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

Answers for most odd numbered exercises are at the end of the book. Answers to exercises whose numbers are surrounded by a box can be accessed in the e-Text by clicking on the box. Complete worked solutions to certain exercises are also available in the e-Text. These are indicated in the Answers to Selected Exercises section by a box around the exercise number. Exercises are also available for the text sections that appear on CD only. These exercises may be found within the e-Text immediately following the section they accompany.

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.

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, $x = 10$.

EXAMPLE 3-2 In a semiconductor manufacturing process, two wafers from a lot are tested. Each wafer is classified as *pass* or *fail*. Assume that the probability that a wafer passes the test is 0.8 and that wafers are independent. The sample space for the experiment and associated probabilities are shown in Table 3-1. For example, because of the independence, the probability of the outcome that the first wafer tested passes and the second wafer tested fails, denoted as *pf*, is

$$P(pf) = 0.8(0.2) = 0.16$$

The random variable X is defined to be equal to the number of wafers that pass. The last column of the table shows the values of X that are assigned to each outcome in the experiment.

EXAMPLE 3-3 Define the random variable X to be the number of contamination particles on a wafer in semiconductor manufacturing. Although wafers possess a number of characteristics, the random variable X summarizes the wafer only in terms of the number of particles.

The possible values of X are integers from zero up to some large value that represents the maximum number of particles that can be found on one of the wafers. If this maximum number is very large, we might simply assume that the range of X is the set of integers from zero to infinity.

Note that more than one random variable can be defined on a sample space. In Example 3-3, we might define the random variable Y to be the number of chips from a wafer that fail the final test.

Table 3-1 Wafer Tests

Outcome			
Wafer 1	Wafer 2	Probability	x
Pass	Pass	0.64	2
Fail	Pass	0.16	1
Pass	Fail	0.16	1
Fail	Fail	0.04	0

EXERCISES FOR SECTION 3-1

For each of the following exercises, determine the range (possible values) of the random variable.

3-1. The random variable is the number of nonconforming solder connections on a printed circuit board with 1000 connections.

3-2. In a voice communication system with 50 lines, the random variable is the number of lines in use at a particular time.

3-3. An electronic scale that displays weights to the nearest pound is used to weigh packages. The display shows only five digits. Any weight greater than the display can indicate is shown as 99999. The random variable is the displayed weight.

3-4. A batch of 500 machined parts contains 10 that do not conform to customer requirements. The random variable is the number of parts in a sample of 5 parts that do not conform to customer requirements.

3-5. A batch of 500 machined parts contains 10 that do not conform to customer requirements. Parts are selected successively, without replacement, until a nonconforming part is obtained. The random variable is the number of parts selected.

3-6. The random variable is the moisture content of a lot of raw material, measured to the nearest percentage point.

3-7. The random variable is the number of surface flaws in a large coil of galvanized steel.

3-8. The random variable is the number of computer clock cycles required to complete a selected arithmetic calculation.

3-9. An order for an automobile can select the base model or add any number of 15 options. The random variable is the number of options selected in an order.

3-10. Wood paneling can be ordered in thicknesses of 1/8, 1/4, or 3/8 inch. The random variable is the total thickness of paneling in two orders.

3-11. A group of 10,000 people are tested for a gene called Ifi202 that has been found to increase the risk for lupus. The random variable is the number of people who carry the gene.

3-12. A software program has 5000 lines of code. The random variable is the number of lines with a fatal error.

3-2 PROBABILITY DISTRIBUTIONS AND PROBABILITY MASS FUNCTIONS

Random variables are so important in random experiments that sometimes we essentially ignore the original sample space of the experiment and focus on the probability distribution of the random variable. For example, in Example 3-1, our analysis might focus exclusively on the integers $\{0, 1, \ldots, 48\}$ in the range of X. In Example 3-2, we might summarize the random experiment in terms of the three possible values of X, namely $\{0, 1, 2\}$. In this manner, a random variable can simplify the description and analysis of a random experiment.

The **probability distribution** of a random variable X is a description of the probabilities associated with the possible values of X. For a discrete random variable, the distribution is often specified by just a list of the possible values along with the probability of each. In some cases, it is convenient to express the probability in terms of a formula.

EXAMPLE 3-4

There is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equal the number of bits in error in the next four bits transmitted. The possible values for X are $\{0, 1, 2, 3, 4\}$. Based on a model for the errors that is presented in the following section, probabilities for these values will be determined. Suppose that the probabilities are

$$P(X = 0) = 0.6561 \qquad P(X = 1) = 0.2916 \qquad P(X = 2) = 0.0486$$
$$P(X = 3) = 0.0036 \qquad P(X = 4) = 0.0001$$

The probability distribution of X is specified by the possible values along with the probability of each. A graphical description of the probability distribution of X is shown in Fig. 3-1.

Suppose a loading on a long, thin beam places mass only at discrete points. See Fig. 3-2. The loading can be described by a function that specifies the mass at each of the discrete points. Similarly, for a discrete random variable X, its distribution can be described by a function that specifies the probability at each of the possible discrete values for X.

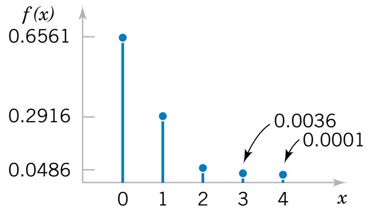


Figure 3-1 Probability distribution for bits in error.

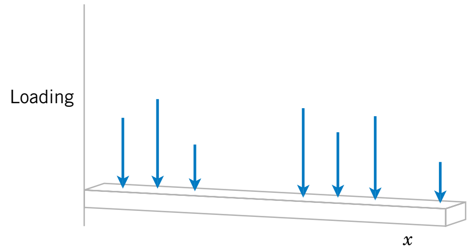


Figure 3-2 Loadings at discrete points on a long, thin beam.

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)

For example, in Example 3-4, $f(0) = 0.6561$, $f(1) = 0.2916$, $f(2) = 0.0486$, $f(3) = 0.0036$, and $f(4) = 0.0001$. Check that the sum of the probabilities in Example 3-4 is 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

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

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.

EXERCISES FOR SECTION 3-2

3-13. The sample space of a random experiment is $\{a, b, c, d, e, f\}$, and each outcome is equally likely. A random variable is defined as follows:

outcome	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

Determine the probability mass function of X.

3-14. Use the probability mass function in Exercise 3-11 to determine the following probabilities:
(a) $P(X = 1.5)$ (b) $P(0.5 < X < 2.7)$
(c) $P(X > 3)$ (d) $P(0 \le X < 2)$
(e) $P(X = 0 \quad \text{or} \quad X = 2)$

Verify that the following functions are probability mass functions, and determine the requested probabilities.

3-15.

x	-2	-1	0	1	2
$f(x)$	1/8	2/8	2/8	2/8	1/8

(a) $P(X \le 2)$ (b) $P(X > -2)$
(c) $P(-1 \le X \le 1)$ (d) $P(X \le -1 \quad \text{or} \quad X = 2)$

3-16. $f(x) = (8/7)(1/2)^x, \quad x = 1, 2, 3$
(a) $P(X \le 1)$ (b) $P(X > 1)$
(c) $P(2 < X < 6)$ (d) $P(X \le 1 \quad \text{or} \quad X > 1)$

3-17. $f(x) = \dfrac{2x + 1}{25}, \quad x = 0, 1, 2, 3, 4$
(a) $P(X = 4)$ (b) $P(X \le 1)$
(c) $P(2 \le X < 4)$ (d) $P(X > -10)$

3-18. $f(x) = (3/4)(1/4)^x, \quad x = 0, 1, 2, \ldots$
(a) $P(X = 2)$ (b) $P(X \le 2)$
(c) $P(X > 2)$ (d) $P(X \ge 1)$

3-19. Marketing estimates that a new instrument for the analysis of soil samples will be very successful, moderately successful, or unsuccessful, with probabilities 0.3, 0.6, and 0.1, respectively. The yearly revenue associated with a very successful, moderately successful, or unsuccessful product is \$10 million, \$5 million, and \$1 million, respectively. Let the random variable X denote the yearly revenue of the product. Determine the probability mass function of X.

3-20. A disk drive manufacturer estimates that in five years a storage device with 1 terabyte of capacity will sell with probability 0.5, a storage device with 500 gigabytes capacity will sell with a probability 0.3, and a storage device with 100 gigabytes capacity will sell with probability 0.2. The revenue associated with the sales in that year are estimated to be \$50 million, \$25 million, and \$10 million, respectively. Let X be the revenue of storage devices during that year. Determine the probability mass function of X.

3-21. An optical inspection system is to distinguish among different part types. The probability of a correct classification of any part is 0.98. Suppose that three parts are inspected and that the classifications are independent. Let the random variable X denote the number of parts that are correctly classified. Determine the probability mass function of X.

3-22. In a semiconductor manufacturing process, three wafers from a lot are tested. Each wafer is classified as *pass* or *fail*. Assume that the probability that a wafer passes the test is 0.8 and that wafers are independent. Determine the probability mass function of the number of wafers from a lot that pass the test.

3-23. The distributor of a machine for cytogenics has developed a new model. The company estimates that when it is introduced into the market, it will be very successful with a probability 0.6, moderately successful with a probability 0.3, and not successful with probability 0.1. The estimated yearly profit associated with the model being very successful is \$15 million and being moderately successful is \$5 million; not successful would result in a loss of \$500,000. Let X be the yearly profit of the new model. Determine the probability mass function of X.

3-24. An assembly consists of two mechanical components. Suppose that the probabilities that the first and second components meet specifications are 0.95 and 0.98. Assume that the components are independent. Determine the probability mass function of the number of components in the assembly that meet specifications.

3-25. An assembly consists of three mechanical components. Suppose that the probabilities that the first, second, and third components meet specifications are 0.95, 0.98, and 0.99. Assume that the components are independent. Determine the probability mass function of the number of components in the assembly that meet specifications.

3-3 CUMULATIVE DISTRIBUTION FUNCTIONS

EXAMPLE 3-6

In Example 3-4, we might be interested in the probability of three or fewer bits being in error. This question can be expressed as $P(X \le 3)$.

The event that $\{X \le 3\}$ is the union of the events $\{X = 0\}$, $\{X = 1\}$, $\{X = 2\}$, and

$\{X = 3\}$. Clearly, these three events are mutually exclusive. Therefore,

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.6561 + 0.2916 + 0.0486 + 0.0036 = 0.9999 \end{aligned}$$

This approach can also be used to determine

$$P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.0036$$

Example 3-6 shows that it is sometimes useful to be able to provide **cumulative probabilities** such as $P(X \leq x)$ and that such probabilities can be used to find the probability mass function of a random variable. Therefore, using cumulative probabilities is an alternate method of describing the probability distribution of a random variable.

In general, for any discrete random variable with possible values $x_1, x_2, \ldots, x_n$, the events $\{X = x_1), \{X = x_2), \ldots, \{X = x_n)$ are mutually exclusive. Therefore, $P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.

Definition

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

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

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

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

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

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

Like a probability mass function, a cumulative distribution function provides probabilities. Notice that even if the random variable X can only assume integer values, the cumulative distribution function can be defined at noninteger values. In Example 3-6, $F(1.5) = P(X \leq 1.5) = P\{X = 0\} + P(X = 1) = 0.6561 + 0.2916 = 0.9477$. Properties (1) and (2) of a cumulative distribution function follow from the definition. Property (3) follows from the fact that if $x \leq y$, the event that $\{X \leq x\}$ is contained in the event $\{X \leq y\}$.

The next example shows how the cumulative distribution function can be used to determine the probability mass function of a discrete random variable.

EXAMPLE 3-7

Determine the probability mass function of X from the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \leq x < 0 \\ 0.7 & 0 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

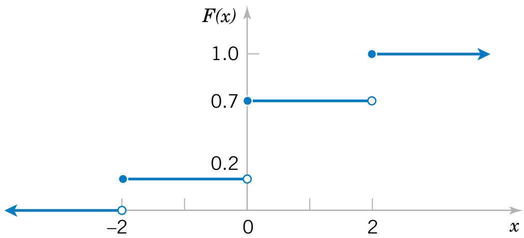


Figure 3-3 Cumulative distribution function for Example 3-7.

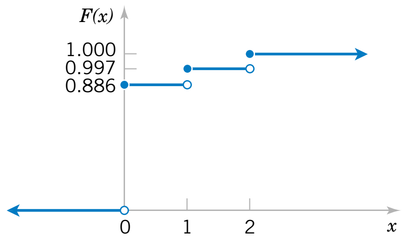


Figure 3-4 Cumulative distribution function for Example 3-8.

Figure 3-3 displays a plot of $F(x)$. From the plot, the only points that receive nonzero probability are -2, 0, and 2. The probability mass function at each point is the change in the cumulative distribution function at the point. Therefore,

$$f(-2) = 0.2 - 0 = 0.2 \qquad f(0) = 0.7 - 0.2 = 0.5 \qquad f(2) = 1.0 - 0.7 = 0.3$$

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,

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

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

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

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.

EXERCISES FOR SECTION 3-3

3-26. Determine the cumulative distribution function of the random variable in Exercise 3-13.

3-27. Determine the cumulative distribution function for the random variable in Exercise 3-15; also determine the following probabilities:
(a) $P(X \le 1.25)$ (b) $P(X \le 2.2)$
(c) $P(-1.1 < X \le 1)$ (d) $P(X > 0)$

3-28. Determine the cumulative distribution function for the random variable in Exercise 3-17; also determine the following probabilities:
(a) $P(X < 1.5)$ (b) $P(X \le 3)$
(c) $P(X > 2)$ (d) $P(1 < X \le 2)$

3-29. Determine the cumulative distribution function for the random variable in Exercise 3-19.

3-30. Determine the cumulative distribution function for the random variable in Exercise 3-20.

3-31. Determine the cumulative distribution function for the random variable in Exercise 3-22.

3-32. Determine the cumulative distribution function for the variable in Exercise 3-23.

Verify that the following functions are cumulative distribution functions, and determine the probability mass function and the requested probabilities.

3-33.

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.5 & 1 \le x < 3 \\ 1 & 3 \le x \end{cases}$$

(a) $P(X \le 3)$ (b) $P(X \le 2)$
(c) $P(1 \le X \le 2)$ (d) $P(X > 2)$

3-34. Errors in an experimental transmission channel are found when the transmission is checked by a certifier that detects missing pulses. The number of errors found in an eight-bit byte is a random variable with the following distribution:

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.7 & 1 \le x < 4 \\ 0.9 & 4 \le x < 7 \\ 1 & 7 \le x \end{cases}$$

Determine each of the following probabilities:
(a) $P(X \le 4)$ (b) $P(X > 7)$
(c) $P(X \le 5)$ (d) $P(X > 4)$
(e) $P(X \le 2)$

3-35.

$$F(x) = \begin{cases} 0 & x < -10 \\ 0.25 & -10 \le x < 30 \\ 0.75 & 30 \le x < 50 \\ 1 & 50 \le x \end{cases}$$

(a) $P(X \le 50)$ (b) $P(X \le 40)$
(c) $P(40 \le X \le 60)$ (d) $P(X < 0)$
(e) $P(0 \le X < 10)$ (f) $P(-10 < X < 10)$

3-36. The thickness of wood paneling (in inches) that a customer orders is a random variable with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < 1/8 \\ 0.2 & 1/8 \le x < 1/4 \\ 0.9 & 1/4 \le x < 3/8 \\ 1 & 3/8 \le x \end{cases}$$

Determine the following probabilities:
(a) $P(X \le 1/18)$ (b) $P(X \le 1/4)$
(c) $P(X \le 5/16)$ (d) $P(X > 1/4)$
(e) $P(X \le 1/2)$

3-4 MEAN AND VARIANCE OF A DISCRETE RANDOM VARIABLE

Two numbers are often used to summarize a probability distribution for a random variable X. The mean is a measure of the center or middle of the probability distribution, and the variance is a measure of the dispersion, or variability in the distribution. These two measures do not uniquely identify a probability distribution. That is, two different distributions can have the same mean and variance. Still, these measures are simple, useful summaries of the probability distribution of X.

Definition

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

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

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 $\sigma = \sqrt{\sigma^2}$.

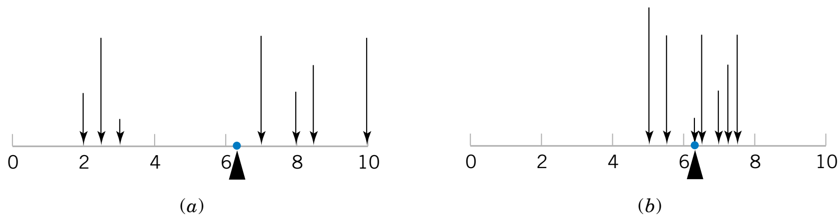


Figure 3-5 A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

The mean of a discrete random variable X is a weighted average of the possible values of X, with weights equal to the probabilities. If $f(x)$ is the probability mass function of a loading on a long, thin beam, $E(X)$ is the point at which the beam balances. Consequently, $E(X)$ describes the "center" of the distribution of X in a manner similar to the balance point of a loading. See Fig. 3-5.

The variance of a random variable X is a measure of dispersion or scatter in the possible values for X. The variance of X uses weight $f(x)$ as the multiplier of each possible squared deviation $(x - \mu)^2$. Figure 3-5 illustrates probability distributions with equal means but different variances. Properties of summations and the definition of μ can be used to show the equality of the formulas for variance.

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

$$= \sum_x x^2 f(x) - 2\mu^2 + \mu^2 = \sum_x x^2 f(x) - \mu^2$$

Either formula for $V(x)$ can be used. Figure 3-6 illustrates that two probability distributions can differ even though they have identical means and variances.

EXAMPLE 3-9

In Example 3-4, there is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equal the number of bits in error in the next four bits transmitted. The possible values for X are $\{0, 1, 2, 3, 4\}$. Based on a model for the errors that is presented in the following section, probabilities for these values will be determined. Suppose that the probabilities are

$$P(X = 0) = 0.6561 \qquad P(X = 2) = 0.0486 \qquad P(X = 4) = 0.0001$$
$$P(X = 1) = 0.2916 \qquad P(X = 3) = 0.0036$$

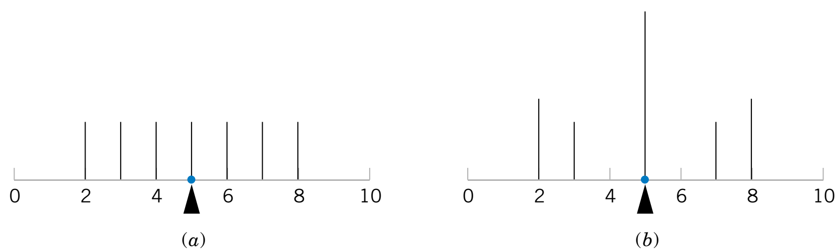


Figure 3-6 The probability distributions illustrated in Parts (a) and (b) differ even though they have equal means and equal variances.

Now

$$\begin{aligned}\mu = E(X) &= 0f(0) + 1f(1) + 2f(2) + 3f(3) + 4f(4)\\ &= 0(0.6561) + 1(0.2916) + 2(0.0486) + 3(0.0036) + 4(0.0001)\\ &= 0.4\end{aligned}$$

Although X never assumes the value 0.4, the weighted average of the possible values is 0.4.

To calculate $V(X)$, a table is convenient.

x	$x - 0.4$	$(x - 0.4)^2$	$f(x)$	$f(x)(x - 0.4)^2$
0	−0.4	0.16	0.6561	0.104976
1	0.6	0.36	0.2916	0.104976
2	1.6	2.56	0.0486	0.124416
3	2.6	6.76	0.0036	0.024336
4	3.6	12.96	0.0001	0.001296

$$V(X) = \sigma^2 = \sum_{i=1}^{5} f(x_i)(x_i - 0.4)^2 = 0.36$$

The alternative formula for variance could also be used to obtain the same result.

EXAMPLE 3-10

Two new product designs are to be compared on the basis of revenue potential. Marketing feels that the revenue from design A can be predicted quite accurately to be \$3 million. The revenue potential of design B is more difficult to assess. Marketing concludes that there is a probability of 0.3 that the revenue from design B will be \$7 million, but there is a 0.7 probability that the revenue will be only \$2 million. Which design do you prefer?

Let X denote the revenue from design A. Because there is no uncertainty in the revenue from design A, we can model the distribution of the random variable X as \$3 million with probability 1. Therefore, $E(X) = \$3$ million.

Let Y denote the revenue from design B. The expected value of Y in millions of dollars is

$$E(Y) = \$7(0.3) + \$2(0.7) = \$3.5$$

Because $E(Y)$ exceeds $E(X)$, we might prefer design B. However, the variability of the result from design B is larger. That is,

$$\begin{aligned}\sigma^2 &= (7 - 3.5)^2(0.3) + (2 - 3.5)^2(0.7)\\ &= 5.25 \text{ millions of dollars squared}\end{aligned}$$

Because the units of the variables in this example are millions of dollars, and because the variance of a random variable squares the deviations from the mean, the units of σ^2 are millions of dollars squared. These units make interpretation difficult.

Because the units of standard deviation are the same as the units of the random variable, the standard deviation σ is easier to interpret. In this example, we can summarize our results as "the average deviation of Y from its mean is \$2.29 million."

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.

$$E(X) = 10(0.08) + 11(0.15) + \cdots + 15(0.07) = 12.5$$
$$V(X) = 10^2(0.08) + 11^2(0.15) + \cdots + 15^2(0.07) - 12.5^2 = 1.85$$
$$\sigma = \sqrt{V(X)} = \sqrt{1.85} = 1.36$$

The variance of a random variable X can be considered to be the expected value of a specific function of X, namely, $h(X) = (X - \mu)^2$. In general, the expected value of any function $h(X)$ of a discrete random variable is defined in a similar manner.

Expected Value of a Function of a Discrete Random Variable

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

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

EXAMPLE 3-12

In Example 3-9, X is the number of bits in error in the next four bits transmitted. What is the expected value of the square of the number of bits in error? Now, $h(X) = X^2$. Therefore,

$$E[h(X)] = 0^2 \times 0.6561 + 1^2 \times 0.2916 + 2^2 \times 0.0486 + 3^2 \times 0.0036 + 4^2 \times 0.0001 = 0.52$$

In the previous example, the expected value of X^2 does not equal $E(X)$ squared. However, in the special case that $h(X) = aX + b$ for any constants a and b, $E[h(X)] = aE(X) + b$. This can be shown from the properties of sums in the definition in Equation 3-4.

EXERCISES FOR SECTION 3-4

3-37. If the range of X is the set $\{0, 1, 2, 3, 4\}$ and $P(X = x) = 0.2$ determine the mean and variance of the random variable.

3-38. Determine the mean and variance of the random variable in Exercise 3-13.

3-39. Determine the mean and variance of the random variable in Exercise 3-15.

3-40. Determine the mean and variance of the random variable in Exercise 3-17.

3-41. Determine the mean and variance of the random variable in Exercise 3-19.

3-42. Determine the mean and variance of the random variable in Exercise 3-20.

3-43. Determine the mean and variance of the random variable in Exercise 3-22.

3-44. Determine the mean and variance of the random variable in Exercise 3-23.

3-45. The range of the random variable X is $[0, 1, 2, 3, x]$, where x is unknown. If each value is equally likely and the mean of X is 6, determine x.

3-5 DISCRETE UNIFORM DISTRIBUTION

The simplest discrete random variable is one that assumes only a finite number of possible values, each with equal probability. A random variable X that assumes each of the values $x_1, x_2, \ldots, x_n$, with equal probability $1/n$, is frequently of interest.

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,

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

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.

Suppose the range of the discrete random variable X is the consecutive integers a, $a + 1, a + 2, \ldots, b$, for $a \leq b$. The range of X contains $b - a + 1$ values each with probability $1/(b - a + 1)$. Now,

$$\mu = \sum_{k=a}^{b} k\left(\frac{1}{b - a + 1}\right)$$

The algebraic identity $\sum_{k=a}^{b} k = \frac{b(b + 1) - (a - 1)a}{2}$ can be used to simplify the result to $\mu = (b + a)/2$. The derivation of the variance is left as an exercise.

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

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

The variance of X is

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

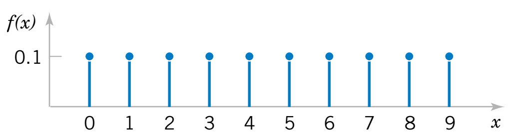


Figure 3-7 Probability mass function for a discrete uniform random variable.

EXAMPLE 3-14 As in Example 3-1, let the random variable X denote the number of the 48 voice lines that are in use at a particular time. Assume that X is a discrete uniform random variable with a range of 0 to 48. Then,

$$E(X) = (48 + 0)/2 = 24$$

and

$$\sigma = \{[(48 - 0 + 1)^2 - 1]/12\}^{1/2} = 14.14$$

Equation 3-6 is more useful than it might first appear. If all the values in the range of a random variable X are multiplied by a constant (without changing any probabilities), the mean and standard deviation of X are multiplied by the constant. You are asked to verify this result in an exercise. Because the variance of a random variable is the square of the standard deviation, the variance of X is multiplied by the constant squared. More general results of this type are discussed in Chapter 5.

EXAMPLE 3-15 Let the random variable Y denote the proportion of the 48 voice lines that are in use at a particular time, and X denotes the number of lines that are in use at a particular time. Then, $Y = X/48$. Therefore,

$$E(Y) = E(X)/48 = 0.5$$

and

$$V(Y) = V(X)/48^2 = 0.087$$

EXERCISES FOR SECTION 3-5

3-46. Let the random variable X have a discrete uniform distribution on the integers $0 \le x \le 100$. Determine the mean and variance of X.

3-47. Let the random variable X have a discrete uniform distribution on the integers $1 \le x \le 3$. Determine the mean and variance of X.

3-48. Let the random variable X be equally likely to assume any of the values 1/8, 1/4, or 3/8. Determine the mean and variance of X.

3-49. Thickness measurements of a coating process are made to the nearest hundredth of a millimeter. The thickness measurements are uniformly distributed with values 0.15, 0.16, 0.17, 0.18, and 0.19. Determine the mean and variance of the coating thickness for this process.

3-50. Product codes of 2, 3, or 4 letters are equally likely. What is the mean and standard deviation of the number of letters in 100 codes?

3-51. The lengths of plate glass parts are measured to the nearest tenth of a millimeter. The lengths are uniformly distributed, with values at every tenth of a millimeter starting at 590.0 and continuing through 590.9. Determine the mean and variance of lengths.

3-52. Suppose that X has a discrete uniform distribution on the integers 0 through 9. Determine the mean, variance, and standard deviation of the random variable $Y = 5X$ and compare to the corresponding results for X.

3-53. Show that for a discrete uniform random variable X, if each of the values in the range of X is multiplied by the constant c, the effect is to multiply the mean of X by c and the variance of X by c^2. That is, show that $E(cX) = cE(X)$ and $V(cX) = c^2V(X)$.

3-54. The probability of an operator entering alphanumeric data incorrectly into a field in a database is equally likely. The random variable X is the number of fields on a data entry form with an error. The data entry form has 28 fields. Is X a discrete uniform random variable? Why or why not?

3-6 BINOMIAL DISTRIBUTION

Consider the following random experiments and random variables:

1. Flip a coin 10 times. Let X = number of heads obtained.
2. A worn machine tool produces 1% defective parts. Let X = number of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. Let X = the number of air samples that contain the rare molecule in the next 18 samples analyzed.
4. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X = the number of bits in error in the next five bits transmitted.
5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X = the number of questions answered correctly.
6. In the next 20 births at a hospital, let X = the number of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X = the number of patients who experience improvement.

These examples illustrate that a general probability model that includes these experiments as particular cases would be very useful.

Each of these random experiments can be thought of as consisting of a series of repeated, random trials: 10 flips of the coin in experiment 1, the production of 25 parts in experiment 2, and so forth. The random variable in each case is a count of the number of trials that meet a specified criterion. The outcome from each trial either meets the criterion that X counts or it does not; consequently, each trial can be summarized as resulting in either a success or a failure. For example, in the multiple choice experiment, for each question, only the choice that is correct is considered a success. Choosing any one of the three incorrect choices results in the trial being summarized as a failure.

The terms *success* and *failure* are just labels. We can just as well use A and B or 0 or 1. Unfortunately, the usual labels can sometimes be misleading. In experiment 2, because X counts defective parts, the production of a defective part is called a success.

A trial with only two possible outcomes is used so frequently as a building block of a random experiment that it is called a **Bernoulli trial.** It is usually assumed that the trials that constitute the random experiment are **independent.** This implies that the outcome from one trial has no effect on the outcome to be obtained from any other trial. Furthermore, it is often reasonable to assume that the **probability of a success in each trial is constant.** In the multiple choice experiment, if the test taker has no knowledge of the material and just guesses at each question, we might assume that the probability of a correct answer is 1/4 *for each question.*

Factorial notation is used in this section. Recall that $n!$ denotes the product of the integers less than or equal to n:

$$n! = n(n-1)(n-2)\cdots(2)(1)$$

For example,

$$5! = (5)(4)(3)(2)(1) = 120 \qquad 1! = 1$$

and by definition $0! = 1$. We also use the combinatorial notation

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

For example,

$$\binom{5}{2} = \frac{5!}{2!\,3!} = \frac{120}{2 \cdot 6} = 10$$

See Section 2-1.4, CD material for Chapter 2, for further comments.

EXAMPLE 3-16 The chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Also, assume that the transmission trials are independent. Let X = the number of bits in error in the next four bits transmitted. Determine $P(X = 2)$.

Let the letter E denote a bit in error, and let the letter O denote that the bit is okay, that is, received without error. We can represent the outcomes of this experiment as a list of four letters that indicate the bits that are in error and those that are okay. For example, the outcome *OEOE* indicates that the second and fourth bits are in error and the other two bits are okay. The corresponding values for x are

Outcome	x	Outcome	x
OOOO	0	*EOOO*	1
OOOE	1	*EOOE*	2
OOEO	1	*EOEO*	2
OOEE	2	*EOEE*	3
OEOO	1	*EEOO*	2
OEOE	2	*EEOE*	3
OEEO	2	*EEEO*	3
OEEE	3	*EEEE*	4

The event that $X = 2$ consists of the six outcomes:

$$\{EEOO, EOEO, EOOE, OEEO, OEOE, OOEE\}$$

Using the assumption that the trials are independent, the probability of $\{EEOO\}$ is

$$P(EEOO) = P(E)P(E)P(O)P(O) = (0.1)^2(0.9)^2 = 0.0081$$

Also, any one of the six mutually exclusive outcomes for which $X = 2$ has the same probability of occurring. Therefore,

$$P(X = 2) = 6(0.0081) = 0.0486$$

In general,

$$P(X = x) = (\text{number of outcomes that result in } x \text{ errors}) \text{ times } (0.1)^x(0.9)^{4-x}$$

To complete a general probability formula, only an expression for the number of outcomes that contain x errors is needed. An outcome that contains x errors can be constructed by partitioning the four trials (letters) in the outcome into two groups. One group is of size x and contains the errors, and the other group is of size $n - x$ and consists of the trials that are okay. The number of ways of partitioning four objects into two groups, one of which is of size x, is $\binom{4}{x} = \frac{4!}{x!(4-x)!}$. Therefore, in this example

$$P(X = x) = \binom{4}{x}(0.1)^x(0.9)^{4-x}$$

Notice that $\binom{4}{2} = 4!/[2!\,2!] = 6$, as found above. The probability mass function of X was shown in Example 3-4 and Fig. 3-1.

The previous example motivates the following result.

Definition

A random experiment consists of n Bernoulli trials such that

(1) The trials are independent

(2) Each trial results in only two possible outcomes, labeled as "success" and "failure"

(3) The probability of a success in each trial, denoted as p, remains constant

The random variable X that equals the number of trials that result in a success has a **binomial random variable** with parameters $0 < p < 1$ and $n = 1, 2, \ldots$. The probability mass function of X is

$$f(x) = \binom{n}{x} p^x(1-p)^{n-x} \qquad x = 0, 1, \ldots, n \tag{3-7}$$

As in Example 3-16, $\binom{n}{x}$ equals the total number of different sequences of trials that contain x successes and $n - x$ failures. The total number of different sequences that contain x successes and $n - x$ failures times the probability of each sequence equals $P(X = x)$.

The probability expression above is a very useful formula that can be applied in a number of examples. The name of the distribution is obtained from the *binomial expansion*. For constants a and b, the binomial expansion is

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Let p denote the probability of success on a single trial. Then, by using the binomial expansion with $a = p$ and $b = 1 - p$, we see that the sum of the probabilities for a binomial random variable is 1. Furthermore, because each trial in the experiment is classified into two outcomes, {success, failure}, the distribution is called a "bi"-nomial. A more

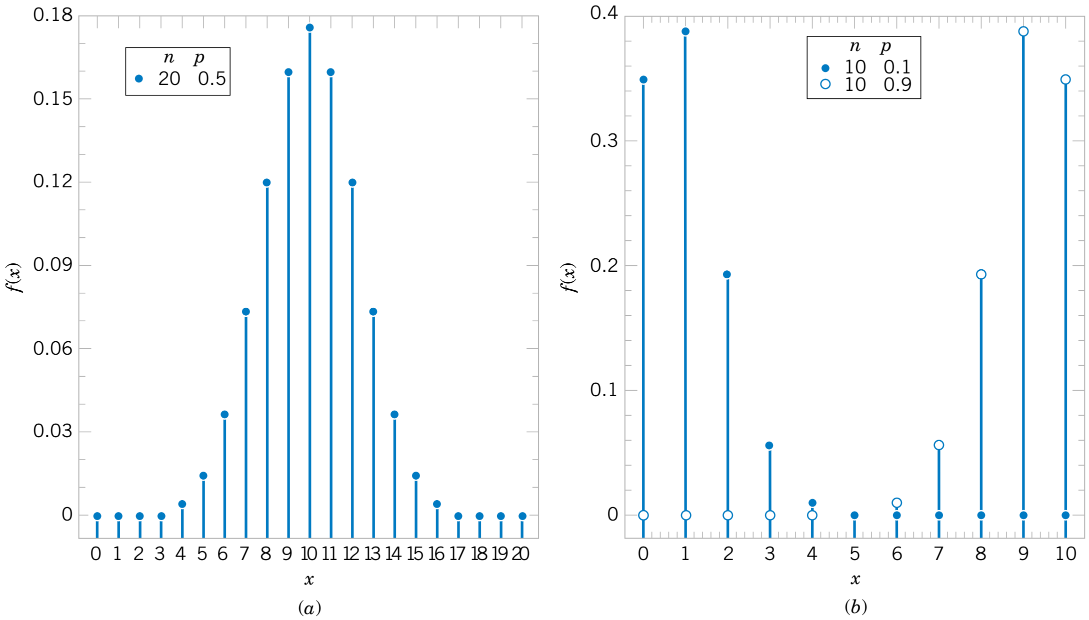


Figure 3-8 Binomial distributions for selected values of n and p.

general distribution, which includes the binomial as a special case, is the multinomial distribution.

Examples of binomial distributions are shown in Fig. 3-8. For a fixed n, the distribution becomes more symmetric as p increases from 0 to 0.5 or decreases from 1 to 0.5. For a fixed p, the distribution becomes more symmetric as n increases.

EXAMPLE 3-17

Several examples using the binomial coefficient $\binom{n}{x}$ follow.

$$\binom{10}{3} = 10!/[3!\ 7!] = (10 \cdot 9 \cdot 8)/(3 \cdot 2) = 120$$

$$\binom{15}{10} = 15!/[10!\ 5!] = (15 \cdot 14 \cdot 13 \cdot 12 \cdot 11)/(5 \cdot 4 \cdot 3 \cdot 2) = 3003$$

$$\binom{100}{4} = 100!/[4!\ 96!] = (100 \cdot 99 \cdot 98 \cdot 97)/(4 \cdot 3 \cdot 2) = 3{,}921{,}225$$

EXAMPLE 3-18

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

Let X = the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with $p = 0.1$ and $n = 18$. Therefore,

$$P(X = 2) = \binom{18}{2}(0.1)^2(0.9)^{16}$$

Now $\binom{18}{2} = 18!/[2!\ 16!] = 18(17)/2 = 153$. Therefore,

$$P(X = 2) = 153(0.1)^2(0.9)^{16} = 0.284$$

Determine the probability that at least four samples contain the pollutant. The requested probability is

$$P(X \geq 4) = \sum_{x=4}^{18} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$

However, it is easier to use the complementary event,

$$P(X \geq 4) = 1 - P(X < 4) = 1 - \sum_{x=0}^{3} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$
$$= 1 - [0.150 + 0.300 + 0.284 + 0.168] = 0.098$$

Determine the probability that $3 \leq X < 7$. Now

$$P(3 \leq X < 7) = \sum_{x=3}^{6} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$
$$= 0.168 + 0.070 + 0.022 + 0.005$$
$$= 0.265$$

The mean and variance of a binomial random variable depend only on the parameters p and n. Formulas can be developed from moment generating functions, and details are provided in Section 5-8, part of the CD material for Chapter 5. The results are simply stated here.

Definition

If X is a binomial random variable with parameters p and n,

$$\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = V(X) = np(1 - p) \qquad (3\text{-}8)$$

EXAMPLE 3-19

For the number of transmitted bits received in error in Example 3-16, $n = 4$ and $p = 0.1$, so

$$E(X) = 4(0.1) = 0.4 \qquad \text{and} \qquad V(X) = 4(0.1)(0.9) = 0.36$$

and these results match those obtained from a direct calculation in Example 3-9.

EXERCISES FOR SECTION 3-6

3-55. For each scenario described below, state whether or not the binomial distribution is a reasonable model for the random variable and why. State any assumptions you make.

(a) A production process produces thousands of temperature transducers. Let X denote the number of nonconforming transducers in a sample of size 30 selected at random from the process.

(b) From a batch of 50 temperature transducers, a sample of size 30 is selected without replacement. Let X denote the number of nonconforming transducers in the sample.

(c) Four identical electronic components are wired to a controller that can switch from a failed component to one of the remaining spares. Let X denote the number of components that have failed after a specified period of operation.
(d) Let X denote the number of accidents that occur along the federal highways in Arizona during a one-month period.
(e) Let X denote the number of correct answers by a student taking a multiple choice exam in which a student can eliminate some of the choices as being incorrect in some questions and all of the incorrect choices in other questions.
(f) Defects occur randomly over the surface of a semiconductor chip. However, only 80% of defects can be found by testing. A sample of 40 chips with one defect each is tested. Let X denote the number of chips in which the test finds a defect.
(g) Reconsider the situation in part (f). Now, suppose the sample of 40 chips consists of chips with 1 and with 0 defects.
(h) A filling operation attempts to fill detergent packages to the advertised weight. Let X denote the number of detergent packages that are underfilled.
(i) Errors in a digital communication channel occur in bursts that affect several consecutive bits. Let X denote the number of bits in error in a transmission of 100,000 bits.
(j) Let X denote the number of surface flaws in a large coil of galvanized steel.

3-56. The random variable X has a binomial distribution with $n = 10$ and $p = 0.5$. Sketch the probability mass function of X.
(a) What value of X is most likely?
(b) What value(s) of X is(are) least likely?

3-57. The random variable X has a binomial distribution with $n = 10$ and $p = 0.5$. Determine the following probabilities:
(a) $P(X = 5)$ (b) $P(X \le 2)$
(c) $P(X \ge 9)$ (d) $P(3 \le X < 5)$

3-58. Sketch the probability mass function of a binomial distribution with $n = 10$ and $p = 0.01$ and comment on the shape of the distribution.
(a) What value of X is most likely?
(b) What value of X is least likely?

3-59. The random variable X has a binomial distribution with $n = 10$ and $p = 0.01$. Determine the following probabilities.
(a) $P(X = 5)$ (b) $P(X \le 2)$
(c) $P(X \ge 9)$ (d) $P(3 \le X < 5)$

3-60. Determine the cumulative distribution function of a binomial random variable with $n = 3$ and $p = 1/2$.

3-61. Determine the cumulative distribution function of a binomial random variable with $n = 3$ and $p = 1/4$.

3-62. An electronic product contains 40 integrated circuits. The probability that any integrated circuit is defective is 0.01, and the integrated circuits are independent. The product operates only if there are no defective integrated circuits. What is the probability that the product operates?

3-63. Let X denote the number of bits received in error in a digital communication channel, and assume that X is a binomial random variable with $p = 0.001$. If 1000 bits are transmitted, determine the following:
(a) $P(X = 1)$ (b) $P(X \ge 1)$
(c) $P(X \le 2)$ (d) mean and variance of X

3-64. The phone lines to an airline reservation system are occupied 40% of the time. Assume that the events that the lines are occupied on successive calls are independent. Assume that 10 calls are placed to the airline.
(a) What is the probability that for exactly three calls the lines are occupied?
(b) What is the probability that for at least one call the lines are not occupied?
(c) What is the expected number of calls in which the lines are all occupied?

3-65. Batches that consist of 50 coil springs from a production process are checked for conformance to customer requirements. The mean number of nonconforming coil springs in a batch is 5. Assume that the number of nonconforming springs in a batch, denoted as X, is a binomial random variable.
(a) What are n and p?
(b) What is $P(X \le 2)$?
(c) What is $P(X \ge 49)$?

3-66. **A statistical process control chart example.** Samples of 20 parts from a metal punching process are selected every hour. Typically, 1% of the parts require rework. Let X denote the number of parts in the sample of 20 that require rework. A process problem is suspected if X exceeds its mean by more than three standard deviations.
(a) If the percentage of parts that require rework remains at 1%, what is the probability that X exceeds its mean by more than three standard deviations?
(b) If the rework percentage increases to 4%, what is the probability that X exceeds 1?
(c) If the rework percentage increases to 4%, what is the probability that X exceeds 1 in at least one of the next five hours of samples?

3-67. Because not all airline passengers show up for their reserved seat, an airline sells 125 tickets for a flight that holds only 120 passengers. The probability that a passenger does not show up is 0.10, and the passengers behave independently.
(a) What is the probability that every passenger who shows up can take the flight?
(b) What is the probability that the flight departs with empty seats?

3-68. This exercise illustrates that poor quality can affect schedules and costs. A manufacturing process has 100 customer orders to fill. Each order requires one component part that is purchased from a supplier. However, typically, 2% of the components are identified as defective, and the components can be assumed to be independent.
(a) If the manufacturer stocks 100 components, what is the probability that the 100 orders can be filled without reordering components?

(b) If the manufacturer stocks 102 components, what is the probability that the 100 orders can be filled without reordering components?
(c) If the manufacturer stocks 105 components, what is the probability that the 100 orders can be filled without reordering components?

3-69. A multiple choice test contains 25 questions, each with four answers. Assume a student just guesses on each question.
(a) What is the probability that the student answers more than 20 questions correctly?
(b) What is the probability the student answers less than 5 questions correctly?

3-70. A particularly long traffic light on your morning commute is green 20% of the time that you approach it. Assume that each morning represents an independent trial.
(a) Over five mornings, what is the probability that the light is green on exactly one day?
(b) Over 20 mornings, what is the probability that the light is green on exactly four days?
(c) Over 20 mornings, what is the probability that the light is green on more than four days?

3-7 GEOMETRIC AND NEGATIVE BINOMIAL DISTRIBUTIONS

3-7.1 Geometric Distribution

Consider a random experiment that is closely related to the one used in the definition of a binomial distribution. Again, assume a series of Bernoulli trials (independent trials with constant probability p of a success on each trial). However, instead of a fixed number of trials, trials are conducted until a success is obtained. Let the random variable X denote the number of trials until the first success. In Example 3-5, successive wafers are analyzed until a large particle is detected. Then, X is the number of wafers analyzed. In the transmission of bits, X might be the number of bits transmitted until an error occurs.

EXAMPLE 3-20 The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable X denote the number of bits transmitted *until* the first error.

Then, $P(X = 5)$ is the probability that the first four bits are transmitted correctly and the fifth bit is in error. This event can be denoted as $\{OOOOE\}$, where O denotes an okay bit. Because the trials are independent and the probability of a correct transmission is 0.9,

$$P(X = 5) = P(OOOOE) = 0.9^4 0.1 = 0.066$$

Note that there is some probability that X will equal any integer value. Also, if the first trial is a success, $X = 1$. Therefore, the range of X is $\{1, 2, 3, \ldots\}$, that is, all positive integers.

Definition

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until the first success. Then X is a **geometric random variable** with parameter $0 < p < 1$ and

$$f(x) = (1 - p)^{x-1} p \qquad x = 1, 2, \ldots \tag{3-9}$$

Examples of the probability mass functions for geometric random variables are shown in Fig. 3-9. Note that the height of the line at x is $(1 - p)$ times the height of the line at $x - 1$. That is, the probabilities decrease in a geometric progression. The distribution acquires its name from this result.

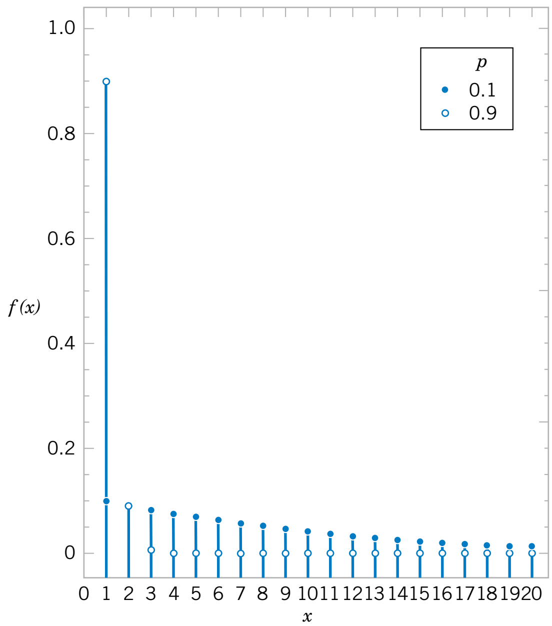


Figure 3-9 Geometric distributions for selected values of the parameter p.

EXAMPLE 3-21 The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent, what is the probability that exactly 125 wafers need to be analyzed before a large particle is detected?

Let X denote the number of samples analyzed until a large particle is detected. Then X is a geometric random variable with $p = 0.01$. The requested probability is

$$P(X = 125) = (0.99)^{124}0.01 = 0.0029$$

The derivation of the mean and variance of a geometric random variable is left as an exercise. Note that $\sum_{k=1}^{\infty} k(1-p)^{k-1}p$ can be shown to equal $1/p$. The results are as follows.

> If X is a geometric random variable with parameter p,
>
> $$\mu = E(X) = 1/p \qquad \text{and} \qquad \sigma^2 = V(X) = (1-p)/p^2 \qquad (3\text{-}10)$$

EXAMPLE 3-22 Consider the transmission of bits in Example 3-20. Here, $p = 0.1$. The mean number of transmissions until the first error is $1/0.1 = 10$. The standard deviation of the number of transmissions before the first error is

$$\sigma = [(1 - 0.1)/0.1^2]^{1/2} = 9.49$$

Lack of Memory Property

A geometric random variable has been defined as the number of trials until the first success. However, because the trials are independent, the count of the number of trials until the next success can be started at any trial without changing the probability distribution of the random variable. For example, in the transmission of bits, if 100 bits are transmitted, the probability that the first error, after bit 100, occurs on bit 106 is the probability that the next six outcomes are *OOOOOE*. This probability is $(0.9)^5(0.1) = 0.059$, which is identical to the probability that the initial error occurs on bit 6.

The implication of using a geometric model is that the system presumably will not wear out. The probability of an error remains constant for all transmissions. In this sense, the geometric distribution is said to lack any memory. The **lack of memory property** will be discussed again in the context of an exponential random variable in Chapter 4.

EXAMPLE 3-23 In Example 3-20, the probability that a bit is transmitted in error is equal to 0.1. Suppose 50 bits have been transmitted. The mean number of bits until the next error is $1/0.1 = 10$—the same result as the mean number of bits until the first error.

3-7.2 Negative Binomial Distribution

A generalization of a geometric distribution in which the random variable is the number of Bernoulli trials required to obtain r successes results in the **negative binomial distribution**.

EXAMPLE 3-24 As in Example 3-20, suppose the probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable X denote the number of bits transmitted until the *fourth* error.

Then, X has a negative binomial distribution with $r = 4$. Probabilities involving X can be found as follows. The $P(X = 10)$ is the probability that exactly three errors occur in the first nine trials and then trial 10 results in the fourth error. The probability that exactly three errors occur in the first nine trials is determined from the binomial distribution to be

$$\binom{9}{3}(0.1)^3(0.9)^6$$

Because the trials are independent, the probability that exactly three errors occur in the first 9 trials and trial 10 results in the fourth error is the product of the probabilities of these two events, namely,

$$\binom{9}{3}(0.1)^3(0.9)^6(0.1) = \binom{9}{3}(0.1)^4(0.9)^6$$

The previous result can be generalized as follows.

Definition

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until r successes occur. Then X is a **negative binomial random variable** with parameters $0 < p < 1$ and $r = 1, 2 3, \ldots,$ and

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \qquad x = r, r+1, r+2, \ldots. \tag{3-11}$$

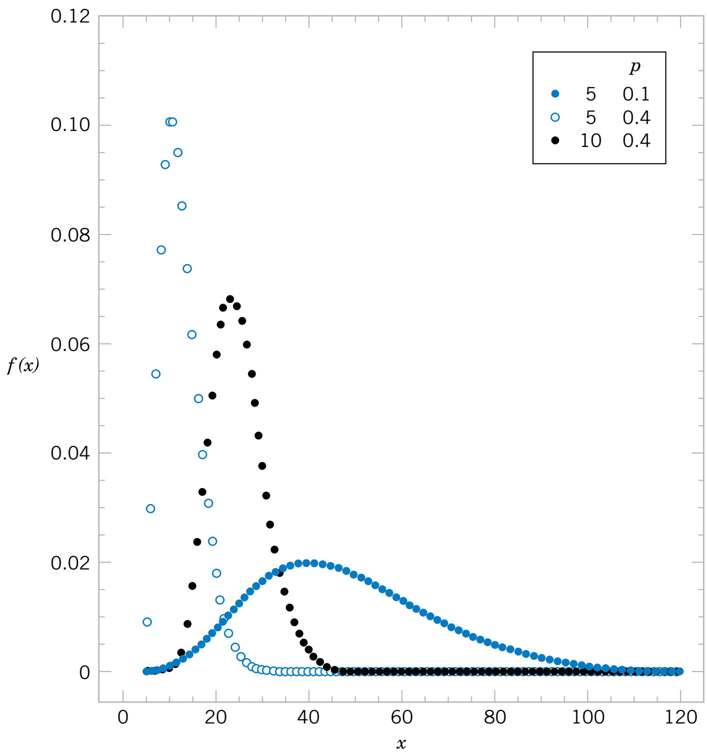


Figure 3-10 Negative binomial distributions for selected values of the parameters r and p.

Because at least r trials are required to obtain r successes, the range of X is from r to ∞. In the special case that $r = 1$, a negative binomial random variable is a geometric random variable. Selected negative binomial distributions are illustrated in Fig. 3-10.

The lack of memory property of a geometric random variable implies the following. Let X denote the total number of trials required to obtain r successes. Let X_1 denote the number of trials required to obtain the first success, let X_2 denote the number of extra trials required to obtain the second success, let X_3 denote the number of extra trials to obtain the third success, and so forth. Then, the total number of trials required to obtain r successes is $X = X_1 + X_2 + \cdots + X_r$. Because of the lack of memory property, each of the random variables $X_1, X_2, \ldots, X_r$ has a geometric distribution with the same value of p. Consequently, a negative binomial random variable can be interpreted as the sum of r geometric random variables. This concept is illustrated in Fig. 3-11.

Recall that a binomial random variable is a count of the number of successes in n Bernoulli trials. That is, the number of trials is predetermined, and the number of successes is random. A negative binomial random variable is a count of the number of trials required to

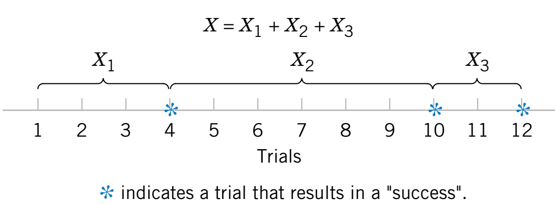


Figure 3-11 Negative binomial random variable represented as a sum of geometric random variables.

obtain r successes. That is, the number of successes is predetermined, and the number of trials is random. In this sense, a negative binomial random variable can be considered the opposite, or negative, of a binomial random variable.

The description of a negative binomial random variable as a sum of geometric random variables leads to the following results for the mean and variance. Sums of random variables are studied in Chapter 5.

If X is a negative binomial random variable with parameters p and r,

$$\mu = E(X) = r/p \quad \text{and} \quad \sigma^2 = V(X) = r(1-p)/p^2 \qquad (3\text{-}12)$$

EXAMPLE 3-25

A Web site contains three identical computer servers. Only one is used to operate the site, and the other two are spares that can be activated in case the primary system fails. The probability of a failure in the primary computer (or any activated spare system) from a request for service is 0.0005. Assuming that each request represents an independent trial, what is the mean number of requests until failure of all three servers?

Let X denote the number of requests until all three servers fail, and let X_1, X_2, and X_3 denote the number of requests before a failure of the first, second, and third servers used, respectively. Now, $X = X_1 + X_2 + X_3$. Also, the requests are assumed to comprise independent trials with constant probability of failure $p = 0.0005$. Furthermore, a spare server is not affected by the number of requests before it is activated. Therefore, X has a negative binomial distribution with $p = 0.0005$ and $r = 3$. Consequently,

$$E(X) = 3/0.0005 = 6000 \text{ requests}$$

What is the probability that all three servers fail within five requests? The probability is $P(X \le 5)$ and

$$\begin{aligned} P(X \le 5) &= P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.0005^3 + \binom{3}{2} 0.0005^3(0.9995) + \binom{4}{2} 0.0005^3(0.9995)^2 \\ &= 1.25 \times 10^{-10} + 3.75 \times 10^{-10} + 7.49 \times 10^{-10} \\ &= 1.249 \times 10^{-9} \end{aligned}$$

EXERCISES FOR SECTION 3-7

3-71. Suppose the random variable X has a geometric distribution with $p = 0.5$. Determine the following probabilities:
(a) $P(X = 1)$ (b) $P(X = 4)$
(c) $P(X = 8)$ (d) $P(X \le 2)$
(e) $P(X > 2)$

3-72. Suppose the random variable X has a geometric distribution with a mean of 2.5. Determine the following probabilities:
(a) $P(X = 1)$ (b) $P(X = 4)$
(c) $P(X = 5)$ (d) $P(X \le 3)$
(e) $P(X > 3)$

3-73. The probability of a successful optical alignment in the assembly of an optical data storage product is 0.8. Assume the trials are independent.
(a) What is the probability that the first successful alignment requires exactly four trials?
(b) What is the probability that the first successful alignment requires at most four trials?
(c) What is the probability that the first successful alignment requires at least four trials?

3-74. In a clinical study, volunteers are tested for a gene that has been found to increase the risk for a disease. The probability that a person carries the gene is 0.1.
(a) What is the probability 4 or more people will have to be tested before 2 with the gene are detected?
(b) How many people are expected to be tested before 2 with the gene are detected?

3-75. Assume that each of your calls to a popular radio station has a probability of 0.02 of connecting, that is, of not obtaining a busy signal. Assume that your calls are independent.
(a) What is the probability that your first call that connects is your tenth call?
(b) What is the probability that it requires more than five calls for you to connect?
(c) What is the mean number of calls needed to connect?

3-76. In Exercise 3-70, recall that a particularly long traffic light on your morning commute is green 20% of the time that you approach it. Assume that each morning represents an independent trial.
(a) What is the probability that the first morning that the light is green is the fourth morning that you approach it?
(b) What is the probability that the light is not green for 10 consecutive mornings?

3-77. A trading company has eight computers that it uses to trade on the New York Stock Exchange (NYSE). The probability of a computer failing in a day is 0.005, and the computers fail independently. Computers are repaired in the evening and each day is an independent trial.
(a) What is the probability that all eight computers fail in a day?
(b) What is the mean number of days until a specific computer fails?
(c) What is the mean number of days until all eight computers fail in the same day?

3-78. In Exercise 3-66, recall that 20 parts are checked each hour and that X denotes the number of parts in the sample of 20 that require rework.
(a) If the percentage of parts that require rework remains at 1%, what is the probability that hour 10 is the first sample at which X exceeds 1?
(b) If the rework percentage increases to 4%, what is the probability that hour 10 is the first sample at which X exceeds 1?
(c) If the rework percentage increases to 4%, what is the expected number of hours until X exceeds 1?

3-79. Consider a sequence of independent Bernoulli trials with $p = 0.2$.
(a) What is the expected number of trials to obtain the first success?
(b) After the eighth success occurs, what is the expected number of trials to obtain the ninth success?

3-80. Show that the probability density function of a negative binomial random variable equals the probability density function of a geometric random variable when $r = 1$. Show that the formulas for the mean and variance of a negative binomial random variable equal the corresponding results for geometric random variable when $r = 1$.

3-81. Suppose that X is a negative binomial random variable with $p = 0.2$ and $r = 4$. Determine the following:
(a) $E(X)$ (b) $P(X = 20)$
(c) $P(X = 19)$ (d) $P(X = 21)$
(e) The most likely value for X

3-82. The probability is 0.6 that a calibration of a transducer in an electronic instrument conforms to specifications for the measurement system. Assume the calibration attempts are independent. What is the probability that at most three calibration attempts are required to meet the specifications for the measurement system?

3-83. An electronic scale in an automated filling operation stops the manufacturing line after three underweight packages are detected. Suppose that the probability of an underweight package is 0.001 and each fill is independent.
(a) What is the mean number of fills before the line is stopped?
(b) What is the standard deviation of the number of fills before the line is stopped?

3-84. A fault-tolerant system that processes transactions for a financial services firm uses three separate computers. If the operating computer fails, one of the two spares can be immediately switched online. After the second computer fails, the last computer can be immediately switched online. Assume that the probability of a failure during any transaction is 10^{-8} and that the transactions can be considered to be independent events.
(a) What is the mean number of transactions before all computers have failed?
(b) What is the variance of the number of transactions before all computers have failed?

3-85. Derive the expressions for the mean and variance of a geometric random variable with parameter p. (Formulas for infinite series are required.)

3-8 HYPERGEOMETRIC DISTRIBUTION

In Example 3-8, 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 day's production. That is, selected units are not replaced before the next selection is made. Let A and B denote the events that the first and second parts are nonconforming, respectively. In Chapter 2, we found $P(B|A) = 49/849$ and $P(A) = 50/850$. Consequently, knowledge that the first part is nonconforming suggests that it is less likely that the second part selected is nonconforming.

This experiment is fundamentally different from the examples based on the binomial distribution. In this experiment, the trials are not independent. Note that, in the unusual case that each unit selected is replaced before the next selection, the trials are independent and there is a constant probability of a nonconforming part on each trial. Then, the number of nonconforming parts in the sample is a binomial random variable.

Let X equal the number of nonconforming parts in the sample. Then

$$P(X = 0) = P(\text{both parts conform}) = (800/850)(799/849) = 0.886$$

$$\begin{aligned} P(X = 1) &= P(\text{first part selected conforms and the second part selected} \\ &\quad \text{does not, or the first part selected does not and the second part} \\ &\quad \text{selected conforms}) \\ &= (800/850)(50/849) + (50/850)(800/849) = 0.111 \end{aligned}$$

$$P(X = 2) = P(\text{both parts do not conform}) = (50/850)(49/849) = 0.003$$

As in this example, samples are often selected without replacement. Although probabilities can be determined by the reasoning used in the example above, a general formula for computing probabilities when samples are selected without replacement is quite useful. The counting rules presented in Section 2-1.4, part of the CD material for Chapter 2, can be used to justify the formula given below.

Definition

A set of N objects contains

K objects classified as successes

$N - K$ objects classified as failures

A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$.

Let the random variable X denote the number of successes in the sample. Then X is a **hypergeometric random variable** and

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \qquad x = \max\{0, n + K - N\} \text{ to } \min\{K, n\} \qquad \text{(3-13)}$$

The expression $\min\{K, n\}$ is used in the definition of the range of X because the maximum number of successes that can occur in the sample is the smaller of the sample size, n,

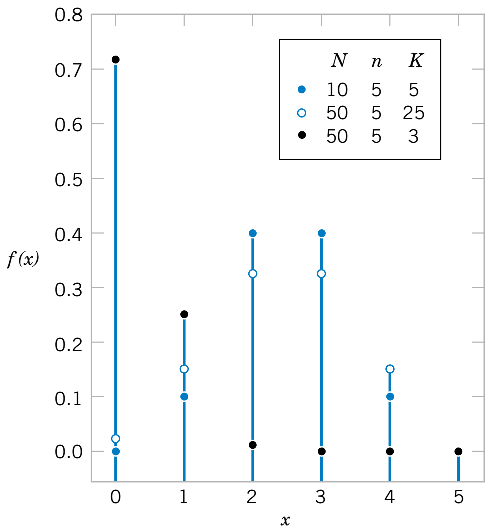


Figure 3-12 Hypergeometric distributions for selected values of parameters N, K, and n.

and the number of successes available, K. Also, if $n + K > N$, at least $n + K - N$ successes must occur in the sample. Selected hypergeometric distributions are illustrated in Fig. 3-12.

EXAMPLE 3-26

The example at the start of this section can be reanalyzed by using the general expression in the definition of a hypergeometric random variable. That is,

$$P(X = 0) = \frac{\binom{50}{0}\binom{800}{2}}{\binom{850}{2}} = \frac{319600}{360825} = 0.886$$

$$P(X = 1) = \frac{\binom{50}{1}\binom{800}{1}}{\binom{850}{2}} = \frac{40000}{360825} = 0.111$$

$$P(X = 2) = \frac{\binom{50}{2}\binom{800}{0}}{\binom{850}{2}} = \frac{1225}{360825} = 0.003$$

EXAMPLE 3-27

A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in the next state. If four parts are selected randomly and without replacement, what is the probability they are all from the local supplier?

Let X equal the number of parts in the sample from the local supplier. Then, X has a hypergeometric distribution and the requested probability is $P(X = 4)$. Consequently,

$$P(X = 4) = \frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}} = 0.0119$$

What is the probability that two or more parts in the sample are from the local supplier?

$$P(X \geq 2) = \frac{\binom{100}{2}\binom{200}{2}}{\binom{300}{4}} + \frac{\binom{100}{3}\binom{200}{1}}{\binom{300}{4}} + \frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}$$
$$= 0.298 + 0.098 + 0.0119 = 0.408$$

What is the probability that at least one part in the sample is from the local supplier?

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{100}{0}\binom{200}{4}}{\binom{300}{4}} = 0.804$$

The mean and variance of a hypergeometric random variable can be determined from the trials that comprise the experiment. However, the trials are not independent, and so the calculations are more difficult than for a binomial distribution. The results are stated as follows.

If X is a hypergeometric random variable with parameters N, K, and n, then

$$\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = V(X) = np(1 - p)\left(\frac{N - n}{N - 1}\right) \tag{3-14}$$

where $p = K/N$.

Here p is interpreted as the proportion of successes in the set of N objects.

EXAMPLE 3-28 In the previous example, the sample size is 4. The random variable X is the number of parts in the sample from the local supplier. Then, $p = 100/300 = 1/3$. Therefore,

$$E(X) = 4(100/300) = 1.33$$

and

$$V(X) = 4(1/3)(2/3)[(300 - 4)/299] = 0.88$$

For a hypergeometric random variable, $E(X)$ is similar to the mean a binomial random variable. Also, $V(X)$ differs from the result for a binomial random variable only by the term shown below.

Finite Population Correction Factor

The term in the variance of a hypergeometric random variable

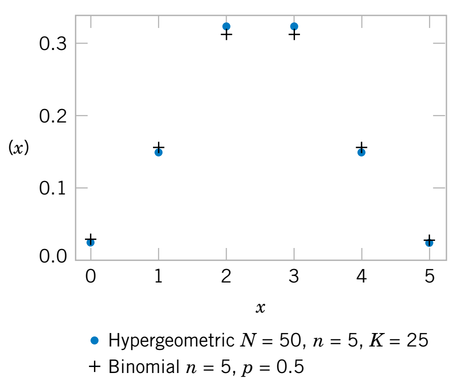
$$\frac{N - n}{N - 1}$$

is called the finite population correction factor.

Sampling with replacement is equivalent to sampling from an infinite set because the proportion of success remains constant for every trial in the experiment. As mentioned previously, if sampling were done with replacement, X would be a binomial random variable and its variance would be $np(1 - p)$. Consequently, the finite population correction represents the correction to the binomial variance that results because the sampling is without replacement from the finite set of size N.

If n is small relative to N, the correction is small and the hypergeometric distribution is similar to the binomial. In this case, a binomial distribution can effectively approximate the distribution of the number of units of a specified type in the sample. A case is illustrated in Fig. 3-13.

EXAMPLE 3-29 A listing of customer accounts at a large corporation contains 1000 customers. Of these, 700 have purchased at least one of the corporation's products in the last three months. To evaluate a new product design, 50 customers are sampled at random from the corporate listing. What is



	0	1	2	3	4	5
Hypergeometric probability	0.025	0.149	0.326	0.326	0.149	0.025
Binomial probability	0.031	0.156	0.321	0.312	0.156	0.031

Figure 3-13 Comparison of hypergeometric and binomial distributions.

the probability that more than 45 of the sampled customers have purchased from the corporation in the last three months?

The sampling is without replacement. However, because the sample size of 50 is small relative to the number of customer accounts, 1000, the probability of selecting a customer who has purchased from the corporation in the last three months remains approximately constant as the customers are chosen.

For example, let A denote the event that the first customer selected has purchased from the corporation in the last three months, and let B denote the event that the second customer selected has purchased from the corporation in the last three months. Then, $P(A) = 700/1000 = 0.7$ and $P(B|A) = 699/999 = 0.6997$. That is, the trials are approximately independent.

Let X denote the number of customers in the sample who have purchased from the corporation in the last three months. Then, X is a hypergeometric random variable with $N = 1000$, $n = 50$, and $K = 700$. Consequently, $p = K/N = 0.7$. The requested probability is $P(X > 45)$. Because the sample size is small relative to the batch size, the distribution of X can be approximated as binomial with $n = 50$ and $p = 0.7$. Using the binomial approximation to the distribution of X results in

$$P(X > 45) = \sum_{x=46}^{50} \binom{50}{x} 0.7^x (1 - 0.7)^{50-x} = 0.00017$$

The probability from the hypergeometric distribution is 0.000166, but this requires computer software. The result agrees well with the binomial approximation.

EXERCISES FOR SECTION 3-8

3-86. Suppose X has a hypergeometric distribution with $N = 100$, $n = 4$, and $K = 20$. Determine the following:
(a) $P(X = 1)$ (b) $P(X = 6)$
(c) $P(X = 4)$ (d) Determine the mean and variance of X.

3-87. Suppose X has a hypergeometric distribution with $N = 20$, $n = 4$, and $K = 4$. Determine the following:
(a) $P(X = 1)$ (b) $P(X = 4)$
(c) $P(X \leq 2)$ (d) Determine the mean and variance of X.

3-88. Suppose X has a hypergeometric distribution with $N = 10$, $n = 3$, and $K = 4$. Sketch the probability mass function of X.

3-89. Determine the cumulative distribution function for X in Exercise 3-88.

3-90. A lot of 75 washers contains 5 in which the variability in thickness around the circumference of the washer is unacceptable. A sample of 10 washers is selected at random, without replacement.
(a) What is the probability that none of the unacceptable washers is in the sample?
(b) What is the probability that at least one unacceptable washer is in the sample?
(c) What is the probability that exactly one unacceptable washer is in the sample?
(d) What is the mean number of unacceptable washers in the sample?

3-91. A company employs 800 men under the age of 55. Suppose that 30% carry a marker on the male chromosome that indicates an increased risk for high blood pressure.
(a) If 10 men in the company are tested for the marker in this chromosome, what is the probability that exactly 1 man has the marker?
(b) If 10 men in the company are tested for the marker in this chromosome, what is the probability that more than 1 has the marker?

3-92. Printed circuit cards are placed in a functional test after being populated with semiconductor chips. A lot contains 140 cards, and 20 are selected without replacement for functional testing.
(a) If 20 cards are defective, what is the probability that at least 1 defective card is in the sample?
(b) If 5 cards are defective, what is the probability that at least 1 defective card appears in the sample?

3-93. Magnetic tape is slit into half-inch widths that are wound into cartridges. A slitter assembly contains 48 blades. Five blades are selected at random and evaluated each day for sharpness. If any dull blade is found, the assembly is replaced with a newly sharpened set of blades.

(a) If 10 of the blades in an assembly are dull, what is the probability that the assembly is replaced the first day it is evaluated?

(b) If 10 of the blades in an assembly are dull, what is the probability that the assembly is not replaced until the third day of evaluation? [*Hint:* Assume the daily decisions are independent, and use the geometric distribution.]

(c) Suppose on the first day of evaluation, two of the blades are dull, on the second day of evaluation six are dull, and on the third day of evaluation, ten are dull. What is the probability that the assembly is not replaced until the third day of evaluation? [*Hint:* Assume the daily decisions are independent. However, the probability of replacement changes every day.]

3-94. A state runs a lottery in which 6 numbers are randomly selected from 40, without replacement. A player chooses 6 numbers before the state's sample is selected.

(a) What is the probability that the 6 numbers chosen by a player match all 6 numbers in the state's sample?

(b) What is the probability that 5 of the 6 numbers chosen by a player appear in the state's sample?

(c) What is the probability that 4 of the 6 numbers chosen by a player appear in the state's sample?

(d) If a player enters one lottery each week, what is the expected number of weeks until a player matches all 6 numbers in the state's sample?

3-95. Continuation of Exercises 3-86 and 3-87.

(a) Calculate the finite population corrections for Exercises 3-86 and 3-87. For which exercise should the binomial approximation to the distribution of X be better?

(b) For Exercise 3-86, calculate $P(X = 1)$ and $P(X = 4)$ assuming that X has a binomial distribution and compare these results to results derived from the hypergeometric distribution.

(c) For Exercise 3-87, calculate $P(X = 1)$ and $P(X = 4)$ assuming that X has a binomial distribution and compare these results to the results derived from the hypergeometric distribution.

3-96. Use the binomial approximation to the hypergeometric distribution to approximate the probabilities in Exercise 3-92. What is the finite population correction in this exercise?

3-9 POISSON DISTRIBUTION

We introduce the Poisson distribution with an example.

EXAMPLE 3-30

Consider the transmission of n bits over a digital communication channel. Let the random variable X equal the number of bits in error. When the probability that a bit is in error is constant and the transmissions are independent, X has a binomial distribution. Let p denote the probability that a bit is in error. Let $\lambda = pn$. Then, $E(x) = pn = \lambda$ and

$$P(X = x) = \binom{n}{x} P^x(1 - p)^{n-x} = \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Now, suppose that the number of bits transmitted increases and the probability of an error decreases exactly enough that pn remains equal to a constant. That is, n increases and p decreases accordingly, such that $E(X) = \lambda$ remains constant. Then, with some work, it can be shown that

$$\lim_{n\to\infty} P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Also, because the number of bits transmitted tends to infinity, the number of errors can equal any nonnegative integer. Therefore, the range of X is the integers from zero to infinity.

The distribution obtained as the limit in the above example is more useful than the derivation above implies. The following example illustrates the broader applicability.

EXAMPLE 3-31 Flaws occur at random along the length of a thin copper wire. Let X denote the random variable that counts the number of flaws in a length of L millimeters of wire and suppose that the average number of flaws in L millimeters is λ.

The probability distribution of X can be found by reasoning in a manner similar to the previous example. Partition the length of wire into n subintervals of small length, say, 1 micrometer each. If the subinterval chosen is small enough, the probability that more than one flaw occurs in the subinterval is negligible. Furthermore, we can interpret the assumption that flaws occur at random to imply that every subinterval has the same probability of containing a flaw, say, p. Finally, if we assume that the probability that a subinterval contains a flaw is independent of other subintervals, we can model the distribution of X as approximately a binomial random variable. Because

$$E(X) = \lambda = np$$

we obtain

$$p = \lambda/n$$

That is, the probability that a subinterval contains a flaw is λ/n. With small enough subintervals, n is very large and p is very small. Therefore, the distribution of X is obtained as in the previous example.

Example 3-31 can be generalized to include a broad array of random experiments. The interval that was partitioned was a length of wire. However, the same reasoning can be applied to any interval, including an interval of time, an area, or a volume. For example, counts of (1) particles of contamination in semiconductor manufacturing, (2) flaws in rolls of textiles, (3) calls to a telephone exchange, (4) power outages, and (5) atomic particles emitted from a specimen have all been successfully modeled by the probability mass function in the following definition.

Definition

Given an interval of real numbers, assume counts occur at random throughout the interval. If the interval can be partitioned into subintervals of small enough length such that

(1) the probability of more than one count in a subinterval is zero,
(2) the probability of one count in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and
(3) the count in each subinterval is independent of other subintervals, the random experiment is called a **Poisson process.**

The random variable X that equals the number of counts in the interval is a **Poisson random variable** with parameter $0 < \lambda$, and the probability mass function of X is

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \ldots \tag{3-15}$$

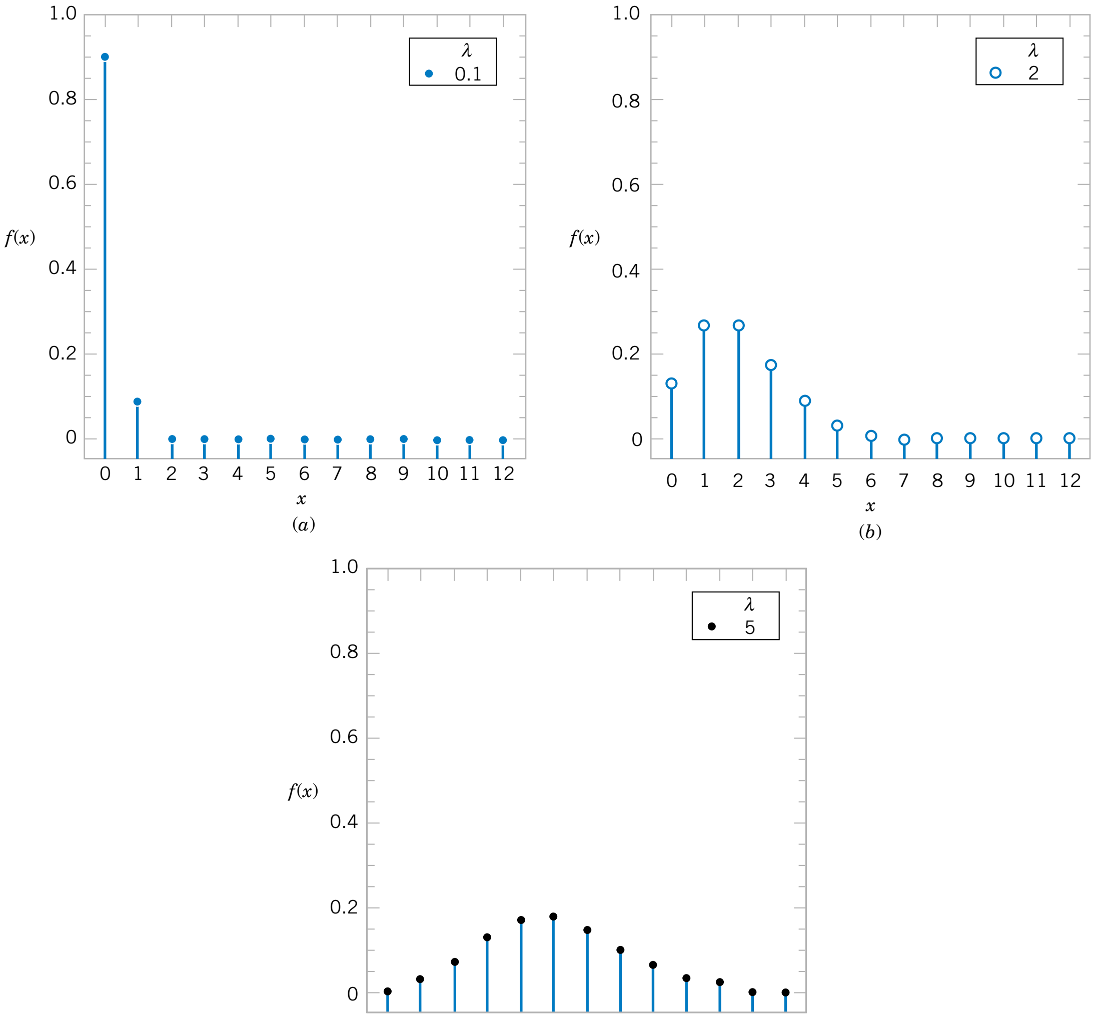


Figure 3-14 Poisson distributions for selected values of the parameters.

Historically, the term *process* has been used to suggest the observation of a system over time. In our example with the copper wire, we showed that the Poisson distribution could also apply to intervals such as lengths. Figure 3-14 provides graphs of selected Poisson distributions.

It is important to **use consistent units** in the calculation of probabilities, means, and variances involving Poisson random variables. The following example illustrates unit conversions. For example, if the

average number of flaws per millimeter of wire is 3.4, then the

average number of flaws in 10 millimeters of wire is 34, and the

average number of flaws in 100 millimeters of wire is 340.

If a Poisson random variable represents the number of counts in some interval, the mean of the random variable must equal the expected number of counts in the same length of interval.

EXAMPLE 3-32

For the case of the thin copper wire, suppose that the number of flaws follows a Poisson distribution with a mean of 2.3 flaws per millimeter. Determine the probability of exactly 2 flaws in 1 millimeter of wire.

Let X denote the number of flaws in 1 millimeter of wire. Then, $E(X) = 2.3$ flaws and

$$P(X = 2) = \frac{e^{-2.3} 2.3^2}{2!} = 0.265$$

Determine the probability of 10 flaws in 5 millimeters of wire. Let X denote the number of flaws in 5 millimeters of wire. Then, X has a Poisson distribution with

$$E(X) = 5 \text{ mm} \times 2.3 \text{ flaws/mm} = 11.5 \text{ flaws}$$

Therefore,

$$P(X = 10) = e^{-11.5} \frac{11.5^{10}}{10!} = 0.113$$

Determine the probability of at least 1 flaw in 2 millimeters of wire. Let X denote the number of flaws in 2 millimeters of wire. Then, X has a Poisson distribution with

$$E(X) = 2 \text{ mm} \times 2.3 \text{ flaws/mm} = 4.6 \text{ flaws}$$

Therefore,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-4.6} = 0.9899$$

EXAMPLE 3-33

Contamination is a problem in the manufacture of optical storage disks. The number of particles of contamination that occur on an optical disk has a Poisson distribution, and the average number of particles per centimeter squared of media surface is 0.1. The area of a disk under study is 100 squared centimeters. Find the probability that 12 particles occur in the area of a disk under study.

Let X denote the number of particles in the area of a disk under study. Because the mean number of particles is 0.1 particles per cm^2

$$E(X) = 100 \text{ cm}^2 \times 0.1 \text{ particles/cm}^2 = 10 \text{ particles}$$

Therefore,

$$P(X = 12) = \frac{e^{-10} 10^{12}}{12!} = 0.095$$

The probability that zero particles occur in the area of the disk under study is

$$P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$$

Determine the probability that 12 or fewer particles occur in the area of the disk under study. The probability is

$$P(X \leq 12) = P(X = 0) + P(X = 1) + \cdots + P(X = 12) = \sum_{i=0}^{12} \frac{e^{-10} 10^i}{i!}$$

Because this sum is tedious to compute, many computer programs calculate cumulative Poisson probabilities. From one such program, $P(X \le 12) = 0.791$.

The derivation of the mean and variance of a Poisson random variable is left as an exercise. The results are as follows.

If X is a Poisson random variable with parameter λ, then

$$\mu = E(X) = \lambda \quad \text{and} \quad \sigma^2 = V(X) = \lambda \tag{3-16}$$

The mean and variance of a Poisson random variable are equal. For example, if particle counts follow a Poisson distribution with a mean of 25 particles per square centimeter, the variance is also 25 and the standard deviation of the counts is 5 per square centimeter. Consequently, information on the variability is very easily obtained. Conversely, if the variance of count data is much greater than the mean of the same data, the Poisson distribution is not a good model for the distribution of the random variable.

EXERCISES FOR SECTION 3-9

3-97. Suppose X has a Poisson distribution with a mean of 4. Determine the following probabilities:
(a) $P(X = 0)$ (b) $P(X \le 2)$
(c) $P(X = 4)$ (d) $P(X = 8)$

3-98. Suppose X has a Poisson distribution with a mean of 0.4. Determine the following probabilities:
(a) $P(X = 0)$ (b) $P(X \le 2)$
(c) $P(X = 4)$ (d) $P(X = 8)$

3-99. Suppose that the number of customers that enter a bank in an hour is a Poisson random variable, and suppose that $P(X = 0) = 0.05$. Determine the mean and variance of X.

3-100. The number of telephone calls that arrive at a phone exchange is often modeled as a Poisson random variable. Assume that on the average there are 10 calls per hour.
(a) What is the probability that there are exactly 5 calls in one hour?
(b) What is the probability that there are 3 or less calls in one hour?
(c) What is the probability that there are exactly 15 calls in two hours?
(d) What is the probability that there are exactly 5 calls in 30 minutes?

3-101. The number of flaws in bolts of cloth in textile manufacturing is assumed to be Poisson distributed with a mean of 0.1 flaw per square meter.
(a) What is the probability that there are two flaws in 1 square meter of cloth?
(b) What is the probability that there is one flaw in 10 square meters of cloth?
(c) What is the probability that there are no flaws in 20 square meters of cloth?
(d) What is the probability that there are at least two flaws in 10 square meters of cloth?

3-102. When a computer disk manufacturer tests a disk, it writes to the disk and then tests it using a certifier. The certifier counts the number of missing pulses or errors. The number of errors on a test area on a disk has a Poisson distribution with $\lambda = 0.2$.
(a) What is the expected number of errors per test area?
(b) What percentage of test areas have two or fewer errors?

3-103. The number of cracks in a section of interstate highway that are significant enough to require repair is assumed to follow a Poisson distribution with a mean of two cracks per mile.
(a) What is the probability that there are no cracks that require repair in 5 miles of highway?
(b) What is the probability that at least one crack requires repair in 1/2 mile of highway?
(c) If the number of cracks is related to the vehicle load on the highway and some sections of the highway have a heavy load of vehicles whereas other sections carry a light load, how do you feel about the assumption of a Poisson distribution for the number of cracks that require repair?

3-104. The number of failures for a cytogenics machine from contamination in biological samples is a Poisson random variable with a mean of 0.01 per 100 samples.
(a) If the lab usually processes 500 samples per day, what is the expected number of failures per day?

(b) What is the probability that the machine will not fail during a study that includes 500 participants? (Assume one sample per participant.)

3-105. The number of surface flaws in plastic panels used in the interior of automobiles has a Poisson distribution with a mean of 0.05 flaw per square foot of plastic panel. Assume an automobile interior contains 10 square feet of plastic panel.
(a) What is the probability that there are no surface flaws in an auto's interior?
(b) If 10 cars are sold to a rental company, what is the probability that none of the 10 cars has any surface flaws?
(c) If 10 cars are sold to a rental company, what is the probability that at most one car has any surface flaws?

3-106. The number of failures of a testing instrument from contamination particles on the product is a Poisson random variable with a mean of 0.02 failure per hour.
(a) What is the probability that the instrument does not fail in an 8-hour shift?
(b) What is the probability of at least one failure in a 24-hour day?

Supplemental Exercises

3-107. A shipment of chemicals arrives in 15 totes. Three of the totes are selected at random, without replacement, for an inspection of purity. If two of the totes do not conform to purity requirements, what is the probability that at least one of the nonconforming totes is selected in the sample?

3-108. The probability that your call to a service line is answered in less than 30 seconds is 0.75. Assume that your calls are independent.
(a) If you call 10 times, what is the probability that exactly 9 of your calls are answered within 30 seconds?
(b) If you call 20 times, what is the probability that at least 16 calls are answered in less than 30 seconds?
(c) If you call 20 times, what is the mean number of calls that are answered in less than 30 seconds?

3-109. Continuation of Exercise 3-108.
(a) What is the probability that you must call four times to obtain the first answer in less than 30 seconds?
(b) What is the mean number of calls until you are answered in less than 30 seconds?

3-110. Continuation of Exercise 3-109.
(a) What is the probability that you must call six times in order for two of your calls to be answered in less than 30 seconds?
(b) What is the mean number of calls to obtain two answers in less than 30 seconds?

3-111. The number of messages sent to a computer bulletin board is a Poisson random variable with a mean of 5 messages per hour.
(a) What is the probability that 5 messages are received in 1 hour?
(b) What is the probability that 10 messages are received in 1.5 hours?
(c) What is the probability that less than two messages are received in one-half hour?

3-112. A Web site is operated by four identical computer servers. Only one is used to operate the site; the others are spares that can be activated in case the active server fails. The probability that a request to the Web site generates a failure in the active server is 0.0001. Assume that each request is an independent trial. What is the mean time until failure of all four computers?

3-113. The number of errors in a textbook follow a Poisson distribution with a mean of 0.01 error per page. What is the probability that there are three or less errors in 100 pages?

3-114. The probability that an individual recovers from an illness in a one-week time period without treatment is 0.1. Suppose that 20 independent individuals suffering from this illness are treated with a drug and 4 recover in a one-week time period. If the drug has no effect, what is the probability that 4 or more people recover in a one-week time period?

3-115. Patient response to a generic drug to control pain is scored on a 5-point scale, where a 5 indicates complete relief. Historically the distribution of scores is

1	2	3	4	5
0.05	0.1	0.2	0.25	0.4

Two patients, assumed to be independent, are each scored.
(a) What is the probability mass function of the total score?
(b) What is the probability mass function of the average score?

3-116. In a manufacturing process that laminates several ceramic layers, 1% of the assemblies are defective. Assume that the assemblies are independent.
(a) What is the mean number of assemblies that need to be checked to obtain five defective assemblies?
(b) What is the standard deviation of the number of assemblies that need to be checked to obtain five defective assemblies?

3-117. Continuation of Exercise 3-116. Determine the minimum number of assemblies that need to be checked so that the probability of at least one defective assembly exceeds 0.95.

3-118. Determine the constant c so that the following function is a probability mass function: $f(x) = cx$ for $x = 1, 2, 3, 4$.

3-119. A manufacturer of a consumer electronics product expects 2% of units to fail during the warranty period. A sample of 500 independent units is tracked for warranty performance.
(a) What is the probability that none fails during the warranty period?
(b) What is the expected number of failures during the warranty period?
(c) What is the probability that more than two units fail during the warranty period?

3-120. Messages that arrive at a service center for an information systems manufacturer have been classified on the basis of the number of keywords (used to help route messages) and the type of message, either email or voice. Also, 70% of the messages arrive via email and the rest are voice.

number of keywords	0	1	2	3	4
email	0.1	0.1	0.2	0.4	0.2
voice	0.3	0.4	0.2	0.1	0

Determine the probability mass function of the number of keywords in a message.

3-121. The random variable X has the following probability distribution:

x	2	3	5	8
probability	0.2	0.4	0.3	0.1

Determine the following:
(a) $P(X \leq 3)$ (b) $P(X > 2.5)$
(c) $P(2.7 < X < 5.1)$ (d) $E(X)$
(e) $V(X)$

3-122. Determine the probability mass function for the random variable with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < 2 \\ 0.2 & 2 \leq x < 5.7 \\ 0.5 & 5.7 \leq x < 6.5 \\ 0.8 & 6.5 \leq x < 8.5 \\ 1 & 8.5 \leq x \end{cases}$$

3-123. Each main bearing cap in an engine contains four bolts. The bolts are selected at random, without replacement, from a parts bin that contains 30 bolts from one supplier and 70 bolts from another.
(a) What is the probability that a main bearing cap contains all bolts from the same supplier?
(b) What is the probability that exactly three bolts are from the same supplier?

3-124. Assume the number of errors along a magnetic recording surface is a Poisson random variable with a mean of one error every 10^5 bits. A sector of data consists of 4096 eight-bit bytes.
(a) What is the probability of more than one error in a sector?
(b) What is the mean number of sectors until an error is found?

3-125. An installation technician for a specialized communication system is dispatched to a city only when three or more orders have been placed. Suppose orders follow a Poisson distribution with a mean of 0.25 per week for a city with a population of 100,000 and suppose your city contains a population of 800,000.
(a) What is the probability that a technician is required after a one-week period?
(b) If you are the first one in the city to place an order, what is the probability that you have to wait more than two weeks from the time you place your order until a technician is dispatched?

3-126. From 500 customers, a major appliance manufacturer will randomly select a sample without replacement. The company estimates that 25% of the customers will provide useful data. If this estimate is correct, what is the probability mass function of the number of customers that will provide useful data?
(a) Assume that the company samples 5 customers.
(b) Assume that the company samples 10 customers.

3-127. It is suspected that some of the totes containing chemicals purchased from a supplier exceed the moisture content target. Samples from 30 totes are to be tested for moisture content. Assume that the totes are independent. Determine the proportion of totes from the supplier that must exceed the moisture content target so that the probability is 0.90 that at least one tote in the sample of 30 fails the test.

3-128. Messages arrive to a computer server according to a Poisson distribution with a mean rate of 10 per hour. Determine the length of an interval of time such that the probability that no messages arrive during this interval is 0.90.

3-129. Flaws occur in the interior of plastic used for automobiles according to a Poisson distribution with a mean of 0.02 flaw per panel.
(a) If 50 panels are inspected, what is the probability that there are no flaws?
(b) What is the expected number of panels that need to be inspected before a flaw is found?
(c) If 50 panels are inspected, what is the probability that the number of panels that have one or more flaws is less than or equal to 2?

MIND-EXPANDING EXERCISES

3-130. Derive the mean and variance of a hypergeometric random variable (difficult exercise).

3-131. Show that the function $f(x)$ in Example 3-5 satisfies the properties of a probability mass function by summing the infinite series.

3-132. Derive the formula for the mean and standard deviation of a discrete uniform random variable over the range of integers $a, a + 1, \ldots, b$.

3-133. A company performs inspection on shipments from suppliers in order to defect nonconforming products. Assume a lot contains 1000 items and 1% are nonconforming. What sample size is needed so that the probability of choosing at least one nonconforming item in the sample is at least 0.90? Assume the binomial approximation to the hypergeometric distribution is adequate.

3-134. A company performs inspection on shipments from suppliers in order to detect nonconforming products. The company's policy is to use a sample size that is always 10% of the lot size. Comment on the effectiveness of this policy as a general rule for all sizes of lots.

3-135. Surface flaws in automobile exterior panels follow a Poisson distribution with a mean of 0.1 flaw per panel. If 100 panels are checked, what is the probability that fewer than five panels have any flaws?

3-136. A large bakery can produce rolls in lots of either 0, 1000, 2000, or 3000 per day. The production cost per item is \$0.10. The demand varies randomly according to the following distribution:

demand for rolls	0	1000	2000	3000
probability of demand	0.3	0.2	0.3	0.2

Every roll for which there is a demand is sold for \$0.30. Every roll for which there is no demand is sold in a secondary market for \$0.05. How many rolls should the bakery produce each day to maximize the mean profit?

3-137. A manufacturer stocks components obtained from a supplier. Suppose that 2% of the components are defective and that the defective components occur independently. How many components must the manufacturer have in stock so that the probability that 100 orders can be completed without reordering components is at least 0.95?

IMPORTANT TERMS AND CONCEPTS

In the E-book, click on any term or concept below to go to that subject.

Bernoulli trial
Binomial distribution
Cumulative probability distribution function-discrete random variable
Discrete uniform distribution
Expected value of a function of a random variable
Finite population correction factor
Geometric distribution
Hypergeometric distribution
Lack of memory property-discrete random variable
Mean-discrete random variable
Mean-function of a discrete random variable
Negative binomial distribution
Poisson distribution
Poisson process
Probability distribution-discrete random variable
Probability mass function
Standard deviation-discrete random variable
Variance-discrete random variable