41. a. Draw a line graph of the pmf of X in Exercise 34. Then determine the pmf of $-X$ and draw its line graph. From these two pictures, what can you say about $V(X)$ and $V(-X)$?
    b. Use the proposition involving $V(aX + b)$ to establish a general relationship between $V(X)$ and $V(-X)$.

42. Use the definition of variance to prove that $V(aX+b) = a^2\sigma_X^2$. [*Hint:* With $Y = aX + b$, $E(Y) = a\mu + b$ where $\mu = E(X)$.]

43. Suppose $E(X) = 5$ and $E[X(X - 1)] = 27.5$. What is
    a. $E(X^2)$? [*Hint*: $E[X(X - 1)] = E[X^2 - X] = E(X^2) - E(X)$.]
    b. $V(X)$?
    c. The general relationship among the quantities $E(X)$, $E[X(X - 1)]$, and $V(X)$?

44. Write a general rule for $E(X - c)$ where c is a constant. What happens when you let $c = \mu$, the expected value of X?

45. A result called **Chebyshev's inequality** states that for any probability distribution of a rv X and any number k that is at least 1, $P(|X - \mu| \geq k\sigma) \leq 1/k^2$. In words, the probability that the value of X lies at least k standard deviations from its mean is at most $1/k^2$.
    a. What is the value of the upper bound for $k = 2$? $k = 3$? $k = 4$? $k = 5$? $k = 10$?
    b. Compute μ and σ for the distribution given in Exercise 13. Then evaluate $P(|X - \mu| \geq k\sigma)$ for the values of k given in part (a). What does this suggest about the upper bound relative to the corresponding probability?
    c. Let X have possible values, -1, 0, and 1, with probabilities 1/18, 8/9, and 1/18, respectively. What is $P(|X - \mu| \geq 3\sigma)$, and how does its value compare to the corresponding bound?
    d. Give a distribution for which $P(|X - \mu| \geq 5\sigma) = .04$.

3.4 Moments and Moment Generating Functions

The expected values of integer powers of X and $X - \mu$ are often referred to as *moments*, terminology borrowed from physics. In this section, we'll discuss the general topic of moments and develop a shortcut for computing them.

DEFINITION The ***k*th moment** of a random variable X is $E(X^k)$, while the ***k*th moment about the mean** (or ***k*th central moment**) of X is $E[(X - \mu)^k]$, where $\mu = E(X)$.

For example, $\mu = E(X)$ is the "first moment" of X and corresponds to the center of mass of the distribution of X. Similarly, $V(X) = E[(X - \mu)^2]$ is the second moment of X about the mean, which is known in physics as the *moment of inertia.*

Example 3.26 A popular brand of dog food is sold in 5, 10, 15, and 20 lb bags. Let X be the weight of the next bag purchased, and suppose the pmf of X is

x	5	10	15	20
$p(x)$	.1	.2	.3	.4

The first moment of X is its mean:

$$\mu = E(X) = \sum_{x \in D} xp(x) = 5(.1) + 10(.2) + 15(.3) + 20(.4) = 15 \text{ lbs}$$

The second moment about the mean is the variance:

$$\begin{aligned}\sigma^2 = E[(X-\mu)^2] &= \sum_{x \in D} (x-\mu)^2 p(x) \\ &= (5-15)^2(.1) + (10-15)^2(.2) + (15-15)^2(.3) + (20-15)^2(.4) = 25,\end{aligned}$$

for a standard deviation of 5 lbs. The third central moment of X is

$$\begin{aligned}E[(X-\mu)^3] &= \sum_{x \in D} (x-\mu)^3 p(x) \\ &= (5-15)^3(.1) + (10-15)^3(.2) + (15-15)^3(.3) + (20-15)^3(.4) = -75\end{aligned}$$

We'll discuss an interpretation of this last number next. ■

It is not difficult to verify that the third moment about the mean is 0 if the pmf of X is symmetric. So, we would like to use $E[(X-\mu)^3]$ as a measure of lack of symmetry, but it depends on the scale of measurement. If we switch the unit of weight in Example 3.26 from pounds to ounces or kilograms, the value of the third moment about the mean (as well as the values of all the other moments) will change. But we can achieve scale independence by dividing the third moment about the mean by σ^3:

$$\frac{E[(X-\mu)^3]}{\sigma^3} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \tag{3.14}$$

Expression (3.14) is our measure of departure from symmetry, called the **skewness coefficient**. The skewness coefficient for a symmetric distribution is 0 because its third moment about the mean is 0. However, in the foregoing example the skewness coefficient is $E[(X-\mu)^3]/\sigma^3 = -75/5^3 = -0.6$. When the skewness coefficient is negative, as it is here, we say that the distribution is *negatively skewed* or that it is *skewed to the left*. Generally speaking, it means that the distribution stretches farther to the left of the mean than to the right.

If the skewness coefficient were positive, then we would say that the distribution is *positively skewed* or that it is *skewed to the right*. For example, reverse the order of the probabilities in the pmf of Example 3.26, so the probabilities of the values 5, 10, 15, 20 are now .4, .3, .2, and .1 (customers now favor much smaller bags of dog food). Exercise 61 shows that this changes the sign but not the magnitude of the skewness coefficient, so it becomes +0.6 and the distribution is skewed right. Both distributions are illustrated in Figure 3.8.

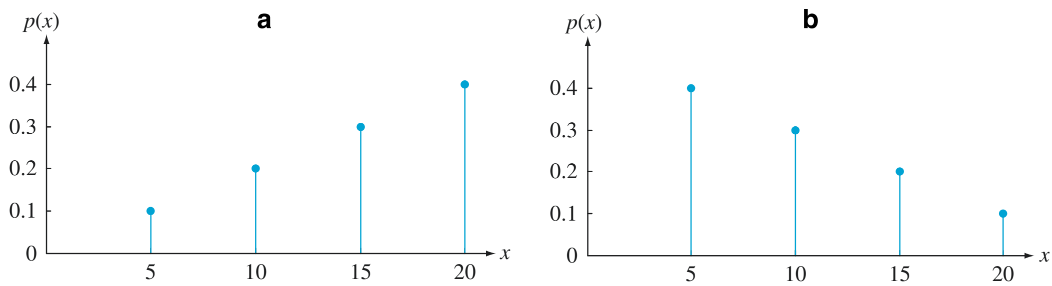

Figure 3.8 Departures from symmetry: (**a**) skewness coefficient < 0 (skewed left); (**b**) skewness coefficient > 0 (skewed right)

The Moment Generating Function

Calculation of the mean, variance, skewness coefficient, etc., for a particular discrete rv requires extensive, sometimes tedious, summation. Mathematicians have developed a tool, the *moment generating function*, that will allow us to determine the moments of a distribution with less effort. Moreover, this function will allow us to derive properties of several important probability distributions in subsequent sections of the book.

Note first that $e^{1.7X}$ is a particular function of X; its expected value is $E(e^{1.7X}) = \sum e^{1.7x} \cdot p(x)$. The number 1.7 in the foregoing expression can be replaced by any other number—2.5, 179, –3.25, etc. Now consider replacing 1.7 by the letter t. Then the expected value depends on the numerical value of t; that is, $E(e^{tX})$ is a function of t. It is this function that will generate moments for us.

DEFINITION The **moment generating function (mgf)** of a discrete random variable X is defined to be

$$M_X(t) = E(e^{tX}) = \sum_{x \in D} e^{tx} p(x)$$

where D is the set of possible X values. The moment generating function exists iff $M_X(t)$ is defined for an interval that includes zero as well as positive and negative values of t.

For any random variable X, the mgf evaluated at $t = 0$ is

$$M_X(0) = E(e^{0X}) = \sum_{x \in D} e^{0x} p(x) = \sum_{x \in D} 1 p(x) = 1$$

That is, $M_X(0)$ is the sum of all the probabilities, so it must always be 1. However, in order for the mgf to be useful in generating moments, it will need to be defined for an interval of values of t including 0 in its interior. The moment generating function fails to exist in cases when moments themselves fail to exist (see Example 3.30 below).

Example 3.27 The simplest example of an mgf is for a Bernoulli distribution, where only the X values 0 and 1 receive positive probability. Let X be a Bernoulli random variable with $p(0) = 1/3$ and $p(1) = 2/3$. Then

$$M_X(t) = E(e^{tX}) = \sum_{x \in D} e^{tx} p(x) = e^{t \cdot 0} \frac{1}{3} + e^{t \cdot 1} \frac{2}{3} = \frac{1}{3} + e^t \frac{2}{3}$$

A Bernoulli random variable will always have an mgf of the form $p(0) + p(1)e^t$, a well-defined function for all values of t. ■

A key property of the mgf is its "uniqueness," the fact that it completely characterizes the underlying distribution.

MGF UNIQUENESS THEOREM

If the mgf exists and is the same for two distributions, then the two distributions are identical. That is, the moment generating function uniquely specifies the probability distribution; there is a one-to-one correspondence between distributions and mgfs.

The proof of this theorem, originally due to Laplace, requires some sophisticated mathematics and is beyond the scope of this textbook.

Example 3.28 Let X, the number of claims submitted on a renter's insurance policy in a given year, have mgf $M_X(t) = .7 + .2e^t + .1e^{2t}$. It follows that X must have the pmf $p(0) = .7$, $p(1) = .2$, and $p(2) = .1$—because if we use this pmf to obtain the mgf, we get $M_X(t)$, and the distribution is uniquely determined by its mgf. ■

Example 3.29 Consider testing individuals' blood samples one by one in order to find someone whose blood type is Rh+. The rv X = the number of tested samples should follow the pmf specified in Example 3.10 with $p = .85$:

$$p(x) = .85(.15)^{x-1} \quad \text{for } x = 1, 2, 3, \ldots.$$

Determining the moment generating function here requires using the formula for the sum of a geometric series: $1 + r + r^2 + \cdots = 1/(1 - r)$ for $|r| < 1$. The moment generating function is

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \sum_{x \in D} e^{tx} p(x) = \sum_{x=1}^{\infty} e^{tx} .85(.15)^{x-1} = .85e^t \sum_{x=1}^{\infty} e^{t(x-1)} (.15)^{x-1} \\ &= .85e^t \sum_{x=1}^{\infty} (.15e^t)^{x-1} = .85e^t [1 + .15e^t + (.15e^t)^2 + \cdots] = \frac{.85e^t}{1 - .15e^t} \end{aligned}$$

The condition on r requires $|.15e^t| < 1$. Dividing by .15 and taking logs gives $t < -\ln(.15) \approx 1.90$; i.e., this function is defined in the interval $(-\infty, 1.90)$. The result is an interval of values that includes 0 in its interior, so the mgf exists. As a check, $M_X(0) = .85/(1 - .15) = 1$, as required. ■

Example 3.30 Reconsider Example 3.20, where $p(x) = k/x^2$, $x = 1, 2, 3, \ldots$. Recall that $E(X)$ does not exist for this distribution, portending a problem for the existence of the mgf:

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} \frac{k}{x^2}$$