- If empirical distributions are used in the *usual* way (see Sec. 6.2.4), it is not possible to generate values outside the range of the observed data in the simulation (see Sec. 8.3.12). This is unfortunate, since many measures of performance for simulated systems depend heavily on the probability of an "extreme" event's occurring, e.g., generation of a very large service time. With a fitted theoretical distribution, on the other hand, values outside the range of the observed data can be generated.
- There may be a compelling physical reason in some situations for using a certain theoretical distribution form as a model for a particular input random variable. Even when we are fortunate enough to have this kind of information, it is a good idea to use observed data to provide empirical support for the use of this particular distribution.
- A theoretical distribution is a compact way of representing a set of data values. Conversely, if n data values are available from a continuous distribution, then $2n$ values (data and corresponding cumulative probabilities) must be entered and stored in the computer to represent an empirical distribution in many simulation languages. Thus, use of an empirical distribution will be cumbersome if the data set is large.

There are definitely situations for which no theoretical distribution will provide an adequate fit for the observed data. In these cases we recommend using an empirical distribution. It should also be mentioned that a fourth approach for using observed data to specify a distribution has been proposed by several authors [see, for example, Swain, Venkatraman, and Wilson (1988)]. This approach involves using a general four-parameter family of distributions (e.g., the Johnson translation system) to model all sources of system randomness.

The remainder of this chapter discusses various topics related to the selection of input distributions. Section 6.2 discusses how theoretical distributions are parameterized, provides a compendium of relevant facts on most of the commonly used continuous and discrete distributions, and discusses how empirical distributions can be specified. In Sec. 6.3 we present techniques for determining whether the data are independent observations from some underlying distribution, which is a requirement of many of the statistical procedures in this chapter. Sections 6.4 through 6.6 discuss the three basic activities in specifying a theoretical distribution on the basis of observed data; a comprehensive example illustrating these methods is given in Sec. 6.7. It should be mentioned, however, that choosing a distribution is not necessarily a strictly sequential procedure. Based on the results from one activity, it may be necessary to go back to an earlier activity. In Sec. 6.8 we discuss how certain of the theoretical continuous distributions, e.g., gamma, Weibull, and lognormal, can be "shifted" away from 0 to make them better fit our observed data in some cases. Section 6.9 discusses possible methods for specifying input distributions when no data are available. Several useful probabilistic models for describing the manner in which "customers" arrive to a system are given in Sec. 6.10, while Sec. 6.11 presents techniques for determining whether observations from different sources can be "pooled."

Most of the graphical plots and goodness-of-fit tests presented in this chapter were developed using the UniFit II distribution-fitting package [see Law and Vincent (1990)].

6.2 USEFUL PROBABILITY DISTRIBUTIONS

The purpose of this section is to discuss a variety of distributions that have been found useful in simulation modeling and to provide a unified listing of relevant properties of these distributions [see also Hastings and Peacock (1975)]. Section 6.2.1 provides a short discussion of common methods by which continuous distributions are defined, or parameterized. Then, Secs. 6.2.2 and 6.2.3 contain compilations of several continuous and discrete distributions. Finally, Sec. 6.2.4 suggests how the data themselves can be used directly to define an empirical distribution.

6.2.1 Parameterization of Continuous Distributions

For a given family of continuous distributions, e.g., normal or gamma, there are usually several alternative ways to define, or *parameterize*, the probability density function. However, if the parameters are defined correctly, they can be classified, on the basis of their physical or geometric interpretation, as being one of three basic types: location, scale, or shape parameters.

A *location parameter* γ specifies an abscissa (x axis) location point of a distribution's range of values; usually γ is the midpoint (e.g., the mean μ for a normal distribution) or lower endpoint (see Sec. 6.8) of the distribution's range. (In the latter case, location parameters are sometimes called *shift parameters*.) As γ changes, the associated distribution merely shifts left or right without otherwise changing. A *scale parameter* β determines the scale (or unit) of measurement of the values in the range of the distribution. A change in β compresses or expands the associated distribution without altering its basic form. A *shape parameter* α determines, distinct from location and scale, the basic form or shape of a distribution within the general family of distributions of interest. A change in α generally alters a distribution's properties (e.g., skewness) more fundamentally than a change in location or scale. Some distributions, e.g., exponential and normal, do not have a shape parameter, while others may have several (the beta distribution has two).

6.2.2 Continuous Distributions

Table 6.3 gives information relevant to simulation modeling applications for 10 continuous distributions. Possible applications are given first to indicate some (certainly not all) uses of the distribution [see Hahn and Shapiro (1967) and

TABLE 6.3
Continuous distributions

Uniform	**U(*a*,*b*)**
Possible applications	Used as a "first" model for a quantity that is felt to be randomly varying between a and b but about which little else is known. The U(0,1) distribution is essential in generating random values from all other distributions (see Chaps. 7 and 8)
Density (see Fig. 6.1)	$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$
Distribution	$F(x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \le x \le b \\ 1 & \text{if } b < x \end{cases}$
Parameters	a and b real numbers with $a < b$; a is a location parameter, $b - a$ is a scale parameter
Range	$[a,b]$
Mean	$\dfrac{a+b}{2}$
Variance	$\dfrac{(b-a)^2}{12}$
Mode	Does not uniquely exist
MLE	$\hat{a} = \min_{1 \le i \le n} X_i$, $\hat{b} = \max_{1 \le i \le n} X_i$
Comments	1. The U(0,1) distribution is a special case of the beta distribution (when $\alpha_1 = \alpha_2 = 1$) 2. If $X \sim U(0,1)$ and $[x, x + \Delta x]$ is a subinterval of $[0,1]$ with $\Delta x \ge 0$, $P(X \in [x, x+\Delta x]) = \int_x^{x+\Delta x} 1\,dy = (x + \Delta x) - x = \Delta x$ which justifies the name "uniform"

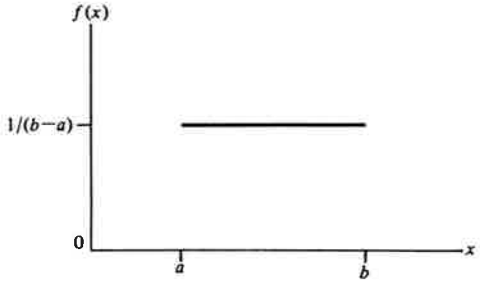

FIGURE 6.1
U(a,b) density function.

Exponential	**expo(β)**
Possible applications	Interarrival times of "customers" to a system that occur at a constant rate
Density (see Fig. 6.2)	$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-x/\beta} & \text{if } x \ge 0 \\ 0 & \text{otherwise} \end{cases}$

TABLE 6.3 *(continued)*

Exponential	**expo(β)**
Distribution	$F(x) = \begin{cases} 1 - e^{-x/\beta} & \text{if } x \ge 0 \\ 0 & \text{otherwise} \end{cases}$
Parameter	Scale parameter $\beta > 0$
Range	$[0,\infty)$
Mean	β
Variance	β^2
Mode	0
MLE	$\hat{\beta} = \bar{X}(n)$
Comments	1. The expo(β) distribution is a special case of both the gamma and Weibull distributions (for shape parameter $\alpha = 1$ and scale parameter β in both cases) 2. If $X_1, X_2, \ldots, X_m$ are independent expo(β) random variables, then $X_1 + X_2 + \cdots + X_m \sim \text{gamma}(m,\beta)$, also called the *m-Erlang distribution* 3. The exponential distribution is the only continuous distribution with the memoryless property (see Prob. 4.26)

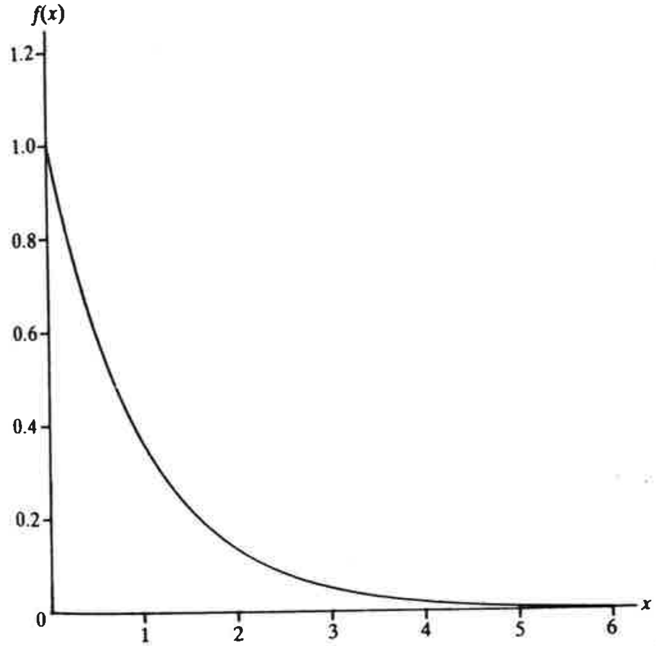

FIGURE 6.2
expo(1) density function.

Gamma	**gamma(α,β)**
Possible applications	Time to complete some task, e.g., customer service or machine repair

TABLE 6.3 *(continued)*

Gamma	gamma(α,β)
Density (see Fig. 6.3)	$f(x)=\begin{cases}\dfrac{\beta^{-\alpha}x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)} & \text{if } x>0\\ 0 & \text{otherwise}\end{cases}$ where $\Gamma(\alpha)$ is the *gamma function*, defined by $\Gamma(z)=\int_0^\infty t^{z-1}e^{-t}\,dt$ for any real number $z>0$. Some properties of the gamma function: $\Gamma(z+1)=z\Gamma(z)$ for any $z>0$, $\Gamma(k+1)=k!$ for any nonnegative integer k, $\Gamma(k+\frac{1}{2})=\sqrt{\pi}\cdot 1\cdot 3\cdot 5\cdots(2k-1)/2^k$ for any positive integer k, $\Gamma(1/2)=\sqrt{\pi}$
Distribution	If α is not an integer, there is no closed form. If α is a positive integer, then $F(x)=\begin{cases}1-e^{-x/\beta}\displaystyle\sum_{j=0}^{\alpha-1}\frac{(x/\beta)^j}{j!} & \text{if } x>0\\ 0 & \text{otherwise}\end{cases}$
Parameters	Shape parameter $\alpha>0$, scale parameter $\beta>0$
Range	$[0,\infty)$
Mean	$\alpha\beta$
Variance	$\alpha\beta^2$
Mode	$\beta(\alpha-1)$ if $\alpha\geq 1$, 0 if $\alpha<1$
MLE	The following two equations must be satisfied: $\ln\hat{\beta}+\Psi(\hat{\alpha})=\dfrac{\sum_{i=1}^{n}\ln X_i}{n}, \qquad \hat{\alpha}\hat{\beta}=\bar{X}(n)$ which could be solved numerically. [$\Psi(\hat{\alpha})=\Gamma'(\hat{\alpha})/\Gamma(\hat{\alpha})$ and is called the *digamma function*; Γ' denotes the derivative of Γ.] Alternatively, approximations to $\hat{\alpha}$ and $\hat{\beta}$ can be obtained by letting $T=[\ln\bar{X}(n)-\sum_{i=1}^{n}\ln X_i/n]^{-1}$, using Table 6.19 (see App. 6A) to obtain $\hat{\alpha}$ as a function of T, and letting $\hat{\beta}=\bar{X}(n)/\hat{\alpha}$. [See Choi and Wette (1969) for the derivation of this procedure and of Table 6.19]
Comments	1. The expo(β) and gamma($1,\beta$) distributions are the same 2. For a positive integer m, the gamma(m,β) distribution is called the m-Erlang(β) distribution 3. The chi-square distribution with k df is the same as the gamma($k/2$, 2) distribution 4. If $X_1, X_2, \ldots, X_m$ are independent random variables with $X_i\sim$ gamma(α_i,β), then $X_1+X_2+\cdots+X_m\sim$ gamma($\alpha_1+\alpha_2+\cdots+\alpha_m, \beta$) 5. If X_1 and X_2 are independent random variables with $X_i\sim$ gamma(α_i,β), then $X_1/(X_1+X_2)\sim$ beta(α_1,α_2) 6. $X\sim$ gamma(α,β) if and only if $Y=1/X$ has a Pearson type V distribution with shape and scale parameters α and $1/\beta$, denoted PT5($\alpha,1/\beta$) 7. $\lim_{x\to 0} f(x)=\begin{cases}\infty & \text{if } \alpha<1\\ \dfrac{1}{\beta} & \text{if } \alpha=1\\ 0 & \text{if } \alpha>1\end{cases}$

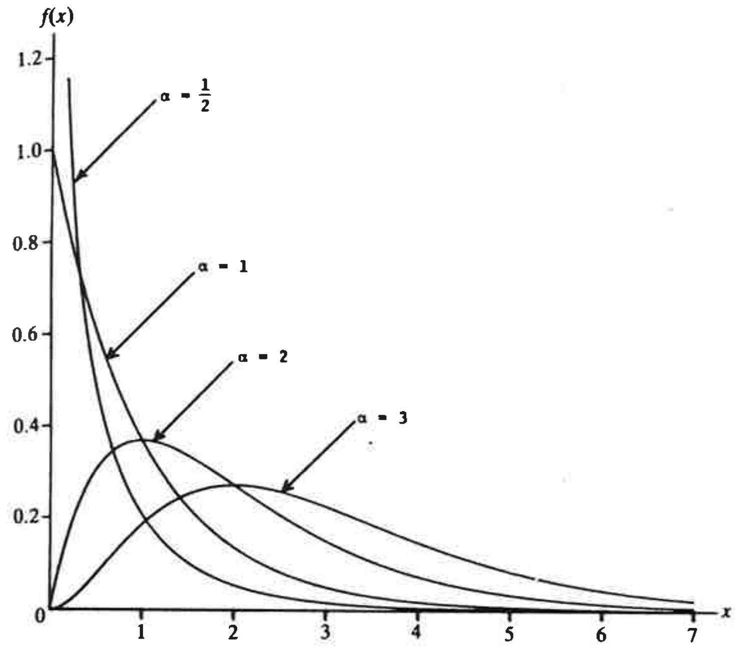

FIGURE 6.3
gamma(α,1) density functions.

TABLE 6.3 *(continued)*

Weibull	Weibull(α,β)
Possible applications	Time to complete some task (density takes on shapes similar to gamma densities), time to failure of a piece of equipment
Density (see Fig. 6.4)	$f(x)=\begin{cases}\alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha} & \text{if } x>0\\ 0 & \text{otherwise}\end{cases}$
Distribution	$F(x)=\begin{cases}1-e^{-(x/\beta)^\alpha} & \text{if } x>0\\ 0 & \text{otherwise}\end{cases}$
Parameters	Shape parameter $\alpha>0$, scale parameter $\beta>0$
Range	$[0,\infty)$
Mean	$\dfrac{\beta}{\alpha}\Gamma\left(\dfrac{1}{\alpha}\right)$
Variance	$\dfrac{\beta^2}{\alpha}\left\{2\Gamma\left(\dfrac{2}{\alpha}\right)-\dfrac{1}{\alpha}\left[\Gamma\left(\dfrac{1}{\alpha}\right)\right]^2\right\}$
Mode	$\begin{cases}\beta\left(\dfrac{\alpha-1}{\alpha}\right)^{1/\alpha} & \text{if } \alpha\geq 1\\ 0 & \text{if } \alpha<1\end{cases}$

TABLE 6.3 *(continued)*

Weibull	**Weibull(α,β)**

MLE

The following two equations must be satisfied:

$$\frac{\sum_{i=1}^{n} X_i^{\hat{\alpha}} \ln X_i}{\sum_{i=1}^{n} X_i^{\hat{\alpha}}} - \frac{1}{\hat{\alpha}} = \frac{\sum_{i=1}^{n} \ln X_i}{n}, \qquad \hat{\beta} = \left(\frac{\sum_{i=1}^{n} X_i^{\hat{\alpha}}}{n} \right)^{1/\hat{\alpha}}$$

The first can be solved for $\hat{\alpha}$ numerically by Newton's method, and the second equation then gives $\hat{\beta}$ directly. The general recursive step for the Newton iterations is

$$\hat{\alpha}_{k+1} = \hat{\alpha}_k + \frac{A + 1/\hat{\alpha}_k - C_k/B_k}{1/\hat{\alpha}_k^2 + (B_k H_k - C_k^2)/B_k^2}$$

where

$$A = \frac{\sum_{i=1}^{n} \ln X_i}{n}, \qquad B_k = \sum_{i=1}^{n} X_i^{\hat{\alpha}_k}, \qquad C_k = \sum_{i=1}^{n} X_i^{\hat{\alpha}_k} \ln X_i$$

and

$$H_k = \sum_{i=1}^{n} X_i^{\hat{\alpha}_k} (\ln X_i)^2$$

[See Thoman, Bain, and Antle (1969) for these formulas, as well as for confidence intervals on the true α and β.] As a starting point for the iterations, the estimate

$$\hat{\alpha}_0 = \left\{ \frac{\frac{6}{\pi^2}\left[\sum_{i=1}^{n} (\ln X_i)^2 - \left(\sum_{i=1}^{n} \ln X_i \right)^2 \Big/ n \right]}{n-1} \right\}^{-1/2}$$

[due to Menon (1963) and suggested in Thoman, Bain, and Antle (1969)] may be used. With this choice of $\hat{\alpha}_0$, it was reported in Thoman, Bain, and Antle (1969) that an average of only 3.5 Newton iterations were needed to achieve four-place accuracy.

Comments

1. The expo(β) and Weibull($1,\beta$) distributions are the same
2. $X \sim \text{Weibull}(\alpha,\beta)$ if and only if $X^{\alpha} \sim \text{expo}(\beta^{\alpha})$ (see Prob. 6.2)
3. The (natural) logarithm of a Weibull random variable has a distribution known as the *extreme-value* or *Gumbel distribution* [see Law and Vincent (1990), Lawless (1982), and Prob. 8.1(*b*)]
4. The Weibull($2, \beta$) distribution is also called a *Rayleigh distribution* with parameter β, denoted Rayleigh(β). If Y and Z are independent normal random variables with mean 0 and variance β^2 (see the normal distribution), then $X = (Y^2 + Z^2)^{1/2} \sim \text{Rayleigh}(2^{1/2}\beta)$
5. As $\alpha \to \infty$, the Weibull distribution becomes degenerate at β. Thus, Weibull densities for large α have a sharp peak at the mode
6. $$\lim_{x \to 0} f(x) = \begin{cases} \infty & \text{if } \alpha < 1 \\ \frac{1}{\beta} & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha > 1 \end{cases}$$

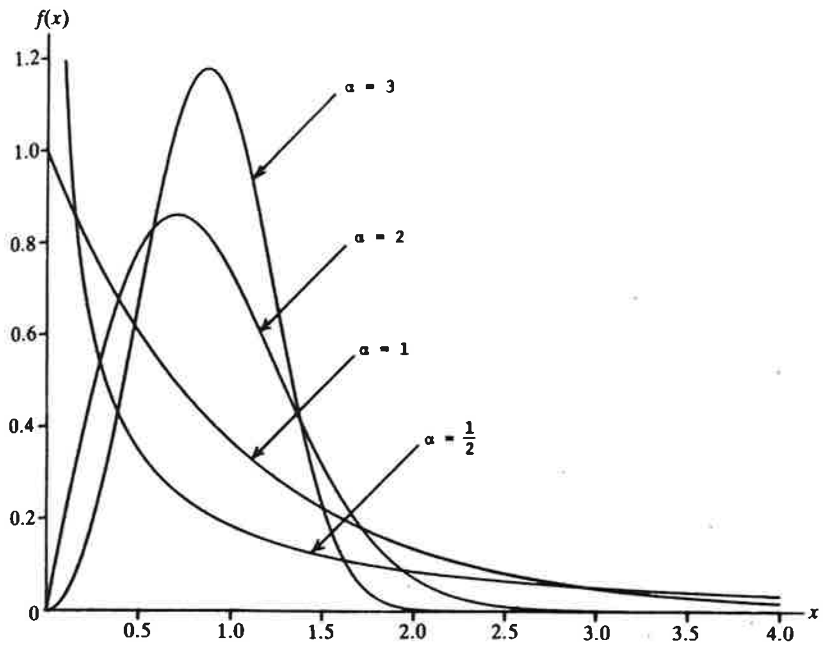

FIGURE 6.4
Weibull(α,1) density functions.

TABLE 6.3 *(continued)*

Normal	**N(μ,σ^2)**
Possible applications	Errors of various types, e.g., in the impact point of a bomb; quantities that are the sum of a large number of other quantities (by virtue of central limit theorems)
Density (see Fig. 6.5)	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$ for all real numbers x
Distribution	No closed form
Parameters	Location parameter $\mu \in (-\infty,\infty)$, scale parameter $\sigma > 0$
Range	$(-\infty,\infty)$
Mean	μ
Variance	σ^2
Mode	μ
MLE	$\hat{\mu} = \bar{X}(n), \quad \hat{\sigma} = \left[\frac{n-1}{n} S^2(n) \right]^{1/2}$
Comments	1. If two jointly distributed normal random variables are uncorrelated, they are also independent. For distributions other than normal, this implication is not true in general

TABLE 6.3 *(continued)*

Normal	N(μ,σ^2)
	2. Suppose that the joint distribution of $X_1, X_2, \ldots, X_m$ is multivariate normal and let $\mu_i = E(X_i)$ and $C_{ij} = \mathrm{Cov}(X_i, X_j)$. Then for any real numbers $a, b_1, b_2, \ldots, b_m$, the random variable $a + b_1X_1 + b_2X_2 + \cdots + b_mX_m$ has a normal distribution with mean $\mu = a + \Sigma_{i=1}^{m} b_i\mu_i$ and variance $$\sigma^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} b_ib_jC_{ij}$$ Note that we need *not* assume independence of the X_i's. If the X_i's *are* independent, then $$\sigma^2 = \sum_{i=1}^{m} b_i^2\,\mathrm{Var}(X_i)$$
	3. The N(0,1) distribution is often called the *standard* or *unit normal distribution*
	4. If $X_1, X_2, \ldots, X_k$ are independent standard normal random variables, then $X_1^2 + X_2^2 + \cdots + X_k^2$ has a chi-square distribution with k df, which is also the gamma($k/2$,2) distribution
	5. If $X \sim \mathrm{N}(\mu,\sigma^2)$, then e^X has the *lognormal distribution* with parameters μ and σ, denoted LN(μ,σ^2)
	6. If $X \sim \mathrm{N}(0,1)$, if Y has a chi-square distribution with k df, and if X and Y are independent, then $X/\sqrt{Y/k}$ has a t distribution with k df (sometimes called *Student's t distribution*)
	7. If the normal distribution is used to represent a nonnegative quantity (e.g., time), then its density should be truncated at $x = 0$ (see Sec. 6.8)
	8. As $\sigma \to 0$, the normal distribution becomes degenerate at μ

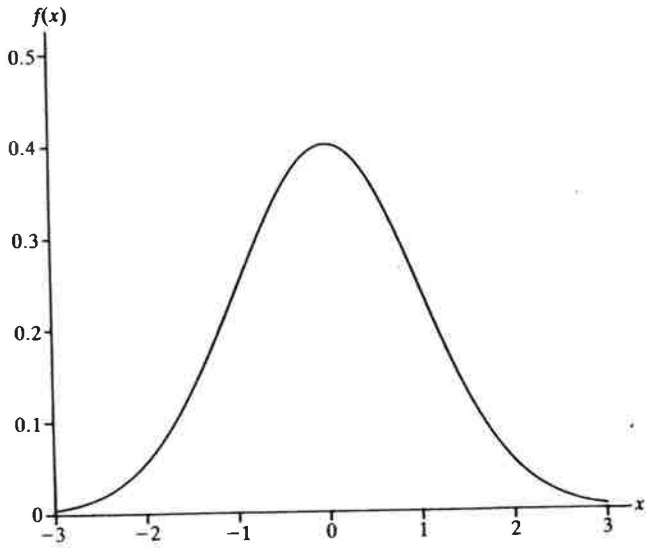

FIGURE 6.5
N(0,1) density function.

TABLE 6.3 *(continued)*

Lognormal	LN(μ,σ^2)
Possible applications	Time to perform some task [density takes on shapes similar to gamma(α,β) and Weibull(α,β) densities for $\alpha > 1$, but can have a large "spike" close to $x = 0$ that is often useful]; quantities that are the product of a large number of other quantities (by virtue of central limit theorems)
Density (see Fig. 6.6)	$f(x) = \begin{cases} \dfrac{1}{x\sqrt{2\pi\sigma^2}} \exp \dfrac{-(\ln x - \mu)^2}{2\sigma^2} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$
Distribution	No closed form
Parameters	Shape parameter $\sigma > 0$, scale parameter $\mu \in (-\infty,\infty)$
Range	$[0,\infty)$
Mean	$e^{\mu+\sigma^2/2}$
Variance	$e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$
Mode	$e^{\mu-\sigma^2}$
MLE	$\hat{\mu} = \dfrac{\sum_{i=1}^{n} \ln X_i}{n}, \quad \hat{\sigma} = \left[\dfrac{\sum_{i=1}^{n} (\ln X_i - \hat{\mu})^2}{n}\right]^{1/2}$
Comments	1. $X \sim \mathrm{LN}(\mu,\sigma^2)$ if and only if $\ln X \sim \mathrm{N}(\mu,\sigma^2)$. Thus, if one has data $X_1, X_2, \ldots, X_n$ that are thought to be lognormal, the logarithms of the data points, $\ln X_1, \ln X_2, \ldots, \ln X_n$, can be treated as normally distributed data for purposes of hypothesizing a distribution, parameter estimation, and goodness-of-fit testing
	2. As $\sigma \to 0$, the lognormal distribution becomes degenerate at e^μ. Thus, lognormal densities for small σ have a sharp peak at the mode
	3. $\lim_{x\to 0} f(x) = 0$, regardless of the parameter values

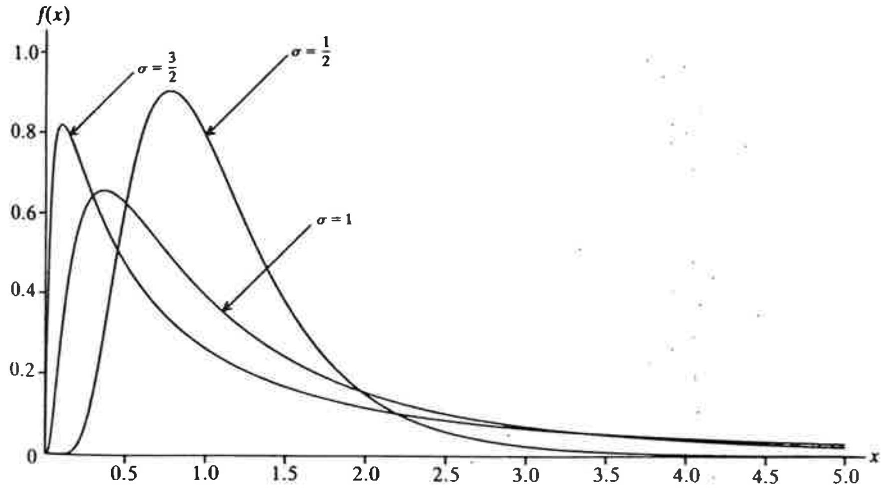

FIGURE 6.6
LN(0,σ^2) density functions.

TABLE 6.3 (*continued*)

Beta	$\text{beta}(\alpha_1,\alpha_2)$
Possible applications	Used as a rough model in the absence of data (see Sec. 6.9); distribution of a random proportion, such as the proportion of defective items in a shipment; time to complete a task, e.g., in a PERT network
Density (see Fig. 6.7)	$f(x) = \begin{cases} \dfrac{x^{\alpha_1-1}(1-x)^{\alpha_2-1}}{B(\alpha_1,\alpha_2)} & \text{if } 0<x<1 \\ 0 & \text{otherwise} \end{cases}$ where $B(\alpha_1,\alpha_2)$ is the *beta function*, defined by $B(z_1,z_2) = \int_0^1 t^{z_1-1}(1-t)^{z_2-1}\,dt$ for any real numbers $z_1>0$ and $z_2>0$. Some properties of the beta function: $B(z_1,z_2)=B(z_2,z_1)\,, \qquad B(z_1,z_2)=\dfrac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1+z_2)}$
Distribution	No closed form, in general. If either α_1 or α_2 is a positive integer, a binomial expansion can be used to obtain $F(x)$, which will be a polynomial in x, and the powers of x will be, in general, positive real numbers ranging from 0 through $\alpha_1+\alpha_2-1$
Parameters	Shape parameters $\alpha_1>0$ and $\alpha_2>0$
Range	$[0,1]$
Mean	$\dfrac{\alpha_1}{\alpha_1+\alpha_2}$
Variance	$\dfrac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}$
Mode	$\begin{cases} \dfrac{\alpha_1-1}{\alpha_1+\alpha_2-2} & \text{if } \alpha_1>1,\ \alpha_2>1 \\ 0 \text{ and } 1 & \text{if } \alpha_1<1,\ \alpha_2<1 \\ 0 & \text{if } (\alpha_1<1,\ \alpha_2\ge 1) \text{ or if } (\alpha_1=1,\ \alpha_2>1) \\ 1 & \text{if } (\alpha_1\ge 1,\ \alpha_2<1) \text{ or if } (\alpha_1>1,\ \alpha_2=1) \\ \text{does not uniquely exist} & \text{if } \alpha_1=\alpha_2=1 \end{cases}$
MLE	The following two equations must be satisfied: $\Psi(\hat\alpha_1)-\Psi(\hat\alpha_1+\hat\alpha_2)=\ln G_1\,, \qquad \Psi(\hat\alpha_2)-\Psi(\hat\alpha_1+\hat\alpha_2)=\ln G_2$ where Ψ is the digamma function, $G_1=(\Pi_{i=1}^n X_i)^{1/n}$, and $G_2=[\Pi_{i=1}^n(1-X_i)]^{1/n}$ [see Gnanadesikan, Pinkham, and Hughes (1967)]; note that $G_1+G_2\le 1$. These equations could be solved numerically [see Beckman and Tietjen (1978)], or approximations to $\hat\alpha_1$ and $\hat\alpha_2$ can be obtained from Table 6.20 (see App. 6A), which was computed for particular (G_1,G_2) pairs by modifications of the methods in Beckman and Tietjen (1978)
Comments	1. The U(0,1) and beta(1,1) distributions are the same 2. If X_1 and X_2 are independent random variables with $X_i \sim \text{gamma}(\alpha_i,\beta)$, then $X_1/(X_1+X_2)\sim\text{beta}(\alpha_1,\alpha_2)$ 3. A beta random variable X on $[0,1]$ can be rescaled and relocated to obtain a beta random variable on $[a,b]$ of the same shape by the transformation $a+(b-a)X$

TABLE 6.3 (*continued*)

Beta	$\text{beta}(\alpha_1,\alpha_2)$
	4. $X\sim\text{beta}(\alpha_1,\alpha_2)$ if and only if $1-X\sim\text{beta}(\alpha_2,\alpha_1)$ 5. $X\sim\text{beta}(\alpha_1,\alpha_2)$ if and only if $Y=X/(1-X)$ has a Pearson type VI distribution with shape parameters α_1, α_2 and scale parameter 1, denoted $\text{PT6}(\alpha_1,\alpha_2,1)$ 6. The beta(1,2) density is a left triangle, and the beta(2,1) density is a right triangle 7. $\lim_{x\to 0} f(x) = \begin{cases} \infty & \text{if } \alpha_1<1 \\ \alpha_2 & \text{if } \alpha_1=1 \\ 0 & \text{if } \alpha_1>1 \end{cases}, \qquad \lim_{x\to 1} f(x) = \begin{cases} \infty & \text{if } \alpha_2<1 \\ \alpha_1 & \text{if } \alpha_2=1 \\ 0 & \text{if } \alpha_2>1 \end{cases}$ 8. The density is symmetric about $x=\frac{1}{2}$ if and only if $\alpha_1=\alpha_2$. Also, the mean and the mode are equal if and only if $\alpha_1=\alpha_2$

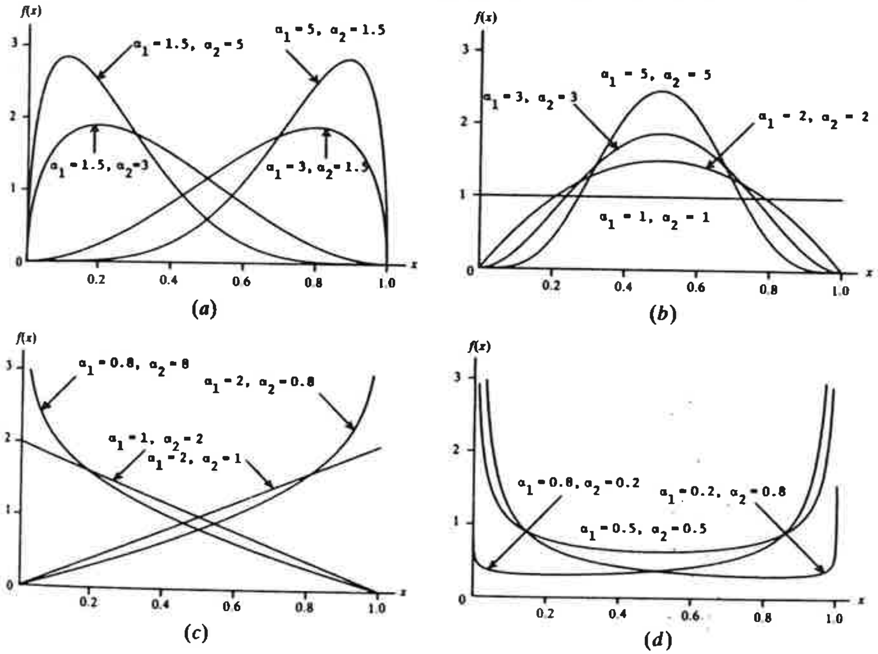

FIGURE 6.7
$\text{beta}(\alpha_1,\alpha_2)$ density functions.

Pearson type V	$\text{PT5}(\alpha,\beta)$
Possible applications	Time to perform some task (density takes on shapes similar to lognormal, but can have a larger "spike" close to $x=0$)
Density (see Fig. 6.8)	$f(x) = \begin{cases} \dfrac{x^{-(\alpha+1)}e^{-\beta/x}}{\beta^{-\alpha}\Gamma(\alpha)} & \text{if } x>0 \\ 0 & \text{otherwise} \end{cases}$

TABLE 6.3 *(continued)*

Pearson type V	**PT5(α,β)**
Distribution	$F(x) = \begin{cases} 1 - F_G\left(\frac{1}{x}\right) & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$ where $F_G(x)$ is the distribution function of a gamma($\alpha,1/\beta$) random variable
Parameters	Shape parameter $\alpha > 0$, scale parameter $\beta > 0$
Range	$[0,\infty)$
Mean	$\dfrac{\beta}{\alpha - 1}$ for $\alpha > 1$
Variance	$\dfrac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$
Mode	$\dfrac{\beta}{\alpha+1}$
MLE	If one has data $X_1, X_2, \ldots, X_n$, then fit a gamma(α_G,β_G) distribution to $1/X_1, 1/X_2, \ldots, 1/X_n$, resulting in the maximum-likelihood estimators $\hat{\alpha}_G$ and $\hat{\beta}_G$. Then the maximum-likelihood estimators for the PT5(α,β) are $\hat{\alpha} = \hat{\alpha}_G$ and $\hat{\beta} = 1/\hat{\beta}_G$ (see comment 1 below)
Comments	1. $X \sim$ PT5(α,β) if and only if $Y = 1/X \sim$ gamma($\alpha,1/\beta$). Thus, the Pearson type V distribution is sometimes called the *inverted gamma distribution* 2. Note that the mean and variance exist only for certain values of the shape parameter

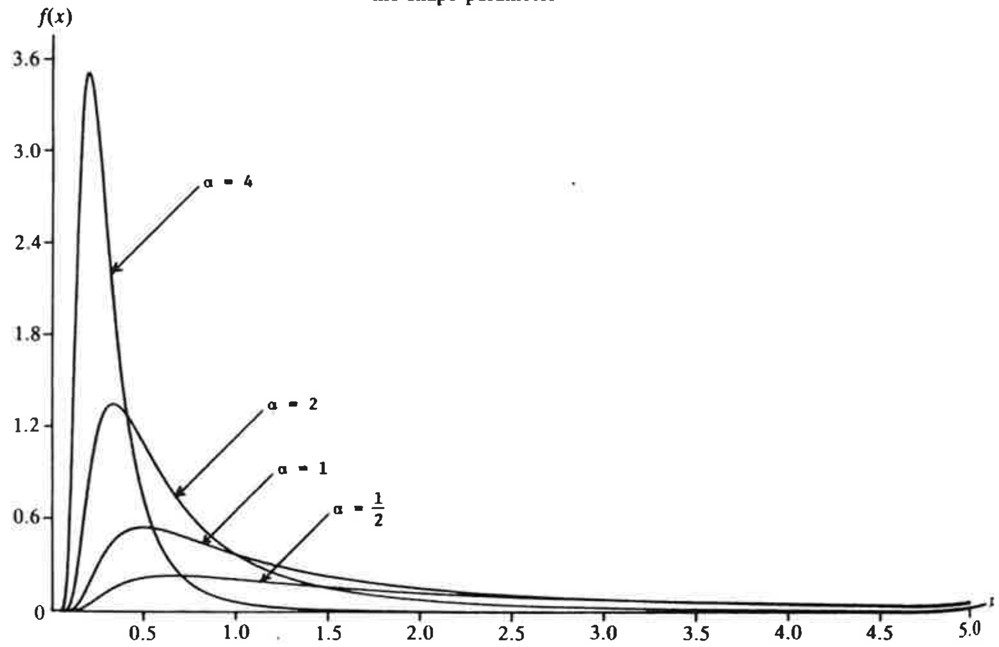

FIGURE 6.8
PT5(α,1) density functions.

TABLE 6.3 *(continued)*

Pearson type VI	**PT6(α_1,α_2,β)**
Possible applications	Time to perform some task
Density (see Fig. 6.9)	$f(x) = \begin{cases} \dfrac{(x/\beta)^{\alpha_1-1}}{\beta B(\alpha_1,\alpha_2)[1+(x/\beta)]^{\alpha_1+\alpha_2}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$
Distribution	$F(x) = \begin{cases} F_B\left(\dfrac{x}{x+\beta}\right) & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$ where $F_B(x)$ is the distribution function of a beta(α_1,α_2) random variable
Parameters	Shape parameters $\alpha_1 > 0$ and $\alpha_2 > 0$, scale parameter $\beta > 0$
Range	$[0,\infty)$
Mean	$\dfrac{\beta\alpha_1}{\alpha_2 - 1}$ for $\alpha_2 > 1$
Variance	$\dfrac{\beta^2\alpha_1(\alpha_1+\alpha_2-1)}{(\alpha_2-1)^2(\alpha_2-2)}$ for $\alpha_2 > 2$
Mode	$\begin{cases} \dfrac{\beta(\alpha_1-1)}{\alpha_2+1} & \text{if } \alpha_1 \ge 1 \\ 0 & \text{otherwise} \end{cases}$
MLE	If one has data $X_1, X_2, \ldots, X_n$ that are thought to be PT6(α_1,α_2,1), then fit a beta(α_1,α_2) distribution to $X_i/(1+X_i)$ for $i = 1, 2, \ldots, n$, resulting in the maximum-likelihood estimators $\hat{\alpha}_1$ and $\hat{\alpha}_2$. Then the maximum-likelihood estimators for the PT6(α_1,α_2,1) (note that $\beta = 1$) distribution are also $\hat{\alpha}_1$ and $\hat{\alpha}_2$ (see comment 1 below)
Comments	1. $X \sim$ PT6(α_1,α_2,1) if and only if $Y = X/(1+X) \sim$ beta(α_1,α_2) 2. If X_1 and X_2 are independent random variables with $X_i \sim$ gamma(α_i,β), then $Y = X_1/X_2 \sim$ PT6(α_1,α_2,β) (see Prob. 6.3) 3. Note that the mean and variance exist only for certain values of the shape parameter α_2

Triangular	**triang(a,b,c)**
Possible applications	Used as a rough model in the absence of data (see Sec. 6.9)
Density (see Fig. 6.10)	$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)} & \text{if } a \le x \le c \\ \dfrac{2(b-x)}{(b-a)(b-c)} & \text{if } c < x \le b \\ 0 & \text{otherwise} \end{cases}$
Distribution	$F(x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{(x-a)^2}{(b-a)(c-a)} & \text{if } a \le x \le c \\ 1 - \dfrac{(b-x)^2}{(b-a)(b-c)} & \text{if } c < x \le b \\ 1 & \text{if } b < x \end{cases}$
Parameters	a, b, and c real numbers with $a < c < b$. a is a location parameter, $b - a$ is a scale parameter, c is a shape parameter
Range	$[a,b]$

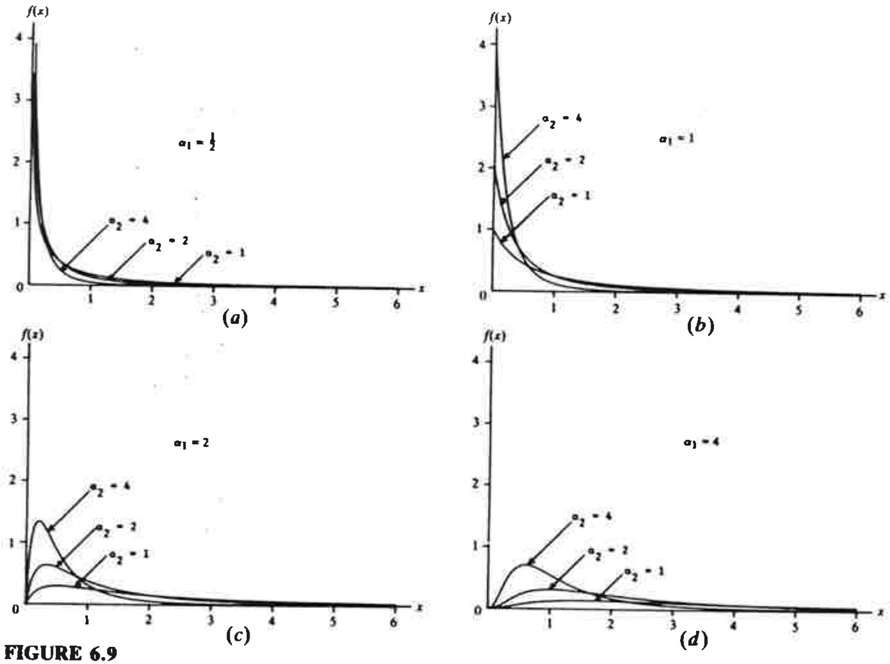

FIGURE 6.9
$PT6(\alpha_1,\alpha_2,1)$ density functions.

TABLE 6.3 *(continued)*

Triangular	**triang(a,b,c)**
Mean	$\dfrac{a+b+c}{3}$
Variance	$\dfrac{a^2+b^2+c^2-ab-ac-bc}{18}$
Mode	c
MLE	Our use of the triangular distribution, as described in Sec. 6.9, is as a rough model when there are no data. Thus, MLEs are not relevant
Comment	The limiting cases as $c \to b$ and $c \to a$ are called the *right triangular* and *left triangular distributions*, respectively, and are discussed in Prob. 8.7. For $a=0$ and $b=1$, both the left and right triangular distributions are special cases of the beta distribution

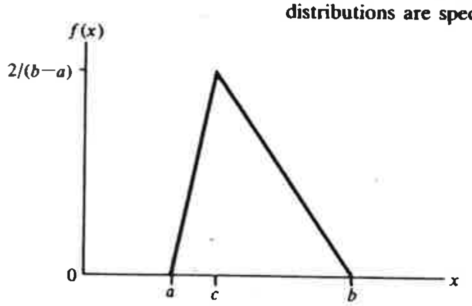

FIGURE 6.10
triang(a,b,c) density functions.

Lawless (1982) for other applications]. Then the density function and distribution function (if it exists in simple closed form) are listed. Next is a short description of the parameters, including their possible values. The range indicates the interval where the associated random variable can take on values. Also listed are the mean (expected value), variance, and mode, i.e., the value at which the density function is maximized. MLE refers to the maximum-likelihood estimator(s) of the parameter(s), treated later in Sec. 6.5. General comments include relationships of the distribution under study to other distributions. Graphs are given of the density functions for each distribution. The notation following the name of each distribution is our abbreviation for that distribution, which includes the parameters. The symbol ~ is read "is distributed as."

Note that we have included the less familiar Pearson type V and Pearson type VI distributions, because we have found that these distributions often provide a better fit to data sets whose histograms are skewed to the right (see Fig. 6.19) than standard distributions such as gamma, Weibull, and lognormal.

6.2.3 Discrete Distributions

The descriptions of the six discrete distributions in Table 6.4 follow the same pattern as for the continuous distributions in Table 6.3.

TABLE 6.4
Discrete distributions

Bernoulli	**Bernoulli(p)**
Possible applications	Random occurrence with two possible outcomes; used to generate other discrete random variates, e.g., binomial, geometric, and negative binomial
Mass (see Fig. 6.11)	$p(x) = \begin{cases} 1-p & \text{if } x=0 \\ p & \text{if } x=1 \\ 0 & \text{otherwise} \end{cases}$
Distribution	$F(x) = \begin{cases} 0 & \text{if } x<0 \\ 1-p & \text{if } 0 \le x < 1 \\ 1 & \text{if } 1 \le x \end{cases}$
Parameter	$p \in (0,1)$
Range	$\{0,1\}$
Mean	p
Variance	$p(1-p)$
Mode	$\begin{cases} 0 & \text{if } p < \frac{1}{2} \\ 0 \text{ and } 1 & \text{if } p = \frac{1}{2} \\ 1 & \text{if } p > \frac{1}{2} \end{cases}$
MLE	$\hat{p} = \bar{X}(n)$
Comments	1. A Bernoulli(p) random variable X can be thought of as the outcome of an experiment that either "fails" or "succeeds." If the probability of success is p, and we let $X=0$ if the experiment