9	9	29	20	7	5
9.	Wyoming	26	3	26	37	9	10
10.	Montana	25	7	18	16	13	18
11.	Washington	11	11	22	16	10	13
12.	Oregon	14	14	15	5	18	24
13.	Nuevo México	22	12	17	43	2	8
14.	Wisconsin	7	12	31	17	31	14
15.	Dakota del Sur	14	10	7	23	27	29
16.	Iowa	3	19	3	30	45	14
17.	Maine	19	17	25	12	45	23
18.	California	19	26	31	8	4	7
19.	Massachusetts	13	23	38	1	37	21
20.	Alaska	46	21	47	40	3	3
21.	Indiana	26	18	18	33	31	31
22.	Arizona	21	25	37	28	5	16
23.	Oklahoma	31	20	12	42	18	32
24.	New Hampshire*	14	15	49	11	37	9
25.	Kansas*	7	35	5	21	31	19
26.	Texas*	31	26	29	45	5	11
27.	Pennsylvania*	34	24	24	12	50	39
28.	Connecticut	11	36	38	1	27	20
29.	Kentucky	41	22	14	25	37	36
30.	New Jersey	22	28	38	19	18	28
31.	Missouri	26	31	13	33	45	32
32.	Ohio	35	30	26	25	37	30
33.	Virginia	36	29	38	31	18	27
34.	Arkansas	29	32	7	35	37	44
35.	West Virginia	43	34	9	28	45	47
36.	Illinois	37	37	31	24	13	36
37.	Nueva York	29	42	36	7	13	38
38.	Louisiana*	50	33	28	47	13	40
39.	Tennessee*	39	41	10	35	37	40
40.	Rhode Island	18	43	45	9	45	26
41.	Carolina del Norte	42	40	22	43	27	40
42.	Alabama	45	39	10	46	31	48
43.	Maryland	37	49	44	14	18	34
44.	Florida	22	46	48	15	31	45

*Estos estados se unieron

Lugar	Estado	Esperanza de vida	Enfermedad	Estilo de vida	Disponibilidad	Tiempo perdido	Mortalidad
45.	Georgia	46	47	31	48	18	45
46.	Carolina del Sur	49	45	18	49	13	49
47.	Nevada	44	38	50	32	8	24
48.	Michigan	31	50	42	40	18	35
49.	Mississippi	48	44	16	50	18	50
50.	Delaware	40	48	46	38	27	40

Calcule el coeficiente de correlación de Pearson para los rangos de:

a) Salud y estilo de vida en un estado.

b) Salud y esperanza de vida en un estado.

c) Estilo de vida y tiempo perdido.

d) Disponibilidad y tiempo perdido.

e) Enfermedad y estilo de vida.

4. Recabe información de 25 conductores con licencia de manejo en el estado donde se localiza su universidad: fecha de nacimiento, codificada de 1 a 366, y los últimos tres dígitos del número de su licencia de manejo.

a) Determine el coeficiente de correlación de Pearson.

b) Calcule la ecuación de regresión para predecir los tres últimos dígitos de la fecha de nacimiento codificada.

c) Analice el significado de sus resultados.

▪ EXAMEN DE CONOCIMIENTOS DEL CAPÍTULO ▪

1. El número de sentadillas $\hat{y}$ que un niño normal y saludable debe ser capaz de realizar, con base en su edad x, está dado por $\hat{y} = 1.4x - 0.9$, donde $4 \leq x \leq 17$.

a) ¿Cuántas sentadillas debe esperarse que haga un niño de diez años de edad?

b) Cuando la edad aumenta, ¿crece o decrece el número de sentadillas?

c) ¿Cuántas sentadillas más debe esperarse que haga un niño por cada año de edad?

d) En este caso, ¿debe tener la intercección con el eje y una interpretación con sentido? Explique.

2. Un estudio de la relación entre la estatura, en pulgadas, y el peso, en libras, de los universitarios hombres produjo los datos dados aquí.

Estatura (x)	64	72	73	68	66	67
Peso (y)	165	158	173	125	125	139

Determine lo siguiente:

a) SS_x

b) SS_y

c) SS_{xy}

d) la pendiente, m,

e) la intercección con el eje y, b,

f) la ecuación de regresión,

g) el coeficiente de correlación de Pearson,

h) SSE,

i) $\hat{y}$ cuando $x = 65$,

j) el error e para el hombre en la muestra cuya estatura es $x = 71$ pulgadas.

3. Existe una fuerte correlación positiva entre los salarios de los maestros y el consumo anual de cerveza en Estados Unidos.

a) ¿Significa esto que el aumento en el consumo de cerveza ha sido causa del incremento a los salarios de los maestros? ¿O que mientras más bebe un maestro más pago conseguirá? Explique.

b) ¿Qué factor(es) adicional(es) pueden causar que los salarios de los maestros y el consumo de cerveza se incrementen simultáneamente?

4. Como $\hat{y} = b + mx$ y $b = \bar{y} - m\bar{x}$, tenemos $\hat{y} = (\bar{y} - m\bar{x}) + mx$. Por tanto, $\hat{y} - \bar{y} = m(x - \bar{x})$. Con la ecuación de regresión escrita en esta forma, demuestre que el centroide $(\bar{x}, \bar{y})$ está en la recta de regresión.

5. Encuentre SSE para los datos bivariados siguientes (x, y):

x	1	3	5
y	2	4	0

6. Localice la covarianza muestral para los datos pareados:

x	1	3	5	15	16
y	11	15	21	25	28

UNIDAD DOS

Probabilidad básica

5 Introducción a la probabilidad elemental

DESCRIPCIÓN

OBJETIVOS DEL CAPÍTULO

En este capítulo estudiaremos:

- *Experimentos.*
- *Espacios muestrales.*
- *Eventos.*
- *Qué es un evento compuesto y cómo formar eventos compuestos.*
- *Eventos complementarios.*
- *Cómo usar diagramas de Venn para representar eventos.*
- *Eventos mutuamente excluyentes.*
- *Probabilidad.*
- *Los métodos de asignar probabilidades.*
- *Qué son las posibilidades matemáticas y cómo calcularlas.*
- *Teorema fundamental del conteo.*
- *Cómo usar el teorema fundamental del conteo para encontrar probabilidades.*
- *Qué son las permutaciones y cómo usarlas para contar.*
- *Qué son las combinaciones y cómo usarlas para contar.*
- *Algunas reglas fundamentales de probabilidad.*
- *Eventos independientes.*
- *Variables aleatorias.*
- *Cómo encontrar la media y la varianza de una variable aleatoria.*

MOTIVADOR 5

El problema médico del síndrome de inmunodeficiencia adquirida (SIDA) ha creado mucha actividad dentro de la comunidad de investigadores para identificar a aquellas personas que tienen el virus. ¿Sabía usted que si la probabilidad de que una persona con el virus del SIDA de un examen positivo es 98/100, eso no necesariamente significa que la probabilidad de que una persona sin el virus del SIDA de un examen positivo sea 2/100?

La prueba usada más frecuentemente para discernir si la sangre donada está contaminada con el virus del SIDA es la prueba de ELISA, desarrollada para mantener el SIDA fuera del suministro de sangre; cuando la sangre contiene el virus del SIDA, la prueba de ELISA es positiva con probabilidad de 98/100 y negativa con probabilidad de 2/100; si la sangre no contiene los anticuerpos para el virus, la prueba es positiva con probabilidad 7/100 y negativa con probabilidad 93/100.

La alta probabilidad de que los resultados positivos sean falsos, se ve como aceptable debido a la baja probabilidad de que sea falso un resultado negativo.

Si la prueba de ELISA se aplica a 10,000 muestras de sangre de una población de alto riesgo, cuya probabilidad de tener la enfermedad es 5/1000, entonces debemos esperar obtener 50 muestras contaminadas; **de éstas, 0.98 × 50 = 49 darán una prueba positiva; si las quitamos de las 10,000 muestras esperaríamos que 9,950 de ellas no estén contaminadas, y de éstas 0.07 × 9950 = 696.5 darán una prueba positiva falsa, si usamos la prueba de ELISA. El número de falsas negativas es 0.02 × 50 = 1. El alto número de positivas falsas, 696.5; en este caso** debería considerarse aceptable debido al bajo número de negativas falsas, 1, asociado con la prueba de ELISA.

La probabilidad de que una persona tenga SIDA dado que su sangre dé una prueba positiva es condicional, un tema que se estudiará en la sección 5.5. Idealmente, se necesita una prueba con un alto valor predictivo para resolver positivo o negativo; es decir, si la prueba es positiva entonces es altamente probable que la persona tenga SIDA. Después de estudiar los capítulos siguientes, deberá quedar razonablemente claro por qué es imposible desarrollar una prueba de SIDA con un valor predictivo del 100 por ciento.[32]

Panorama del capítulo

El propósito de este capítulo es desarrollar las ideas básicas que se necesitarán para una adecuada comprensión de la estadística inferencial. La estadística inferencial es un cuerpo de conocimientos que trata los métodos para caracterizar una población usando información calculada de muestras extraídas de la misma. Estos métodos siempre conllevan cierto grado de incertidumbre. Por ejemplo, podemos hacer las preguntas siguientes:

1. ¿Cuál es la posibilidad de que cierre la AJAX Company?
2. ¿Cuál es el peso promedio de los bebés de un mes de nacidos?
3. ¿Es mejor la marca A que la marca B?

Todos los días enfrentamos tomas de decisiones y planteamientos probabilísticos. Los planteamientos que contienen las palabras posibilidad, plausibilidad, oportunidad, parecido, esperado, posible, incierto y probabilidad, se refieren todos al mismo tema: la incertidumbre. A diario, hacemos u oímos planteamientos como los siguientes:

1. ¿Cuál es la probabilidad de que tengamos un examen hoy?
2. Las oportunidades de lo golpee un poste del alumbrado son de 1 en 2 millones.
3. El trabajo se terminaría plausiblemente a tiempo.
4. Las posibilidades de que llueva hoy son del 50 por ciento.
5. Si se arroja una moneda, hay una posibilidad de 50-50 para que salga cara.

6. ¿Cuál es la probabilidad de que el nuevo método propuesto lleve a mejores resultados?
7. Tengo confianza en que puedo aprobar este curso.

Contamos con un buen concepto intuitivo de la probabilidad; partiendo de esta base, exploraremos algunas de las propiedades no tan obvias de la teoría de la probabilidad para ayudarnos a desarrollar una mejor comprensión de la estadística inferencial. La probabilidad nos ofrece el fundamento para desarrollar la ciencia de la estadística inferencial; mediante la teoría de la probabilidad, podemos deducir la posibilidad de que aparezcan ciertas muestras con propiedades específicas. Tal información nos permitirá obtener inferencias sobre una población. Comenzamos nuestro estudio con una exposición sobre los experimentos.

SECCIÓN 5.1 Experimentos y eventos

Experimentos

Un **experimento** es cualquier proceso planeado que da lugar a observaciones o a recolección de datos; un ejemplo muy simple de un experimento es arrojar un dado y observar el número de la cara que queda hacia arriba cuando se detiene; para este experimento habría seis posibles resultados registrados, 1, 2, 3, 4, 5 o 6. También, el resultado de este experimento puede registrarse como un número par o impar, por eso es importante definir con cuidado la forma del registro, pues la misma situación básica produciría datos que pueden registrarse desde varios puntos de vista. Como otro ejemplo de un experimento, digamos que probamos una muestra de un producto terminado, seleccionada de una línea de ensamble para determinar si el producto es defectuoso o no; este experimento tiene dos resultados: el producto es defectuoso o no, pero en cualquiera de estas dos situaciones podemos repetir el experimento muchas veces. Para nuestros fines, estaremos interesados principalmente en experimentos que pueden repetirse o que se acepte que pueden repetirse. Experimentos como observar si lloverá mañana o determinar quién ganará la serie mundial el año próximo no son repetibles y no los consideraremos en esta sección, sino que estaremos interesados con frecuencia en los resultados obtenidos de repetir un experimento un cierto número de veces.

Todos los experimentos tienen resultados y la mayor parte de ellos son inciertos y dependen del azar; los resultados de un experimento forman un conjunto llamado espacio muestral.

> Un **espacio muestral** de un experimento es la colección de todos los resultados posibles.

El experimento más simple referente a incertidumbre es uno que tiene dos resultados y un espacio muestral único, como vemos en el ejemplo 5.1. Sin embargo, un experimento puede tener más de un espacio muestral; es decir, se puede usar más de un espacio muestral para describir los resultados de un experimento. En general, es deseable elegir un espacio muestral que proporcione la máxima información referente al experimento. Considere los ejemplos 5.2 y 5.3 y la aplicación 5.1.

EJEMPLO 5.1

Observar el sexo del siguiente bebé que nazca en el Memorial Hospital es un experimento con dos resultados; un espacio muestral para este experimento consiste en el conjunto $S = \{H, M\}$, donde H representa a un hombre, M representa a una mujer y las llaves se usan para indicar colección o conjunto.

EJEMPLO 5.2

Si observamos los nacimientos de los siguientes dos bebés nacidos en el Memorial Hospital, entonces un espacio muestral para el experimento podría ser el conjunto $S_1 = \{HH, HM, MH, MM\}$, donde, por ejemplo, MH indica que el primer bebé en nacer fue una mujer y el segundo fue un hombre. Otro espacio muestral para este experimento puede consistir en el número de posibles nacimientos de hombres: $S_2 = \{0, 1, 2\}$; note que estos dos espacios muestrales proporcionan información diferente. Del resultado de sólo un hombre en S_2, no podemos determinar si fue el primero o el segundo en nacer; para este experimento referente al mismo tema, hemos registrado los datos en dos formas distintas.

EJEMPLO 5.3

Para el experimento de arrojar un dado, $S_1 = \{1, 2, 3, 4, 5, 6\}$ es un espacio muestral y el conjunto $S_2 = \{\text{número par, número impar}\}$ también lo es. Saber que el resultado de un lanzamiento de un dado es par, no posibilita determinar si fue 2, 4 o 6.

APLICACIÓN 5.1

Liste un espacio muestral para cada uno de los siguientes experimentos:

a) Lanzar una moneda de diez centavos y otra de veinte en ese orden, y observar cómo caen.
b) Lanzar una moneda de un peso, una de cinco centavos y otra de diez centavos, en ese orden, y observar cómo caen.
c) Seleccionar a una estudiante universitaria al azar y preguntarle cuántos años tiene.
d) Lanzar primero una moneda, de un lado y observar cómo cae.
e) Lanzar una moneda hasta que salga águila.

Solución:

a) $S = \{SA, AS, AA, SS\}$. El resultado SA significa que la moneda de diez centavos muestra el "sol" y la de veinte el "águila".
b) $S = \{SSS, SSA, SAS, SAA, ASS, AAS, ASA, AAA\}$. El resultado SAS significa que el peso muestra el "sol", la moneda de cinco centavos el águila y la moneda de diez, el "sol".
c) $S = \{10, 11, 12, 13, ..., 98, 99, 100\}$.
d) $S = \{S1, S2, S3, S4, S5, S6, A1, A2, A3, A4, A5, A6\}$.
e) $S = \{S, AS, AAS, AAAS, AAAAS, etc.)$. El resultado AAAS significa que salió "sol" en el cuarto lanzamiento. ■

Eventos

Para un cierto experimento, podemos estar interesados en determinar la probabilidad de que ocurra una colección de resultados, en lugar de la probabilidad de que se dé uno solo. Por ejemplo, cuando se lanzan tres monedas a la vez, podemos estar interesados en los resultados que indiquen que al menos han salido dos "soles"; esta colección de resultados, (SSA, SAS, ASS, SSS) se llama un **evento**.

> Un **evento** es cualquier subcolección (o **subconjunto**) de un espacio muestral S.

EJEMPLO 5.4

Suponga que el experimento es lanzar primero un peso y luego una moneda de diez centavos; un espacio muestral para este experimento podría ser $S = \{SS, SA, AS, AA\}$. Algunos eventos posibles son:

$$E_1 = \{SS\} \quad E_4 = \{SS, AA\}$$
$$E_2 = \{SA\} \quad E_5 = \{SA, AS\}$$
$$E_3 = \{AS\} \quad E_6 = \{AA\}$$

Hay 16 eventos posibles. El **conjunto vacío** y el espacio muestral S también son eventos; el evento E_6 puede describirse como la obtención de un águila en el peso y un águila en la moneda de diez centavos, y el evento E_4 puede describirse como la obtención de dos "soles" o dos "águilas".

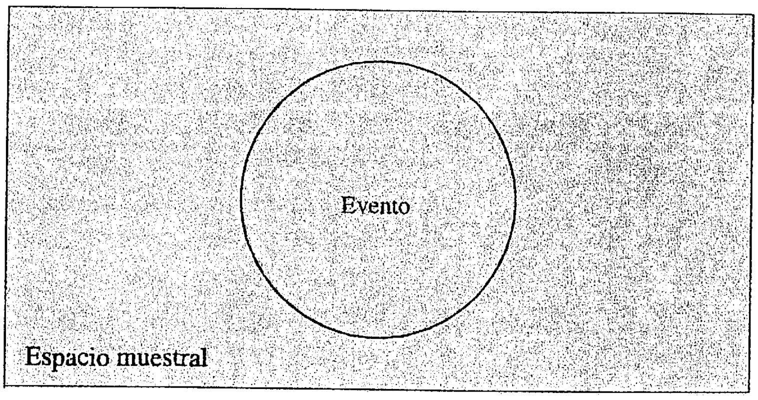
> Un **evento simple** es un evento que contiene sólo un resultado.

EJEMPLO 5.5

El evento $E_1 = \{SS\}$ del ejemplo 5.4 es un evento simple, mientras que el E_4 no lo es.

Recuerde que un evento es siempre una colección de resultados del universo de todos los resultados identificados como el espacio muestral. Para representar gráficamente espacios muestrales y relaciones entre eventos se puede usar un **diagrama de Venn**; se acostumbra que un rectángulo denote el espacio muestral y que los eventos se representen como círculos dentro del rectángulo, como se indica en la figura 5.1.

FIGURA 5.1
Diagrama de Venn que representa el espacio muestral y un evento



APLICACIÓN 5.2

Suponga que un experimento consiste en examinar tres fusibles; cada fusible puede ser defectuoso (D) o no defectuoso (N), los ocho resultados posibles son {NDD, NDN, NND, NDD, DND, DNN, DDN, DDD}. Anote los resultados que conforman cada uno de los eventos siguientes:

a) E_1 = El primer fusible está defectuoso.
b) E_2 = el primer fusible y el último están defectuosos.
c) E_3 = todos los fusibles son buenos.
d) E_4 = al menos un fusible está defectuoso.
e) E_5 = a lo más un fusible está defectuoso.

Solución:

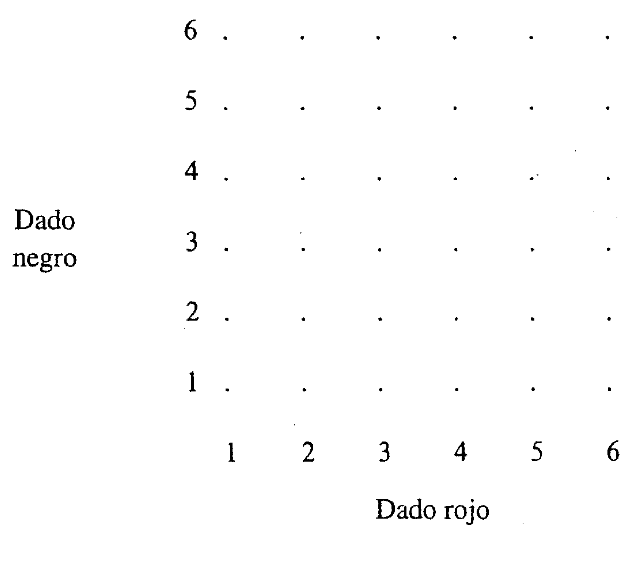
a) E_1 = {DNN, DND, DDN, DDD}
b) E_2 = {DDD, DND}
c) E_3 = {NNN}
d) Note que *al menos uno* significa uno o más. E_4 = {NDD, NDN, NND, DNN, DND, DDN, DDD}
e) Note que *a lo más uno* significa uno o menos. E_5 = {NDN, NND, NNN, DNN}. ■

APLICACIÓN 5.3

Considere el experimento de tirar un dado rojo y uno negro, y observar cómo caen; un espacio muestral de 36 resultados posibles es como sigue, donde la primera entrada es el resultado del dado rojo y la segunda el del dado negro.

$$
\begin{array}{llllll}
\{(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)\}
\end{array}
$$

El espacio muestral S se puede representar por el siguiente diagrama:



Dé una descripción para los eventos:

a) $\{(1, 1), (2, 1), (3, 1), (4, 1), (5, 1), (6, 1)\}$
b) $\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$
c) $\{(3, 4), (4, 3), (5, 2), (2, 5), (6, 1), (1, 6)\}$
d) $\{(5, 6), (6, 5)\}$
e) $\{(1, 1)\}$
f) $\{(4, 4)\}$

Solución:

a) El dado negro muestra el 1.
b) Los dos dados coinciden.
c) La suma de los dados es igual a 7.
d) La suma de los dados es igual a 11.
e) Ambos dados muestran 1 (a un par de unos se le llama "ojo de serpiente").
f) Ambos dados muestran un 4 (a un par de cuatros se le conoce como "furgón").

APLICACIÓN 5.4

Para el experimento de lanzar los dados en la aplicación 5.3, liste los resultados de los eventos siguientes:

a) La suma es par.
b) La suma es divisible entre 5.
c) La suma es un número primo. (Un número primo es un número mayor que 1, divisible sólo entre 1 y él mismo.)
d) El número del dado negro es dos unidades mayor que el número del dado rojo.
e) La suma es impar.
f) La suma no es divisible exactamente entre 5.

Solución:

a) Los pares siguientes tienen una suma que es par:

$$\begin{matrix} \{(1,1) & (1,3) & (1,5) & (2,2) & (2,4) & (2,6) \\ (3,1) & (3,3) & (3,5) & (4,2) & (4,4) & (4,6) \\ (5,1) & (5,3) & (5,5) & (6,2) & (6,4) & (6,6)\} \end{matrix}$$

b) Estos pares tienen una suma divisible entre 5:

$$\{(1,4),\quad (4,1),\quad (3,2),\quad (2,3),\quad (5,5),\quad (6,4),\quad (4,6)\}$$

c) Y éstos, una suma que es un número primo:

$$\begin{matrix} \{(1,2), & (2,1), & (1,4), & (4,1), & (1,6), & (6,1), & (2,5), & (5,2) \\ (3,4), & (4,3), & (5,6), & (6,5), & (2,3), & (3,2)\} \end{matrix}$$

d) Los pares siguientes tienen un número en el dado negro que es dos unidades mayor que el del dado rojo:

$$\{(1,3),\quad (2,4),\quad (3,5),\quad (4,6)\}$$

e) Estos pares tienen una suma que es impar:

$$\begin{matrix}\{(1,2) & (1,4) & (1,6) & (2,1) & (2,3) & (2,5) \\ (3,2) & (3,4) & (3,6) & (4,1) & (4,3) & (4,5) \\ (5,2) & (5,4) & (5,6) & (6,1) & (6,3) & (6,5)\}\end{matrix}$$

f) Y éstos, una suma que no es divisible entre 5:

$$\begin{matrix}\{(1,1) & (1,2) & (1,3) & (1,5) & (1,6) & (2,1) \\ (2,2) & (2,4) & (2,5) & (2,6) & (3,1) & (3,3) \\ (3,4) & (3,5) & (3,6) & (4,2) & (4,3) & (4,4) \\ (4,5) & (5,1) & (5,2) & (5,3) & (5,4) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,5) & (6,6)\} & \end{matrix}$$ ■

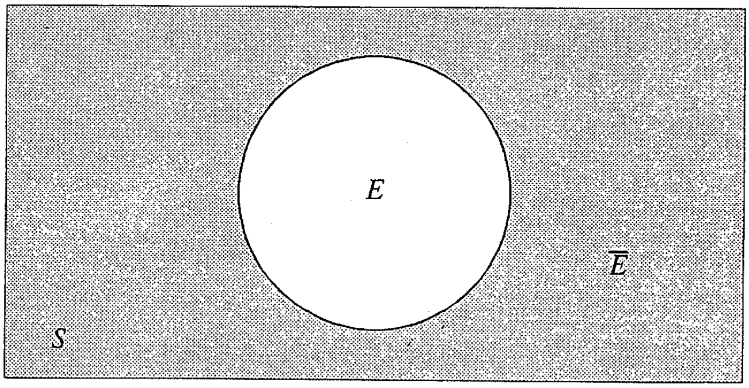
Evento (no E)

Si E es un evento contenido en un espacio muestral S, entonces el evento no E, denotado por $\overline{E}$, es el que contiene todos los resultados en S que no están contenidos en E.

En el diagrama de Venn de la figura 5.2, $\overline{E}$ es el área sombreada dentro de S y fuera de E.

FIGURA 5.2

Diagrama de Venn del evento "no E"

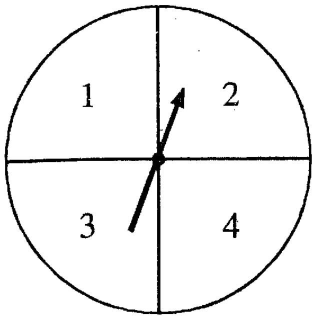


EJEMPLO 5.6

Considere el experimento de lanzar un dado. Si E es el evento de obtener un 4 o un 6, entonces el evento $\overline{E}$ contiene los resultados 1, 2, 3 y 5; esto es, si $E = \{4, 6\}$, entonces $\overline{E} = \{1, 2, 3, 5\}$.

Suponga que E es un evento de algún experimento; como un evento E ocurre o no ocurre, un espacio muestral para el experimento es $S = \{E, \overline{E}\}$; en consecuencia, cualquier experimento tiene un espacio muestral con sólo dos resultados, E y $\overline{E}$. Considere la aplicación 5.5.

APLICACIÓN 5.5



Para cada uno de los siguientes experimentos, liste un espacio muestral con sólo dos resultados.

a) Lanzar una moneda cinco veces y observar el número de caras.
b) Lanzar una moneda de diez centavos seguida de otra.
c) Girar la aguja y observar dónde se detiene. (Entra esquema, al margen.)

Solución: Para cada experimento anotamos dos de los espacios muestrales posibles con sólo dos resultados. Hay otros espacios muestrales que el estudiante puede obtener.

a) S_1 = {obtener 5 caras, no obtener 5 caras}
S_2 = {obtener 2 caras, no obtener 2 caras}
b) S_1 = {la suma es par, la suma es impar}
S_2 = {obtener dos veces 3, no obtener dos veces 3}
c) S_1 = {se detiene en 1, no se detiene en 1}
S_2 = {se detiene en 2, no se detiene en 2} ■

Eventos compuestos

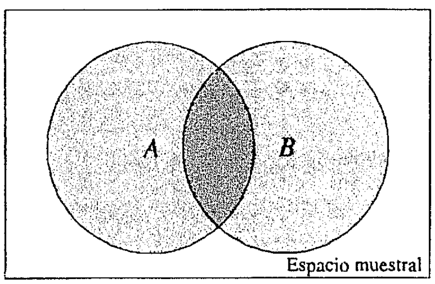
Como los eventos son conjuntos, los operadores de unión y de intersección pueden usarse para formar **eventos compuestos**; si E y F son eventos, entonces $(E \cup F)$ y $(E \cap F)$ son ejemplos de eventos compuestos.

$(E \cup F)$ es el evento de que ocurran E, o F, o ambos.

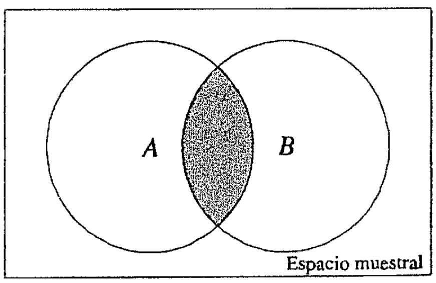
$(E \cap F)$ es el evento de que tanto E como F ocurran al mismo tiempo.

Se puede usar diagramas de Venn para ilustrar eventos compuestos; en los de la figura 5.3, los eventos compuestos $(A \cup B)$ y $(A \cap B)$ están representados por las regiones sombreadas.

FIGURA 5.3
Diagrama de Venn de eventos compuestos



El evento $A \cup B$



El evento $A \cap B$

APLICACIÓN 5.6

Considere el experimento de lanzar los dados de la aplicación 5.3. Sean: E el evento de que el número en el dado rojo sea 4 y F que la suma de los números mostrados sea 7; encuentre los resultados que forman los eventos siguientes:

a) E
b) F
c) $E \cup F$
d) $E \cap F$

Solución: Vea la aplicación 5.3 para la lista de S.

a) $E = \{(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6)\}$.
b) $F = \{(1. 6), (6, 1), (5, 2), (2, 5), (3, 4), (4, 3)\}$.
c) $E \cup F = \{(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), (1, 6), (6, 1), (5, 2), (2, 5), (3, 4)\}$. Note que (4, 3) no se incluye en dos ocasiones en el conjunto.
d) $E \cap F = \{(4, 3)\}$. ■

APLICACIÓN 5.7

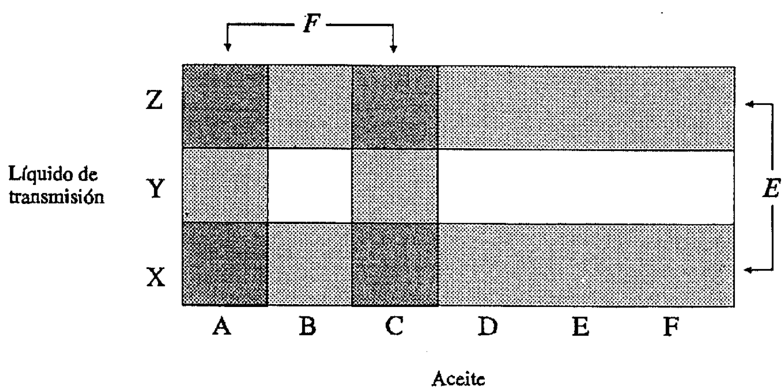
Javier va a una estación de servicio automotriz rápido a comprar un cuarto de líquido para transmisión y un cuarto de aceite para su coche; si hay disponibles tres marcas de líquido para transmisión (X, Y, Z) y seis marcas de aceite (A, B, C, D, E, F), Javier puede hacer 18 compras. Suponga además que las tres marcas de líquido de transmisión y las seis de aceite coinciden cada una en el precio; como Javier no sabe mucho sobre el aceite que usa su coche, decide hacer sus compras adivinando. Sea E el evento de que Javier compre la marca X o la Z de líquido de transmisión, y F el evento de que compre la marca de aceite A o la B. Liste cada uno de los siguientes:

a) Un espacio muestral S para el experimento de escoger un cuarto de líquido de transmisión y un cuarto de aceite.
b) El evento $\overline{E}$.
c) El evento $(E \cup F)$.
d) El evento $(E \cap F)$.

Solución: La figura 5.4 ilustra un espacio muestral que contiene los eventos $\overline{E}$, $(E \cup F)$ y $(E \cap F)$. Note que S tiene 18 resultados, el evento E 12 resultados y el F 6.

FIGURA 5.4

Espacio muestral y eventos para la aplicación 5.7



a) $S = \{XA, XB, XC, XD, XE, XF, YA, YB, YC, YD, YE, YF, ZA, ZB, ZC, ZD, ZE, ZF\}$.
b) $\overline{E} = \{YA, YB, YC, YD, YE, YF\}$.
c) $(E \cup F) = \{XA, XB, XC, XD, XE, XF, YA, YC, ZA, ZB, ZC, ZD, ZE, ZF\}$.
d) $(E \cap F) = \{XA, ZA, XC, ZC\}$. ■

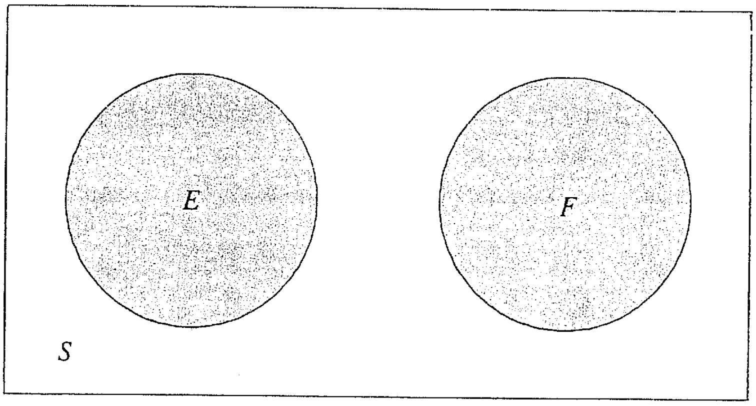
Eventos mutuamente excluyentes

Si E y F son eventos que no tienen resultados en común, entonces se denominan **eventos mutuamente excluyentes**. Esto es:

> los eventos E y F son mutuamente excluyentes si $(E \cap F) = \emptyset$.

Los eventos mutuamente excluyentes pueden ilustrarse por un diagrama de Venn como en la figura 5.5.

FIGURA 5.5
Eventos mutuamente excluyentes



APLICACIÓN 5.8

Suponga que un experimento consiste en lanzar un peso y a continuación una moneda de veinte centavos. ¿Son mutuamente excluyentes los pares de los eventos siguientes?

a) $E =$ dos "soles", $F =$ dos "águilas"
b) $E = \{SA, AA\}$, $F = \{SA, AS\}$
c) $E \neq \emptyset$, $F = \{AA\}$
d) $E, \overline{E}$

Solución:

a) Sí, porque no tienen resultados en común.
b) No, porque tienen en común el resultado SA.
c) Sí, porque no tienen resultados en común.
d) Sí, porque no tienen resultados en común. ■

APLICACIÓN 5.9 Suponga que una ciudad tiene tres distribuidoras de automóviles: Ford, GM y Chrysler; el distribuidor GM vende las marcas Pontiac, Oldsmobile y Cadillac; el distribuidor Ford vende Ford y Mercury; y el de Chrysler vende Dodge, Plymouth y Chrysler. El experimento consiste en observar el orden en que se venden los siguientes dos coches en la ciudad; si A es el evento de que los dos coches sean productos Ford, B es el evento de que los siguientes dos coches sean productos GM y C que los subsiguientes dos coches vendidos sean un Pontiac y un Dodge; encuentre el conjunto de resultados que conforman cada uno de los eventos:

a) A
b) B
c) C
d) $A \cap B$
e) $A \cup C$

¿Son mutuamente excluyentes los eventos A y B?

Solución: Suponga que F indica que el siguiente coche vendido es un Ford, M que es un Mercury, O que es un Oldsmobile, P que es un Pontiac, C que es un Cadillac y D que es un Dodge.

a) $A = \{\text{FM, MF}\}$.
b) $B = \{\text{PO, OP, PC, CP, CO, OC}\}$.
c) $C = \{\text{PD, DP}\}$.
d) $A \cap B = \varnothing$.
e) $A \cup C = \{\text{FM, MF, PD, DP}\}$.

Los eventos A y B son eventos mutuamente excluyentes, porque $\{A \cap B\} = \varnothing$. ■

GRUPO DE EJERCICIOS 5.1

Habilidades básicas

1. Liste un espacio muestral para cada uno de los siguientes experimentos:
   a) Lanzar una moneda de cinco centavos y otra de un peso, en ese orden, y registrar la forma en que caen.
   b) Lanzar un dado y una moneda de diez centavos, en ese orden, y registrar la forma en que caen.
   c) Seleccionar al azar a un estudiante universitario hombre y preguntarle si posee un coche.

2. Anote un espacio muestral para cada uno de los siguientes experimentos:
   a) Escoger dos monedas entre una de diez centavos, otra de veinte y otra más de cinco, y registrarlas.
   b) Seleccionar a dos adultos y preguntarles si están casados.
   c) Lanzar un dado hasta que muestre un 1.

3. Describa al menos dos espacios muestrales diferentes para el experimento de elegir un grupo de dos estudiantes de una clase de cinco y registre los resultados.

4. Represente al menos dos espacios muestrales diferentes para el experimento de seleccionar una carta de un conjunto de diez, numeradas del 1 al 10, y registre los resultados.

5. Sea C el evento de que el clima de mañana será caluroso, y D que mañana llueva. Describa los eventos compuestos siguientes:
   a) $C \cup D$ b) $C \cap D$ c) $\overline{C} \cap D$ d) $\overline{C} \cup \overline{D}$

6. Use los datos del inciso anterior para describir
   a) $\overline{C} \cup D$ b) $\overline{C} \cap \overline{D}$ c) $\overline{C \cup D}$ d) $\overline{C \cap D}$

7. Para el espacio muestral de 36 resultados posibles cuando se lanzan dos dados (véase la aplicación 5.3), encuentre el número de resultados para los cuales:
   a) ambos dados muestran un número par.
   b) exactamente un dado muestra un número par.
   c) a lo más un dado muestra un número par.

8. Para los 36 resultados posibles cuando se lanzan dos dados (véase la aplicación 5.3), identifique los resultados en que:
   a) ningún dado muestra un número par.
   b) la suma de los números mostrados es par.
   c) el cociente del número del dado rojo, dividido entre el número del dado negro es un número entero.

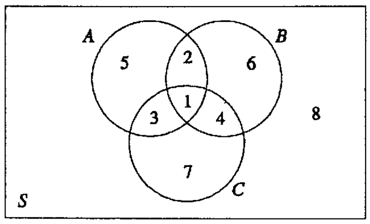
9. Se lanza un peso, una moneda de cinco centavos y otra de diez, en ese orden.
   a) Elabore un espacio muestral de ocho resultados.
   b) Anote los resultados del evento consistente en que la moneda de diez centavos caiga "sol".
   c) Cuente el número de resultados para los cuales la moneda de diez centavos muestra "sol".
   d) Liste los resultados para el evento en que una de las monedas de cinco o diez centavos muestra "sol".
   e) Cuente el número de resultados en los que una de las monedas de cinco o diez centavos muestra "sol".
   f) ¿En cuántos resultados una de las monedas de cinco y diez centavos muestra "sol", pero no las dos a la vez?
   g) Diga el número de resultados para los cuales el peso y la moneda de cinco concuerdan mostrando "sol" o "águila".
   h) ¿En qué resultados el peso concuerda con la moneda de cinco, pero no con la de diez?

10. Una moneda de diez centavos se lanza cuatro veces.
    a) Liste un espacio muestral de 16 resultados.
    b) Cuente el número de resultados para los cuales hay exactamente tres "soles".
    c) Anote el número de resultados para los cuales los cuatro primeros lanzamientos caen "sol".
    d) ¿Cuántos resultados dan dos "soles" y dos "águilas"?
    e) Diga el número de resultados en que todas las monedas caen águila o todas caen "sol".
    f) ¿Cuántas veces caen tres "soles" y un "águila"?

11. Refiérase al experimento de lanzar dos dados de la aplicación 5.3 y considere el espacio muestral de 36 resultados. Ilustre tres pares distintos de eventos mutuamente excluyentes.

12. Con los mismos datos de la aplicación 5.3 y considerando el espacio muestral obtenido, ilustre tres pares distintos de eventos que no sean mutuamente excluyentes.

Use el diagrama de Venn adjunto que contiene eventos A, B y C para los ejercicios 13 y 14. Sea $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $A = \{1, 2, 3, 5\}$, $B = \{1, 2, 4, 6\}$ y $C = \{1, 3, 4, 7\}$.



13. Liste los números que conforman los siguientes eventos:
    a) $A \cap B$ b) $A \cap B \cap C$
    c) $A \cup B$ d) $(\overline{A} \cup B) \cap C$
    e) $\overline{A} \cap B$ f) $\overline{A}$

14. Igual que el anterior.
    a) $\overline{A \cup B \cup C}$ b) $(A \cap B) \cup (\overline{C} \cap A)$
    c) $\overline{A} \cap C$ d) $\overline{A \cup B}$
    e) $(A \cap B) \cup C$ f) $A \cup (B \cap C)$

15. ¿Cuáles de los siguientes pares de eventos son mutuamente excluyentes?
    a) E = La señora Smith dio a luz gemelos.
       F = Una madre dio a luz una niña.
    b) E = Enrique falló en el último examen de estadística.
       F = Enrique aprobó el curso.
    c) E = Obtener un "sol" y un "águila" en dos lanzamientos de una moneda legal o balanceada.
       F = Obtener dos soles o dos "águilas" de una moneda legal, es decir, balanceada.

16. ¿Qué eventos son mutuamente excluyentes?
    a) E = José va al cine.
       F = José come una barra de dulce.
    b) E = El señor Doe llena la forma 1040 EZ de impuestos.
       F = El señor Doe llena la forma larga 1040 de impuestos.
    c) E = Nuestro equipo perdió el último juego de beisbol.
       F = Nuestro equipo perdió el último torneo de beisbol.

Más aplicaciones

17. Un experimento consiste en preguntar al azar a tres consumidores si compraron la marca A de mantequilla de cacahuate. Use S para sí y N para no; por ejemplo, SSN denota el evento simple de que las dos primeras personas compraron la marca A y que la tercera no lo hizo. Liste un espacio muestral para este experimento.

18. Refiérase al experimento del ejercicio 17 y sean E el evento de que al menos dos personas digan sí y F el que la primera persona encuestada diga no. Anote los eventos simples que conforman los siguientes:
   a) E b) F
   c) $\overline{E}$ d) $E \cup F$
   e) $E \cap F$ f) $E \cap \overline{F}$

Un paso más allá

19. Si un espacio muestral contiene n resultados, ¿cuántos eventos posibles hay?

20. Si se lanza una moneda n veces, ¿cuál es el espacio muestral si se registran todos los "soles" y todas las "águilas"?

21. Use diagramas de Venn para ilustrar cada uno de los eventos compuestos del ejercicio 5.

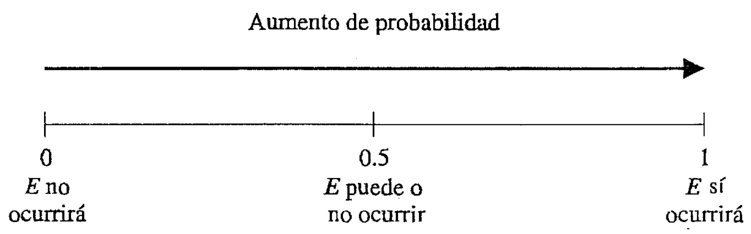
22. Lo mismo que el anterior pero del ejercicio 6.

SECCIÓN 5.2 *Concepto de probabilidad*

Es difícil que pase un día sin encontrarnos con algún aspecto de la probabilidad o del azar; las predicciones del clima anuncian un 80% de posibilidades de lluvia; según un cronista deportivo, los Gigantes de San Francisco tienen la mitad de las posibilidades de ganar la Serie Mundial; usted tiene buenas oportunidades de pasar el curso de Cálculo I si estudia seriamente.

La probabilidad de un evento es un número entre 0 y 1, inclusive, que se asocia al evento; si E es un evento, entonces $P(E)$ denota la probabilidad de E, que es la posibilidad del suceso del evento E. Si la probabilidad es 0, entonces el evento no ocurre, si es 1, el evento ocurre; mientras más cercano a 1 sea $P(E)$, más posibilidad hay de que ocurra, y mientras más cercano a cero sea $P(E)$, menos probable es que suceda, como lo muestra la figura 5.6.

FIGURA 5.6
Interpretación de $P(E)$



La probabilidad satisface las siguientes propiedades:

1. $P(E_i) \geq 0$
2. $P(E_i) \leq 1$
3. $\Sigma\, P(E_i) = 1$

donde $\Sigma P(E_i)$ es la suma de las probabilidades para todos los resultados (eventos simples) en el espacio muestral.

La probabilidad de un evento A se define como la suma de las probabilidades para los resultados contenidos en A. La aplicación 5.10 muestra cómo asignar probabilidades a eventos una vez que se conocen las de los resultados en el espacio muestral.

APLICACIÓN 5.10

Suponga que se lanza un dado una vez y la probabilidad de cualquier cara de quedar hacia arriba es 1/6; si E es el evento de sacar un número par y F el de uno impar, encuentre:

a) $P(E)$.
b) $P(F)$.
c) $P(E \cup F)$.
d) $P(E \cap F)$.

Solución: El espacio muestral es $S = \{1, 2, 3, 4, 5, 6\}$, el evento E es $\{2, 4, 6\}$ y el evento $F = \{1, 3, 5\}$. Por tanto, tenemos:

a) $P(E) = P(2) + P(4) + P(6) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$
b) $P(F) = P(1) + P(3) + P(5) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$
c) $P(E \cup F) = P(S) = 1$
d) $P(E \cap F) = 0$, puesto que $E \cap F = \emptyset$ y $P(\emptyset) = 0$. ■

Cómo asignar probabilidades a eventos

Las tres propiedades que satisfacen las probabilidades no nos dicen cómo asignarlas a los resultados en un espacio muestral; lo que hacen es normar ciertas asignaciones que no son consistentes con nuestras nociones intuitivas. Hay dos métodos generales para asignar probabilidades a eventos: el método objetivo y el método subjetivo. El **método objetivo** implica asignar probabilidades a eventos con base en el conteo o en la experimentación repetida; el **método subjetivo**, por otro lado, nos permite asignar probabilidades con fundamento en la intuición o en la creencia personal; cuando se utiliza el método subjetivo, dos conocedores pueden no concordar en sus asignaciones.

Métodos objetivos

EJEMPLO 5.7

Se lanza una moneda, los resultados serán cara (C) y águila (A). ¿Qué número P(C) le asignaríamos a C y cuál P(A) le asignaríamos a A? Suponga que le asignamos 0.7 a C y 0.3 a A; ¿es ésta una asignación válida de probabilidades? La respuesta es sí, con base en las propiedades anteriores porque: a) ambos números son mayores que cero; b) ambos números son menores que 1, y c) la suma de 0.7 y 0.3 es 1; pero estas asignaciones van en contra de nuestra intuición si sabemos que la moneda es legal y muchos de nosotros convendríamos en que lo correcto sería 0.5 para C y 0.5 para A. Si lanzamos una moneda un gran número N de veces y encontramos la frecuencia f para el suceso de una cara, esperaríamos que la **frecuencia relativa** f/N para la ocurrencia de una cara sea cercana a 0.5. La tabla 5.1 contiene el número de caras obtenidas en cada caso. Como C y A son los únicos resultados del experimento, no cuenta que la moneda caiga sobre el canto; entonces P(A) debe ser 0.5 pues así $0.5 + 0.5 = 1$.

TABLA 5.1
Frecuencia y frecuencia relativa de caras en el lanzamiento

Núm. lanzamientos N	Núm. "soles" f	Frecuencia relativa f/N
1	0	0.000
2	1	0.500
3	1	0.333
4	1	0.250
5	3	0.600
6	3	0.500
7	4	0.571
8	3	0.375
9	6	0.667
10	4	0.400
.	.	.
.	.	.
.	.	.
100	51	0.510
1000	447	0.447
10000	5047	0.505

El valor límite de la frecuencia relativa f/N de la obtención de un "sol" cuando una moneda legal se lanza N veces, se aproximará a 0.5 cuando N sea grande, como lo ilustra la figura 5.7.

Como 0.7 no es igual al valor límite de la frecuencia relativa del suceso de un "sol", que es 0.5, no recomendaríamos asignar 0.7 como probabilidad del resultado C. Con el valor 0.5 asignado a C, ¿qué probabilidad P(A) debe asignársele a A? Como consecuencia de la tercera propiedad, sabemos que:

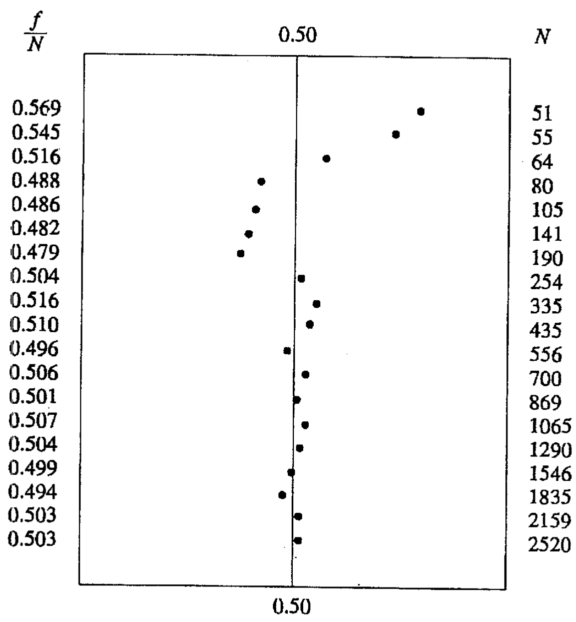
$$P(\mathrm{C}) + P(\mathrm{A}) = 1$$

o

$$0.5 + P(\mathrm{A}) = 1$$

Por tanto, $P(\mathrm{A}) = 0.5$ debe ser la probabilidad asignada a la obtención de un "águila".

FIGURA 5.7
Valores límites de f/N cuando N se vuelve grande



N = número de veces que una moneda es lanzada

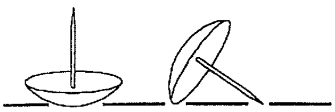
Si un experimento es repetible podemos asignar probabilidades a los resultados de acuerdo con los límites de las frecuencias relativas, en forma análoga a como lo hicimos antes para la moneda; el único problema de hacer esto es que no siempre se conocen los valores límite de las frecuencias relativas; para usar este método, necesitamos tener disponibles muchos datos repetitivos y aún entonces sólo se pueden encontrar aproximaciones de los límites de las frecuencias relativas.

Probabilidad empírica

Cuando la probabilidad se basa en nuestra experiencia y se desconocen los valores exactos de los límites de las frecuencias relativas, deben usarse aproximaciones a estos valores límite; al hacerlo, al método objetivo de asignación de probabilidades se le llama empírico. Según el método de **probabilidad empírica**, si E es un evento, $P(E)$ es aproximadamente igual a f/N, donde f es el número de resultados favorables y N el de repeticiones del experimento. En este caso, tenemos:

$$P(E) \approx \frac{f}{N}$$

EJEMPLO 5.8



Consideremos el experimento de bailar una pirinola que puede detenerse de dos maneras.

¿Cómo podemos determinar la probabilidad de que la pirinola se detenga hacia arriba? Podríamos preguntarle a alguien, pero, ¿alguien sabría la respuesta? Quizá nadie; podemos bailar la pirinola diez veces y anotar el número de veces que queda hacia arriba. Esa frecuencia relativa podría servir como una estimación de la probabilidad deseada; más aún, para tener una estimación mejor, podríamos bailar la pirinola 100, 1000 veces o más y anotar la frecuencia de que apunte hacia arriba, de esta manera obtendríamos una buena aproximación de la probabilidad de que la pirinola apunte hacia arriba.

EJEMPLO 5.9

Según cambia nuestra experiencia, cambia la frecuencia relativa. Por ejemplo, si lanzamos una moneda seis veces y obtenemos tres caras, estimaríamos que la posibilidad de obtener una cara es 3/6 = 0.5; si la moneda fuera lanzada una vez más y mostrara una cara, entonces la posibilidad estimada de obtener una cara sería 4/7 = 0.5714, o si fuese cruz la posibilidad estimada de obtener cara sería 3/7 = 0.4286; este cambio en la frecuencia relativa refleja el cambio de nuestro conocimiento. Como lo señalamos antes, cuando la repetición aumenta, la frecuencia relativa cambia muy poco y el valor límite al que se aproxima se le denomina **probabilidad**.

Si usted tiene acceso a un programa de computadora como MINITAB, puede simular el lanzamiento de una moneda tantas veces como quiera; a esto se le llama **simulación por computadora**. Las primeras dos líneas en la pantalla 5.1 usan MINITAB para hacer que la computadora simule un proceso aleatorio de 150 lanzamientos de una moneda legal, y exhiba los resultados en una columna etiquetada con C1; entonces los resultados de los lanzamientos simulados se imprimen, seguidos de un recuento de todos ellos. Los resultados de los primeros 15 lanzamientos se muestran en la primera columna de unos y ceros; cada 1 representa un "sol" y cada 0 un "águila"; si usamos sólo los primeros 15 lanzamientos de la moneda para estimar la probabilidad de obtener un "sol" en un solo lanzamiento, la $P(\text{sol})$ estimada sería $7/15 \approx 0.467$. Si leemos de la cuenta dada al final de la pantalla, vemos que la simulación produjo 69 soles y 81 "águilas" basados en 150 lanzamientos, $P(\text{sol}) \approx 69/150 = 0.46$. Otra simulación produciría, en general, resultados diferentes.

Pantalla 5.1

```
MTB> RANDOM 150 C1;
SUBC> BERNOULLI P = 0.5.
MTB> PRINT C1

C1
 1  0  0  1  0  0  1  0  1  0  1  0  1  1  0
 0  1  0  0  1  0  0  1  0  1  1  0  1  0  0
 1  1  1  0  0  0  1  0  1  0  1  0  0  1  0
 1  1  0  1  0  1  0  0  0  1  1  1  1  1  1
 0  0  0  0  0  0  0  1  0  0  1  1  1  0  0
 1  0  1  1  1  0  1  0  1  0  0  0  1  0  1
 1  1  1  1  0  0  0  0  0  0  0  0  0  1  1
 1  1  0  0  0  0  0  1  0  1  1  0  0  0  1
 1  0  1  0  1  0  1  1  0  0  1  0  0  0  0
 0  1  0  1  1  1  0  1  1  1  1  0  1  0  0

MTB>TALLY C1;
SUBC> ALL.
  C1   COUNT   CUMCNT   PERCENT   CUMPCT
  0     81       81      54.00     54.00
  1     69      150      46.00    100.00
N = 150

MTB>
```

APLICACIÓN 5.11

Una compañía de seguros quiere estimar la probabilidad de que le ocurra un accidente a un coche de la policía en una cierta ciudad durante un periodo de un mes; el último mes, 7 de 20 coches de la policía tuvieron accidentes.

a) ¿Cuál sería la estimación dada por usted de la probabilidad deseada?
b) ¿Cuál es la probabilidad de que un coche de la policía no participe en un accidente?

Solución:

a) $7/20 = 0.24$
b) $0.65 = 1 - 0.35$. ■

APLICACIÓN 5.12

Las calificaciones del SAT en matemáticas de los estudiantes de una gran universidad se muestran en la tabla siguiente de frecuencias agrupadas.

SAT	*f*
200-299	3,600
300-399	11,900
400-499	12,000
500-599	5,500
600-699	1,500
700-799	500

Si se elige a un estudiante al azar, ¿cuál es la probabilidad de que la calificación del estudiante en matemáticas en el SAT

a) exceda de 399?
b) sea a lo más 599?
c) esté entre 600 y 699, inclusive?
d) no esté entre 400 y 499, inclusive?
e) sea menor o igual que 699?

Solución: Primero construimos una tabla de frecuencias relativas agrupadas, tabla 5.2. Los cálculos se aproximaron hasta la tercera cifra decimal; recuerde que la frecuencia relativa de una clase se encuentra dividiendo la frecuencia f de la clase entre el número total de medidas N.

Tabla 5.2
Tabla de frecuencias relativas agrupada para la aplicación 5.12

SAT	*f*	Rel *F*
200-299	3,600	0.103
300-399	11,900	0.340
400-499	12,000	0.343
500-599	5,500	0.157
600-699	1,500	0.043
700-799	500	0.014
	35,000	

a) $P(\text{SAT} > 399) = 0.343 + 0.157 + 0.043 + 0.014 = 0.557$
b) $P(\text{SAT} \leq 599) = 0.103 + 0.340 + 0.343 + 0.157 = 0.943$
c) $P(600 \leq \text{SAT} \leq 699) = 0.043$.
d) $P(\text{SAT} < 400$ o $\text{SAT} > 499) = 1 - 0.343 = 0.657$ (pues la suma de las frecuencias relativas es 1).
e) $P(\text{SAT} \leq 699) = 1 - 0.014 = 0.986$ (pues la suma de las frecuencias relativas es 1). ∎

APLICACIÓN 5.13

En una pequeña ciudad, se clasificó a cada persona de acuerdo con su religión y su afiliación a una partido político. Los resultados se resumen en la siguiente tabla:

Religión	Partido político: Demócrata	Republicano	Independiente
Protestante	10,000	8,000	2,000
Judío	5,500	6,000	500
Católico	8,500	9,500	1,500

Si se escoge al azar a una persona de la ciudad, ¿cuál es la probabilidad de que la persona sea

a) republicana?
b) católica?
c) protestante y republicana?
d) católica e independiente?

Solución: Primero encontramos el total de cada renglón y columna, como se indica en la tabla 5.3.

TABLA 5.3
Totales por columna y renglón para la aplicación 5.13

Religión	Partido político: Demócrata	Republicano	Independiente	Total
Protestante	10,000	8,000	2,000	20,000
Judío	5,500	6,000	500	12,000
Católico	8,500	9,500	1,500	19,500
Total	24,00	23,500	4,000	51,500

a) Hay 23,500 republicanos del total de 51,500 personas. Por lo tanto, $P(\text{R}) = 23{,}500/51{,}500 = 0.456$.
b) Hay 19,500 católicos de ese total. Por tanto, $P(\text{C}) = 19{,}500/51{,}500 = 0.379$.
c) Hay 8,000 personas del mismo total que son protestantes y republicanos. Por tanto, $P(P \text{ y } R) = 8{,}000/51{,}500 = 0.155$.
d) Hay 1,500 personas de las 51,500 que son católicos e independientes. Por tanto, $P(\text{C e I}) = 1{,}500/51{,}500 = 0.029$. ■

Probabilidad clásica

Si en un experimento todos los resultados tienen la misma probabilidad de ocurrir, se dice que son **resultados igualmente posibles**. Si un experimento tiene n resultados que creemos igualmente posibles, podemos asignar a cada uno en el espacio muestral S, un valor de probabilidad de $1/n$; ésta es una consecuencia de la propiedad 3 de la probabilidad, que establece que la suma de las probabilidades para un espacio muestral debe ser 1; entonces si E es un evento que contiene f resultados de un espacio muestral constituido por n resultados igualmente posibles, la probabilidad de que ocurra E es simplemente f/n. En consecuencia, tenemos el hecho básico siguiente:

Si S es un espacio muestral de resultados igualmente posibles y E es un evento, entonces:

$$P(E) = \frac{f}{n} \tag{5.1}$$

donde f es el número de resultados en E y n es el número de resultados en S. Se acostumbra llamar **resultados favorables** a los resultados contenidos en E. El método objetivo de asignación de probabilidades usando espacios muestrales de resultados igualmente posibles, se denomina **método probabilístico clásico.**

EJEMPLO 5.10

Suponga que se lanzan juntos una vez, un peso y una moneda de diez centavos; pueden caer de cuatro maneras. Un espacio muestral para este experimento es $S = \{SS, AS, SA, AA\}$. La primera letra de cada par corresponde al resultado del peso y la segunda corresponde al resultado de la moneda de diez centavos. Los resultados posibles pueden visualizarse como sigue:

		2ª moneda	
		S	A
Primera	S	SS	SA
moneda	A	AS	AA

Resultados posibles

Los cuatro resultados tienen la misma posibilidad de ocurrir si ambas monedas se lanzan de modo que el resultado de una no influya en, ni sea influenciado por el resultado de la otra; en consecuencia, asignamos un valor de 1/4 a la probabilidad de cada uno de los cuatro resultados. Si E es el evento $\{AS, SA\}$, entonces tenemos los dos métodos siguientes de asignar un valor a la probabilidad de E.

1. Sumar las probabilidades de los resultados contenidos en E. Por tanto, $P(E) = 1/4 + 1/4 = 0.5$.

2. Usar la fórmula (5.1). El número de resultados favorables es $f = 2$, el número de resultados en E y el número de resultados igualmente posibles en S es $n = 4$. En consecuencia,
$$P(E) = \frac{f}{n} = \frac{2}{4} = \frac{1}{2}$$

APLICACIÓN 5.14

Considere el experimento previo de lanzar un peso y una moneda de diez centavos. Cuando se lanzan dos monedas, podemos registrar la aparición de 0, 1 o 2 "soles"; así, el espacio muestral es $S = \{0S, 1S, 2S\}$. Como hay tres resultados en S podemos estar tentados a concluir que $P(1S) = 1/3$, pero por el argumento anterior, $P(1S) = P(\{SA, AS\}) = 1/2$. Resuelva esta contradicción aparente.

Solución: Los tres resultados en S no tienen la misma posibilidad de ocurrir. El resultado 1S tiene el doble de posibilidades de ocurrir que 0S o 2S; por tanto, la fórmula (5.1) no es válida para calcular $P(1S)$ y tenemos $P(1S) = 1/2$. ■

APLICACIÓN 5.15

Para el experimento de lanzar dos dados en la aplicación 5.3, encuentre la probabilidad de que:

a) Ambos muestren un número par.

b) Los números tengan una suma igual a 7.
c) Los números tengan una suma igual a 7 u 11.
d) Ambos muestren un número primo (un número primo es un número cuyos únicos divisores son 1 y él mismo).
e) El dado rojo muestre un 2.
f) Los números mostrados tengan una suma igual a 13.

Solución: Un espacio muestral es como sigue:

$$\begin{matrix} \{(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,5) & (5,6) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)\} \end{matrix}$$

Hay 36 resultados igualmente probables en S.

a) Representemos por E_1 el evento de que ambos dados muestren un número par; un resultado favorable es aquel en que ambos dados muestran un número par: (2, 2), (2, 4), (2, 6), (4, 2), (4, 4), (4, 6), (6, 2), (6, 4) y (6, 6); hay nueve resultados favorables, en consecuencia:

$$P(E_1) = \frac{f}{n} = \frac{9}{36} = \frac{1}{4}$$

b) Llamemos E_2 al evento de que la suma de los dos números mostrados sea 7; el conjunto de resultados favorables es $E_2 = \{(3, 4), (4, 3), (5, 2), (2, 5), (6, 1), (1, 6)\}$. Por tanto,

$$P(E_2) = \frac{f}{n} = \frac{6}{36} = \frac{1}{6}$$

c) Señalemos por E_3 que la suma de los números mostrados sea 7 u 11. El conjunto de resultados favorables es $E_3 = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1), (5, 6) \text{ y } (6, 5)\}$. De este modo:

$$P(E_3) = \frac{f}{n} = \frac{8}{36} = \frac{2}{9}$$

d) E_4 será el evento de que ambos dados muestren un número primo. Como $E_4 = \{(1,1),(1,2),(1,3),(1,5),(2,1),(2,2),(2,3),(2,5),(3,1),(3,2),(3,3),(3,5),(5,1),(5,2),(5,3),(5,5)\}$,

$$P(E_4) = \frac{f}{n} = \frac{16}{36}$$

e) El evento E_5 resulta de que el dado rojo muestre un 2. El conjunto de resultados favorables es: $E_5 = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\}$. Entonces,

$$P(E_5) = \frac{f}{n} = \frac{6}{36} = \frac{1}{6}$$

f) Representemos por E_6 el evento de que la suma de los números mostrados sea 13. Como una suma de 13 es imposible, $E_6 = \emptyset$. Así,

$$P(E_6) = 0 \quad ■$$

El experimento de los dados de la aplicación 5.15 puede simularse usando MINITAB; se lanzan dos dados 1000 veces; los resultados del primer dado se guardan en C1 y los del segundo en C2; la suma de cada par se guarda en C3 y los resultados se pueden ver a lo largo de la pantalla 5.2 junto con las órdenes.

Pantalla 5.2

```
MTB> RANDOM 1000 C1 C2;
SUBC> INTEGER 1 6.
MTB>RSUM C1 C2 C3
MTB> TALLY C3;
SUBC> PERCENT.

      C3    PERCENT
       2       2.40
       3       5.70
       4      10.00
       5      11.70
       6      13.60
       7      16.90
       8      14.60
       9       8.70
      10       8.70
      11       5.40
      12       2.30

MTB>
```

Note que la probabilidad de que se muestren números cuya suma sea 7 es aproximadamente 0.169 y la probabilidad de que la suma de los números mostrados sea 7 u 11 es 0.169 + 0.054 = 0.223, aproximadamente. Estos dos resultados se comparan favorablemente con los que vimos en la aplicación 5.15.

APLICACIÓN 5.16

¿Cuál es la probabilidad de adivinar correctamente todas las respuestas de un examen tipo verdadero/falso, que consta de cuatro preguntas?

Solución: Denotemos verdadero por V y falso por F. Como sólo uno de estos resultados representa un acierto, $f = 1$. Un espacio muestral de resultados igualmente probables consta de 16 respuestas:

VVVV	VFFV	FFVF	FVFV
VVFV	VFFF	FFFV	FVVF
VVVF	VFVV	FFVV	FVFF
VVFF	VFVF	FFFF	FVVV

VVVV corresponde a responder correctamente todas las preguntas y consta de sólo un resultado de los 16 listados. Usando la fórmula (5.1), tenemos:

$$P(\text{VVVV}) = \frac{f}{n} = \frac{1}{16} \quad \blacksquare$$

Método subjetivo

En muchas situaciones tenemos poca o ninguna información sobre los resultados de un experimento, no tendremos datos disponibles de frecuencia si la situación no ha ocurrido antes. Un médico que trata a un paciente de una enfermedad rara, debe elaborar un diagnóstico basándose en su experiencia con el paciente y en el registro médico completo del enfermo; un candidato que aspira a un puesto público por primera vez, estimará sus posibilidades de triunfo apoyado en su intuición y en los comentarios que escuche. Cuando se asigna probabilidad a eventos con base en la intuición y consideraciones personales, el método de asignación se llama *subjetivo*; desde luego, cuando se usa ese método las asignaciones dependerán de la forma individual de hacerlos y pueden variar ampliamente de un individuo a otro.

Histogramas de probabilidad

Para los datos exhibidos en una tabla de frecuencias agrupadas, se puede construir un tipo especial de histograma de frecuencias relativas, y usarlo para estimar la probabilidad de que una medida caiga en una clase particular. Si en lugar de que la i-ésima barra de un histograma de frecuencias relativas tenga una altura de f_i/N, como se hizo en la sección 2.3, permitimos que sea:

$$h = \frac{f_i}{wN}$$

donde f_i es la frecuencia asociada con la i-ésima clase, w el ancho de clase y N el número total de medidas, entonces el área de la barra aproxima la probabilidad de que una medida caiga en esa clase particular. Además, el área total del histograma es entonces igual a 1, la suma de las probabilidades; por esta razón lo llamamos **histograma de probabilidad**. Advierta que el área de un histograma de frecuencias relativas, como el construido en la sección 2.3, tiene un área igual a w, el ancho de clase.

APLICACIÓN 5.17

Construya un histograma de probabilidad para los datos adjuntos de la aplicación 5.12.

SAT	f
200-299	3,600
300-399	11,900
400-499	12,000
500-599	5,500
600-699	1,500
700-799	500

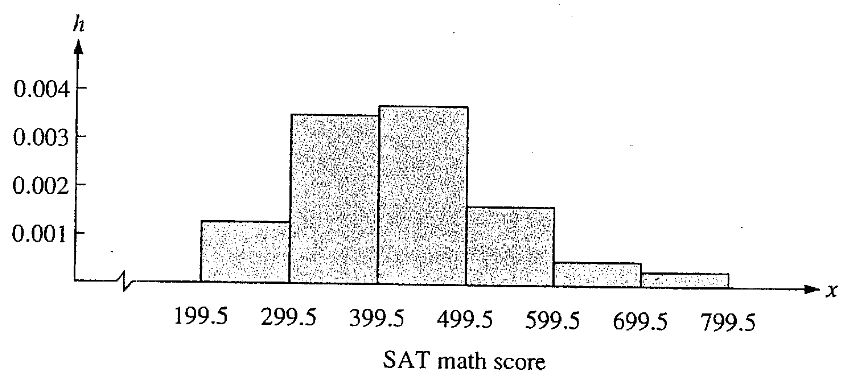
Solución: La tabla 5.4 contiene las fronteras de clase, las alturas de las barras y sus áreas.

TABLA 5.4
Valores para el histograma de probabilidad para la aplicación 5.17

Frontera	h	Área de la barra
199.5–299.5	0.00103	0.103
299.5–399.5	0.00340	0.340
399.5–499.5	0.00343	0.343
499.5–599.5	0.00157	0.157
599.5–699.5	0.00043	0.043
699.5–799.5	0.00014	0.014
		1.000

Las alturas h se encontraron dividiendo las frecuencias relativas entre $w = 100$, dando la medida $h = f_i/(wN)$. Por ejemplo, la frecuencia relativa de la primera clase en la tabla 5.4 es $f_i/N = 3{,}600/35{,}000 = 0.103$. Si dividimos 0.103 entre $w = 100$, obtendremos $(0.103)/(100) = 0.00103$. La gráfica del histograma de probabilidad se muestra en la figura 5.8; la suma de las áreas de las dos barras más altas representan la probabilidad $P(299.5 \le x \le 499.5) = 0.683$. ■

FIGURA 5.8
Histograma de probabilidad para la aplicación 5.17



Oportunidades matemáticas

En el caso de apuestas, como carreras de caballos y pronósticos deportivos, a menudo se determina la posibilidad del suceso de un evento usando probabilidad subjetiva, y se establece comúnmente en términos de **oportunidades**. Si E es un evento, las oportunidades a favor de E, denotadas por Ops(E), se definen así:

Oportunidades a favor de E

$$\text{Ops}(E) = \frac{P(E)}{P(\overline{E})}$$

Las *oportunidades en contra de E*, Ops $(\overline{E})$, se definen:

Oportunidades en contra de E

$$\text{Ops}(\overline{E}) = \frac{P(\overline{E})}{P(E)} = \frac{1}{\text{Ops}(E)}$$

EJEMPLO 5.11

Si E es el evento de que un caballo favorito gane la carrera y $P(E) = 1/3$, entonces las oportunidades a favor de ganar, Ops(E), son:

$$\text{Ops}(E) = \frac{P(E)}{P(\overline{E})} = \frac{1/3}{2/3} = \frac{1}{2}$$

Este resultado se escribe frecuentemente como 1:2.

Las oportunidades se pueden convertir en probabilidades. Como regla general tenemos:

Si Ops(E) = $a{:}b$, entonces

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

EJEMPLO 5.12

Si las oportunidades en favor de que los Piratas de Pittsburgh ganen el juego de beisbol son de 2:3, entonces su probabilidad de ganar es:

$$\begin{aligned} P(E) &= \frac{a}{a+b} \\ &= \frac{2}{2+3} \\ &= \frac{2}{5} = 0.4 \end{aligned}$$

APLICACIÓN 5.18

Si las oportunidades son de 3:2 en contra de que el caballo favorito gane la carrera, ¿cuál es la probabilidad de que

a) pierda la carrera?
b) gane?

Solución: Sea E el evento de que el caballo gane la carrera; entonces, $\overline{E}$ es el evento de que la pierda.

a) $P(\overline{E}) = 3/(2+3) = 3/5$.
b) Como $P(E) + P(\overline{E}) = 1$, $P(E) = 1 - P(\overline{E})$, es decir,

$$P(E) = \frac{1-3}{5} = \frac{2}{5}$$

También podemos obtener el mismo resultado usando la regla de las oportunidades:

$$P(E) = \frac{2}{2+3} = \frac{2}{5} \quad \blacksquare$$

APLICACIÓN 5.19

Durante la guerra de Vietnam en 1967, aproximadamente 15 millones de hombres eran elegibles para el reclutamiento militar en cualquier mes; si cada mes se alistaron 20,000 hombres, ¿cuáles eran las oportunidades en contra de que Juan Pérez fuera reclutado en abril de ese año?

Solución: Sea E el evento de que Juan fuese reclutado. Entonces,

$$P(E) = \frac{20{,}000}{15{,}000{,}000} = \frac{1}{750} \text{ y } P(\overline{E}) = \frac{749}{750}$$

Por tanto, las oportunidades en contra de que Juan fuera llamado eran:

$$\text{Ops}(\overline{E}) = \frac{P(\overline{E})}{P(E)} = \frac{749/750}{1/750} = 749{:}1 \quad \blacksquare$$

GRUPO DE EJERCICIOS 5.2

Habilidades básicas

1. Asigne una de las probabilidades siguientes a una de las afirmaciones dadas:

   0 0.01 0.3 0.99 1

   a) El evento es imposible. Nunca podrá ocurrir.
   b) El evento es seguro.
   c) El evento es muy poco probable, pero ocurrirá alguna vez si hay gran número de ensayos.
   d) Son más las veces que ocurre el evento de las que no.

2. ¿Cuáles de los números siguientes no pueden ser la probabilidad de algún evento?

a) 0.74	d) 0.5	g) 2/3
b) −1	e) 0	h) 0.999...
c) 1.02	f) 1	i) 0.67

3. Cada una de las siguientes situaciones, exprésela usando un valor de probabilidad en forma decimal.
   a) Hay una posibilidad de 30-70 de conseguir fondos para nuestro proyecto.
   b) Las oportunidades en contra de encontrar petróleo son de 100 a 1.

4. Exprese cada una de las siguientes situaciones con un valor de probabilidad en forma decimal.
   a) Hay un 75% de posibilidad de que se necesite cirugía para corregir el problema.
   b) Las oportunidades de que los Gigantes ganen el juego de mañana son de 5 a 3.

Más aplicaciones

5. Un meteorólogo predice lluvia con una probabilidad de 70 %.

a) ¿Cuáles son las oportunidades a favor de que llueva?
b) ¿En contra de ella?

6. Un meteorólogo predice que las oportunidades de que llueva son de 7 a 13.
a) ¿Cuál es la probabilidad de que llueva?
b) ¿Y de que no llueva?

7. Una bolsa contiene tres canicas rojas, dos blancas y cinco azules; se elige una canica al azar. ¿Cuál es la probabilidad de que la canica resulte:
a) roja?
b) blanca? o
c) azul?

8. Una caja contiene tres canicas azules, cuatro amarillas y dos verdes; si se elige una canica al azar, ¿cuál es la probabilidad de sacar:
a) azul?
b) amarilla?
c) verde?

9. Refiérase al experimento del ejercicio 7. Liste dos espacios muestrales, uno que contenga resultados igualmente probables y otro que no; enumere en cada caso las probabilidades asociadas con los resultados.

10. Del ejercicio 8, ubique dos espacios muestrales para el experimento, uno con resultados igualmente probables y otro sin ellos. ¿Cuáles son las probabilidades asociadas con los resultados en cada caso?

11. Las oportunidades en contra de que nos repartan una tercia en un juego de póker de cinco cartas son de 49:1. ¿Cuál es la probabilidad de que nos toque una tercia?

12. Las oportunidades de que un alumno de sexto grado continúe y se gradúe de la preparatoria, son de tres a uno. ¿Cuál es la probabilidad para un alumno de sexto grado de que no se gradúe de la preparatoria?

13. Una ruleta estadounidense contiene compartimentos numerados del 1 al 36 más 0 y 00; de los 38 compartimentos 0 y 00 están pintados de verde, 18 compartimentos de rojo y 18 son negros; se lanza un balín en dirección contraria al movimiento de la rueda y se apuesta a cuál será el número en el que se detendrá el balín. Suponga que la ruleta es legal.
a) ¿Cuál es la probabilidad de que el resultado sea negro?
b) ¿Cuáles son las oportunidades en contra de un resultado rojo?
c) ¿Cuál es la probabilidad de un resultado rojo?
d) ¿Cuáles son las oportunidades a favor de un resultado rojo?
e) ¿Qué probabilidad existe de que el resultado sea un número impar?

14. Conteste con base en la información del ejercicio 13.
a) ¿Cuál es la probabilidad de un resultado rojo o negro?
b) ¿Cuál la de un 0?
c) ¿Cuáles son las oportunidades a favor de un resultado verde?
d) ¿Cuántas son las oportunidades en contra de un resultado verde?
e) ¿Qué probabilidad hay de que el balín señale un número par?

15. Para estimar el número de libros requeridos por curso para una universidad pequeña, el consejo de estudiantes consideró una muestra de 100 cursos y obtuvo los resultados siguientes:

Núm. de libros requeridos (x)	Núm. de cursos (f)
0	5
1	48
2	21
3	12
4	9
5	3
6	2

Use probabilidades empíricas para encontrar:
a) el valor de la probabilidad para cada x.
b) $\Sigma P(x)$.
c) $P(x$ es al menos 5).
d) $P(x$ es 2 o 4).

16. Se preguntó el nombre del refresco favorito de cada persona en una muestra de cuarenta. Las respuestas fueron:

Refresco	Núm. de preferencias
Pepsi-Cola	14
Coca-Cola	12
Sprite	8
Seven-Up	3
Dr. Pepper	2
Nehi Orange	1

Mediante probabilidades empíricas, ¿cuál es la probabilidad de que el refresco favorito de una persona sea:
a) Pepsi-Cola?
b) Coca-Cola?
c) Dr. Pepper?

d) Mr. Pibb?
e) Pepsi-Cola o Coca-Cola?

17. Los maestros de una universidad se clasificaron de acuerdo con el sexo y rango académico. Los resultados se muestran en la tabla siguiente:

	Rango			
Sexo	Instructor	Profesor Asistente	Profesor sociado	Profesor
Hombre	300	400	700	300
Mujer	350	450	300	200

Si de esta universidad se elige a un maestro al azar, ¿cuál es la probabilidad de que sea:
a) hombre?
b) mujer?
c) profesor?
d) profesor hombre?
e) profesor asistente mujer?

18. Un estudio sobre los empleados de una gran organización proporcionó la estadística siguiente según el sexo y el estado civil:

	Estado civil			
Sexo	Casado	Soltero	Divorciado	Viudo
Hombre	25%	11%	10%	3%
Mujer	30%	8%	7%	6%

Suponga que se elige a un empleado al azar. ¿Cuál es la probabilidad de que resulte:
a) casado?
b) mujer y divorciada?
c) viudo?
d) hombre?
e) casado o soltero?

19. Los salarios anuales promedio, en miles de dólares, para trabajadores de los 50 estados estadounidenses que cuentan con seguro de desempleo son:[33]

28.7	18.0	17.1	16.0	15.2
19.7	17.9	16.9	15.9	15.0
19.7	17.8	16.8	15.9	14.8
19.3	17.3	16.8	15.5	14.7
17.8	17.3	16.7	15.5	14.6
18.8	17.2	16.5	15.5	14.6
18.7	17.2	16.5	15.5	14.1
18.1	17.2	16.4	15.4	14.1
18.1	17.1	16.1	15.2	13.9
18.1	17.1	16.1	15.2	13.2

Construya un histograma de probabilidad con ocho barras.

Un paso más allá

20. Las huellas digitales se clasifican en ocho tipos genéricos y poseen múltiples características; una de ellas es el número de estrías, útil en el trabajo de investigación criminalística. La tabla de frecuencias agrupadas adjunta muestra el número de estrías para huellas digitales de 800 hombres.

Número de estrías	f
0-19	10
20-39	12
40-59	24
60-79	40
80-99	73
100-119	100
120-139	90
140-159	112
160-179	124
180-199	95
200-219	67
220-239	36
240-259	10
260-279	4
280-299	3
	800

a) Suponga que hay huellas en la escena de un crimen y se determina que el número de estrías es al menos 220; si un sospechoso tiene un número de estrías de 241, ¿qué se puede concluir? ¿Por qué?
b) ¿Cuál es la probabilidad de que un hombre tenga un número de estrías de al menos 200?
c) ¿Qué probabilidad existe de que un hombre tenga al menos 120 estrías y no más de 199?

21. Tire un par de dados 100 veces y registre la suma de los números que hayan quedado hacia arriba en cada tirada. ¿Cuál es la frecuencia relativa de una suma de 8? (Por lógica matemática, en un gran número de tiradas la frecuencia relativa aproximará la probabilidad de una suma de 8 a 0.14 si los dados son legales.)

22. Se lanzan dos dados y se registra el número mayor de los dos que quedan hacia arriba; si $\{1, 2, 3, 4, 5, 6\}$ es un espacio muestral para el experimento, encuentre la probabilidad para cada uno de los seis resultados. ¿Son igualmente probables los resultados? ¿Por qué sí o por qué no?

23. Si un paquete de 52 cartas se baraja bien, ¿cuál sería, según usted, la probabilidad de que entre las tres cartas de arriba haya un rey o una reina, o de que esas figuras estén juntas entre las 49 cartas restantes? Ponga a prueba su predicción simulando el experimento 20 veces.

24. ¿Puede un experimento tener dos espacios muestrales de resultados igualmente probables? Explique su respuesta.

25. Para decidir cuándo "tocar" en un juego de beisbol, un director técnico estudia los datos adjuntos obtenidos de varios cientos de juegos de beisbol.[34]

Bases ocupadas	Núm. de "outs"	Proporción de casos en que no se anota carrera en una entrada	Núm. promedio de carreras en una entrada	Núm. de casos observados
Primera	0	0.604	0.813	1728
Segunda	1	0.610	0.671	657
Segunda	0	0.381	1.194	294

a) Si un jugador está en primera base sin "outs" ¿en cuántos casos anota carrera?
b) Para la situación anterior encuentre la probabilidad de que se anote al menos una carrera.
c) Si un jugador está en primera base sin "outs" y se realiza un toque de sacrificio como la forma normal de hacer avanzar al jugador en primera, ¿es ésta la mejor situación que puede tenerse, en cuanto al número de carreras anotadas?
d) Si un jugador está en segunda base sin "outs", ¿qué probabilidad tiene de anotar al menos una carrera?
e) Si un jugador está en segunda base con un "out", ¿en cuántos casos no se anotará una carrera en esa entrada?

26. Cuando dos personas se sientan en una mesa cuadrada con una silla en cada uno de los lados, no es raro que se sienten en lados contiguos en lugar de ocupar asientos opuestos; si dos personas eligen asientos al azar en una mesa así, determine la probabilidad de que ocupen asientos contiguos en la mesa.

SECCIÓN 5.3 Conteo

Si un experimento contiene un gran número de resultados, puede ser difícil contarlos en un evento; para tales experimentos o problemas difíciles de cuenta, necesitamos utilizar técnicas especiales de conteo. Los ejemplos 5.13 y 5.14 ilustran problemas de este tipo; ambos requieren un tratamiento sistemático y cuidadoso para poder enlistar los resultados apropiadamente.

EJEMPLO 5.13

Suponga que se lanzan tres dados y nos piden determinar el número de formas en que pueden caer. Como veremos después, un espacio muestral contiene 216 resultados igualmente probables.

EJEMPLO 5.14

Considere el experimento de ir con un distribuidor de coches nuevos a comprar un modelo nuevo; una vez allí, usted encuentra que puede escoger entre cuatro modelos, cada uno con 15 opciones de potencia, con opción de cinco colores exteriores y ocho interiores ¿Cuántas elecciones distintas puede usted hacer? La respuesta es 2 400; en este caso, los resultados pueden no ser igualmente probables, a menos que usted escoja aleatoriamente.

Las técnicas especiales de conteo que estudiaremos, que pueden usarse para contar en problemas similares a los de los ejemplos 5.13 y 5.14, son: el teorema fundamental del conteo, el conteo de permutaciones y el conteo de combinaciones. Estudiaremos primero el teorema fundamental del conteo.

Teorema fundamental del conteo

El *teorema fundamental del conteo (TFC)* se establece como sigue:

> Si un evento puede ocurrir de m formas y si después que ha sucedido puede seguir un segundo evento que puede ser de cualquiera de n formas, entonces los dos eventos pueden ocurrir simultáneamente en el orden establecido de $(m)\,(n)$ formas. (Esta regla se puede extender a cualquier cantidad de eventos.)

Como los tres dados del ejemplo 5.13 no están relacionados en forma alguna cuando se lanzan y como cada uno puede caer de seis formas distintas, el número total de formas en que pueden caer, uno después del otro, es $(6)(6)(6) = 216$ por el *TFC*.

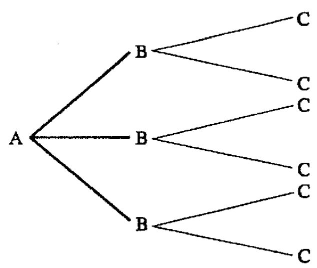
Para el ejemplo 5.15 pueden ocurrir cuatro eventos, uno después del otro. Son los siguientes:

E_1 = Escoger uno de los cuatro modelos.
E_2 = Elegir una de las 15 opciones de potencia.
E_3 = Preferir uno de los cinco colores exteriores.
E_4 = Seleccionar uno de los ocho colores interiores.

El primer evento puede ocurrir de cuatro formas, el segundo de 15, el tercero de cinco y el cuarto de ocho. Por tanto, como consecuencia del *TFC*, los cuatro eventos uno seguido del otro, pueden ocurrir en $(4)(15)(5)(8) = 2400$ maneras distintas. Note que éste es un ejemplo muy simplificado porque, en realidad, algunas opciones están relacionadas por los mismos fabricantes y distribuidores de cochos. Veamos una aplicación simple.

APLICACIÓN 5.20

Hay tres caminos de la ciudad A a la ciudad B y dos caminos desde la ciudad B hasta la ciudad C. ¿De cuántas formas distintas puede viajar una persona desde A hasta C pasando por B?



Solución: Considere el diagrama, llamado un **diagrama de árbol**.

Hay dos decisiones que deben tomarse:

a) Desde A, ¿cómo ir a B?
b) Desde B, ¿cómo ir a C?

La primera decisión puede hacerse de tres maneras y después de hacerla, realizar la segunda en dos formas. De acuerdo con el *TFC*, el total de formas para ir desde A hasta C es $(3)(2) = 6$. Esto puede concluirse también contando las ramas del diagrama que van desde A hasta C. ■

APLICACIÓN 5.21

Hay cinco libros diferentes que un maestro quiere ordenar en su escritorio de izquierda a derecha. ¿Cuántos arreglos son posibles?

Solución: Esta circunstancia puede verse de dos maneras; como un problema de colocación o como un problema de selección.

Un problema de colocación

Los cinco eventos son:

E_1 = Colocar el primer libro en uno de los cinco espacios.

E_2 = En seguida, colocar el segundo libro en uno de los cuatro espacios restantes.

E_3 = Después de los dos primeros libros, el tercero irá en uno de los tres espacios que sobran.

E_4 = De los dos espacios aún vacíos, uno será para el cuarto libro.

E_5 = El último libro va en el único lugar disponible.

E_1 puede hacerse de cinco formas, E_2 de cuatro, E_3 de tres, E_4 de dos maneras y E_5 sólo de una. Por tanto, por el teorema fundamental de conteo, el número total de arreglos posibles es (5)(4)(3)(2)(1) = 120.

Un problema de selección

Los cinco eventos son:

E_1 = Seleccionar un libro para el primer lugar puede hacerse de cinco maneras.

E_2 = Después de llenar el primer espacio, se elige el segundo libro de entre los cuatro restantes.

E_3 = Llenados los dos primeros espacios, elegimos el tercer libro de los restantes. Esto puede hacerse de tres formas.

E_4 = Luego de ocupar tres espacios, se selecciona el cuarto libro de entre los que quedan. Esto puede hacerse de dos formas.

E_5 = Colocar el quinto libro en el último espacio. Esto puede hacerse de una forma.

Por el *TFC*, el número total de selecciones es (5)(4)(3)(2)(1) = 120. ■

APLICACIÓN 5.22

¿De cuántas formas pueden elegirse tres artículos de un grupo de siete y acomodarlos en un estante de izquierda a derecha?

Solución: Hay tres eventos.

E_1 = Escoger un artículo de entre los siete para el primer espacio. Esto puede hacerse de siete formas.

E_2 = Seleccionar de entre los seis artículos restantes para ocupar el segundo espacio. Esto puede hacerse de seis maneras.

E_3 = Decidir de entre los cinco artículos restantes para el tercer espacio. Sólo de cinco maneras es posible hacerlo.

Mediante el *TFC*, vemos que el número total de formas es (7)(6)(5) = 210. ■

Si examinamos las aplicaciones anteriores, notamos que hay dos preguntas que deben plantearse cuando resolvemos cualquier problema de conteo:

1. ¿Cuántos eventos hay?
2. ¿De cuántas formas pueden ocurrir?

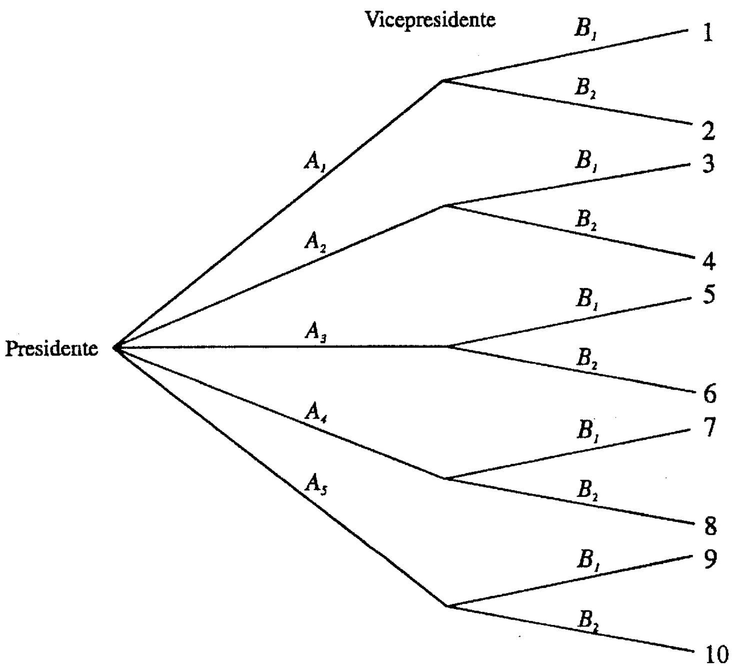
Después de encontrar la respuesta a estas preguntas usamos el *TFC* para resolver el problema.

APLICACIÓN 5.23

Hay cinco candidatos para presidente de un club y dos candidatos para vicepresidente; use un diagrama de árbol para determinar el número de maneras en que pueden cubrirse los dos puestos.

Solución: Hay dos eventos de interés: elegir un presidente y elegir un vicepresidente. El presidente se puede elegir de cinco maneras y el vicepresidente de dos; como consecuencia del *TFC*, los dos puestos pueden cubrirse de (5)(2) = 10 formas distintas. Se puede construir un diagrama de árbol para ilustrar las 10 posibilidades. Representemos con A_1, A_2, A_3, A_4 y A_5 los candidatos para presidente, y con B_1 y B_2 los candidatos a vicepresidente. Así tenemos el diagrama de árbol en la figura 5.9.

FIGURA 5.9
Diagrama de árbol para la aplicación 5.23



■

Permutaciones

Un arreglo ordenado de n objetos se llama **permutación**. Hay seis permutaciones de tres letras, A, B y C. Son:

ABC ACB BAC BCA CAB CBA

Podemos usar el *TFC* para determinar el número de permutaciones de n objetos, denotadas por P_n^n. Para hacerlo, imaginamos que los n objetos se colocan en n cajas dispuestas en fila: la primera caja se puede llenar de n formas, la segunda de $(n-1)$ formas, la tercera caja se puede llenar de $(n-2)$ formas y así sucesivamente hasta la última caja. Por tanto, el número de permutaciones de n objetos es igual a:

$$P_n^n = n(n-1)(n-2)(n-3)\ldots(2)(1) \tag{5.2}$$

La expresión del lado derecho de la ecuación 5.2, se llama ***n* factorial** y se denota por $n!$ Por definición, $0! = 1$. En consecuencia, $3! = (3)(2)(1) = 6$ y $5! = (5)(4)(3)(2)(1) = 120$.

Número de permutaciones de *n* objetos tomados de *n* a la vez

$$P_n^n = n! \tag{5.3}$$

EJEMPLO 5.15

El número de permutaciones de cinco objetos es

$$\begin{aligned} P_n^n &= n! \\ P_5^5 &= 5! \\ &= (5)(4)(3)(2)(1) = 120 \end{aligned}$$

EJEMPLO 5.16

Suponga que deseamos determinar el número de permutaciones de cinco objetos tomados tres a la vez; esto se escribe P_3^5. Podemos pensar este problema como el de colocar cinco objetos distintos en tres cajas; la primera caja puede llenarse de cinco maneras. Después de hacer esto, la segunda caja puede llenarse de cuatro formas, luego la tercera se llena de tres maneras. Como consecuencia del *TFC*, el número de permutaciones de cinco objetos distintos tomados tres a la vez es $(5)(4)(3) = 60$. Por tanto, $P_3^5 = 60$.

Para desarrollar una fórmula de cálculo del valor de P_r^n, hagamos las cuatro observaciones siguientes relativas a P_3^5.

1. De los cinco objetos, tres deben colocarse o permutarse y $(5 - 3) = 2$ no deben colocarse.
2. El número de permutaciones de los objetos que deben colocarse es $(5)(4)(3) = 60$, y el de aquellos objetos que no deben colocarse es $2! = 2$. Por cada una de las 60 permutaciones de los objetos que debemos colocar, tenemos dos que no deben reacomodarse.
3. Por el *TFC*, el número total de colocaciones de cinco objetos distintos es $(5)(4)(3)(2)(1) = 5!$ Por tanto, por la observación 2 tenemos $5! = 60 \cdot 2 = 120$ o $P_3^5 \cdot (5 - 3)! = 5!$ Si dividimos ambos lados de la ecuación entre $(5 - 3)!$, tenemos la siguiente relación: $P_3^5 = 5!/(5 - 3)!$

Las observaciones del ejemplo 5.16 sugieren la siguiente fórmula general para calcular P_r^n, el número de permutaciones de n objetos distintos de r en r $(n \geq r)$:

Número de permutaciones de *n* objetos tomados de *r* en *r*

$$P_r^n = \frac{n!}{(n-r)!} \tag{5.4}$$

La fórmula (5.4) se denomina **regla de permutaciones**.

APLICACIÓN 5.24

Suponga que diez estudiantes están disponibles para tres tareas distintas en el campus. ¿De cuántas formas pueden cubrirse los trabajos?

Solución: Necesitamos determinar cuántas formas hay de asignar las tres tareas entre diez estudiantes, o el número de acomodos de diez objetos tomados de tres en tres. Por la regla 5.4, el número de permutaciones de diez objetos tomados de tres en tres será:

$$P_3^{10} = \frac{10!}{(10-3)!}$$

$$= \frac{10!}{7!} = \frac{(10)(9)(8)(7!)}{7!}$$

$$= (10)(9)(8) = 720 \quad ■$$

Combinaciones

Cuando tratamos con permutaciones de objetos, el orden de selección o de colocación es importante; hay ocasiones en que nos interesa considerar colecciones de objetos donde el orden no es importante; cuando ocurre esto, la selección se llama **combinación**. En otras palabras, una selección de r objetos de una colección de n objetos distintos, sin importar el orden en que los r objetos son seleccionados, se llama **combinación**, y el número de combinaciones de n objetos tomados de r en r se denota por $\binom{n}{r}$. El símbolo $\binom{n}{r}$ se llama también un *coeficiente binomial.* Por ejemplo, supongamos que en un experimento se requiere seleccionar a un comité de tres mujeres de un grupo de ocho; aquí no importa el orden de selección de una combinación. El número de combinaciones (o selecciones) de ocho mujeres tomando tres a un tiempo se denota por $\binom{8}{3}$.

El número de combinaciones de n objetos tomados de r en r está relacionado con el número de permutaciones de n objetos tomados de r en r; cada combinación de r objetos puede arreglarse de $r!$ maneras distintas. Si aplicamos el *TFC*, el número total de permutaciones de n objetos distintos tomados de r en r, es igual al productos de r, y el número de combinaciones de n objetos distintos tomados de r en r. Es decir,

$$P_r^n = r!\binom{n}{r}$$

En consecuencia, el número de combinaciones de n objetos distintos tomados de r en r está dado por

$$\binom{n}{r} = \frac{P_r^n}{r!} \qquad (5.5)$$

EJEMPLO 5.17

El número de maneras de elegir a tres mujeres de un grupo de ocho es:

$$\binom{n}{r} = \frac{P_r^n}{r!}$$
$$\binom{8}{3} = \frac{P_3^8}{3!} = \frac{(8)(7)(6)}{6} = 56$$

Podemos expresar la fórmula (5.5) mediante la fórmula (5.4); así:

Número de combinaciones de n objetos tomados de r en r

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \qquad (5.6)$$

APLICACIÓN 5.25

Calcule los coeficientes binomiales siguientes.

a) $\binom{6}{4}$

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

c) $\binom{5}{3}$

d) $\binom{5}{2}$

Solución: Si usamos la fórmula (5.6), tenemos

a) $$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$
$$\binom{6}{4} = \frac{6!}{4!(6-4)!}$$
$$= \frac{(6)(5)(4!)}{4!2!}$$
$$= \frac{(6)(5)}{2} = 15$$

b) $$\binom{6}{2} = \frac{6!}{2!\,4!} = 15$$

c) $$\binom{5}{3} = \frac{5!}{3!\,2!}$$
$$= \frac{(5)(4)(3!)}{3!\,2!}$$
$$= \frac{(5)(4)}{2} = 10$$

d) $$\binom{5}{2} = \frac{5!}{2!3!} = 10,$$

por el inciso c). ■

Se puede demostrar en general que el número de combinaciones de n objetos tomados de r en r, es igual al número de combinaciones de n objetos tomados de $n - r$ en $n - r$. Esto es,

$$\binom{n}{r} = \binom{n}{n-r}$$

Si utilizamos este hecho, a menudo podremos simplificar cálculos relativos a coeficientes binomiales.

EJEMPLO 5.18

a) $\binom{10}{8} = \binom{10}{2} = 10!/(2!8!) = (10)(9)(8!)/(2!8!) = 45$. Para obtener 45, multiplicamos los dos factores más grandes de 10!, 10 y 9, y dividimos su producto entre $2! = (2)(1) = 2$. Note que tanto el numerador como el denominador de $\binom{10}{2} = (10)(9)/(2)(1)$ tienen el mismo número de factores, que es 2.

b) $\binom{10}{7} = \binom{10}{3} = 10!/(3!7!) = (10)(9)(8)(7!)/(3!7!) = 120$. Para obtener 120, multiplicamos los tres factores más grandes de 10!, 10, 9 y 8, y dividimos su producto entre $3! = (3)(2)(1) = 6$. Observe nuevamente que tanto el numerador como el denominador de $\binom{10}{3} = (10)(9)(8)/(3)(2)(1)$ tienen tres factores.

APLICACIÓN 5.26

Se realiza un estudio para determinar la opinión de los profesores de una universidad respecto al aborto; si se elige una muestra de cuatro profesores de un total de 45, ¿cuántas muestras distintas pueden seleccionarse?

Solución: Como en una muestra el orden no es importante, nos interesa determinar el número de combinaciones de 45 profesores tomados de cuatro en cuatro. Por la fórmula (5.6), tenemos:

$$\begin{aligned}\binom{45}{4} &= \frac{45!}{(45-4)!4!}\\ &= \frac{45!}{41!4!}\\ &= \frac{(45)(44)(43)(42)(41!)}{41!4!}\\ &= \frac{(45)(44)(43)(42)}{4!} = \frac{3{,}575{,}880}{24} = 148{,}995 \quad \blacksquare\end{aligned}$$

Triángulo de Pascal

Los coeficientes binomiales se pueden obtener de un arreglo triangular de números llamado **triángulo de Pascal**. El arreglo siguiente es un triángulo parcial de siete renglones:

	x resultados						
n ensayos	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
6	1	6	15	20	15	6	1

Es fácil construir el triángulo; la primera columna ($x = 0$) tiene sólo unos. Para obtener un número de cualquier renglón, exceptuando los de la primera columna, basta sumar el número que se encuentra arriba del que nos interesa, más el que se encuentra arriba del que está a la izquierda.

EJEMPLO 5.19

Para determinar $\binom{6}{2}$, vamos al renglón correspondiente a $n = 6 - 1 = 5$ y sumamos 5 y 10. Por tanto, $\binom{6}{2} = 15$. Para encontrar $\binom{6}{5}$ sumamos 5 y 1. El proceso de sumar números del renglón anterior puede continuarse para construir tantos renglones como se desee.

n ensayos	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	$\boxed{5}$	$\boxed{10}$	10	5	1	
6	1	6	15	20	15	6	1

Las entradas en el renglón correspondiente a $n = 7$ son:

1 7 21 35 35 21 7 1

GRUPO DE EJERCICIOS 5.3

Habilidades básicas

1. Evalúe las expresiones siguientes:

a) P_2^4 b) $\binom{5}{2}$ c) P_3^6

d) P_5^5 e) $\binom{7}{3}$ f) $\binom{9}{4}$

2. Calcule las expresiones siguientes:

a) P_1^5 b) $\binom{6}{4}$ c) P_2^7

d) P_6^8 e) $\binom{9}{2}$ f) $\binom{10}{3}$

Más aplicaciones

3. Si suponemos que todas las combinaciones son posibles, ¿cuántas casas distintas puede edificar un constructor que ofrece cinco planes básicos, tres tipos de techo y dos acabados exteriores?

4. Una escuela comercial da cursos de mecanografía, taquigrafía, transcripción, inglés comercial, edición y contabilidad. ¿Entre cuántas formas puede elegir un estudiante tres cursos que tomará durante tres periodos de clases?

5. ¿De cuántas formas puede ordenarse a diez monos en una fila para un experimento de genética?

6. En un experimento de interacción social, se sentará a seis personas en una fila de seis sillas ¿Cuántos resultados podemos obtener?

7. Al montar equipo electrónico, seis alambres se conectan a una caja que tiene seis terminales. ¿De cuántas formas pueden conectarse los alambres a las terminales si sólo entra un alambre en cada terminal?

8. ¿De cuántas maneras puede un juez adjudicar el primero, el segundo y el tercer lugar en un examen de 15 entradas?

9. Una tienda de ropa vende calcetines de algodón o lana, cada uno en cinco colores y siete medidas. ¿Cuántos artículos debe haber en el almacén para tener disponible un surtido completo?

10. Un hombre intenta escoger al ganador de cada uno de ocho partidos de futbol; sin contar los empates, ¿cuántas predicciones distintas son posibles? ¿Y si se incluyen los empates?

11. Responda el ejercicio 6 para el caso de que las seis personas se sienten en torno a una mesa circular, siendo sus posiciones relativas las que determinen una diferencia y no las sillas que ocupan.

12. Si usted está haciendo un examen de falso-verdadero con diez preguntas, ¿de cuántas formas distintas puede llenar su hoja de respuestas si desconoce todas las respuestas y adivina en cada una?

13. ¿De cuántas formas puede asignarse a nueve hombres en nueve tareas distintas?

14. ¿Y cinco tareas a nueve hombres?

15. Si se debe elegir cinco cartas en forma sucesiva, sin reemplazo, de un paquete común de 52, ¿cuántas selecciones diferentes son posibles?

16. Refiérase al ejercicio 15. Si las cartas se eligen con remplazo, ¿cuántas selecciones distintas resultarían?

17. Un teclado especial de computadora puede comprarse en seis casas distribuidoras diferentes. ¿De cuántas formas pueden elegirse cuatro distribuidoras de entre las seis?

Un paso más allá

18. La Upper Crust Pizza Shop anuncia diez ingredientes distintos para las pizzas. ¿Cuántas pizzas pueden ordenarse sin repetir los ingredientes?

19. Refiérase al ejercicio 4. ¿De cuántas formas distintas puede elegirse una lista de tres cursos?

SECCIÓN 5.4 *Determinación de probabilidades mediante el teorema fundamental del conteo*

El teorema fundamental del conteo (*TFC*) se puede usar para resolver algunos de los problemas de probabilidad más difíciles. Las aplicaciones 5.27 a 5.30 muestran el uso del *TFC* para resolver problemas de probabilidad.

APLICACIÓN 5.27

Si se lanzan cuatro monedas, ¿cuál es la probabilidad de obtener cuatro "soles"?

Solución: Primero determinamos el número de formas en que pueden caer las cuatro monedas cuando se lanzan una tras otra. Denotemos por E_1, E_2, E_3 y E_4 el lanzamiento de las cuatro monedas, respectivamente; cada una puede caer de dos formas. Por el teorema fundamental del conteo hay $2^4 = 16$ formas distintas; de ellas, sólo una es un resultado favorable, SSSS. En consecuencia,

$$P(\text{SSSS}) = \frac{1}{16} \quad \blacksquare$$

APLICACIÓN 5.28

Si se lanzan cuatro monedas, ¿cuál es la probabilidad de que las dos primeras monedas muestren "sol"?

Solución: El número total de resultados es 16, según la aplicación 5,27; un resultado favorable es aquel en el que las dos primeras monedas caen en "sol"; la primera moneda debe mostrar el "sol", la segunda también, y la tercera y la cuarta pueden mostrar "sol" o águila. Por el teorema fundamental del conteo, el número total de resultados favorables es (1)(1)(2)(2) = 4; por tanto, la probabilidad de que las dos primeras monedas sean "soles" es 4/16 = 1/4. ■

APLICACIÓN 5.29

De un conjunto de 100 tarjetas numeradas del 1 al 100, se elige una al azar. ¿Cuál es la probabilidad de que el número en ella,

a) sea divisible exactamente entre 5?
b) termine en 1 o en 2?

Solución:

a) Representemos por E al evento de que el número de la tarjeta sea divisible entre 5; un número es divisible entre 5 si termina en 0 o en 5; los números de un dígito divisibles entre 5 son iguales a 1; los de dos dígitos divisibles entre 5 son (9)(2) = 18; y hay un número de tres dígitos divisibles entre 5. Por tanto, el número total de tarjetas con números divisibles entre 5 es $f = 1 + 18 + 1 = 20$. Como hay 100 tarjetas, tenemos

$$P(E) = \frac{f}{n} = \frac{20}{100} = \frac{1}{5}$$

b) Representemos por F al evento de que el número de la tarjeta termine en 1 o en 2; un número que termine en 1 o en 2 es un número de uno o de dos dígitos; hay $f_1 = 2$ tarjetas numeradas con un dígito que termina en 1 o en 2.
Para las tarjetas numeradas con dos dígitos, el dígito de las unidades puede ser uno de los dos valores y el de las decenas puede ser cualquiera de los nueve valores. Por el *TFC*, el número de tarjetas con dos dígitos que terminan en 1 o en 2 es $f_2 = (9)(2) = 18$. En consecuencia,

$$P(F) = \frac{f_1 + f_2}{n} = \frac{2 + 18}{100} = \frac{20}{100} = \frac{1}{5} \quad \blacksquare$$

APLICACIÓN 5.30

Hay diez libros distintos en un estante, de los cuales siete son de matemáticas y tres de otras ciencias; si se colocan aleatoriamente, ¿cuál es la probabilidad de que los libros que no son de matemáticas queden juntos?

Solución: Sea E el evento de que los libros de las otras ciencias queden juntos. Por el TFC, los diez libros pueden colocarse de (10)(9)(8)(7)(6)(5)(4)(3)(2)(1) = 3.628,800 formas distintas. Para encontrar el número de arreglos de los 10 libros en que los libros de las otras ciencias quedan juntos, identificamos los eventos siguientes:

E_1 = Colocar los libros de matemáticas.
E_2 = Ubicar los libros de las otras ciencias.
E_3 = Insertar los libros de las otras ciencias entre los libros de matemáticas.

El número de formas en que puede ocurrir E_1 es $f_1 = (7)(6)(5)(4)(3)(2)(1) = 5040$; el de E_2 es $f_2 = (3)(2)(1) = 6$; el de E_3 es $f_3 = 8$ (poner los libros de otras ciencias antes de los de matemáticas, después del primer libro de matemáticas, en seguida del segundo, luego del tercer libro de matemáticas y así sucesivamente). Por el TFC, el número de formas de arreglar los diez libros de modo que los de las otras ciencias queden juntos es:

$$\begin{aligned} f &= (f_1)(f_2)(f_3) \\ &= (5040)(6)(8) = 241{,}920 \end{aligned}$$

Así, la probabilidad de que los libros de las otras ciencias queden juntos será:

$$\begin{aligned} P(E) &= \frac{f}{n} \\ &= \frac{241{,}920}{3{,}628{,}800} \approx 0.06667 \end{aligned}$$ ■

GRUPO DE EJERCICIOS 5.4

Más aplicaciones

1. En una caja hay boletos numerados del 1 al 20 inclusive, y se sacan dos, uno a continuación del otro sin remplazar el primero. ¿Cuál es la probabilidad de que:
   a) ambos sean pares?
   b) el primero sea par y el segundo impar?
   c) ambos sean pares o impares?
2. Refiérase al ejercicio 1. Resuelva las partes a) y c) para el caso de si el primer boleto se remplaza antes de sacar el segundo.
3. Cuatro tornillos y cuatro tuercas están revueltos. Si se eligen dos piezas al azar, ¿cuál es la probabilidad de que:
   a) ambas sean tornillos?
   b) la primera sea tornillo y la segunda sea tuerca?
   c) una sea tornillo y la otra tuerca?
   d) ambas sean tornillos o tuercas?
4. Las letras de la palabra *Chance* se han escrito en tarjetas, una letra en cada tarjeta; las tarjetas son barajadas y se voltean una después de otra. ¿Cuál es la probabilidad de que se haya escrito la palabra *Chance*? (Tenga en cuenta que la primera letra es mayúscula.)
5. ¿Cuál es la probabilidad de que un estudiante escriba la palabra *MATE* si coloca aleatoriamente las letras M, A, T y E?

6. ¿Qué probabilidad hay de que una palabra de cuatro letras elegidas de entre la palabra *figura* comience con una consonante?
7. Si diez pelotas numeradas, cada una con un único número entre el 1 y el 10; si se colocan a lo largo de un lado de una mesa de juego, ¿cuál es la probabilidad de que la 5 y la 6 queden juntas?
8. ¿Qué probabilidad existe de que un número de tres dígitos formado con 1, 2, 3, 4 y 5 sea par? ¿Y de que sea impar? (Suponga en ambos casos que los dígitos pueden repetirse.)
9. Resuelva el ejercicio 8 bajo la hipótesis de que los dígitos no pueden repetirse.

Un paso más allá

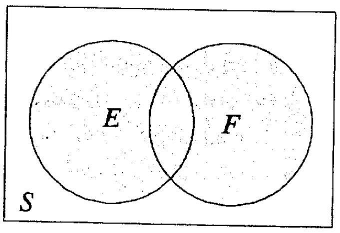
10. Si cuatro hombres y cuatro mujeres se colocan en fila, ¿cuál es la probabilidad de que un arreglo aleatorio de los ocho individuos tenga:
    a) hombres y mujeres alternativamente?
    b) a los hombres uno junto a otro?
11. Si cuatro hombres y cuatro mujeres van a sentarse en torno a una mesa redonda, ¿cuál es la probabilidad de que una colocación al azar de los ocho individuos tenga:
    a) hombres y mujeres alternadamente?
    b) a los hombres uno junto al otro?
12. De un conjunto de tarjetas numeradas del 1 al 10,000, se elige una al azar. ¿Cuál es la probabilidad de que el número en ella:
    a) sea divisible entre 5 exactamente?
    b) termine en un número par?
13. En un motor de ocho cilindros, los cilindros con número par están del lado izquierdo y los de número impar del lado derecho; un buen orden de encendido es un arreglo en el cual los dos lados del motor se alternan al encender, comenzando con el cilindro 1. (Por ejemplo, 1, 4, 5, 8, 3, 2, 7 y 6 es un buen orden de encendido.) Si el motor lo interconecta un novato que no sabe nada de lo que está haciendo, determine la probabilidad de que haya logrado un buen orden de encendido.

SECCIÓN 5.5 *Algunas reglas de probabilidad*

Probabilidad de E o F $P(E \cup F)$

Si E y F son dos eventos, quisiéramos desarrollar una fórmula para $P(E \cup F)$; examinemos la figura 5.10, un diagrama de Venn para $(E \cup F)$. Representemos con $P(E)$ el área del círculo que contiene a E, y con $P(E)$ el área F. El evento $(E \cup F)$ se ilustra con toda la región sombreada dentro de los dos círculos y $P(E \cup F)$ es el área de esa región. También está representada el área común a las dos regiones E y F, que es $P(E \cap F)$. La región $E \cap F$ está contada dos veces, una por causa de E y otra para F; así que necesitamos restar lo que contamos de más. Entonces

FIGURA 5.10
Diagrama de Venn para $(E \cup F)$



$$P(E \cup F) = (\text{Área } E) + (\text{Área } F) - (\text{Área común a } E \text{ y } F)$$

$$= P(E) + P(F) - P(E \cap F)$$

En consecuencia, la probabilidad de $(E \cup F)$ puede encontrarse usando la fórmula:

Regla de la suma

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) \tag{5.7}$$

La regla (5.7) se denomina **regla de la suma** y la aplicación 5.31 ilustra su uso.

APLICACIÓN 5.31

Si se toma una sola carta de una baraja, encuentre la probabilidad de que sea roja o figura (sota, reina y rey).

Solución: Representemos por E la selección de una carta roja y por F la de una con figura. Entonces $P(E) = 26/52$, $P(F) = 12/52$ y $P(E \cap F) = 6/52$. Usando la regla de la suma tenemos:

$$\begin{aligned} P(E \cup F) &= P(E) + P(F) - P(E \cap F) \\ &= \frac{26}{52} + \frac{12}{52} - \frac{6}{52} = \frac{32}{52} = \frac{8}{13} \quad \blacksquare \end{aligned}$$

Si E y F son eventos mutuamente excluyentes, entonces $(E \cap F) = \emptyset$, el conjunto vacío. En tal caso, $P(E \cap F) = 0$ y la regla de la suma, (5.7), toma la forma $P(E \cup F) = P(E) + P(F)$, de lo que se sigue la regla:

Si E y F son eventos mutuamente excluyentes, entonces

$$P(E \cup F) = P(E) + P(F) \tag{5.8}$$

Probabilidad de no E, $P(\overline{E})$

Como E y $\overline{E}$ son eventos mutuamente excluyentes, podemos aplicar la regla (5.8) para obtener:

$$P(E \cup \overline{E}) = P(E) + P(\overline{E})$$

Además, como $P(E \cup \overline{E}) = P(S) = 1$, tenemos $P(E) + P(\overline{E}) = 1$, es decir, $P(\overline{E}) = 1 - P(E)$. De ello inferimos la regla:

Probabilidad de $\overline{E}$

$$P(\overline{E}) = 1 - P(E) \tag{5.9}$$

APLICACIÓN 5.32

La probabilidad de que Roberto haya terminado un artículo es de 3/7. Encuentre la probabilidad de que no lo haya terminado.

Solución: Sea E el evento de que Roberto haya terminado el artículo; entonces $\overline{E}$ es el evento de que no lo haya terminado. Como P(E) = 3/7, usando la regla (5.9) tenemos:

$$\begin{aligned} P(\overline{E}) &= 1 - P(E) \\ &= 1 - \frac{3}{7} = \frac{4}{7} \quad \blacksquare \end{aligned}$$

Probabilidad condicional

Ahora queremos encontrar una regla para calcular $P(E \cap F)$, la probabilidad de $(E \cap F)$; para hacerlo, necesitamos la notación de probabilidad condicional. A la probabilidad de que ocurra el evento E cuando sabemos que ya ha ocurrido el evento F, se le llama **probabilidad condicional** y se escribe como $P(E|F)$. $P(E|F)$ significa la probabilidad de que ocurra el evento E dada la condición de que el evento F ya haya sucedido; o simplemente, la probabilidad de E dado F. Las aplicaciones 5.33 a 5.35 ejemplifican la probabilidad condicional mediante la selección de problemas típicos; en estas aplicaciones, **selección con remplazo** significa que el primer objeto se devuelve antes de elegir un segundo objeto; análogamente, **selección sin remplazo** quiere decir que el primer objeto no se devuelve antes de elegir el segundo. Observe en particular, que las aplicaciones 5.33 y 5.34 enseñan que el tipo de muestreo tiene un efecto en los resultados.

APLICACIÓN 5.33

Dos pelotas se sacan sin remplazo de una bolsa que contiene tres pelotas blancas y dos negras. Encuentre la probabilidad de que:

a) La segunda pelota sea negra dado que la primera pelota es blanca.
b) La segunda pelota sea negra dado que la primera pelota es negra.

Solución:

a) Si la primera pelota elegida es blanca, quedan cuatro en la bolsa y de ellas dos son negras. Por tanto,

$$P(\text{N}|\text{B}) = \frac{2}{4} = \frac{1}{2}$$

Vea el diagrama siguiente:

$$\begin{array}{l} \quad\;\; \cancel{\text{B}} \\ 4 \rightarrow \left\{\begin{array}{l} \text{B} \\ \text{B} \\ \left.\begin{array}{l} \text{N} \\ \text{N} \end{array}\right\} \leftarrow 2 \end{array}\right. \end{array}$$

b) Si la primera pelota elegida es negra, quedan cuatro en la bolsa y de ellas una es negra. En consecuencia,

$$\text{P}(\text{N}|\text{N}) = \frac{1}{4}$$

Vea el diagrama:

$$\begin{array}{l} 4 \rightarrow \left\{\begin{array}{l} \text{B} \\ \text{B} \\ \text{B} \\ \left.\text{N}\right\} \leftarrow 1 \end{array}\right. \\ \quad\;\; \cancel{\text{N}} \end{array}$$

■

APLICACIÓN 5.34

Suponga que las pelotas de la aplicación 5.33 se sacan con remplazo; dos pelotas se sacan con remplazo de una bolsa que contiene tres blancas y dos negras. Encuentre la probabilidad de que:

a) La segunda pelota sea negra dado que la primera es blanca.
b) La segunda pelota sea negra dado que la primera es negra.

Solución:

a) Como la pelota blanca se regresa a la bolsa antes de sacar la segunda, no tiene efecto en la selección de la pelota siguiente. Como hay cinco pelotas en la bolsa y dos de ellas son negras, la probabilidad de seleccionar una negra es de 2/5.
b) Como la pelota negra se devuelve a la bolsa antes de elegir la siguiente, no tiene efecto en la selección de la siguiente. Como en la parte a, la probabilidad de seleccionar una pelota negra es de 2/5. ■

APLICACIÓN 5.35

En una universidad se realiza un estudio para determinar qué relación existe, en caso de haberla, entre la habilidad matemática y el interés por las matemáticas. Se determina la habilidad y el interés de 150 estudiantes, con los resultados de la tabla siguiente:

	Interés		
Habilidad	Escaso	Promedio	Mucho
Escasa	40	8	12
Promedio	15	17	18
Mucha	5	10	25

Si se escoge al azar uno de los participantes en el estudio, ¿cuál es la probabilidad

a) de elegir a una persona que tenga escaso interés en las matemáticas?
b) de seleccionar a una persona con habilidad promedio?
c) de que la persona tenga mucha habilidad para las matemáticas dado que manifieste mucho interés por esa disciplina?
d) de que la persona tenga mucho interés en las matemáticas dado que posea una habilidad promedio?

Solución: Primero determinamos el renglón y la columna de los totales, como vemos en la tabla 5.5.

TABLA 5.5
Renglones y columnas de los totales para la aplicación 5.35

	Interés			
Habilidad	Escaso	Medio	Mucho	Total
Escasa	40	8	12	60
Promedio	15	17	18	5
Mucha	5	10	25	40
Total	60	35	55	150

a) Como hay 60 participantes con escaso interés de un total de 150, la probabilidad es de 60/150 = 2/5.

b) Como hay 50 interesados con habilidad promedio de un total de 150, la probabilidad es de 50/150 = 1/3.
c) De los 55 participantes con mucho interés, 25 tienen gran habilidad. Por tanto, la probabilidad es 25/55 = 5/11.
d) Como 50 participantes tienen habilidad promedio y de ellos 18 tienen mucho interés, la probabilidad es de 18/50 = 9/25. ■

Note que las partes c) y d) se refieren a probabilidad condicional. Para c), sabemos que la persona escogida tiene mucho interés en las matemáticas; hay 55 estudiantes con mucho interés en las matemáticas; de ellos, 25 tienen gran habilidad, denotemos por MH tener mucha habilidad, por MI mucho interés y por HP habilidad promedio. Entonces:

$$P(\mathrm{MH}|\mathrm{MI}) = \frac{25}{55} = \frac{5}{11}$$

Esta probabilidad se puede escribir como:

$$\begin{aligned} P(\mathrm{MH}|\mathrm{MI}) &= \frac{P(\mathrm{MH} \cap \mathrm{MI})}{P(\mathrm{MI})} \\ &= \frac{25/150}{55/150} \\ &= \frac{25}{55} = \frac{5}{11} \end{aligned}$$

También, para la parte d de la aplicación 5.35 encontramos:

$$\begin{aligned} P(\mathrm{MI}|\mathrm{HP}) &= \frac{P(\mathrm{MI} \cap \mathrm{HP})}{P(\mathrm{HP})} \\ &= \frac{18/150}{50/150} \\ &= \frac{18}{50} = \frac{9}{25} \end{aligned}$$

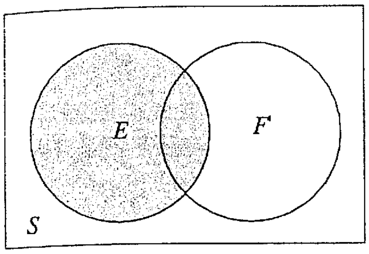
Estos resultados sugieren la fórmula siguiente para calcular $P(E|F)$.

Fórmula para la probabilidad condicional

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \qquad (5.10)$$
$$P(F) \neq 0$$

Advierta que la condición F en $E \mid F$ tiene el efecto de reducir el tamaño del espacio muestral. En la aplicación 5.35, parte c), el tamaño del espacio muestral es 150 y, después de que se estipula la condición MI, ese tamaño del espacio muestral se reduce a sólo aquellos con mucho interés; es decir, a 55.

FIGURA 5.11
Diagrama de Venn para ilustrar la probabilidad condicional



La probabilidad condicional $P(E \mid F)$ puede ilustrarse usando un diagrama de Venn en el que las probabilidades se interpretan como áreas (véase la figura 5.11). Si estipulamos la condición F, estamos restringidos a los resultados en ella; por eso los elementos en F comprenden el espacio muestral reducido. Entonces, estamos interesados en los resultados de E que también están contenidos en F, que son precisamente los resultados en E y en F. La probabilidad de E dado F se obtiene entonces dividiendo el área de $(E \cap F)$ entre el área de F. Así,

$$P(E|F) = \frac{\text{Área } (E \cap F)}{\text{Área } F} = \frac{P(E \cap F)}{P(F)} \qquad (5.11)$$

Probabilidad de E y F, P(E ∩ F)

La probabilidad de $E \cap F$ está dada por la fórmula:

Regla del producto

$$P(E \cap F) = P(F)\, P(E/F) \qquad (5.12)$$

Esta fórmula (5.12) se obtuvo multiplicando ambos lados de la fórmula (5.10) por $P(F)$. La regla (5.12) suele llamarse **regla del producto** y se ilustra en la aplicación 5.36. Note que $(E \cap F)$ es lógicamente equivalente a $(F \cap E)$; por eso tenemos:

$$P(E \cap F) = P(F)\, P(E \mid F) = P(E)\, P(F \mid E)$$

Un punto crucial que debe recordarse al usar la regla del producto para calcular $P(E \cap F)$ es la colocación del símbolo del evento que ocupa la posición "*" en la siguiente expresión:

$$P(E \cap F) = P(*)\, P(\ \mid *) \qquad (5.13)$$

Por ejemplo, si * se remplaza por E, entonces la expresión (5.13) toma la forma

$$P(E \cap F) = P(E)\, P(F \mid E)$$

Si * se remplaza por F, la expresión (5.13) adopta la forma

$$P(E \cap F) = P(F)\, P(E \mid F)$$

APLICACIÓN 5.36

Si dos pelotas se sacan sin remplazo de una bolsa que contiene tres rojas, dos negras y una blanca, ¿cuál es la probabilidad de obtener dos pelotas rojas?

Solución: Sea R_1 el evento de sacar una pelota roja en la primera oportunidad y denotemos por R_2 al evento de sacar una pelota roja en la segunda oportunidad. Mediante la regla (5.12) tenemos:

$$P(R_1 \cap R_2) = P(R_1)P(R_2 \mid R_1) = \frac{3}{6}\,\frac{2}{5} = \frac{1}{5} \quad \blacksquare$$

APLICACIÓN 5.37

Localice la probabilidad de sacar dos pelotas rojas en la aplicación 5.36 si se sacan con remplazo.

Solución: Mediante la regla (5.12) concluimos:

$$P(R_1 \cap R_2) = P(R_1)\,P(R_2 \mid R_1) = \frac{3}{6}\,\frac{3}{6} = \frac{1}{4}$$

Note que en este caso

$$P(R_2 | R_1) = P(R_2) \quad \blacksquare$$

APLICACIÓN 5.38

Remítase a la aplicación 5.35.

a) Encuentre la probabilidad de elegir a una persona con mucha habilidad y gran interés en matemáticas.
b) Localice la probabilidad de elegir a una persona con escaso interés y gran habilidad en matemáticas.

Solución:

a) Mediante la regla del producto, tenemos:

$$P(MH \cap MI) = P(MH)\,P(MI | MH) = \frac{40}{150}\,\frac{25}{40} = \frac{25}{150} = \frac{1}{6}$$

b) Mediante la regla del producto:

$$P(EI \cap MH) = P(MH)\,P(EI | MH) = \left(\frac{40}{150}\right)\left(\frac{5}{40}\right) = \frac{5}{150} = \frac{1}{30} \quad \blacksquare$$

Para calcular probabilidades asociadas con eventos compuestos, deberán tenerse en mente las dos siguientes observaciones:

1. Para calcular $P(E \cup F)$, asociamos $\cup$ con la regla de la suma.
2. Para calcular $P(E \cap F)$, asociamos $\cap$ con la regla del producto.

GRUPO DE EJERCICIOS 5.5

Habilidades básicas

1. Sean $P(E) = 0.4$, $P(F \mid E) = 0.3$, y $P(E \mid F) = 0.4$.
   a) Encuentre lo siguiente:
      i. $P(E \cap F)$
      ii. $P(F)$
      iii. $P(E \cup F)$
   b) ¿Son eventos mutuamente excluyentes E y F?

2. Sean $P(E) = 0.5$, $P(F) = 0.4$ y $P(E \cup F) = 0.8$.
   a) Encuentre:
      i. $P(E \cap F)$
      ii. $P(F \mid E)$
      iii. $P(E \mid F)$
   b) ¿Son eventos mutuamente excluyentes E y F?

3. Suponga que $S = \{1, 2, 3, 4, 5, 6\}$ es un espacio muestral para un experimento; la tabla adjunta enlista las probabilidades asociadas con cada elemento del espacio muestral.

x	1	2	3	4	5	6
$P(x)$	0.1	0.3	0.1	0.2	0.1	0.2

   Sean $E = \{1, 2, 5\}$ y $F = \{1, 6\}$; encuentre las probabilidades siguientes:
   a) $P(E)$ b) $P(F)$ c) $P(E \cup F)$
   d) $P(E \cap F)$ e) $P(E \mid F)$ f) $P(\overline{E})$
   g) $P(\overline{F} \cap E)$ h) $P(\overline{E} \cap \overline{F})$ i) $P(F \mid \overline{E})$

4. Sea $S = \{1, 2, 3, 4, 5\}$, un espacio muestral para un experimento; en seguida adjuntamos las probabilidades asociadas con cada elemento del espacio muestral.

x	1	2	3	4	5
$P(x)$	0.2	0.2	0.1	0.3	0.2

   Sea $E = \{1, 2\}$ y $F = \{2, 3\}$, calcule las probabilidades:
   a) $P(E)$ b) $P(F)$ c) $P(E \cup F)$
   d) $P(E \cap F)$ e) $P(E \mid F)$ f) $P(\overline{E})$
   g) $P(\overline{F} \cap E)$ h) $P(\overline{E} \cap \overline{F})$ i) $P(F \mid \overline{E})$

Más aplicaciones

5. Si se lanza un dado legal, qué probabilidad existe de haber sacado:
   a) Un 2, si se sabe que el número mostrado es impar.
   b) Un 4, si se sabe que el número mostrado es par.
   c) Un par, si sabemos que el número mostrado es 6.

6. Si se saca una pelota de una bolsa que contiene tres pelotas rojas, dos blancas y cinco negras, encuentre la probabilidad de que:
   a) sea roja
   b) no sea roja
   c) sea roja o negra
   d) ni sea roja ni blanca

7. Se sacan dos pelotas sin remplazo de una bolsa que contiene dos pelotas bancas y dos negras; calcule la probabilidad de que:
   a) La segunda pelota sea blanca dado que la primera fue blanca.
   b) La segunda pelota sea blanca si la primera fue negra.
   c) Ambas pelotas sean blancas.
   d) La primera pelota sea blanca y la segunda negra.
   e) Una pelota sea blanca y la otra negra.

8. Si se sacan dos cartas sin remplazo de un paquete de 52, qué probabilidad hay de que:
   a) La segunda carta sea un corazón si la primera fue un corazón.
   b) Ambas cartas sean corazones.
   c) La segunda carta sea negra cuando la primera fue una espada.
   d) La segunda carta sea figura dado que la primera fue una sota.
   e) La primera carta sea una sota y la segunda una figura.

9. Resuelva el ejercicio 7 si la primera pelota se remplaza antes de sacar la segunda.

10. Resuelva el ejercicio 8 cuando la primera carta se remplaza antes de sacar la segunda.

11. De una jarra con cuatro canicas negras y seis blancas, se sacan dos canicas sin remplazo. Encuentre la probabilidad de que:
   a) ambas sean blancas
   b) las dos sean negras
   c) la segunda sea blanca dado que la primera es negra
   d) difieran en color

12. Resuelva el ejercicio 11 si la primera canica se remplaza antes de sacar la segunda.

13. Suponga que en un hospital un estudio indica que el 35% de los pacientes admitidos tienen alta la presión sanguínea, el 53% molestias cardiacas y un 25% sufren los dos padecimientos. ¿Cuál es la probabilidad de que un paciente elegido al azar tenga:

a) presión sanguínea alta, molestias cardiacas o ambas?
b) presión sanguínea alta dado que tiene molestias cardiacas?
c) molestias cardiacas dado que tiene presión sanguínea alta?
d) ni molestias cardiacas ni presión sanguínea alta?

14. La probabilidad de que una persona nade es 0.45 y la de que una persona cace es 0.58; si la probabilidad de que una persona nade sabiendo que también caza es 0.21, encuentre la probabilidad de que:
a) cace y nade
b) cace si también nada
c) cace y no nade
d) cace o nade

15. La tabla adjunta muestra las frecuencias relativas para el daltonismo en hombres y mujeres, donde H representa hombres, M mujeres, D daltónico y $\overline{\text{D}}$ no daltónico.

	H	M
D	0.042	0.007
$\overline{\text{D}}$	0.485	0.466

Si se escoge a una persona al azar, use la tabla para determinar las probabilidades siguientes:
a) P(H)
b) P(H | D)
c) P(M | D)
d) P(D)
e) P(F | $\overline{\text{D}}$)
f) P(H ∩ D)
g) P(D | H)

16. Un banco ha observado que la mayor parte de los clientes solicitan a los empleados de las ventanillas cambiar un cheque o hacer un depósito. La tabla muestra las operaciones realizadas por el empleado A en un día.

	Cambiar cheque	No cambiar cheque
Depositar	50	20
No depositar	30	10

Representemos por C cambiar un cheque y por D hacer un depósito. Se escoge al azar a un cliente del empleado A. Exprese cada una de las probabilidades siguientes en palabras y encuentre su valor.
a) P(C)
b) P(D)
c) P(C ∩ D)
d) P(C ∪ D)
e) P(C | D)
f) P($\overline{\text{C}}$ | D)
g) P($\overline{\text{D}}$ | D)

17. En un estudio de actitudes sobre una legislación para el control estricto de armas de fuego efectuado con 800 estadounidenses adultos, se obtuvo:

	Posición	
	A favor	En contra
Ha disparado un arma	75	200
Nunca ha disparado un arma	425	100

Si elige al azar uno de los 800 adultos, use frecuencias relativas para aproximar probabilidades y determine:
a) P(a favor)
b) P(ha disparado un arma de fuego y está en contra)
c) P(en contra | nunca ha disparado un arma)
d) P(ha disparado un arma y está a favor)
e) P(ha disparado un arma)

Un paso más allá

18. Sea $P(E) = 0.2$ y $P(F) = 0.3$. Responda si puede ser cierto en cada una de las preguntas siguientes; cuando sea apropiado, dé un ejemplo.
a) ¿$P(E \cup F) = 0.5$?
b) ¿$P(E \cup F) = 0.7$?
c) ¿$P(E \cup F) = 0.4$?
d) ¿$P(E \cap F) = 0.2$?
e) ¿$P(E \cap F) = 0.3$?
f) ¿$P(E \cap F) = 0.1$?
g) ¿$P(E \cap F) = 0.4$?

19. Durante una clase, un maestro de historia afirmó que la probabilidad de que tanto Israel como Siria manden representantes a una conferencia de paz es 0.8; después, en la misma clase, afirmó que la probabilidad de que Siria mande a un representante es 0.5. ¿Cree usted ambas afirmaciones? Explique.

20. Si $P(E) = 0.2$, $P(F) = 0.3$ y $P(E \mid F) = 0.5$, ordene los eventos siguientes de acuerdo con una probabilidad creciente: E, F, $\overline{E}$, $(E \cap F)$, $(E \cup F)$, $(E \cup \overline{E})$ y $(\overline{F} \cap F)$.

21. Dé argumentos convincentes para estos dos hechos e ilustre cada uno utilizando un ejemplo.
a) Si $P(E)<P(F)$, entonces $P(E \cup F)$ es al menos tan grande como $P(E)$.
b) Si $P(E)<P(F)$, entonces $P(E \cap F)$ es cuando mucho, tan grande como $P(E)$.

22. Determine una fórmula para $P(E \cap \overline{F})$ que no utilice probabilidad condicional. (*Sugerencia:* dibuje un diagrama de Venn.)

23. Refiérase al motivador 5. Determine la probabilidad de que una persona:
a) que se haga la prueba del SIDA tenga SIDA.

b) tenga SIDA y dé una prueba positiva.
c) tenga SIDA y dé una prueba negativa.
d) si obtuvo un resultado positivo en la prueba, tenga SIDA.
e) que obtuvo un resultado negativo en la prueba tenga SIDA.

24. Refiérase al motivador 5. Suponga que la probabilidad de tener SIDA para la población de alto riesgo es 0.003. De 25,000 muestras de sangre de esta población,
a) ¿cuántos resultados positivos falsos esperaría usted?
b) ¿cuántos resultados negativos falsos calcularía?
c) ¿cuántas personas con SIDA esperaría usted?
d) ¿cuántas personas con SIDA que den un resultado positivo cabría esperar?

SECCIÓN 5.6 *Eventos independientes*

Si E y F son eventos tales que la ocurrencia de F no influye en forma alguna en la de E, entonces E y F se llaman **eventos independientes**. Dicho de otra forma, E y F son eventos independientes si la probabilidad de que ocurra E dado que ha sucedido el evento F, es igual a la probabilidad del suceso del evento E.

E y F son eventos independientes si

$$P(E \mid F) = P(E) \tag{5.14}$$

Si dos eventos no son independientes, entonces son **eventos dependientes**.

Como una consecuencia de la regla (5.12), tenemos también:

$$P(E \cap F) = P(F)P(E \mid F)$$

Si E y F son eventos independientes, entonces:

$$P(F)P(E \mid F) = P(F)P(E)$$

En consecuencia, tenemos la siguiente **regla de multiplicación** para eventos independientes.

Regla de multiplicación

$$P(E \cap F) = P(E)P(F) \tag{5.15}$$

Note que, como regla general, el muestro con remplazo asegura que dos eventos serán independientes, mientras que el muestreo sin remplazo produce dos eventos que serán dependientes.

APLICACIÓN 5.39 ¿Cuáles de los pares de eventos siguientes son independientes?

a) E = sacar un "sol" al lanzar un peso.
F = sacar un "sol" al lanzar una moneda de diez centavos.

b) E = el primer bebé de María será un niño.
F = el segundo bebé de María será una niña.

c) E = hoy lloverá en Frostburg.
F = Juan reprobará hoy su examen de matemáticas 101.

Solución:

a) Como los eventos no están relacionados, son independientes.
b) Los dos eventos son independientes.
c) Estos dos eventos dan toda la impresión de no estar relacionados; en consecuencia, son eventos independientes. ▪

APLICACIÓN 5.40

Se sacan dos pelotas de una bolsa que contiene tres blancas y dos negras. Denotemos por B_1 el evento de sacar una pelota blanca en la primera oportunidad y por N_2 el de sacar una pelota negra en la segunda oportunidad.

a) Si las pelotas se sacan con remplazo, determine si B_1 y N_2 son eventos independientes.
b) Si las pelotas se sacan sin remplazo, determine si B_1 y N_2 son eventos independientes.

Solución: Usaremos la expresión 5.14 para determinar si los eventos son independientes.

a) Como la primera pelota se devuelve a la bolsa después de sacarla y hay dos pelotas negras dentro, $P(N_1 | B_1) = 2/5 = P(N_2)$. En consecuencia, los eventos son independientes.
b) La probabilidad de elegir una pelota negra en la segunda selección dado que en la primera se eligió una pelota blanca es $P(N_2 \mid B_1) = 2/4 = 1/2$, porque hay cuatro pelotas en la bolsa y dos de ellas son negras. Hay dos posibilidades de elegir una pelota negra en la segunda oportunidad: la primera fue blanca o negra.

   Por tanto,

$$P(N_2) = P[(B_1 \cap N_2) \cup (N_1 \cap N_2)]$$

   Como los eventos $(B_1 \cap N_2)$ y $(N_1 \cap N_2)$ son mutuamente excluyentes, tenemos:

$$\begin{aligned} P[(B_1 \cap N_2) \cup (N_1 \cap N_2)] &= P(B_1 \cap N_2) + P(N_1 \cap N_2) \\ &= P(N_2 | B_1)P(B_1) + P(N_2 | N_1)P(N_1) \\ &= \left(\frac{2}{4}\right)\left(\frac{3}{5}\right) + \left(\frac{1}{4}\right)\left(\frac{2}{5}\right) = \frac{8}{20} = \frac{2}{5} \end{aligned}$$

   La probabilidad de elegir una pelota negra en la segunda oportunidad dado que en la primera se eligió una blanca es $P(N_2 | B_1) = 2/4 = 1/2$. Puesto que $2/5 \neq 1/2$, los dos eventos no son independientes. ▪

APLICACIÓN 5.41

Si se lanzan dos monedas, encuentre $P(\text{SA})$.

Solución: La probabilidad $P(\text{SA})$ es la de obtener "sol" en la primera moneda y águila en la segunda. Como obtener un "sol" en una moneda y un "águila" en otra son eventos independientes, usando la regla (5.15) tenemos:

$$P(\text{SA}) = P(\text{S})\,P(\text{A})$$
$$= \frac{1}{2}\frac{1}{2} = \frac{1}{4} \quad \blacksquare$$

APLICACIÓN 5.42

Refiérase al problema ilustrado en la aplicación 5.35. Si se elige a un estudiante y E es el evento de que tenga escasa habilidad en matemáticas y F es el evento de que tenga mucho interés, ¿son independientes los eventos E y F?

Solución: No. La probabilidad de E dado F es:

$$P(E \mid F) = \frac{12}{55}$$

y

$$P(E) = \frac{60}{150} = \frac{2}{5}$$

Como una consecuencia de la regla (5.14), E y F no son eventos independientes, pues

$$P(E|F) \neq P(E)$$

Esto es,

$$\frac{12}{55} \neq \frac{2}{5} \quad \blacksquare$$

Se debe tener cuidado de no confundir los conceptos de eventos mutuamente excluyentes y eventos independientes; no existe una relación general entre los dos tipos de eventos. Los eventos pueden ser mutuamente excluyentes sin ser independientes, o viceversa (véase los ejercicios 13-16).

GRUPO DE EJERCICIOS 5.6

Habilidades básicas

1. Sean $P(E) = 0.5$, $P(F) = 0.6$, y $P(E \cap F) = 0.1$.
   a) ¿Son eventos independientes E y F? ¿Por qué?
   b) ¿Son eventos mutuamente excluyentes? ¿Por qué?
2. Sean $P(E) = 0.3$, $P(F) = 0.2$ y $P(E \cap F) = 0.06$.
   a) ¿Son eventos independientes E y F? ¿Por qué
   b) ¿Son eventos mutuamente excluyentes E y F? ¿Por qué?
3. Si los eventos E y F son eventos mutuamente excluyentes con $P(E) = 0.2$ y $P(F) = 0.5$, encuentre:
   a) $P(E \cup F)$.
   b) $P(E|F)$.
4. Si los eventos E y F son eventos mutuamente excluyentes con $P(E) = 0.3$ y $P(F) = 0.4$, encuentre:
   a) $P(E \cup F)$.
   b) $P(E \mid F)$.
5. Si E y F son eventos tales que $P(E) = 0.4$, $P(F) = 0.3$ y $P(E \mid F) = 0.4$, ¿son eventos independientes E y F? Explique.

6. Si E y F son eventos tales que $P(E) = 0.3$, $P(F) = 0.2$ y $P(E|F) = 0.3$, ¿son eventos independientes E y F. Explique.

Más aplicaciones

7. Si dos cartas se sacan con remplazo de un paquete de 52, encuentre la probabilidad de que:
   a) la segunda carta sea un corazón, dado que la primera es un corazón;
   b) ambas cartas sean corazones;
   c) la segunda sea negra, si la primera fue una espada;
   d) la segunda sea una figura, cuando la primera es una sota;
   e) la primera carta sea un basto o la segunda un as.

8. Dos canicas se sacan con remplazo de una vasija con cuatro canicas negras y seis blancas. Encuentre las probabilidades:
   a) ambas resultan blancas;
   b) las dos son negras;
   c) la segunda es blanca, dado que la primera fue negra;
   d) la primera es negra y la segunda blanca;
   e) una sale negra y la otra blanca.

9. La tabla adjunta muestra las frecuencias para el daltonismo, donde D representa el evento de que una persona sea daltónica, $\overline{D}$ el evento de cuando no lo sea, H significa que una persona sea hombre y $\overline{H}$ que sea mujer.

	H	$\overline{H}$
D	0.042	0.007
$\overline{D}$	0.485	0.466

   ¿Son eventos independientes D y H?

10. Si tres monedas se lanzan juntas encuentre:
    a) P(SSS).
    b) P(ASA) (nota: ASA significa obtener un “águila” para la primera moneda, un “sol” para la segunda y un “águila” para la tercera.
    c) P(exactamente un “sol”).
    d) P(al menos un “sol”).
    e) P(al menos dos “soles”).

11. Los principiantes de una pequeña universidad se clasifican de acuerdo con su promedio (GPA) en el bachillerato y la calificación en el SAT en matemáticas. Los resultados se resumen en la tabla siguiente:

	GPA en el bachillerato		
Calificación en el SAT	Bajo	Medio	Alto
Baja	50	30	50
Media	20	30	20
Alta	30	40	30

Se elige a un estudiante al azar. Si E es el evento de que el estudiante tenga un GPA medio en el bachillerato y F es que tenga una calificación baja en matemáticas en el SAT, ¿son eventos independientes E y F? Explique.

12. Los trabajadores de una planta industrial pequeña se clasifican de acuerdo con su religión y sexo; los resultados están en la tabla; si E es el evento de que una persona sea hombre y F de que sea judío, ¿son eventos independientes E y F?

	Sexo	
Religión	Hombre	Mujer
Protestante	30	20
Católico	45	30
Judío	45	30
Otros	7	8

Un paso más allá

13. Suponga que se lanza una moneda una vez. Sean $E = \varnothing$ y $F = \{S\}$.
    a) ¿Son eventos independientes E y F? Explique.
    b) ¿Son esos eventos mutuamente excluyentes? ¿Por qué?

14. Se lanzan dos monedas. Representemos por E el evento de obtener un “sol” en la primera moneda y por F el de obtener dos “soles” o dos águilas. Demuestre que:
    a) E y F son eventos independientes.
    b) E y F no son eventos mutuamente excluyentes.

15. Suponga que una moneda se lanza dos veces. Sean $E = \{SS\}$ y $F = \{AA\}$.
    a) ¿Son eventos mutuamente excluyentes E y F? Explique.
    b) Son eventos independientes? ¿Por qué?

16. Suponga que dos dados se lanzan una vez y sean $E = \{(1, 2,), (2, 1)\}$ y $F = \{(1, 2), (3, 4)\}$.
    a) ¿Son eventos independientes E y F? Explique.
    b) Son eventos mutuamente excluyentes E y F? Explique.

17. Si E y F son eventos mutuamente excluyentes, ¿cuál es el valor de $P(E \mid F)$?

18. Suponga que E y F son eventos para los cuales $P(E) > 0$ y $P(F) > 0$. Si E y F son eventos independientes, ¿pueden ser eventos mutuamente excluyentes?

19. Un cierto proceso de fabricación produce un 5% de partes defectuosas; el veinte por ciento de todas las partes se producen en la máquina A; hay un 10% de

probabilidades de que una parte sea defectuosa dado que fue producida por la máquina A.

a) ¿Cuál es la probabilidad de que una pieza que se pruebe resulte defectuosa y la haya producido la máquina A?

b) ¿Cuál es la probabilidad de que una parte provenga de la máquina A dado que es defectuosa?

20. Suponga que un espacio muestral es igual al evento $(E_1 \cup E_2 \cup E_3)$ y que E_1, E_2 y E_3 no tienen resultados comunes (los tres eventos se llaman *mutuamente excluyentes* y *exhaustivos*). Demuestre que para cualquier otro evento B,

$$P(B) = \Sigma[P(B \mid E_i)\, P(E_i)]$$

Este resultado se llama la *ley de probabilidad total.* (*Sugerencia*: véase la aplicación 5.42.)

21. En el estacionamiento de un gran aeropuerto, 60 de los coches están fabricados en Estados Unidos y el 10% de ellos son coches deportivos; 20% están fabricados en Japón y de ésos un 10% son deportivos; finalmente, 20% son europeos y de ellos 20% son deportivos. Si se elige un coche al azar, ¿cuál es la probabilidad de que resulte un modelo deportivo?

22. Sean E_1, E_2 y E_3 eventos mutuamente excluyentes y exhaustivos, y suponga que A es un evento tal que $P(A) > 0$. Entonces, para cualquier evento E_k ($k = 1, 2$ o 3), demuestre que:

$$P(E_k \mid A) = \frac{P(A \mid E_k)\, P(E_k)}{\Sigma\, [P(A \mid E_i)\, P(E_i)]}$$

Este resultado se conoce como el *teorema de Bayes.* (*Sugerencia*: use la ley de probabilidad total del ejercicio 20.)

23. Tres cajas contienen pelotas rojas, blancas y azules. La caja 1 tiene siete pelotas rojas, una blanca y cinco azules; la caja 2 tres pelotas rojas, cuatro blancas y dos azules; y la 3 contiene una pelota roja, seis pelotas blancas y dos azules; se toma una caja al azar y se saca de ella una pelota aleatoriamente. Dado que la pelota es roja, ¿cuál es la probabilidad de que provenga de la caja 3?

24. Suponga que participa en un espectáculo de concurso y se le da a elegir entre tres puertas. Tras una de ellas hay un coche y tras las otras hay cabras; usted elige la puerta número 1 y el conductor del espectáculo, conocedor de lo que hay tras las tres puertas, abre la número 3, que tiene una cabra; entonces él pregunta si quiere usted abrir la puerta número 2.

a) ¿Debe usted cambiar la elección?

b) ¿Cuál es la probabilidad de ganar el coche si lo hace?

c) ¿Cuál es la probabilidad de ganar el coche si no lo hace?

SECCIÓN 5.7 *Variables aleatorias*

Es conveniente para el trabajo futuro, saber relacionar los resultados de un experimento con números reales; ya que cuando los resultados de un experimento se pueden asociar con números reales, son más fáciles de analizar. Desgraciadamente, no todos los experimentos dan como resultado números reales. Suponga que está usted trabajando para un fabricante de microcomputadoras en el departamento de control de calidad, donde se le ha asignado la tarea de revisar cuatro computadoras elegidas al azar del último lote producido y clasificarlas como defectuosa (D) o no defectuosa (N). Los resultados de su labor son el siguiente espacio muestral de datos cualitativos:

DNDN DDND NDND NNDN
DNND DDDN NDDN NNND
DNDD DDNN NDDD NNDD
DNNN DDDD NDNN NNNN

Estos resultados no son números reales, pero si cada uno se asocia con el número de microcomputadoras defectuosas, podemos asociar un único número real a cada resultado. Por ejemplo, al resultado DNDN se le puede asignar el número 2, igual que al DNND, al NDNN se le da el número 1 y al resultado NNNN se le puede asociar el número 0; al hecho de asociar los resultados de un espacio muestral de un experimento con números reales únicos se le llama **variable aleatoria.** La variable aleatoria en el ejemplo de control de calidad es el número de microcomputadoras defectuosas en el paquete de cuatro y tiene los cinco valores posibles: 0, 1, 2, 3 y 4.

Variable aleatoria

Una *variable aleatoria* es una regla, o fución, que asigna un único número real a cada resultado de un espacio muestral en un experimento. Los siguientes son otros ejemplos de variables aleatorias:
ejemplos de variables aleatorias.

EJEMPLO 5.20

Considere los siguientes experimentos y las correspondientes variables aleatorias.

Experimento	Variable aleatoria
1. Lanzar cinco monedas.	Número de "soles" obtenidos.
2. Observar a los clientes de un barco durante una hora.	Número de clientes.
3. Comprar 12 computadoras.	Cantidad de computadoras defectuosas.
4. Pesar a una persona.	Peso en libras.
5. Golpear una pelota de golf.	Distancia recorrida.
6. Revisar diez impuestos a las ganancias.	Impuestos con error.
7. Revisar diez tasas de línea 33.	Número de errores en la impuestos.
8. Observar el trabajo de un empleado durante en trabajo ocho horas.	Tiempo utilizado por el empleado improductivo.
9. Ecuestar a diez personas de sobre la oportunidad de que sea elegido el candidato A.	Proporción en favor A.
10. Dar la pastilla A a 50 personas con dolor de cabeza.	Número de curaciones.

Los valores de una variable aleatoria suelen identificarse con ella; en general, esto no causa confusión y tiene sus ventajas. Para el experimento 1 propuesto anteriormente, si denotamos por X el número de soles obtenidos al lanzar cinco monedas, los valores posibles de la variables aleatoria son 0, 1, 2, 3, 4 y 5; representamos lo anterior por $X = \{0, 1, 2, 3, 4, 5\}$. Si se escribe con propiedad, esto es incorrecto porque X es la regla, función, y $\{0, 1, 2, 3, 4, 5\}$ el conjunto de sus valores. A lo largo de esta sección, las letras mayúsculas como X, Y y Z denotarán variables aleatorias y las letras minúsculas como x, y y z sus valores, elementos de la imagen.

Si es posible contar los valores de una variable aleatoria, ésta se denomina **variable aleatoria discreta**; si los valores no se pueden contar, a la variable se le llama **variable aleatoria continua**. Los experimentos 4, 5 y 8 de la lista anterior presentan variables aleatorias continuas, mientras que las restantes son discretas. El número de unidades vendidas, de computadoras defectuosas y cualquier otra variable que cuente, son variables discretas; cualquier variable aleatoria que mida, como el peso, la temperatura y la distancia, es una variable continua; el experimento 8 se refiere a una variable continua porque el número de horas de trabajo improductivo puede ser cualquier valor en el intervalo de tiempo 0-8.

Distribuciones de probabilidad

Es posible asociar probabilidades a los valores de una variables aleatoria discreta. Si X es una variable aleatoria discreta y estamos interesados en determinar la probabilidad de los resultados asociados con k, un valor de la variable aleatoria X, escribimos $P(X = k)$; cuando no hay posibilidad de confusión escribimos $P(X = k)$ como $P(k)$.

EJEMPLO 5.21

La variable aleatoria X representa el número de "soles" obtenidos al lanzar dos monedas; los valores posibles son 0, 1 y 2. Entonces

$$P(X=0) = P(0) = P(\text{AA}) = \frac{1}{4}$$

$$P(X=1) = P(1) = P(\text{SA, AS}) = \frac{1}{2}$$

$$P(X=2) = P(2) = P(\text{SS}) = \frac{1}{4}$$

Una *tabla de distribución de probabilidad* o **tabla de probabilidades** para una variable aleatoria discreta es una tabla de dos columnas; una columna representa los valores de la variable aleatoria y la otra las probabilidades asociadas. La tabla 5.6 deja ver la distribución de probabilidad para el ejemplo 5.21.

TABLA 5.6
Distribución de probabilidad del ejemplo 5.21

x	$P(x)$
0	1/4
1	1/2
2	1/4
	1

Note que la suma de la columna de probabilidades siempre debe ser 1.

Funciones de probabilidad

A veces es conveniente expresar la relación entre los valores de una variable aleatoria y sus probabilidades asociadas, en términos de una regla; esa regla se llama **función de probabilidad** o función de distribución de probabilidad. Considere la aplicación 5.43.

APLICACIÓN 5.43

Suponga que cinco pelotas del mismo tamaño se numeran del 1 al 5 y se colocan en una bolsa; un experimento pide seleccionar una pelota al azar. Denotemos por la variable aleatoria X el valor de la pelota seleccionada.

a) Construya una tabla de probabilidad para X.

b) Encuentre la función de probabilidad para X y llámela g.

Solución: Las cinco pelotas tienen la misma probabilidad de ser elegidas. La siguiente es la tabla de probabilidad para la variable aleatoria discreta X.

a)

x	$P(x)$
1	1/5
2	1/5
3	1/5
4	1/5
5	1/5

b) La función que relaciona los valores de x con las probabilidades es g, definida por $g(x) = 1/5$; $x = 1, 2, 3, 4, 5$. Para cualquier valor de x no contenido en el conjunto $\{1, 2, 3, 4, 5\}$, $g(x)$ se define como cero. ∎

Gráficas de probabilidad

Se puede construir gráficas para funciones de probabilidad. Cuando la variable aleatoria es discreta, la gráfica de la función de probabilidad puede construirse usando segmentos de rectas verticales. Los valores de la variable aleatoria se localizan en el eje horizontal y las probabilidades en el eje vertical; en cada valor se construye un segmento de recta vertical de altura igual a la probabilidad de la variable aleatoria (véase el ejemplo 5.22). Advierta que la suma de las longitudes de los segmentos verticales debe ser igual a 1. Por otro lado, para representar probabilidades de variables aleatorias continuas usamos áreas en vez de rectas verticales, como en el ejemplo 5.23.

EJEMPLO 5.22

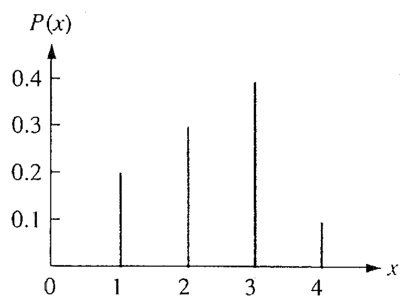
Suponga que la tabla siguiente corresponde a una variable aleatoria discreta X.

x	$P(x)$
1	0.2
2	0.3
3	0.4
4	0.1

La gráfica de la función de probabilidad correspondiente se muestra en la figura 5.12. Note que la suma de las longitudes de los segmentos verticales es 1.

FIGURA 5.12

Gráfica de la función de probabilidad para una variable aleatoria discreta



EJEMPLO 5.23

Suponga que los valores de la variable aleatoria X son todos números reales entre 0 y 1 y que la función de probabilidad de X es f, definida por:

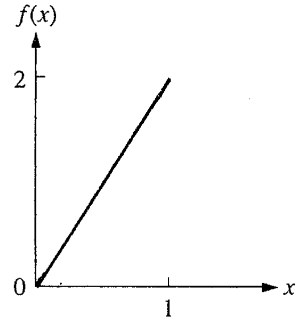
$$f(x) = \begin{cases} 2x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{para cualquier otro } x \end{cases}$$

Una gráfica de la función de probabilidad f se ve en la figura 5.13; el área de la región acotada por la recta $f(x) = 2x$, el eje x y las rectas $x = 0$ y $x = 1$. La región es un triángulo rectángulo cuya área se encuentra como sigue:

$$A = \left(\frac{1}{2}\right)(\text{base})(\text{altura})$$
$$= \left(\frac{1}{2}\right)(1)(2) = 1$$

FIGURA 5.13

Gráfica de la función de probabilidad para una variable aleatoria continua



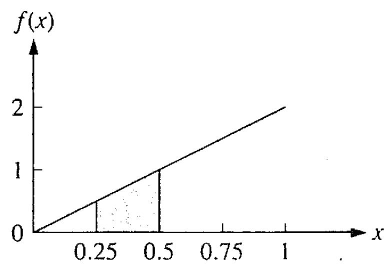
Como las probabilidades de variables aleatorias continuas se asocian con áreas, se pueden encontrar sólo para intervalos como $a < x < b$. La probabilidad de cualquier valor particular de x para una variable aleatoria continua debe ser necesariamente igual a 0, porque un segmento de recta vertical sobre el valor x y de altura igual al valor de la función de probabilidad debe tener área 0; en consecuencia, si X es una variable aleatoria continua, $P(a < X < b) = P(a \leq X \leq b)$. Un estudio completo de variables aleatorias continuas involucra cálculo y se encuentra fuera del alcance de este libro; por ello, después del ejemplo 5.24 y para el resto de esta sección, trataremos únicamente variables aleatorias discretas y sus probabilidades.

EJEMPLO 5.24

Para la variable aleatoria continua de la figura 5.13, suponga que queremos encontrar la probabilidad de que el valor de la X esté entre 0.25 y 0.5; esto es, $P(0.25 \leq x \leq 0.5)$; esta probabilidad es igual al área bajo la recta cuya ecuación es $f(x) = 2x$ acotada por el eje x y las rectas $x = 0.25$ y $x = 0.5$, como lo ilustra la figura 5.14. La figura formada es un trapezoide; recuerde que el área de un trapezoide está dada por $A = h(b_1 + b_2)/2$, donde h es la altura y b_1 y b_2 son las bases. Esta fórmula da el área o probabilidad para la región sombreada.

FIGURA 5.14

$P(0.25 \leq X \leq 0.5)$



$$A = \frac{(0.25)(0.5+1)}{2}$$
$$= \frac{(0.25)\,(1.5)}{2} = 0.1875$$

APLICACIÓN 5.44

Considere la siguiente función de probabilidad para una variable aleatoria discreta X.

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

a) ¿Cuáles son los valores de la variable aleatoria X?
b) ¿Cuál probabilidad está asociada con cada valor de X?
c) Construya una tabla de probabilidad para X.
d) Trace una gráfica de probabilidad para X.

Solución:

a) Los valores de X son 1, 2, 3 y 4.
b) La probabilidad asociada con $x = 1$ es:

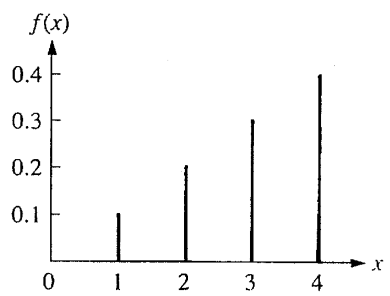
$$P(X = 1) = f(1) = \frac{1}{10} = 0.1$$

Análogamente, las probabilidades para $x = 2$, 3 y 4 resultan ser 0.2, 0.3 y 0.4, respectivamente.

c) Una tabla de probabilidad es como sigue:

x	$P(x) = f(x)$
1	0.1
2	0.2
3	0.3
4	0.4
	1

d)



■

Media de una variable aleatoria discreta

Recuerde que una fórmula para calcular el valor de la media poblacional μ es:

$$\mu = \frac{\Sigma\,(fx)}{N}$$

donde f es la frecuencia de una medida particular x y N el tamaño de la población. Esta fórmula puede reescribirse como:

$$\mu = \Sigma\left[x\frac{f}{N}\right]$$

Como la frecuencia relativa f/N representa la probabilidad de que ocurra x, $P(x)$, la media poblacional puede escribirse como $\mu = \Sigma[x \cdot P(x)]$. Como consecuencia de estas observaciones, si X es una variable aleatoria, definimos la **media** de X como sigue:

Media de una variable aleatoria X

$$\mu_x = \Sigma[xP(x)] \tag{5.16}$$

donde μ_x representa la media de la variable aleatoria X.

EJEMPLO 5.25

Una planta industrial grande realiza una campaña para promover el uso compartido del automóvil entre sus empleados; los datos en la tabla 5.7 se registraron entre todos los empleados de la planta para conocer los efectos de la campaña.

TABLA 5.7
Datos de uso compartido del automóvil

Núm. de ocupantes (x) por automóvil	f	xf	Frecuencia relativa f/N
1	425	425	0.450
2	235	470	0.249
3	205	615	0.217
4	52	208	0.055
5	22	110	0.023
6	6	36	0.006
	945	1864	1.000

La media poblacional es $\mu = \frac{\Sigma(fx)}{N} = \frac{1864}{945} = 1.97$

Ahora escojamos un coche al azar que transporte empleados al trabajo y contemos el número de ocupantes; este número representa una variable aleatoria X, cuyos valores son 1, 2, 3, 4, 5 y 6, con las probabilidades 0.45, 0.249, 0.217, 0.555, 0.023 y 0.006 respectivamente. La media de esta variable aleatoria es entonces:

$$\begin{aligned}\mu_x &= \Sigma(xP(x)) = (1)P(1) + (2)P(2) + (3)P(3) + (4)P(4) + (5)P(5) + (6)P(6) \\ &= (1)(0.450) + (2)(0.249) + (3)(0.217) \\ &\quad + (4)(0.055) + (5)(0.023) + (6)(0.006) \\ &= 1.97\end{aligned}$$

Vea que esto concuerda con el valor calculado anteriormente.

La media de una variable aleatoria X se llama también el **valor esperado** de X y se denota por $E(X)$; en consecuencia, tenemos los resultados siguientes para una variable aleatoria discreta:

Media de una variable aleatoria discreta X

$$\mu_X = E(X) = \Sigma(xP(x))$$

También existen fórmulas análogas para el caso de variables aleatorias continuas.

APLICACIÓN 5.45

Un gran hotel ha encontrado que el número de unidades de aire acondicionado que debe remplazar cada verano tiene la tabla de probabilidad siguiente.

Núm. remplazado	Probabilidad
0	0.35
1	0.30
2	0.20
3	0.10
4	0.05

Encuentre el número esperado de acondicionadores de aire que debe remplazarse cada verano e interprete su respuesta.

Solución: Denotemos por X el número de unidades remplazadas; calcularemos el producto de x por $P(x)$ para cada valor de X. Los resultados están organizados en la tabla 5.8, de donde podemos ver que $E(X) = \Sigma(xP(x)) = 1.20$. Deduciéndose que si se registra por bastantes años el número de unidades de aire acondicionado que debe sustituirse, el promedio por año debe ser 1.20.

TABLA 5.8

Cálculo de $E(X)$ para la aplicación 5.45

x	$P(x)$	$xP(x)$
0	0.35	0
1	0.30	0.30
2	0.20	0.40
3	0.10	0.30
4	0.05	0.20
		1.20

■

Varianza de una variable aleatoria discreta

Recuerde del capítulo 3 que la varianza poblacional se puede definir como:

$$\sigma^2 = \frac{SS}{N} = \frac{\Sigma f(x-\mu)^2}{N}$$

donde f es la frecuencia asociada con la medida x y N es el tamaño de la población. Esto puede reescribirse:

$$\sigma^2 = \Sigma\left[(x-\mu)^2 \frac{f}{N}\right]$$

De nuevo, si representamos por f/N la probabilidad $P(x)$ de ocurrencia de x, entonces tenemos la fórmula siguiente para la **varianza de una variable aleatoria** X.

Varianza de una variable aleatoria X.

$$\sigma_x^2 = \Sigma[(x-\mu)^2 P(x)] \tag{5.17}$$

donde σ_x^2 es la varianza de la variable aleatoria X y μ es la media de X.

EJEMPLO 5.26

Para los datos del uso compartido del automóvil presentados en la tabla 5.7, la media poblacional es $\mu = 1.97$ y la varianza poblacional es:

$$\begin{aligned}\sigma^2 &= \frac{\Sigma[(x-\mu)^2 f]}{N}\\ &= [(1-1.97)^2(425) + (2-1.97)^2(235) + (3-1.97)^2(205) + (4-1.97)^2(52)\\ &\quad + (5-1.97)^2(22) + (6-1.97)^2(6)]/945\\ &= 1.197\end{aligned}$$

Estos cálculos pueden reescribirse como:

$$\begin{aligned}\sigma^2 &= (1-1.97)^2\frac{425}{945} + (2-1.97)^2\frac{235}{945} + (3-1.97)^2\frac{205}{945} + (4-1.97)^2\frac{52}{945}\\ &\quad + (5-1.97)^2\frac{2}{945} + (6-1.97)^2\frac{6}{945}\\ &= 1.197\\ &= (1-1.97)^2P(1) + (2-1.97)^2P(2) + (3-1.97)^2P(3)\\ &\quad + (4-1.97)^2P(4) + (5-1.97)^2P(5) + (6-1.97)^2P(6)\end{aligned}$$

Mas esto es precisamente un ejemplo de la fórmula (5.17).

Desviación estándar de una variable aleatoria discreta

La **desviación estándar de una variable aleatoria** se define como la raíz cuadrada de la varianza y se denota por σ.

Desviación estándar de una variable aleatoria X

$$\sigma_x = \sqrt{\text{varianza}}$$

APLICACIÓN 5.46

Para la variable aleatoria X de la aplicación 5.45, encuentre la varianza σ_x^2 y la desviación estándar σ_x.

Solución: En la aplicación 5.45, se encontró que μ es 1.2. Organizamos nuestros cálculos para σ_x^2 en la tabla siguiente.

x	$P(x)$	$x-\mu$	$(x-\mu)^2$	$(x-\mu)^2P(x)$
0	0.35	−1.2	1.44	0.504
1	0.30	−0.2	0.04	0.012
2	0.20	0.8	0.64	0.128
3	0.10	1.8	3.24	0.324
4	0.05	2.8	7.84	0.392
				1.36

Por tanto, $\sigma_x^2 = 1.36$ y $\sigma_x = \sqrt{1.36} = 1.17$. ▪

Para encontrar la varianza σ_x^2 de una variable aleatoria X es más sencillo realizar los cálculos con la siguiente fórmula que con la fórmula (5.17), y es matemáticamente equivalente.

Fórmula para calcular σ_x^2

$$\sigma_x^2 = \Sigma\,[x^2P(x)] - \mu_x^2 \qquad (5.18)$$

EJEMPLO 5.27

Usemos la fórmula (5.18) para encontrar σ_x^2 para la variable aleatoria X ilustrada en la aplicación 5.45; los cálculos se organizan en esta tabla:

x	x^2	$P(x)$	$x^2P(x)$
0	0	0.35	0
1	1	0.30	0.30
2	4	0.20	0.80
3	9	0.10	0.90
4	16	0.05	0.80
			2.80

Encontramos previamente que la media μ_x es $\mu_x = 1.2$. Así, la varianza de la variable aleatoria X es:

$$\begin{aligned}\sigma_x^2 &= \Sigma\,[x^2P(x)] - \mu_x^2 \\ &= 2.80 - (1.2)^2 = 1.36\end{aligned}$$

Note que esto concuerda con el resultado obtenido en la aplicación 5.46.

GRUPO DE EJERCICIOS 5.7

Habilidades básicas

1. Suponga que un experimento consiste en lanzar una moneda cuatro veces y la variable aleatoria X denota el número de "soles" obtenidos.
   a) Construya una tabla de probabilidad para X.
   b) ¿Es X una variable aleatoria discreta o continua?
   c) Trace una gráfica de probabilidad para X.
2. Un experimento consiste en lanzar dos dados juntos y la variable aleatoria Y significa la suma de los números que quedan hacia arriba.

a) Construya una tabla de probabilidad para Y.
b) ¿Es Y una variable discreta o continua?
c) Trace una gráfica de probabilidad para Y.

3. Dada la función de probabilidad f definida por

$$f(x) = \frac{5-x}{10} \quad \text{para } x = 1, 2, 3, 4$$

para una variable aleatoria x, encuentre la media y la desviación estándar.

4. Dada la función de probabilidad g definida por $g(y) = 0.2$ si $y = 2, 3, 4, 5, 6$ para una variable aleatoria Y, encuentre la media y la desviación estándar de Y.

5. Representemos por x el número de "soles" obtenidos al lanzar tres monedas. Encuentre μ_x y σ_x.

6. Dada la función de probabilidad h definida por:

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

Para una variable aleatoria x, encuentre μ_x y σ_x.

7. Suponga que X es una variable aleatoria discreta con función de probabilidad f definida por:

$$f(x) = \frac{x}{55} \quad \text{para } x = 1, 2, 3, 4, ..., 10.$$

a) Demuestre que $\Sigma P(x) = 1$, y encuentre:
b) $P(X = 4)$.
c) $P(X \leq 3)$.
d) $P(X \leq 9)$.
e) $P(2 \leq X \leq 4)$.
f) $P(2 < X < 4)$.
g) μ_x.
h) σ_x.

8. Sea X una variable aleatoria discreta con función de probabilidad f definida por:

$$f(x) = \frac{x}{15} \quad \text{para } x = 1, 2, 3, 4, 5$$

a) Demuestre que $\Sigma P(x) = 1$, y encuentre:
b) $P(X = 4)$.
c) $P(X \geq 3)$.
d) $P(X \leq 5)$.
e) $P(2 \leq X \leq 4)$.
f) $P(2 < X < 4)$.
g) μ_x.
h) σ_x.

Más aplicaciones

9. La tabla de probabilidad para el número de llamadas telefónicas recibidas por el señor Pérez en un día, es:

x	$P(x)$
0	0.40
1	0.23
2	0.17
3	0.09
4	0.11

Encuentre:
a) $P(X = 1)$.
b) $P(1 < X \leq 3)$.
c) $P(X \geq 1)$.
d) $E(X)$.
e) σ_x^2.

10. El número X de platillos de pescado vendidos en una hora en un restaurante local se describe con la tabla de probabilidad siguiente:

x	0	1	2	3	4	5	6
$P(x)$	0.14	0.16	0.30	0.14	0.13	0.08	0.05

Construya una gráfica de probabilidad y encuentre:
a) $P(X \leq 1)$.
b) $P(X > 1)$.
c) $P(2 \leq X \leq 4)$.
d) $P(X \leq 5)$.
e) μ_x.
f) σ_x.

11. Con base en su historia, la tabla de frecuencias adjunta registra X, el número de automóviles vendidos por día por un distribuidor.

X	Núm. de días
0	44
1	87
2	128
3	234
4	297
5	155
6	30
7	25

a) Construya una tabla de probabilidad para la variable aleatoria X.
b) Encuentre la media de X.
c) Localice la desviación estándar de X.
d) Calcule $P(X \leq 5)$.
e) ¿Cuál es el número esperado de coches vendidos? Interprete el resultado.

12. Una rifa ofrece un primer premio de 1000 dólares, dos segundos premios de 500 y 20 premios de 20 dólares cada uno. Si se venden 10,000 boletos a medio dólar cada uno, encuentre:
   a) el beneficio neto esperado (ganancia) si se compra un boleto.
   b) la varianza del beneficio neto si se compra un boleto.

13. Un constructor está considerando realizar un trabajo que le traerá un beneficio de 25,000 dólares con una probabilidad de 0.8 o una pérdida, debido al mal tiempo, huelgas y demás, de 10,000 dólares con una probabilidad de 0.2.
   a) ¿Cuál es la ganancia esperada del constructor?
   b) ¿Cuál es la desviación estándar de la ganancia?

14. Para promover sus productos, la Pepsi-Cola Company anunció un juego de diversión Nintendo; se ofrecieron premios instantáneos a los consumidores afortunados que encontraran mensajes especiales en las tapas de las botellas de ciertos productos marcados. La información sobre los premios está contenida en la tabla adjunta.

Premio instantáneo	Oportunidades al menudeo	Valor estimado
2L de Pepsi	1:25.5	$1.49
Cupón de $ Mario Money	1:1,400	$5.00
Paquete de juego Nintendo	1:23,000	$30.00
Colección de videojuegos	1:70,000	$70.00
Juego Nintendo para niños	1:140,000	$90.00
Paquete de acción Nintendo	1:140,000	$100.00

Calcule el valor esperado al menudeo de un premio recibido por un consumidor que compra una botella de un producto Pepsi de la promoción.

15. Encuentre el número esperado de varones en una familia con cinco hijos; suponga que los nacimientos de hombres y de mujeres tienen la misma probabilidad de ocurrir.

Un paso más allá

16. Usted debe estar en la escuela en 20 minutos y hay dos rutas que puede tomar para llegar allí; el tiempo medio para llegar es de 12 y 16 minutos, respectivamente. ¿Es la mejor ruta para llegar la de 12 minutos? Explique.

17. El señor Baker quiere asegurar su casa por 80,000 dólares. La compañía de seguros estima en 0.004 la probabilidad de que ocurra una pérdida total; en 0.02 la de que sufra una pérdida del 50%, y en 0.08 que la pérdida sea del 25 por ciento. Si la aseguradora no pagará indemnización por otras pérdida parciales, ¿qué prima debe pagar el señor Baker cada año si la empresa de seguros quiere lograr una ganancia promedio de 500 dólares anuales en todas las pólizas de ese tipo?

18. Si se lanzan dos dados y se registra la suma de los números obtenidos, denotemos por X la variable aleatoria correspondiente a la suma. Encuentre μ y σ para X.

19. Si X es una variable aleatoria continua, con $P(X \leq 3) = 0.45$ y $P(X \geq 4) = 0.40$, encuentre:
   a) $P(X < 3)$.
   b) $P(4 < X)$.
   c) $P(3 \leq X \leq 4)$.

20. Si se extraen dos cartas con remplazo de un paquete común de 52 cartas, encuentre el número esperado de diamantes que puede salir.

21. Demuestre que $\sigma^2 = \sum[x^2P(x)] - \mu^2$.

22. El problema siguiente se conoce como la paradoja de San Petersburgo. Pedro acuerda lanzar una moneda hasta que salga "sol"; si sale en el primer lanzamiento, Pedro le paga a usted 1 dólar, si no, el convenio es darle a usted 2 dólares si sale "sol" en el segundo lanzamiento, 4 si sale "sol" en el tercero, 8 dólares si es en el cuarto, y así sucesivamente; el número de dólares que le pagará a usted se duplica con cada lanzamiento adicional. Para quedar a mano, ¿cuánto debe pagar usted por el juego? Explique.

RESUMEN DEL CAPÍTULO

En este capítulo se introdujeron los conceptos y principios de la probabilidad elemental; el teorema fundamental del conteo, permutaciones y combinaciones para ayudar a calcular las probabilidades difíciles; vimos que se pueden usar variables aleatorias para proporcionar descripciones numéricas de resultados experimentales; aprendimos cómo asociar distribuciones de probabilidad con variables aleatorias y cómo calcular medias, varianzas y desviaciones estándar de variables aleatorias discretas.