9

Some Important Continuous Random Variables

9.1 Introduction

In this chapter we shall pursue the task we set for ourselves in Chapter 8, and study in considerable detail, a number of important continuous random variables and their characteristics. As we have pointed out before, in many problems it becomes mathematically simpler to consider an "idealized" range space for a random variable X, in which *all* possible real numbers (in some specified interval or set of intervals) may be considered as possible outcomes. In this manner we are led to continuous random variables. Many of the random variables we shall now introduce have important applications and we shall defer until a later chapter the discussion of some of these applications.

9.2 The Normal Distribution

One of the most important continuous random variables is the following.

Definition. The random variable X, assuming all real values $-\infty < x < \infty$, has a *normal* (*or Gaussian*) distribution if its pdf is of the form

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right), \qquad -\infty < x < \infty. \tag{9.1}$$

The parameters μ and σ must satisfy the conditions $-\infty < \mu < \infty, \sigma > 0$. Since we shall have many occasions to refer to the above distribution we shall use the following notation: X has distribution $N(\mu, \sigma^2)$ if and only if its probability distribution is given by Eq. (9.1). [We shall frequently use the notation $\exp(t)$ to represent e^t.]

We shall delay until Chapter 12 discussing the reason for the great importance of this distribution. Let us simply state now that the *normal distribution serves as an excellent approximation to a large class of distributions* which have great practical importance. Furthermore, this distribution has a number of very desirable mathematical properties which make it possible to derive important theoretical results.

9.3 Properties of the Normal Distribution

(a) Let us state that f is a legitimate pdf. Obviously $f(x) \geq 0$. We must still check that $\int_{-\infty}^{+\infty} f(x)\, dx = 1$. We observe that by letting $t = (x - \mu)/\sigma$, we may write $\int_{-\infty}^{+\infty} f(x)\, dx$ as $(1/\sqrt{2\pi}) \int_{-\infty}^{+\infty} e^{-t^2/2}\, dt = I$.

The "trick" used to evaluate this integral (and it *is* a trick) is to consider, instead of I, the square of this integral, namely I^2. Thus

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-t^2/2}\, dt \int_{-\infty}^{+\infty} e^{-s^2/2}\, ds$$
$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(s^2+t^2)/2}\, ds\, dt.$$

Let us introduce polar coordinates to evaluate this double integral:

$$s = r\cos\alpha, \qquad t = r\sin\alpha.$$

Hence the element of area $ds\, dt$ becomes $r\, dr\, d\alpha$. As s and t vary between $-\infty$ and $+\infty$, r varies between 0 and ∞, while α varies between 0 and 2π. Thus

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} re^{-r^2/2}\, dr\, d\alpha$$
$$= \frac{1}{2\pi} \int_0^{2\pi} -e^{-r^2/2}\Big|_0^{\infty}\, d\alpha$$
$$= \frac{1}{2\pi} \int_0^{2\pi} d\alpha = 1.$$

Hence $I = 1$ as was to be shown.

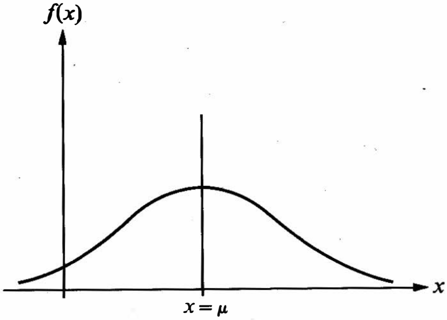

FIGURE 9.1

(b) Let us consider the appearance of the graph of f. It has the well-known bell shape indicated in Fig. 9.1. Since f depends on x only through the expression $(x - \mu)^2$, it is evident that the graph of f will be *symmetric* with respect to μ. For example, if $x = \mu + 2$, $(x - \mu)^2 = (\mu + 2 - \mu)^2 = 4$, while for $x = \mu - 2$, $(x - \mu)^2 = (\mu - 2 - \mu)^2 = 4$ also.

The parameter σ may also be interpreted geometrically. We note that at $x = \mu$ the graph of f is concave downward. As $x \to \pm\infty$, $f(x) \to 0$, asymptotically. Since $f(x) \geq 0$ for all x, this means that for large (positive or negative) values of x, the graph of f is concave upward. The point at which the concavity changes is called a point of inflection, and it is located by solving the equation $f''(x) = 0$. When we do this we find that the points of inflection occur at $x = \mu \pm \sigma$. That is, σ units to the right and to the left of μ the graph of f changes concavity. Thus if σ is relatively large, the graph of f tends to be "flat," while if σ is small the graph of f tends to be quite "peaked."

(c) In addition to the geometric interpretation of the parameters μ and σ, the following important probabilistic meaning may be associated with these quantities. Consider

$$E(X) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{+\infty} x \exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right) dx.$$

Letting $z = (x - \mu)/\sigma$ and noting that $dx = \sigma\, dz$, we obtain

$$E(X) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} (\sigma z + \mu)e^{-z^2/2}\, dz$$

$$= \frac{1}{\sqrt{2\pi}}\,\sigma\int_{-\infty}^{+\infty} ze^{-z^2/2}\, dz + \mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-z^2/2}\, dz.$$

The first of the above integrals equals zero since the integrand, say $g(z)$, has the property that $g(z) = -g(-z)$, and hence g is an odd function. The second integral (without the factor μ) represents the total area under the normal pdf and hence equals unity. Thus $E(X) = \mu$.

Consider next

$$E(X^2) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right) dx.$$

Again letting $z = (x - \mu)/\sigma$, we obtain

$$E(X^2) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} (\sigma z + \mu)^2 e^{-z^2/2}\, dz$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \sigma^2 z^2 e^{-z^2/2}\, dz + 2\mu\sigma\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} ze^{-z^2/2}\, dz$$

$$+ \mu^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-z^2/2}\, dz.$$

The second integral again equals zero by the argument used above. The last integral (without the factor μ^2) equals unity. To evaluate $(1/\sqrt{2\pi})\int_{-\infty}^{+\infty} z^2 e^{-z^2/2}\, dz$, we integrate by parts letting $ze^{-z^2/2} = dv$ and $z = u$. Hence $v = -e^{-z^2/2}$ while $dz = du$. We obtain

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z^2 e^{-z^2/2}\, dz = \frac{-ze^{-z^2/2}}{\sqrt{2\pi}}\bigg|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-z^2/2}\, dz = 0 + 1 = 1.$$

Therefore, $E(X^2) = \sigma^2 + \mu^2$, and hence $V(X) = E(X^2) - (E(X))^2 = \sigma^2$. *Thus we find that the two parameters μ and σ^2 which characterize the normal distribution are the expectation and variance of X, respectively.* To put it differently, if we know that X is normally distributed we know only that its probability distribution is of a certain type (or belongs to a certain family.) If in addition, we know $E(X)$ and $V(X)$, the distribution of X is completely specified. As we mentioned above, the graph of the pdf of a normally distributed random variable is symmetric about μ. The steepness of the graph is determined by σ^2 in the sense that if X has distribution $N(\mu, \sigma_1^2)$ and Y has distribution $N(\mu, \sigma_2^2)$, where $\sigma_1^2 > \sigma_2^2$, then their pdf's would have the *relative* shapes shown in Fig. 9.2.

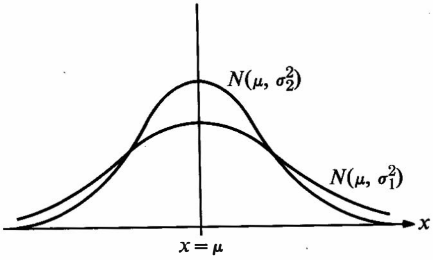

FIGURE 9.2

(d) If X has distribution $N(0, 1)$ we say that X has a *standardized normal* distribution. That is, the pdf of X may be written as

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{9.2}$$

(We shall use the letter φ exclusively for the pdf of the above random variable X.) The importance of the standardized normal distribution is due to the fact that it is *tabulated*. Whenever X has distribution $N(\mu, \sigma^2)$ we can always obtain the standardized form by simply taking a *linear* function of X as the following theorem indicates.

Theorem 9.1. If X has the distribution $N(\mu, \sigma^2)$ and if $Y = aX + b$, then Y has the distribution $N(a\mu + b, a^2\sigma^2)$.

Proof: The fact that $E(Y) = a\mu + b$ and that $V(Y) = a^2\sigma^2$ follows immediately from the properties of expectation and variance discussed in Chapter 7. To show that in fact Y is normally distributed, we may apply Theorem 5.1, since $aX + b$ is either a decreasing or increasing function of X, depending on the sign of a. Hence if g is the pdf of Y, we have

$$g(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}\left[\frac{y-b}{a} - \mu\right]^2\right)\left|\frac{1}{a}\right|$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma|a|} \exp\left(-\frac{1}{2\sigma^2 a^2}[y - (a\mu + b)]^2\right),$$

which represents the pdf of a random variable with distribution $N(a\mu + b, a^2\sigma^2)$.

Corollary. If X has distribution $N(\mu, \sigma^2)$ and if $Y = (X - \mu)/\sigma$, then Y has distribution $N(0, 1)$.

Proof: It is evident that Y is a linear function of X, and hence Theorem 9.1 applies.

Note: The importance of this corollary is that by *changing the units* in which the variable is measured we can obtain the standardized distribution (see d). By doing this we obtain a distribution with no unspecified parameters, a most desirable situation from the point of view of tabulating the distribution (see the next section).

9.4 Tabulation of the Normal Distribution

Suppose that X has distribution $N(0, 1)$. Then

$$P(a \leq X \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, dx.$$

This integral cannot be evaluated by ordinary means. (The difficulty stems from the fact that we cannot apply the Fundamental Theorem of the Calculus since we cannot find a function whose derivative equals $e^{-x^2/2}$.) However, methods of numerical integration can be used to evaluate integrals of the above form, and in fact $P(X \leq s)$ has been *tabulated.*

The cdf of the standardized normal distribution will be consistently denoted by Φ. That is,

$$\Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-x^2/2}\, dx. \tag{9.3}$$

(See Fig. 9.3.) The function Φ has been extensively tabulated, and an excerpt of such a table is given in the Appendix. We may now use the tabulation of the function Φ in order to evaluate $P(a \leq X \leq b)$, where X has the standardized $N(0, 1)$ distribution since

$$P(a \leq X \leq b) = \Phi(b) - \Phi(a).$$

The particular usefulness of the above tabulation is due to the fact that if X has *any* normal distribution $N(\mu, \sigma^2)$, the tabulated function Φ may be used to evaluate probabilities associated with X.

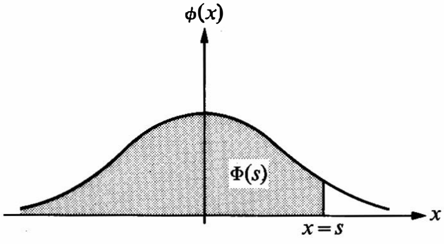

FIGURE 9.3

We simply use Theorem 9.1 to note that if X has distribution $N(\mu, \sigma^2)$, then $Y = (X - \mu)/\sigma$ has distribution $N(0, 1)$. Hence

$$P(a \leq X \leq b) = P\left(\frac{a - \mu}{\sigma} \leq Y \leq \frac{b - \mu}{\sigma}\right)$$

$$= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right). \tag{9.4}$$

It is also evident from the definition of Φ (see Fig. 9.3) that

$$\Phi(-x) = 1 - \Phi(x). \tag{9.5}$$

This relationship is particularly useful since in most tables the function Φ is tabulated only for positive values of x.

Finally let us compute $P(\mu - k\sigma \leq X \leq \mu + k\sigma)$, where X has distribution $N(\mu, \sigma^2)$. The above probability may be expressed in terms of the function Φ by writing

$$\begin{aligned} P(\mu - k\sigma \leq X \leq \mu + k\sigma) &= P\left(-k \leq \frac{X - \mu}{\sigma} \leq k\right) \\ &= \Phi(k) - \Phi(-k). \end{aligned}$$

Using Eq. (9.5), we have for $k > 0$,

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) = 2\Phi(k) - 1. \tag{9.6}$$

Note that the above probability is independent of μ and σ. In words: The probability that a random variable with distribution $N(\mu, \sigma^2)$ takes values within k standard deviations of the expected value depends only on k and is given by Eq. (9.6).

Note: We shall have many occasions to refer to "tabulated functions." In a sense when an expression can be written in terms of tabulated functions, the problem is "solved." (With the availability of modern computing facilities, many functions which are not tabulated may be readily evaluated. Although we do not expect that everyone has easy access to a computer, it does not seem too unreasonable to suppose that certain common tables are available.) Thus we should feel as much at ease with the function $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-s^2/2}\, ds$ as with the function $f(x) = \sqrt{x}$. Both these functions are tabulated, and in either case we might experience some difficulty in evaluating the function directly for $x = 0.43$, for example. In the Appendix various tables are listed of some of the most important functions we shall encounter in our work. Occasional references will be given to other tables not listed in this text.

EXAMPLE 9.1. Suppose that X has distribution $N(3, 4)$. We want to find a number c such that

$$P(X > c) = 2P(X \leq c).$$

We note that $(X - 3)/2$ has distribution $N(0, 1)$. Hence

$$P(X > c) = P\left(\frac{X - 3}{2} > \frac{c - 3}{2}\right) = 1 - \Phi\left(\frac{c - 3}{2}\right).$$

Also,

$$P(X \leq c) = P\left(\frac{X - 3}{2} \leq \frac{c - 3}{2}\right) = \Phi\left(\frac{c - 3}{2}\right).$$

The above condition may therefore be written as $1 - \Phi[(c - 3)/2] = 2\Phi[(c - 3)/2]$. This becomes $\Phi[(c - 3)/2] = \frac{1}{3}$. Hence (from the tables of the normal distribution) we find that $(c - 3)/2 = -0.43$, yielding $c = 2.14$.

EXAMPLE 9.2. Suppose that the breaking strength of cotton fabric (in pounds), say X, is normally distributed with $E(X) = 165$ and $V(X) = 9$. Assume furthermore that a sample of this fabric is considered to be defective if $X < 162$. What is the probability that a fabric chosen at random will be defective?

We must compute $P(X < 162)$. However

$$P(X < 162) = P\left(\frac{X - 165}{3} < \frac{162 - 165}{3}\right)$$
$$= \Phi(-1) = 1 - \Phi(1) = 0.159.$$

Note: An immediate objection to the use of the normal distribution may be raised here. For it is obvious that X, the strength of cotton fabric, cannot assume negative values, while a normally distributed random variable may assume all positive and negative values. However, the above model (apparently not valid in view of the objection just raised) assigns negligible probability to the event $\{X < 0\}$. That is,

$$P(X < 0) = P\left(\frac{X - 165}{3} < \frac{0 - 165}{3}\right) = \Phi(-55) \simeq 0.$$

The point raised here will occur frequently: A certain random variable X which we know cannot assume negative values (say) will be assumed to have a normal distribution, thus taking on (theoretically, at least) both positive and negative values. So long as the parameters μ and σ^2 are chosen so that $P(X < 0)$ is essentially zero, such a representation is perfectly valid.

The problem of finding the pdf of a function of a random variable, say $Y = H(X)$, as discussed in Chapter 5, occurs in the present context in which the random variable X is normally distributed.

EXAMPLE 9.3. Suppose that the radius R of a ball bearing is normally distributed with expected value 1 and variance 0.04. Find the pdf of the volume of the ball bearing.

The pdf of the random variable R is given by

$$f(r) = \frac{1}{\sqrt{2\pi}\,(0.2)} \exp\left(-\frac{1}{2}\left[\frac{r - 1}{0.2}\right]^2\right).$$

Since V is a monotonically increasing function of R, we may directly apply Theorem 5.1 for the pdf of $V = \frac{4}{3}\pi R^3$, and obtain $g(v) = f(r)(dr/dv)$, where r is everywhere expressed in terms of v. From the above relationship, we obtain $r = \sqrt[3]{3v/4\pi}$. Hence $dr/dv = (1/4\pi)(3v/4\pi)^{-2/3}$. By substituting these expressions into the above equation, we obtain the desired pdf of V.

EXAMPLE 9.4. Suppose that X, the inside diameter (millimeters) of a nozzle, is a normally distributed random variable with expectation μ and variance 1. If X does not meet certain specifications, a loss is assessed to the manufacturer. Specifically, suppose that the profit T (per nozzle) is the following function of X:

$$\begin{aligned} T &= C_1 \text{ (dollars)} \quad \text{if} \quad 10 \leq X \leq 12, \\ &= -C_2 \quad \text{if} \quad X < 10, \\ &= -C_3 \quad \text{if} \quad X > 12. \end{aligned}$$

Hence the expected profit (per nozzle) may be written as

$$\begin{aligned} E(T) &= C_1[\Phi(12 - \mu) - \Phi(10 - \mu)] \\ &\qquad - C_2[\Phi(10 - \mu)] - C_3[1 - \Phi(12 - \mu)] \\ &= (C_1 + C_3)\Phi(12 - \mu) - (C_1 + C_2)\Phi(10 - \mu) - C_3. \end{aligned}$$

Suppose that the manufacturing process can be adjusted so that different values of μ may be achieved. For what value of μ is the expected profit maximum? We must compute $dE(T)/(d\mu)$ and set it equal to zero. Denoting, as usual, the pdf of the $N(0, 1)$ distribution by φ, we have

$$\frac{dE(T)}{d\mu} = (C_1 + C_3)[-\varphi(12 - \mu)] - (C_1 + C_2)[-\varphi(10 - \mu)].$$

Hence

$$-(C_1 + C_3)\frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(12 - \mu)^2\right) + (C_1 + C_2)\frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(10 - \mu)^2\right) = 0.$$

Or

$$e^{22-2\mu} = \frac{C_1 + C_3}{C_1 + C_2}.$$

Thus

$$\mu = 11 - \tfrac{1}{2} \ln \left(\frac{C_1 + C_3}{C_1 + C_2}\right).$$

[It is an easy matter for the reader to check that the above yields a maximum value for $E(T)$.]

Notes: (a) If $C_2 = C_3$, that is, if too large or too small a diameter X is an equally serious defect, then the value of μ for which the maximum value of $E(T)$ is attained is $\mu = 11$. If $C_2 > C_3$, the value of μ is <11, while if $C_2 < C_3$, the value of μ is >11. As $\mu \to +\infty$, $E(T) \to -C_3$, while if $\mu \to -\infty$, $E(T) \to -C_2$.

(b) Consider the following cost values: $C_1 = \$10$, $C_2 = \$3$, and $C_3 = \$2$. Hence the value of μ for which $E(T)$ is maximized equals $\mu = 11 - \frac{1}{2} \ln [\frac{12}{13}] = \$11.04$. Thus the maximum value attained by $E(T)$ equals \$6.04 per nozzle.

9.5 The Exponential Distribution

Definition. A continuous random variable X assuming all nonnegative values is said to have an *exponential distribution* with parameter $\alpha > 0$ if its pdf is given by

$$f(x) = \alpha e^{-\alpha x}, \quad x > 0$$
$$= 0, \quad \text{elsewhere.} \tag{9.7}$$

(See Fig. 9.4.) [A straightforward integration reveals that

$$\int_0^\infty f(x)\,dx = 1$$

and hence Eq. (9.7) does represent a pdf.]

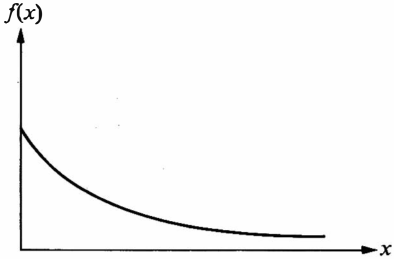

FIGURE 9.4

The exponential distribution plays an important role in describing a large class of phenomena, particularly in the area of reliability theory. We shall devote Chapter 11 to some of these applications. For the moment, let us simply investigate some of the properties of the exponential distribution.

9.6 Properties of the Exponential Distribution

(a) The cdf F of the exponential distribution is given by

$$F(x) = P(X \leq x) = \int_0^x \alpha e^{-\alpha t}\,dt = 1 - e^{-\alpha x}, \quad x \geq 0 \tag{9.8}$$
$$= 0, \quad \text{elsewhere.}$$

[Hence $P(X > x) = e^{-\alpha x}$.]

(b) The expected value of X is obtained as follows:

$$E(X) = \int_0^\infty x\alpha e^{-\alpha x}\,dx.$$

Integrating by parts and letting $\alpha e^{-\alpha x}\,dx = dv$, $x = u$, we obtain $v = -e^{-\alpha x}$, $du = dx$. Thus,

$$E(X) = [-xe^{-\alpha x}]\big|_0^\infty + \int_0^\infty e^{-\alpha x}\,dx = \frac{1}{\alpha}. \tag{9.9}$$

Thus the expected value equals the reciprocal of the parameter α. [By simply relabeling the parameter $\alpha = 1/\beta$, we could have written the pdf of X as $f(x) = (1/\beta)e^{-x/\beta}$. In this form, the parameter β equals the expected value of X. However, we shall continue to use the form of Eq. (9.7).]

(c) The variance of X may be obtained by a similar integration. We find that $E(X^2) = 2/\alpha^2$ and therefore

$$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{\alpha^2}. \tag{9.10}$$

(d) The exponential distribution has the following interesting property, analogous to Eq. (8.6) described for the geometric distribution. Consider for any $s, t > 0$, $P(X > s + t \mid X > s)$. We have

$$P(X > s + t \mid X > s) = \frac{P(X > s + t)}{P(X > s)} = \frac{e^{-\alpha(s+t)}}{e^{-\alpha s}} = e^{-\alpha t}.$$

Hence

$$P(X > s + t \mid X > s) = P(X > t). \tag{9.11}$$

Thus we have shown that the exponential distribution also has the property of having "no memory" as did the geometric distribution. (See Note following Theorem 8.4.) We shall make considerable use of this property in applying the exponential distribution to fatigue models in Chapter 11.

Note: As was true in the case of the geometric distribution, the *converse* of Property (d) also holds. The only continuous random variable X assuming nonnegative values for which $P(X > s + t \mid X > s) = P(X > t)$ for all $s, t > 0$, is an exponentially distributed random variable. [Although we shall not prove this here, it might be pointed out that the crux of the argument involves the fact that the only continuous function G having the property that $G(x + y) = G(x)G(y)$ for all $x, y > 0$, is $G(x) = e^{-kx}$. It is easily seen that if we define $G(x) = 1 - F(x)$, where F is the cdf of X, then G will satisfy this condition.]

EXAMPLE 9.5. Suppose that a fuse has a life length X which may be considered as a continuous random variable with an exponential distribution. There are two processes by which the fuse may be manufactured. Process I yields an expected life length of 100 hours (that is, the parameter equals 100^{-1}), while process II yields an expected life length of 150 hours (that is, the parameter equals 150^{-1}). Suppose that process II is twice as costly (per fuse) as process I, which costs C dollars per fuse. Assume, furthermore, that if a fuse lasts less than 200 hours, a loss of K dollars is assessed against the manufacturer. Which process should be used? Let us compute the *expected* cost for each process. For process I, we have

$$\begin{aligned} C_{\text{I}} = \text{cost (per fuse)} &= C \quad \text{if } X > 200 \\ &= C + K \quad \text{if } X \leq 200. \end{aligned}$$

Therefore,

$$\begin{aligned} E(C_{\text{I}}) &= CP(X > 200) + (C + K)P(X \leq 200) \\ &= Ce^{-(1/100)200} + (C + K)(1 - e^{-(1/100)200}) \\ &= Ce^{-2} + (C + K)(1 - e^{-2}) = K(1 - e^{-2}) + C. \end{aligned}$$

By a similar computation we find that

$$E(C_{\text{II}}) = K(1 - e^{-4/3}) + 2C.$$

Thus

$$E(C_{\text{II}}) - E(C_{\text{I}}) = C + K(e^{-2} - e^{-4/3}) = C - 0.13K.$$

Hence we prefer process I, provided that $C > 0.13K$.

EXAMPLE 9.6. Suppose that X has an exponential distribution with parameter α. Then $E(X) = 1/\alpha$. Let us compute the probability that X exceeds its expected value (Fig. 9.5). We have

$$P\left(X > \frac{1}{\alpha}\right) = e^{-\alpha(1/\alpha)} = e^{-1} < \tfrac{1}{2}.$$

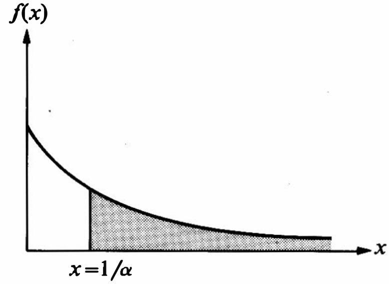

FIGURE 9.5

EXAMPLE 9.7. Suppose that T, the time to failure of a component, is exponentially distributed. Hence $f(t) = \alpha e^{-\alpha t}$. If n such components are installed, what is the probability that one-half or more of these components are still functioning at the end of t hours? The required probability is

$$\sum_{k=n/2}^{n} \binom{n}{k} (1 - e^{-\alpha t})^{n-k}(e^{-\alpha tk}) \quad \text{if } n \text{ is even;}$$

$$\sum_{k=(n+1)/2}^{n} \binom{n}{k} (1 - e^{-\alpha t})^{n-k}(e^{-\alpha tk}) \quad \text{if } n \text{ is odd.}$$

EXAMPLE 9.8. Suppose that the life length in hours, say T, of a certain electronic tube is a random variable with exponential distribution with parameter β. That is, the pdf is given by $f(t) = \beta e^{-\beta t}$, $t > 0$. A machine using this tube costs C_1 dollars/hour to run. While the machine is functioning, a profit of C_2 dollars/hour is realized. An operator must be hired for a *prearranged* number of hours, say H, and he gets paid C_3 dollars/hour. For what value of H is the *expected profit* greatest?

Let us first get an expression for the profit, say R. We have

$$\begin{aligned} R &= C_2H - C_1H - C_3H \quad &\text{if } T > H \\ &= C_2T - C_1T - C_3H \quad &\text{if } T \le H. \end{aligned}$$

Note that R is a random variable since it is a function of T. Hence

$$\begin{aligned} E(R) &= H(C_2 - C_1 - C_3)P(T > H) - C_3HP(T \le H) \\ &\quad + (C_2 - C_1)\int_0^H t\beta e^{-\beta t}\,dt \\ &= H(C_2 - C_1 - C_3)e^{-\beta H} - C_3H(1 - e^{-\beta H}) \\ &\quad + (C_2 - C_1)[\beta^{-1} - e^{-\beta H}(\beta^{-1} + H)] \\ &= (C_2 - C_1)[He^{-\beta H} + \beta^{-1} - e^{-\beta H}(\beta^{-1} + H)] - C_3H. \end{aligned}$$

To obtain the maximum value of $E(R)$ we differentiate it with respect to H and set the derivative equal to zero. We have

$$\frac{dE(R)}{dH} = (C_2 - C_1)[H(-\beta)e^{-\beta H} + e^{-\beta H} - e^{-\beta H} + (\beta^{-1} + H)(\beta)e^{-\beta H}] - C_3$$
$$= (C_2 - C_1)e^{-\beta H} - C_3.$$

Hence $dE(R)/dH = 0$ implies that

$$H = -\left(\frac{1}{\beta}\right)\ln\left[\frac{C_3}{C_2 - C_1}\right].$$

[In order for the above solution to be meaningful, we must have $H > 0$ which occurs if and only if $0 < C_3/(C_2 - C_1) < 1$, which in turn is equivalent to $C_2 - C_1 > 0$ and $C_2 - C_1 - C_3 > 0$. However, the last condition simply requires that the cost figures be of such a magnitude that a profit may be realized.]

Suppose in particular that $\beta = 0.01$, $C_1 = \$3$, $C_2 = \$10$, and $C_3 = \$4$. Then $H = -100 \ln [\frac{4}{7}] = 55.9$ hours $\simeq 56$ hours. Thus, the operator should be hired for 56 hours in order to achieve the maximum profit. (For a slight modification of the above example, see Problem 9.18.)

9.7 The Gamma Distribution

Let us first introduce a function which is most important not only in probability theory but in many areas of mathematics.

Definition. The *Gamma function*, denoted by Γ, is defined as follows:

$$\Gamma(p) = \int_0^\infty x^{p-1}e^{-x}\,dx, \qquad \text{defined for } p > 0. \tag{9.12}$$

[It can be shown that the above improper integral exists (converges) whenever $p > 0$.] If we integrate the above by parts, letting $e^{-x}\,dx = dv$ and $x^{p-1} = u$, we obtain

$$\Gamma(p) = -e^{-x}x^{p-1}|_0^\infty - \int_0^\infty [-e^{-x}(p-1)x^{p-2}\,dx]$$
$$= 0 + (p-1)\int_0^\infty e^{-x}x^{p-2}\,dx$$
$$= (p-1)\Gamma(p-1). \tag{9.13}$$

Thus we have shown that the Gamma function obeys an interesting recursion relationship. Suppose that p is a *positive integer*, say $p = n$. Then applying Eq. (9.13) repeatedly, we obtain

$$\Gamma(n) = (n-1)\Gamma(n-1)$$
$$= (n-1)(n-2)\Gamma(n-2) = \cdots$$
$$= (n-1)(n-2)\cdots\Gamma(1).$$

However, $\Gamma(1) = \int_0^\infty e^{-x}\,dx = 1$, and hence we have

$$\Gamma(n) = (n-1)! \tag{9.14}$$

(if n is a positive integer). (Thus we may consider the Gamma function to be a generalization of the factorial function.) It is also easy to verify that

$$\Gamma(\tfrac{1}{2}) = \int_0^{\infty} x^{-1/2} e^{-x}\, dx = \sqrt{\pi}. \tag{9.15}$$

(See Problem 9.19.)

With the aid of the Gamma function we can now introduce the Gamma probability distribution.

Definition. Let X be a continuous random variable assuming only nonnegative values. We say that X has a *Gamma probability distribution* if its pdf is given by

$$f(x) = \frac{\alpha}{\Gamma(r)} (\alpha x)^{r-1} e^{-\alpha x}, \qquad x > 0$$

$$= 0, \qquad \text{elsewhere.} \tag{9.16}$$

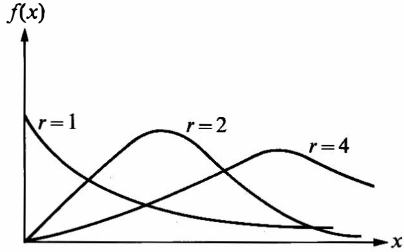

FIGURE 9.6

This distribution depends on *two parameters*, r and α, of which we require $r > 0$, $\alpha > 0$. [Because of the definition of the Gamma function, it is easy to see that $\int_{-\infty}^{+\infty} f(x)\, dx = 1$.] Figure 9.6 shows graphs of the pdf Eq. (9.16) for various values of r and $\alpha = 1$.

9.8 Properties of the Gamma Distribution

(a) If $r = 1$, Eq. (9.16) becomes $f(x) = \alpha e^{-\alpha x}$. Hence the *exponential distribution* is a *special case of the Gamma distribution.* (If r is a positive integer > 1, the Gamma distribution is also related to the exponential distribution but in a slightly different way. We shall refer to this in Chapter 10.)

(b) In most of our applications, the parameter r will be a positive *integer.* In this case, an interesting relationship between the cdf of the Gamma distribution and the Poisson distribution exists which we shall now develop.

Consider the integral $I = \int_a^{\infty} (e^{-y} y^r / r!)\, dy$, where r is a positive integer and $a > 0$. Then $r!I = \int_a^{\infty} e^{-y} y^r\, dy$. Integrating by parts, letting $u = y^r$ and $dv = e^{-y}\, dy$, will yield $du = ry^{r-1}\, dy$ and $v = -e^{-y}$. Hence $r!I = e^{-a}a^r + r\int_a^{\infty} e^{-y} y^{r-1}\, dy$. The integral in this expression is exactly of the same form as the original integral with r replaced by $(r - 1)$. Thus continuing to integrate by parts we obtain, since r is a positive integer,

$$r!I = e^{-a}[a^r + ra^{r-1} + r(r-1)a^{r-2} + \cdots + r!].$$

Therefore

$$I = e^{-a}[1 + a + a^2/2! + \cdots + a^r/r!]$$
$$= \sum_{k=0}^{r} P(Y = k),$$

where Y has a Poisson distribution with parameter a.

We now consider the cdf of the random variable whose pdf is given by Eq. (9.16). Since r is a positive integer, Eq. (9.16) may be written as

$$f(x) = \frac{\alpha}{(r-1)!}(\alpha x)^{r-1}e^{-\alpha x}, \qquad 0 < x$$

and consequently the cdf of X becomes

$$F(x) = 1 - P(X > x)$$
$$= 1 - \int_x^\infty \frac{\alpha}{(r-1)!}(\alpha s)^{r-1}e^{-\alpha s}\,ds, \qquad x > 0.$$

Letting $(\alpha s) = u$, we find that this becomes

$$F(x) = 1 - \int_{\alpha x}^\infty \frac{u^{r-1}e^{-u}}{(r-1)!}\,du, \qquad x > 0.$$

This integral is precisely of the form considered above, namely I (with $a = \alpha x$), and thus

$$F(x) = 1 - \sum_{k=0}^{r-1} e^{-\alpha x}(\alpha x)^k/k!, \qquad x > 0. \tag{9.17}$$

Hence, *the cdf of the Gamma distribution may be expressed in terms of the tabulated cdf of the Poisson distribution.* (We recall that this is valid if the parameter r is a positive integer.)

Note: The result stated in Eq. (9.17), relating the cdf of the Poisson distribution to the cdf of the Gamma distribution, is not as surprising as it might first appear, as the following discussion will indicate.

First of all, recall the relationship between the binomial and Pascal distributions (see Note (b), Section 8.6). A similar relationship exists between the Poisson and Gamma distribution except that the latter is a continuous distribution. When we deal with a Poisson distribution we are essentially concerned about the number of occurrences of some event during a fixed time period. And, as will be indicated, the Gamma distribution arises when we ask for distribution of the *time* required to obtain a specified number of occurrences of the event.

Specifically, suppose X = number of occurrences of the event A during $(0, t]$. Then, under suitable conditions (e.g., satisfying assumptions A_1 through A_5 in Section 8.3) X has a Poisson distribution with parameter αt, where α is the expected number of occurrences of A during a unit time interval. Let T = time required to observe r occurrences of A. We have:

$$\begin{aligned} H(t) = P(T \leq t) &= 1 - P(T > t) \\ &= 1 - P(\text{fewer than } r \text{ occurrences of } A \text{ occur in } (0, t]) \\ &= 1 - P(X < r) \\ &= 1 - \sum_{k=0}^{r-1} \frac{e^{-\alpha t}(\alpha t)^k}{k!}. \end{aligned}$$

Comparing this with Eq. (9.17) establishes the desired relationship.

(c) If X has a Gamma distribution given by Eq. (9.16), we have

$$E(X) = r/\alpha, \qquad V(X) = r/\alpha^2. \tag{9.18}$$

Proof: See Problem 9.20.

9.9 The Chi-Square Distribution

A special, very important, case of the Gamma distribution Eq. (9.16) is obtained if we let $\alpha = \frac{1}{2}$ and $r = n/2$, where n is a positive integer. We obtain a one-parameter family of distributions with pdf

$$f(z) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1}e^{-z/2}, \qquad z > 0. \tag{9.19}$$

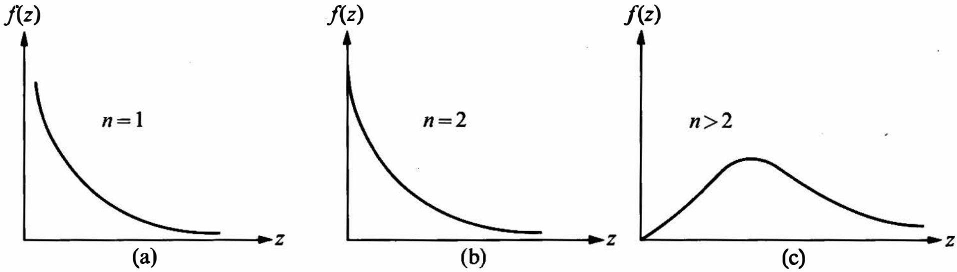

FIGURE 9.7

A random variable Z having pdf given by Eq. (9.19) is said to have a *chi-square distribution with n degrees of freedom* (denoted by χ_n^2). In Fig. 9.7, the pdf for $n = 1, 2$ and $n > 2$ is shown. It is an immediate consequence of Eq. (9.18) that if Z has pdf Eq. (9.19), we have

$$E(Z) = n, \quad V(Z) = 2n. \tag{9.20}$$

The chi-square distribution has many important applications in statistical inference, some of which we shall refer to later. Because of its importance, the chi-square distribution is tabulated for various values of the parameter n. (See Appendix.) Thus we may find in the table that value, denoted by χ_α^2, satisfying $P(Z \leq \chi_\alpha^2) = \alpha$, $0 < \alpha < 1$ (Fig. 9.8). Example 9.9 deals with a special case of

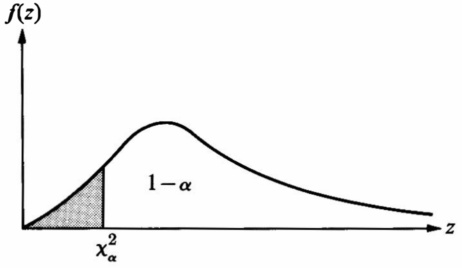

FIGURE 9.8

a general characterization of the chi-square distribution which we shall study in a later chapter.

EXAMPLE 9.9. Suppose that the velocity V of an object has distribution $N(0, 1)$ Let $K = mV^2/2$ be the kinetic energy of the object. To find the pdf of K, let us first find the pdf of $S = V^2$. Applying Theorem 5.2 directly we have

$$g(s) = \frac{1}{2\sqrt{s}}[\varphi(\sqrt{s}) + \varphi(-\sqrt{s})]$$
$$= s^{-1/2}\frac{1}{\sqrt{2\pi}}e^{-s/2}.$$

If we compare this with Eq. (9.19) and recall that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, we note that S has a χ_1^2-distribution. Thus we find that *the square of a random variable with distribution* $N(0, 1)$ *has a* χ_1^2*-distribution*. (It is this result which we shall generalize later.)

We can now obtain the pdf h of the kinetic energy K. Since K is a monotone function of V^2 whose pdf is given by g above, we have directly

$$h(k) = \frac{2}{m}g\left(\frac{2}{m}k\right) = \frac{2}{m}\frac{1}{\sqrt{2\pi}}\left(\frac{2}{m}k\right)^{-1/2}e^{-k/m}, \qquad k > 0.$$

In order to evaluate $P(K \le 5)$ for example, we need not use the pdf of K but may simply use the tabulated chi-square distribution as follows.

$$P(K \le 5) = P(m/2)V^2 \le 5) = P(V^2 \le 10/m).$$

This latter probability can be obtained directly from the tables of the chi-square distribution (if m is known) since V^2 has a χ_1^2-distribution. Since $E(V^2) = 1$ and variance $(V^2) = 2$ [see Eq. (9.20)], we find directly

$$E(K) = m/2 \qquad \text{and} \qquad V(K) = m^2/2.$$

Note: The tabulation of the chi-square distribution as given in the Appendix only tabulates those values for which n, the number of degrees of freedom, is less than or equal to 45. The reason for this is that if n is large, we may approximate the chi-square distribution with the normal distribution, as indicated by the following theorem.

Theorem 9.2. Suppose that the random variable Y has distribution χ_n^2. Then for sufficiently large n the random variable $\sqrt{2Y}$ has approximately the distribution $N(\sqrt{2n-1}, 1)$. (The proof is not given here.)

This theorem may be used as follows. Suppose that we require $P(Y \leq t)$, where Y has distribution χ_n^2, and n is so large that the above probability cannot be directly obtained from the table of the chi-square distribution. Using Theorem 9.2 we may write,

$$\begin{aligned} P(Y \leq t) &= P(\sqrt{2Y} \leq \sqrt{2t}) \\ &= P(\sqrt{2Y} - \sqrt{2n-1} \leq \sqrt{2t} - \sqrt{2n-1}) \\ &\simeq \Phi(\sqrt{2t} - \sqrt{2n-1}). \end{aligned}$$

The value of Φ may be obtained from the tables of the normal distribution.

9.10 Comparisons among Various Distributions

We have by now introduced a number of important probability distributions, both discrete and continuous: the binomial, Pascal, and Poisson among the discrete ones, and the exponential, geometric, and gamma among the continuous ones. We shall not restate the various assumptions which led to these distributions. Our principal concern here is to point out certain similarities (and differences) among the random variables having these distributions.

1. Assume that independent Bernoulli trials are being performed.

   (a) random variable: number of occurrences of event A in a *fixed* number of trials
   distribution: binomial

   (b) random variable: number of Bernoulli trials required to obtain first occurrence of A
   distribution: geometric

   (c) random variable: number of Bernoulli trials required to obtain rth occurrence of A
   distribution: Pascal

2. Assume a Poisson process (see note (c) preceeding Example 8.5).

   (d) random variable: number of occurrences of event A during a *fixed* time interval
   distribution: Poisson

   (e) random variable: time required until first occurrence of A
   distribution: exponential

(f) random variable: time required until rth occurrence of A
distribution: gamma

Note: Observe the similarity between (a) and (d), (b) and (e), and finally (c) and (f).

9.11 The Bivariate Normal Distribution

All the continuous random variables we have discussed have been one-dimensional random variables. As we mentioned in Chapter 6, higher-dimensional random variables play an important role in describing experimental outcomes. One of the most important continuous two-dimensional random variables, a direct generalization of the one-dimensional normal distribution, is defined as follows.

Definition. Let (X, Y) be a two-dimensional, continuous random variable assuming all values in the euclidean plane. We say that (X, Y) has a *bivariate normal distribution* if its joint pdf is given by the following expression.

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$$
$$\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\},$$
$$-\infty < x < \infty, \quad -\infty < y < \infty. \qquad (9.21)$$

The above pdf depends on 5 parameters. For f to define a legitimate pdf [that is, $f(x, y) \geq 0$, $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\,dx\,dy = 1$], we must place the following restrictions on the parameters: $-\infty < \mu_x < \infty$; $-\infty < \mu_y < \infty$; $\sigma_x > 0$; $\sigma_y > 0$; $-1 < \rho < 1$. The following properties of the bivariate normal distribution may easily be checked.

Theorem 9.3. Suppose that (X, Y) has pdf as given by Eq. (9.21). Then

(a) the marginal distributions of X and of Y are $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$, respectively;

(b) the parameter ρ appearing above is the correlation coefficient between X and Y;

(c) the conditional distributions of X (given that $Y = y$) and of Y (given that $X = x$) are respectively

$$N\left[\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y), \sigma_x^2(1-\rho^2)\right], \qquad N\left[\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1-\rho^2)\right].$$

Proof: See Problem 9.21.

Notes: (a) The converse of (a) of Theorem 9.3 is not true. It is possible to have a joint pdf which is not bivariate normal and yet the marginal pdf's of X and of Y are one-dimensional normal.

(b) We observe from Eq. (9.21) that if $\rho = 0$, the joint pdf of (X, Y) may be factored and hence X and Y are independent. Thus we find that in the case of a bivariate normal distribution, zero correlation and independence *are* equivalent.

(c) Statement (c) of the above theorem shows that both regression functions of the mean are *linear*. It also shows that the variance of the conditional distribution is reduced in the same proportion as $(1 - \rho^2)$. That is, if ρ is close to zero, the conditional variance is essentially the same as the unconditional variance, while if ρ is close to ± 1, the conditional variance is close to zero.

The bivariate normal pdf has a number of interesting properties. We shall state some of these as a theorem, leaving the proof to the reader.

Theorem 9.4. Consider the surface $z = f(x, y)$, where f is the bivariate normal pdf given by Eq. (9.3).

(a) $z = c$ (const) cuts the surface in an *ellipse*. (These are sometimes called contours of constant probability density.)

(b) If $\rho = 0$ and $\sigma_x = \sigma_y$, the above ellipse becomes a circle. (What happens to the above ellipse as $\rho \to \pm 1$?)

Proof: See Problem 9.22.

Note: Because of the importance of the bivariate normal distribution, various probabilities associated with it have been tabulated. (See D. B. Owen, *Handbook of Statistical Tables*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962.)

9.12 Truncated Distributions

EXAMPLE 9.10. Suppose that a certain type of bolt is manufactured and its length, say Y, is a random variable with distribution $N(2.2, 0.01)$. From a large lot of such bolts a new lot is obtained by discarding all those bolts for which $Y > 2$. Hence if X is the random variable representing the length of the bolts in the new lot, and if F is its cdf we have

$$\begin{aligned} F(x) &= P(X \le x) \\ &= P(Y \le x \mid Y \le 2) = 1 \qquad \text{if } x > 2, \\ &= P(Y \le x)/P(Y \le 2) \qquad \text{if } x \le 2. \end{aligned}$$

(See Fig. 9.9.) Thus f, the pdf of X is given by

$$\begin{aligned} f(x) &= F'(x) = 0 \qquad \text{if } x > 2, \\ &= \frac{\dfrac{1}{\sqrt{2\pi}\,(0.1)} \exp\left(-\dfrac{1}{2}\left[\dfrac{x - 2.2}{0.1}\right]^2\right)}{\Phi(-2)} \qquad \text{if } x \le 2 \end{aligned}$$

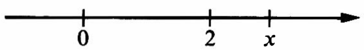

FIGURE 9.9

since

$$P(Y \leq 2) = P\left(\frac{Y - 2.2}{0.1} \leq \frac{2 - 2.2}{0.1}\right) = \Phi(-2).$$

[Φ, as usual, is the cdf of the distribution $N(0, 1)$.]

The above is an illustration of a *truncated normal distribution* (specifically, truncated to the right at $X = 2$). This example may be generalized as follows.

Definition. We say that the random variable X has a normal distribution *truncated to the right* at $X = \tau$ if its pdf f is of the form

$$\begin{aligned} f(x) &= 0 \quad \text{if} \quad x > \tau, \\ &= K \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left[\frac{x - \mu}{\sigma}\right]^2\right) \quad \text{if} \quad x \leq \tau. \end{aligned} \tag{9.22}$$

We note that K is determined from the condition $\int_{-\infty}^{+\infty} f(x)\, dx = 1$ and hence

$$K = \frac{1}{\Phi[(\tau - \mu)/\sigma]} = \frac{1}{P(Z \leq \tau)}$$

where Z has distribution $N(\mu, \sigma^2)$. Analogously to the above we have the following definition.

Definition. We say that the random variable X has a normal distribution *truncated to the left* at $X = \gamma$, if its pdf f is of the form

$$\begin{aligned} f(x) &= 0 \quad \text{if} \quad x < \gamma, \\ &= \frac{K}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left[\frac{x - \mu}{\sigma}\right]^2\right) \quad \text{if} \quad x \geq \gamma. \end{aligned} \tag{9.23}$$

Again, K is determined from the condition $\int_{-\infty}^{+\infty} f(x)\, dx = 1$ and thus

$$K = \left[1 - \Phi\left(\frac{\gamma - \mu}{\sigma}\right)\right]^{-1}.$$

The concepts introduced above for the normal distribution may be extended in an obvious way to other distributions. For example, an exponentially distributed random variable X, truncated to the left at $X = \gamma$, would have the following pdf:

$$\begin{aligned} f(x) &= 0 \quad \text{if} \quad x < \gamma, \\ &= C\alpha e^{-\alpha x} \quad \text{if} \quad x \geq \gamma. \end{aligned} \tag{9.24}$$

Again, C is determined from the condition $\int_{-\infty}^{+\infty} f(x)\, dx = 1$ and hence

$$C = e^{\alpha\gamma}.$$

We can consider also a truncated random variable in the discrete case. For instance, if a Poisson-distributed random variable X (with parameter λ) is truncated to the right at $X = k + 1$, it means that X has the following distribution:

$$P(X = i) = 0 \quad \text{if} \quad i \geq k + 1,$$
$$= C\frac{\lambda^i}{i!}e^{-\lambda} \quad \text{if} \quad i = 0, 1, \ldots, k. \tag{9.25}$$

We determine C from the condition $\sum_{i=0}^{\infty} P(X = i) = 1$ and find

$$C = \frac{1}{\sum_{j=0}^{k} (\lambda^j/j!)e^{-\lambda}}.$$

Thus

$$P(X = i) = \frac{\lambda^i}{i!}\frac{1}{\sum_{j=0}^{k} (\lambda^j/j!)}, \qquad i = 0, 1, \ldots, k \quad \text{and} \quad 0 \text{ elsewhere.}$$

Truncated distributions may arise in many important applications. We shall consider a few examples below.

EXAMPLE 9.11. Suppose that X represents the life length of a component. If X is normally distributed with

$$E(X) = 4 \quad \text{and} \quad V(X) = 4,$$

we find that

$$P(X < 0) = \Phi(-2) = 0.023.$$

Thus this model is not very meaningful since it assigns probability 0.023 to an event which we know *cannot* occur. We might consider, instead, the above random variable X truncated to the left at $X = 0$. Hence we shall suppose that the pdf of the random variable X is given by

$$f(x) = 0 \quad \text{if} \quad x \leq 0,$$
$$= \frac{1}{\sqrt{(2\pi)}\,(2)} \exp\left[-\frac{1}{2}\left(\frac{x-4}{2}\right)^2\right]\frac{1}{\Phi(2)} \quad \text{if} \quad x > 0.$$

Note: We have indicated that we often use the normal distribution to represent a random variable X about which we *know* that it cannot assume negative values. (For instance, time to failure, the length of a rod, etc.) For certain parameter values $\mu = E(X)$ and $\sigma^2 = V(X)$ the value of $P(X < 0)$ will be negligible. However, if this is not the case (as in Example 9.11) we should consider using the normal distribution truncated to the left at $X = 0$.

EXAMPLE 9.12. Suppose that a system is made up of n components which function independently, each having the same probability p of functioning properly. Whenever the system malfunctions, it is inspected in order to discover which and how many components are faulty. Let the random variable X be defined as the number of components that are found to be faulty in a system which has broken down. If we suppose that the system fails if and only if at least one component fails, then X has a *binomial distribution truncated to the left* at $X = 0$. For the very fact that the system *has* failed precludes the possibility that $X = 0$. Specifically we have

$$P(X = k) = \frac{\binom{n}{k}(1-p)^k p^{n-k}}{P\text{ (system fails)}}, \qquad k = 1, 2, \ldots, n.$$

Since P (system fails) $= 1 - p^n$, we may write

$$P(X = k) = \frac{\binom{n}{k}(1-p)^k p^{n-k}}{1 - p^n}, \qquad k = 1, 2, \ldots, n.$$

EXAMPLE 9.13. Suppose that particles are emitted from a radioactive source according to the Poisson distribution with parameter λ. A counting device, recording these emissions only functions if fewer than three particles arrive. (That is, if more than three particles arrive during a specified time period, the device ceases to function because of some "locking" that takes place.) Hence if Y is the number of particles recorded during the specified time interval, Y has possible values 0, 1, and 2. Thus

$$\begin{aligned} P(Y = k) &= \frac{e^{-\lambda}}{k!}\,\frac{\lambda^k}{e^{-\lambda}[1 + \lambda + (\lambda^2/2)]}, \qquad k = 0, 1, 2, \\ &= 0, \qquad \text{elsewhere.} \end{aligned}$$

Since the truncated normal distribution is particularly important, let us consider the following problem associated with this distribution.

Suppose that X is a normally distributed random variable truncated to the right at $X = \tau$. Hence the pdf f is of the form

$$\begin{aligned} f(x) &= 0 \qquad \text{if } x \geq \tau, \\ &= \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\frac{1}{\Phi[(\tau-\mu)/\sigma]} \qquad \text{if } x \leq \tau. \end{aligned}$$

We therefore have

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x\, f(x)\, dx = \frac{1}{\Phi[(\tau-\mu)/\sigma]}\int_{-\infty}^{\tau}\frac{x}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]dx \\ &= \frac{1}{\Phi[(\tau-\mu)/\sigma]}\,\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(\tau-\mu)/\sigma}(s\sigma + \mu)e^{-s^2/2}\,ds \\ &= \frac{1}{\Phi[(\tau-\mu)/\sigma]}\left[\mu\Phi\left(\frac{\tau-\mu}{\sigma}\right) + \sigma\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(\tau-\mu)/\sigma} se^{-s^2/2}\,ds\right] \\ &= \mu + \frac{\sigma}{\Phi[(\tau-\mu)/\sigma]}\,\frac{1}{\sqrt{2\pi}}\,e^{-s^2/2}(-1)\Big|_{-\infty}^{(\tau-\mu)/\sigma} \\ &= \mu - \frac{\sigma}{\Phi[(\tau-\mu)/\sigma]}\,\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\tau-\mu}{\sigma}\right)^2\right]. \end{aligned}$$

Note that the expression obtained for $E(X)$ is expressed in terms of tabulated functions. The function Φ is of course the usual cdf of the distribution $N(0, 1)$, while $(1/\sqrt{2\pi})e^{-x^2/2}$ is the ordinate of the pdf of the $N(0, 1)$ distribution and is also tabulated. In fact the quotient

$$\frac{(1/\sqrt{2\pi})e^{-x^2/2}}{\Phi(x)}$$

is tabulated. (See D. B. Owen, *Handbook of Statistical Tables*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962.)

Using the above result, we may now ask the following question: For given μ and σ where should the truncation occur (that is, what should the value of τ be) so that the expected value *after* truncation has some preassigned value, say A? We can answer this question with the aid of the tabulated normal distribution. Suppose that $\mu = 10, \sigma = 1$, and we require that $A = 9.5$. Hence we must solve

$$9.5 = 10 - \frac{1}{\Phi(\tau - 10)} \frac{1}{\sqrt{2\pi}} e^{-(\tau-10)^2/2}.$$

This becomes

$$\frac{1}{2} = \frac{(1/\sqrt{2\pi})e^{-(\tau-10)^2/2}}{\Phi(\tau - 10)}.$$

Using the tables referred to above, we find that $\tau - 10 = 0.52$. and hence $\tau = 10.52$.

Note: The problem raised above may be solved only for certain values of μ, σ, and A. That is, for given μ and σ, it may not be possible to obtain a specified value of A. Consider the equation which must be solved:

$$\mu - A = \frac{\sigma}{\Phi[(\tau - \mu)/\sigma]} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\tau - \mu}{\sigma}\right)^2\right].$$

The right-hand side of this equation is obviously positive. Hence we must have $(\mu - A) > 0$ in order for the above problem to have a solution. This condition is not very surprising since it says simply that the expected value (after truncation on the *right*) must be less than the original expected value.

PROBLEMS

9.1. Suppose that X has distribution $N(2, 0.16)$. Using the table of the normal distribution, evaluate the following probabilities.

(a) $P(X \geq 2.3)$ (b) $P(1.8 \leq X \leq 2.1)$

9.2. The diameter of an electric cable is normally distributed with mean 0.8 and variance 0.0004. What is the probability that the diameter will exceed 0.81 inch?

9.3. Suppose that the cable in Problem 9.2 is considered defective if the diameter differs from its mean by more than 0.025. What is the probability of obtaining a defective cable?

9.4. The errors in a certain length-measuring device are known to be normally distributed with expected value zero and standard deviation 1 inch. What is the probability that the error in measurement will be greater than 1 inch? 2 inches? 3 inches?

9.5. Suppose that the life lengths of two electronic devices, say D_1 and D_2, have distributions $N(40, 36)$ and $N(45, 9)$, respectively. If the electronic device is to be used for a 45-hour period, which device is to be preferred? If it is to be used for a 48-hour period, which device is to be preferred?

9.6. We may be interested only in the magnitude of X, say $Y = |X|$. If X has distribution $N(0, 1)$, determine the pdf of Y, and evaluate $E(Y)$ and $V(Y)$.

9.7. Suppose that we are measuring the position of an object in the plane. Let X and Y be the errors of measurement of the x- and y-coordinates, respectively. Assume that X and Y are independently and identically distributed, each with distribution $N(0, \sigma^2)$. Find the pdf of $R = \sqrt{X^2 + Y^2}$. (The distribution of R is known as the *Rayleigh distribution*.) [*Hint:* Let $X = R \cos \psi$ and $Y = R \sin \psi$. Obtain the joint pdf of (R, ψ) and then obtain the marginal pdf of R.]

9.8. Find the pdf of the random variable $Q = X/Y$, where X and Y are distributed as in Problem 9.7. (The distribution of Q is known as the *Cauchy* distribution.) Can you compute $E(Q)$?

9.9. A distribution closely related to the normal distribution is the *lognormal distribution*. Suppose that X is normally distributed with mean μ and variance σ^2. Let $Y = e^X$. Then Y has the lognormal distribution. (That is, Y is lognormal if and only if $\ln Y$ is normal.) Find the pdf of Y. *Note:* The following random variables may be represented by the above distribution: the diameter of small particles after a crushing process, the size of an organism subject to a number of small impulses, and the life length of certain items.

9.10. Suppose that X has distribution $N(\mu, \sigma^2)$. Determine c (as a function of μ and σ) such that $P(X \leq c) = 2P(X > c)$.

9.11. Suppose that temperature (measured in degrees centigrade) is normally distributed with expectation 50° and variance 4. What is the probability that the temperature T will be between 48° and 53° centigrade?

9.12. The outside diameter of a shaft, say D, is specified to be 4 inches. Consider D to be a normally distributed random variable with mean 4 inches and variance 0.01 inch^2. If the actual diameter differs from the specified value by more than 0.05 inch but less than 0.08 inch, the loss to the manufacturer is \$0.50. If the actual diameter differs from the specified diameter by more than 0.08 inch, the loss is \$1.00. The loss, L, may be considered as a random variable. Find the probability distribution of L and evaluate $E(L)$.

9.13. Compare the *upper bound* on the probability $P[|X - E(X)| \geq 2\sqrt{V(X)}]$ obtained from Chebyshev's inequality with the exact probability in each of the following cases.

(a) X has distribution $N(\mu, \sigma^2)$.

(b) X has Poisson distribution with parameter λ.
(c) X has exponential distribution with parameter α.

9.14. Suppose that X is a random variable for which $E(X) = \mu$ and $V(X) = \sigma^2$. Suppose that Y is uniformly distributed over the interval (a, b). Determine a and b so that $E(X) = E(Y)$ and $V(X) = V(Y)$.

9.15. Suppose that X, the breaking strength of rope (in pounds), has distribution $N(100, 16)$. Each 100-foot coil of rope brings a profit of \$25, provided $X > 95$. If $X \leq 95$, the rope may be used for a different purpose and a profit of \$10 per coil is realized. Find the expected profit per coil.

9.16. Let X_1 and X_2 be independent random variables each having distribution $N(\mu, \sigma^2)$. Let $Z(t) = X_1 \cos \omega t + X_2 \sin \omega t$. This random variable is of interest in the study of random signals. Let $V(t) = dZ(t)/dt$. (ω is assumed to be constant.)

(a) What is the probability distribution of $Z(t)$ and $V(t)$ for any fixed t?
(b) Show that $Z(t)$ and $V(t)$ are uncorrelated. [*Note:* One can actually show that $Z(t)$ and $V(t)$ are independent but this is somewhat more difficult to do.]

9.17. A rocket fuel is to contain a certain percent (say X) of a particular compound. The specifications call for X to be between 30 and 35 percent. The manufacturer will make a net profit on the fuel (per gallon) which is the following function of X:

$$\begin{aligned} T(X) &= \$0.10 \text{ per gallon} \quad \text{if } 30 < X < 35, \\ &= \$0.05 \text{ per gallon} \quad \text{if } 35 \leq X < 40 \text{ or } 25 < X \leq 30, \\ &= -\$0.10 \text{ per gallon otherwise.} \end{aligned}$$

(a) If X has distribution $N(33, 9)$, evaluate $E(T)$.
(b) Suppose that the manufacturer wants to increase his expected profit, $E(T)$, by 50 percent. He intends to do this by increasing his profit (per gallon) on those batches of fuel meeting the specifications, $30 < X < 35$. What must his new net profit be?

9.18. Consider Example 9.8. Suppose that the operator is paid C_3 dollars/hour while the machine is operating and C_4 dollars/hour ($C_4 < C_3$) for the remaining time he has been hired after the machine has failed. Again determine for what value of H (the number of hours the operator is being hired), the expected profit is maximized.

9.19. Show that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. (See 9.15.) [*Hint:* Make the change of variable $x = u^2/2$ in the integral $\Gamma(\frac{1}{2}) = \int_0^\infty x^{-1/2}e^{-x}\,dx$.]

9.20. Verify the expressions for $E(X)$ and $V(X)$, where X has a Gamma distribution [see Eq. (9.18)].

9.21. Prove Theorem 9.3.

9.22. Prove Theorem 9.4.

9.23. Suppose that the random variable X has a chi-square distribution with 10 degrees of freedom. If we are asked to find two numbers a and b such that $P(a < x < b) = 0.85$, say, we should realize that there are many pairs of this kind.

(a) Find two different sets of values (a, b) satisfying the above condition.
(b) Suppose that in addition to the above, we require that

$$P(X < a) = P(X > b).$$

How many sets of values are there?

9.24. Suppose that V, the velocity (cm/sec) of an object having a mass of 1 kg, is a random variable having distribution $N(0, 25)$. Let $K = 1000V^2/2 = 500V^2$ represent the kinetic energy (KE) of the object. Evaluate $P(K < 200)$, $P(K > 800)$.

9.25. Suppose that X has distribution $N(\mu, \sigma^2)$. Using Theorem 7.7, obtain an approximation expression for $E(Y)$ and $V(Y)$ if $Y = \ln X$.

9.26. Suppose that X has a normal distribution truncated to the right as given by Eq. (9.22). Find an expression for $E(X)$ in terms of tabulated functions.

9.27. Suppose that X has an exponential distribution truncated to the left as given by Eq. (9.24). Obtain $E(X)$.

9.28. (a) Find the probability distribution of a binomially distributed random variable (based on n repetitions of an experiment) truncated to the right at $X = n$; that is, $X = n$ cannot be observed.

(b) Find the expected value and variance of the random variable described in (a).

9.29. Suppose that a normally distributed random variable with expected value μ and variance σ^2 is truncated to the left at $X = \tau$ and to the right at $X = \gamma$. Find the pdf of this "doubly truncated" random variable.

9.30. Suppose that X, the length of a rod, has distribution $N(10, 2)$. Instead of measuring the value of X, it is only specified whether certain requirements are met. Specifically, each manufactured rod is classified as follows: $X < 8$, $8 \leq X < 12$, and $X \geq 12$. If 15 such rods are manufactured, what is the probability that an equal number of rods fall into each of the above categories?

9.31. The annual rainfall at a certain locality is known to be a normally distributed random variable with mean value equal to 29.5 inches and standard deviation 2.5 inches. How many inches of rain (annually) is exceeded about 5 percent of the time?

9.32. Suppose that X has distribution $N(0, 25)$. Evaluate $P(1 < X^2 < 4)$.

9.33. Let X_t be the number of particles emitted in t hours from a radioactive source and suppose that X_t has a Poisson distribution with parameter βt. Let T equal the number of hours until the first emission. Show that T has an exponential distribution with parameter β. [*Hint:* Find the equivalent event (in terms of X_t) to the event $T > t$.]

9.34. Suppose that X_t is defined as in Problem 9.33 with $\beta = 30$. What is the probability that the time between successive emissions will be >5 minutes? >10 minutes? < 30 seconds?

9.35. In some tables for the normal distribution, $H(x) = (1/\sqrt{2\pi})\int_0^x e^{-t^2/2}\,dt$ is tabulated for positive values of x (instead of $\Phi(x)$ as given in the Appendix). If the random variable X has distribution $N(1, 4)$ express each of the following probabilities in terms of *tabulated* values of the function H.

(a) $P[|X| > 2]$ (b) $P[X < 0]$

9.36. Suppose that a satellite telemetering device receives two kinds of signals which may be recorded as real numbers, say X and Y. Assume that X and Y are independent, continuous random variables with pdf's f and g, respectively. Suppose that during any specified period of time only one of these signals may be received and hence transmitted back to earth, namely that signal which arrives first. Assume furthermore that the signal giving rise to the value of X arrives first with probability p and hence the signal giving rise to Y arrives first with probability $1 - p$. Let Z denote the random variable whose value is actually received and transmitted.

(a) Express the pdf of Z in terms of f and g.

(b) Express $E(Z)$ in terms of $E(X)$ and $E(Y)$.

(c) Express $V(Z)$ in terms of $V(X)$ and $V(Y)$.

(d) Suppose that X has distribution $N(2, 4)$ and that Y has distribution $N(3, 3)$. If $p = \frac{2}{3}$, evaluate $P(Z > 2)$.

(e) Suppose that X and Y have distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. Show that if $\mu_1 = \mu_2$, the distribution of Z is "uni-modal," that is, the pdf of Z has a unique relative maximum.

9.37. Assume that the number of accidents in a factory may be represented by a Poisson process averaging 2 accidents per week. What is the probability that (a) the time from one accident to the next will be more than 3 days, (b) the time from one accident to the third accident will be more than a week? [*Hint:* In (a), let T = time (in days) and compute $P(T > 3)$.]

9.38. On the average a production process produces one defective item among every 300 manufactured. What is the probability that the *third* defective item will appear:

(a) before 1000 pieces have been produced?

(b) as the 1000th piece is produced?

(c) after the 1000th piece has been produced?

[*Hint:* Assume a Poisson process.]