

Unidad 5

- Distribuciones probabilísticas discontinuas.

4

distribuciones probabilísticas discontinuas

Objetivos del capítulo

Después de estudiar este capítulo, se deberá estar en condiciones de:

1. Explicar qué es una distribución probabilística.
2. Definir una variable aleatoria y dar ejemplos.
3. Explicar en qué forma puede servir de modelo una distribución probabilística.
4. Enunciar los supuestos propios de la distribución binomial.
5. Establecer los supuestos relacionados con la distribución de Poisson.
6. Utilizar las tablas y fórmulas binomiales para obtener probabilidades.
7. Emplear las tablas y fórmulas de Poisson para calcular probabilidades.
8. Utilizar la distribución de Poisson para aproximar probabilidades binomiales.
9. Resolver problemas sencillos que comprendan probabilidades binomiales o de Poisson.

Sinopsis del capítulo

Introducción

Variables aleatorias

Valor esperado de una variable aleatoria

Sumas de variables aleatorias

Distribuciones probabilísticas

Distribuciones discontinuas

Distribución binomial

Fórmula binomial

Tablas binomiales

Probabilidades binomiales individuales

Tabla binomial acumulativa

Características de las distribuciones binomiales

Distribución de Poisson

Fórmula de Poisson

Aplicación que comprende el tiempo

Aplicación que comprende el área

Tablas de Poisson

Probabilidades individuales de Poisson

Tabla acumulativa de Poisson

La distribución de Poisson como una aproximación y la binomial

Otras distribuciones discretas.

INTRODUCCIÓN

¿POR QUE CUANDO se tira al aire una moneda, algunas veces cae cara y otras veces cruz? ¿Por qué cuando se arroja un dado, éste cae sobre una cara en lugar de caer sobre cualquier otra? Se dice que éstas y otras situaciones semejantes son regidas por el azar, pero, ¿qué es exactamente el azar?

La mejor explicación sería considerar el azar como la interacción de un gran número —quizá en extremo grande— de factores que influyen colectivamente en el resultado de un experimento o muestra. No parece aventurado suponer, en el caso del dado, que la fuerza con la que se arroja, las corrientes de aire, el ángulo que forma al caer, el “coeficiente de rebote” respecto de la superficie sobre la que cae, las veces que se haya agitado antes de lanzarse, etc., influyen en el resultado. Como sería virtualmente imposible controlar o regular todos estos factores, o predecir cómo interactúan en cualquier tirada para afectar el resultado, no se puede especificar con precisión cuál será el resultado que se obtendrá en determinado tiro. Además, la misma incapacidad para saber por anticipado qué resultado, de un conjunto posible, se obtendrá en cualquier intento es una característica inherente de todo proceso en el que interviene el azar, como en el caso de barajar naipes, sacar conejos de un sombrero o del *muestreo*.

Por otra parte, si se supone que los mismos factores interactúan de igual o semejante forma con respecto a muchas observaciones repetidas, se observa que existe *predicibilidad* “en el largo plazo”. En otras palabras, determinados resultados pueden ser más factibles que otros, lo cual sería más evidente, dado un número considerable de observaciones.

VARIABLES ALEATORIAS

La variable que tiene resultados o valores que tienden a variar de observación en observación debido a los factores relacionados con el azar, recibe el nombre de *variable aleatoria*. Es muy conveniente, dada su importancia práctica, definir una variable aleatoria asociada con una muestra o experimento, de tal manera que sus resultados posibles sean *numéricos*. Por ejemplo, el experimento de tirar la moneda una sola vez tiene dos posibles resultados, C y Cr*, los cuales no son numéricos. De otra manera, el “número de caras de una tirada” se consideraría como la variable aleatoria que tiene como posibles valores *numéricos* 0 y 1. Por la misma razón, nuestra variable aleatoria podría ser “el número de cruces en una tirada”. En el caso de una moneda que se tira

N. del S. Se adoptan estos símbolos para designar al anverso o *cara* y al reverso o *cruz*, respectivamente de una moneda.

dos veces al aire, la variable aleatoria podría ser “el número de *caras* en dos tiradas”, y tener como posibles resultados 0, 1, 2. Otra variable aleatoria podría ser el número de clientes que entran a una tienda de discos en un lapso de 20 minutos: 0, 1, 2, 3, 4,.... Aun más, otra variable aleatoria podría ser la estatura de los estudiantes de una clase en la Universidad, con una variedad continua de valores que quedaría comprendida entre 1.20 m y 2.10 m.

Una *variable aleatoria* es una función valorada numéricamente, cuyo valor está regida por factores en los que interviene el azar.

Las variables aleatorias pueden ser discretas o continuas.

Una variable aleatoria se considera *discreta* si los valores que asume se pueden *contar*.

Como ejemplos representativos de variables aleatorias discretas se encuentran el número de accidentes que ocurren durante una semana, número de defectos de los zapatos, cantidad de cosechas perdidas, número de terremotos, número de juegos perdidos por inasistencia y la cantidad de libros que hay en un estante.

Una variable aleatoria se considera *continua* si puede asumir *cualquier* valor dentro de un determinado intervalo.

Una variable continua tiene un número infinito de valores posibles. Como ejemplos típicos tenemos: peso de las cajas de naranjas, altura de los pinos, duración de una conversación telefónica y el tiempo que se requiere para llevar a cabo un examen de tipo ensayo.

La diferencia que existe entre variables aleatorias discretas y continuas es muy importante, debido a que los diferentes modelos de probabilidad (distribuciones) se utilizan según el tipo de variable aleatoria en consideración.

Valor esperado de una variable aleatoria

Si una variable aleatoria x asume los valores $x_1, x_2, x_3, \dots, x_n$, con las probabilidades correspondientes $p_1, p_2, p_3, \dots, p_n$, entonces el *valor esperado* de la variable aleatoria $E(x)$ es

$$p_1x_1 + p_2x_2 + p_3x_3 + \cdots + p_nx_n$$

Por tanto*

$$E(x) = \sum_{i=1}^n p_i x_i$$

Supóngase que una tienda de aparatos electrodomésticos ha reunido los siguientes datos sobre ventas de congeladores:

x_i Cantidad de congeladores vendidos	$P(x)$ Frecuencia relativa
0	0.20
1	0.30
2	0.30
3	0.15
4	0.05
	1.00

$$E(x) = 0.20(0) + 0.30(1) + 0.30(2) + 0.15(3) + 0.05(4) = 1.55$$

Como obviamente la tienda no puede vender en realidad 1.55 congeladores en un día determinado (dado que la cantidad vendida es una variable discreta que consta de los enteros 0, 1, 2, 3 y 4), la pregunta obvia es cómo interpretar dicha cifra. Es muy simple, el valor esperado es un promedio de largo plazo.

En forma semejante, si se tira un dado no cargado, ¿cuál es el valor esperado de una tirada? Hay seis resultados que tienen la misma probabilidad, y el valor esperado es

$$\frac{1}{6}(1) + \frac{1}{6}(2) + \frac{1}{6}(3) + \frac{1}{6}(4) + \frac{1}{6}(5) + \frac{1}{6}(6) = 3.5$$

Una vez más, 3.5 es un evento imposible en lo referente a una *sola* tirada, pero ciertamente es razonable en términos de un promedio calculado sobre muchas pruebas.

El *valor esperado* de un experimento es un promedio, y se puede calcular como sigue

$$E(x) = \sum_{i=1}^n p_i x_i$$

Es interesante observar que el valor esperado se puede calcular aun cuando no se hayan llevado a cabo observaciones muestrales, como en el caso del dado, y el valor esperado se puede *estimar* a partir de datos de muestreo, como en el ejemplo de las ventas de congeladores.

*Cabe observar que esto es idéntico a la determinación de la media de una distribución de frecuencias utilizando frecuencias relativas.

Ejemplo 1 Un inversionista se da cuenta de que tiene una probabilidad de 0.40 de obtener una utilidad de \$25 000, y una probabilidad de 0.60 de perder \$15 000 en una inversión. Su ganancia esperada es

$$0.40(25\,000) + 0.60(-15\,000) = \$1\,000$$

Obsérvese que la pérdida de \$15 000 tiene el signo menos.

Ejemplo 2 Un contratista hace las siguientes estimaciones:

Probabilidad	Tiempo de terminación
0.30	10 días
0.20	15 días
0.50	22 días

El número esperado de días para la terminación del proyecto, según estas estimaciones, es

$$0.30(10) + 0.20(15) + 0.50(22) = 17 \text{ días}$$

Los cálculos de valor esperado pueden comprender el número de ocurrencias, como: el número de errores cantidad de productos defectuosos, número de accidentes, etc., así como determinados resultados financieros, como utilidades, pérdidas o ganancias, rendimientos de inversión, etc.

Cuando las decisiones en finanzas se basan en valores esperados, se supone que existe una *valuación lineal* de las cantidades de dinero. Es decir, se supone que el valor de una cantidad de \$ 2000 es el doble del valor de \$ 1000 para el tomador de decisiones. Pero éste no siempre es el caso. Si una persona necesita una moneda de 10 centavos para hacer una llamada telefónica, nueve centavos no se pueden considerar como el 90% de la moneda de 10. En forma similar, si una pequeña compañía constructora necesita \$ 50 000 para seguir laborando, debe considerar dos ofertas: un 10% de probabilidad de obtener \$ 50 000 en una obra [valor esperado = $0.10(50\,000) = \$ 5000$], o bien, un 90% de probabilidad de obtener \$ 30 000 en otra [valor esperado = $0.90(30\,000) = 27\,000$] posiblemente seleccionaría el trabajo más riesgoso y que tiene el valor esperado más bajo, ya que necesita los \$ 50 000 para seguir su negocio. En este caso se diría que hay una *valuación no lineal*.

Sumas de variables aleatorias

Existe cierto número de situaciones en las que se deseará considerar una variable aleatoria que sea por sí misma la suma de dos o más variables aleatorias. En estos casos se debe poder determinar la media y la desviación estándar respecto a la variable aleatoria resultante. Supóngase que se tienen dos variables aleatorias, x , y , y que se conoce la media y la desviación estándar de cada una. A partir de esta información, se puede determinar la media y la desviación estándar de la suma de las dos variables aleatorias. Si para x se tiene μ_x y σ_x , y para y , μ_y y σ_y , entonces para $x + y$ se tiene

$$\mu_{x+y} = \mu_x + \mu_y \quad \text{y} \quad \sigma_{x+y} = \sqrt{\sigma_x^2 + \sigma_y^2}$$

Nótese que las desviaciones estándar *nunca* se suman, solamente se pueden sumar las variancias. Por tanto, para encontrar la variancia del total, se suman las variancias. Posteriormente se saca la raíz cuadrada de la variancia total si se quiere obtener la desviación estándar del total. Por ejemplo, se conectarán dos tramos de tubería, que provienen de una distribución que tiene una media de 10.0 pies y una desviación estándar de 3 pies. Es posible calcular la media y la desviación estándar de dicha sección.

$$\begin{aligned} \mu_x &= \mu_y = 10 & \sigma_x &= \sigma_y = 3 \\ \mu_{x+y} &= 10 + 10 = 20 & \sigma_{x+y} &= \sqrt{3^2 + 3^2} = 4.24 \end{aligned}$$

Si se hubiera querido unir cuatro tramos de tubería, la media y la desviación estándar de la longitud total hubieran sido

$$\mu = 10 + 10 + 10 + 10 = 40 \quad \sigma = \sqrt{3^2 + 3^2 + 3^2 + 3^2} = 6.0$$

La *media* de la suma de dos o más variables aleatorias es igual a la suma de las medias de dichas variables.

La *variancia* de la suma de dos o más variables aleatorias es la suma de las variancias de dichas variables.

EJERCICIOS

1. Clasifique cada una de las variables aleatorias siguientes, ya sea como discretas o continuas:
 - a. Las edades de niños que están en una cafetería.
 - b. El número de niños que se encuentran en una cafetería.
 - c. Litros de gasolina vendidos el martes en una gasolinera.
 - d. Número de actores de una obra de teatro.
2. El 10% de autos usados de un lote tienen problemas con el sistema eléctrico. Si hay 82 automóviles en el lote, ¿cuál es el número esperado de autos del lote que tienen problemas con el sistema eléctrico?
3. El número de llamadas telefónicas que se registran en un conmutador, y sus respectivas probabilidades en un intervalo de tres minutos, se ilustra en la siguiente tabla:

	Total					
Número de llamadas	0	1	2	3	4	5
Frecuencia relativa	0.60	0.20	0.10	0.04	0.03	0.03
						1.00

Según el promedio ¿cuántas llamadas se podrían registrar en un intervalo de tres minutos?

4. Una compañía está trabajando en cuatro proyectos independientes, *A*, *B*, *C* y *D*, con utilidades esperadas de \$ 4000, \$ 5000, \$ 10 000 y \$ 20 000, y con desviaciones estándar de \$ 100, \$ 300, \$ 200 y \$ 400. Determinar la utilidad esperada total de esos cuatro proyectos y la desviación estándar de dicho total.
5. Una variable aleatoria x tiene una media de 15 y una variancia de 2, en tanto que una variable aleatoria y tiene una media de 6 y una variancia de 1.
 - a. Encuentre μ_{x+y}
 - b. Determine σ_{x+y}
6. La Secretaría de Salud Pública informa que alrededor del 15% de los adultos de la nación se verán afectados por determinada enfermedad en los siguientes 12 meses. En el caso de una ciudad con una población de 250 000 adultos, ¿cuántos se espera que surgirán con dicha enfermedad?
7. Una pastelería llevó un registro de los pedidos (ver la tabla) de pasteles de chocolate de 7 pisos. Encuentre el número esperado de los pedidos de pasteles.

Total

Número de pasteles/día	0	1	2	3	4	5	6	7	8	9	
Frecuencia relativa	0.02	0.07	0.09	0.12	0.20	0.20	0.18	0.10	0.01	0.01	1.00

8. Un proceso de fabricación que consiste de tres pasos requiere de un tiempo promedio de 15 minutos para la terminación del primer paso, 25 minutos para el segundo y 30 minutos para el tercero. Las desviaciones estándar respectivas son de tres, cuatro y cinco minutos. Encuentre la media y la variancia del tiempo total de terminación del proceso.
9. Un billete de lotería tiene 0.00001 de probabilidad de ganar \$ 100 000, 0.0002 de ganar \$ 50 000 y 0.004 de ganar 25 000 ¿Cuál sería el precio justo que se debería pagar por el billete?
10. Se selecciona un varón de un numeroso grupo. El peso promedio del grupo es de 81 kg y la desviación estándar es de 9 kg. También se escogerá una mujer de un grupo semejante, que tiene una media de 63.5 kg y una desviación estándar de 6 kg. Obtenga la media y la variancia de los pesos combinados de un hombre y una mujer.
11. Repita el ejercicio 10, pero ahora considerando los pesos combinados de tres hombres.

DISTRIBUCIONES PROBABILÍSTICAS

Son las de distribuciones de frecuencias para los resultados de un espacio muestral (por ejemplo, para los resultados de una variable aleatoria). Las frecuencias son *frecuencias relativas*, o *probabilidades*. De este modo, las probabilidades indican el porcentaje de veces respecto a un gran número de observaciones en que se espera que se presenten los diversos resultados de una variable aleatoria. A menudo se utilizan tablas o gráficas para mostrar cómo la probabilidad total asignada a un espacio muestral (100%) *se distribuye* en relación con los resultados de dicho espacio.

Una *distribución probabilística* es una distribución de frecuencias relativa respecto a resultados del espacio muestral; señala la proporción de veces en que la variable aleatoria tiende a adoptar diversos valores.

Considérese la variable aleatoria “número de caras al tirar dos veces al aire una moneda”. La lista de los puntos en el espacio muestral y los valores correspondientes de la

Resultado	Valor de la variable aleatoria
CrCr	0
CrC	1
CCr	1
CC	2

Si la moneda no está cargada, $P(C) = P(Cr) = \frac{1}{2}$. Las probabilidades de los diferentes resultados son

Resultado	Probabilidad del resultado	Número de caras	$P(x)$
Cr Cr	$\frac{1}{2}(\frac{1}{2}) = \frac{1}{4}$	0	0.25
1 cara { Cr C	$\frac{1}{2}(\frac{1}{2}) = \frac{1}{4}$	1	0.50
CCr	$\frac{1}{2}(\frac{1}{2}) = \frac{1}{4}$		
C C	$\frac{1}{2}(\frac{1}{2}) = \frac{1}{4}$	2	0.25

De esta manera la distribución de probabilidad para el número de caras al tirar dos veces una moneda común es

Número de caras	$P(x)$
0	0.25
1	0.50
2	0.25
	1.00

Obsérvese que las probabilidades suman 1.00, ya que los resultados muestran que son mutuamente excluyentes y colectivamente exhaustivos. La misma distribución se puede mostrar en forma *acumulada*.

Número de caras	$P(x \text{ o menos})$
0	0.25
1	0.75
2	1.00

Gráficamente, esto se presentaría según se muestra en la figura 4.1.

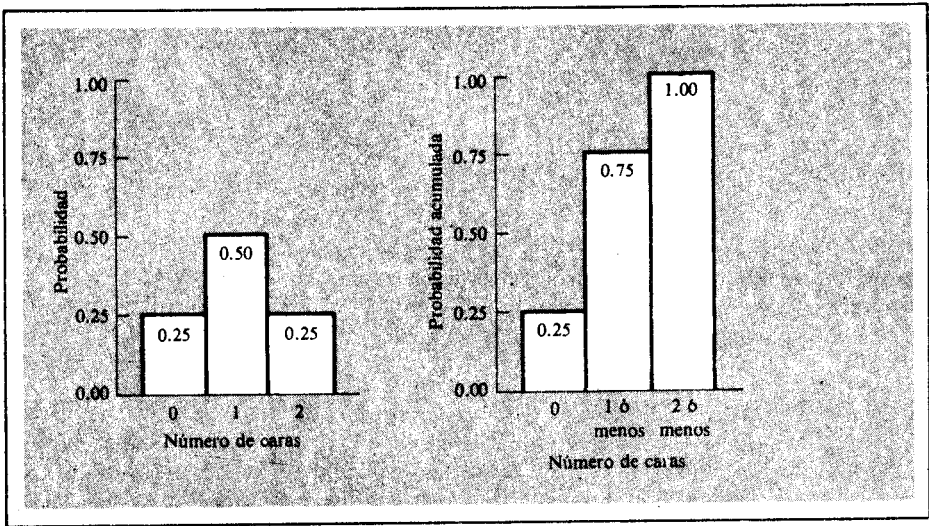


FIGURA 4.1 Distribución probabilística individual y distribución probabilística acumulada para el "número de caras al tirar una moneda dos veces".

Supóngase que una moneda se lanza en una situación donde $P(C) = 0.60$ (y así, $P(Cr) = 0.40$). Por tanto, se tiene una distribución probabilística *diferente* para el número de caras al tirar dos veces la moneda.

Probabilidad	Número de caras	Probabilidad
CrCr 0.40(.40) = 0.16	0	0.16
CrC 0.40(.60) = 0.24	1	0.48
CCr 0.60(.40) = 0.24		
CC 0.60(.60) = 0.36	2	0.36
1.00		1.00

Gráficamente, esta distribución se representaría como se muestra en la figura 4.2.

Nótese que, *dada* una distribución de probabilidad, es fácilmente evidente que algunos resultados de una variable aleatoria sean más probables que otros. Además, la probabilidad de un determinado resultado, o grupo de resultados, se puede determinar sin mucho esfuerzo. En términos prácticos, por lo general no es necesario molestarse en calcular cada una de las probabilidades para obtener una distribución probabilística. Para ello se dispone de tablas y fórmulas. En consecuencia, el problema real no es "¿cómo se *derivan* los valores?", sino ¿"cómo se *utilizan* las distribuciones para resolver problemas?"

Ahora, además del hecho de que las distribuciones probabilísticas proporcionan un método sencillo para la determinación de ciertas probabilidades, los tipos de distribu-

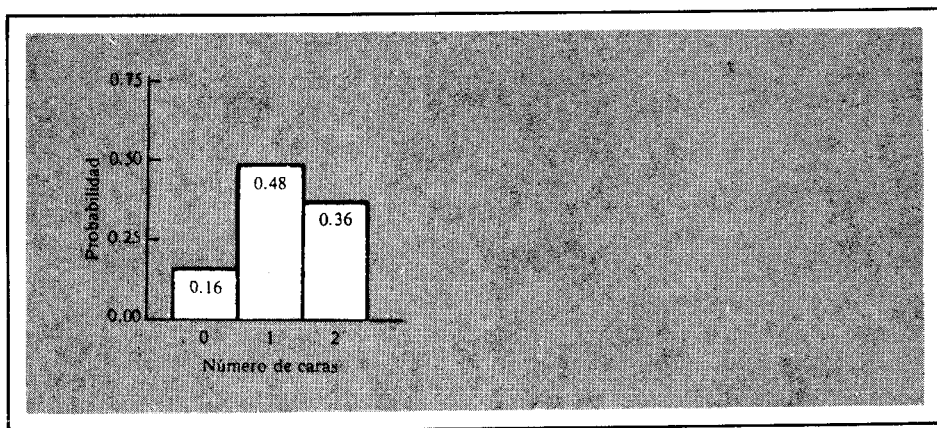


FIGURA 4.2 Gráfica de la distribución probabilística para el "número de caras en dos lanzamientos de una moneda" cuando $P(\text{caras}) = 0.60$ en cada tirada.

ción se pueden considerar como *modelos* que describen situaciones que comprenden resultados generados aleatoriamente.

En el Capítulo 1 se explicó el uso de modelos como base para la toma de decisiones racionales. Los modelos comprenden establecer supuestos simplificadores y eliminar detalles sin importancia y a menudo se puede utilizar para hacer que los problemas complejos sean más cortos y fáciles de manejar.

Un resultado de esta simplificación en el uso de modelos es que en sus formas más puras, pocos problemas son realmente únicos. En consecuencia, con mucha frecuencia se puede utilizar un reducido número de modelos básicos para obtener soluciones para una amplia variedad de problemas que, a primera vista, parecen no estar relacionados. Esto es especialmente cierto en la labor estadística. Algunos modelos básicos se pueden utilizar para resolver la mayoría de los problemas. Por ejemplo, el tirar al aire monedas, encontrar artículos defectuosos en una remesa de productos manufacturados, y el adivinar las respuestas en un examen, generalmente se manejan mediante el mismo tipo de distribución probabilística. La identificación de esto ha dado lugar al desarrollo de un conjunto de técnicas estandarizadas útiles para la obtención de soluciones para muchos problemas "diferentes". Esto permite a un administrador u otra clase de analista emplear la estadística para resolver un determinado problema sin tener que empezar cada vez desde el principio.

En estadística existe una gran variedad de tipos de distribuciones probabilísticas, cada una de las cuales tiene su propio conjunto de supuestos que definen las condiciones en las que cada tipo de distribución se puede emplear eficazmente. La clave del uso de distribuciones probabilísticas es hacer corresponder los supuestos de un tipo de distribución con las características de la situación real. Una vez que se lleva a cabo la correspondencia, el análisis se transforma en una cosa relativamente simple, ya que las distribuciones probabilísticas se pueden utilizar para resolver una clase de problemas, y todos ellos, dentro de cada clase, se manejan esencialmente de la misma manera.

En este capítulo se presentan algunos tipos de distribuciones probabilísticas, básicas pero muy importantes. A medida que se avance, el lector deberá concentrarse en las preguntas siguientes:

1. ¿Qué supuestos o restricciones básicas requiere cada una de las distribuciones probabilísticas? Este conocimiento es vital para lograr una correspondencia satisfactoria entre una variable aleatoria y la situación real.
2. ¿Cómo se pueden utilizar las distribuciones probabilísticas en la resolución de problemas?

La validez de utilizar determinada distribución para un problema, dependerá de qué tan cerca se aproxime la situación de éste al conjunto de condiciones supuestas por la distribución probabilística. En general, cuanto mejor sea la correspondencia, mejor será la respuesta.

La esencia de un análisis estadístico es hacer corresponder la suposiciones de cierta distribución probabilística con los datos específicos de un problema determinado.

La siguiente explicación de las distribuciones probabilísticas está organizada en dos partes: en este capítulo se estudian las distribuciones discontinuas, en tanto que las distribuciones continuas se considerarán en el siguiente.

DISTRIBUCIONES DISCONTINUAS

Las distribuciones probabilísticas discontinuas comprenden variables aleatorias para el conteo de datos, como el número de nacimientos por muestra o la cantidad de ocurrencias por unidad con respecto a un intervalo de tiempo, área o distancia. En páginas siguientes se hablará acerca de dos distribuciones discontinuas muy importantes: la binomial y la de Poisson.

DISTRIBUCIÓN BINOMIAL

El término “binomial” se utiliza para designar situaciones en las que los resultados de una variable aleatoria se pueden agrupar en dos clases o categorías. Por tanto, los datos son *nominales*. Las categorías deben ser mutuamente excluyentes, de manera que es evidente a qué clase pertenece una observación en particular, y las clases deben ser colectivamente exhaustivas, por lo que no es posible obtener ningún otro resultado.

Hay muchos ejemplos de variables aleatorias que se pueden clasificar como binomiales: respuestas a un examen de verdadero-falso, respuestas afirmativas o negativas a un cuestionario, productos manufacturados clasificados como defectuosos o satis-

factorios, pequeños de un jardín de niños que han sido vacunados o no, y sistemas de calificación de aprobado-no aprobado. Además, en general, las variables con resultados múltiples se pueden, tratar como binomiales cuando sólo uno de los resultados es de interés. De ahí que las respuestas a un examen de opción múltiple pueden ser correctas o incorrectas; en una urna puede haber canicas de cinco colores, pero si lo que deseamos es sacar sólo una canica verde, el resultado se puede considerar como verde- no verde; o bien, puede haber cinco aspirantes para una sola vacante de empleo, y el resultado final se puede establecer en términos de aceptado o no aceptado. De igual forma, la correspondencia puede ser local o para el extranjero, y las llamadas telefónicas pueden ser locales o de larga distancia. Aun los resultados de una variable continua se pueden reducir a dos clases mutuamente excluyentes. Por ejemplo, la velocidad de un automóvil se puede considerar (y comúnmente así ocurre) como dentro del límite legal de velocidad o por encima de dicho límite. Similarmente, se puede decir si un atleta corre 1 km en menos de cuatro minutos o si no lo logró, si una persona tiene una estatura mayor de 1.80 o no, etc.

Es completamente normal referirse a las dos categorías de una distribución binomial como “éxito” y “fracaso”, aunque para los objetivos del cálculo no importa a cuál se le dé un nombre u otro, dado que las dos son complementos. Por ejemplo, si interviene en un juego de azar con un amigo, el éxito de usted representa el fracaso de este último.

Con frecuencia, a las observaciones de un experimento binomial se les conoce como “ensayos”. Por ejemplo, un problema puede requerir la determinación de la probabilidad de cinco resultados satisfactorios en siete ensayos (observaciones).

La distribución binomial sirve para determinar la probabilidad de cierto número de resultados satisfactorios en un número dado de observaciones. Por ejemplo, supóngase que se sabe que el 80% de los votantes registrados en un distrito tienen más de 30 años de edad. Si alguien quiere saber la probabilidad —en una muestra de 10 votantes registrados— de encontrar siete o más votantes mayores de 30 años, se diría entonces que el éxito es igual a número de votantes registrados, y que $P(\text{éxito}) = 0.80$.

Ahora, para utilizar la distribución binomial es necesario satisfacer ciertos supuestos, los cuales se enumerarán a continuación:

1. Existen n observaciones o ensayos idénticos.
2. Cada ensayo tiene dos posibles resultados, uno denominado “éxito” y el otro “fracaso”.
3. Las probabilidades de éxito p y de fracaso $1 - p$ se mantienen constantes para todos los ensayos.
4. Los resultados de los ensayos son independientes entre sí.

Existen dos métodos para obtener las probabilidades de una variable aleatoria que está distribuida binomialmente. Uno de ellos consiste en utilizar la fórmula binomial, y el otro en consultar una tabla de probabilidades binomiales.

Fórmula binomial

Para calcular la probabilidad binomial es necesario especificar n , que es el número de ensayos; x , que es el número de éxitos y p , que es la probabilidad de éxito en cada ensayo. Supóngase que $p = 0.80$ [y por tanto, $p(\text{fracaso}) = 0.20$], y que se quiere determinar la probabilidad de obtener tres éxitos (y un fracaso) en cuatro observaciones. Ahora hay cuatro formas de obtener exactamente tres éxitos de cuatro observaciones. Cada una de las formas se muestra a continuación, junto con su respectiva probabilidad.

Ordenamiento	Probabilidad
EEEE	$(0.8)(0.8)(0.8)(0.2) = 0.1024$
EEFE	$(0.8)(0.8)(0.2)(0.8) = 0.1024$
EFEE	$(0.8)(0.2)(0.8)(0.8) = 0.1024$
FE EE	$(0.2)(0.8)(0.8)(0.8) = 0.1024$
	0.4096

La probabilidad de obtener tres éxitos y un fracaso, es la suma de todas las formas de obtener tres éxitos en cuatro observaciones. En este caso, la suma es 0.4096. Se debe observar que cada situación tiene la misma probabilidad de ocurrir, ya que los *mismos* números se multiplican juntos: solamente su *orden* es diferente. Esto *siempre* se cumple.

Esta observación conduce a las siguientes pautas. Las probabilidades de obtener resultados binomiales se pueden determinar matemáticamente, tomando en consideración dos cosas: el *número* de formas en que puede ocurrir la situación y la *probabilidad de una* de dichas formas. Nótese también que la primera es realmente el número de permutaciones que se pueden distinguir,

$$\binom{n}{x}$$

Más aún, en el Apéndice B se presenta una tabla (tabla E) de valores seleccionados de coeficiente binomiales, es decir

$$\binom{n}{x}$$

—por lo que generalmente es innecesario calcular esos valores.

Quizá el método más sencillo para determinar la probabilidad de una de las situaciones es considerar el caso en el que todos los éxitos ocurren primero y todos los fracasos después. De tener tres éxitos y un fracaso, (es decir, EEEF) y la probabilidad es entonces, $(0.8)(0.8)(0.8)(0.2)$, o bien, $(0.8)^3(0.2)^1$. De este modo, $P(x = 3) = 4(0.8)^3(0.2)^1 = 0.4096$.

Uniando estas dos ideas —número de formas y probabilidad de una de ellas— se obtiene lo siguiente:

$$P(x) = \binom{n}{x} P(\text{éxito})^x P(\text{fracaso})^{n-x}$$

en donde,

$\binom{n}{x}$ = número de formas de obtener x éxitos y $n-x$ fracasos en n ensayos

Nótese que si se quiere obtener x éxitos en n observaciones, entonces se deben esperar $n - x$ fracasos, ya que $x + n - x = n$, que es el total de observaciones. (Siempre se deberá comprobar que los exponentes se sumen a n , al utilizar este método.)

A continuación se presentan algunos ejemplos de cómo utilizar la fórmula para realizar los cálculos.

n	x	p	$\binom{n}{x} p^x (1 - p)^{n-x}$
5	3	0.30	$\binom{5}{3} (0.30)^3 (0.70)^2$
8	6	0.11	$\binom{8}{6} (0.11)^6 (0.89)^2$
9	5	0.44	$\binom{9}{5} (0.44)^5 (0.56)^4$
10	4	0.85	$\binom{10}{4} (0.85)^4 (0.15)^6$

La fórmula también se puede emplear para obtener probabilidades *acumuladas*, al calcular y posteriormente *sumar* las probabilidades individuales. He aquí algunos ejemplos.

n	p	x	Comprende	Cálculos	$P(x)$
3	0.4	1	1	$\binom{3}{1} (0.4)^1 (0.6)^2$	0.4320
5	0.2	0	0	$\binom{5}{0} (0.2)^0 (0.8)^5$	0.3277
5	0.2	1 o menos	0, 1	$\binom{5}{0} (0.2)^0 (0.8)^5 + \binom{5}{1} (0.2)^1 (0.8)^4$	0.7373
10	0.5	8 o más	8, 9, 10	$\binom{10}{8} (0.5)^8 (0.5)^2 + \binom{10}{9} (0.5)^9 (0.5)^1 + \binom{10}{10} (0.5)^{10} (0.5)^0$	0.0547
10	0.7	6	6	$\binom{10}{6} (0.7)^6 (0.3)^4$	0.2001

EJERCICIOS

Utilice la fórmula binomial para contestar las siguientes preguntas.

1. Una empresa que fabrica mesas de billar sospecha que el 2% de su producción está defectuosa en alguna forma. Si esta sospecha es correcta, encuentre la probabilidad de que en una muestra de nueve mesas:
 - a. haya por lo menos una defectuosa.
 - b. no haya mesas defectuosas.
2. De los alumnos de una universidad, el 41% fuma. Se eligen seis alumnos para conocer sus opiniones sobre el cigarro.
 - a. Encuentre la probabilidad de que ninguno de ellos fume.
 - b. Obtenga la probabilidad de que todos fumen.
 - c. Determine la probabilidad de que por lo menos la mitad de los seis fumen.
3. El 12% de los que hacen reservaciones para un vuelo en avioneta a menudo no llegan a tiempo para abordarla. Dicha avioneta tiene capacidad para 15 pasajeros.
 - a. Obtenga la probabilidad de que las 15 personas que hicieron reservaciones aborden la avioneta.
 - b. Si se anotaron 16 reservaciones, encuentre la probabilidad de que:
 - 1) se quede una persona.
 - 2) no se quede ninguna.
 - 3) se quede más de una.
4. Un vendedor de autos nuevos observa que el 80% de los autos vendidos son regresados al departamento de servicio para corregir diversos defectos de fabricación en los primeros 25 días después de su compra. De los 11 autos que se vendieron en un periodo de cinco días, ¿cuál es la probabilidad de que:
 - a. todos regresen en el lapso de 25 días para recibir servicio?
 - b. solo uno no regrese?
5. Suponga que el 8% de los emparedados que se venden en un estadio de béisbol se pidan sin mayonesa. Si siete personas ordenan emparedados encuentre la probabilidad de que:
 - a. todas los quieran con mayonesa.
 - b. solo una lo quiera con mayonesa.

TABLAS BINOMIALES

Las tablas de probabilidad ofrecen un método muy práctico para el análisis estadístico; proporcionan probabilidades con muy poco esfuerzo. Hay dos tipos de tablas binomiales. Uno proporciona las probabilidades de los resultados únicos o *individuales* de una variable aleatoria, en tanto que el otro establece las probabilidades de un *conjunto* de resultados. Aun cuando las dos tablas contienen esencialmente la misma información, algunos problemas utilizan más un tipo de tabla que otro, por lo que se habrán de considerar lo dos tipos.

Probabilidades binomiales individuales

Cuando el interés radica en la determinación de la probabilidad de un valor único en una distribución binomial (como la probabilidad de obtener *exactamente* cuatro resultados satisfactorios en seis observaciones), entonces puede ser muy útil una tabla de probabilidades binomiales individuales. Como sucede con la fórmula, se requieren tres factores de información: n , que es el número de observaciones, p , que es la probabilidad del éxito, y x , que es la cantidad específica de éxitos.

En la tabla 4.1 se muestra parte de una tabla binomial. En la parte superior de dicha tabla se indican los valores seleccionados de p , los cuales varían en incrementos de 0.05. En la parte inferior del lado izquierdo de la tabla se encuentran los tamaños de la muestra n . Obsérvese que, para cada n , el número posible de éxitos x se listan (de 0 a n). En el Apéndice se proporciona una tabla de probabilidades binomiales más extensa.

Utilícese la tabla 4.1 para obtener la probabilidad. Supóngase que se quiere encontrar la probabilidad de 5 éxitos ($x = 5$) en 8 observaciones ($n = 8$) cuando la probabilidad de éxito es 0.30. La tabla se emplea como sigue:

1. En la parte superior se encuentran los valores de p , allí encontrará el que busca.
2. En la parte lateral de arriba hacia abajo, se localiza el valor n . Encuentre el número deseado de éxitos x en la lista de resultados respecto a ese subconjunto.
3. La probabilidad de x éxitos se encuentra en la intersección de la fila situada en la parte 2 y la columna de la parte 1.

De este modo, la probabilidad de obtener exactamente 5 éxitos en 8 observaciones, cuando la probabilidad de lograr éxitos en cada observación es de 0.30 y 0.0467. En la tabla 4.1 este valor se encuentra encerrado en un círculo.

La siguiente tabla ilustra las probabilidades que se encontraron en la tabla A del Apéndice. Puede ser útil verificar en la tabla los números presentados.

n	Probabilidad de éxito P	x	$P(x)$
5	0.20	0	0.3277
8	0.60	3	0.1238
11	0.30	5	0.1321

Tabla binomial acumulativa

Muchos problemas en estadística requieren del empleo de la probabilidad combinada de un grupo de resultados en lugar de un resultado único. Por lo general, los resultados de interés son aquéllos que son mayores a cierto número especificado. Por ejemplo, el interés puede concentrarse en la probabilidad de obtener 5 o menos caras en 10 tiradas de una moneda. Tales preguntas se pueden contestar utilizando una tabla de probabilidades binomiales individuales, no obstante, este método requerirá buscar las probabilidades en la tabla, y después sumarlas: $P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$.

TABLA 4.1 Parte de una tabla de probabilidades binomiales individuales

n	x	p									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
→ 8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312
	2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094
	3	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039
	9	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	0	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	1	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	2	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	3	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	4	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	5	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	6	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	7	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	8	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

N. del S. Todas las tablas fueron tomadas del original en inglés, por lo tanto, no se puso el cero para indicar entero.

Una alternativa más eficaz es utilizar una tabla de probabilidades acumulativas, ya que en una tabla de este tipo los valores individuales ya están sumados (acumulados), lo que ahorra tiempo, y reduce la posibilidad de cometer errores en el cálculo.

El formato de una tabla binomial acumulativa es casi idéntica a la tabla de probabilidades binomiales individuales. Los valores seleccionados de p y la probabilidad de éxito se enumeran en la parte superior de la tabla, en tanto que el número posible de ocurrencias (éxitos) para varios tamaños muestrales se indica en sentido descendente en uno de los lados de la misma. Las probabilidades que se muestran son para x o *menos* éxitos, en lugar de para la probabilidad de *exactamente* x éxitos, como se observa en la tabla de probabilidades binomiales individuales. En la tabla 4.2. se muestra una parte de una tabla acumulativa. El valor encerrado en un círculo 0.9360, es la probabilidad de que en tres observaciones haya dos o menos éxitos (es decir, 0 ó 1 ó 2) cuando la probabilidad del éxito es 0.40.

Una tabla acumulativa se puede utilizar en diferentes formas. Puede emplearse para encontrar directamente la probabilidad de que x sea igual a cierto número específico de éxitos o menor a éste. Y se puede utilizar indirectamente para obtener, tanto la probabilidad de que x sea *mayor* que cierto número de éxitos, como la probabilidad de obtener *exactamente* x éxitos. Esto se debe al hecho de que los resultados son mutuamente excluyentes y colectivamente exhaustivos; sus probabilidades en cualquier ejemplo dado sumarán 1.00. Por tanto, si se encuentra $P(x \leq 6) = 0.72$, entonces $P(x > 6) = 1.00 - 0.72 = 0.28$.

Un pequeño truco que ayudará a evitar cierta confusión es enumerar o listar los resultados posibles para una situación dada, y subrayar los resultados para los que se desea obtener las probabilidades. Por ejemplo, supóngase que se quiere determinar la probabilidad de 3 o menos éxitos en 7 observaciones. Primero se deberán enumerar los éxitos posibles:

0 1 2 3 4 5 6 7

A continuación, se subrayarán los de interés

0 1 2 3 4 5 6 7
(3 o menos)

Como la tabla muestra la probabilidad de x o *menos* éxitos, se puede buscar directamente la probabilidad de 3 o menos [dados n y $P(\text{éxitos})$]. Es decir

0 1 2 3 4 5 6 7

la tabla muestra la probabilidad de cualquiera de este número de éxitos.

Para obtener la probabilidad de *más de* 3 (es decir, 4, 5, 6 ó 7), se deberá encontrar la probabilidad de tres o menos y restarla de 1.00:

$\xrightarrow{100\%}$
 0 1 2 3 4 5 6 7
 └────────┬────────┘
 valor tabular 1 — valor tabular

También se puede utilizar una tabla acumulativa para determinar las probabilidades individuales. Por ejemplo, para encontrar la probabilidad de exactamente 4 éxitos, reste $P(3 \text{ o menos})$ de $P(4 \text{ o menos})$:

$$P(x \leq 4) - P(x \leq 3) = P(x = 4)$$

Considérese este ejemplo. Si $P(\text{éxito}) = 0.50$ y $n = 4$ utilice la tabla acumulativa para encontrar $P(x = 3)$:

$P(x \leq 3)$ comprende	1	2	3	y es igual a	0.9375
$P(x \leq 2)$ comprende	1	2		y es igual a	0.6875
$P(x = 3)$ comprende			3	y es igual a	0.2500

En la tabla 4.3 se ilustran los diversos usos de la tabla acumulativa, y la forma como se pueden obtener las probabilidades deseadas. Si se estudia con cuidado, se estará en condiciones de contestar la mayoría de los problemas que puedan resultarle difíciles.

TABLA 4.3 Uso de una tabla binomial acumulativa

A. Lista de resultados deseados

		Resultados											Leer en la tabla acumulativa
$n = 10$		0	1	2	3	4	5	6	7	8	9	10	
$P(X \leq 6)$	6 o menos	0	1	2	3	4	5	6					directamente: $P(6)$
$P(X < 6)$	menos de 6	0	1	2	3	4	5						directamente: $P(5)$
$P(X \geq 6)$	6 o más							6	7	8	9	10	$1 - P(5)$
$P(X > 6)$	más de 6								7	8	9	10	$1 - P(6)$
$P(X = 6)$	6							6					$P(6) - P(5)$

B. Probabilidades para lo anterior, utilizando $p = 0.30$ y $n = 10$. Las probabilidades provienen de la tabla B del Apéndice.

$P(X \leq 6)$			0.9894
$P(X < 6)$			0.9527
$P(X \geq 6)$	1	-0.9527	0.0473
$P(X > 6)$	1	-0.9894	0.0106
$P(X = 6)$	0.9894	-0.9527	0.0367

Las tablas ofrecen el método más sencillo y práctico para determinar probabilidades, no obstante, hay situaciones en las que las probabilidades deseadas no se pueden obtener a partir de tablas como las que se han venido empleando. Por ejemplo, supóngase que $P(\text{éxito}) = 0.12$. La tabla no incluye este valor. Para mayor seguridad, se dispone de tablas más grandes y extensas, y muchas bibliotecas las tienen en sus secciones de referencia. Pero no siempre es práctico hacer un viaje hasta la biblioteca. Por éstas y otras razones, será útil emplear la fórmula binomial para obtener las probabilidades deseadas.

Características de las distribuciones binomiales

Cualquier tipo de distribución probabilística presenta ciertas características que la diferencian de otros tipos de distribución de probabilidades. Muy a menudo conocer estas diferencias proporciona una mejor comprensión de cuándo utilizar un tipo de distribución, un mejor entendimiento de cómo se pueden emplear algunas distribuciones a fin de aproximar probabilidades para otros tipos de distribución y, en general, una mejor apreciación de los conceptos del análisis estadístico.

Como los datos están en forma de *cuentas*, el número de éxitos siempre deberá ser un valor entero (0, 1, 2, 3, . . .). Sin embargo, algunas veces es útil expresar el número de éxitos como un porcentaje del número de observaciones. Por ejemplo, 2 éxitos en 10 observaciones equivaldrían a 0.20 ó 20%. Asimismo, otro método de representar una distribución binomial es utilizar una gráfica de barras en lugar de un histograma. La gráfica de barras es especialmente útil para ilustrar la distribución cuando se trabaja con un porcentaje de éxitos en lugar de con una cantidad de éxitos. Véase la figura 4.3.

La *media* de una distribución binomial es el promedio a largo plazo, o el *valor esperado* de una variable aleatoria binomial. La desviación estándar de una distribución binomial indica el grado en que los valores muestrales tenderán a variar a partir de la media de la distribución. En el caso del binomio, tanto la media como la desviación estándar se pueden expresar en términos del *número* o el *porcentaje* de éxitos. Las fórmulas se presentarán a continuación.

	Media	Desviación estándar
Número de éxitos	np	$\sqrt{np(1 - p)}$
Porcentaje de éxitos	p	$\sqrt{p(1 - p)/n}$

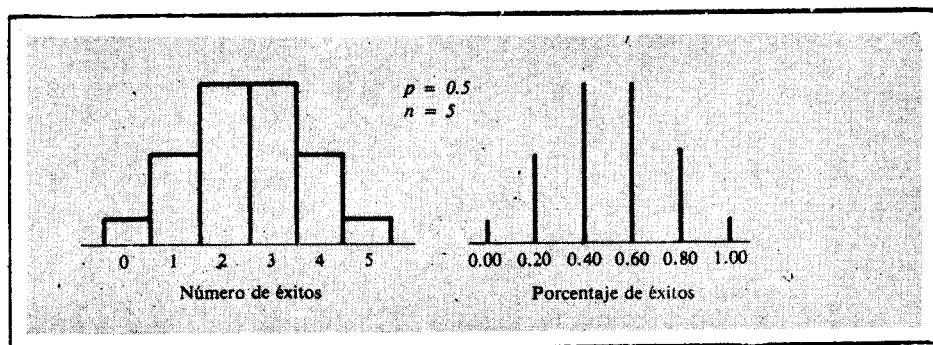


FIGURA 4.3 Generalmente se utiliza un histograma cuando la variable aleatoria es el *número* de éxitos; en lo referente al *porcentaje* de éxitos se utiliza una gráfica de barras.

Ejemplo 3 Supóngase que la probabilidad de éxito es 0.10 y el número de observaciones es 100. Encuentre la media y la desviación estándar de la distribución probabilística, tanto para el número como para el porcentaje de éxitos.

Solución:

	Media	Desviación estándar
Número de éxitos	$100(0.10) = 10$	$\sqrt{100(0.10)(0.90)} = 3$
Porcentaje de éxitos	0.10	$\sqrt{0.10(0.90)/100} = 0.03$

La media de 10 en lo referente a número de éxitos se interpreta como el número promedio de éxitos a largo plazo en muestras de 100 observaciones. En forma semejante, el porcentaje de éxitos a largo plazo en muestras de 100 es 0.10. Las dos desviaciones es-

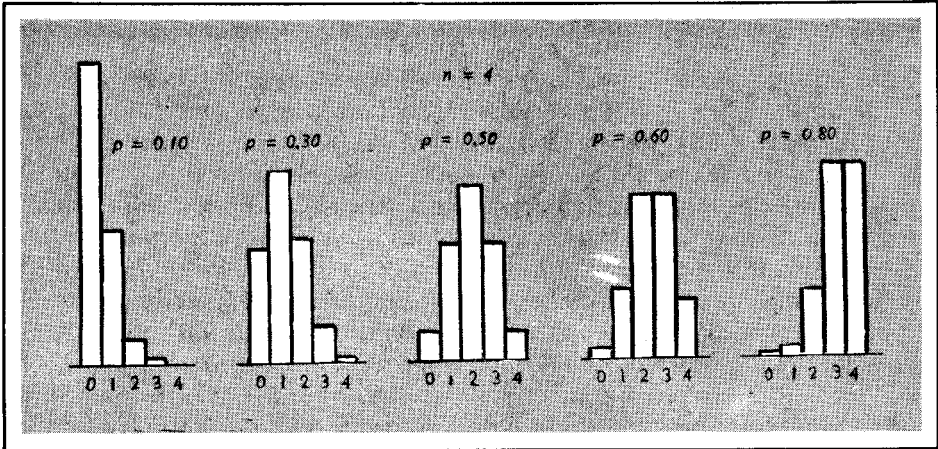


FIGURA 4.4 Una distribución binomial será simétrica para $p = 0.5$, sesgada hacia la derecha para $p < 0.5$ y sesgada hacia la izquierda para $p > 0.5$.

tándar reflejan la variabilidad que presentarán las muestras individuales. Cabe observar que el 3% de 100 es 3; las dos formas alternativas son equivalentes.

La distribución binomial tiene dos parámetros: p , que es la probabilidad de éxito, y n , que es el número de observaciones. Cada combinación de p y n se refiere a una distribución única o espacio muestral. Obsérvese cómo los valores sombreados en la tabla binomial se utilizaron para elaborar las distribuciones de la figura 4.4.

En el caso de cualquier tamaño muestral n , una distribución de probabilidad binomial será simétrica para $p = 0.50$, pero sesgada hacia la derecha si p es mayor que 0.50, y hacia la izquierda si p es menor de 0.50, como se muestra en la figura 4.4. Esta tendencia a sesgarse cuando p no es 0.50 disminuye a medida que n aumenta.

EJERCICIOS

- Utilice una tabla binomial (ya sea individual o acumulativa) para determinar la probabilidad de x éxitos:

	Número de observaciones	$P(\text{éxitos})$	x
a.	5	0.2	0
b.	5	0.3	1
c.	5	0.5	1 o menos
d.	10	0.1	2 o menos
e.	10	0.1	2 o más
f.	10	0.9	8 o menos
g.	10	0.7	2 ó 3
h.	12	0.7	3 o menos

- Los registros de una pequeña compañía de servicios indican que el 40% de las facturas que envían son pagadas después de la fecha de vencimiento. Si se envían 14 facturas, encuentre la probabilidad de que:

- a. ninguna se pague con retraso.
 - b. cuando menos dos se paguen con retraso.
 - c. cuando menos la mitad se paguen con retraso.
3. Una compañía de exploración petrolera observa que en casi el 5% de los pozos de prueba que perfora, encuentra un depósito de gas natural. Si perfora seis pozos, obtenga la probabilidad de que al menos en uno se encuentre gas.
 - ★ 4. Una prueba de opción múltiple presenta cuatro alternativas por pregunta y 14 preguntas. Si la calificación aprobatoria depende de obtener nueve o más respuestas correctas, ¿cuál es la probabilidad de que un estudiante que adivina todas las preguntas apruebe el examen?
 5. Una compañía de bienes raíces observa que 1 de 10 compradores potenciales de casas prometen comprar una si regresan por segunda vez. En 10 de estos casos, encuentre la probabilidad de que ninguno haga una oferta.
 6. En una encuesta reciente se concluyó que únicamente el 15% de los médicos de un área rural fuman. Se observó que dos de los ocho médicos seleccionados de una lista suministrada por el directorio médico local, también fuman. Suponiendo que la encuesta esté en lo correcto, ¿cuál es la probabilidad de obtener este resultado?
 7. Las investigaciones médicas señalan que el 20% de la población general sufre efectos negativos colaterales al ingerir un nuevo fármaco. Si un médico receta dicho fármaco a cuatro pacientes, ¿cuál es la probabilidad de que:
 - a. ninguno sufra efectos colaterales?
 - b. todos los tengan?
 - c. al menos uno presente efectos colaterales?
 8. En una reciente encuesta gubernamental se encontró que el 80% de las familias que viven en una comunidad suburbana, y cuyos ingresos brutos son superiores a los \$ 15 000, poseían dos autos. Suponiendo que el estudio esté en lo cierto si se selecciona una muestra de 10 familias de esta categoría, obtenga la probabilidad de que exactamente el 18% de los que integran dicha muestra tenga dos autos.
 9. Un televisor que presenta 10 series de circuitos tiene uno defectuoso. Ocho de las series son difíciles de remplazar. Encuentre la probabilidad de que la serie defectuosa no sea una de ellas.
 10. Los informes del tránsito indican que el 25% de los vehículos que se detienen en una autopista interestatal no son sometidos a una revisión de seguridad. Si se detienen 16 vehículos, encuentre la probabilidad de que:
 - a. 2 o más no satisfagan las normas de seguridad.
 - b. 4 o más no las satisfagan.
 - c. 9 o más no las cumplan.
 11. Un comentarista de deportes acierta al señalar al ganador en 6 de 10 partidos de béisbol. Si una persona simplemente está adivinando, ¿cuál es la probabilidad de que pueda igualar o superar dicha marca?
 12. Una moneda común se lanza tres veces al aire acertando siempre en su resultado. ¿Cuál es la probabilidad de repetir este hecho?
 13. Según los archivos universitarios, de los estudiantes de una escuela de enseñanza media superior, el 15% cambia de especialidad por lo menos una vez durante su

primer año de estudios. Si se seleccionan 11 estudiantes de los grupos de primer año, encuentre la probabilidad de que:

- a. todos cambien de especialidad por lo menos una vez.
- b. por lo menos nueve hayan cambiado.
- c. más de la mitad hayan cambiado.

14. Una gran urna contiene 10 000 canicas de colores:

blancas	5000
verdes	3000
rojas	1500
negras	500
	10 000

- a. De escogerse 20 canicas, obtenga $P(2 \text{ verdes})$.
 - b. Si se seleccionan 19, ¿cuál es la probabilidad de que por lo menos 3 sean verdes?
 - c. Si se seleccionan 19, ¿cuál es la probabilidad de que por lo menos 3 sean blancas?
15. Remesas de 500 productos cada una son aprobadas si en una muestra aleatoria de 10 se hallan menos de dos defectuosos. Si en realidad una remesa tiene el 5% de artículos defectuosos, ¿cuál es la probabilidad de que sea aprobada?
16. En el caso del ejercicio 15, ¿qué porcentaje de lotes con 10% de productos defectuosos producirán muestras de 10 artículos en las que ninguno estuviese defectuoso?
17. Un mecánico sabe, con base en su experiencia, que el 90% de las refacciones que desecha pueden usarse de nuevo. Si para un trabajo se necesitan cinco partes reutilizables, ¿cuál es el número mínimo de refacciones que deberá obtener si desea que la probabilidad de devolver partes sobrantes sea menor de 0.12?
18. Calcule la media y la desviación estándar para el número de éxitos en los siguientes casos:
- a. $n = 25, p = 0.5$
 - b. $n = 50, p = 0.2$
 - c. $n = 80, p = 0.4$
19. Repita el ejercicio 8 respecto al *porcentaje* de éxitos.
20. A veces es imposible que el porcentaje de éxitos de la muestra iguale el porcentaje de la media. Por ejemplo, cuando $n = 4$ y $p = 0.3$, el porcentaje nunca puede ser 0.30.
- a. Enumere los posibles porcentajes de la muestra para este caso.
 - b. ¿Cómo se interpreta la media de 0.3?
21. Trace dos series de gráficas, una para el número de éxitos y otra para el porcentaje de éxitos, respecto a las siguientes situaciones:
- a. $n = 5, p = .10$
 - b. $n = 10, p = 0.70$

La distribución de Poisson suele ser útil para describir las probabilidades del número de acaecimientos con respecto a un campo o intervalo continuo (generalmente de tiempo o espacio). Algunos ejemplos de variables aleatorias que se pueden representar o modelar por la distribución de Poisson, son los defectos por centímetro cuadrado,

accidentes por día, clientes por hora, llamadas telefónicas por minuto, mosquitos por kilómetro cuadrado, cabezas de ganado por hectárea, y así sucesivamente. Obsérvese que la unidad de medición (por ejemplo, tiempo, área) es continua, pero que la variable aleatoria, el número de acaecimientos, es discreta. Además, cabe notar que es imposible contar las fallas o fracasos. El número de accidentes que *no* ocurrieron no se puede contar, así como el número de llamadas que no se realizaron o el número de defectos que no se presentaron en un centímetro cuadrado. Véase figura 4.5.

El uso de la distribución de Poisson supone lo siguiente:

1. La probabilidad de un acaecimiento u ocurrencia es la misma a través de todo el campo de observación.
2. La probabilidad de más de un acaecimiento en cualquier punto único es aproximadamente cero.
3. El número de ocurrencias en cualquier intervalo es independiente del número de acaecimientos en otros intervalos.

El límite inferior del número de acaecimientos en todas esas situaciones es cero, y el límite superior —por lo menos teóricamente— es infinito, aun cuando en la mayoría de los ejemplos anteriores sería difícil imaginar un número ilimitado de ocurrencias. De este modo, la distribución de Poisson, que va de cero a un número infinito de ocurrencias por unidad, puede no representar *exactamente* cualquiera de los procesos aleatorios mencionados en páginas anteriores. No obstante, por lo general, la distribución de Poisson se puede utilizar como *modelo* para aproximar estrechamente tales procesos.

En la figura 4.6 se ilustra una típica distribución de frecuencias de observaciones. Obsérvese que casi toda la probabilidad se concentra cerca del origen, y que la probabili-

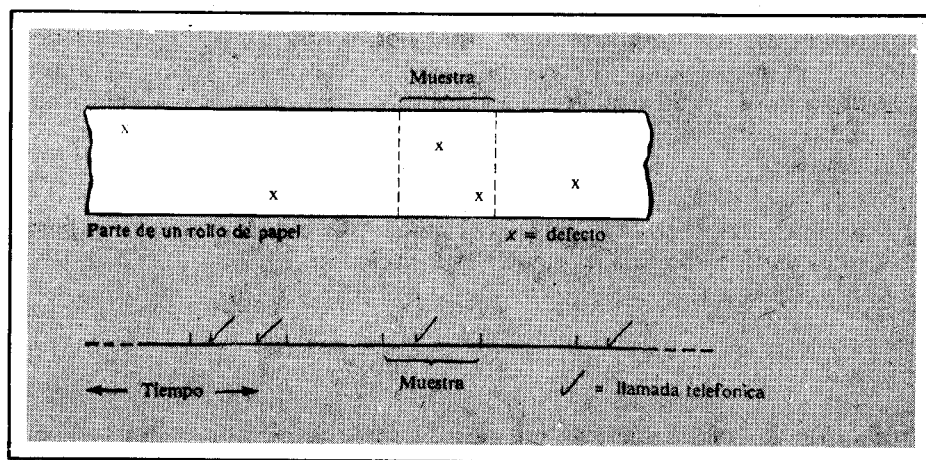


FIGURA 4.5 La distribución de Poisson se utiliza para el número de ocurrencias respecto a un campo continuo, como el área o el tiempo.

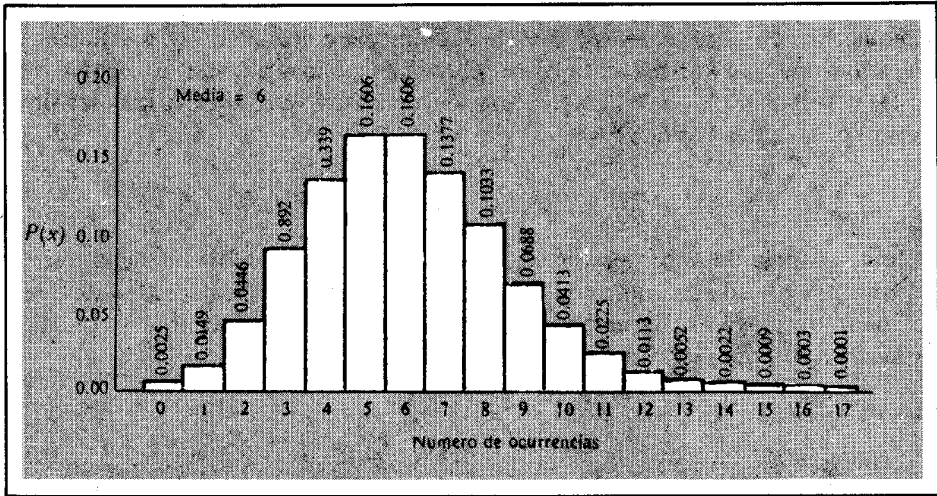


FIGURA 4.6 Histograma de una distribución de Poisson típica.

dad de observar grandes valores de la variable aleatoria es muy pequeña. Cualquier diferencia entre lo que teóricamente pueda ocurrir y lo que sea probable que suceda es, por tanto, insignificante. En consecuencia, cuando se dice que una variable aleatoria está distribuida mediante el método de Poisson, a menudo se quiere decir que la distribución de frecuencias de acaecimientos respecto a esa variable, se puede aproximar razonablemente mediante una distribución de Poisson.

La distribución de Poisson se puede describir en su totalidad por medio de un solo parámetro, la media, según se señala en la figura 4.6.* De esta manera, sabiendo que una variable aleatoria tiene resultados que son distribuidos mediante el método de Poisson, y conociendo el número promedio de ocurrencias por unidad, se puede determinar la probabilidad para cualquiera o todos los resultados posibles. Como sucede con la distribución binomial, hay dos métodos para llevar a cabo esto. Hay tablas de probabilidad individuales y acumulativas a partir de las que se pueden obtener probabilidades para medias *seleccionadas*, y se cuenta con una fórmula que se puede utilizar en el cálculo de la probabilidad de *cualquier* media.

Fórmula de Poisson

Si una variable aleatoria se describe mediante una distribución de Poisson, entonces la probabilidad de considerar (observar) cualquier número dado de ocurrencias por unidad de medición (minuto, hora, centímetro, metro cuadrado, etc.) se puede obtener mediante la fórmula

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

*La variancia de una distribución de Poisson es igual a la media: media = μ y variancia = μ .

en la cual x es el número de ocurrencias; e es la base de logaritmos naturales (la tabla F del Apéndice tiene valores para $e^{-\mu}$); λ es la razón media por unidad; y t es el número de unidades. La cantidad λt representa la cantidad media de ocurrencias respecto al intervalo t . De esta manera, $\mu = \lambda t$. La fórmula se puede escribir en una forma más simple sustituyendo la media λt por μ :

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

Ejemplo 4 Mediante un proceso mecánico se producen alfombras de lana que presentan un promedio de dos defectos por metro. Encuentre la probabilidad de que un metro cuadrado tenga exactamente un defecto, suponiendo que el proceso puede ser aproximado mediante una distribución de Poisson.

Solución:

Se sabe que $\mu = 2$, y a partir de la tabla F del Apéndice, se observa que $e^{-2} = 0.135$. De tal manera que

$$P(x = 1) = \frac{e^{-2}(2)^1}{1!} = \frac{0.135(2)}{1} = 0.270$$

Aplicación que comprende el tiempo

Considérese este ejemplo. Suponga que los barcos arriban a un puerto a razón de $\lambda =$ dos barcos/hora, y que esta proporción está bien aproximada mediante el proceso de Poisson. Si se observa este proceso durante un periodo de media hora ($t = \frac{1}{2}$), encuentre la probabilidad de que (a) no arribe ningún barco y (b) arriben tres barcos.

Primeramente se encuentra μ

$$\mu = \lambda t = 2\left(\frac{1}{2}\right) = 1$$

De la tabla F se obtiene que $e^{-1} = 0.368$.

$$(a) \quad P(x = 0) = \frac{e^{-1}(1)^0}{0!} = \frac{e^{-1}(1)}{1} = 0.368$$

$$(b) \quad P(x = 3) = \frac{e^{-1}(1)^3}{3!} = \frac{0.368}{3 \cdot 2 \cdot 1} = 0.061$$

Aplicación que comprende el área

Supóngase que los defectos de la hilaza se pueden aproximar mediante un proceso de Poisson con una media de 0.2 defectos/metro ($\lambda = 0.2$). Si se inspeccionan longitudes de seis metros ($t = 6$) calcule la probabilidad de hallar menos de dos defectos (es decir, 0 ó 1).

Una vez más, primero se determina μ

$$\mu = \lambda t = 0.2(6) = 1.2$$

De la tabla F se obtiene $e^{-1.2} = 0.301$. De este modo

$$\begin{aligned} P(x \leq 1) &= P(0) + P(1) = \frac{e^{-1.2}(1.2)^0}{0!} + \frac{e^{-1.2}(1.2)^1}{1!} \\ &= 0.301 + 0.301(1.2) = 0.6622 \end{aligned}$$

EJERCICIOS

- ★ 1. Las llamadas de emergencia registradas en el conmutador de una estación de policía a razón de 4/hora, un fin de semana cualquiera (no en época de vacaciones) se pueden aproximar mediante la distribución de Poisson.
 - a. En un lapso de 30 minutos, ¿cuántas llamadas de emergencia se espera recibir?
 - b. En un lapso de 30 minutos, ¿cuál es la probabilidad de que no se registren llamadas?
 - c. En un lapso de 30 minutos, ¿cuál es la probabilidad de que haya por lo menos dos llamadas?
2. El número promedio de radios que vende una compañía por día se aproxima mediante el método de Poisson, con una media de 1.5. Calcule la probabilidad de que la compañía venda por lo menos cuatro radios durante un:
 - a. periodo de dos días b. periodo de tres días c. periodo de cuatro días
3. Los defectos en un rollo fotográfico de color promedian 0.1 defectos/rollo, y la distribución que se sigue para determinar el número de defectos es la de Poisson. Obtenga la probabilidad de que cualquier rollo fotográfico de color presente uno o más defectos.
4. Los clientes llegan a una exhibición a razón de 6.5 clientes/hora (Poisson). Calcule la probabilidad de que en cualquier hora dada:
 - a. no llegue ningún cliente b. por lo menos lleguen cinco
 - c. llegue más de uno d. lleguen exactamente 6.5

TABLAS DE POISSON

Las tablas de probabilidad de Poisson proporcionan un método conveniente para obtener las probabilidades con un mínimo de tiempo y esfuerzo. Por esta razón, se deberán utilizar siempre que sea posible, en lugar del método por fórmulas.

Probabilidades individuales de Poisson

En realidad, las tablas de Poisson son muy semejantes a las tablas binomiales, aunque a primera vista parecen ser diferentes. Como la distribución de Poisson es una función sólo de la media del proceso, las tablas están diseñadas para proporcionar las probabi-

lidades con base en la media del proceso. Parte de la tabla de Poisson presentada en el Apéndice se muestra en la tabla 4.4. Los valores seleccionados de μ , la media del proceso (número promedio de ocurrencias por unidad), se enumeran en la parte superior de la tabla, en tanto que los resultados posibles se listan hacia abajo. En la tabla se muestra la probabilidad de obtener *exactamente* x ocurrencias por unidad.

Por ejemplo, supóngase que se tiene un proceso de Poisson con un promedio de tres ocurrencias por hora, y se quiere saber la probabilidad de obtener exactamente una ocurrencia en cualquier hora dada. La respuesta se puede encontrar en la intersección de $\mu = 3$ y $x = 1$, según se muestra en la tabla 4.4.

TABLA 4.4 Parte de una tabla de Poisson

	μ										
x	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0	
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498	
→ 1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494	
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240	
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240	
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680	
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008	
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504	
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216	
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081	
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027	
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008	
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	

A continuación se presentan algunos ejemplos de probabilidades de Poisson, obtenidas a partir de la tabla de probabilidades individuales.

μ	x	$P(x)$
2.1	0	0.1225
2.4	1	0.2177
3.0	2	0.2240
2.2	2	0.2681
3.0	5	0.1008
2.8	10	0.0005

Obsérvese en la tabla que para cada media, la suma de las probabilidades de los diversos resultados posibles es igual a 1.00.

Tabla acumulativa de Poisson

Dicha tabla proporciona sumas de probabilidades, al igual que la tabla binomial acumulativa. La tabla muestra probabilidades para x o menos ocurrencias, dada la media

del proceso. Parte de una tabla de Poisson acumulativa se muestra en la tabla 4.5. Una vez más, enumere todos los resultados posibles y subraye los resultados para lo que se quiere conocer la probabilidad, esto simplificará en gran medida trabajar con la tabla acumulativa.

TABLA 4.5 Parte de una tabla acumulativa de Poisson.

	μ									
	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
0	.0498	.0450	.0408	.0369	.0334	.0302	.0273	.0247	.0224	.0202
1	.1991	.1847	.1712	.1586	.1468	.1359	.1257	.1162	.1074	.0992
2	.4232	.4012	.3799	.3594	.3397	.3208	.3027	.2854	.2689	.2531
3	.6472	.6248	.6025	.5803	.5584	.5366	.5152	.4942	.4735	.4532
4	.8153	.7982	.7806	.7626	.7442	.7254	.7064	.6872	.6678	.6484
5	.9161	.9057	.8946	.8829	.8705	.8576	.8441	.8301	.8156	.8006
6	.9665	.9612	.9554	.9490	.9421	.9347	.9267	.9182	.9091	.8995
7	.9881	.9858	.9832	.9802	.9769	.9733	.9692	.9648	.9599	.9546
8	.9962	.9953	.9943	.9931	.9917	.9901	.9883	.9863	.9840	.9815
9	.9989	.9986	.9982	.9978	.9973	.9967	.9960	.9952	.9942	.9931
10	.9997	.9996	.9995	.9994	.9992	.9990	.9987	.9984	.9981	.9977
11	.9999	.9999	.9999	.9998	.9998	.9997	.9996	.9995	.9994	.9993
12	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999	.9998	.9998
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Algunas probabilidades obtenidas a partir de la tabla acumulativa de Poisson que se encuentra en el Apéndice, se muestran en la tabla 4.6.

TABLA 4.6 Uso de la tabla acumulativa de Poisson.

μ	Probabilidad deseada para	incluir los resultados	Cálculos	$P(x)$
.8	$x \leq 1$	0, 1	se lee directamente	0.809
1.2	$x < 3$	0, 1, 2	se lee $P(x \leq 2)$	0.879
1.5	$x = 0$	0	se lee directamente	0.223
2.0	$x > 3$	4, 5, 6, ...	$1 - P(x \leq 3)$	0.143
2.6	$1 < x \leq 4$	2, 3, 4	$P(x \leq 4) - P(x \leq 1)$	0.610
3.8	$1 \leq x \leq 4$	1, 2, 3, 4	$P(x \leq 4) - P(x = 0)$	0.646
5.6	$1 \leq x \leq 4$	1, 2, 3, 4	$P(x \leq 4) - P(x = 0)$	0.338
6.0	$x \geq 5$	5, 6, 7, ...	$1 - P(x \leq 4)$	0.715

Para valores de μ que no se encuentran en la tabla, se puede encontrar una más amplia, o interpolar (en el caso de valores aproximados) en la tabla, o recurrir a una fórmula matemática para probabilidades de Poisson.

La distribución de Poisson como una aproximación a la binomial

En determinadas circunstancias la distribución de Poisson se puede utilizar para aproximar probabilidades binomiales. La aproximación es más adecuada cuando el número de observaciones n es mayor, y la probabilidad del éxito p está cerca de 0 ó 1.00. Es deseable contar con un método alternativo para obtener probabilidades binomiales, por las siguientes razones:

1. La distribución binomial describe adecuadamente muchas situaciones interesantes.
2. La mayoría de las tablas se limitan a $n \leq 20$.
3. La fórmula binomial puede requerir de un gran esfuerzo para obtener una solución exacta.

Las ventajas de la aproximación son que la exactitud se altera muy poco, y que el esfuerzo comprendido es considerablemente menor. Para utilizar la aproximación, es necesario determinar solamente la media o el valor esperado de la distribución binomial. Este valor se utiliza como la media del proceso para la distribución de Poisson. Es decir, el promedio del proceso μ es igual al promedio binomial np .

Ejemplo 5 Obtenga la probabilidad de encontrar cuatro artículos defectuosos de una muestra de 300 tomada de un enorme lote, en el que se dice que hay un 2% de artículos defectuosos.

Solución:

Como los valores $n = 300$ y $p = 0.02$ están más allá de la variedad de valores presentados en la tabla binomial, las alternativas serían buscar una tabla más extensa, y recurrir a la fórmula binomial

$$P(X = 4) = \binom{300}{4} (0.02)^4 (0.98)^{296}$$

(lo que representaría una tarea muy difícil), o bien, utilizar la aproximación de Poisson, con $\mu = np$. De esta manera

$$\mu = np = 300(0.02) = 6$$

A partir de la fórmula de Poisson se observa que:

$$P(X = 4) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{6^4 e^{-6}}{4!} = 0.135$$

El mismo resultado se puede obtener utilizando una tabla de Poisson: cuando $\mu = 6$, a partir de la tabla $P(X = 4)$ es 0.135.

En la tabla 4.7 se hace una comparación entre algunas de las características fundamentales de las distribuciones de Poisson y binomial.

TABLA 4.7 Comparación entre las distribuciones binomial y de Poisson

	Binomial	De Poisson
Resultados posibles	enteros 0 a n	enteros 0 a $+\infty$
Observaciones	conteo de éxitos o de fracasos	conteo sólo de los éxitos
Parámetros	n y p	μ

EJERCICIOS

- Utilice una tabla de Poisson para determinar cada una de las siguientes probabilidades:

Media	Probabilidad de
1	1
1.5	0
2	1 o menos
3	1 o menos
3	más de 3
3	3
4	3
4.2	más de 5

- Vuelva a desarrollar los siguientes ejercicios en los que se utilizó la fórmula de Poisson, pero ahora utilizando la tabla acumulativa de Poisson.
 - Ejercicio 1 (pág. 141)
 - Ejercicio 3 (página 141)
 - Ejercicio 2a (página 141)
 - Ejercicio 2c (página 141)
- A una construcción llegan camiones de carga a una razón media de 2.8 camiones/hora. Obtenga la probabilidad de tener tres o más camiones que lleguen en un:
 - lapso de 30 minutos.
 - lapso de una hora.
 - lapso de dos horas.
- En un conmutador se registran llamadas telefónicas a razón de 4.6 llamadas por minuto. Calcule la probabilidad de que se presente cada una de estas ocurrencias en un intervalo de un minuto:
 - Exactamente dos llamadas.
 - Por lo menos dos llamadas.
 - Ninguna llamada.
 - De dos a seis llamadas.

5. Los accidentes de una enorme fábrica se aproximan mediante la distribución de Poisson, con una media de tres accidentes/mes. Estime la probabilidad de que en un determinado mes haya:
 - a. cero accidentes.
 - b. un accidente.
 - c. tres o cuatro accidentes.
6. Si el 3% de las personas que viven en una gran ciudad son empleados del gobierno, determine la probabilidad de encontrar personas que no trabajen para el gobierno en una muestra aleatoria de 50 habitantes de dicha ciudad. ¿Cuál es la probabilidad de encontrar 3 o menos en la muestra?
7. El 2% de las cartas que se envían a cierta ciudad no llevan los timbres postales correctos. En 400 de dichas cartas:
 - a. ¿Cuántos timbres incorrectos se esperaría encontrar?
 - b. ¿Cuál es la probabilidad de hallar 5 o menos cartas con timbres incorrectos?
 - c. ¿Cuál es la probabilidad de encontrar más de 5 cartas con timbres incorrectos?
 - d. ¿Cuál es la probabilidad de encontrar 5 o más cartas con timbres incorrectos?
8. La probabilidad de vender un seguro de vida a personas que contesten un anuncio especial se estima que es de 0.01. Sobre esta base, si 1000 personas contestan el anuncio, ¿cuál es la probabilidad de que:
 - a. nadie compre un seguro?
 - b. por lo menos uno compre un seguro?
 - c. más de 10 compren seguros?
9. ¿En qué se diferenciarían las respuestas del ejercicio 7, si el porcentaje de cartas con timbres postales incorrectos fuera del 0.4%?

OTRAS DISTRIBUCIONES DISCRETAS

Ahora vale la pena mencionar otras dos distribuciones, las cuales son variantes de la distribución binomial.

La distribución *multinomial* se utiliza en situaciones en las que hay *más de dos* resultados mutuamente excluyentes. Como en el caso de la binomial, se tienen los requisitos de ensayos independientes con probabilidad constante. Un ejemplo de una variable multinomial es la tirada de un dado, la cual tiene seis resultados posibles. A diferencia de los ejemplos anteriores, que tratan este ejemplo clasificando los seis resultados en *dos* clases, el método multinomial considera de tres a seis clases, dependiendo de la forma cómo se categoricen los resultados. Por ejemplo, se podrían presentar las siguientes tres clases: 1 y 2, 3 y 4, y 5 y 6. Otro ejemplo es sacar canicas de una urna que contiene muchas de diversos colores. Cada color se puede considerar como una categoría, o se pueden agrupar ciertos colores (por ejemplo, rojo y verde se pueden conjuntar en la misma categoría).

La probabilidad multinomial que, en n observaciones, tiene como resultado E_1 ocurre x_1 veces, E_2 sucede x_2 veces, . . . y en E_r se presenta x_r veces, se expresa mediante la fórmula

$$\frac{n!}{n_1! n_2! \cdots n_k!} (P_1^{n_1} P_2^{n_2} \cdots P_k^{n_k})$$

en la cual

$$n = \sum_{i=1}^k n_i$$

y P_i es la probabilidad del evento E_i en cada observación.

Ejemplo 6 Un administrador sabe que el 80% de la producción de una máquina es aceptable, el 15% requiere de cierto trabajo adicional y el 5% se debe descartar. En una muestra de $n = 10$, ¿cuál es la probabilidad de obtener 8 partes buenas, 2 que necesiten ser repetidas, y ninguna que se deba descartar?

$$\frac{10!}{8! 2! 0!} (0.80)^8 (0.15)^2 (0.05)^0 = 0.17$$

La distribución *hipergeométrica* se relaciona con situaciones con dos o más resultados, en las que la probabilidad de tener éxito, *cambie* de ensayo a ensayo. Por ejemplo, suponga que hay 10 nombres escritos en pequeños pedazos de papel, los cuales se colocan en un sombrero. Su nombre se encuentra en uno de ellos. Si se mezclan los papeles y se saca un nombre cualquiera la probabilidad de que sea el suyo es $\frac{1}{10}$. Al leerse el nombre en voz alta, resulta no ser el suyo. A continuación, se saca un segundo nombre de los nueve restantes. Ahora la probabilidad de que sea el suyo es de $\frac{1}{9}$, mientras que anteriormente era de $\frac{1}{10}$. De ahí que, la probabilidad de obtener dicho resultado en el segundo intento es *condicional* (dependiente) de los resultados del primero. Por ejemplo, si ese primer nombre hubiera sido el suyo, entonces la probabilidad de que el segundo nombre fuera suyo se reduciría a cero, suponiendo que los nombres no se vuelven a colocar en el sombrero una vez que se sacan.

En general, las probabilidades hipergeométricas se pueden obtener a partir de la fórmula

$$P(x|N, r, n) = \frac{\binom{N-r}{n-x} \binom{r}{x}}{\binom{N}{n}}$$

en la cual N es el tamaño de la población, n es el tamaño muestral, r es el número de éxitos de la población y x es el número de éxitos de la muestra.

Ejemplo 7 En una caja de 10 fusibles, dos de ellos están defectuosos. Si se examina una muestra aleatoria de cuatro fusibles, ¿cuál es la probabilidad de encontrar: a) ninguno defectuoso; b) uno defectuoso; c) uno o menos defectuosos?

Solución:

Se obtiene que $N = 10$, $r = 2$ y $n = 4$.

(a)
$$P(x = 0) = \frac{\binom{10-2}{4-0} \binom{2}{0}}{\binom{10}{4}} = \frac{\binom{8}{4} \binom{2}{0}}{\binom{10}{4}} = 0.333$$

(b)
$$P(x = 1) = \frac{\binom{10-2}{4-1} \binom{2}{1}}{\binom{10}{4}} = \frac{\binom{8}{3} \binom{2}{1}}{\binom{10}{4}} = 0.533$$

(c)
$$P(x \leq 1) = P(0) + P(1) = 0.333 + 0.533 = 0.866$$

En la tabla 4.8 se enumeran los modelos de probabilidad que se han estudiado, y las situaciones en las que dichos modelos resultan más útiles.

TABLA 4.8 Aplicaciones de distribuciones discretas probabilísticas

Distribución		Ejemplos	Supuesto
binomial	dos resultados	lanzamiento de una moneda prueba de verdadero-falso defectuoso-no defectuoso	observación independiente probabilidad constante
de Poisson	solamente ocurrencias	accidentes/años defectos/metro llamadas/minuto	observación independiente probabilidad constante
hipergeométrica	dos o más resultados	muestreo sin sustitución	observaciones dependientes
multinomial	más de dos resultados	prueba de opción múltiple	observación independiente probabilidad constante

EJERCICIOS

- ★ 1. Una reciente encuesta nacional señaló que a las 9 p.m. del sábado, 40% del teleauditorio sintoniza el canal A, 30%, el canal B y 30%, el canal C.
 - a. En una muestra aleatoria de 10 televidentes, ¿cuántos se esperaría que estarían viendo cada canal?
 - b. ¿Cuál es la probabilidad de que todos sintonicen el canal A?
 - c. ¿Cuál es la probabilidad de que 4 sintonicen el canal A, 3 el B y 3 el C?
- 2. Tres individuos venden refacciones. El vendedor A provee del 50% de las partes, el B del 40%, y el C del 10%. Si se seleccionan aleatoriamente 5 refacciones del suministro total y se le buscan defectos,
 - a. ¿cuál es la probabilidad de que los 5 los haya proporcionado el vendedor A?
 - b. ¿Cuál es la probabilidad de que 2 sean del vendedor A, 2 del B y 1 del C?
- 3. El 30% de los alumnos de una universidad son de primer año, el 30%, de segundo, el 20%, de tercero y el otro 20%, de cuarto año. De una lista se toma una