Chapter 2

Random Variables

2.1 INTRODUCTION

In this chapter, the concept of a random variable is introduced. The main purpose of using a random variable is so that we can define certain probability functions that make it both convenient and easy to compute the probabilities of various events.

2.2 RANDOM VARIABLES

A. Definitions:

Consider a random experiment with sample space S. A *random variable* $X(\zeta)$ is a single-valued real function that assigns a real number called the *value* of $X(\zeta)$ to each sample point ζ of S. Often, we use a single letter X for this function in place of $X(\zeta)$ and use r.v. to denote the random variable.

Note that the terminology used here is traditional. Clearly a random variable is not a variable at all in the usual sense, and it is a function.

The sample space S is termed the *domain* of the r.v. X, and the collection of all numbers [values of $X(\zeta)$] is termed the *range* of the r.v. X. Thus the range of X is a certain subset of the set of all real numbers (Fig. 2-1).

Note that two or more different sample points might give the same value of $X(\zeta)$, but two different numbers in the range cannot be assigned to the same sample point.

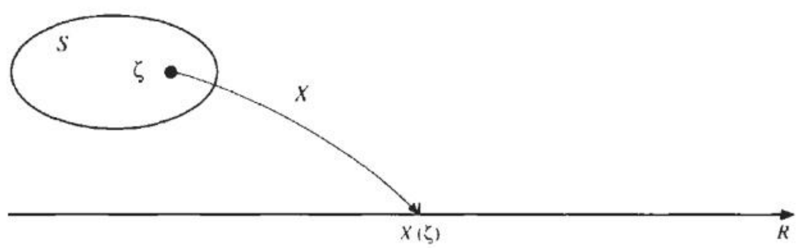

Fig. 2-1 Random variable X as a function.

EXAMPLE 2.1 In the experiment of tossing a coin once (Example 1.1), we might define the r.v. X as (Fig. 2-2)

$$X(H) = 1 \qquad X(T) = 0$$

Note that we could also define another r.v., say Y or Z, with

$$Y(H) = 0,\ Y(T) = 1 \quad \text{or} \quad Z(H) = 0,\ Z(T) = 0$$

B. Events Defined by Random Variables:

If X is a r.v. and x is a fixed real number, we can define the event $(X = x)$ as

$$(X = x) = \{\zeta: X(\zeta) = x\} \tag{2.1}$$

Similarly, for fixed numbers x, x_1, and x_2, we can define the following events:

$$\begin{aligned} (X \le x) &= \{\zeta: X(\zeta) \le x\} \\ (X > x) &= \{\zeta: X(\zeta) > x\} \\ (x_1 < X \le x_2) &= \{\zeta: x_1 < X(\zeta) \le x_2\} \end{aligned} \tag{2.2}$$

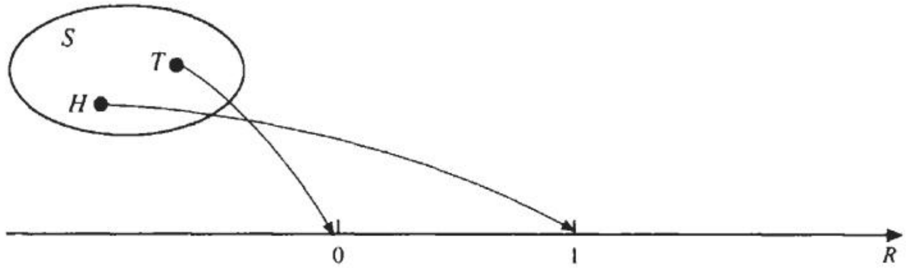

Fig. 2-2 One random variable associated with coin tossing.

These events have probabilities that are denoted by

$$\begin{aligned} P(X = x) &= P\{\zeta: X(\zeta) = x\} \\ P(X \le x) &= P\{\zeta: X(\zeta) \le x\} \\ P(X > x) &= P\{\zeta: X(\zeta) > x\} \\ P(x_1 < X \le x_2) &= P\{\zeta: x_1 < X(\zeta) \le x_2\} \end{aligned} \tag{2.3}$$

EXAMPLE 2.2 In the experiment of tossing a fair coin three times (Prob. 1.1), the sample space S_1 consists of eight equally likely sample points $S_1 = \{HHH, \ldots, TTT\}$. If X is the r.v. giving the number of heads obtained, find (*a*) $P(X = 2)$; (*b*) $P(X < 2)$.

(*a*) Let $A \subset S_1$ be the event defined by $X = 2$. Then, from Prob. 1.1, we have

$$A = (X = 2) = \{\zeta: X(\zeta) = 2\} = \{HHT, HTH, THH\}$$

Since the sample points are equally likely, we have

$$P(X = 2) = P(A) = \tfrac{3}{8}$$

(*b*) Let $B \subset S_1$ be the event defined by $X < 2$. Then

$$B = (X < 2) = \{\zeta: X(\zeta) < 2\} = \{HTT, THT, TTH, TTT\}$$

and
$$P(X < 2) = P(B) = \tfrac{4}{8} = \tfrac{1}{2}$$

2.3 DISTRIBUTION FUNCTIONS

A. Definition:

The *distribution function* [or *cumulative distribution function* (cdf)] of X is the function defined by

$$F_X(x) = P(X \le x) \qquad -\infty < x < \infty \tag{2.4}$$

Most of the information about a random experiment described by the r.v. X is determined by the behavior of $F_X(x)$.

B. Properties of $F_X(x)$:

Several properties of $F_X(x)$ follow directly from its definition (*2.4*).

1. $0 \le F_X(x) \le 1$ (*2.5*)
2. $F_X(x_1) \le F_X(x_2) \quad \text{if } x_1 < x_2$ (*2.6*)
3. $\lim\limits_{x \to \infty} F_X(x) = F_X(\infty) = 1$ (*2.7*)
4. $\lim\limits_{x \to -\infty} F_X(x) = F_X(-\infty) = 0$ (*2.8*)
5. $\lim\limits_{x \to a^+} F_X(x) = F_X(a^+) = F_X(a) \qquad a^+ = \lim\limits_{0 < \varepsilon \to 0} a + \varepsilon$ (*2.9*)

Property 1 follows because $F_X(x)$ is a probability. Property 2 shows that $F_X(x)$ is a nondecreasing function (Prob. 2.5). Properties 3 and 4 follow from Eqs. (*1.22*) and (*1.26*):

$$\lim_{x \to \infty} P(X \le x) = P(X \le \infty) = P(S) = 1$$

$$\lim_{x \to -\infty} P(X \le x) = P(X \le -\infty) = P(\varnothing) = 0$$

Property 5 indicates that $F_X(x)$ is *continuous on the right*. This is the consequence of the definition (*2.4*).

Table 2.1

x	$(X \le x)$	$F_X(x)$
-1	$\varnothing$	0
0	(TTT)	$\frac{1}{8}$
1	$\{TTT, TTH, THT, HTT\}$	$\frac{4}{8} = \frac{1}{2}$
2	$\{TTT, TTH, THT, HTT, HHT, HTH, THH\}$	$\frac{7}{8}$
3	S	1
4	S	1

EXAMPLE 2.3 Consider the r.v. X defined in Example 2.2. Find and sketch the cdf $F_X(x)$ of X.

Table 2.1 gives $F_X(x) = P(X \le x)$ for $x = -1, 0, 1, 2, 3, 4$. Since the value of X must be an integer, the value of $F_X(x)$ for noninteger values of x must be the same as the value of $F_X(x)$ for the nearest smaller integer value of x. The $F_X(x)$ is sketched in Fig. 2-3. Note that $F_X(x)$ has jumps at $x = 0, 1, 2, 3$, and that at each jump the upper value is the correct value for $F_X(x)$.

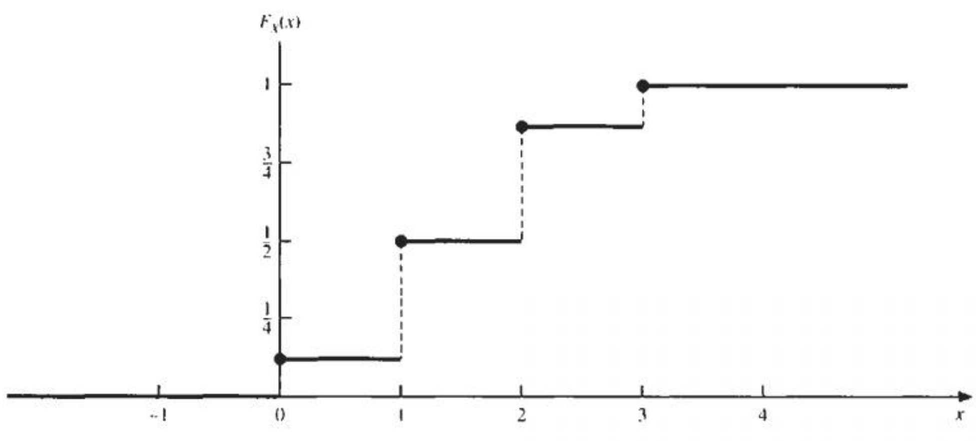

Fig. 2-3

C. Determination of Probabilities from the Distribution Function:

From definition (*2.4*), we can compute other probabilities, such as $P(a < X \le b)$, $P(X > a)$, and $P(X < b)$ (Prob. 2.6):

$$P(a < X \le b) = F_X(b) - F_X(a) \tag{2.10}$$

$$P(X > a) = 1 - F_X(a) \tag{2.11}$$

$$P(X < b) = F_X(b^-) \qquad b^- = \lim_{0 < \varepsilon \to 0} b - \varepsilon \tag{2.12}$$

2.4 DISCRETE RANDOM VARIABLES AND PROBABILITY MASS FUNCTIONS

A. Definition:

Let X be a r.v. with cdf $F_X(x)$. If $F_X(x)$ changes values only in jumps (at most a countable number of them) and is constant between jumps—that is, $F_X(x)$ is a staircase function (see Fig. 2-3)— then X is called a *discrete* random variable. Alternatively, X is a discrete r.v. only if its range contains a finite or countably infinite number of points. The r.v. X in Example 2.3 is an example of a discrete r.v.

B. Probability Mass Functions:

Suppose that the jumps in $F_X(x)$ of a discrete r.v. X occur at the points $x_1, x_2, \ldots$, where the sequence may be either finite or countably infinite, and we assume $x_i < x_j$ if $i < j$.

Then
$$F_X(x_i) - F_X(x_{i-1}) = P(X \le x_i) - P(X \le x_{i-1}) = P(X = x_i) \tag{2.13}$$

Let
$$p_X(x) = P(X = x) \tag{2.14}$$

The function $p_X(x)$ is called the *probability mass function* (pmf) of the discrete r.v. X.

Properties of $p_X(x)$:

1. $0 \le p_X(x_k) \le 1 \qquad k = 1, 2, \ldots$ (2.15)
2. $p_X(x) = 0 \qquad \text{if } x \ne x_k \ (k = 1, 2, \ldots)$ (2.16)
3. $\sum_k p_X(x_k) = 1$ (2.17)

The cdf $F_X(x)$ of a discrete r.v. X can be obtained by

$$F_X(x) = P(X \le x) = \sum_{x_k \le x} p_X(x_k) \tag{2.18}$$

2.5 CONTINUOUS RANDOM VARIABLES AND PROBABILITY DENSITY FUNCTIONS

A. Definition:

Let X be a r.v. with cdf $F_X(x)$. If $F_X(x)$ is continuous and also has a derivative $dF_X(x)/dx$ which exists everywhere except at possibly a finite number of points and is piecewise continuous, then X is called a *continuous* random variable. Alternatively, X is a continuous r.v. only if its range contains an interval (either finite or infinite) of real numbers. Thus, if X is a continuous r.v., then (Prob. 2.18)

$$P(X = x) = 0 \tag{2.19}$$

Note that this is an example of an event with probability 0 that is not necessarily the impossible event $\varnothing$.

In most applications, the r.v. is either discrete or continuous. But if the cdf $F_X(x)$ of a r.v. X possesses features of both discrete and continuous r.v.'s, then the r.v. X is called the *mixed* r.v. (Prob. 2.10).

B. Probability Density Functions:

Let
$$f_X(x) = \frac{dF_X(x)}{dx} \tag{2.20}$$

The function $f_X(x)$ is called the *probability density function* (pdf) of the continuous r.v. X.

Properties of $f_X(x)$:

1. $f_X(x) \geq 0$ (2.21)

2. $\displaystyle\int_{-\infty}^{\infty} f_X(x)\, dx = 1$ (2.22)

3. $f_X(x)$ is piecewise continuous.

4. $\displaystyle P(a < X \leq b) = \int_a^b f_X(x)\, dx$ (2.23)

The cdf $F_X(x)$ of a continuous r.v. X can be obtained by

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(\xi)\, d\xi \tag{2.24}$$

By Eq. (2.19), if X is a continuous r.v., then

$$\begin{aligned} P(a < X \leq b) &= P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b) \\ &= \int_a^b f_X(x)\, dx = F_X(b) - F_X(a) \end{aligned} \tag{2.25}$$

2.6 MEAN AND VARIANCE

A. Mean:

The *mean* (or *expected value*) of a r.v. X, denoted by μ_X or $E(X)$, is defined by

$$\mu_X = E(X) = \begin{cases} \sum_k x_k\, p_X(x_k) & X\text{: discrete} \\ \int_{-\infty}^{\infty} x f_X(x)\, dx & X\text{: continuous} \end{cases} \tag{2.26}$$

B. Moment:

The nth *moment* of a r.v. X is defined by

$$E(X^n) = \begin{cases} \sum_k x_k{}^n p_X(x_k) & X\text{: discrete} \\ \int_{-\infty}^{\infty} x^n f_X(x)\, dx & X\text{: continuous} \end{cases} \tag{2.27}$$

Note that the mean of X is the first moment of X.

C. Variance:

The *variance* of a r.v. X, denoted by $\sigma_X{}^2$ or $\text{Var}(X)$, is defined by

$$\sigma_X{}^2 = \text{Var}(X) = E\{[X - E(X)]^2\} \tag{2.28}$$

Thus,

$$\sigma_X{}^2 = \begin{cases} \sum_k (x_k - \mu_X)^2 p_X(x_k) & X\text{: discrete} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx & X\text{: continuous} \end{cases} \tag{2.29}$$

Note from definition (*2.28*) that

$$\text{Var}(X) \geq 0 \tag{2.30}$$

The *standard deviation* of a r.v. X, denoted by σ_X, is the positive square root of Var(X).
Expanding the right-hand side of Eq. (*2.28*), we can obtain the following relation:

$$\text{Var}(X) = E(X^2) - [E(X)]^2 \tag{2.31}$$

which is a useful formula for determining the variance.

2.7 SOME SPECIAL DISTRIBUTIONS

In this section we present some important special distributions.

A. Bernoulli Distribution:

A r.v. X is called a *Bernoulli* r.v. with parameter p if its pmf is given by

$$p_X(k) = P(X = k) = p^k(1 - p)^{1-k} \qquad k = 0, 1 \tag{2.32}$$

where $0 \leq p \leq 1$. By Eq. (*2.18*), the cdf $F_X(x)$ of the Bernoulli r.v. X is given by

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \tag{2.33}$$

Figure 2-4 illustrates a Bernoulli distribution.

Fig. 2-4 Bernoulli distribution.

The mean and variance of the Bernoulli r.v. X are

$$\mu_X = E(X) = p \tag{2.34}$$
$$\sigma_X{}^2 = \text{Var}(X) = p(1 - p) \tag{2.35}$$

A Bernoulli r.v. X is associated with some experiment where an outcome can be classified as either a "success" or a "failure," and the probability of a success is p and the probability of a failure is $1 - p$. Such experiments are often called *Bernoulli trials* (Prob. 1.61).

B. Binomial Distribution:

A r.v. X is called a *binomial* r.v. with parameters (n, p) if its pmf is given by

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \qquad k = 0, 1, \ldots, n \tag{2.36}$$

where $0 \le p \le 1$ and

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

which is known as the binomial coefficient. The corresponding cdf of X is

$$F_X(x) = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} \qquad n \le x < n + 1 \tag{2.37}$$

Figure 2-5 illustrates the binomial distribution for $n = 6$ and $p = 0.6$.

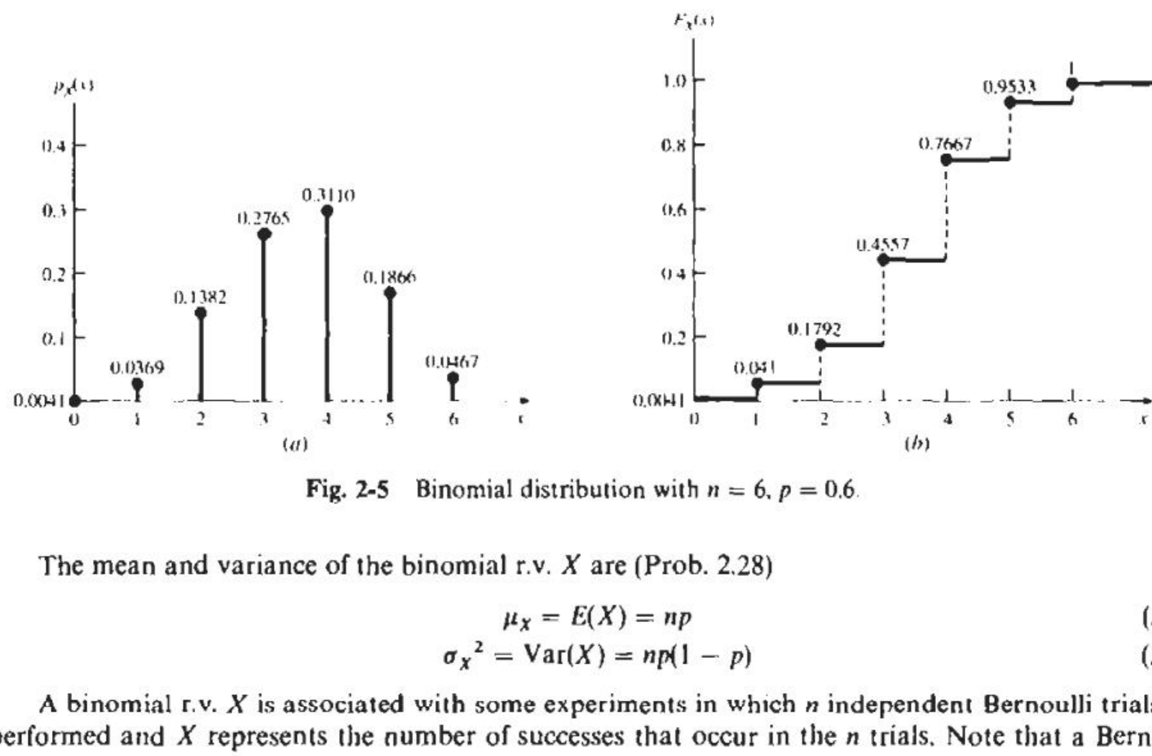

Fig. 2-5 Binomial distribution with $n = 6$, $p = 0.6$.

The mean and variance of the binomial r.v. X are (Prob. 2.28)

$$\mu_X = E(X) = np \tag{2.38}$$

$$\sigma_X{}^2 = \mathrm{Var}(X) = np(1 - p) \tag{2.39}$$

A binomial r.v. X is associated with some experiments in which n independent Bernoulli trials are performed and X represents the number of successes that occur in the n trials. Note that a Bernoulli r.v. is just a binomial r.v. with paramcters $(1, p)$.

C. Poisson Distribution:

A r.v. X is called a *Poisson* r.v. with parameter λ (>0) if its pmf is given by

$$p_X(k) = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad k = 0, 1, \ldots \tag{2.40}$$

The corresponding cdf of X is

$$F_X(x) = e^{-\lambda} \sum_{k=0}^{n} \frac{\lambda^k}{k!} \qquad n \le x < n + 1 \tag{2.41}$$

Figure 2-6 illustrates the Poisson distribution for $\lambda = 3$.

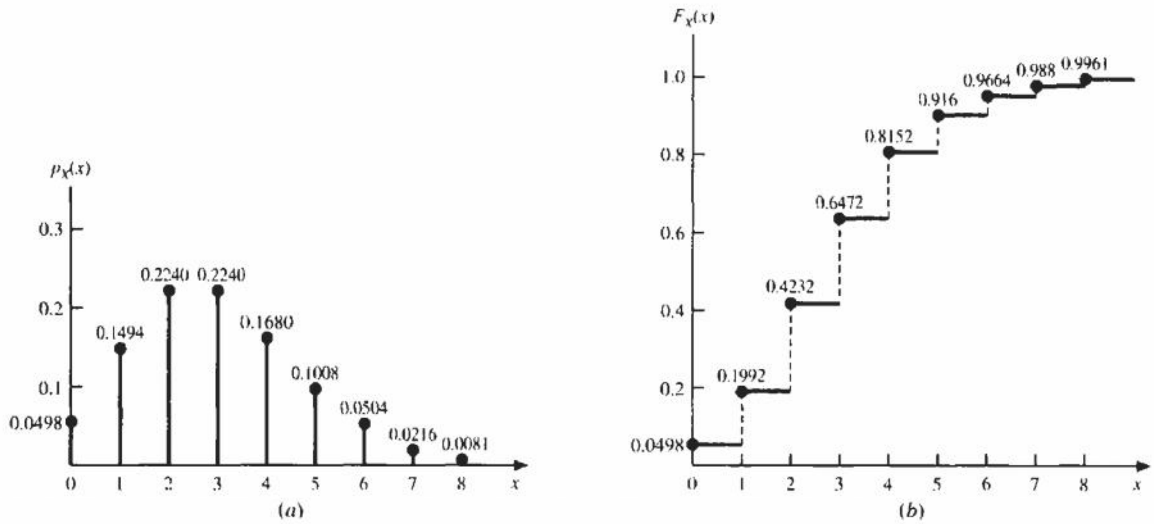

Fig. 2-6 Poisson distribution with $\lambda = 3$.

The mean and variance of the Poisson r.v. X are (Prob. 2.29)

$$\mu_X = E(X) = \lambda \tag{2.42}$$

$$\sigma_X{}^2 = \text{Var}(X) = \lambda \tag{2.43}$$

The Poisson r.v. has a tremendous range of applications in diverse areas because it may be used as an approximation for a binomial r.v. with parameters (n, p) when n is large and p is small enough so that np is of a moderate size (Prob. 2.40).

Some examples of Poisson r.v.'s include

1. The number of telephone calls arriving at a switching center during various intervals of time
2. The number of misprints on a page of a book
3. The number of customers entering a bank during various intervals of time

D. Uniform Distribution:

A r.v. X is called a *uniform* r.v. over (a, b) if its pdf is given by

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases} \tag{2.44}$$

The corresponding cdf of X is

$$F_X(x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases} \tag{2.45}$$

Figure 2-7 illustrates a uniform distribution.

The mean and variance of the uniform r.v. X are (Prob. 2.31)

$$\mu_X = E(X) = \frac{a+b}{2} \tag{2.46}$$

$$\sigma_X{}^2 = \text{Var}(X) = \frac{(b-a)^2}{12} \tag{2.47}$$

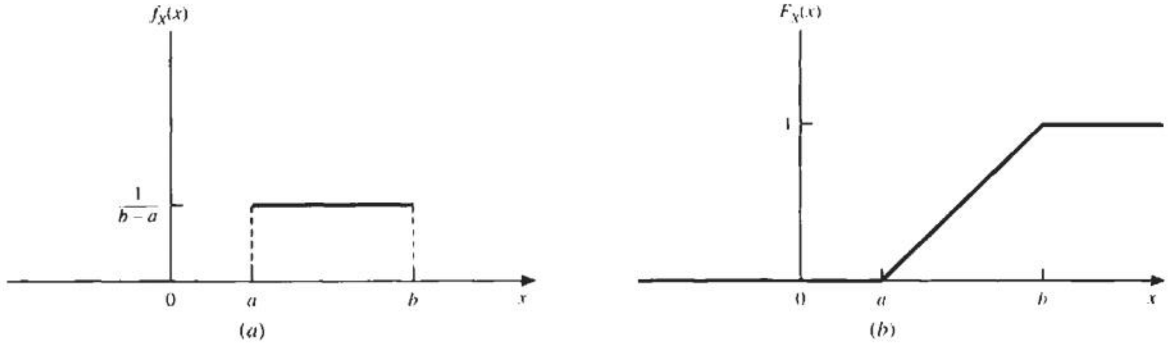

Fig. 2-7 Uniform distribution over (a, b).

A uniform r.v. X is often used where we have no prior knowledge of the actual pdf and all continuous values in some range seem equally likely (Prob. 2.69).

E. Exponential Distribution:

A r.v. X is called an *exponential* r.v. with parameter λ (>0) if its pdf is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x < 0 \end{cases} \tag{2.48}$$

which is sketched in Fig. 2-8(*a*). The corresponding cdf of X is

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{2.49}$$

which is sketched in Fig. 2-8(*b*).

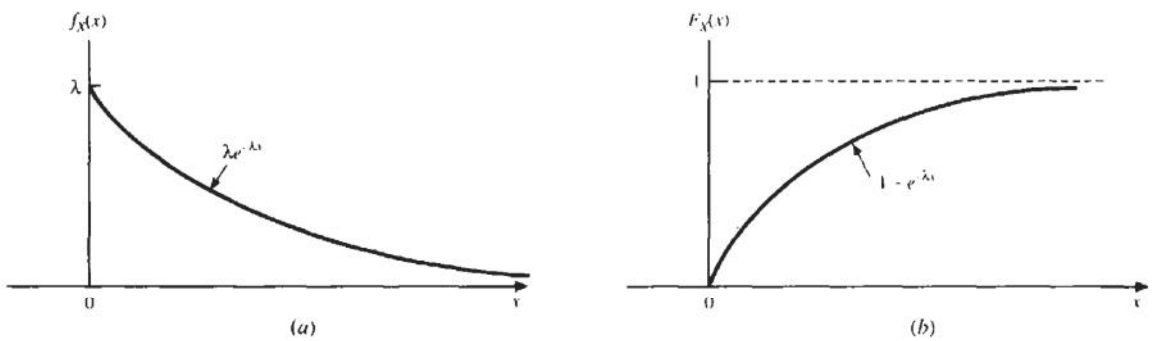

Fig. 2-8 Exponential distribution.

The mean and variance of the exponential r.v. X are (Prob. 2.32)

$$\mu_X = E(X) = \frac{1}{\lambda} \tag{2.50}$$

$$\sigma_X{}^2 = \text{Var}(X) = \frac{1}{\lambda^2} \tag{2.51}$$

The most interesting property of the exponential distribution is its "memoryless" property. By this we mean that if the lifetime of an item is exponentially distributed, then an item which has been in use for some hours is as good as a new item with regard to the amount of time remaining until the item fails. The exponential distribution is the only distribution which possesses this property (Prob. 2.53).

F. Normal (or Gaussian) Distribution:

A r.v. X is called a *normal* (or *gaussian*) r.v. if its pdf is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \tag{2.52}$$

The corresponding cdf of X is

$$F_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-(\xi-\mu)^2/(2\sigma^2)}\, d\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-\xi^2/2}\, d\xi \tag{2.53}$$

This integral cannot be evaluated in a closed form and must be evaluated numerically. It is convenient to use the function $\Phi(z)$, defined as

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\xi^2/2}\, d\xi \tag{2.54}$$

to help us to evaluate the value of $F_X(x)$. Then Eq. (*2.53*) can be written as

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{2.55}$$

Note that

$$\Phi(-z) = 1 - \Phi(z) \tag{2.56}$$

The function $\Phi(z)$ is tabulated in Table A (Appendix A). Figure 2-9 illustrates a normal distribution.

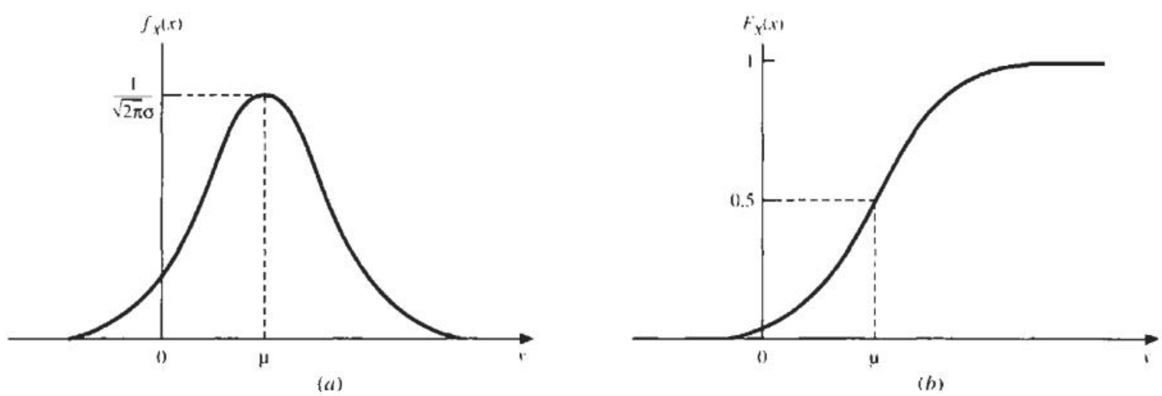

Fig. 2-9 Normal distribution.

The mean and variance of the normal r.v. X are (Prob. 2.33)

$$\mu_X = E(X) = \mu \tag{2.57}$$

$$\sigma_X{}^2 = \text{Var}(X) = \sigma^2 \tag{2.58}$$

We shall use the notation $N(\mu;\, \sigma^2)$ to denote that X is normal with mean μ and variance σ^2. A normal r.v. Z with zero mean and unit variance—that is, $Z = N(0;\, 1)$—is called a *standard normal* r.v. Note that the cdf of the standard normal r.v. is given by Eq. (*2.54*). The normal r.v. is probably the most important type of continuous r.v. It has played a significant role in the study of random phenomena in nature. Many naturally occurring random phenomena are approximately normal. Another reason for the importance of the normal r.v. is a remarkable theorem called the *central limit theorem*. This theorem states that the sum of a large number of independent r.v.'s, under certain conditions, can be approximated by a normal r.v. (see Sec. 4.8C).

2.8 CONDITIONAL DISTRIBUTIONS

In Sec. 1.6 the conditional probability of an event A given event B is defined as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) > 0$$

The conditional cdf $F_X(x \mid B)$ of a r.v. X given event B is defined by

$$F_X(x \mid B) = P(X \leq x \mid B) = \frac{P\{(X \leq x) \cap B\}}{P(B)} \tag{2.59}$$

The conditional cdf $F_X(x \mid B)$ has the same properties as $F_X(x)$. (See Prob. 1.37 and Sec. 2.3.) In particular,

$$F_X(-\infty \mid B) = 0 \qquad F_X(\infty \mid B) = 1 \tag{2.60}$$

$$P(a < X \leq b \mid B) = F_X(b \mid B) - F_X(a \mid B) \tag{2.61}$$

If X is a discrete r.v., then the conditional pmf $p_X(x_k \mid B)$ is defined by

$$p_X(x_k \mid B) = P(X = x_k \mid B) = \frac{P\{(X = x_k) \cap B\}}{P(B)} \tag{2.62}$$

If X is a continuous r.v., then the conditional pdf $f_X(x \mid B)$ is defined by

$$f_X(x \mid B) = \frac{dF_X(x \mid B)}{dx} \tag{2.63}$$

Solved Problems

RANDOM VARIABLES

2.1. Consider the experiment of throwing a fair die. Let X be the r.v. which assigns 1 if the number that appears is even and 0 if the number that appears is odd.

(*a*) What is the range of X?

(*b*) Find $P(X = 1)$ and $P(X = 0)$.

The sample space S on which X is defined consists of 6 points which are equally likely:

$$S = \{1, 2, 3, 4, 5, 6\}$$

(*a*) The range of X is $R_X = \{0, 1\}$.

(*b*) $(X = 1) = \{2, 4, 6\}$. Thus, $P(X = 1) = \frac{3}{6} = \frac{1}{2}$. Similarly, $(X = 0) = \{1, 3, 5\}$, and $P(X = 0) = \frac{1}{2}$.

2.2. Consider the experiment of tossing a coin three times (Prob. 1.1). Let X be the r.v. giving the number of heads obtained. We assume that the tosses are independent and the probability of a head is p.

(*a*) What is the range of X?

(*b*) Find the probabilities $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, and $P(X = 3)$.

The sample space S on which X is defined consists of eight sample points (Prob. 1.1):

$$S = \{HHH, HHT, \ldots, TTT\}$$

(*a*) The range of X is $R_X = \{0, 1, 2, 3\}$.