8

Random-Variate Generation

This chapter deals with procedures for sampling from a variety of widely used continuous and discrete distributions. Previous discussions and examples indicated the usefulness of statistical distributions to model activities that are generally unpredictable or uncertain. For example, interarrival times and service times at queues, and demands for a product, are quite often unpredictable in nature, at least to a certain extent. Usually, such variables are modeled as random variables with some specified statistical distribution, and standard statistical procedures exist for estimating the parameters of the hypothesized distribution and for testing the validity of the assumed statistical model. Such procedures are discussed in Chapter 9.

In this chapter it is assumed that a distribution has been completely specified, and ways are sought to generate samples from this distribution to be used as input to a simulation model. The purpose of the chapter is to explain and illustrate some widely used techniques for generating random variates, not to give a state-of-the-art survey of the most efficient techniques. In practice, most simulation modelers will use existing routines available in programming libraries, or the routines built into the simulation language being used. However, some programming languages do not have built-in routines for all of the regularly used distributions, and some computer installations do not have random-variate-generation libraries, in which case the modeler must construct an acceptable routine. Even though the chance of this happening is small, it is nevertheless worthwhile to understand how random-variate generation occurs.

This chapter discusses the inverse transform technique, the convolution method, and, more briefly, the acceptance-rejection technique. Another technique, the composition method, is discussed by Fishman [1978] and Law and

Kelton [2000]. All the techniques in this chapter assume that a source of uniform (0,1) random numbers, $R_1, R_2, \ldots$ is readily available, where each R_i has pdf

$$f_R(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

and cdf

$$F_R(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

Throughout this chapter R and $R_1, R_2, \ldots$ represent random numbers uniformly distributed on (0,1) and generated by one of the techniques in Chapter 7 or taken from a random-number table such as Table A.1 described in Chapter 2.

8.1 Inverse Transform Technique

The inverse transform technique can be used to sample from the exponential, the uniform, the Weibull, and the triangular distributions and empirical distributions. Additionally, it is the underlying principle for sampling from a wide variety of discrete distributions. The technique will be explained in detail for the exponential distribution and then applied to other distributions. It is the most straightforward, but not always the most efficient, technique computationally.

8.1.1 Exponential Distribution

The exponential distribution, discussed in Section 5.4, has probability density function (pdf) given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

and cumulative distribution function (cdf) given by

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

The parameter λ can be interpreted as the mean number of occurrences per time unit. For example, if interarrival times $X_1, X_2, X_3, \ldots$ had an exponential distribution with rate λ, then λ could be interpreted as the mean number of arrivals per time unit, or the arrival rate. Notice that for any i

$$E(X_i) = \frac{1}{\lambda}$$

so that $1/\lambda$ is the mean interarrival time. The goal here is to develop a procedure for generating values $X_1, X_2, X_3, \ldots$ which have an exponential distribution.

The inverse transform technique can be utilized, at least in principle, for any distribution, but it is most useful when the cdf, $F(x)$, is of such simple form that its inverse, F^{-1}, can be easily computed.[1] A step-by-step procedure for the inverse transform technique, illustrated by the exponential distribution, is as follows:

Step 1. Compute the cdf of the desired random variable X. For the exponential distribution, the cdf is $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$.

Step 2. Set $F(X) = R$ on the range of X. For the exponential distribution, it becomes $1 - e^{-\lambda X} = R$ on the range $x \geq 0$. Since X is a random variable (with the exponential distribution in this case), it follows that $1 - e^{-\lambda X}$ is also a random variable, here called R. As will be shown later, R has a uniform distribution over the interval (0, 1).

Step 3. Solve the equation $F(X) = R$ for X in terms of R. For the exponential distribution, the solution proceeds as follows:

$$\begin{aligned} 1 - e^{-\lambda X} &= R \\ e^{-\lambda X} &= 1 - R \\ -\lambda X &= \ell n(1 - R) \\ X &= -\frac{1}{\lambda}\ell n(1 - R) \qquad (8.1) \end{aligned}$$

Equation (8.1) is called a random-variate generator for the exponential distribution. In general, Equation (8.1) is written as $X = F^{-1}(R)$. Generating a sequence of values is accomplished through step 4.

Step 4. Generate (as needed) uniform random numbers $R_1, R_2, R_3, \ldots$ and compute the desired random variates by

$$X_i = F^{-1}(R_i)$$

For the exponential case, $F^{-1}(R) = (-1/\lambda)\ell n(1 - R)$ by Equation (8.1), so that

$$X_i = -\frac{1}{\lambda}\ell n(1 - R_i) \qquad (8.2)$$

for $i = 1, 2, 3, \ldots$. One simplification that is usually employed in Equation (8.2) is to replace $1 - R_i$ by R_i to yield

$$X_i = -\frac{1}{\lambda}\ell n R_i \qquad (8.3)$$

which is justified since both R_i and $1 - R_i$ are uniformly distributed on (0, 1).

[1] The notation F^{-1} denotes the solution of the equation $r = F(x)$ in terms of r, not $1/F$.

Table 8.1. Generation of Exponential Variates X_i with Mean 1, Given Random Numbers R_i

i	1	2	3	4	5
R_i	0.1306	0.0422	0.6597	0.7965	0.7696
X_i	0.1400	0.0431	1.078	1.592	1.468

$X_i = -\frac{1}{\lambda} \ln(1-R_i)$

EXAMPLE 8.1

Table 8.1 gives a sequence of random numbers from Table A.1 and the computed exponential variates, X_i, given by Equation (8.2) with a value of $\lambda = 1$. Figure 8.1(a) is a histogram of 200 values, $R_1, R_2, \ldots, R_{200}$ from the uniform dis-

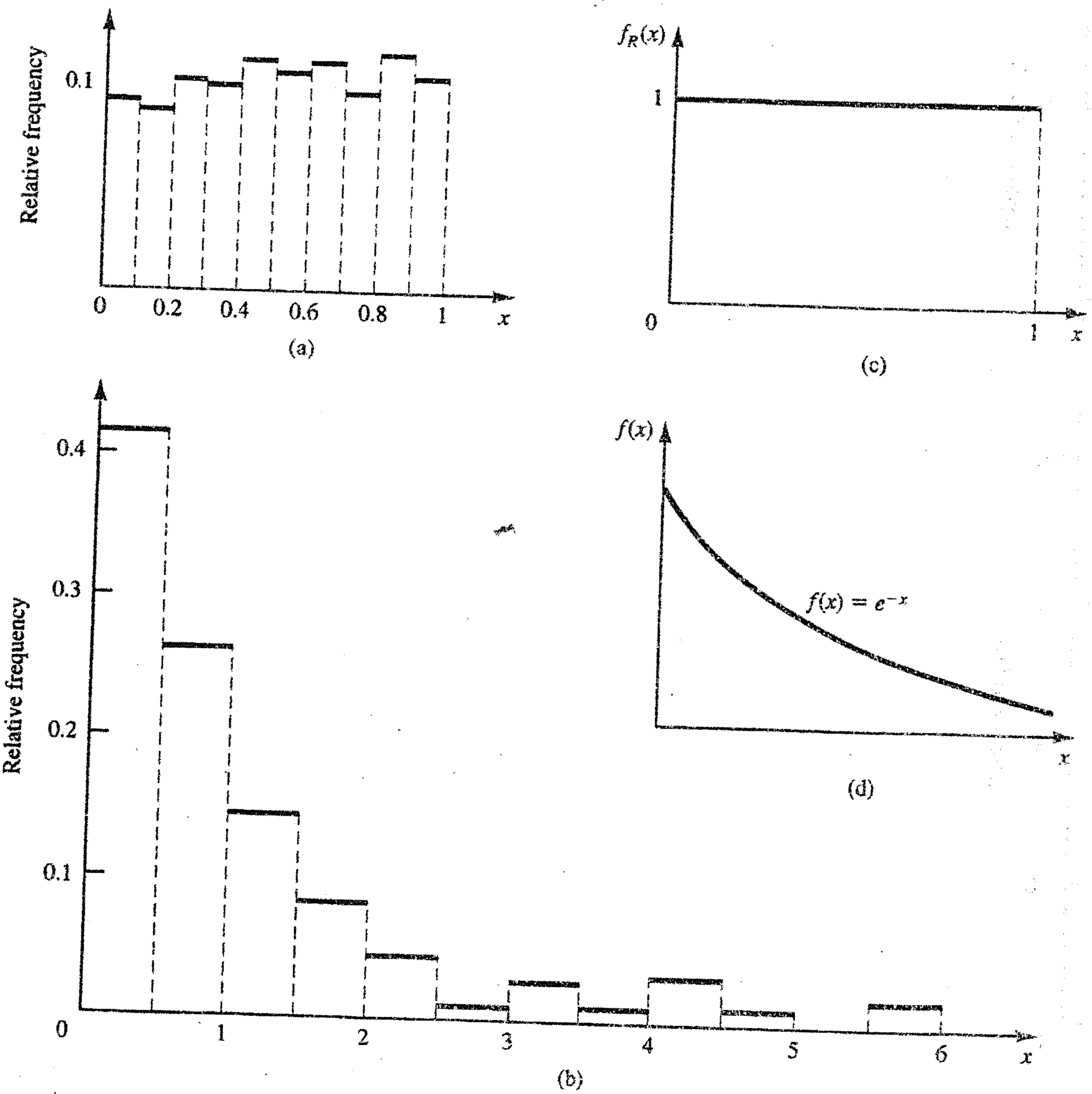

Figure 8.1. (a) Empirical histogram of 200 uniform random numbers; (b) empirical histogram of 200 exponential variates; (c) theoretical uniform density on (0, 1); (d) theoretical exponential density with mean 1.

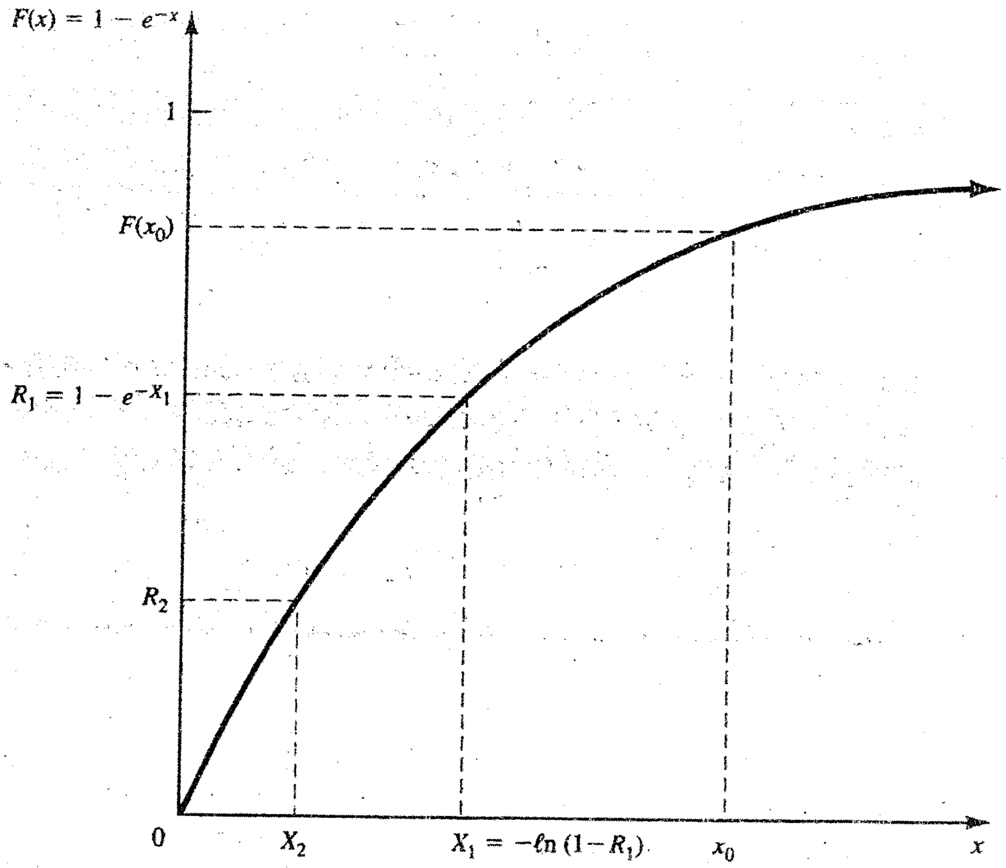

Figure 8.2. Graphical view of the inverse transform technique.

tribution, and Figure 8.1(b) is a histogram of the 200 values, $X_1, X_2, \ldots, X_{200}$, computed by Equation (8.2). Compare these empirical histograms with the theoretical density functions in Figure 8.1(c) and (d). As illustrated here, a histogram is an estimate of the underlying density function. (This fact is used in Chapter 9 as a way to identify distributions.) ◀

Figure 8.2 gives a graphical interpretation of the inverse transform technique. The cdf shown is $F(x) = 1 - e^{-x}$, an exponential distribution with rate $\lambda = 1$. To generate a value X_1 with cdf $F(x)$, first a random number R_1 between 0 and 1 is generated, a horizontal line is drawn from R_1 to the graph of the cdf, then a vertical line is dropped to the x-axis to obtain X_1, the desired result. Notice the inverse relation between R_1 and X_1, namely

$$R_1 = 1 - e^{-X_1}$$

and

$$X_1 = -\ell n(1 - R_1)$$

In general, the relation is written as

$$R_1 = F(X_1)$$

and

$$X_1 = F^{-1}(R_1)$$

Why does the random variable X_1 generated by this procedure have the desired distribution? Pick a value x_0 and compute the cumulative probability

$$P(X_1 \le x_0) = P(R_1 \le F(x_0)) = F(x_0) \tag{8.4}$$

To see the first equality in Equation (8.4), refer to Figure 8.2, where the fixed numbers x_0 and $F(x_0)$ are drawn on their respective axes. It can be seen that $X_1 \le x_0$ when and only when $R_1 \le F(x_0)$. Since $0 \le F(x_0) \le 1$, the second equality in Equation (8.4) follows immediately from the fact that R_1 is uniformly distributed on (0, 1). Equation (8.4) shows that the cdf of X_1 is F; hence, X_1 has the desired distribution.

8.1.2 Uniform Distribution

Consider a random variable X that is uniformly distributed on the interval $[a, b]$. A reasonable guess for generating X is given by

$$X = a + (b - a)R \tag{8.5}$$

[Recall that R is always a random number on (0, 1).] The pdf of X is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

The derivation of Equation (8.5) follows steps 1 through 3 of Section 8.1.1:

Step 1. The cdf is given by

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

Step 2. Set $F(X) = (X - a)/(b - a) = R$.

Step 3. Solving for X in terms of R yields $X = a + (b - a)R$, which agrees with Equation (8.5).

8.1.3 Weibull Distribution

The Weibull distribution was introduced in Section 5.4 as a model for time to failure for machines or electronic components. When the location parameter ν is set to 0, its pdf is given by Equation (5.46) as

$$f(x) = \begin{cases} \dfrac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(x/\alpha)^\beta}, & x \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\alpha > 0$ and $\beta > 0$ are the scale and shape parameters of the distribution. To generate a Weibull variate, follow steps 1 through 3 of Section 8.1.1:

Step 1. The cdf is given by $F(X) = 1 - e^{-(x/\alpha)^\beta}, x \ge 0$.

Step 2. Let $F(X) = 1 - e^{-(X/\alpha)^\beta} = R$.
Step 3. Solving for X in terms of R yields

$$X = \alpha[-\ell n(1 - R)]^{1/\beta} \tag{8.6}$$

The derivation of Equation (8.6) is left as Exercise 10 for the reader. By comparing Equations (8.6) and (8.1), it can be seen that if X is a Weibull variate, then X^β is an exponential variate with mean α^β. Conversely, if Y is an exponential variate with mean μ, then $Y^{1/\beta}$ is a Weibull variate with shape parameter β and scale parameter $\alpha = \mu^{1/\beta}$.

8.1.4 Triangular Distribution

Consider a random variable X which has pdf

$$f(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2 - x, & 1 < x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

as shown in Figure 8.3. This distribution is called a triangular distribution with endpoints (0, 2) and mode at 1. Its cdf is given by

$$F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{x^2}{2}, & 0 < x \le 1 \\ 1 - \dfrac{(2-x)^2}{2}, & 1 < x \le 2 \\ 1, & x > 2 \end{cases}$$

For $0 \le X \le 1$,

$$R = \frac{X^2}{2} \tag{8.7}$$

and for $1 \le X \le 2$,

$$R = 1 - \frac{(2 - X)^2}{2} \tag{8.8}$$

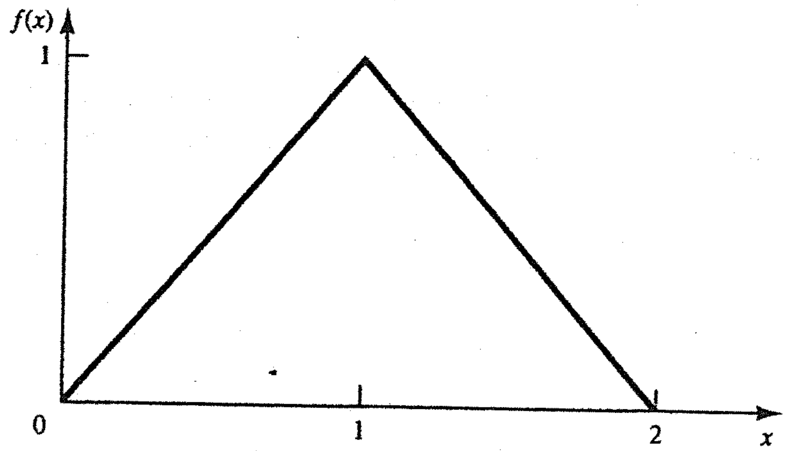

Figure 8.3. Density function for a triangular distribution.

By Equation (8.7), $0 \le X \le 1$ implies that $0 \le R \le \frac{1}{2}$, in which case $X = \sqrt{2R}$. By Equation (8.8), $1 \le X \le 2$ implies that $\frac{1}{2} \le R \le 1$, in which case $X = 2 - \sqrt{2(1-R)}$. Thus, X is generated by

$$X = \begin{cases} \sqrt{2R}, & 0 \le R \le \frac{1}{2} \\ 2 - \sqrt{2(1-R)}, & \frac{1}{2} < R \le 1 \end{cases} \tag{8.9}$$

Exercises 2, 3, and 4 give the student practice in dealing with other triangular distributions. Notice that if the pdf and cdf of the random variable X come in parts (i.e., require different formulas over different parts of the range of X), then the application of the inverse transform technique for generating X will result in separate formulas over different parts of the range of R, as in Equation (8.9). A general form of the triangular distribution was discussed in Section 5.4.

8.1.5 *Empirical Continuous Distributions*

If the modeler has been unable to find a theoretical distribution that provides a good model for the input data, then it may be necessary to use the empirical distribution of the data. One possibility is to simply resample the observed data itself. This is known as using the *empirical distribution*, and it makes particularly good sense when the input process is known to take on a finite number of values. See Section 8.1.7 for an example of this type of situation and a method for generating random inputs.

On the other hand, if the data are drawn from what is believed to be a continuous-valued input process, then it makes sense to interpolate between the observed data points to fill in the gaps. This section describes a method for defining and generating data from a continuous empirical distribution.

EXAMPLE 8.2

Five observations of fire crew response times (in minutes) to incoming alarms have been collected to be used in a simulation investigating possible alternative staffing and crew scheduling policies. The data are

2.76 1.83 0.80 1.45 1.24

Before collecting more data, it is desired to develop a preliminary simulation model which uses a response-time distribution based on these five observations. Thus, a method for generating random variates from the response-time distribution is needed. Initially, it will be assumed that response times X have a range $0 \le X \le c$, where c is unknown, but will be estimated by $\widehat{c} = \max\{X_i\colon i = 1, \ldots, n\} = 2.76$, where $\{X_i, i = 1, \ldots, n\}$ are the raw data and $n = 5$ is the number of observations.

Arrange the data from smallest to largest and let $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ denote these sorted values. Since the smallest possible value is believed to be 0, define $x_{(0)} = 0$. Assign a probability of $1/n = 1/5$ to each interval

Table 8.2. Summary of Fire Crew Response-Time Data

i	*Interval,* $x_{(i-1)} < x \le x_{(i)}$	*Probability,* $1/n$	*Cumulative Probability,* i/n	*Slope,* a_i
1	$0.0 < x \le 0.80$	0.2	0.2	4.00
2	$0.80 < x \le 1.24$	0.2	0.4	2.20
3	$1.24 < x \le 1.45$	0.2	0.6	1.05
4	$1.45 < x \le 1.83$	0.2	0.8	1.90
5	$1.83 < x \le 2.76$	0.2	1.0	4.65

$x_{(i-1)} < x \le x_{(i)}$, as shown in Table 8.2. The resulting empirical cdf, $\widehat{F}(x)$, is illustrated in Figure 8.4. The slope of the ith line segment is given by

$$a_i = \frac{x_{(i)} - x_{(i-1)}}{i/n - (i-1)/n} = \frac{x_{(i)} - x_{(i-1)}}{1/n}$$

The inverse cdf is calculated by

$$X = \widehat{F}^{-1}(R) = x_{(i-1)} + a_i\left(R - \frac{(i-1)}{n}\right) \tag{8.10}$$

when $(i-1)/n < R \le i/n$.

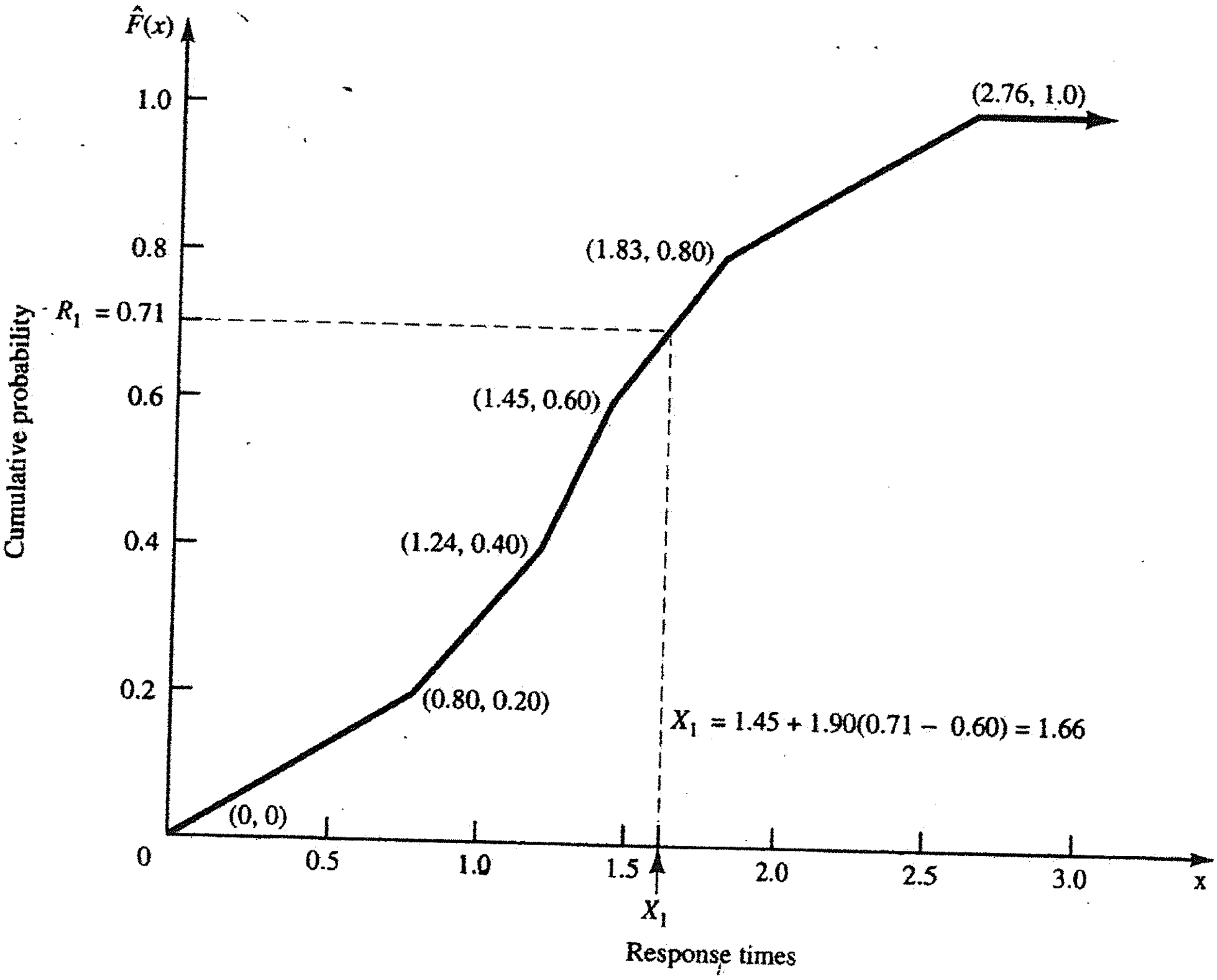

Figure 8.4. Empirical cdf of fire crew response times.

For example, if a random number $R_1 = 0.71$ is generated, then R_1 is seen to lie in the fourth interval (between $3/5 = 0.60$ and $4/5 = 0.80$), so that by Equation (8.10),

$$\begin{aligned} X_1 &= x_{(4-1)} + a_4(R_1 - (4-1)/n) \\ &= 1.45 + 1.90(0.71 - 0.60) \\ &= 1.66 \end{aligned}$$

The reader is referred to Figure 8.4 for a graphical view of the generation procedure. ◀

In Example 8.2 each data point was represented in the empirical cdf. If a large sample of data is available (and sample sizes from several hundred to tens of thousands are possible with modern, automated data collection), then it may be more convenient and computationally efficient to first summarize the data into a frequency distribution with a much smaller number of intervals and then fit a continuous empirical cdf to the frequency distribution. Only a slight generalization of Equation (8.10) is required to accomplish this. Now the slope of the ith line segment is given by

$$a_i = \frac{x_{(i)} - x_{(i-1)}}{c_i - c_{i-1}}$$

where c_i is the cumulative probability of the first i intervals of the frequency distribution and $x_{(i-1)} < x \le x_{(i)}$ is the ith interval. The inverse cdf is calculated by

$$X = \hat{F}^{-1}(R) = x_{(i-1)} + a_i\,(R - c_{i-1}) \tag{8.11}$$

when $c_{i-1} < R \le c_i$.

Example 8.3

Suppose that 100 broken-widget repair times have been collected. The data are summarized in Table 8.3 in terms of the number of observations in various intervals. For example, there were 31 observations between 0 and 0.5 hour, 10 between 0.5 and 1 hour, and so on.

Suppose it is known that all repairs take at least 15 minutes, so that $X \ge 0.25$ hour always. Then we set $x_{(0)} = 0.25$, as shown in Table 8.3 and Figure 8.5.

Table 8.3. Summary of Repair-Time Data

i	Interval (Hours)	Frequency	Relative Frequency	Cumulative Frequency, c_i	Slope, a_i
1	$0.25 \le x \le 0.5$	31	0.31	0.31	0.81
2	$0.5 < x \le 1.0$	10	0.10	0.41	5.0
3	$1.0 < x \le 1.5$	25	0.25	0.66	2.0
4	$1.5 < x \le 2.0$	34	0.34	1.00	1.47

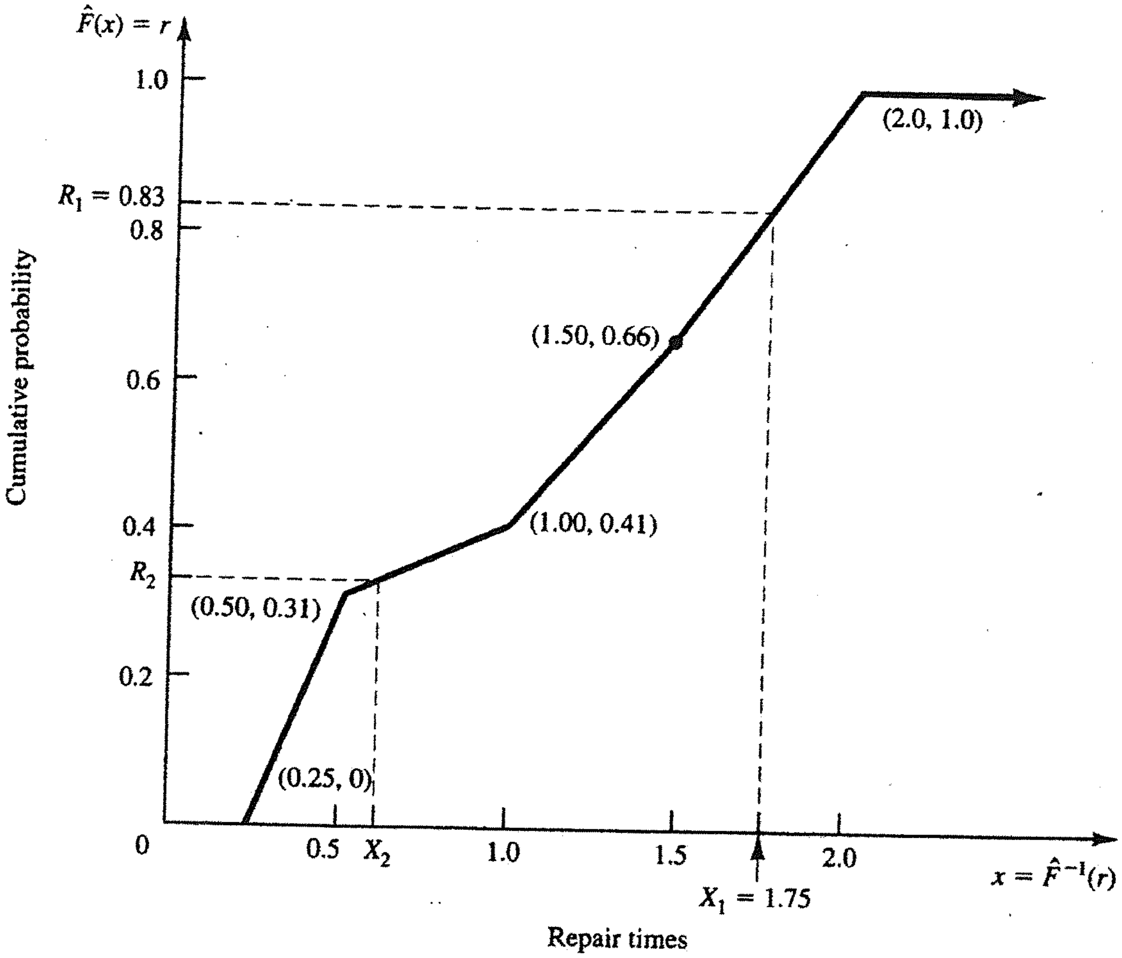

Figure 8.5. Generating variates from the empirical distribution function for repair-time data ($X \geq 0.25$).

For example, suppose the first random number generated is $R_1 = 0.83$. Then since R_1 is between $c_3 = 0.66$ and $c_4 = 1.00$, X_1 is

$$X_1 = x_{(4-1)} + a_4(R_1 - c_{4-1}) = 1.5 + 1.47(0.83 - 0.66) = 1.75 \quad (8.12)$$

As another illustration, suppose that $R_2 = 0.33$. Since $c_1 = 0.31 < R_2 \leq 0.41 = c_2$,

$$\begin{aligned} X_2 &= x_{(1)} + a_2(R_2 - c_1) \\ &= 0.5 + 5.0(0.33 - 0.31) \\ &= 0.6 \end{aligned}$$

The point ($R_2 = 0.33$, $X_2 = 0.6$) is also shown in Figure 8.5. ◀

Now reconsider the data of Table 8.3. The data are restricted in the range $0.25 \leq X \leq 2.0$, but the underlying distribution may have a wider range. This provides one important reason for attempting to find a theoretical statistical distribution (such as the gamma or Weibull) for the data, since these distributions allow a wider range, namely $0 \leq X < \infty$. On the other hand, an empirical distribution adheres closely to what is present in the data itself, and the data are often the best source of information available.

When data are summarized in terms of frequency intervals, it is recommended that relatively short intervals be used, as this results in a more accurate portrayal of the underlying cdf. For example, for the repair-time data of Table 8.3, for which there were $n = 100$ observations, a much more accurate estimate could have been obtained by using 10 to 20 intervals, certainly not an excessive number, rather than the four fairly wide intervals actually used here for purposes of illustration.

Several comments are in order:

1. A computerized version of the procedure will become more inefficient as the number of intervals, n, increases. A systematic computerized version is often called a table-lookup generation scheme, because given a value of R, the computer program must search an array of c_i values to find the interval i in which R lies, namely the interval i satisfying

$$c_{i-1} < R \leq c_i$$

The more intervals there are, the longer on the average the search will take if it is implemented in the crude way described here. The analyst should consider this trade-off between accuracy of the estimating cdf and computational efficiency when programming the procedure. If a large number of observations are available, the analyst may well decide to group the observations from 20 to 50 intervals (say) and then use the procedure of Example 8.3. Or a more efficient table-lookup procedure may be used, such as the one described in Law and Kelton [2000].

2. In Example 8.2 it was assumed that response times X satisfied $0 \leq X \leq 2.76$. This assumption led to the inclusion of the points $x_{(0)} = 0$ and $x_{(5)} = 2.76$ in Figure 8.4 and Table 8.2. If it is known a priori that X falls in some other range, for example, if it is known that response times are always between 15 seconds and 3 minutes, that is,

$$0.25 \leq X \leq 3.0$$

then the points $x_{(0)} = 0.25$ and $x_{(6)} = 3.0$ would be used to estimate the empirical cdf of response times. Notice that because of inclusion of the new point $x_{(6)}$ there are now six intervals instead of five and each interval is assigned probability $1/6 = 0.167$. Exercise 12 illustrates the use of these additional assumptions.

8.1.6 Continuous Distributions without a Closed-Form Inverse

A number of useful continuous distributions do not have a closed form expression for their cdf or its inverse; examples include the normal, gamma, and beta distributions. For this reason, it is often stated that the inverse transform technique for random-variate generation cannot be used for these distributions. This is not true, provided we are willing to *approximate* the inverse cdf, or numerically integrate and search the cdf. Although this may sound imprecise, notice that even a closed-form inverse requires approximation in order to evaluate it on a computer. For example, generating exponentially distributed

Table 8.4. Comparison of Approximate Inverse with Exact Values (to Four Decimal Places) for the Standard Normal Distribution

R	*Approximate Inverse*	*Exact Inverse*
0.01	−2.3263	−2.3373
0.10	−1.2816	−1.2813
0.25	−0.6745	−0.6713
0.50	0.0000	0.0000
0.75	0.6745	0.6713
0.90	1.2816	1.2813
0.99	2.3263	2.3373

random variates via the inverse cdf $X = F^{-1}(R) = -\ell n(1 - R)/\lambda$ requires a numerical approximation for the logarithm function. Thus, there is no essential difference between using an approximate inverse cdf and approximately evaluating a closed-form inverse. The problem with using an approximate inverse cdf is that some of them are computationally slow to evaluate.

To illustrate the idea, consider the simple approximation to the inverse cdf of the standard normal distribution proposed by Schmeiser [1979]:

$$X = F^{-1}(R) \approx \frac{R^{0.135} - (1 - R)^{0.135}}{0.1975}$$

This approximation gives at least one-decimal-place accuracy for $0.0013499 \leq R \leq 0.9986501$. Table 8.4 compares the approximation with exact values (to four decimal places) obtained by numerical integration for several values of R. Much more accurate approximations exist that are only slightly more complicated. A good source of these approximations for a number of distributions is Bratley, Fox, and Schrage [1987].

8.1.7 Discrete Distributions

All discrete distributions can be generated using the inverse transform technique, either numerically through a table-lookup procedure, or in some cases algebraically with the final generation scheme in terms of a formula. Other techniques are sometimes used for certain distributions, such as the convolution technique for the binomial distribution. Some of these methods are discussed in later sections. This subsection gives examples covering both empirical distributions and two of the standard discrete distributions, the (discrete) uniform and the geometric. Highly efficient table-lookup procedures for these and other distributions are found in Bratley, Fox, and Schrage [1987] and Ripley [1987].

Table 8.5. Distribution of Number of Shipments, X

x	$p(x)$	$F(x)$
0	0.50	0.50
1	0.30	0.80
2	0.20	1.00

EXAMPLE 8.4 (An Empirical Discrete Distribution)

At the end of the day, the number of shipments on the loading dock of the IHW Company (whose main product is the famous, incredibly huge widget) is either 0, 1, or 2, with observed relative frequency of occurrence of 0.50, 0.30, and 0.20, respectively. Internal consultants have been asked to develop a model to improve the efficiency of the loading and hauling operations, and as part of this model they will need to be able to generate values, X, to represent the number of shipments on the loading dock at the end of each day. The consultants decide to model X as a discrete random variable with distribution as given in Table 8.5 and shown in Figure 8.6.

The probability mass function (pmf), $p(x)$, is given by

$$p(0) = P(X = 0) = 0.50$$
$$p(1) = P(X = 1) = 0.30$$
$$p(2) = P(X = 2) = 0.20$$

and the cdf, $F(x) = P(X \leq x)$, is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.5, & 0 \leq x < 1 \\ 0.8, & 1 \leq x < 2 \\ 1.0, & 2 \leq x \end{cases}$$

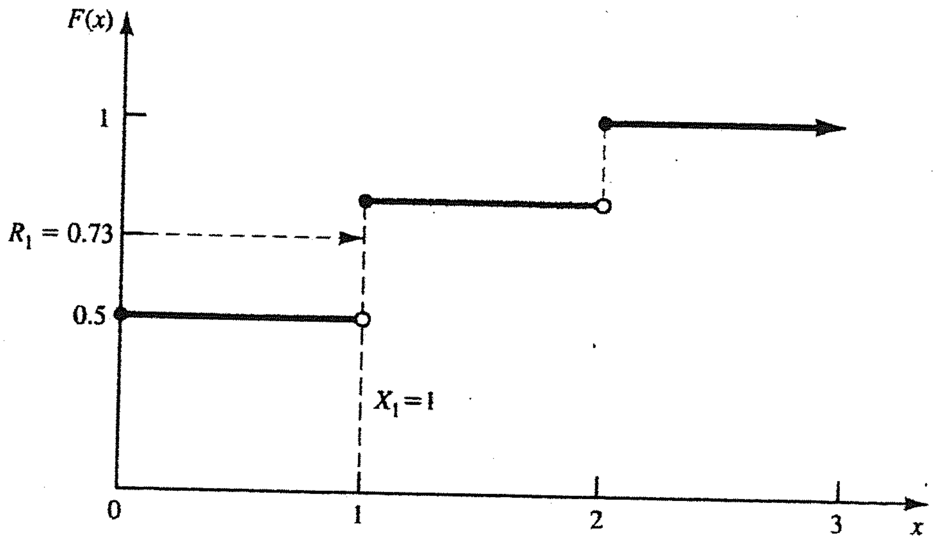

Figure 8.6. The cdf of number of shipments, X.

Table 8.6. Table for Generating the Discrete Variate X

i	*Input,* r_i	*Output,* x_i
1	0.50	0
2	0.80	1
3	1.00	2

Recall that the cdf of a discrete random variable always consists of horizontal line segments with jumps of size $p(x)$ at those points, x, which the random variable can assume. For example, in Figure 8.6 there is a jump of size $p(0) = 0.5$ at $x = 0$, of size $p(1) = 0.3$ at $x = 1$, and of size $p(2) = 0.2$ at $x = 2$.

For generating discrete random variables, the inverse transform technique becomes a table-lookup procedure, but unlike the case of continuous variables, interpolation is not required. To illustrate the procedure, suppose that $R_1 = 0.73$ is generated. Graphically, as illustrated in Figure 8.6, first locate $R_1 = 0.73$ on the vertical axis, next draw a horizontal line segment until it hits a "jump" in the cdf, and then drop a perpendicular to the horizontal axis to get the generated variate. Here $R_1 = 0.73$ is transformed to $X_1 = 1$. This procedure is analogous to the procedure used for empirical continuous distributions in Section 8.1.5 and illustrated in Figure 8.5, except that the final step of linear interpolation is eliminated.

The table-lookup procedure is facilitated by construction of a table such as Table 8.6. When $R_1 = 0.73$ is generated, first find the interval in which R_1 lies. In general, for $R = R_1$, if

$$F(x_{i-1}) = r_{i-1} < R \leq r_i = F(x_i) \tag{8.13}$$

then set $X_1 = x_i$. Here $r_0 = 0$, $x_0 = -\infty$, while $x_1, x_2, \ldots, x_n$ are the possible values of the random variable, and $r_k = p(x_1) + \cdots + p(x_k)$, $k = 1, 2, \ldots, n$. For this example, $n = 3$, $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, and hence $r_1 = 0.5$, $r_2 = 0.8$, and $r_3 = 1.0$. (Notice that $r_n = 1.0$ in all cases.)

Since $r_1 = 0.5 < R_1 = 0.73 \leq r_2 = 0.8$, set $X_1 = x_2 = 1$. The generation scheme is summarized as follows:

$$X = \begin{cases} 0, & R \leq 0.5 \\ 1, & 0.5 < R \leq 0.8 \\ 2, & 0.8 < R \leq 1.0 \end{cases}$$

◀

Example 8.4 illustrates the table-lookup procedure, while the next example illustrates an algebraic approach that can be used for certain distributions.

Example 8.5 (A Discrete Uniform Distribution)

Consider the discrete uniform distribution on $\{1, 2, \ldots, k\}$ with pmf and cdf given by

$$p(x) = \frac{1}{k}, \quad x = 1, 2, \ldots, k$$

and

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{k}, & 1 \le x < 2 \\ \frac{2}{k}, & 2 \le x < 3 \\ \vdots & \vdots \\ \frac{k-1}{k}, & k-1 \le x < k \\ 1, & k \le x \end{cases}$$

Let $x_i = i$ and $r_i = p(1) + \cdots + p(x_i) = F(x_i) = i/k$ for $i = 1, 2, \ldots, k$. Then by using Inequality (8.13) it can be seen that if the generated random number R satisfies

$$r_{i-1} = \frac{i-1}{k} < R \le r_i = \frac{i}{k} \tag{8.14}$$

then X is generated by setting $X = i$. Now, Inequality (8.14) can be solved for i:

$$i - 1 < Rk \le i$$

$$Rk \le i < Rk + 1 \tag{8.15}$$

Let $\lceil y \rceil$ denote the smallest integer $\ge y$. For example, $\lceil 7.82 \rceil = 8$, $\lceil 5.13 \rceil = 6$, and $\lceil -1.32 \rceil = -1$. For $y \ge 0$, $\lceil y \rceil$ is a function that rounds up. This notation and Inequality (8.15) yield a formula for generating X, namely

$$X = \lceil Rk \rceil \tag{8.16}$$

For example, consider generating a random variate X, uniformly distributed on $\{1, 2, \ldots, 10\}$. The variate, X, might represent the number of pallets to be loaded onto a truck. Using Table A.1 as a source of random numbers, R, and Equation (8.16) with $k = 10$ yields

$$R_1 = 0.78, \quad X_1 = \lceil 7.8 \rceil = 8$$

$$R_2 = 0.03, \quad X_2 = \lceil 0.3 \rceil = 1$$

$$R_3 = 0.23, \quad X_3 = \lceil 2.3 \rceil = 3$$

$$R_4 = 0.97, \quad X_4 = \lceil 9.7 \rceil = 10$$

The procedure discussed here can be modified to generate a discrete uniform random variate with any range consisting of consecutive integers. Exercise 13 asks the student to devise a procedure for one such case. ◀

EXAMPLE 8.6

Consider the discrete distribution with pmf given by

$$p(x) = \frac{2x}{k(k+1)}, \quad x = 1, 2, \ldots, k$$

(This example is taken from Schmidt and Taylor [1970].) For integer values of x in the range $\{1, 2, \ldots, k\}$, the cdf is given by

$$\begin{aligned} F(x) &= \sum_{i=1}^{x} \frac{2i}{k(k+1)} \\ &= \frac{2}{k(k+1)} \sum_{i=1}^{x} i \\ &= \frac{2}{k(k+1)} \frac{x(x+1)}{2} \\ &= \frac{x(x+1)}{k(k+1)} \end{aligned}$$

Generate R and use Inequality (8.13) to conclude that $X = x$ whenever

$$F(x-1) = \frac{(x-1)x}{k(k+1)} < R \le \frac{x(x+1)}{k(k+1)} = F(x)$$

or whenever

$$(x-1)x < k(k+1)R \le x(x+1)$$

To solve this inequality for x in terms of R, first find a value of x that satisfies

$$(x-1)x = k(k+1)R$$

or

$$x^2 - x - k(k+1)R = 0$$

Then by rounding up, the solution is $X = \lceil x - 1 \rceil$. By the quadratic formula, namely

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

with $a = 1, b = -1$, and $c = -k(k+1)R$, the solution to the quadratic equation is

$$x = \frac{1 \pm \sqrt{1 + 4k(k+1)R}}{2} \tag{8.17}$$

The positive root in Equation (8.17) is the correct one to use (why?), so X is generated by

$$X = \left\lceil \frac{1 + \sqrt{1 + 4k(k+1)R}}{2} - 1 \right\rceil \tag{8.18}$$

Exercise 14 asks the student to generate a few values from this distribution. ◀

EXAMPLE 8.7 (The Geometric Distribution)

Consider the geometric distribution with pmf

$$p(x) = p(1-p)^x, \quad x = 0, 1, 2, \ldots$$

where $0 < p < 1$. Its cdf is given by

$$\begin{aligned} F(x) &= \sum_{j=0}^{x} p(1-p)^j \\ &= \frac{p\{1-(1-p)^{x+1}\}}{1-(1-p)} \\ &= 1-(1-p)^{x+1} \end{aligned}$$

for $x = 0, 1, 2, \ldots$ Using the inverse transform technique [i.e., Inequality (8.13)], recall that a geometric random variable X will assume the value x whenever

$$F(x-1) = 1-(1-p)^x < R \le 1-(1-p)^{x+1} = F(x) \tag{8.19}$$

where R is a generated random number assumed $0 < R < 1$. Solving Inequality (8.19) for x proceeds as follows:

$$\begin{aligned} (1-p)^{x+1} &\le 1-R < (1-p)^x \\ (x+1)\ell n(1-p) &\le \ell n(1-R) < x\ell n(1-p) \end{aligned}$$

But $1-p < 1$ implies that $\ell n(1-p) < 0$, so that

$$\frac{\ell n(1-R)}{\ell n(1-p)} - 1 \le x < \frac{\ell n(1-R)}{\ell n(1-p)} \tag{8.20}$$

Thus, $X = x$ for that integer value of x satisfying Inequality (8.20), or, in brief, using the round-up function $\lceil \cdot \rceil$

$$X = \left\lceil \frac{\ell n(1-R)}{\ell n(1-p)} - 1 \right\rceil \tag{8.21}$$

Since p is a fixed parameter, let $\beta = -1/\ell n(1-p)$. Then $\beta > 0$ and, by Equation (8.21), $X = \lceil -\beta \ell n(1-R) - 1 \rceil$. By Equation (8.1), $-\beta \ell n(1-R)$ is an exponentially distributed random variable with mean β, so that one way of generating a geometric variate with parameter p is to generate (by any method) an exponential variate with parameter $\beta^{-1} = -\ell n(1-p)$, subtract one, and round up.

Occasionally, a geometric variate X is needed which can assume values $\{q, q+1, q+2, \ldots\}$ with pmf $p(x) = p(1-p)^{x-q}$ $(x = q, q+1, \ldots)$. Such a variate, X can be generated, using Equation (8.21), by

$$X = q + \left\lceil \frac{\ell n(1-R)}{\ell n(1-p)} - 1 \right\rceil \tag{8.22}$$

One of the most common cases is $q = 1$. ◀

EXAMPLE 8.8

Generate three values from a geometric distribution on the range $\{X \geq 1\}$ with mean 2. Such a geometric distribution has pmf $p(x) = p(1-p)^{x-1}(x = 1, 2, \ldots)$ with mean $1/p = 2$, or $p = 1/2$. Thus, X can be generated by Equation (8.22) with $q = 1$, $p = 1/2$, and $1/\ell n(1-p) = -1.443$. Using Table A.1, $R_1 = 0.932$, $R_2 = 0.105$, and $R_3 = 0.687$, which yields

$$\begin{aligned} X_1 &= 1 + \lceil -1.443 \ell n(1 - 0.932) - 1 \rceil \\ &= 1 + \lceil 3.878 - 1 \rceil = 4 \\ X_2 &= 1 + \lceil -1.443 \ell n(1 - 0.105) - 1 \rceil = 1 \\ X_3 &= 1 + \lceil -1.443 \ell n(1 - 0.687) - 1 \rceil = 2 \end{aligned}$$

Exercise 15 deals with an application of the geometric distribution. ◀

8.2 Direct Transformation for the Normal and Lognormal Distributions

Many methods have been developed for generating normally distributed random variates. The inverse transform technique cannot be applied easily, however, because the inverse cdf cannot be written in closed form. The standard normal cdf is given by

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt, \quad -\infty < x < \infty$$

This section describes an intuitively appealing direct transformation that produces an independent pair of standard normal variates with mean zero and variance 1. The method is due to Box and Muller [1958]. Although not as efficient as many more modern techniques, it is easy to program in a scientific language such as FORTRAN, C, or Pascal. We then show how to transform a standard normal variate into a normal variate with mean μ and variance σ^2. Once we have a method (this or any other) for generating X from a $N(\mu, \sigma^2)$ distribution, then we can generate a lognormal random variate Y with parameters μ and σ^2 using the direct transformation $Y = e^X$ [recall that μ and σ^2 are *not* the mean and variance of the lognormal; see Equations (5.57) and (5.58)].

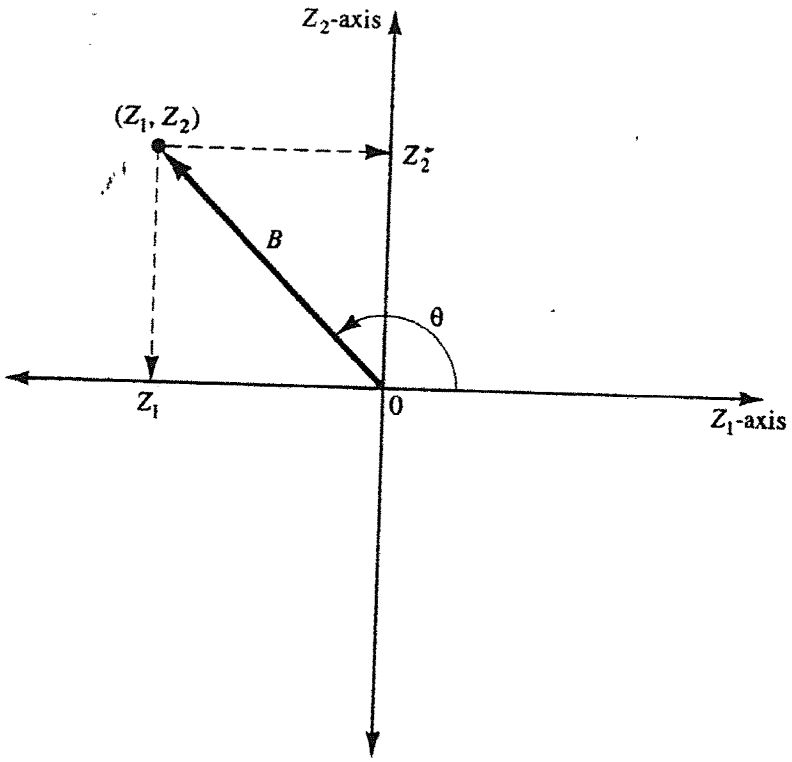

Figure 8.7. Polar representation of a pair of standard normal variables.

Consider two standard normal random variables, Z_1 and Z_2, plotted as a point in the plane as shown in Figure 8.7 and represented in polar coordinates as

$$Z_1 = B \cos \theta$$
$$Z_2 = B \sin \theta \tag{8.23}$$

It is known that $B^2 = Z_1^2 + Z_2^2$ has the chi-square distribution with 2 degrees of freedom, which is equivalent to an exponential distribution with mean 2. Thus, the radius, B, can be generated by use of Equation (8.3):

$$B = (-2\ell n R)^{1/2} \tag{8.24}$$

By the symmetry of the normal distribution, it seems reasonable to suppose, and indeed it is the case, that the angle is uniformly distributed between 0 and 2π radians. In addition, the radius, B, and the angle, θ, are mutually independent. Combining Equations (8.23) and (8.24) gives a direct method for generating two independent standard normal variates, Z_1 and Z_2, from two independent random numbers R_1 and R_2:

$$Z_1 = (-2\ell n R_1)^{1/2} \cos (2\pi R_2)$$
$$Z_2 = (-2\ell n R_1)^{1/2} \sin (2\pi R_2) \tag{8.25}$$

To illustrate the generation scheme, consider Equation (8.25) with $R_1 = 0.1758$ and $R_2 = 0.1489$. Two standard normal random variates are generated as follows:

$$Z_1 = [-2\ell n(0.1758)]^{1/2} \cos(2\pi 0.1489) = 1.11$$
$$Z_2 = [-2\ell n(0.1758)]^{1/2} \sin(2\pi 0.1489) = 1.50$$

To obtain normal variates X_i with mean μ and variance σ^2, we then apply the transformation

$$X_i = \mu + \sigma Z_i \tag{8.26}$$

to the standard normal variates. For example, to transform the two standard normal variates into normal variates with mean $\mu = 10$ and variance $\sigma^2 = 4$ we compute

$$X_1 = 10 + 2(1.11) = 12.22$$
$$X_2 = 10 + 2(1.50) = 13.00$$

8.3 Convolution Method

The probability distribution of a sum of two or more independent random variables is called a convolution of the distributions of the original variables. The convolution method thus refers to adding together two or more random variables to obtain a new random variable with the desired distribution. This technique can be applied to obtain Erlang variates and binomial variates. What is important is not the cdf of the desired random variable, but rather its relation to other more easily generated variates.

8.3.1 Erlang Distribution

As discussed in Section 5.4, an Erlang random variable X with parameters (K, θ) can be shown to be the sum of K independent exponential random variables, $X_i (i = 1, \ldots, K)$, each having mean $1/K\theta$; that is,

$$X = \sum_{i=1}^{K} X_i$$

Since each X_i can be generated by Equation (8.3) with $1/\lambda = 1/K\theta$, an Erlang variate can be generated by

$$X = \sum_{i=1}^{K} -\frac{1}{K\theta} \ell n R_i$$

$$= -\frac{1}{K\theta} \ell n \left(\prod_{i=1}^{K} R_i \right) \tag{8.27}$$

In Equation (8.27), $\prod$ stands for product. It is more efficient computationally to multiply all the random numbers first and then to compute only one logarithm.

EXAMPLE 8.9

Trucks arrive at a large warehouse in a completely random fashion which is modeled as a Poisson process with arrival rate $\lambda = 10$ trucks per hour. The guard at the entrance sends trucks alternately to the north and south docks. An analyst has developed a model to study the loading/unloading process at the south docks and needs a model of the arrival process at the south docks alone. An interarrival time X between successive truck arrivals at the south docks is equal to the sum of two interarrival times at the entrance and thus it is the sum of two exponential random variables, each with mean 0.1 hour, or 6 minutes. Thus, X has the Erlang distribution with $K = 2$ and mean $1/\theta = 2/\lambda = 0.2$ hour. To generate the variate X, first obtain $K = 2$ random numbers from Table A.1, say $R_1 = 0.937$ and $R_2 = 0.217$. Then by Equation (8.27),

$$\begin{aligned} X &= -0.1\ell n[0.937(0.217)] \\ &= 0.159 \text{ hour} = 9.56 \text{ minutes} \end{aligned}$$

◀

In general, Equation (8.27) implies that K uniform random numbers are needed for each Erlang variate generated. If K is large, it is more efficient to generate Erlang variates by other techniques, such as one of the many acceptance-rejection techniques for the gamma distribution given by Bratley, Fox, and Schrage [1987], Fishman [1978], and Law and Kelton [2000].

8.4 Acceptance-Rejection Technique

Suppose that an analyst needed to devise a method for generating random variates, X, uniformly distributed between 1/4 and 1. One way to proceed would be to follow these steps:

Step 1. Generate a random number R.
Step 2a. If $R \geq 1/4$, accept $X = R$, then go to step 3.
Step 2b. If $R < 1/4$, reject R, and return to step 1.
Step 3. If another uniform random variate on [1/4, 1] is needed, repeat the procedure beginning at step 1. If not, stop.

Each time step 1 is executed, a new random number R must be generated. Step 2a is an "acceptance" and step 2b is a "rejection" in this acceptance-rejection technique. To summarize the technique, random variates (R) with some distribution (here uniform on [0, 1]) are generated until some condition ($R > 1/4$) is satisfied. When the condition is finally satisfied, the desired random variate, X (here uniform on [1/4, 1]), can be computed ($X = R$). This procedure can be shown to be correct by recognizing that the accepted values of R are conditioned values; that is, R itself does not have the desired distribution, but R conditioned on the event $\{R \geq 1/4\}$ does have the desired distribution. To show this, take $1/4 \leq a < b \leq 1$; then

$$P(a < R \leq b | 1/4 \leq R \leq 1) = \frac{P(a < R \leq b)}{P(1/4 \leq R \leq 1)} = \frac{b-a}{3/4} \qquad (8.28)$$

which is the correct probability for a uniform distribution on [1/4, 1]. Equation (8.28) says that the probability distribution of R, given that R is between 1/4 and 1 (all other values of R are thrown out), is the desired distribution. Therefore, if $1/4 \leq R \leq 1$, set $X = R$.

The efficiency of an acceptance-rejection technique depends heavily on being able to minimize the number of rejections. In this example, the probability of a rejection is $P(R < 1/4) = 1/4$, so that the number of rejections is a geometrically distributed random variable with probability of "success" being $p = 3/4$ and mean number of rejections $(1/p - 1) = 4/3 - 1 = 1/3$. (Example 8.7 discussed the geometric distribution.) The mean number of random numbers R required to generate one variate X is one more than the number of rejections; hence, it is $4/3 = 1.33$. In other words, to generate 1000 values of X would require approximately 1333 random numbers R.

In the present situation an alternative procedure exists for generating a uniform variate on [1/4, 1], namely Equation (8.5), which reduces to $X = 1/4 + (3/4)R$. Whether the acceptance-rejection technique or an alternative procedure such as the inverse-transform technique [Equation (8.5)] is the more efficient depends on several considerations. The computer being used, the skills of the programmer, and the relative efficiency of generating the additional (rejected) random numbers needed by acceptance-rejection should be compared to the computations required by the alternative procedure. In practice, concern with generation efficiency is left to specialists who conduct extensive tests comparing alternative methods.

For the uniform distribution on [1/4, 1], the inverse transform technique of Equation (8.5) is undoubtedly much easier to apply and more efficient than the acceptance-rejection technique. The main purpose of this example was to explain and motivate the basic concept of the acceptance-rejection technique. However, for some important distributions such as the normal, gamma, and beta, the inverse cdf does not exist in closed form and therefore the inverse transform technique is difficult. These more advanced techniques are summarized by Bratley, Fox, and Schrage [1987], Fishman [1978], and Law and Kelton [2000].

In the following subsections, the acceptance-rejection technique is illustrated for the generation of random variates for the Poisson and gamma distributions.

8.4.1 Poisson Distribution

A Poisson random variable, N, with mean $\alpha > 0$ has pmf

$$p(n) = P(N = n) = \frac{e^{-\alpha}\alpha^n}{n!}, \quad n = 0, 1, 2, \ldots$$

but more important, N can be interpreted as the number of arrivals from a Poisson arrival process in one unit of time. Recall from Section 5.5 that the interarrival times, $A_1, A_2, \ldots$ of successive customers are exponentially distributed with rate α (i.e., α is the mean number of arrivals per unit time); in addition, an exponential variate can be generated by Equation (8.3). Thus there is a relationship between the (discrete) Poisson distribution and the (continuous) exponential distribution, namely:

$$N = n \tag{8.29}$$

if and only if

$$A_1 + A_2 + \cdots + A_n \le 1 < A_1 + \cdots + A_n + A_{n+1} \tag{8.30}$$

Equation (8.29), $N = n$, says there were exactly n arrivals during one unit of time; but relation (8.30) says that the nth arrival occurred before time 1 while the $(n+1)$st arrival occurred after time 1. Clearly, these two statements are equivalent. Proceed now by generating exponential interarrival times until some arrival, say $n+1$, occurs after time 1; then set $N = n$.

For efficient generation purposes, relation (8.30) is usually simplified by first using Equation (8.3), $A_i = (-1/\alpha)\ell n R_i$, to obtain

$$\sum_{i=1}^{n} -\frac{1}{\alpha}\ell n R_i \le 1 < \sum_{i=1}^{n+1} -\frac{1}{\alpha}\ell n R_i$$

Next multiply through by $-\alpha$, which reverses the sign of the inequality, and use the fact that a sum of logarithms is the logarithm of a product, to get

$$\ell n \prod_{i=1}^{n} R_i = \sum_{i=1}^{n} \ell n R_i \ge -\alpha > \sum_{i=1}^{n+1} \ell n R_i = \ell n \prod_{i=1}^{n+1} R_i$$

Finally, use the relation $e^{\ell n x} = x$ for any number x to obtain

$$\prod_{i=1}^{n} R_i \ge e^{-\alpha} > \prod_{i=1}^{n+1} R_i \tag{8.31}$$

which is equivalent to relation (8.30). The procedure for generating a Poisson random variate, N, is given by the following steps:

Step 1. Set $n = 0$, $P = 1$.

Step 2. Generate a random number R_{n+1} and replace P by $P \cdot R_{n+1}$.

Step 3. If $P < e^{-\alpha}$, then accept $N = n$. Otherwise, reject the current n, increase n by one, and return to step 2.

Notice that upon completion of step 2, P is equal to the rightmost expression in relation (8.31). The basic idea of a rejection technique is again exhibited; if $P \ge e^{-\alpha}$ in step 3, then n is rejected and the generation process must proceed through at least one more trial.

How many random numbers will be required, on the average, to generate one Poisson variate, N? If $N = n$, then $n + 1$ random numbers are required, so the average number is given by

$$E(N + 1) = \alpha + 1$$

which is quite large if the mean, α, of the Poisson distribution is large.

EXAMPLE 8.10

Generate three Poisson variates with mean $\alpha = 0.2$. First compute $e^{-\alpha} = e^{-0.2} = 0.8187$. Next get a sequence of random numbers R from Table A.1 and follow steps 1 to 3 above:

Step 1. Set $n = 0$, $P = 1$.
Step 2. $R_1 = 0.4357$, $P = 1 \cdot R_1 = 0.4357$.
Step 3. Since $P = 0.4357 < e^{-\alpha} = 0.8187$, accept $N = 0$.
Step 1–3. ($R_1 = 0.4146$ leads to $N = 0$.)
Step 1. Set $n = 0$, $P = 1$.
Step 2. $R_1 = 0.8353$, $P = 1 \cdot R_1 = 0.8353$.
Step 3. Since $P \geq e^{-\alpha}$, reject $n = 0$ and return to step 2 with $n = 1$.
Step 2. $R_2 = 0.9952$, $P = R_1 R_2 = 0.8313$.
Step 3. Since $P \geq e^{-\alpha}$, reject $n = 1$ and return to step 2 with $n = 2$.
Step 2. $R_3 = 0.8004$, $P = R_1 R_2 R_3 = 0.6654$.
Step 3. Since $P < e^{-\alpha}$, accept $N = 2$.

The calculations required for the generation of these three Poisson random variates are summarized as follows:

n	R_{n+1}	P	*Accept/Reject*	*Result*
0	0.4357	0.4357	$P < e^{-\alpha}$ (accept)	$N = 0$
0	0.4146	0.4146	$P < e^{-\alpha}$ (accept)	$N = 0$
0	0.8353	0.8353	$P \geq e^{-\alpha}$ (reject)	
1	0.9952	0.8313	$P \geq e^{-\alpha}$ (reject)	
2	0.8004	0.6654	$P < e^{-\alpha}$ (accept)	$N = 2$

It took five random numbers, R, to generate three Poisson variates here ($N = 0$, $N = 0$, and $N = 2$), but in the long run to generate, say, 1000 Poisson variates with mean $\alpha = 0.2$ it would require approximately $1000(\alpha + 1)$ or 1200 random numbers. ◀

EXAMPLE 8.11

Buses arrive at the bus stop at Peachtree and North Avenue according to a Poisson process with a mean of one bus per 15 minutes. Generate a random variate, N, which represents the number of arriving buses during a 1-hour time slot. Now, N is Poisson distributed with a mean of four buses per hour. First compute $e^{-\alpha} = e^{-4} = 0.0183$. Using a sequence of 12 random numbers from Table A.1 yields the following summarized results:

n	R_{n+1}	P	*Accept/Reject*	*Result*
0	0.4357	0.4357	$P \geq e^{-\alpha}$ (reject)	
1	0.4146	0.1806	$P \geq e^{-\alpha}$ (reject)	
2	0.8353	0.1508	$P \geq e^{-\alpha}$ (reject)	
3	0.9952	0.1502	$P \geq e^{-\alpha}$ (reject)	
4	0.8004	0.1202	$P \geq e^{-\alpha}$ (reject)	
5	0.7945	0.0955	$P \geq e^{-\alpha}$ (reject)	
6	0.1530	0.0146	$P < e^{-\alpha}$ (accept)	$N = 6$

It is immediately seen that a larger value of α (here $\alpha = 4$) usually requires more random numbers; if 1000 Poisson variates were desired, approximately $1000(\alpha + 1) = 5000$ random numbers would be required. ◀

When α is large, say $\alpha \geq 15$, the rejection technique outlined here becomes quite expensive, but fortunately an approximate technique based on the normal distribution works quite well. When the mean, α, is large,

$$Z = \frac{N - \alpha}{\sqrt{\alpha}}$$

is approximately normally distributed with mean zero and variance 1, which suggests an approximate technique. First generate a standard normal variate Z, by Equation (8.25), then generate the desired Poisson variate, N, by

$$N = \lceil \alpha + \sqrt{\alpha} Z - 0.5 \rceil \tag{8.32}$$

where $\lceil \cdot \rceil$ is the round-up function described in Section 8.1.7. (If $\alpha + \sqrt{\alpha}Z - 0.5 < 0$, then set $N = 0$.) The "0.5" used in the formula makes the round-up function become a "round to the nearest integer" function. Equation (8.32) is not an acceptance-rejection technique, but, used as an alternative to the acceptance rejection method, it provides a fairly efficient and accurate method for generating Poisson variates with a large mean.

8.4.2 Gamma Distribution

Several acceptance-rejection techniques for generating gamma random variates have been developed (Bratley, Fox, and Schrage [1987]; Fishman [1978]; Law and Kelton [2000]). One of the most efficient is by Cheng [1977]; the mean number of trials is between 1.13 and 1.47 for any value of the shape parameter $\beta \geq 1$.

If the shape parameter β is an integer, say $\beta = k$, one possibility is to use the convolution technique in Section 8.3.1, since the Erlang distribution is a special case of the more general gamma distribution. On the other hand, the acceptance-rejection technique described here would be a highly efficient method for the Erlang distribution, especially if $\beta = k$ were large. The routine generates gamma random variates with scale parameter θ and shape parameter β — that is, with mean $1/\theta$ and variance $1/\beta\theta^2$. The steps are as follows:

Step 1. Compute $a = (2\beta - 1)^{1/2}$, $b = 2\beta - \ell n4 + 1/a$.
Step 2. Generate R_1 and R_2.
Step 3. Compute $X = \beta[R_1/(1 - R_1)]^a$.
Step 4a. If $X > b - \ell n(R_1^2 R_2)$, reject X and return to step 2.
Step 4b. If $X \leq b - \ell n(R_1^2 R_2)$, use X as the desired variate. The generated variates from step 4b will have mean and variance both equal to β. If it is desired to have mean $1/\theta$ and variance $1/\beta\theta^2$ as in Section 5.4, then include
Step 5. Replace X by $X/\beta\theta$.

The basic idea of all acceptance-rejection methods is again illustrated here, but the proof of this example is beyond the scope of this book. In step 3, $X = \beta[R_1/(1 - R_1)]^a$ is not gamma distributed, but rejection of certain values of X in step 4a guarantees that the accepted values in step 4b do have the gamma distribution.

EXAMPLE 8.12

Downtimes for a high-production candy-making machine have been found to be gamma distributed with mean 2.2 minutes and variance 2.10 minutes². Thus, $1/\theta = 2.2$ and $1/\beta\theta^2 = 2.10$, which implies that $\beta = 2.30$ and $\theta = 0.4545$.

Step 1. $a = 1.90$, $b = 3.74$.
Step 2. Generate $R_1 = 0.832$, $R_2 = 0.021$.
Step 3. Compute $X = 2.3(0.832/0.168)^{1.9} = 48.1$.
Step 4. $X = 48.1 > 3.74 - \ell n[(0.832)^2 0.021] = 7.97$, so reject X and return to step 2.
Step 2. Generate $R_1 = 0.434$, $R_2 = 0.716$.
Step 3. Compute $X = 2.3(0.434/0.566)^{1.9} = 1.389$.
Step 4. Since $X = 1.389 \leq 3.74 - \ell n[(0.434)^2 0.716] = 5.74$, accept X.
Step 5. Divide X by $\beta\theta = 1.045$ to get $X = 1.329$.

This example took two trials (i.e., one rejection) to generate an acceptable gamma-distributed random variate, but on the average to generate, say, 1000 gamma variates, the method will require between 1130 and 1470 trials, or equivalently, between 2260 and 2940 random numbers. The method is somewhat cumbersome for hand calculations, but is easy to program on the computer and is one of the most efficient gamma generators known. ◀

8.5 Summary

The basic principles of random-variate generation using the inverse transform technique, the convolution method, and acceptance-rejection techniques have been introduced and illustrated by examples. Methods for generating many of the important continuous and discrete distributions, as well as empirical distributions, have been given. See Schmeiser [1980] for an excellent survey; for a state-of-the-art treatment, the reader is referred to Devroye [1986] or Dagpunar [1988].

REFERENCES

BOX, G. E. P., AND M. F. MULLER [1958], "A Note on the Generation of Random Normal Deviates," *Annals of Mathematical Statistics*, Vol. 29, pp. 610–11.

BRATLEY, P., B. L. FOX, AND L. E. SCHRAGE [1987], *A Guide to Simulation*, 2d ed., Springer-Verlag, New York.

CHENG, R. C. H. [1977], "The Generation of Gamma Variables," *Applied Statistician*, Vol. 26, No. 1, pp. 71–75.

DAGPUNAR, J. [1988], *Principles of Random Variate Generation*, Clarendon Press, Oxford.

DEVROYE, L. [1986], *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

FISHMAN, G. S. [1978], *Principles of Discrete Event Simulation*, John Wiley, New York.

LAW, A. M., AND W. D. KELTON [2000], *Simulation Modeling & Analysis*, 3d ed., McGraw-Hill, New York.

RIPLEY, B. D. [1987], *Stochastic Simulation*, John Wiley, New York.

SCHMEISER, B. W. [1979], "Approximations to the Inverse Cumulative Normal Function for Use on Hand Calculators," *Applied Statistics*, Vol. 28, pp. 175–176.

SCHMEISER, B. W. [1980], "Random Variate Generation: A Survey," in *Simulation with Discrete Models: A State of the Art View*, eds. T. I. Ören, C. M. Shub, and P. F. Roth, IEEE.

SCHMIDT, J. W., AND R. E. TAYLOR [1970], *Simulation and Analysis of Industrial Systems*, Irwin, Homewood, Ill.

EXERCISES

1. Develop a random-variate generator for a random variable X with the pdf

$$f(x) = \begin{cases} e^{2x}, & -\infty < x \leq 0 \\ e^{-2x}, & 0 < x < \infty \end{cases}$$

2. Develop a generation scheme for the triangular distribution with pdf

$$f(x) = \begin{cases} \frac{1}{2}(x-2), & 2 \leq x \leq 3 \\ \frac{1}{2}\left(2 - \frac{x}{3}\right), & 3 < x \leq 6 \\ 0, & \text{otherwise} \end{cases}$$

   Generate 10 values of the random variate, compute the sample mean, and compare it to the true mean of the distribution.

3. Develop a generator for a triangular distribution with range (1, 10) and mode at $x = 4$.

4. Develop a generator for a triangular distribution with range (1, 10) and a mean of 4.

5. Given the following cdf for a continuous variable with range −3 to 4, develop a generator for the variable.

$$F(x) = \begin{cases} 0, & x \le -3 \\ \frac{1}{2} + \frac{x}{6}, & -3 < x \le 0 \\ \frac{1}{2} + \frac{x^2}{32}, & 0 < x \le 4 \\ 1, & x > 4 \end{cases}$$

6. Given the cdf $F(x) = x^4/16$ on $0 \le x \le 2$, develop a generator for this distribution.

7. Given the pdf $f(x) = x^2/9$ on $0 \le x \le 3$, develop a generator for this distribution.

8. Develop a generator for a random variable whose pdf is

$$f(x) = \begin{cases} \frac{1}{3}, & 0 \le x \le 2 \\ \frac{1}{24}, & 2 < x \le 10 \\ 0, & \text{otherwise} \end{cases}$$

9. The cdf of a discrete random variable X is given by

$$F(x) = \frac{x(x+1)(2x+1)}{n(n+1)(2n+1)}, \quad x = 1, 2, \ldots, n$$

When $n = 4$, generate three values of X using $R_1 = 0.83$, $R_2 = 0.24$, and $R_3 = 0.57$.

10. Times to failure for an automated production process have been found to be randomly distributed with a Weibull distribution with parameters $\beta = 2$ and $\alpha = 10$. Derive Equation (8.6) and then use it to generate five values from this Weibull distribution, using five random numbers taken from Table A.1.

11. Data have been collected on service times at a drive-in bank window at the Shady Lane National Bank. This data are summarized into intervals as follows:

Interval (Seconds)	*Frequency*
15–30	10
30–45	20
45–60	25
60–90	35
90–120	30
120–180	20
180–300	10

Set up a table like Table 8.2 for generating service times by the table-lookup method and generate five values of service time using four-digit random numbers.

12. In Example 8.2, assume that fire crew response times satisfy $0.25 \le x \le 3$. Modify Table 8.2 to accommodate this assumption. Then generate five values of response time using four-digit uniform random numbers from Table A.1.

13. For a preliminary version of a simulation model, the number of pallets, X, to be loaded onto a truck at a loading dock was assumed to be uniformly distributed between 8 and 24. Devise a method for generating X, assuming that the loads on successive trucks are independent. Use the technique of Example 8.5 for discrete uniform distributions. Finally, generate loads for 10 successive trucks by using four-digit random numbers.

14. After collecting more data, it was found that the distribution of Example 8.6 was a better approximation to the number of pallets loaded than was the uniform distribution, as was assumed in Exercise 13. Using Equation (8.18) generate loads for 10 successive trucks using the same random numbers as were used in Exercise 13. Compare the results to the results of Exercise 13.

15. The weekly demand, X, for a slow-moving item has been found to be well approximated by a geometric distribution on the range $\{0, 1, 2, \ldots\}$ with mean weekly demand of 2.5 items. Generate 10 values of X, demand per week, using random numbers from Table A.1. [*Hint*: For a geometric distribution on the range $\{q, q+1, \ldots\}$ with parameter p, the mean is $1/p + q - 1$.]

16. In Exercise 15, suppose that the demand has been found to have a Poisson distribution with mean 2.5 items per week. Generate 10 values of X, demand per week, using random numbers from Table A.1. Discuss the differences between the geometric and the Poisson distributions.

17. Lead times have been found to be exponentially distributed with mean 3.7 days. Generate five random lead times from this distribution.

18. Regular maintenance of a production routine has been found to vary and has been modeled as a normally distributed random variable with mean 33 minutes and variance 4 minutes^2. Generate five random maintenance times, with the given distribution.

19. A machine is taken out of production if it fails, or after 5 hours, whichever comes first. By running similar machines until failure, it has been found that time to failure, X, has the Weibull distribution with $\alpha = 8$, $\beta = 0.75$, and $\nu = 0$ (refer to Sections 5.4 and 8.1.3). Thus, the time until the machine is taken out of production can be represented as $Y = \min(X, 5)$. Develop a step-by-step procedure for generating Y.

20. The time until a component is taken out of service is uniformly distributed on 0 to 8 hours. Two such independent components are put in series, and the whole system goes down when one of the components goes down. If X_i $(i = 1, 2)$ represents the component runtimes, then $Y = \min(X_1, X_2)$ represents the system lifetime. Devise two distinct ways to generate Y. [*Hint*: One way is relatively straightforward. For a second method, first compute the cdf of Y: $F_Y(y) = P(Y \leq y) = 1 - P(Y > y)$, for $0 \leq y \leq 8$. Use the equivalence $\{Y > y\} = \{X_1 > y \text{ and } X_2 > y\}$ and the independence of X_1, and X_2. After finding $F_Y(y)$, proceed with the inverse transform technique.]

21. In Exercise 20, component lifetimes are exponentially distributed, one with mean 2 hours and the other with mean 6 hours. Rework Exercise 20 under this new assumption. Discuss the relative efficiency of the two generation schemes devised.

22. Develop a technique for generating a binomial random variable, X, using the convolution technique. [*Hint*: X can be represented as the number of successes in n independent Bernoulli trials, each success having probability p. Thus, $X = \sum_{i=1}^{n} X_i$, where $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$.]

23. Develop an acceptance-rejection technique for generating a geometric random variable, X, with parameter p on the range $\{0, 1, 2, \ldots\}$. (*Hint*: X can be thought of as the number of trials before the first success occurs in a sequence of independent Bernoulli trials.)

24. Write a computer routine to generate standard normal variates by the exact method discussed in this chapter. Use it to generate 1000 values. Compare the true probability, $\Phi(z)$, that a value lies in $(-\infty, z)$ to the actual observed relative frequency that values were $\leq z$, for $z = -4, -3, -2, -1, 0, 1, 2, 3$ and 4.

25. Write a computer routine to generate gamma variates with shape parameter β and scale parameter θ. Generate 1000 values with $\beta = 2.5$ and $\theta = 0.2$ and compare the true mean, $1/\theta = 5$, to the sample mean.

26. Write a computer routine to generate 200 values from one of the variates in Exercises 1 to 23. Make a histogram of the 200 values and compare it to the theoretical density function (or probability mass function for discrete random variables).

27. Many spreadsheet, symbolic calculation, and statistical analysis programs have built-in routines for generating random variates from standard distributions. Try to find out what variate-generation methods are used in one of these packages by looking at the documentation. Should you trust a variate generator if the method is not documented?

28. Suppose that somehow we have available a source of exponentially distributed random variates with mean 1. Write an algorithm to generate random variates with a triangular distribution by transforming the exponentially distributed random variates. [*Hint*: First transform to obtain uniformly distributed random variates.]