Unidad 6

- Distribuciones probabilísticas contínuas

Distribuciones probabilísticas continuas

Objetivos del capítulo

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

1. Enunciar las diferencias básicas entre distribuciones continuas y discretas probabilísticas.
2. Explicar por qué las probabilidades continuas se relacionan con las áreas.
3. Describir las características principales de una distribución normal.
4. Convertir cualquier distribución normal en una distribución normal estándar.
5. Utilizar la tabla de áreas de distribución normal para obtener probabilidades.
6. Calcular el área bajo una curva normal entre dos puntos cualesquiera.
7. Utilizar la distribución normal para aproximar probabilidades binomiales.
8. Resolver problemas en los que intervenga la distribución uniforme o exponencial.

Sinopsis del capítulo

- Introducción
- Distribución uniforme
- Distribuciones normales
  - Características de las distribuciones normales
  - Distribución normal como modelo
  - Distribución normal estándar
  - Tabla normal estándar
  - La distribución normal como una aproximación a la distribución binomial
- Distribución exponencial
- Resumen

INTRODUCCIÓN

CUANDO UNA variable aleatoria discreta presenta un gran número de resultados posibles, o cuando la variable aleatoria que se está considerando, es continua, no se pueden utilizar las distribuciones probabilísticas discretas, como la de Poisson y la binomial, para obtener probabilidades importantes. Una variable discreta con muchos resultados posibles requeriría una tabla muy extensa, o un esfuerzo muy grande para utilizar una fórmula a fin de obtener probabilidades. Una variable continua, debido a que los resultados incluyen valores enteros y no enteros, no se puede manejar en forma adecuada mediante una distribución discreta.

La flecha giratoria que se encuentra en el centro de la circunferencia, mostrada en la figura 5.1, ilustra el concepto de una variable continua. Una vez que se hace girar, la punta de la flecha podrá detenerse en cualquier punto dentro del círculo, y no precisamente en alguno de los valores enteros. Aun aceptando las limitaciones propias de hacer mediciones alrededor del círculo, existe un número extremadamente grande de posibles puntos de reposo.

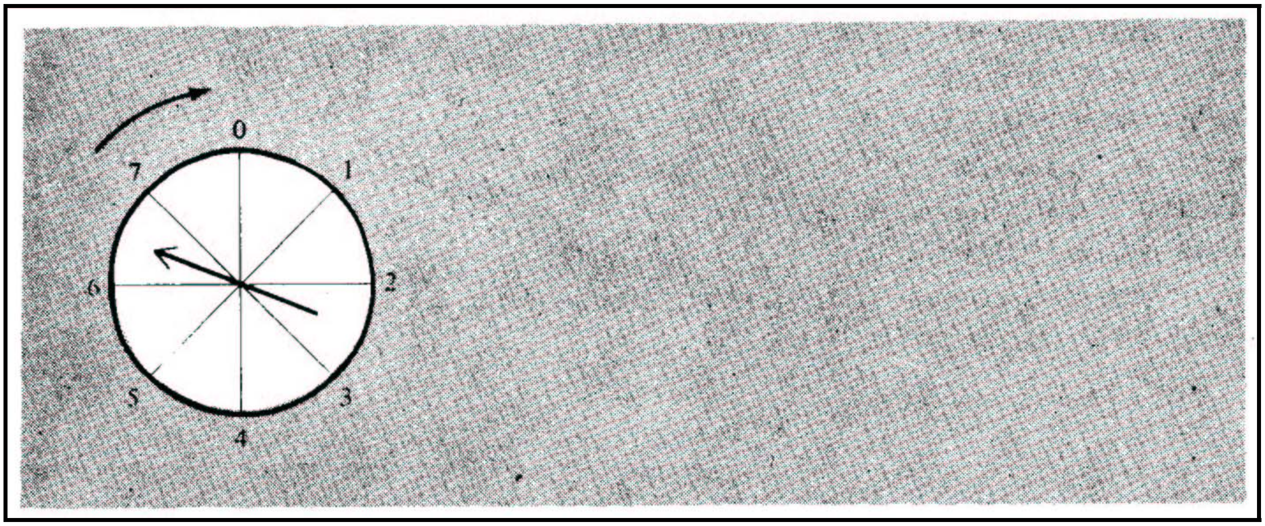

FIGURA 5.1 La flecha giratoria puede quedar en reposo en cualquiera de un número ilimitado de posiciones.

Por ejemplo, imagínese que el círculo se divide en 8000 pares iguales en lugar de ocho. Si se supone que cada posición tiene la misma probabilidad de ser el punto de reposo, se llegará a la siguiente conclusión: Dado que hay muchos resultados posibles, la probabilidad de que la punta señale un valor determinado es tan pequeña, que, para objetivos prácticos se deberá considerar como aproximadamente igual a cero.* De hecho, la tecnología moderna nos permite identificar por lo menos un millón de posiciones diferentes, por lo que la probabilidad de que la flecha se detenga en determinado punto sería 1/1 000 000 6 0.000001.

Debido a esta peculiaridad, realmente no tiene sentido hablar en términos de la probabilidad de obtener algún resultado específico, como se hizo al comentar sobre las distribuciones probabilísticas discontinuas. Por tanto, el análisis de variables continuas

* *Esto significa que es virtualmente imposible predecir el punto exacto donde se detendrá la flecha.*

tiende a concentrarse en la probabilidad de que una variable aleatoria asuma un valor dentro de algún intervalo. De este modo, la probabilidad de que la flecha se detenga, ya sea en 3 6 4, es casi de cero, en tanto que la probabilidad de que lo haga entre estos dos valores, no lo es. Como el círculo está dividido en ocho áreas iguales, convendrá asignarle una probabilidad de $^1/_8$ al resultado "se detuvo entre 3 y 4", según se ilustra en la figura 5.2. Además, como $P(x = 3)$ y $P(x = 4)$ son ambos aproximadamente igual a cero, no es necesario distinguir entre $3 < x < 4$ y $3 \leq x \leq 4$, como sucedió en el caso de las probabilidades discretas.

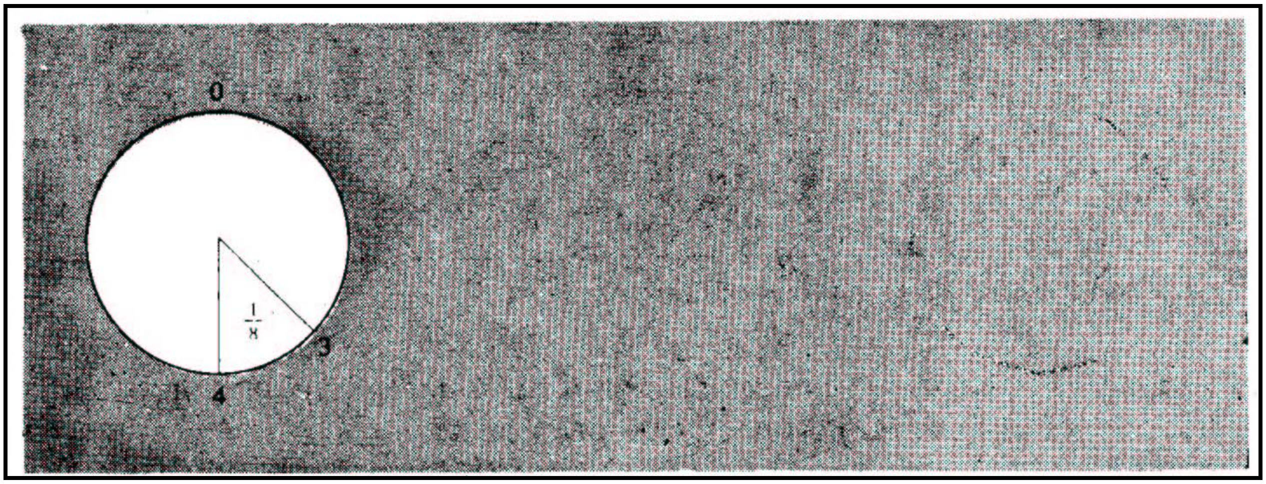

FIGURA 5.2 La probabilidad de que la punta de la flecha se detenga entre dos puntos es igual al porcentaje del área entre ambos puntos.

En forma similar, se asignaría una probabilidad de 25% al evento "la flecha se detiene entre los puntos 4 y 6 (un cuarto del círculo)". Y no hay razón para limitar los intervalos a valores enteros, más que por comodidad. Por ejemplo, la probabilidad de observar un valor entre 3.217 y 4.217 (nota: 4.217 - 3.217 = 1) sería también de $^1/_8$, y la probabilidad de observar un valor entre 3.5 y 4 [es decir, 4 - 3.5 = 0.5] es $^1/_{16}$. De este modo, la probabilidad de que la flecha se detenga entre dos puntos cualesquiera es igual al porcentaje del área correspondiente a esos dos puntos. (Ver figura 5.3.) Además, un círculo dos veces mayor que éste, tendría las mismas probabilidades si su perímetro estuviera numerado de la misma manera (ver figura 5.4). Por tanto, en el caso de una variable aleatoria continua, se determina la probabilidad al obtener el porcentaje del área entre dos valores.

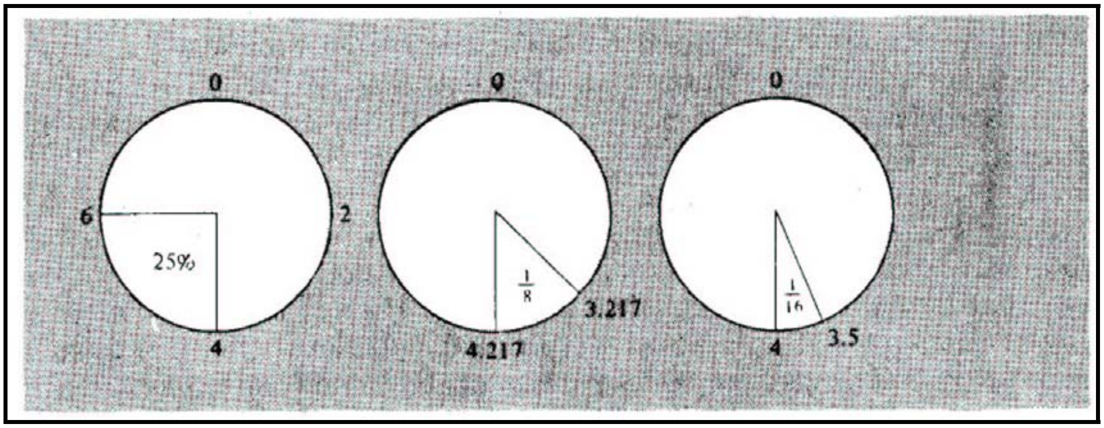

FIGURA 5.3 Otros ejemplos de la probabilidad y el área entre dos puntos.

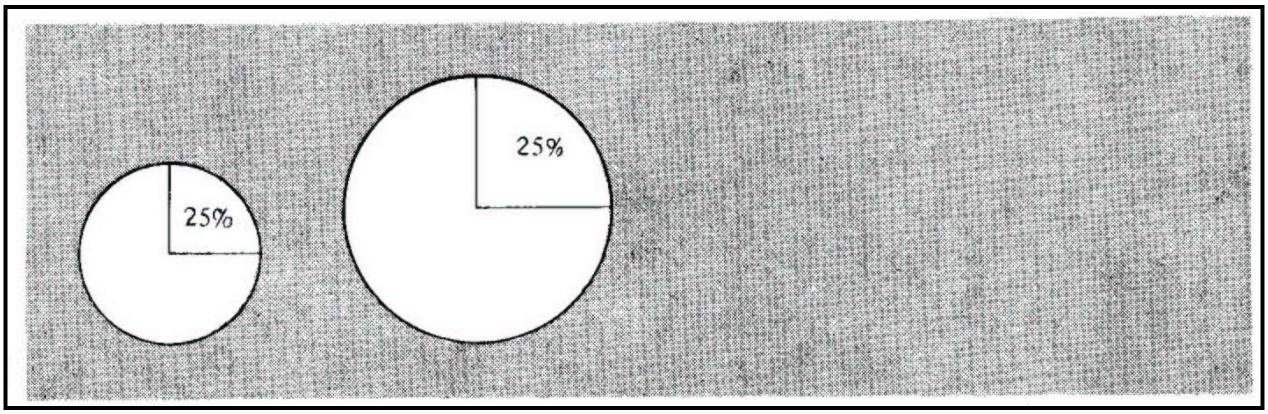

FIGURA 5.4 El tamaño del círculo es indiferente.

En las páginas siguientes se comentarán tres distribuciones continuas probabilísticas: las distribuciones uniformes, normales y exponenciales.

DISTRIBUCIÓN UNIFORME

Cuando una variable aleatoria asume cualquier valor en una escala continua entre dos puntos, de tal forma que ningún valor sea más probable que otro, entonces las probabilidades asociadas con la variable aleatoria se pueden describir mediante la distribución uniforme. El ejemplo anterior de la flecha pertenece a esta categoría: todos los puntos situados en la circunferencia del círculo tiene la misma probabilidad. Gráficamente, la distribución uniforme se representa como un rectángulo limitado por los puntos a y b, los cuales constituyen el intervalo de resultados posibles. Ver figura 5.5(a). La altura del rectángulo se considera igual a 1.00 y el área del mismo es igual a 100%. En consecuencia, el área bajo el rectángulo entre dos puntos cualesquiera c y d es igual al porcentaje del intervalo total incluido entre c y d.

$$P(c \le x \le d) = \frac{d-c}{b-a}$$

Ver figura 5.5(b)

Por ejemplo, supóngase que un vendedor telefonea a las oficinas centrales entre 3 y 4 de la tarde, y, según el registro, en ningún momento de ese lapso hay más probabilidades de que el vendedor llame. Dado que el tiempo se mide en una escala continua, la probabilidad de que se registre una llamada entre dos puntos cualesquiera en el tiempo es igual a la razón de ese tiempo al intervalo de 1 hora. De ahí que, la probabilidad de que se registre la llamada entre 3:00 y 3:15 es $^{15}/_{60} = 0.25$. La probabilidad de que ocurra exactamente a las 3:15 se considera aproximadamente igual a cero, ya que la llamada puede presentarse en un minuto, así como en cualquiera de los puntos intermedios (infinitos), como, 3:15 y 0.00333 segundos. Cuando se dice que la probabilidad de que llame exactamente a las 3:15 es igual a cero, no significa que una llamada no pueda ocurrir en ese momento, sino que sólo con una gama infinita de posibilidades, sería imposible señalar la hora exacta en la que se pueda predecir que se registrará la llamada. Ver figura 5.5(c).

Para ciertas aplicaciones, es necesario utilizar la media y la variancia de una distri bución probabilística.

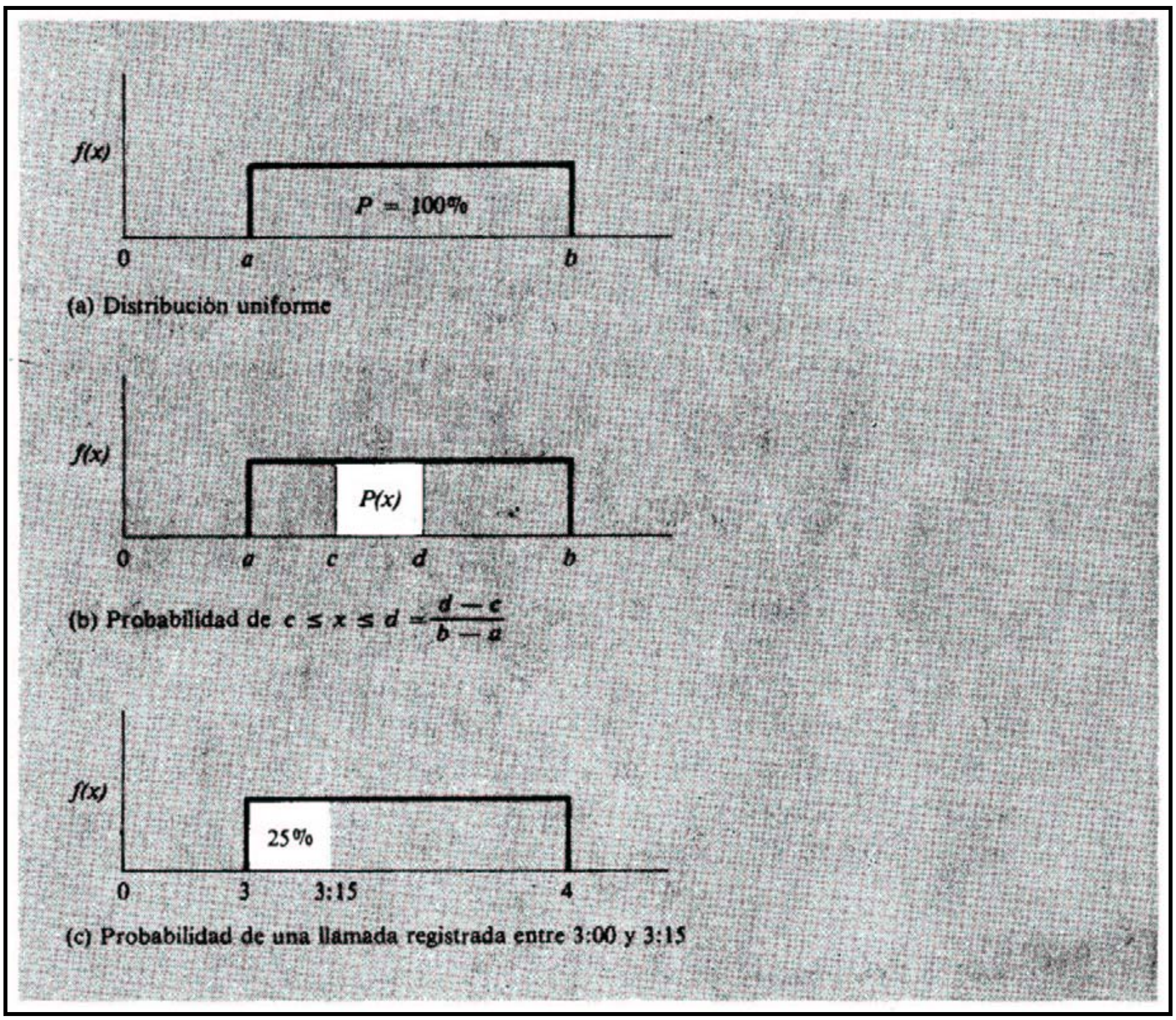

FIGURA 5.5

La media de una distribución uniforme con puntos extremos a y b es

$$\mu = \frac{a + b}{2}$$

y la variancia es

$$\sigma^2 = \frac{(b - a)^2}{12}$$

EJERCICIOS

1. Las ventas de combustible en una gasolinería tienen una media de 40 000 litros por día y un mínimo de 30 000 litros por día. Suponga que una distribución uniforme es apropiada.
   a. Determine las ventas máximas diarias.
   b. ¿Qué porcentaje de días, las ventas excederán de 34 000 litros?
2. Una pequeña compañía corta y vende leños (para chimeneas), cuya longitud varía uniformemente entre 2 y 3 pies.
   a. ¿Cuál es la longitud promedio de un leño?
   b. ¿Cuál es la probabilidad de que cualquier leño dado sea:
      1) mayor que 2.6 pies?
      2) mayor que 3 pies?
      3) menor que el promedio?
      4) exactamente de 2 pies?
      5) entre 2 y 3 pies?
3. Suponga que la temperatura más alta diaria durante el mes de enero en una área rural ha variado uniformemente entre 0°C y 6°C, según registros anteriores.
   a. ¿Qué porcentaje de días se puede esperar que alcancen una temperatura máxima de 3.5°C?
   b. Si el meteréologo quiere minimizar el, error en la predicción, ¿qué temperatura deberá predecir?
   c. ¿Cuál es la probabilidad de que en cualquier día de enero la temperatura no exceda de 1°C?
4. Se sabe que la cantidad de helado vendido el martes en una fuente de sodas, está distribuida uniformemente y varía entre 20 y 50 litros.
   a. ¿Cuál es la probabilidad de vender 40 o más litros el martes?
   b. ¿Cuál es la probabilidad de vender 40 o más litros el lunes?
   c. Si la fuente de sodas obtiene una utilidad de $ 0.30 por litro, ¿cuánto se espera obtener de las ventas del martes?
   d. ¿Cuál es la probabilidad de que la utilidad del martes sea menor de $ 7.50?

DISTRIBUCIONES NORMALES

Las distribuciones normales ocupan un lugar importante, tanto en la estadística teórica como en la aplicada, por numerosas razones. Una de ellas es que suelen coincidir muy cercanamente con las distribuciones de frecuencia observadas de muchas mediciones naturales y físicas. Otro motivo es que se pueden utilizar para aproximar probabilidades binomiales cuando n es grande. Sin embargo, lo que hace más importante a la distribución normal es que las distribuciones de medias muestrales y proporciones de grandes muestras tienden a distribuirse normalmente, lo que tiene repercusiones importantes en el muestreo.

Las distribuciones normales fueron "descubiertas" por primera vez en el siglo XVIII. Los astronómos y otros científicos observaron, con cierto asombro, que las mediciones repetidas de la misma cantidad (como la distancia a la Luna o la masa de un objeto) tendían a variar, y que al reunir grandes cantidades de estas mediciones en una distribución de frecuencias, tendía a reaparecer constantemente un perfil semejante al presentado en la figura 5.6. Como esta forma se relacionaba con errores de medición, pronto se le conoció como "la distribución normal de los errores" o simplemente, como la "distribución normal". Posteriormente se descubrió que la distribución se puede aproximar mucho mediante una distribución matemática continua, como la de la figura 5.7. A esta distribución a veces se le conoce como distribución gausiana, en reconocimiento a las aportaciones del Karl Gauss (1777-1855) a la teoría matemática de la distribución normal.

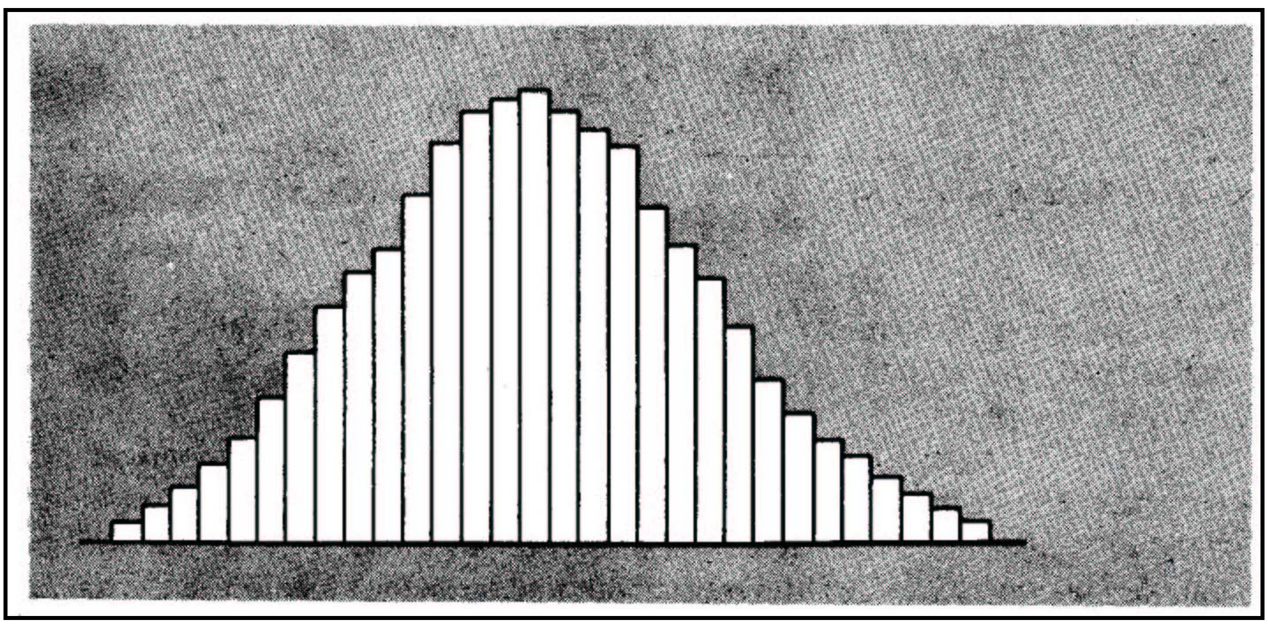

FIGURA 5.6 Distribuciones de frecuencia de observaciones que generalmente tienen la misma forma.

Características de las distribuciones normales

Las curvas normales tienen ciertas características especiales en términos de su configuración y de la forma como están especificadas y como se utilizan para obtener probabilidades.

La gráfica de una distribución normal, según se ha visto, se asemeja mucho a una campana: es suave, unimodal y simétrica con respecto a su media. Menos evidente es el hecho de que la curva se extiende hacia el infinito en ambas direcciones a partir de la media.

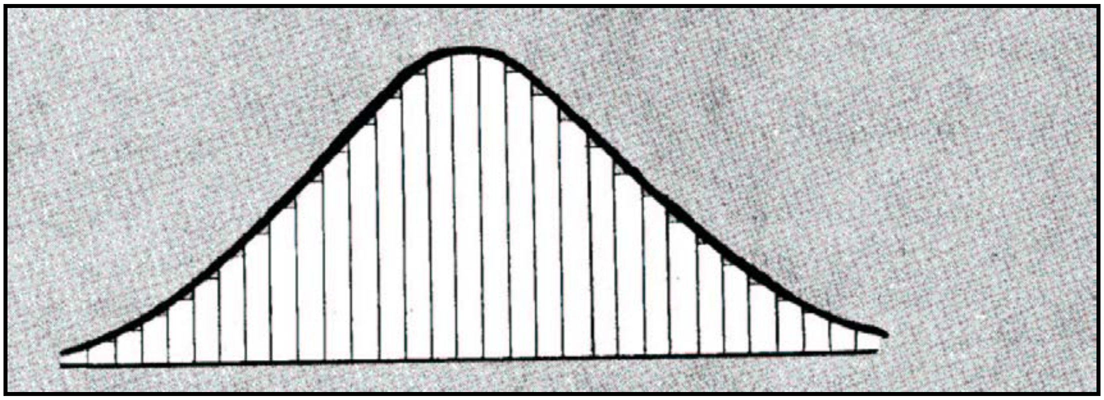

FIGURA 5.7 Curva continua que aproxima la distribución de frecuencias observadas.

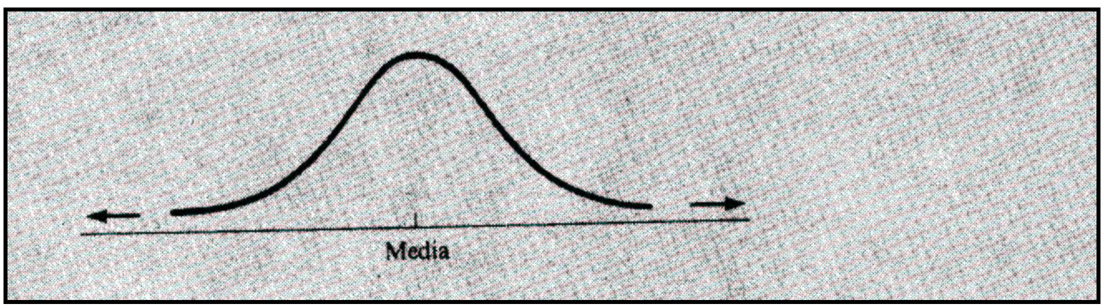

FIGURA 5.8 Curva normal típica.

Se acerca más al eje horizontal a medida que aumenta la distancia con respecto a la media, pero nunca toca realmente el eje. Teóricamente, los valores posibles van de $-\infty$ a $+\infty$. Ver figura 5.8.

Otra característica importante es que es posible especificar ampliamente una distribución normal por medio de dos parámetros: la media y la desviación estándar. En otras palabras, existe una distribución normal única para cada combinación de una media y una desviación estándar. Diferentes combinaciones de la media y la desviación estándar producen curvas normales distintas. Como las medias y las desviaciones estándar se miden en una escala continua, el número posible de distribuciones normales o curvas es infinito. Algunas de estas posibilidades se ilustran en la figura 5.9.

El área total bajo cualquier curva normal representa el 100% de la probabilidad relacionada con dicha variable. Además, como la curva es simétrica respecto a su media, la probabilidad de obtener un valor menor que la media es del 50%, al igual que la de observar un valor mayor que la media. La probabilidad de predecir con exactitud cualquier valor es cero, ya que la escala de medición es continua. De ahí que, la probabilidad de observar un valor que sea exactamente igual a la media, también sea cero.

La probabilidad de que una variable aleatoria que está distribuida normalmente, asuma un valor entre dos puntos cualesquiera es igual al área bajo la curva normal entre esos dos puntos, según se ilustra en la figura 5.10.

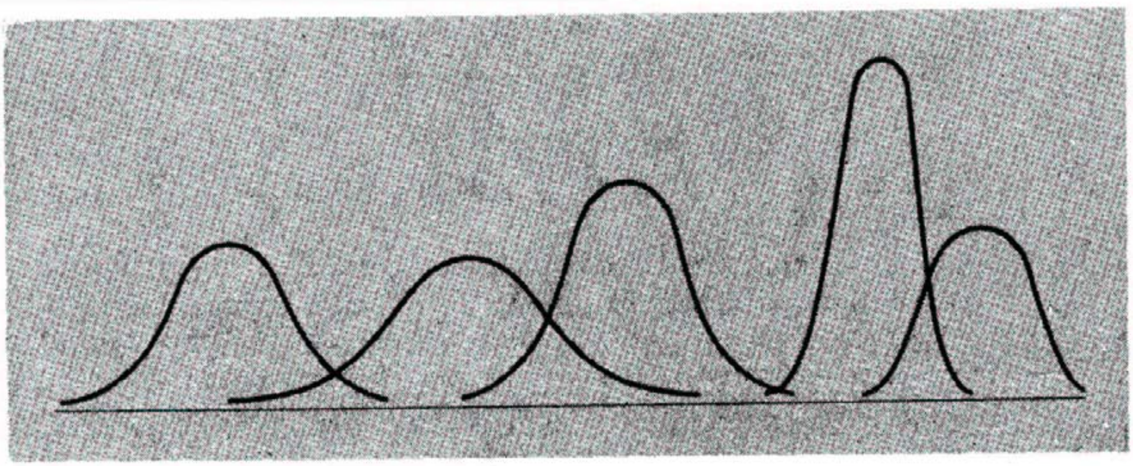

FIGURA 5.9 Son infinitas las combinaciones de media y desviación estándar.

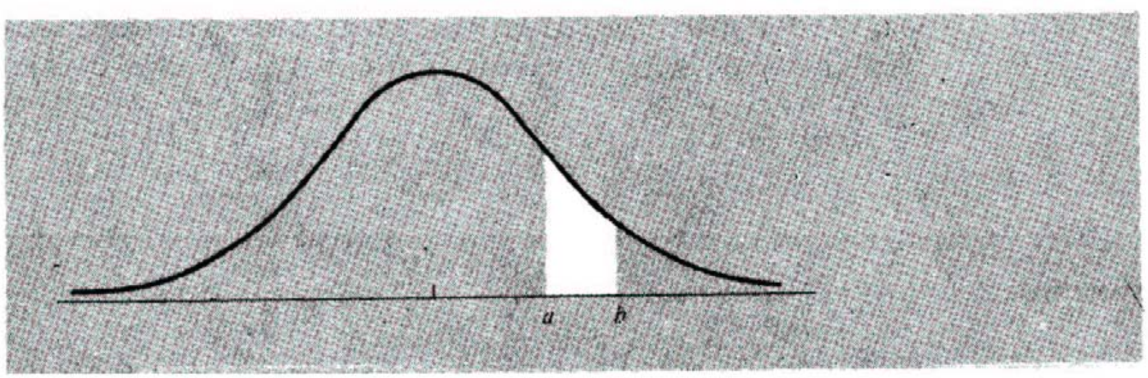

FIGURA 5.10 P(a < x < b) = área bajo la curva entre a y b.

La probabilidad de que una variable aleatoria tenga un valor entre dos puntos cualesquiera es igual al área bajo la curva normal entre esos dos puntos.

Una consecuencia importante del hecho de que una curva normal se pueda especificar completamente por su media y su desviación estándar, es que el área bajo la curva entre cualquier punto y la media es una función sólo del número de desviaciones estándar que el punto dista de la media. Esta es la clave para calcular probabilidades para la curva normal.

En resumen, las curvas normales tienen las siguientes características:

1. La curva normal tiene forma de campana.
2. Es simétrica con respecto a la media de la distribución.
3. Se extiende de - ∞ a + ∞.
4. Cada distribución normal es completamente especificada por su media y desviación estándar; existe una distribución normal diferente para cada combinación de media y desviación estándar.
5. El área total bajo la curva normal se considera que es el 100%.
6. El área bajo la curva entre dos puntos es la probabilidad de que una variable distribuida normalmente asuma un valor entre ellos.
7. Dado que existe un número ilimitado de valores en el intervalo que va de - ∞ a + ∞, la probabilidad de que una variable aleatoria distribuida con normalidad sea exactamente igual a cualquier valor dado es casi de cero. Por tanto, las probabilidades siempre serán para un intervalo de valores.
8. El área bajo la curva entre la media y cualquier otro punto es una función del número de desviaciones estándar que el punto dista de la media.

La distribución normal como modelo

Es esencial darse cuenta de que una distribución normal es una distribución teórica. En el caso de mediciones físicas que se tengan que agrupar en una distribución de frecuencia, constituye una distribución ideal ya que ningún conjunto de datos reales coincidiría exactamente con ella. Por ejemplo, los datos reales podrían no variar entre - ∞ y + ∞. Y las limitaciones de los instrumentos de medición eliminarían eficazmente muchos otros valores potenciales. Sin embargo, tales deficiencias son superadas por la facilidad de utilizar la distribución normal para obtener probabilidades, y por el hecho de que la distribución normal proporciona aproximaciones muy razonables a los datos reales. De este modo, cuando se afirma que una variable aleatoria (física) está distribuida normalmente, se deberá interpretar como que una distribución de frecuencias de sus resultados posibles se puede aproximar con certeza razonable, utilizando una distribución de probabilidad normal. Por tanto, la curva normal es un modelo.

Distribución normal estándar

En realidad, la distribución normal es una "familia" de distribuciones infinitamente grande, hay una para cada combinación posible de la media y la desviación estándar. En consecuencia, sería inútil intentar elaborar las tablas suficientes para satisfacer las necesidades de los posibles usuarios. Por otra parte, la fórmula para la distribución normal no es muy adecuada para este fin debido a su complejidad.* Existe, sin

---

* *La fórmula para la distribución normal es* $$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-[(1/2)[(x-\mu)/\sigma]^2]}$$

embargo, una alternativa sencilla que evita estos problemas. Conceptualmente se asemeja a la determinación de probabilidades por la flecha giratoria. En dicho problema se dijo que el tamaño del círculo no era importante: sino que la forma era el factor significativo. Siempre y cuando el área total del círculo se considere como el 100%, uno de cualquier tamaño daría probabilidades idénticas. Y lo mismo ocurre con distribuciones normales: el considerar el área bajo la curva como 100% estandariza la curva.

Si una variable está distribuida normalmente, entonces alrededor del 68% de sus valores quedarán dentro de una desviación estándar de la media; 95.5% caerán dentro de dos desviaciones estándares de la media; y casi el 99.7% quedarán dentro de tres desviaciones estándares de la media. Esta idea se ilustra en la figura 5.11. Además, esto es cierto independientemente de que la media y la desviación estándar presenten una determinada distribución normal: esto se cumple en el caso de todas las distribuciones normales.

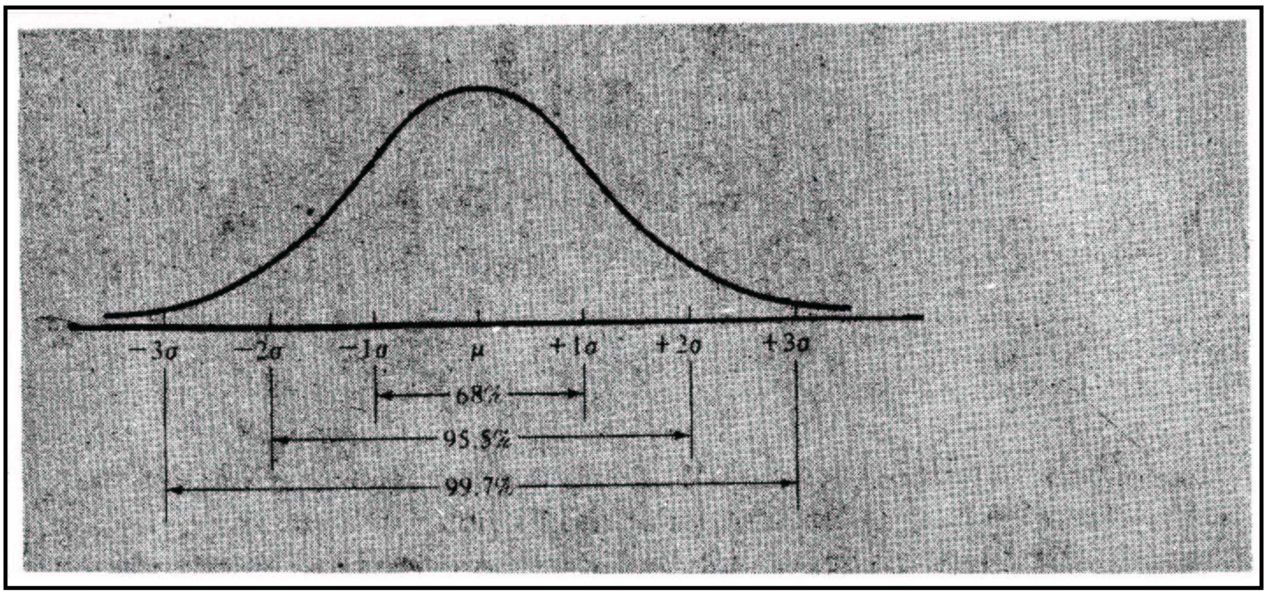

FIGURA 5.11 Area bajo una curva normal dentro de 1, 2 y 3 desviaciones estándar de la media.

A continuación se mostrará la forma cómo estos y otros porcentajes se pueden determinar. Pero por ahora, se debe reflexionar sobre la importancia de este hecho. Esto significa que, el problema de trabajar con una familia infinita de distribuciones normales se puede evitar completamente si es posible manejar valores relativos en lugar de valores reales. Esto equivale a utilizar la media como punto de referencia, y la desviación estándar como una medida de la desviación de dicho punto de referencia. Esta nueva escala comúnmente se conoce como escala z.

Considérese una distribución normal con una media de 100.0, y una desviación estándar de 10.0, según se ilustra en la figura 5.12. Se puede convertir esta escala real a una relativa, o estandarizada, mediante la sustitución de los valores reales por el "número de desviaciones estándar de las medias de distribución".

Sólo unos cuantos valores elegidos se muestran en la figura 5.12, no obstante, es posible aplicar el mismo concepto a cualquier valor de distribución. Por tanto, el valor

90 está - 10 abajo de la media, $^{10}/_{10}$ = - 1 desviación estándar; 120 está + 20 sobre la media, o $^{20}/_{10}$ = 2 desviaciones estándar; y así sucesivamente. El valor 95 sería -0.5 desviaciones estándar $^{5}/_{10}$ desviaciones estándar por debajo de la media) y 107 sería -0.7 desviaciones estándares $^{7}/_{10}$ desviaciones estándar sobre la media).

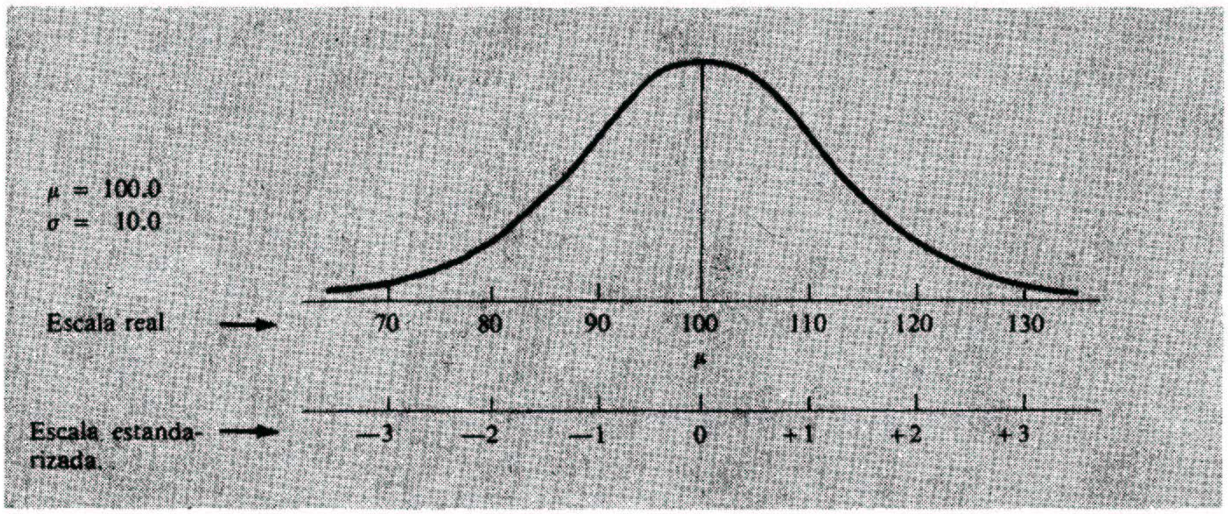

FIGURA 5.12 Comparación de las escalas reales y estandarizadas.

Es posible resumir este procedimiento de la siguiente manera: Convierta la diferencia real entre la media y algún otro valor de distribución a una diferencia relativa, expresando dicha diferencia en términos del número de desviaciones estándar de la media. Algebraicamente esto se puede representar como se muestra en el siguiente cuadro.

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

en donde

z = número de desviaciones estándar a partir de la media

x = algunos valores de interés

μ = media de una distribución normal

σ = desviación estándar

Obsérvese que z tiene signo menos en el caso de valores de x menores que la media, y un signo más para valores mayores que la media.

A continuación se presentan algunos ejemplos de la conversión de la diferencia real entre la media y otro valor a la distancia relativa, en términos del número de desviaciones estándar.

μ Media	σ Desviación estandar	x Valor de interés	$x - \mu$ Diferencia	$(x - \mu)/\sigma = z$ Diferencia relativa
40	1	42	2	+2
25	2	23	−2	−1
30	2.5	37.5	7.5	+3
18	3	13.5	−4.5	−1.5
22	4	22	0	0

Asimismo es necesario ser capaces de trabajar en orden inverso, pasando de los valores de *z* a valores reales. Por ejemplo, se quiere saber qué valor real sería el equivalente de $z = +2$. Suponiendo que se conoce la media y la desviación estándar, y que se considera la distribución normal, la conversión asume la forma

$$\text{valor real} = \mu + z\sigma$$

A continuación se dan algunos ejemplos de lo anterior.

μ Media	σ Desviación estándar	z	$\mu + z\sigma$ Cálculo	Valor real
20	1	+3	20 + 3(1)	23
50	3	−1	50 − 1(3)	47
60	2	−2	60 − 2(2)	56
72	5	+0.3	72 + 0.3(5)	73.5

Existe una gran ventaja en poder pensar y trabajar con valores relativos. Esto significa que en lugar de tener que emplear una familia ilimitada de distribuciones normales, se puede utilizar una sola distribución para todos los valores. Es posible convertir cualquier valor de cualquier distribución normal en un valor z, lo cual indicará a cuántas desviaciones estándar está ese valor de la media de la distribución. Esto permite determinar varias probabilidades con base en la curva normal, mediante el uso de una tabla estandarizada única, ideada sólo para este fin.

Tabla normal estándar

Las áreas bajo la curva para cualquier distribución normal se pueden encontrar, utilizando una tabla normal estándar y cambiando a unidades estándares la escala de unidades reales. La media de la distribución sirve como punto de referencia, y la desviación estándar como la unidad que mide distancias relativas a partir de la media. La tabla normal estándar fue ideada de manera que se pueda leer en unidades de z, que es el número de desviaciones estándar de la media. La tabla muestra el área bajo la curva (es decir, la probabilidad de que un valor quede en ese intervalo) entre la media y valores seleccionados de z. La parte sombreada de la figura 5.13 corresponde al área bajo la curva que se puede leer directamente en la tabla. Obsérvese que la media de la distribución es cero, ya que la media está a cero desviaciones estándar de sí misma.

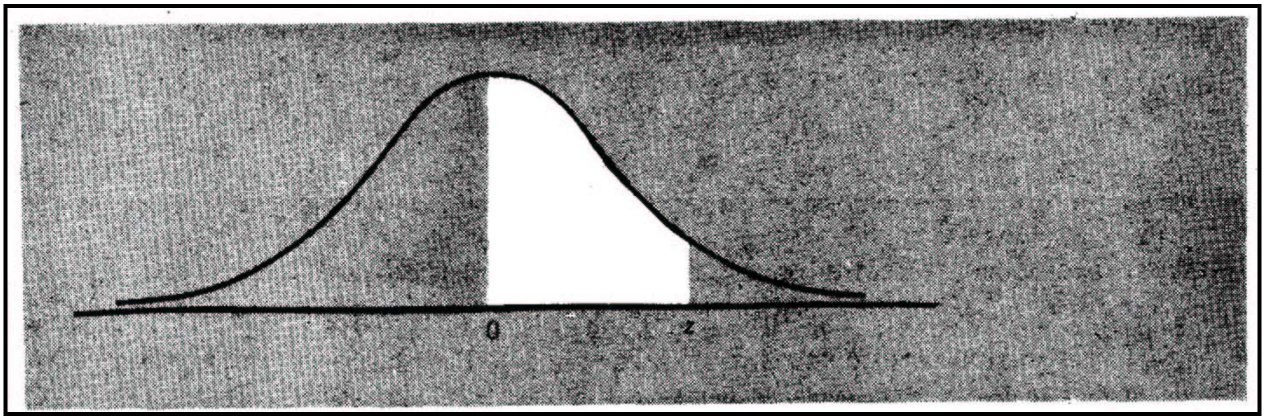

FIGURA 5.13 Área bajo una curva normal que se muestra. en una tabla normal.

Como la distribución normal es simétrica con respecto a su media, el lado izquierdo de la curva es una imagen especular del lado derecho. Y debido a esta simetría, en una tabla se acostumbra proporcionar sólo la mitad de la distribución. En otras palabras, para cada parte del lado izquierdo existe un segmento correspondiente en el lado derecho. Es común proporcionar una tabla para el lado derecho de la distribución. De este modo, si se requiere de una parte del lado izquierdo, estos valores se consideran como desviaciones positivas en lugar de negativas. Por ejemplo, el área bajo la curva entre la media y + 1 desviación estándar es exactamente igual al área bajo la curva entre la media y -1 desviación estándar, según se observa en la figura 5.14.

Ahora se puede considerar la tabla en sí misma. La tabla 5.1 se utilizará para los propósitos de explicación que aquí se planteen; la tabla G del Apéndice es idéntica a la 5.1. Está ordenada en términos de valores de z hasta dos decimales (al centésimo más próximo), como 2.78, 1.04 y 2.45. Una particularidad de una tabla normal típica es que los valores de z se presentan en dos partes, lo que siempre es un poco confuso para principiantes, pero que no causa problemas una vez que se está familiarizado con ella. Los valores de entero y el primer decimal (como 1.3, 2.5, 0.7) se enumeran hacia abajo en el lado izquierdo de la tabla, mientras que el último dígito aparece en la parte superior. Para demostrar el uso de la tabla, calcularemos ciertas áreas bajo la curva entre la media y z.

Supóngase que se quiere obtener el área entre la media y z, cuando z es igual a 1.25. En primer lugar se debe localizar el valor 1.2 en el lado izquierdo de la tabla, y posteriormente el 0.05 (5 es el último dígito), en su parte superior. El área bajo la curva se puede leer en la intersección de la fila z = 1.2 y la columna 0.05. El valor 0.3944 es el

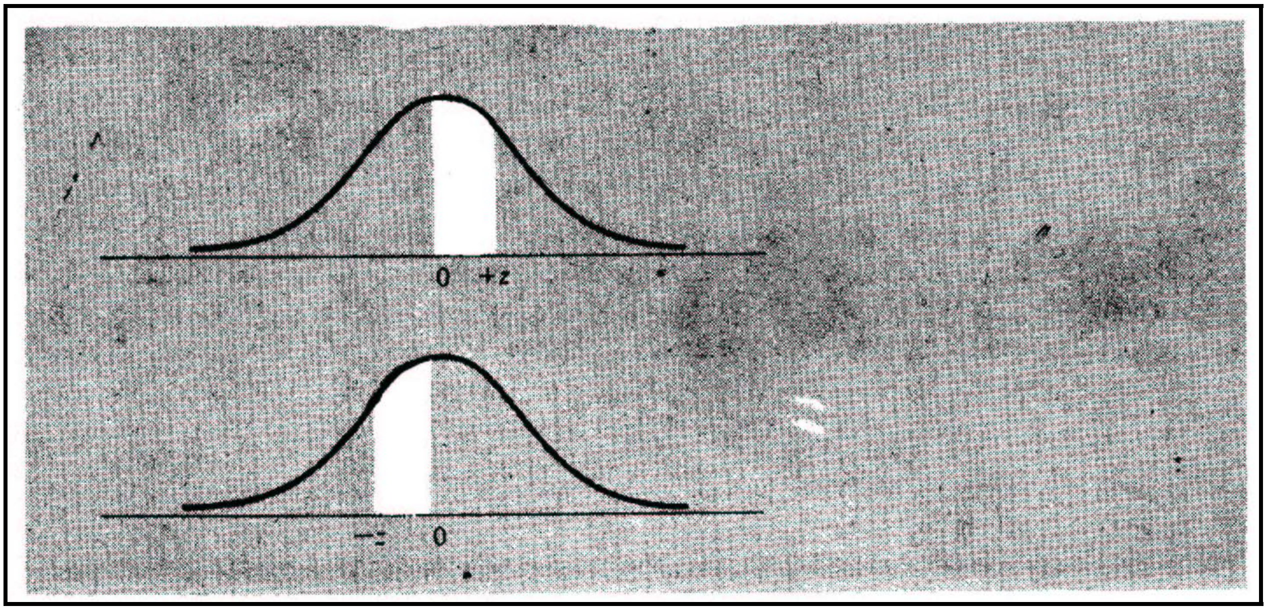

FIGURA 5.14 El área bajo la curva entre la media y + z es igual al área bajo la curva entre la media y - z.

porcentaje de área bajo la curva normal entre la media y z = 1.25 (ver figura 5.15). Evidentemente, este porcentaje equivale a la probabilidad de que una variable aleatoria distribuida normalmente tenga un valor entre la media y un punto 1.25 de desviaciones estándar sobre la media.

A continuación se dan algunos otros ejemplos (ver también figura 5.16).

z	Area entre la media y z
+1.00	0.3413
+1.50	0.4332
+2.13	0.4834
+2.77	0.4972

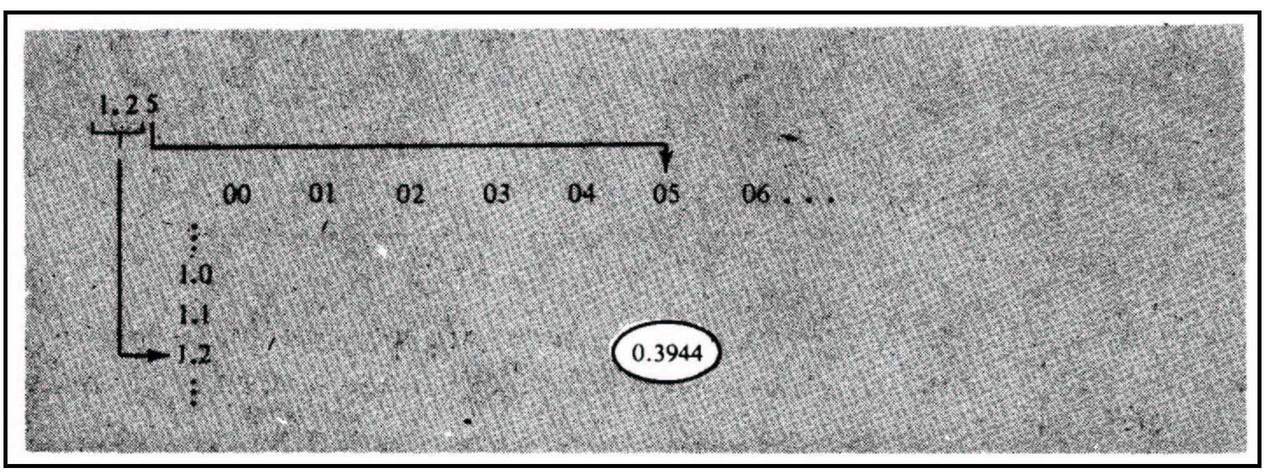

FIGURA 5.15 El área bajo una curva normal entre la media y z = 1.25.

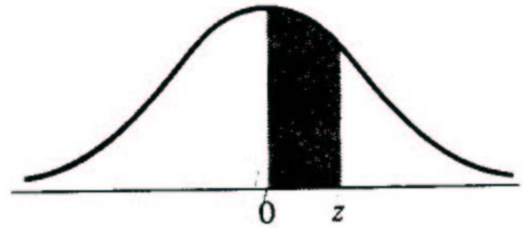

TABLA 5.1 Areas para la distribución probabilística normal estándar.

TABLA 5.1 Areas para la distribución probabilística normal estándar.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4986	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
4.0	0.49997									

Como la parte izquierda de la curva es esencialmente la misma que la de la derecha, si cada uno de los valores de z de la tabla anterior tienen un signo menos, las áreas bajo la curva serán las mismas.

La tabla normal también se puede utilizar para obtener el área bajo la curva más allá de un valor dado de z. La clave en este caso es que la mitad del área es el 50%, y, por tanto, el área más allá de z es 50% - el valor tabular. Por ejemplo, si el valor de la tabla es 30%, el área más allá de z es 50% - 30% = 20%. El área más allá de z = + 1 sería 0.5000 - 0.3413 = 0.1587, dado que el área entre la media y z = + 1 es 0.3413.

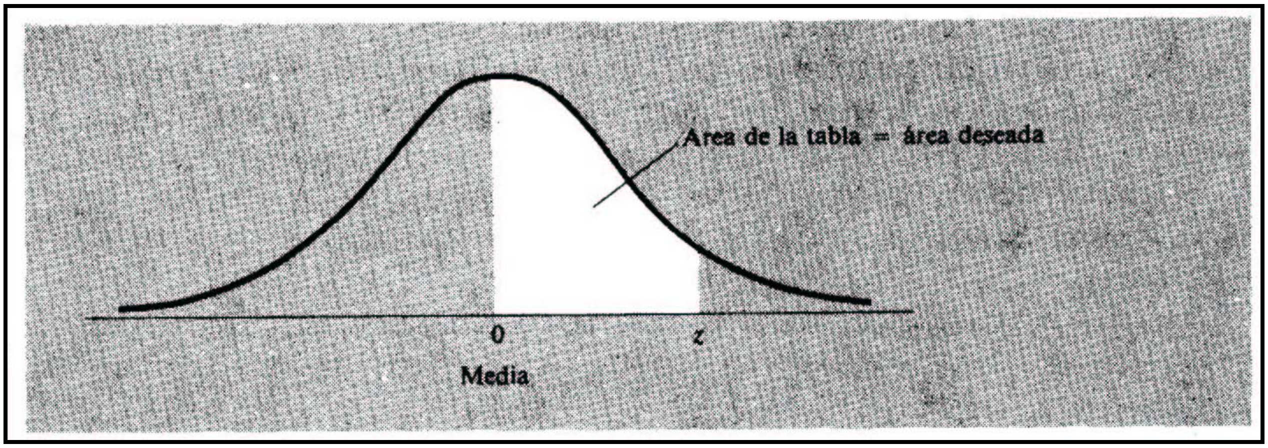

FIGURA 5.16 Una tabla normal proporciona el área bajo la curva normal estándar entre la media y algún valor de z.

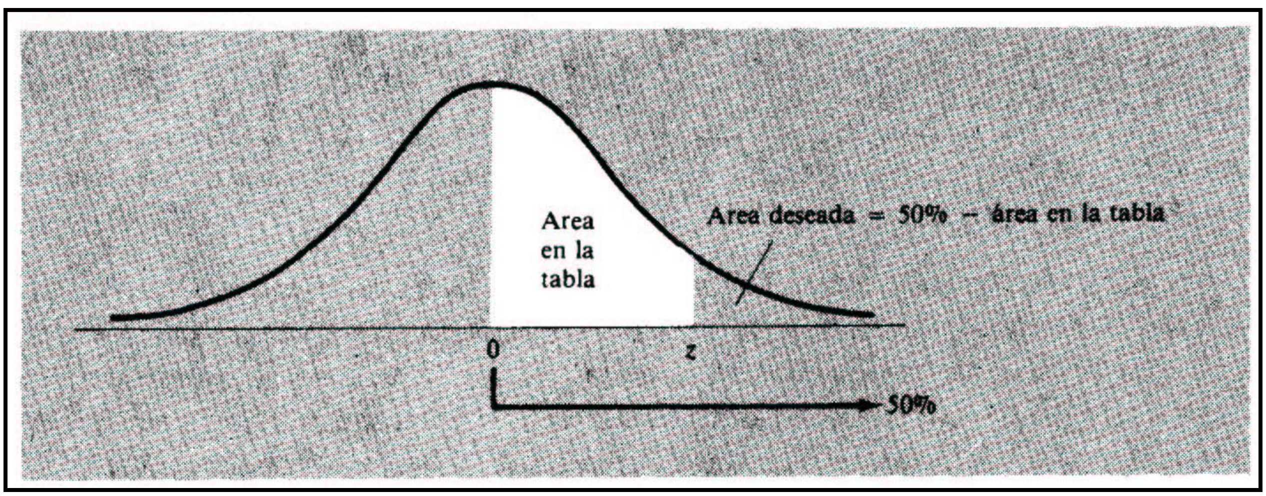

FIGURA 5.17 Es posible calcular el área más allá de z, restando el área entre la media y z de 0.5000.

Este concepto se ilustra en la figura 5.17.

A continuación se presentan algunos ejemplos.

z	$P(0 < x < z)$	$P(x > z) = 0.5000 - P(0 < x < z)$
1.65	0.4505	0.0495
1.96	0.4750	0.0250
2.33	0.4901	0.0099

No nos circunscribimos a situaciones limitadas por la media. Cuando un intervalo o su complemento no está limitado por la media de la distribución, la determinación del área bajo la curva es un proceso de dos etapas. Por ejemplo, supóngase que se quiere obtener el área bajo la curva entre 2 = -1 desviación estándar y z = + 1 desviación estándar. Como la media siempre se utiliza como punto de referencia, se debe calcular el área entre la media y cada frontera o límite. Se acaba de comentar que el área entre la media y z = + 1 es 0.3413. Además, el área entre la media y z = -1 es 0.3413. Mediante la combinación de los dos valores se obtendrá el área total: 0.3413 + 0.3413 = 0.6826. Esto se ilustra en la figura 5.18.

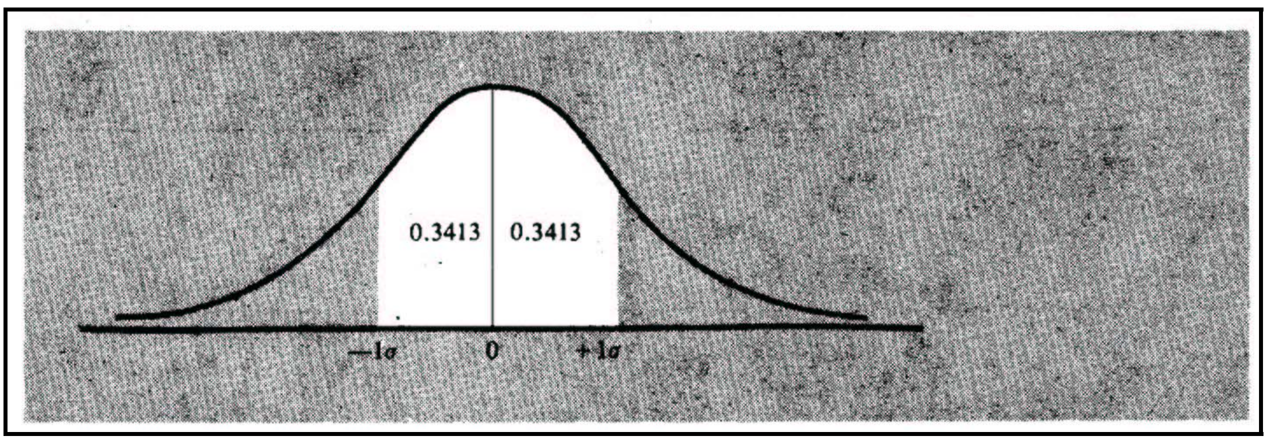

FIGURA 5.18 La determinación del área bajo la curva entre dos valores z es un problema de dos partes.

En forma similar, si los dos límites de un intervalo están en el mismo lado de la media, y se desea calcular el área bajo la curva entre esas dos, una vez más se debe determinar al área entre cada una de ellas y la media. Pero en este caso se quiere conocer la diferencia entre las dos áreas. Por ejemplo, si deseamos encontrar el área entre z = 1 y z = 2 (ver figura 5.19), se debe calcular el área entre z = 1 y la media (0.3413), y restarla después del valor para el área entre z = 2 y la media (0.4772). Así pues, 0.4772 - 0.3413 = 0.1359 es el área entre z = 1 y z = 2.

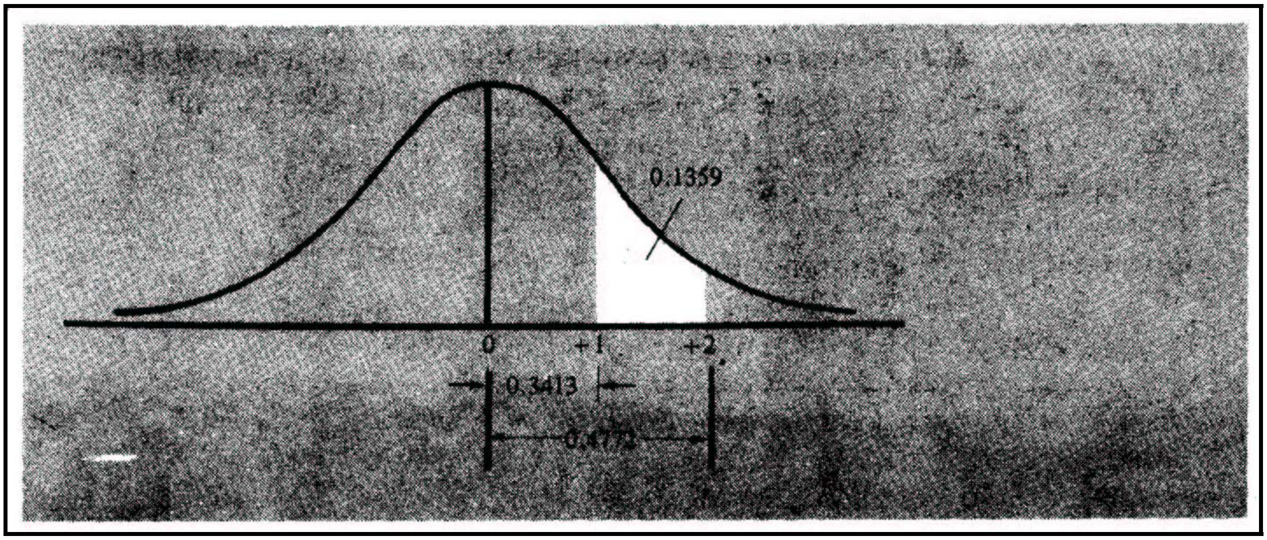

FIGURA 5.19 Cuando dos valores de z tienen el mismo signo y se quiere conocer el área que hay entre ellos, se obtiene la respuesta al determinar primero el área entre cada uno y la media, y al encontrar después la diferencia entre los dos valores.

La distribución normal como una aproximación a la distribución binomial

Muchas situaciones de la vida real son consideradas adecuadamente por la distribución binomial. El problema es que las tablas binomiales rara vez se extienden más allá de n = 20, debido simplemente a que existen tantos resultados que el gran tamaño de las tablas resultantes dificultaría su impresión. Hay tablas más extensas, pero por lo general no es fácil disponer de ellas. Se pueden utilizar fórmulas cuando se cuente con la ayuda de una calculadora; de otro modo, los cálculos serían demasiado engorrosos.

En algunos ejemplos se puede utilizar la distribución normal para obtener buenas aproximaciones a probabilidades binomiales, y ya se ha visto que no es particularmente difícil trabajar con la distribución normal. Anteriormente se comentó en el Capítulo 4 que, en ciertas circunstancias, se podía utilizar la distribución de Poisson para aproximar probabilidades binomiales cuando n es grande y la probabilidad de éxito esté muy cercana a 0 o a 100%. La aproximación normal funciona mejor cuando la probabilidad de éxito está cercana a 0.50, y la 1 aproximación aumenta [y la necesidad de hacer que P(éxito) se aproxime a 0.50 disminuye], a medida que se incrementa n*. En la figura 5.20 se ilustra este concepto.

El uso de la distribución normal para aproximar probabilidades binomiales presenta una dificultad conceptual de que no era una consideración que utilizase aproximaciones de Poisson. La distribución normal es continua, mientras que las distribuciones de Poisson y binomial son discretas. La transición de discreta a continua implica enfrentarse a problemas con valores no enteros que estén asociados a variables continuas, pero no a variables discretas. Por ejemplo, el valor 3.4523 puede

---

* *Una regla generalmente aceptada es que n•p o (1 - p), independientemente de que sea menor o mayor que 5, o igual a dicha cantidad, a fin de utilizar la aproximación normal.*

ser congruente con una variable continua, pero quizá no con una variable discreta, dado que este tipo de variables generalmente comprenden sólo números enteros. Las distribuciones probabilísticas discretas tienen valores de probabilidad en los enteros pero no entre ellos. Las distribuciones continuas, no obstante, son alisadas más que “con protuberancias”, ya que todos los valores (intervalos) tienen probabilidades asociadas a ellos.

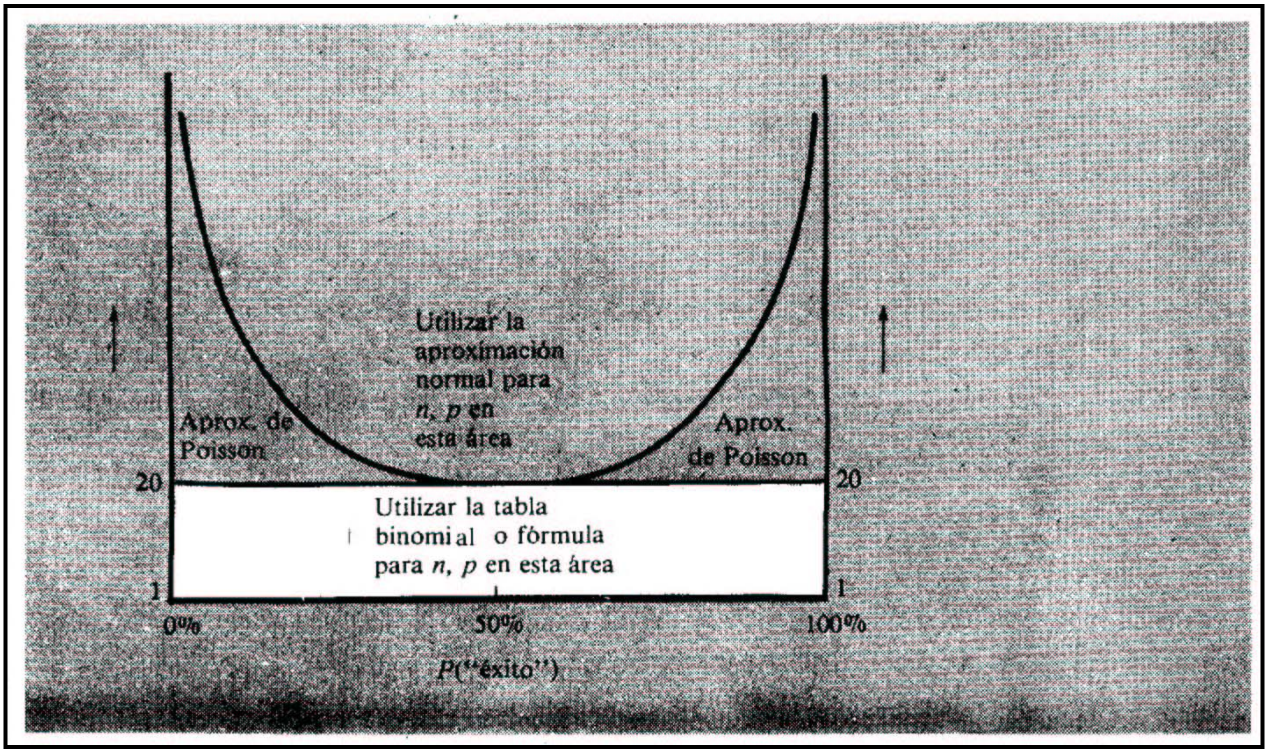

FIGURA 5.20 La decisión de cuando utilizar una distribución binomial en comparación con una aproximación normal o una de Poisson es una función,•tanto de la probabilidad de “éxito' como del número de ensayos u observaciones.

El problema se resuelve asignando intervalos de la distribución continua para representar valores enteros que sean comunes a variables discretas. En esencia, los valores no enteros de una variable continua se redondean al entero más próximo, y las probabilidades relacionadas con los valores no enteros son consideradas probabilidades enteras. Por ejemplo, los valores continuos del intervalo que va de 2.5 a 3.5 se relacionarían con el valor discreto o entero 3; los valores continuos del intervalo que comprende de 6.5 a 7.5 se podrían relacionar con el valor discreto 7; y así sucesivamente. De este modo, para calcular la probabilidad binomial de exactamente 7 éxitos, se utilizará una aproximación normal, basada en la probabilidad (área bajo la curva normal) entre 6.5 y 7.5.

Considérense estos ejemplos.

Ejemplo 1 Supóngase que n = 20 y p = 0.40. Se puede fácilmente utilizar una tabla binomial para obtener varios valores. Por ejemplo, P(x = 3) es 0.0124 empleando la tabla B del apéndice. Intentémoslo ahora utilizando la aproximación normal.

Solución:

Cabe recordar que la distribución normal está expresada por su media y desviación estándar, por lo que, en principio, se debe determinar la media y desviación estándar de esta distribución binomial. La media es np o sea, en este caso, 20 (0.40) = 8 y la desviación estándar de la distribución es

$$\sigma_{np} = \sqrt{np(1-p)} = \sqrt{20(0.4)(0.6)} = 2.2$$

"Exactamente 3" se debe interpretar como el intervalo que va de 2.5 a 3.5 en la distribución normal, como se indica en la figura 5.21.

Como se mencionó antes, se tiene un problema de dos partes, debido a la forma como se elabora la tabla normal. Se debe determinar la probabilidad de un valor entre 2.5 y la media, así como la de un valor entre 3.5 y la media. La diferencia entre esas dos probabilidades es la probabilidad de un valor entre 2.5 y 3.5.

$$z_1 = \frac{2.5-8}{2.2} = \frac{-5.5}{2.2} = -2.50 \qquad P_1 = 0.4938$$

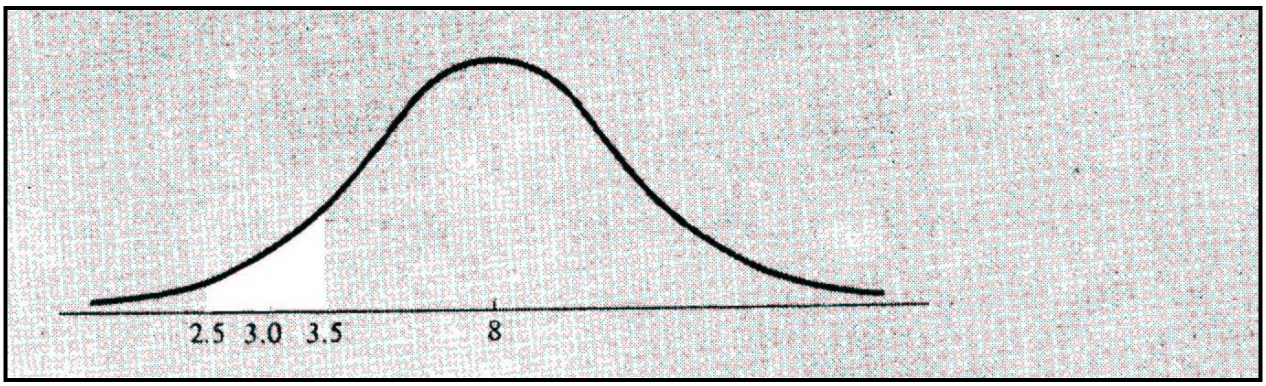

FIGURA 5.21 La aproximación normal a la probabilidad de exactamente 3.

$$z_2 = \frac{3.5-8}{2.2} = \frac{-4.5}{2.2} = -2.05 \qquad P_2 = 0.4798$$

$$P(x=3) = P_1 - P_2 = 0.0140$$

Obsérvese que la aproximación de 0.0140 está muy cercana al valor verdadero de 0.0124; el error es 0.0016, el cual es aceptablemente pequeño.

Asimismo, se puede utilizar la aproximación normal para determinar la probabilidad de un intervalo de resultados.

Ejemplo 2 Mediante el uso de los mismos valores de n y p del ejemplo 1, se tiene n = 20 y p = 0.40.

a. Encuentre la probabilidad de x $\geq$ 10.

b. Obtenga la probabilidad de x = 9, 10 u 11.

(Nota: la media es 8 y la desviación estándar es 2.2, como se determinó en el ejemplo l.)

Solución:

a. x $\geq$ 10 en realidad significa (para una aproximación continua) que x > 9.5. La probabilidad de un valor en este intervalo se encuentra, calculando la probabilidad de observar un valor entre la media y 9.5, y restando después esa probabilidad del 50%, según se ilustra en la figura 5.22(a).

$$P(\mu \leq x \leq 9.5):\ z = \frac{9.5 - 8}{2.2} = \frac{1.5}{2.2} = 0.68$$

A partir de la tabla normal se observa que z = 0.68 que se convierte en una probabilidad de 0.2518. Por tanto

$$P(x > 9.5) = 0.5000 - 0.2518 = 0.2482$$

b. x = 9, 10 u 11 se convierte en el intervalo continuo que va de 8.5 a 11.5.

$$z_1 = \frac{11.5 - 8}{2.2} = \frac{3.5}{2.2} = 1.59 \qquad P_1 = 0.4441$$

$$z_2 = \frac{8.5 - 8}{2.2} = \frac{0.5}{2.2} = 0.23 \qquad P_2 = 0.0910$$

$$P(8.5 \leq x \leq 11.5) = P_1 - P_2 = 0.3531$$

Esto se ilustra en la figura 5.22(b).

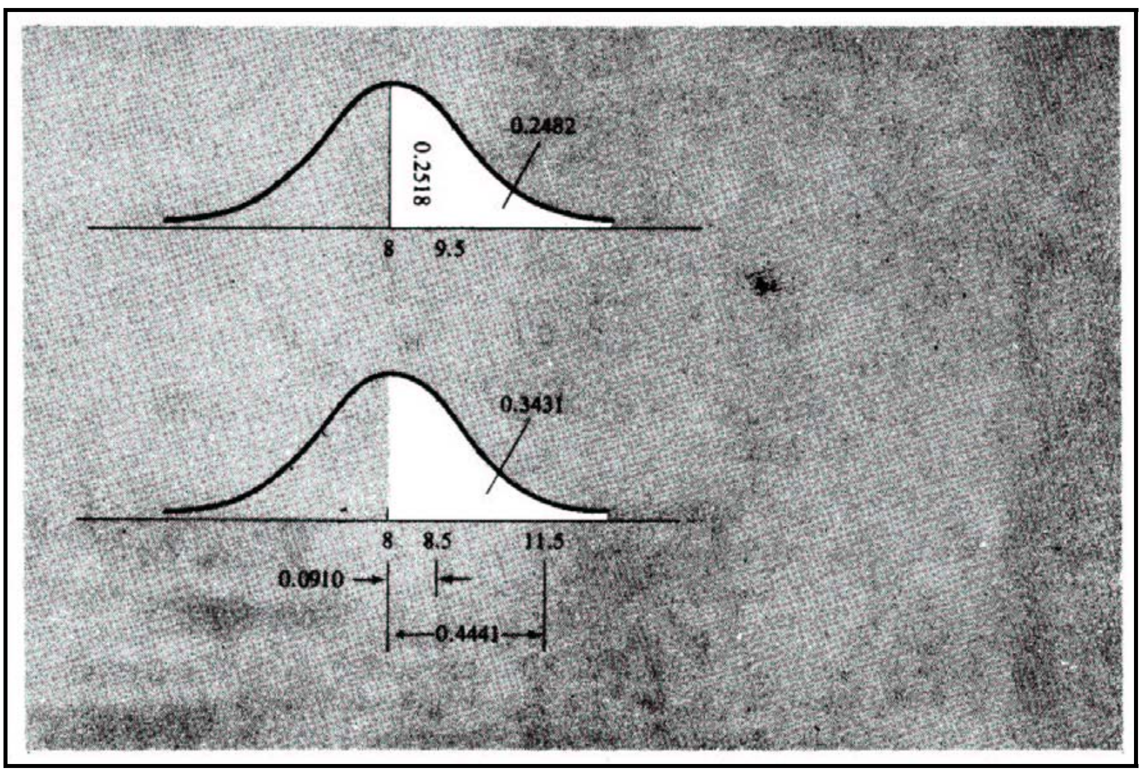

FIGURA 5.22 P(x $\geq$ 10) = 0.2472. (b) P(x = 9, 10, o 11) = 0.3431.

EJERCICIOS

1. Trace una curva normal, sombree el área deseada y obtenga la información requerida a continuación:
   a. Encuentre el área a la derecha de z = 1.0
   b. Obtenga el área a la izquierda de z = 1.0.
   c. Calcule el área a la derecha de z = -0.34.
   d. Determine el área entre z = 0 y z = 1.5.
   e. Halle el área entre z = 0 y - 2.88.
   f. Encuentre el área entre z = -0.56 y z = -0.20
   g. Obtenga el área entre z = -0.49 y + 10.49.
   h. Calcule el área entre z = + 2.5 y z = + 2.8.
2. Trace una curva normal, sombree el área deseada y encuentre el área
   a. a la izquierda de z = - 0.2
   b. a la derecha de z = - 0.2
   c. entre z = - 0.2 y z = 0
   d. entre z = - 0.2 y z = + 0.4
3. Encuentre los valores de z que producirán las siguientes áreas:
   a. El área a la izquierda de z es 0.0505.
   b. El área a la izquierda de z es 0.0228.
   c. El área a la derecha de z es 0.0228.
   d. El área entre 0 y z es 0.4772.
   e. El área entre + z y - z es .0.0240.
   f. El área menor que -z o mayor que +z es 0.9760.
4. Calcule los valores de z que corresponden a estas probabilidades:
   a. El área a la derecha de z es 0.505.
   b. El área a la derecha de z es 0.5000.
   c. El área a la izquierda de z es 0.0107.
   d. El área a la izquierda de z es 0.3520.
   e. El área a la izquierda de z es 0.8051.
   f. El área entre +z y -z es 0.9544.
   g. El área entre + z y - z es 0.6826.

5. Dado que una población con una media de 25 y una desviación estándar de 2.0 está distribuida normalmente, obtenga los valores de z para los siguientes valores de población:

   a. 23.0 b. 23.5 c. 24.0 d. 25.2 e. 25.5

6. Una población normalmente distribuida tiene una media de 40 y una desviación estándar de 3. Calcule los valores reales para los siguientes valores de z:

   a. +0.10 b. + 2.00 c. +0.75

   d. - 2.53 e. - 3.00 f. - 3.20

7. Una distribución normal tiene una media de 50 y una desviación estándar de 5. ¿Qué porcentaje de la población de valores se encuentra en estos intervalos:

   a. de 40 a 50?
   b. de 49 a 50?
   c. de 40 a 45?
   d. de 56 a 60?
   e. de 40 a 65?
   f. de 45 a 55?

8. Después de un curado de 28 días, el cemento Portland común tiene una resistencia promedio a la compresión de 4000 libras por pulgada cuadrada. Suponga que esta resistencia a la compresión está distribuida normalmente, con una desviación estándar de 120 libras por pulgada cuadrada. Obtenga estas probabilidades respecto a la resistencia a la compresión de 28 días:

   a. menor que 3900. b. mayor que 3850.

   c. menor que 3850. d. mayor que 3880.

9. Suponga que el ingreso medio de una gran comunidad se puede aproximar razonablemente mediante una distribución normal que tiene una media de $ 15 000 y una desviación estándar de $ 3000.

   a. ¿Qué porcentaje de la población tendrá ingresos superiores a $ 18 600?
   b. En una muestra aleatoria de 50 empleados, ¿alrededor de cuántas personas se puede esperar que tengan ingresos menores de $ 10 500?

10. Un fabricante de hierro forjado afirma que su producto tiene una resistencia a la tensión casi normal, con una media de 50 000 libras por pulgada cuadrada y una variancia de 810 000 libras por pulgada cuadrada. Si esta aseveración es verdadera, ¿qué porcentaje de las mediciones muestrales sería:

    a. mayor que 50 000 libras por pulgada cuadrada?
    b. menor que 49 550 libras por pulgada cuadrada?
    c. mayor que ± 1350 libras por pulgada cuadrada de 50 000?

11. Un estimador de costos para proyectos gubernamentales descubre que su capacidad para estimar costos está distribuida normalmente alrededor del costo verdadero, con una desviación estándar de $10 000. Si este es el caso, ¿qué porcentaje de las veces sus estimaciones estarían:
    a. dentro de $ 15 000 respecto del costo verdadero?
    b. dentro de $ 20 000 respecto del costo verdadero?
    c. dentro de $ 27 000 respecto del costo verdadero?
12. Mediante un proceso de producción de tubería se fabrican piezas con un diámetro medio de 2.00 y una desviación estándar de 0.01 de pulgada. La tubería con diámetros que varían por más de 0.03 respecto de la media se considera defectuosas. Suponga que esto es normal.
    a. ¿Qué porcentaje de la tubería estará defectuosa?
    b. ¿Cuál es la probabilidad de encontrar dos piezas defectuosas seguidas?
    c. ¿Cuál es la probabilidad de hallar dos piezas seguidas sin defectos?
13. El peso de los pescados atrapados por un barco es aproximadamente normal, con una media de 4.5 kilos y una desviación estándar de 0.5 kilos.
    a. ¿Qué porcentaje de los peces pesarán menos de 4 kilos?
    b. ¿Qué porcentaje del peso de los peces será inferior a un kilogramo del peso promedio?
    c. Si se eligen dos pescados, ¿cuál es la probabilidad de que uno pese más que la media y otro menos?
    d. ¿Cuál es la probabilidad de que dos pescados pesen menos que la media?
14. Una moneda normal se lanza al aire 64 veces. Encuentre las probabilidades para estos resultados:
    a. El número de caras es igual al número de cruces.
    b. El número de caras es mayor que 34. c. El número de caras es menor que 32.
    c. El número de caras está entre 30 y 36, pero no incluye a 30 ó 36.
    d. El número de caras está entre, o incluye, a 30 y 36.
15. Se dice que aproximadamente el 30% de los adultos que viven en Nueva York son accionistas de compañías privadas. Suponiendo que esto sea cierto, determine esas probabilidades para una muestra aleatoria de neoyorkinos.
    a. Menos de 20 en una muestra de 40 posee acciones.
    b. 12 ó menos en una muestra de 50 posee acciones.
    c. 12 ó menos en una muestra de 70 posee acciones.

16. Se sabe que la cantidad de cerveza en una lata de 12 onzas fabricada por cierta embotelladora, está bien aproximada mediante una distribución normal, con una media de 12 onzas y una desviación estándar de 0.25 onzas.
    a. ¿Qué porcentaje de latas podrían tener menos de 11.60 onzas?
    b. ¿Qué porcentaje de latas no variarían en más de 0.30 onzas respecto de la media?
    c. Determine la probabilidad de obtener cuatro latas que contengan menos de 12 onzas si se elige una muestra aleatoria de 4 latas.

DISTRIBUCIÓN EXPONENCIAL

Esta comprende probabilidades acerca de la longitud de tiempo o distancia entre ocurrencias con respecto a un intervalo continuo. Por ejemplo, la distribución exponencial se utiliza para representar o modelar el tiempo entre fallas de equipo eléctrico, el tiempo entre llegadas de clientes a un supermercado, el tiempo entre llamadas para servicio, etc. Existe una estrecha relación entre las distribuciones exponencial y la de Poisson. De hecho, si un proceso de Poisson tiene una media de 7 ocurrencias respecto a un intervalo, el espacio (o tiempo, etc.) entre las ocurrencias en lo referente a ese intervalo será $^1/_7$. Por ejemplo, si las llamadas para servicio son en promedio seis llamadas/ hora, entonces el tiempo promedio entre éstas será $^1/_6$ de hora, o bien, 10 minutos.

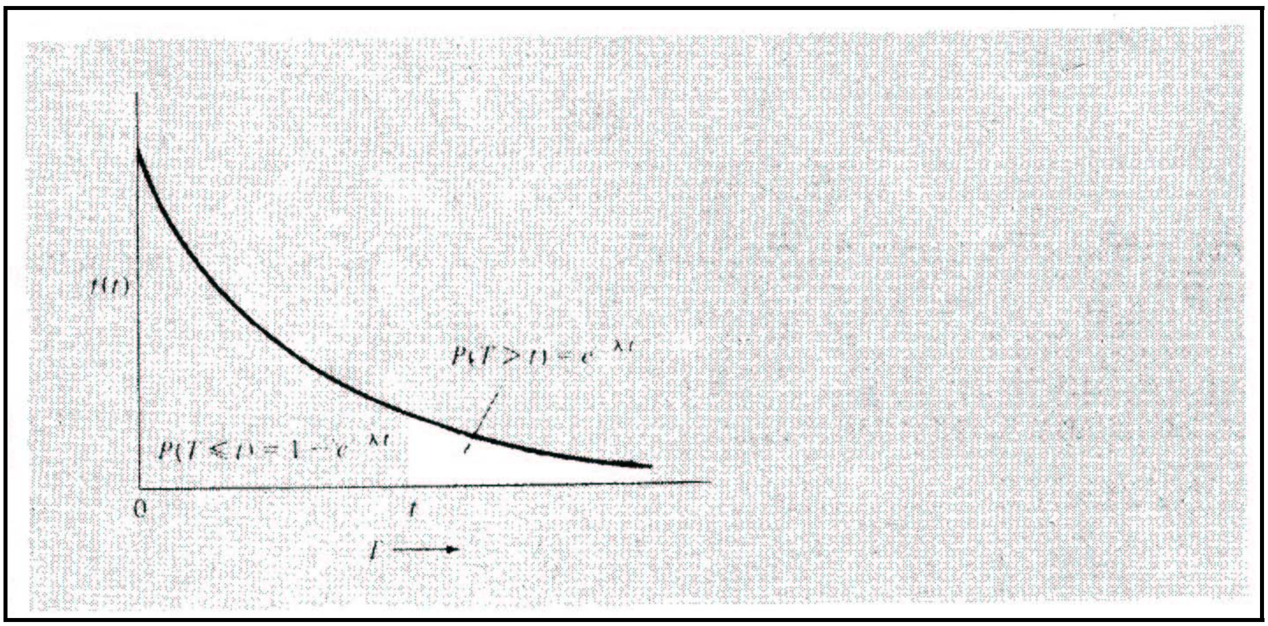

FIGURA 5.23 Distribución exponencial. Las probabilidades se expresan en términos de la probabilidad de una ocurrencia antes o después de algún punto específico t.

Las probabilidades exponenciales se expresan en términos del tiempo o distancia, hasta que un evento u ocurrencia no tiene lugar. Ver figura 5.23. Por ejemplo, se desea determinar la probabilidad de que no haya llamadas durante un periodo de dos horas (t = 2) si la razón promedio (**λ**) es 1.5 llamadas/hora. Es posible utilizar la fórmula que se da a continuación.

$$P(T > t) = e^{-\lambda t}$$

Mediante esta fórmula, se puede calcular la probabilidad de que el espacio (o tiempo) antes de que se presente la primera ocurrencia sea mayor que un espacio dado (o tiempo) t.

La probabilidad de una ocurrencia en t o antes de dicho espacio se obtiene mediante la fórmula:

$$P(T \leq t) = 1 - e^{-\lambda t}$$

Ejemplo 3 Suponga que el tiempo que tardan en recibir su orden después de hacerla en un gran restaurante promedia 10 minutos. Suponga también que ese tiempo que espera en ser atendido se distribuye exponencialmente.

a. Calcule la probabilidad de que su tiempo de espera sea mayor de 10 minutos.
b. Obtenga la probabilidad de que su tiempo de espera sea de 10 minutos o menos.
c. Encontrar la probabilidad de que su tiempo de espera sea de tres minutos o menos.

Solución:

Partiendo de la proposición del problema, se tiene que $\lambda = {}^{1}/_{10} = 0.1$ por minuto.

a.

$$P(T > 10) = e^{-\lambda t} = e^{-0.1(10)} = e^{-1} = 0.368$$

(Nota: en la tabla F del Apéndice se encuentra e^{-1}).

b.

$$P(T \leq 10) = 1 - e^{-\lambda t} = 1 - e^{-1} = 1 - 0.368 = 0.632$$

c.

$$P(T \leq 3) = 1 - e^{-\lambda t} = 1 - e^{-0.1(3)} = 1 - e^{-0.3} = 1 - 0.741 = 0.259$$