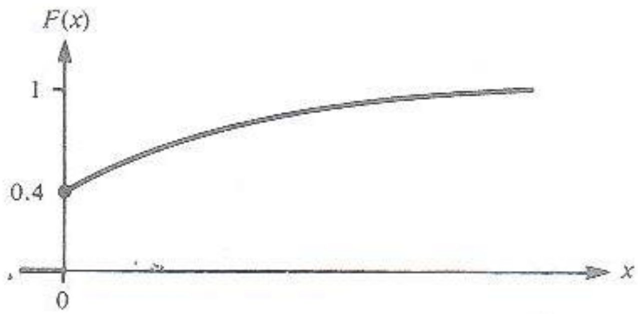
FIGURE 2.9 The CDF of a mixed distribution



minutes is

$$P[X \leqslant 0.5] = 0.4 + 0.6(1 - e^{-0.5}) = 0.636$$

The distribution of X given $0 < X$ corresponds to

$$\begin{aligned} P[X \leqslant x \mid 0 < X] &= \frac{P[0 < X \text{ and } X \leqslant x]}{P[0 < X]} \\ &= \frac{P[0 < X \leqslant x]}{P[0 < X]} \\ &= \frac{F(x) - F(0)}{1 - F(0)} \\ &= \frac{0.4 + 0.6(1 - e^{-x}) - 0.4}{1 - 0.4} \\ &= 1 - e^{-x} \end{aligned}$$

2.4

SOME PROPERTIES OF EXPECTED VALUES

It is useful to consider, more generally, the expected value of a function of X. For example, if the radius of a disc is a random variable X, then the area of the disc, say $Y = \pi X^2$, is a function of X. In general, let X be a random variable with pdf $f(x)$, and denote by $u(x)$ a real-valued function whose domain includes the possible values of X. If we let $Y = u(X)$, then Y is a random variable with its own pdf, say $g(y)$. Suppose, for example, that X is a discrete random variable with pdf $f(x)$. Then $Y = u(X)$ is also a discrete random variable with pdf $g(y)$ and expected value defined in accordance with Definition 2.2.3, namely $E(Y) = \sum_y y g(y)$. Of course, evaluation of $E(Y)$ directly from the definition requires knowing the pdf $g(y)$. The following theorem provides another way to evaluate this expected value. The proof, which requires advanced methods, will be discussed in Chapter 6.

Theorem 2.4.1 If X is a random variable with pdf $f(x)$ and $u(x)$ is a real-valued function whose domain includes the possible values of X, then

$$E[u(X)] = \sum_x u(x) f(x) \qquad \text{if } X \text{ is discrete} \tag{2.4.1}$$

$$E[u(X)] = \int_{-\infty}^{\infty} u(x) f(x)\, dx \qquad \text{if } X \text{ is continuous} \tag{2.4.2}$$

■

It is clear that the expected value will have the "linearity" properties associated with integrals and sums.

Theorem 2.4.2 If X is a random variable with pdf $f(x)$, a and b are constants, and $g(x)$ and $h(x)$ are real-valued functions whose domains include the possible values of X, then

$$E[ag(X) + bh(X)] = aE[g(X)] + bE[h(X)] \tag{2.4.3}$$

Proof

Let X be continuous. It follows that

$$\begin{aligned} E[ag(X) + bh(X)] &= \int_{-\infty}^{\infty} [ag(x) + bh(x)] f(x)\, dx \\ &= a \int_{-\infty}^{\infty} g(x) f(x)\, dx + b \int_{-\infty}^{\infty} h(x) f(x)\, dx \\ &= aE[g(X)] + bE[h(X)] \end{aligned}$$

■

The discrete case is similar.

An obvious result of this theorem is that

$$E(aX + b) = aE(X) + b \tag{2.4.4}$$

An important special expected value is obtained if we consider the function $u(x) = (x - \mu)^2$

Definition 2.4.1

The **variance** of a random variable X is given by

$$\mathrm{Var}(X) = E[(X - \mu)^2] \tag{2.4.5}$$

Other common notations for the variance are σ^2, σ_X^2, or $V(X)$, and a related quantity, called the **standard deviation** of X, is the positive square root of the variance, $\sigma = \sigma_X = \sqrt{\mathrm{Var}(X)}$.

The variance provides a measure of the variability or amount of "spread" in the distribution of a random variable.

Example 2.4.1 In the experiment of Example 2.2.2, $E(X^2) = 2^2(1/2) + 4^2(1/4) + 8^2(1/4) = 22$, and thus $\mathrm{Var}(X) = 22 - 4^2 = 6$ and $\sigma_X = \sqrt{6} = 2.45$. For comparison, consider a slightly different experiment where two chips are labeled with zeros, one with a 4, and one with a 12. If one chip is selected at random, and Y is its number, then $E(Y) = 4$ as in the original example. However, $\mathrm{Var}(Y) = 24$ and $\sigma = 2\sqrt{6} > \sigma_X$, which reflects the fact that the probability distribution of Y h- more spread than that of X.

Certain special expected values, called moments, are useful in characterizing some features of the distribution.

Definition 2.4.2

The **kth moment about the origin** of a random variable X is

$$\mu'_k = E(X^k) \tag{2.4.6}$$

and the **kth moment about the mean** is

$$\mu_k = E[X - E(X)]^k = E(X - \mu)^k \tag{2.4.7}$$

Thus $E(X^k)$ may be considered as the kth moment of X or as the first moment of X^k. The first moment is the mean, and the simpler notation μ, rather than μ'_1, generally is preferred. The first moment about the mean is zero,

$$\mu_1 = E[X - E(X)] = E(X) - E(X) = 0$$

The second moment about the mean is the variance,

$$\mu_2 = E[(X - \mu)^2] = \sigma^2$$

and the second moment about the origin is involved in the following theorem about the variance.

Theorem 2.4.3 If X is a random variable, then

$$\text{Var}(X) = E(X^2) - \mu^2 \tag{2.4.8}$$

Proof

$$\begin{aligned}\text{Var}(X) &= E(X^2 - 2\mu X + \mu^2)\\ &= E(X^2) - 2\mu E(X) + \mu^2\\ &= E(X^2) - 2\mu^2 + \mu^2\end{aligned}$$

which yields the theorem.

It also follows immediately that

$$E(X^2) = \sigma^2 + \mu^2 \tag{2.4.9}$$

As noted previously, the variance provides a measure of the amount of spread in a distribution or the variability among members of a population. A rather extreme example of this occurs when X assumes only one value, say $P[X = c] = 1$. In this case $E(X) = c$ and $\text{Var}(X) = 0$.

The remark following Theorem 2.4.2 dealt with the expected value of a linear function of a random variable. The following theorem deals with the variance.

Theorem 2.4.4 If X is a random variable and a and b are constants, then

$$\text{Var}(aX + b) = a^2\ \text{Var}(X) \tag{2.4.10}$$

Proof

$$\begin{aligned}\text{Var}(aX + b) &= E[(aX + b - a\mu_X - b)^2]\\ &= E[a^2(X - \mu_X)^2]\\ &= a^2\ \text{Var}(X)\end{aligned}$$

This means that the variance is affected by a change of scale, but not by a translation.

Another natural measure of variability would be the **mean absolute deviation**, $E|X - \mu|$, but the variance is generally a more convenient quantity with which to work.

The mean and variance provide a good deal of information about a population distribution, but higher moments and other quantities also may be useful. For example, the third moment about the mean, μ_3, is a measure of asymmetry or "skewness" of a distribution.

Theorem 2.4.5 If the distribution of X is symmetric about the mean $\mu = E(X)$, then the third moment about μ is zero, $\mu_3 = 0$.

Proof

See Exercise 28. ■

We can conclude that if $\mu_3 \neq 0$, then the distribution is not symmetric, but not conversely, because distributions exist that are not symmetric but which do have $\mu_3 = 0$ (see Exercise 29).

BOUNDS ON PROBABILITY

It is possible, in some cases, to find bounds on probabilities based on moments.

Theorem 2.4.6 If X is a random variable and $u(x)$ is a nonnegative real-valued function, then for any positive constant $c > 0$,

$$P[u(X) \geqslant c] \leqslant \frac{E[u(X)]}{c} \tag{2.4.11}$$

Proof

If $A = \{x \mid u(x) \geqslant c\}$, then for a continuous random variable,

$$\begin{aligned} E[u(X)] &= \int_{-\infty}^{\infty} u(x)f(x)\,dx \\ &= \int_A u(x)f(x)\,dx + \int_{A^c} u(x)f(x)\,dx \\ &\geqslant \int_A u(x)f(x)\,dx \\ &\geqslant \int_A cf(x)\,dx \\ &= cP[X \in A] \\ &= cP[u(X) \geqslant c] \end{aligned}$$

A similar proof holds for discrete variables.

A special case, known as the **Markov inequality**, is obtained if $u(x) = |x|^r$ for $r > 0$, namely

$$P[|X| \geqslant c] \leqslant \frac{E(|X|^r)}{c^r} \tag{2.4.12}$$

Another well-known result, the **Chebychev inequality**, is given by the following theorem.

Theorem 2.4.7 If X is a random variable with mean μ and variance σ^2, then for any $k > 0$,

$$P[|X - \mu| \geqslant k\sigma] \leqslant \frac{1}{k^2} \tag{2.4.13}$$

Proof

If $u(X) = (X - \mu)^2$, $c = k^2\sigma^2$, then using equation (2.4.11),

$$P[(X - \mu)^2 \geqslant k^2\sigma^2] \leqslant \frac{E(X - \mu)^2}{k^2\sigma^2} \leqslant \frac{1}{k^2}$$

and the result follows.

An alternative form is

$$P[|X - \mu| < k\sigma] \geqslant 1 - \frac{1}{k^2} \tag{2.4.14}$$

and if we let $\varepsilon = k\sigma$, then

$$P[|X - \mu| < \varepsilon] \geqslant 1 - \frac{\sigma^2}{\varepsilon^2} \tag{2.4.15}$$

and

$$P[|X - \mu| \geqslant \varepsilon] \leqslant \frac{\sigma^2}{\varepsilon^2} \tag{2.4.16}$$

Letting $k = 2$, we see that a random variable will be within two standard deviations of its mean with probability at least 0.75. Although this may not be a tight bound in all cases, it is surprising that such a bound can be found to hold for all possible discrete and continuous distributions. A tighter bound, in general, cannot be obtained, as shown in the following example.

Example 2.4.2 Suppose that X takes on the values -1, 0, and 1 with probabilities 1/8, 6/8, and 1/8, respectively. Then $\mu = 0$ and $\sigma^2 = 1/4$. For $k = 2$,

$$\begin{aligned} P[-2(0.5) < X - 0 < 2(0.5)] &= P[-1 < X < 1] \\ &= P[X = 0] \\ &= \frac{3}{4} = 1 - \frac{1}{k^2} \end{aligned}$$

It also is possible to show that if the variance is zero, the distribution is concentrated at a single value. Such a distribution is called a **degenerate distribution**.

Theorem 2.4.8 Let $\mu = E(X)$ and $\sigma^2 = \mathrm{Var}(X)$. If $\sigma^2 = 0$, then $P[X = \mu] = 1$.

Proof

If $x \neq \mu$ for some observed value x, then $|x - \mu| \geqslant 1/i$ for some integer $i \geqslant 1$, and conversely. Thus,

$$[X \neq \mu] = \bigcup_{i=1}^{\infty} \left[|X - \mu| \geqslant \frac{1}{i} \right]$$

and using Boole's inequality, equation (1.4.5), we have

$$P[X \neq \mu] \leqslant \sum_{i=1}^{\infty} P\left[|X - \mu| \geqslant \frac{1}{i} \right]$$

and using equation (2.4.16) we obtain

$$P[X \neq \mu] \leqslant \sum_{i=1}^{\infty} i^2\sigma^2 = 0$$

which implies that $P[X = \mu] = 1$. ∎

APPROXIMATE MEAN AND VARIANCE

If a function of a random variable, say $H(X)$, can be expanded in a Taylor series, then an expression for the approximate mean and variance of $H(X)$ can be obtained in terms of the mean and variance of X.

For example, suppose that $H(x)$ has derivatives $H'(x)$, $H''(x)$, ... in an open interval containing $\mu = E(X)$. The function $H(x)$ has a Taylor approximation about μ,

$$H(x) \doteq H(\mu) + H'(\mu)(x - \mu) + \tfrac{1}{2}H''(\mu)(x - \mu)^2 \tag{2.4.17}$$

which suggests the approximation

$$E[H(X)] \doteq H(\mu) + \tfrac{1}{2}H''(\mu)\sigma^2 \tag{2.4.18}$$

and, using the first two terms,

$$\mathrm{Var}[H(X)] \doteq [H'(\mu)]^2\sigma^2 \tag{2.4.19}$$

where $\sigma^2 = \mathrm{Var}(X)$.

The accuracy of these approximations depends primarily on the nature of the function $H(x)$ as well as on the amount of variability in the distribution of X.

Example 2.4.3 Let X be a positive-valued random variable, and let $H(x) = \ln x$, so that $H'(x) = 1/x$ and $H''(x) = -1/x^2$. It follows that

$$E[\ln X] \doteq \ln \mu + \left(\frac{1}{2}\right)\left(-\frac{1}{\mu^2}\right)\sigma^2$$

$$= \ln \mu - \frac{\sigma^2}{2\mu^2}$$

and

$$\mathrm{Var}[\ln X] \doteq \left(\frac{1}{\mu}\right)^2\sigma^2 = \frac{\sigma^2}{\mu^2}$$

2.5 MOMENT GENERATING FUNCTIONS

A special expected value that is quite useful is known as the moment generating function.

Definition 2.5.1

If X is a random variable, then the expected value

$$M_X(t) = E(e^{tX}) \tag{2.5.1}$$

is called the **moment generating function** (MGF) of X if this expected value exists for all values of t in some interval of the form $-h < t < h$ for some $h > 0$.

In some situations it is desirable to suppress the subscript and use the simpler notation $M(t)$.

Example 2.5.1 Assume that X is a discrete finite-valued random variable with possible values $x_1, \ldots, x_m$. The MGF is

$$M_X(t) = \sum_{i=1}^{m} e^{tx_i} f_X(x_i)$$

which is a differentiable function of t, with derivative

$$M'_X(t) = \sum_{i=1}^{m} x_i e^{tx_i} f_X(x_i)$$

and, in general, for any positive integer r,

$$M_X^{(r)}(t) = \sum_{i=1}^{m} x_i^r e^{tx_i} f_X(x_i)$$

Notice that if we evaluate $M_X^{(r)}(t)$ at $t = 0$ we obtain

$$M_X^{(r)}(0) = \sum_{i=1}^{m} x_i^r f_X(x_i) = E(X^r)$$

the rth moment about the origin. This also suggests the possibility of expanding in a power series about $t = 0$, $M_X(t) = c_0 + c_1 t + c_2 t^2 + \cdots$, where $c_r = E(X^r)/r!$.

These properties hold for any random variable for which an MGF exists, although a general proof is somewhat harder.

Theorem 2.5.1 If the MGF of X exists, then

$$E(X^r) = M_X^{(r)}(0) \quad \text{for all } r = 1, 2, \ldots \tag{2.5.2}$$

and

$$M_X(t) = 1 + \sum_{r=1}^{\infty} \frac{E(X^r)t^r}{r!} \tag{2.5.3}$$

Proof

We will consider the case of a continuous random variable X. The MGF for a continuous random variable is

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$$

When the MGF exists, it can be shown that the rth derivative exists, and it can be obtained by differentiating under the integral sign,

$$M_X^{(r)}(t) = \int_{-\infty}^{\infty} x^r e^{tx} f_X(x)\, dx$$

from which it follows that for all $r = 1, 2, \ldots$

$$E(X^r) = \int_{-\infty}^{\infty} x^r f_X(x)\,dx = \int_{-\infty}^{\infty} x^r e^{0x} f_X(x)\,dx = M_X^{(r)}(0)$$

When the MGF exists, it also can be shown that a power series expansion about zero is possible, and from standard results about power series, the coefficients have the form $M_X^{(r)}(0)/r!$. We combine this with the above result to obtain

$$M_X(t) = 1 + \sum_{r=1}^{\infty} \frac{M_X^{(r)}(0)t^r}{r!} = 1 + \sum_{r=1}^{\infty} \frac{E(X^r)t^r}{r!}$$

■

The discrete case is similar.

Example 2.5.2 Consider a continuous random variable X with pdf $f(x) = e^{-x}$ if $x > 0$, and zero otherwise. The MGF is

$$\begin{aligned} M_X(t) &= \int_0^{\infty} e^{tx}e^{-x}\,dx \\ &= \int_0^{\infty} e^{-(1-t)x}\,dx \\ &= \frac{1}{1-t}\,e^{-(1-t)x}\Big|_0^{\infty} \\ &= \frac{1}{1-t} \qquad t < 1 \end{aligned}$$

The rth derivative is $M_X^{(r)}(t) = r!(1-t)^{-r-1}$, and thus the rth moment is $E(X^r) = M_X^{(r)}(0) = r!$. The mean is $\mu = E(X) = 1! = 1$, and the variance is $\mathrm{Var}(X) = E(X^2) - \mu^2 = 2 - 1 = 1$.

Example 2.5.3 A discrete random variable X has pdf $f(x) = (1/2)^{x+1}$ if $x = 0, 1, 2, \ldots$, and zero otherwise. The MGF of X is

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{\infty} e^{tx}(1/2)^{x+1} \\ &= (1/2)\sum_{x=0}^{\infty} (e^t/2)^x \end{aligned}$$

We make use of the well-known identity for the geometric series,

$$1 + s + s^2 + s^3 + \cdots = \frac{1}{1-s} \qquad -1 < s < 1$$

with $s = e^t/2$. The resulting MGF is

$$M_X(t) = \frac{1}{2 - e^t} \qquad t < \ln 2$$

The first derivative is $M'_X(t) = e^t(2 - e^t)^{-2}$, and thus $E(X) = M'_X(0) = e^0(2 - e^0)^{-2} = 1$. It is possible to obtain higher derivatives, but the complexity increases with the order of the derivative.

PROPERTIES OF MOMENT GENERATING FUNCTIONS

Theorem 2.5.2 If $Y = aX + b$, then $M_Y(t) = e^{bt}M_X(at)$.

Proof

$$\begin{aligned} M_Y(t) &= E(e^{tY}) \\ &= E(e^{t(aX+b)}) \\ &= E(e^{atX}e^{bt}) \\ &= e^{bt}E(e^{atX}) \\ &= e^{bt}M_X(at) \end{aligned}$$

■

One possible application is in computing the rth moment about the mean, $E[(X - \mu)^r]$. Because $M_{X-\mu}(t) = e^{-\mu t}M_X(t)$,

$$E[(X - \mu)^r] = \frac{d^r}{dt^r}[e^{-\mu t}M_X(t)]\,|_{t=0} \tag{2.5.4}$$

It can be shown that MGFs uniquely determine a distribution.

Example 2.5.4 Suppose, for example, that X and Y are both integer-valued with the same set of possible values—say 0, 1, and 2—and that X and Y have the same MGF,

$$M(t) = \sum_{x=0}^{2} e^{tx}f_X(x) = \sum_{y=0}^{2} e^{ty}f_Y(y)$$

if we let $s = e^t$ and $c_i = f_Y(i) - f_X(i)$ for $i = 0, 1, 2$, then we have $c_0 + c_1 s + c_2 s^2 = 0$ for all $s > 0$. The only possible coefficients are $c_0 = c_1 = c_2 = 0$, which implies that $f_X(i) = f_Y(i)$ for $i = 0, 1, 2$, and consequently X and Y have the same distribution.

In other words, X and Y cannot have the same MGF but different pdf's. Thus, the form of the MGF determines the form of the pdf.

This is true in general, although harder to prove in general.

Theorem 2.5.3 **Uniqueness** If X_1 and X_2 have respective CDFs $F_1(x)$ and $F_2(x)$, and MGFs $M_1(t)$ and $M_2(t)$, then $F_1(x) = F_2(x)$ for all real x if and only if $M_1(t) = M_2(t)$ for all t in some interval $-h < t < h$ for some $h > 0$. ■

For nonnegative integer-valued random variables, the derivation of moments often is made more tractable by first considering another type of expectation known as a factorial moment.

FACTORIAL MOMENTS

Definition 2.5.2

The rth **factorial moment** of X is

$$E[X(X-1)\cdots(X-r+1)] \tag{2.5.5}$$

and the **factorial moment generating function** (FMGF) of X is

$$G_X(t) = E(t^X) \tag{2.5.6}$$

if this expectation exists for all t in some interval of the form $1 - h < t < 1 + h$.

The FMGF is more tractable than the MGF in some problems. Also note that the FMGF sometimes is called the **probability generating function**. This is because for nonnegative integer-valued random variables X, $P[X = r] = G_X^{(r)}(0)/r!$, which means that the FMGF uniquely determines the distribution. Also note the following relationship between the FMGF and MGF:

$$G_X(t) = E(t^X) = E(e^{X \ln t}) = M_X(\ln t)$$

Theorem 2.5.4 If X has a FMGF, $G_X(t)$, then

$$G'_X(1) = E(X) \tag{2.5.7}$$

$$G''_X(1) = E[X(X-1)] \tag{2.5.8}$$

$$G_X^{(r)}(1) = E[X(X-1)\cdots(X-r+1)] \tag{2.5.9}$$

Proof

See Exercise 35. ■

It is possible to compute regular moments from factorial moments. For example, notice that $E[X(X-1)] = E(X^2 - X) = E(X^2) - E(X)$, so that

$$E(X^2) = E(X) + E[X(X-1)] \tag{2.5.10}$$

Example 2.5.5 We consider the discrete distribution of Example 2.5.3. The FMGF of X is

$$G_X(t) = M_X(\ln t)$$
$$= \frac{1}{2-t} \qquad t < 2$$

Notice that higher derivatives are easily obtained for the FMGF, which was not the case for the MGF. In particular, the rth derivative is

$$G_X^{(r)}(t) = r!(2-t)^{-r-1}$$

Consequently, $E(X) = G_X'(1) = 1!(2-1)^{-2} = 1$, and $E[X(X-1)] = G_X''(1) = 2!(2-1)^{-3} = 2$. It follows that $E(X^2) = E(X) + 2 = 3$, and thus, $\text{Var}(X) = 3 - 1^2 = 2$.

SUMMARY

The purpose of this chapter was to develop a mathematical structure for expressing a probability model for the possible outcomes of an experiment when these outcomes cannot be predicted deterministically. A random variable, which is a real-valued function defined on a sample space, and the associated probability density function (pdf) provide a reasonable approach to assigning probabilities when the outcomes of an experiment can be quantified. Random variables often can be classified as either discrete or continuous, and the method of assigning probability to a real event A involves summing the pdf over values of A in the discrete case, and integrating the pdf over the set A in the continuous case. The cumulative distribution function (CDF) provides a unified approach for expressing the distribution of probability to the possible values of the random variable.

The moments are special expected values, which include the mean and variance as particular cases, and also provide descriptive measures for other characteristics such as skewness of a distribution.

Bounds for the probabilities of certain types of events can be expressed in terms of expected values. An important bound of this sort is given by the Chebychev inequality.

EXERCISES

1. Let $e = (i, j)$ represent an arbitrary outcome resulting from two rolls of the four-sided die of Example 2.1.1. Tabulate the discrete pdf and sketch the graph of the CDF for the following random variables:
   (a) $Y(e) = i + j$.

(b) $Z(e) = i - j$.
(c) $W(e) = (i - j)^2$.

2. A game consists of first rolling an ordinary six-sided die once and then tossing an unbiased coin once. The score, which consists of adding the number of spots showing on the die to the number of heads showing on the coin (0 or 1), is a random variable, say X. List the possible values of X and tabulate the values of:
   (a) the discrete pdf.
   (b) the CDF at its points of discontinuity.
   (c) Sketch the graph of the CDF.
   (d) Find $P[X > 3]$.
   (e) Find the probability that the score is an odd integer.

3. A bag contains three coins, one of which has a head on both sides while the other two coins are normal. A coin is chosen at random from the bag and tossed three times. The number of heads is a random variable, say X.
   (a) Find the discrete pdf of X. (*Hint:* Use the Law of Total Probability with B_1 = a normal coin and B_2 = two-headed coin.)
   (b) Sketch the discrete pdf and the CDF of X.

4. A box contains five colored balls, two black and three white. Balls are drawn successively without replacement. If X is the number of draws until the last black ball is obtained, find the discrete pdf $f(x)$.

5. A discrete random variable has pdf $f(x)$.
   (a) If $f(x) = k(1/2)^x$ for $x = 1, 2, 3$, and zero otherwise, find k.
   (b) Is a function of the form $f(x) = k[(1/2)^x - 1/2]$ for $x = 0, 1, 2$ a pdf for any k?

6. Denote by $[x]$ the greatest integer not exceeding x. For the pdf in Example 2.2.1, show that the CDF can be represented as $F(x) = ([x]/12)^2$ for $0 < x < 13$, zero if $x \leqslant 0$, and one if $x \geqslant 13$.

7. A discrete random variable X has a pdf of the form $f(x) = c(8 - x)$ for $x = 0, 1, 2, 3, 4, 5$, and zero otherwise.
   (a) Find the constant c.
   (b) Find the CDF, $F(x)$.
   (c) Find $P[X > 2]$.
   (d) Find $E(X)$.

8. A nonnegative integer-valued random variable X has a CDF of the form $F(x) = 1 - (1/2)^{x+1}$ for $x = 0, 1, 2, \ldots$ and zero if $x < 0$.
   (a) Find the pdf of X.
   (b) Find $P[10 < X \leqslant 20]$.
   (c) Find $P[X \text{ is even}]$.

9. Sometimes it is desirable to assign numerical "code" values to experimental responses that are not basically of numerical type. For example, in testing the color preferences of experimental subjects, suppose that the colors blue, green, and red occur with probabilities

1/4, 1/4, and 1/2, respectively. A different integer value is assigned to each color, and this corresponds to a random variable X that can take on one of these three integer values.

(a) Can $f(x) = (1/4)^{|x|}(1/2)^{1-|x|}$ for $x = -1, 1, 0$ be used as a pdf for this experiment?

(b) Can $f(x) = \binom{2}{x}(1/2)^2$ for $x = 0, 1, 2$ be used?

(c) Can $f(x) = (1 - x)/4$ for $x = -1, 0, 2$ be used?

10. Let X be a discrete random variable such that $P[X = x] > 0$ if $x = 1, 2, 3$, or 4, and $P[X = x] = 0$ otherwise. Suppose the CDF is $F(x) = .05x(1 + x)$ at the values $x = 1, 2, 3$, or 4.

(a) Sketch the graph of the CDF.

(b) Sketch the graph of the discrete pdf $f(x)$.

(c) Find $E(X)$.

11. A player rolls a six-sided die and receives a number of dollars corresponding to the number of dots on the face that turns up. What amount should the player pay for rolling to make this a "fair" game?

12. A continuous random variable X has pdf given by $f(x) = c(1 - x)x^2$ if $0 < x < 1$ and zero otherwise.

(a) Find the constant c.

(b) Find $E(X)$.

13. A function $f(x)$ has the following form:

$$f(x) = kx^{-(k+1)} \qquad 1 < x < \infty$$

and zero otherwise.

(a) For what values of k is $f(x)$ a pdf?

(b) Find the CDF based on (a).

(c) For what values of k does $E(X)$ exist?

14. Determine whether each of the following functions could be a CDF over the indicated part of the domain:

(a) $F(x) = e^{-x};\ 0 \leqslant x < \infty$.

(b) $F(x) = e^{x};\ -\infty < x \leqslant 0$.

(c) $F(x) = 1 - e^{-x};\ -1 \leqslant x < \infty$.

15. Find the pdf corresponding to each of the following CDFs:

(a) $F(x) = (x^2 + 2x + 1)/16;\ -1 \leqslant x \leqslant 3$.

(b) $F(x) = 1 - e^{-\lambda x} - \lambda x e^{-\lambda x};\ 0 \leqslant x < \infty;\ \lambda > 0$.

16. If $f_i(x)$, $i = 1, 2, \ldots, n$, are pdf's, show that

$$\sum_{i=1}^{n} p_i f_i(x) \text{ is a pdf where } p_i \geqslant 0 \text{ and } \sum_{i=1}^{n} p_i = 1$$

17. A random variable X has a CDF such that

$$F(x) = \begin{cases} x/2 & 0 < x \leqslant 1 \\ x - 1/2 & 1 < x \leqslant 3/2 \end{cases}$$

(a) Graph $F(x)$.
(b) Graph the pdf $f(x)$.
(c) Find $P[X \leqslant 1/2]$.
(d) Find $P[X \geqslant 1/2]$.
(e) Find $P[X \leqslant 1.25]$.
(f) What is $P[X = 1.25]$?

18. A continuous random variable X has a pdf of the form $f(x) = 2x/9$ for $0 < x < 3$, and zero otherwise.

(a) Find the CDF of X.
(b) Find $P[X \leqslant 2]$.
(c) Find $P[-1 < X < 1.5]$.
(d) Find a number m such that $P[X \leqslant m] = P[X \geqslant m]$.
(e) Find $E(X)$.

19. A random variable X has the pdf

$$f(x) = \begin{cases} x^2 & \text{if } 0 < x \leqslant 1 \\ 2/3 & \text{if } 1 < x \leqslant 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the median of X.
(b) Sketch the graph of the CDF and show the position of the median on the graph.

20. A continuous random variable X has CDF given by

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ 2(x - 2 + 1/x) & \text{if } 1 \leqslant x \leqslant 2 \\ 1 & \text{if } 2 < x \end{cases}$$

(a) Find the $100 \times p$th percentile of the distribution with $p = 1/3$.
(b) Find the pdf of X.

21. Verify that the following function has the four properties of Theorem 2.2.3, and find the points of discontinuity, if any:

$$F(x) = \begin{cases} 0.25e^x & \text{if } -\infty < x < 0 \\ 0.5 & \text{if } 0 \leqslant x < 1 \\ 1 - e^{-x} & \text{if } 1 \leqslant x < \infty \end{cases}$$

22. For the CDF, $F(x)$, of Exercise 21, find a CDF of discrete type, $F_d(x)$, and a CDF of continuous type, $F_c(x)$, and a number $0 < a < 1$ such that

$$F(x) = aF_d(x) + (1 - a)F_c(x)$$

23. Let X be a random variable with discrete pdf $f(x) = x/8$ if $x = 1, 2, 5$, and zero otherwise. Find:

(a) $E(X)$.

(b) $\text{Var}(X)$.

(c) $E(2X + 3)$.

24. Let X be continuous with pdf $f(x) = 3x^2$ if $0 < x < 1$, and zero otherwise. Find:

(a) $E(X)$.

(b) $\text{Var}(X)$.

(c) $E(X^r)$.

(d) Find $E(3X - 5X^2 + 1)$.

25. Let X be continuous with pdf $f(x) = 1/x^2$ if $1 < x < \infty$, and zero otherwise.

(a) Does $E(X)$ exist?

(b) Does $E(1/X)$ exist?

(c) For what values of k does $E(X^k)$ exist?

26. At a computer store, the annual demand for a particular software package is a discrete random variable X. The store owner orders four copies of the package at \$10 per copy and charges customers \$35 per copy. At the end of the year the package is obsolete and the owner loses the investment on unsold copies. The pdf of X is given by the following table:

x	0	1	2	3	4
$f(x)$	.1	.3	.3	.2	.1

(a) Find $E(X)$.

(b) Find $\text{Var}(X)$.

(c) Express the owner's net profit Y as a linear function of X, and find $E(Y)$ and $\text{Var}(Y)$.

27. The measured radius of a circle, R, has pdf $f(r) = 6r(1 - r)$, $0 < r < 1$. Find:

(a) the expected value of the radius.

(b) the expected circumference.

(c) the expected area.

28. Prove Theorem 2.4.5 for the continuous case. *Hint:* Use the transformation $y = x - \mu$ in the integral and note that $g(y) = yf(\mu + y)$ is an odd function of y.

29. Consider the discrete random variable X with pdf given by the following table:

x	-3	-1	0	2	$2\sqrt{2}$
$f(x)$	1/4	1/4	$(6-3\sqrt{2})/16$	1/8	$3\sqrt{2}/16$

The distribution of X is not symmetric. Why? Show that $\mu_3 = 0$.

30. Let X be a nonnegative continuous random variable with CDF $F(x)$ and $E(X) < \infty$. Use integration by parts to show that

$$E(X) = \int_0^\infty [1 - F(x)]\, dx$$

Note: For any continuous random variable with $E(|X|) < \infty$, this result extends to

$$E(X) = -\int_{-\infty}^0 F(x)\, dx + \int_0^\infty [1 - F(x)]\, dx$$

31. (a) Use Chebychev's inequality to obtain a lower bound on $P[5/8 < X < 7/8]$ in Exercise 24. Is this a useful bound?
    (b) Rework (a) for the probability $P[1/2 < X < 1]$.
    (c) Compare this bound to the exact probability.

32. Consider the random variable X of Example 2.1.1, which represents the largest of two numbers that occur on two rolls of a four-sided die.
    (a) Find the expected value of X.
    (b) Find the variance of X.

33. Suppose $E(X) = \mu$ and $\operatorname{Var}(X) = \sigma^2$. Find the approximate mean and variance of:
    (a) e^X.
    (b) $1/X$ (assuming $\mu \neq 0$).
    (c) $\ln(X)$ (assuming $X > 0$).

34. Suppose that X is a random variable with MGF $M_X(t) = (1/8)e^t + (1/4)e^{2t} + (5/8)e^{5t}$.
    (a) What is the distribution of X?
    (b) What is $P[X = 2]$?

35. Prove Theorem 2.5.4 for a nonnegative integer-valued random variable X.

36. Assume that X is a continuous random variable with pdf

$$f(x) = \exp[-(x + 2)] \text{ if } -2 < x < \infty \text{ and zero otherwise.}$$

    (a) Find the moment generating function of X.
    (b) Use the MGF of (a) to find $E(X)$ and $E(X^2)$.

37. Use the FMGF of Example 2.5.5 to find $E[X(X - 1)(X - 2)]$, and then find $E(X^3)$.

38. In Exercise 26, suppose instead of ordering four copies of the software package, the store owner orders c copies ($0 \leqslant c \leqslant 4$). Then the number sold, say S, is the smaller of c or X.
    (a) Express the net profit Y as a linear function of S.
    (b) Find $E(Y)$ for each value of c and indicate the solution c that maximizes the expected profit.

39. Show that $\sigma^2 = E[X(X - 1)] - \mu(\mu - 1)$.

40. Let $\psi_X(t) = \ln[M_X(t)]$, where $M_X(t)$ is a MGF. The function $\psi_X(t)$ is called the **cumulant generating function** of X, and the value of the rth derivative evaluated at $t = 0$, $\kappa_r = \psi_X^{(r)}(0)$, is called the rth cumulant of X.

(a) Show that $\mu = \psi'_X(0)$.

(b) Show that $\sigma^2 = \psi''_X(0)$.

(c) Use $\psi_X(t)$ to find μ and σ^2 for the random variable of Exercise 36.

(d) Use $\psi_X(t)$ to find μ and σ^2 for the random variable of Example 2.5.5.