3 Discrete Random Variables and Probability Distributions

CHAPTER OUTLINE

LEARNING OBJECTIVES

After careful study of this chapter you should be able to do the following:

1. Determine probabilities from probability mass functions and the reverse
2. Determine probabilities from cumulative distribution functions and cumulative distribution functions from probability mass functions, and the reverse
3. Calculate means and variances for discrete random variables
4. Understand the assumptions for each of the discrete probability distributions presented
5. Select an appropriate discrete probability distribution to calculate probabilities in specific applications
6. Calculate probabilities, determine means and variances for each of the discrete probability distributions presented

3-1 Discrete Random Variables

Many physical systems can be modeled by the same or similar random experiments and random variables. The distribution of the random variables involved in each of these common systems can be analyzed, and the results of that analysis can be used in different applications and examples. In this chapter, we present the analysis of several random experiments and **discrete random variables** that frequently arise in applications. We often omit a discussion of the underlying sample space of the random experiment and directly describe the distribution of a particular random variable.

3-1 Discrete Random Variables

Example 3-1

A voice communication system for a business contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used. Let the random variable X denote the number of lines in use. Then, X can assume any of the integer values 0 through 48. When the system is observed, if 10 lines are in use, $x = 10$.

3-2 Probability Distributions and Probability Mass Functions

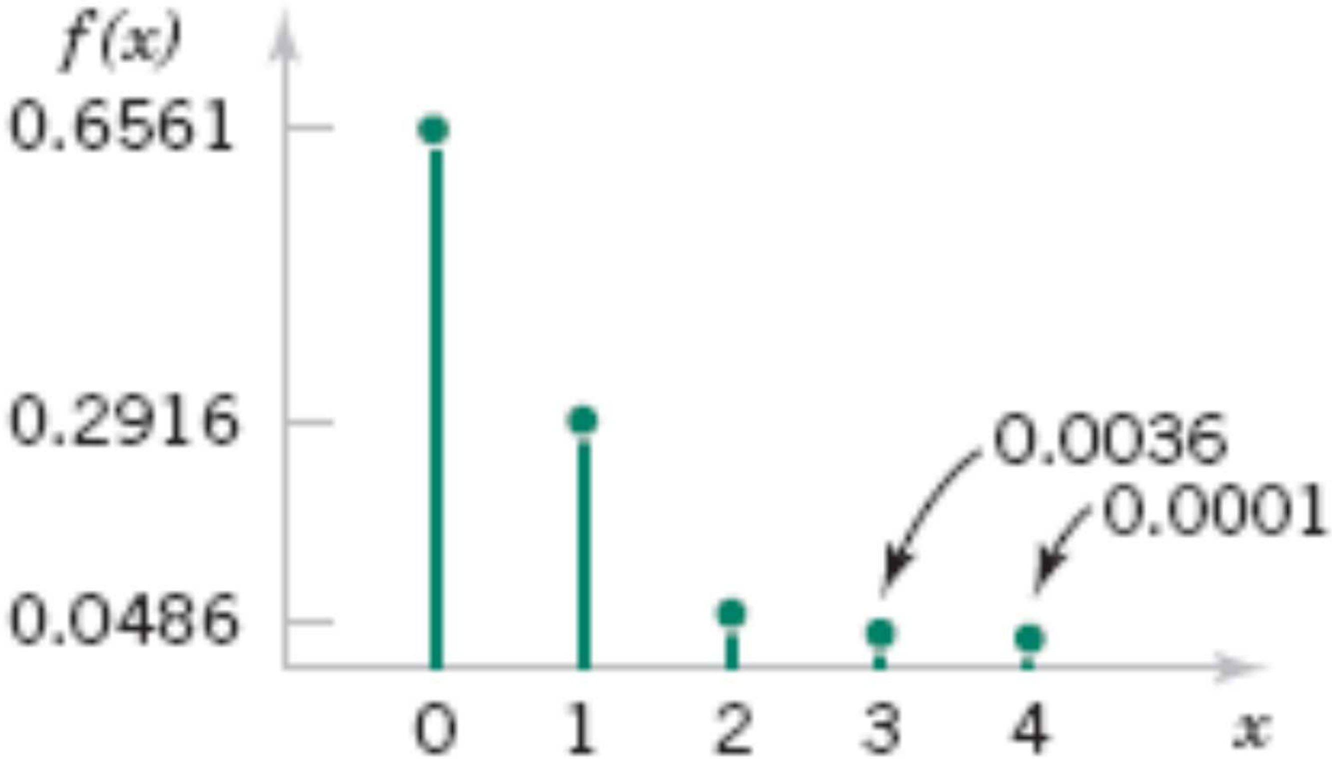

Figure 3-1 Probability distribution for bits in error.

3-2 Probability Distributions and Probability Mass Functions

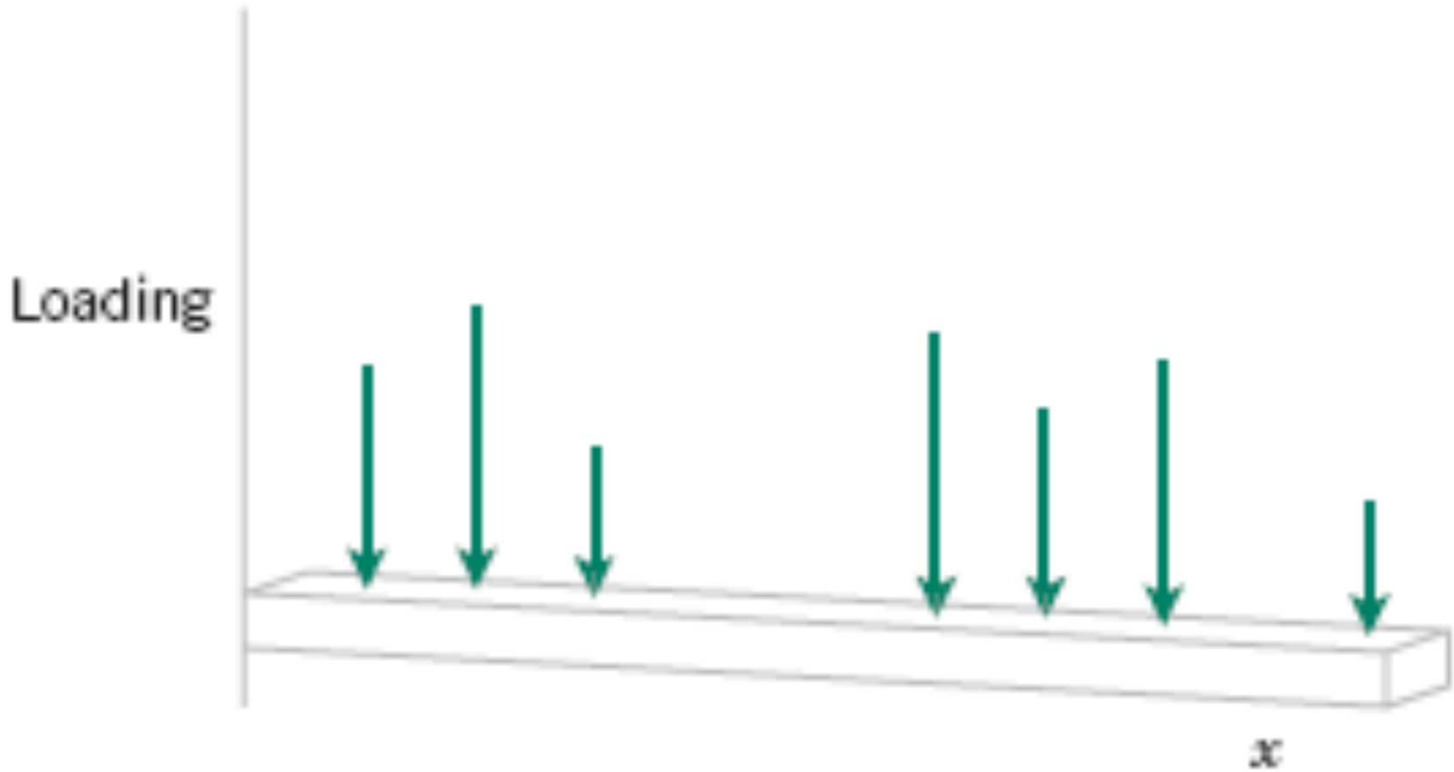

Figure 3-2 Loadings at discrete points on a long, thin beam.

3-2 Probability Distributions and Probability Mass Functions

Definition

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a **probability mass function** is a function such that

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $f(x_i) = P(X = x_i)$ (3-1)

Example 3-5

Let the random variable X denote the number of semiconductor wafers that need to be analyzed in order to detect a large particle of contamination. Assume that the probability that a wafer contains a large particle is 0.01 and that the wafers are independent. Determine the probability distribution of X.

Let p denote a wafer in which a large particle is present, and let a denote a wafer in which it is absent. The sample space of the experiment is infinite, and it can be represented as all possible sequences that start with a string of a's and end with p. That is,

$$s = \{p, ap, aap, aaap, aaaap, aaaaap, \text{and so forth}\}$$

Consider a few special cases. We have $P(X = 1) = P(p) = 0.01$. Also, using the independence assumption

$$P(X = 2) = P(ap) = 0.99(0.01) = 0.0099$$

Example 3-5 (continued)

A general formula is

$$P(X = x) = \underbrace{P(aa \ldots ap)}_{(x-1)a\text{'s}} = 0.99^{x-1}(0.01), \qquad \text{for } x = 1, 2, 3, \ldots$$

Describing the probabilities associated with X in terms of this formula is the simplest method of describing the distribution of X in this example. Clearly $f(x) \geq 0$. The fact that the sum of the probabilities is 1 is left as an exercise. This is an example of a geometric random variable, and details are provided later in this chapter.

3-3 Cumulative Distribution Functions

Definition

The **cumulative distribution function** of a discrete random variable X, denoted as $F(x)$, is

$$F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

(1) $F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i)$

(2) $0 \le F(x) \le 1$

(3) If $x \le y$, then $F(x) \le F(y)$ (3-2)

Example 3-8

Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement, from the batch. Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

The question can be answered by first finding the probability mass function of X.

$$P(X = 0) = \frac{800}{850} \cdot \frac{799}{849} = 0.886$$

$$P(X = 1) = 2 \cdot \frac{800}{850} \cdot \frac{50}{849} = 0.111$$

$$P(X = 2) = \frac{50}{850} \cdot \frac{49}{849} = 0.003$$

Therefore,

$$F(0) = P(X \leq 0) = 0.886$$

$$F(1) = P(X \leq 1) = 0.886 + 0.111 = 0.997$$

$$F(2) = P(X \leq 2) = 1$$

The cumulative distribution function for this example is graphed in Fig. 3-4. Note that $F(x)$ is defined for all x from $-\infty < x < \infty$ and not only for 0, 1, and 2.

Example 3-8

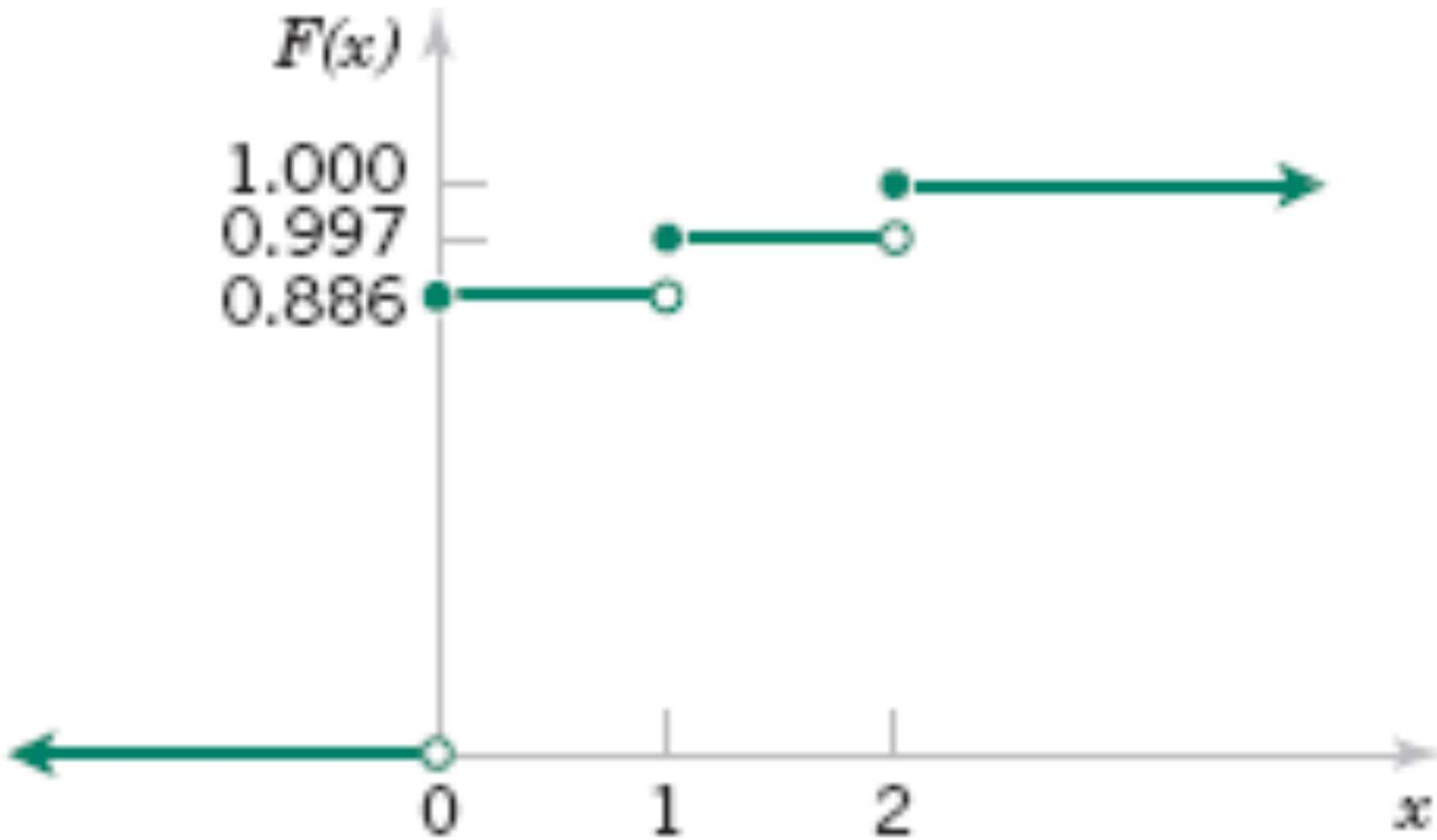

Figure 3-4 Cumulative distribution function for Example 3-8.

3-4 Mean and Variance of a Discrete Random Variable

Definition

The **mean** or **expected value** of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_x x f(x) \tag{3-3}$$

The **variance** of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X - \mu)^2 = \sum_x (x - \mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

The **standard deviation** of X is $\sigma = \sqrt{\sigma^2}$.

3-4 Mean and Variance of a Discrete Random Variable

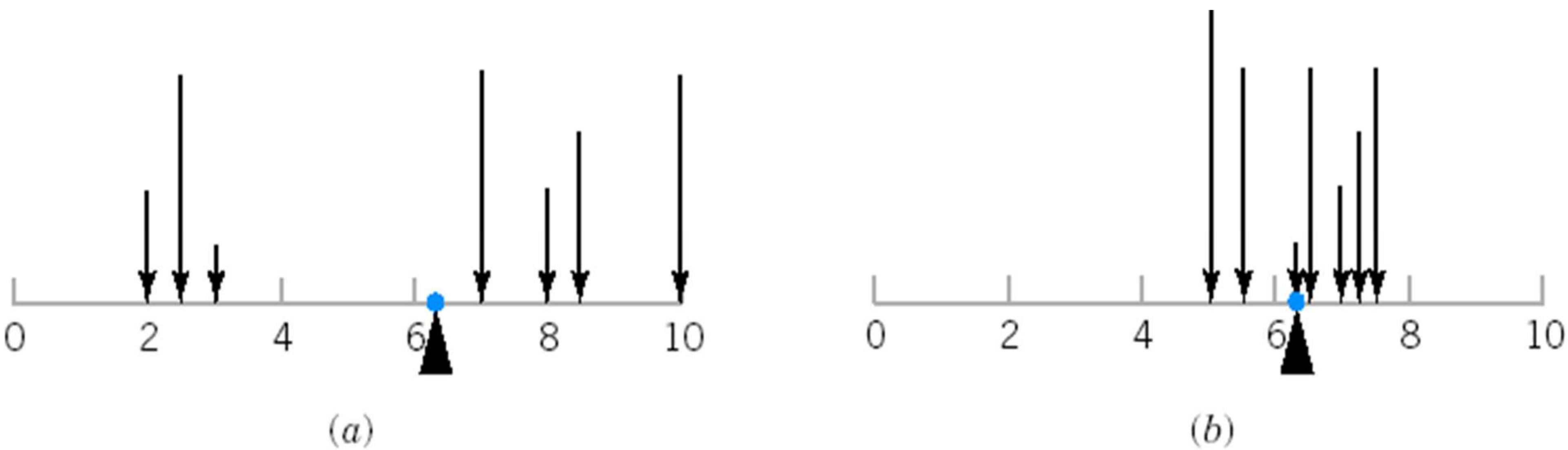

Figure 3-5 A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

3-4 Mean and Variance of a Discrete Random Variable

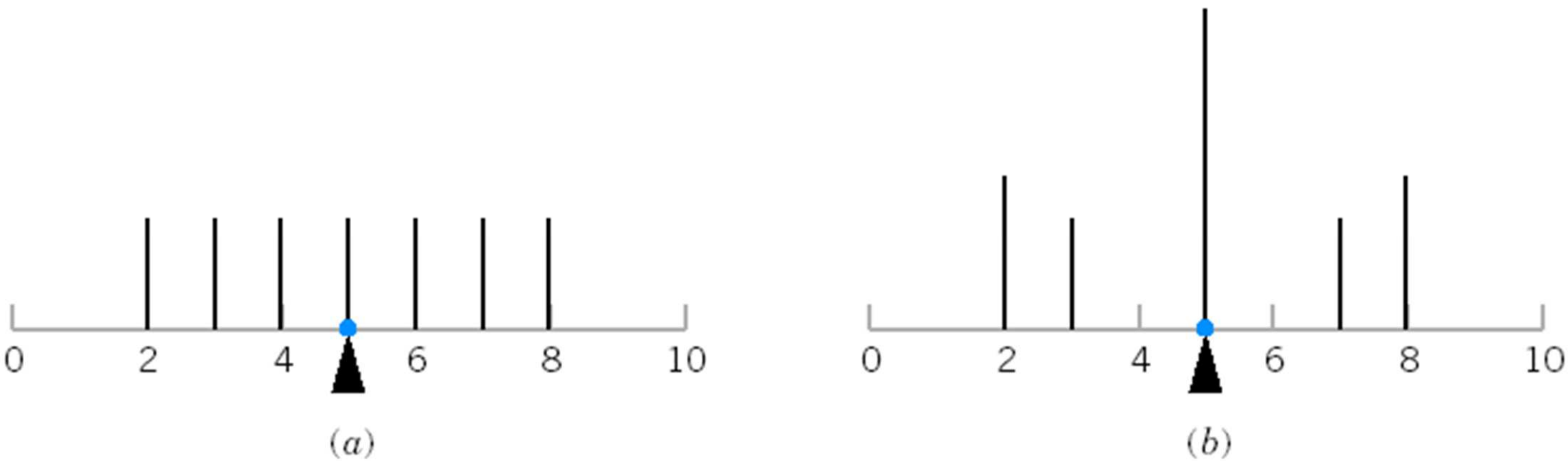

Figure 3-6 The probability distribution illustrated in Parts (a) and (b) differ even though they have equal means and equal variances.

Example 3-11

The number of messages sent per hour over a computer network has the following distribution:

x = number of messages	10	11	12	13	14	15
$f(x)$	0.08	0.15	0.30	0.20	0.20	0.07

Determine the mean and standard deviation of the number of messages sent per hour.

$$E(X) = 10(0.08) + 11(0.15) + \cdots + 15(0.07) = 12.5$$

$$V(X) = 10^2(0.08) + 11^2(0.15) + \cdots + 15^2(0.07) - 12.5^2 = 1.85$$

$$\sigma = \sqrt{V(X)} = \sqrt{1.85} = 1.36$$

3-4 Mean and Variance of a Discrete Random Variable

Expected Value of a Function of a Discrete Random Variable

If X is a discrete random variable with probability mass function $f(x)$,

$$E[h(X)] = \sum_x h(x)f(x) \qquad (3\text{-}4)$$

3-5 Discrete Uniform Distribution

Definition

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$f(x_i) = 1/n \tag{3-5}$$

3-5 Discrete Uniform Distribution

Example 3-13

The first digit of a part's serial number is equally likely to be any one of the digits 0 through 9. If one part is selected from a large batch and X is the first digit of the serial number, X has a discrete uniform distribution with probability 0.1 for each value in $R = \{0, 1, 2, \ldots, 9\}$. That is,

$$f(x) = 0.1$$

for each value in R. The probability mass function of X is shown in Fig. 3-7.

3-5 Discrete Uniform Distribution

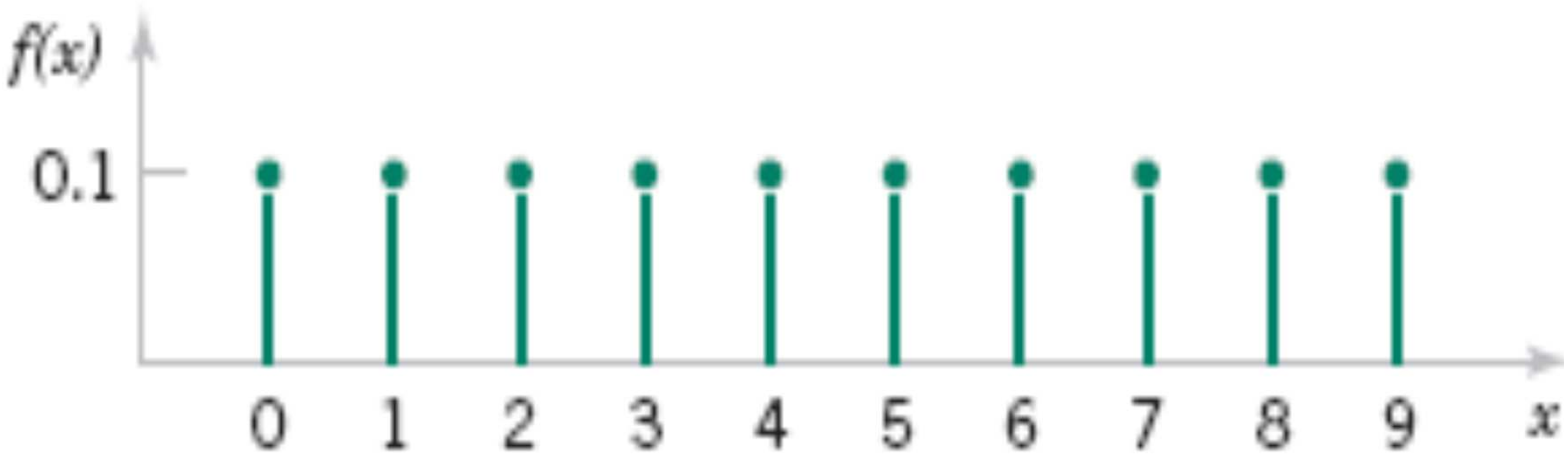

Figure 3-7 Probability mass function for a discrete uniform random variable.

3-5 Discrete Uniform Distribution

Mean and Variance

Suppose X is a discrete uniform random variable on the consecutive integers $a, a+1, a+2, \ldots, b$, for $a \le b$. The mean of X is

$$\mu = E(X) = \frac{b+a}{2}$$

The variance of X is

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12} \tag{3-6}$$