CHAPTER 3

Probability Distributions

CONTENTS

In this chapter we continue the study of probability by introducing the concepts of random variables and probability distributions. We consider the probability distributions of both discrete and continuous random variables. Of particular interest are the binomial, Poisson, uniform, normal, and exponential probability distributions. These probability distributions are of great interest because they are used extensively in practice. The Q.M. in Action, ESPN and Probability, describes how the development and use of a probability distribution helped the organization enhance its audience's enjoyment and understanding of sports.

Q.M. in ACTION

*ESPN AND PROBABILITY**

The Entertainment and Sports Programming Network, known as ESPN since 1985, was originally established as a nationwide cable-television network dedicated to broadcasting and producing sports-related programming. The Bristol, Connecticut–based network, which provides programming 24 hours a day throughout the year, has grown rapidly since its debut on September 7, 1979. At various times its programming has included Major League Baseball, the National Football League (NFL), the National Basketball Association (NBA), the National Hockey League (NHL), NASCAR, NCAA Football and Basketball, Major League Soccer (MLS), the Men's and Women's professional golf associations (PGA and LPGA), and Men's and Women's professional tennis (ATP and WTA). ESPN is now in over 100 million homes in the United States, and ESPN International spans over 200 countries and territories on all seven continents and includes 46 television networks reaching over 350 million subscribers in 16 languages, plus wireless, interactive, print, radio, broadband, event management, and consumer products.

ESPN's rapid growth has coincided with a dramatic increase in sports fans' desire for more sophisticated analyses. The organization responded to this trend by establishing its Production Analytics department, a group of analysts who provide all of ESPN's media platforms with statistical analyses for a wide variety of sports problems. Senior Director of the Production Analytics department Jeff Bennett explains that "ESPN appreciates the sports fan's passion for meaningful analytics, and we are dedicated to creating and providing this content. Basic probability and statistics are critical tools in our analytic arsenal."

Alok Pattani, who is an Analytics Specialist in the Production Analytics department, further describes the department's uses of some specific probability concepts. "For example, we use very basic probability to determine the likelihood an NBA team will win one of the first three picks in the NBA draft. The league holds a Draft Lottery to determine which of its fourteen teams that didn't qualify for the playoffs during the most recent season will receive each of the first three picks, which are considered to be extremely valuable. The lottery is weighted so that teams with worse records have better chances of obtaining early picks, and we use information on how many chances each team is allocated to calculate the probability that any of these fourteen teams will win one of those top picks.

"Conditional probability is also very important; when we look at the probability a team will win a home game or a player will get a hit when playing at night, we are using conditional probability." Alok continues, "We use probability distributions extensively in our work, especially the binomial and normal distributions. We apply the binomial distribution to all kinds of success and failure situations such as wins and losses, field goals and missed shots in basketball, complete and incomplete passes in football, and hits and outs in baseball. The binomial distribution is also useful when estimating the probability of a hitting streak in baseball or a winning streak in any sport."

Bennett adds, "The results of these types of analyses are of great interest to ESPN's base. They enhance the fan's enjoyment and understanding of his or her favorite sports, and that is good business for ESPN."

*The authors are indebted to Jeff Bennett and Alok Pattani of ESPN Inc. for providing input for this Q.M. in Action.

3.1 Random Variables

Recall that in Chapter 2 we defined an experiment as any process that generates well-defined outcomes. We now want to concentrate on the process of assigning *numeric values* to experimental outcomes. To do so we introduce the notion of a random variable.

For any particular experiment a random variable can be defined so that each possible experimental outcome generates exactly one numeric value for the random variable. For example, if we consider the experiment of selling automobiles for one day at a particular dealership, we could describe the experimental outcomes in terms of the *number* of cars sold. In this case, if x = number of cars sold, x is called a random variable. The particular numeric value that the random variable assumes depends on the outcome of the experiment; that is, we will not know the specific value of the random variable until we have observed the experimental outcome. For example, if on a given day three cars are sold, the value of the random variable is 3; if on another day (a repeat of the experiment) four cars are sold, the value is 4. We define a random variable as follows:

Random variables must assume numeric values.

A **random variable** is a numeric description of the outcome of an experiment.

Some additional examples of experiments and their associated random variables are given in Table 3.1. Although many experiments have experimental outcomes that lend themselves quite naturally to numeric values, others do not. For example, for the experiment of tossing a coin one time, the experimental outcome will be either a head or a tail, neither of which has a natural numeric value. However, we still may want to express the outcomes in terms of a random variable. Thus, we need a rule that can be used to assign a numeric value to each of the experimental outcomes. One possibility is to let the random variable $x = 1$ if the experimental outcome is a head and $x = 0$ if the experimental outcome is a tail. Although the numeric values for x are arbitrary, x is a random variable because it describes the experimental outcomes numerically.

A random variable may be classified as either discrete or continuous, depending on the numeric values it may assume. A random variable that may assume only a finite or an infinite sequence (e.g., 1, 2, 3, . . .) of values is a **discrete random variable**. The number of units sold, the number of defects observed, the number of customers that enter a bank during one day of operation, and so on are examples of discrete random variables. The first two and the last random variables in Table 3.1 are discrete. Random variables such as weight, time, and temperature that may assume any value in a certain interval or collection of intervals are **continuous random variables**. For instance, the third random variable in Table 3.1 is a continuous random variable because it may assume any value in the interval from 0 to 100 (for example, 56.33 or 64.223).

Try Problem 1 for practice in identifying discrete and continuous random variables.

TABLE 3.1 EXAMPLES OF RANDOM VARIABLES

Experiment	Random Variable (x)	Possible Values for the Random Variable
Make 100 sales calls	Total number of sales	0, 1, 2, . . . , 100
Inspect a shipment of 70 radios	Number of defective radios	0, 1, 2, . . . , 70
Build a new library	Percentage of project completed after 6 months	$0 \leq x \leq 100$
Operate a restaurant	Number of customers entering in one day	0, 1, 2, . . .

NOTES AND COMMENTS

1. One way to determine whether a random variable is discrete or continuous is to think of the values of the random variable as points on a line. Choose two points representing values the random variable might assume. If the entire line segment between the two points also represents possible values for the random variable, the random variable is continuous. An alternative (but equivalent) way of determining whether a random variable is discrete or continuous is to choose two points that represent values the random variable might assume. If, no matter what two points you initially chose, you can always find a third point between your initial two points that also represents a value of the random variable, the random variable is continuous. For example, if your random variable is the exact weight of a bag of potato chips and you choose 16.0005 ounces and 16.0006 ounces as your original two points, the point 16.00051 (or 16.00052 or 16.00053, etc.) ounces represents a possible value of the exact weight of the bag of potato chips and lies between your intial two points. No matter what two values you initially select, you can find another value that represents a value of the random variable and lies between your initial two points. On the other hand, if your random variable is the number of customers who enter a restaurant in a day and you choose 109 customers and 110 customers as your initial two points, there is no point that represents a value of the random variable and lies between your initial two points (i.e., you cannot have 109.7 customers enter a restaurant on a particular day). This indicates that the random variable is discrete.

3.2 Discrete Random Variables

We can demonstrate the use of a discrete random variable by considering the sales of automobiles at DiCarlo Motors, Inc., in Saratoga, New York. The owner of DiCarlo Motors is interested in the daily sales volume for automobiles. Suppose that we let x be a random variable denoting the number of cars sold on a given day. Sales records show that 5 is the maximum number of cars that DiCarlo has ever sold during one day. The owner believes that the previous history of sales adequately represents what will occur in the future, so we would expect the random variable x to assume one of the numeric values 0, 1, 2, 3, 4, or 5. The possible values of the random variable are finite; thus we would classify x as a discrete random variable.

Probability Distribution of a Discrete Random Variable

Suppose that in checking DiCarlo's sales records we find that over the past year the firm was open for business 300 days. The sales volumes generated and the frequency of their occurrence are summarized in Table 3.2. With these historical data available, the owner of

TABLE 3.2 CARS SOLD PER DAY AT DICARLO MOTORS

Sales Volume	Number of Days
No sales	54
One car	117
Two cars	72
Three cars	42
Four cars	12
Five cars	3
Total	300

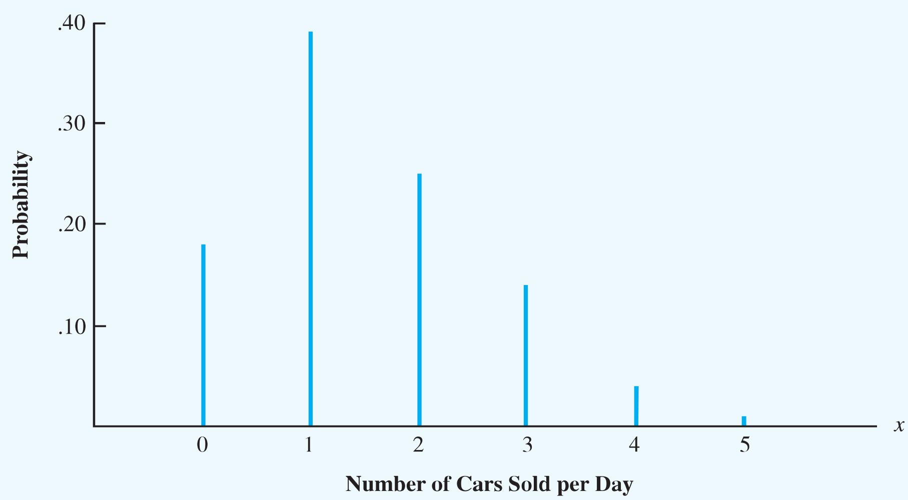
TABLE 3.3 PROBABILITY DISTRIBUTION FOR THE NUMBER OF CARS SOLD PER DAY

x	$f(x)$
0	0.18
1	0.39
2	0.24
3	0.14
4	0.04
5	0.01
Total	1.00

DiCarlo Motors believes that the relative frequency method will provide a reasonable means of assessing the probabilities for the random variable x. The **probability function**, denoted $f(x)$, provides the probability that the random variable x takes on a specific value. Because on 54 of the 300 days of historical data DiCarlo Motors did not sell any cars and because no sales corresponds to $x = 0$, we assign to $f(0)$ the value $^{54}/_{300} = 0.18$. Similarly, $f(1)$ denotes the probability that x takes on the value 1, so we assign to $f(1)$ the value $^{117}/_{300} = 0.39$. After computing the relative frequencies for the other possible values of x, we can develop a table of x and $f(x)$ values. Table 3.3 shows a tabular presentation of the probability distribution of the random variable x.

We can also represent the probability distribution of x graphically. In Figure 3.1 the values of the random variable x are shown on the horizontal axis. The probability that x takes on each of these values is shown on the vertical axis. For many discrete random variables the probability distribution also can be represented as a formula that provides $f(x)$ for every possible value of x. We illustrate this approach in the next section.

FIGURE 3.1 PROBABILITY DISTRIBUTION FOR THE NUMBER OF CARS SOLD PER DAY



In Section 2.2 we defined the two basic requirements of all probability assignments as $0 \leq P(E_i) \leq 1$ and $\Sigma P(E_i) = 1$. Equations (3.1) and (3.2) are the analogs of these basic requirements.

In the development of a **discrete probability distribution**, two requirements must always be satisfied:

$$f(x) \geq 0 \tag{3.1}$$

$$\Sigma f(x) = 1 \tag{3.2}$$

Equation (3.1) specifies that the probabilities associated with each value of x must be greater than or equal to zero, whereas equation (3.2) indicates that the sum of the probabilities for all values of the random variable x must be equal to 1. Table 3.3 shows that equations (3.1) and (3.2) are satisfied. Thus, the probability distribution developed for DiCarlo Motors is a valid discrete probability distribution.

After establishing a random variable and its probability distribution, we can develop a variety of additional probability information, depending on the needs and interests of the decision maker. For example, in the DiCarlo Motors problem the probability distribution shown in Table 3.3 can be used to provide the following information:

1. There is a 0.18 probability that no cars will be sold during a day.
2. The most probable sales volume is 1, with $f(1) = 0.39$.
3. There is a 0.05 probability of an outstanding sales day with four or five cars being sold.

Using probability information such as that just given, DiCarlo's management can understand better the uncertainties associated with the car sales operation. Perhaps this improved understanding can serve as the basis for a new policy or decision that will increase the effectiveness of the firm.

Try Problem 3 for practice in constructing a discrete probability distribution.

Expected Value

After constructing the probability distribution for a random variable, we often want to compute the mean or expected value of the random variable. The **expected value** of a discrete random variable is a weighted average of all possible values of the random variable, where the weights are the probabilities associated with the values. The mathematical formula for computing the expected value of a discrete random variable x is

$$E(x) = \mu = \Sigma x f(x) \tag{3.3}$$

As equation (3.3) shows, both the notations $E(x)$ and μ are used to refer to the expected value of a random variable.

To compute the expected value of a discrete random variable, we must multiply each value of the random variable by its corresponding probability and then add the resulting terms. Calculation of the expected value of the random variable (number of daily sales) for DiCarlo Motors is shown in Table 3.4. The first column contains the values of the random variable x, and the second column contains their associated probabilities $f(x)$. Multiplying each value of x by its probability $f(x)$ provides the $xf(x)$ values in the third column. Following equation (3.3), we sum this column, $\Sigma xf(x)$, to find the expected value of 1.50 cars sold per day.

The expected value of a random variable is the mean, or average, value. For experiments that can be repeated numerous times, the expected value can be interpreted as the "long-run" average value for the random variable. However, the expected value is not necessarily the

TABLE 3.4 EXPECTED VALUE CALCULATION

x	$f(x)$	$xf(x)$
0	0.18	$0(0.18) = 0.00$
1	0.39	$1(0.39) = 0.39$
2	0.24	$2(0.24) = 0.48$
3	0.14	$3(0.14) = 0.42$
4	0.04	$4(0.04) = 0.16$
5	0.01	$5(0.01) = 0.05$
		$E(x) = 1.50$

number that we think the random variable will assume the next time the experiment is conducted. In fact, it is impossible for DiCarlo to sell exactly 1.50 cars on any day. However, if we envision selling cars at DiCarlo Motors for many days into the future, the expected value of 1.50 cars provides the mean, or average, daily sales volume.

The expected value can be important to a manager from both the planning and decision-making points of view. For example, suppose that DiCarlo Motors will be open 60 days during the next three months. How many cars will be sold during this time? Although we can't specify the exact sales for any given day, the expected value of 1.50 cars per day provides an expected or average sales estimate of 60(1.50) = 90 cars for the next three-month period. In terms of setting sales quotas and/or planning orders, the expected value may provide helpful decision-making information.

Variance

The expected value gives us an idea of the average or central value for the random variable, but we often want a measure of the dispersion, or variability, of the possible values of the random variable. For example, if the values of the random variable range from quite large to quite small, we would expect a large value for the measure of variability. If the values of the random variable show only modest variation, we would expect a relatively small value for the measure of variability. The variance is a measure commonly used to summarize the variability in the values of a random variable. The mathematical expression for the *variance* of a discrete random variable is

An alternative formula for the variance of a discrete random variable is $\mathrm{Var}(x) = \sum x^2 f(x) - \mu^2$.

$$\mathrm{Var}(x) = \sigma^2 = \sum (x - \mu)^2 f(x) \tag{3.4}$$

As equation (3.4) shows, an essential part of the variance formula is a *deviation,* $x - \mu$, which measures how far a particular value of the random variable is from the expected value or mean, μ. In computing the variance of a discrete random variable, we square the deviations and then weight them by the corresponding probability. The sum of these weighted squared deviations for all values of the random variable is the **variance**. In other words, the variance is a weighted average of the squared deviations.

The calculation of the variance for the number of daily sales in the DiCarlo Motors problem is summarized in Table 3.5. We see that the variance for the number of cars sold per day is 1.25. A related measure of variability is the **standard deviation**, σ, which is

TABLE 3.5 VARIANCE CALCULATION

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2 f(x)$
0	$0 - 1.50 = -1.50$	2.25	0.18	$2.25(0.18) = 0.4050$
1	$1 - 1.50 = -0.50$	0.25	0.39	$0.25(0.39) = 0.0975$
2	$2 - 1.50 = 0.50$	0.25	0.24	$0.25(0.24) = 0.0600$
3	$3 - 1.50 = 1.50$	2.25	0.14	$2.25(0.14) = 0.3150$
4	$4 - 1.50 = 2.50$	6.25	0.04	$6.25(0.04) = 0.2500$
5	$5 - 1.50 = 3.50$	12.25	0.01	$12.25(0.01) = 0.1225$
				$\sigma^2 = 1.2500$

defined as the positive square root of the variance. For DiCarlo Motors, the standard deviation of the number of cars sold per day is

$$\sigma = \sqrt{1.25} = 1.118$$

For the purpose of easier managerial interpretation, the standard deviation may be preferred over the variance because it is measured in the same units as the random variable ($\sigma = 1.118$ cars sold per day). The variance (σ^2) is measured in squared units and is thus more difficult for a manager to interpret.

Try Problem 4 to be sure you can compute the expected value, variance, and standard deviation.

At this point our interpretation of the variance and the standard deviation is limited to comparisons of the variability of different random variables. For example, if the daily sales data from a second DiCarlo dealership in Albany, New York, provided $\sigma^2 = 2.56$ and $\sigma = 1.6$, we can conclude that the number of cars sold per day at this dealership exhibits more variability than at the first DiCarlo dealership, where $\sigma^2 = 1.25$ and $\sigma = 1.118$. Later in this chapter we discuss the normal distribution. For that probability distribution, we show that the variance and the standard deviation of the random variable are essential for making probability calculations.

3.3 Binomial Probability Distribution

In this section we consider a class of experiments that meet the following conditions:

1. The experiment consists of a sequence of n identical *trials.*
2. Two outcomes are possible on each trial. We refer to one outcome as a *success* and the other as a *failure.*
3. The probabilities of the two outcomes do not change from one trial to the next.
4. The trials are independent (i.e., the outcome of one trial does not affect the outcome of any other trial).

Experiments that satisfy conditions 2, 3, and 4 are said to be generated by a *Bernoulli process.* In addition, if condition 1 is satisfied (there are n identical trials), we have a *binomial experiment.* An important discrete random variable associated with the binomial experiment is the number of outcomes labeled success in the n trials. If we let x denote the value of this random variable, then x can have a value of 0, 1, 2, 3, . . . , n, depending on the number of successes observed in the n trials. The probability distribution associated with this random variable is called the **binomial probability distribution**.

Try Problem 9, parts (a–d), for practice computing binomial probabilities.

In cases where the binomial distribution is applicable, the mathematical formula for computing the probability of any value for the random variable is the binomial probability function

$$f(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \qquad x = 0, 1, \ldots, n \tag{3.5}$$

where

$$\begin{aligned} n &= \text{number of trials} \\ p &= \text{probability of success on one trial} \\ x &= \text{number of successes in } n \text{ trials} \\ f(x) &= \text{probability of } x \text{ successes in } n \text{ trials} \end{aligned}$$

The term $n!$ in the preceding expression is referred to as *n factorial* and is defined as

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

For example, $4! = (4)(3)(2)(1) = 24$. Also, by definition, the special case of zero factorial is $0! = 1$.

Nastke Clothing Store Problem

To illustrate the binomial probability distribution, let us consider the experiment of customers entering the Nastke Clothing Store. To keep the problem relatively small, we restrict the experiment to the next three customers. If, based on experience, the store manager estimates that the probability of a customer making a purchase is 0.30, what is the probability that exactly two of the next three customers make a purchase?

We first want to demonstrate that three customers entering the clothing store and deciding whether to make a purchase can be viewed as a binomial experiment. Checking the four requirements for a binomial experiment, we note the following:

1. The experiment can be described as a sequence of three identical trials, one trial for each of the three customers who will enter the store.
2. Two outcomes—the customer makes a purchase (success) or the customer does not make a purchase (failure)—are possible for each trial.
3. The probabilities of the purchase (0.30) and no purchase (0.70) outcomes are assumed to be the same for all customers.
4. The purchase decision of each customer is independent of the purchase decision of the other customers.

Thus, if we define the random variable x as the number of customers making a purchase (i.e., the number of successes in the three trials), we satisfy the requirements of the binomial probability distribution.

With $n = 3$ trials and the probability of a purchase $p = 0.30$ for each customer, we use equation (3.5) to compute the probability of two customers making a purchase. This probability, denoted $f(2)$, is

$$\begin{aligned} f(2) &= \frac{3!}{2!1!}(0.30)^2(0.70)^1 \\ &= \frac{3 \times 2 \times 1}{2 \times 1 \times 1}(0.30)^2(0.70)^1 = 0.189 \end{aligned}$$

TABLE 3.6 PROBABILITY DISTRIBUTION FOR THE NUMBER OF CUSTOMERS MAKING A PURCHASE

x	$f(x)$
0	0.343
1	0.441
2	0.189
3	0.027
Total	1.000

Try Problem 12 for an application of the binomial distribution.

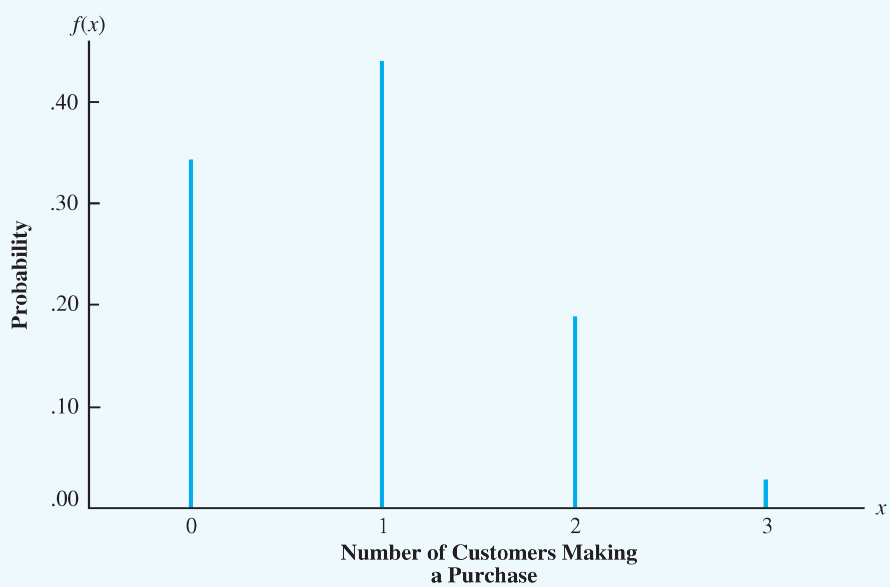
Similarly, the probability of no customers making a purchase, denoted $f(0)$, is

$$f(0) = \frac{3!}{0!3!}(0.30)^0(0.70)^3$$
$$= \frac{3 \times 2 \times 1}{1 \times 3 \times 2 \times 1}(0.30)^0(0.70)^3 = 0.343$$

Equation (3.5) can be used in a similar manner to show that the probabilities of one and three purchases are $f(1) = 0.441$ and $f(3) = 0.027$. Table 3.6 and Figure 3.2 summarize the binomial probability distribution for the Nastke Clothing Store problem.

If we consider any variation of the Nastke problem, such as 10 customers rather than 3 customers entering the store, the binomial probability function given by equation (3.5) still

FIGURE 3.2 PROBABILITY DISTRIBUTION FOR THE NASTKE CLOTHING STORE PROBLEM



applies. For example, the probability that 4 of the 10 customers make a purchase is

$$f(4) = \frac{10!}{4!6!}(0.30)^4(0.70)^6 = 0.2001$$

In this binomial experiment, $n = 10$, $x = 4$, and $p = 0.30$.

With modern computers and calculators, these tables are almost unnecessary. It is easy to evaluate equation (3.5) directly.

With the use of equation (3.5), tables have been developed that provide the probability of x successes in n trials for a binomial experiment. Such a table of binomial probability values is provided in Appendix B. We include a partial binomial table in Table 3.7. In order to use this table, specify the values of n, p, and x for the binomial experiment of interest. Check the use of this table by employing it to verify the probability of four successes in

TABLE 3.7 SELECTED VALUES FROM THE BINOMIAL PROBABILITY TABLE. EXAMPLE: $n = 10$, $x = 4$, $p = 0.30$; $f(4) = 0.2001$

n	*x*	*p*									
		0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020
	1	0.2985	0.3874	0.3679	0.3020	0.2253	0.1556	0.1004	0.0605	0.0339	0.0176
	2	0.0629	0.1722	0.2597	0.3020	0.3003	0.2668	0.2162	0.1612	0.1110	0.0703
	3	0.0077	0.0446	0.1069	0.1762	0.2336	0.2668	0.2716	0.2508	0.2119	0.1641
	4	0.0006	0.0074	0.0283	0.0661	0.1168	0.1715	0.2194	0.2508	0.2600	0.2461
	5	0.0000	0.0008	0.0050	0.0165	0.0389	0.0735	0.1181	0.1672	0.2128	0.2461
	6	0.0000	0.0001	0.0006	0.0028	0.0087	0.0210	0.0424	0.0743	0.1160	0.1641
	7	0.0000	0.0000	0.0000	0.0003	0.0012	0.0039	0.0098	0.0212	0.0407	0.0703
	8	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0013	0.0035	0.0083	0.0176
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0008	0.0020
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.3151	0.3874	0.3474	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098
	2	0.0746	0.1937	0.2759	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439
	3	0.0105	0.0574	0.1298	0.2013	0.2503	0.2668	0.2522	0.2150	0.1665	0.1172
	4	0.0010	0.0112	0.0401	0.0881	0.1460	**0.2001**	0.2377	0.2508	0.2384	0.2051
	5	0.0001	0.0015	0.0085	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461
	6	0.0000	0.0001	0.0012	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051
	7	0.0000	0.0000	0.0001	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172
	8	0.0000	0.0000	0.0000	0.0001	0.0004	0.0014	0.0043	0.0106	0.0229	0.0439
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010
11	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005
	1	0.3293	0.3835	0.3248	0.2362	0.1549	0.0932	0.0518	0.0266	0.0125	0.0054
	2	0.0867	0.2131	0.2866	0.2953	0.2581	0.1998	0.1395	0.0887	0.0531	0.0269
	3	0.0137	0.0710	0.1517	0.2215	0.2581	0.2568	0.2254	0.1774	0.1259	0.0806
	4	0.0014	0.0158	0.0536	0.1107	0.1721	0.2201	0.2428	0.2365	0.2060	0.1611
	5	0.0001	0.0025	0.0132	0.0388	0.0803	0.1321	0.1830	0.2207	0.2360	0.2256
	6	0.0000	0.0003	0.0023	0.0097	0.0268	0.0566	0.0985	0.1471	0.1931	0.2256
	7	0.0000	0.0000	0.0003	0.0017	0.0064	0.0173	0.0379	0.0701	0.1128	0.1611
	8	0.0000	0.0000	0.0000	0.0002	0.0011	0.0037	0.0102	0.0234	0.0462	0.0806
	9	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0018	0.0052	0.0126	0.0269
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0007	0.0021	0.0054
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0005

10 trials for the Nastke Clothing Store problem. Note that the value of $f(4) = 0.2001$ can be read directly from the table of binomial probabilities, making it unnecessary to perform the calculations required by equation (3.5).

Expected Value and Variance for the Binomial Distribution

From the probability distribution in Table 3.6, we can use equation (3.3) to compute the expected value or expected number of customers making a purchase:

$$\mu = \Sigma x f(x) = 0(0.343) + 1(0.441) + 2(0.189) + 3(0.027) = 0.9$$

Note that we could have obtained this same expected value simply by multiplying n (the number of trials) by p (the probability of success on any one trial):

$$np = 3(0.30) = 0.9$$

For the special case of a binomial probability distribution, the expected value of the random variable is given by

$$\mu = np \tag{3.6}$$

Thus, if you know that the probability distribution is binomial, you do not have to make the detailed calculations required by equation (3.3) to compute the expected value.

Suppose that during the next month Nastke's Clothing Store expects 1000 customers to enter the store. What is the expected number of customers who will make a purchase? Using equation (3.6), the answer is $\mu = np = (1000)(0.3) = 300$. To increase the expected number of sales, Nastke's must induce more customers to enter the store and/or somehow increase the probability that any individual customer will make a purchase after entering.

For the special case of a binomial distribution, the variance of the random variable is

Try Problem 9, part (e), for practice computing the expected value, variance, and standard deviation.

$$\sigma^2 = np(1 - p) \tag{3.7}$$

For the Nastke Clothing Store problem with three customers, the variance and standard deviation for the number of customers making a purchase are

$$\sigma^2 = np(1 - p) = 3(0.3)(0.7) = 0.63$$
$$\sigma = \sqrt{0.63} = 0.79$$

3.4 Poisson Probability Distribution

In this section we will consider a discrete random variable that often is useful when we are dealing with the number of occurrences of an event over a specified interval of time or space. For example, the random variable of interest might be the number of arrivals at a car wash in 1 hour, the number of repairs needed in 10 miles of highway, or the number of leaks in 100 miles of pipeline. If the following two assumptions are satisfied, the **Poisson probability distribution** is applicable:

1. The probability of an occurrence of the event is the same for any two intervals of equal length.
2. The occurrence or nonoccurrence of the event in any interval is independent of the occurrence or nonoccurrence in any other interval.

The probability function of the Poisson random variable is given by equation (3.8):

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{for } x = 0, 1, 2, \ldots \tag{3.8}$$

where

λ = mean or average number of occurrences in an interval

e = 2.71828

x = number of occurrences in the interval

$f(x)$ = probability of x occurrences in the interval

Note that equation (3.8) shows no upper limit to the number of possible values that a Poisson random variable can realize. That is, x is a discrete random variable with an infinite sequence of values ($x = 0, 1, 2, \ldots$); the Poisson random variable has no set upper limit.

An Example Involving Time Intervals

Suppose that we are interested in the number of arrivals at the drive-in teller window of a bank during a 15-minute period on weekday mornings. If we assume that the probability of a car arriving is the same for any two time periods of equal length and that the arrival or nonarrival of a car in any time period is independent of the arrival or nonarrival in any other time period, the Poisson probability function is applicable. Then if we assume that an analysis of historical data shows that the average number of cars arriving during a 15-minute interval of time is 10, the Poisson probability function with $\lambda = 10$ applies:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{10^x e^{-10}}{x!} \qquad \text{for } x = 0, 1, 2, \ldots$$

Bell Labs used the Poisson distribution in modeling the arrival of phone calls.

If we wanted to know the probability of five arrivals in 15 minutes, we would set $x = 5$ and obtain[1]

$$f(5) = \frac{10^5 e^{-10}}{5!} = 0.0378$$

Although we determined this probability by evaluating the probability function with $\lambda = 10$ and $x = 5$, the use of Poisson probability distribution tables often is easier. These tables provide probabilities for specific values of x and λ. We included such a table as Appendix C. For convenience we reproduce a portion of it as Table 3.8. To use the table of Poisson probabilities, you need know only the values of x and λ. Thus, from Table 3.8, the probability of five arrivals in a 15-minute period is the value in the row corresponding to $x = 5$ and the column corresponding to $\lambda = 10$. Hence, $f(5) = 0.0378$.

Try Problem 14 for practice computing Poisson probabilities.

An Example Involving Length or Distance Intervals

Suppose that we are concerned with the occurrence of major defects in a section of highway one month after resurfacing. We assume that the probability of a defect is the same for any two intervals of equal length and that the occurrence or nonoccurrence of a defect in any one interval is independent of the occurrence or nonoccurrence in any other interval. Thus, the Poisson probability distribution applies.

[1]Values of $e^{-\lambda}$ are available in Appendix E and can be easily computed with most modern calculators.

TABLE 3.8 SELECTED VALUES FROM THE POISSON PROBABILITY TABLE. EXAMPLE: $\lambda = 10$, $x = 5$; $f(5) = 0.0378$

	λ									
x	**9.1**	**9.2**	**9.3**	**9.4**	**9.5**	**9.6**	**9.7**	**9.8**	**9.9**	**10**
0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
1	0.0010	0.0009	0.0009	0.0008	0.0007	0.0007	0.0006	0.0005	0.0005	0.0005
2	0.0046	0.0043	0.0040	0.0037	0.0034	0.0031	0.0029	0.0027	0.0025	0.0023
3	0.0140	0.0131	0.0123	0.0115	0.0107	0.0100	0.0093	0.0087	0.0081	0.0076
4	0.0319	0.0302	0.0285	0.0269	0.0254	0.0240	0.0226	0.0213	0.0201	0.0189
5	0.0581	0.0555	0.0530	0.0506	0.0483	0.0460	0.0439	0.0418	0.0398	**0.0378**
6	0.0881	0.0851	0.0822	0.0793	0.0764	0.0736	0.0709	0.0682	0.0656	0.0631
7	0.1145	0.1118	0.1091	0.1064	0.1037	0.1010	0.0982	0.0955	0.0928	0.0901
8	0.1302	0.1286	0.1269	0.1251	0.1232	0.1212	0.1191	0.1170	0.1148	0.1126
9	0.1317	0.1315	0.1311	0.1306	0.1300	0.1293	0.1284	0.1274	0.1263	0.1251
10	0.1198	0.1210	0.1219	0.1228	0.1235	0.1241	0.1245	0.1249	0.1250	0.1251
11	0.0991	0.1012	0.1031	0.1049	0.1067	0.1083	0.1098	0.1112	0.1125	0.1137
12	0.0752	0.0776	0.0799	0.0822	0.0844	0.0866	0.0888	0.0908	0.0928	0.0948
13	0.0526	0.0549	0.0572	0.0594	0.0617	0.0640	0.0662	0.0685	0.0707	0.0729
14	0.0342	0.0361	0.0380	0.0399	0.0419	0.0439	0.0459	0.0479	0.0500	0.0521
15	0.0208	0.0221	0.0235	0.0250	0.0265	0.0281	0.0297	0.0313	0.0330	0.0347
16	0.0118	0.0127	0.0137	0.0147	0.0157	0.0168	0.0180	0.0192	0.0204	0.0217
17	0.0063	0.0069	0.0075	0.0081	0.0088	0.0095	0.0103	0.0111	0.0119	0.0128
18	0.0032	0.0035	0.0039	0.0042	0.0046	0.0051	0.0055	0.0060	0.0065	0.0071
19	0.0015	0.0017	0.0019	0.0021	0.0023	0.0026	0.0028	0.0031	0.0034	0.0037
20	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012	0.0014	0.0015	0.0017	0.0019
21	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009
22	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004
23	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

Suppose that major defects occur at the average rate of two per mile. We want to find the probability that no major defects will occur in a particular 3-mile section of the highway. The interval length is 3 miles, so $\lambda = (2 \text{ defects/mile})(3 \text{ miles}) = 6$ represents the expected number of major defects over the 3-mile section of highway. Thus, by using equation (3.8) or Appendix C with $\lambda = 6$ and $x = 0$, we obtain the probability of no major defects of 0.0025. Thus, finding no major defects in the 3-mile section is very unlikely. In fact, there is a $1 - 0.0025 = 0.9975$ probability of at least one major defect in that section of highway.

NOTES AND COMMENTS

1. When working with the Poisson probability distribution, you need to be sure that λ is the mean number of occurrences for the desired interval. For instance, suppose that you know that 30 calls come into a switchboard every 15 minutes. To compute Poisson probabilities for the number of calls coming in over a 5-minute period, you would use $\lambda = 10$; to compute probabilities for the number of calls coming in over a 1-minute period, you would use $\lambda = 2$.

3.5 Continuous Random Variables

In this section we introduce probability distributions for continuous random variables. Recall from Section 3.1 that random variables that may assume any value in a certain interval or collection of intervals are said to be *continuous.* Examples of continuous random variables include the following:

1. The *number of ounces* of soup placed in a can labeled "8 ounces"
2. The *flight time* of an airplane traveling from Chicago to New York
3. The *lifetime* of the monitor of a new laptop computer
4. The *drilling depth* required to reach oil in an offshore drilling operation

To understand the nature of continuous random variables more fully, suppose that, in the first example, one can of soup has 8.2 ounces and another 8.3 ounces. Other cans could weigh 8.25 ounces, 8.225 ounces, and so on. In fact, the actual weight can be any numeric value from 0 ounces for an empty can to, say, 8.5 ounces for a can filled to capacity. Because this interval contains infinitely many values, we can no longer list each value of the random variable and then identify its associated probability. In fact, for continuous random variables we need a new method for computing the probabilities associated with the values of the random variable.

Applying the Uniform Distribution

Let x denote the flight time of an airplane traveling from Chicago to New York. Assume that the minimum time is 2 hours and that the maximum time is 2 hours 20 minutes. Thus, in terms of minutes, the flight time can be any value in the interval from 120 minutes to 140 minutes (e.g., 124 minutes, 125.48 minutes, etc.). As the random variable x can take on any value from 120 to 140 minutes, x is a continuous rather than a discrete random variable. Assume that sufficient actual flight data are available to conclude that the probability of a flight time between 120 and 121 minutes is the same as the probability of a flight time within any other 1-minute interval up to and including 140 minutes. With every 1-minute interval being equally likely, the random variable x has a **uniform probability distribution**. The following **probability density function** describes the uniform probability distribution for the flight time random variable:

$$f(x) = \begin{cases} \dfrac{1}{20} & \text{for } 120 \le x \le 140 \\ 0 & \text{elsewhere} \end{cases} \quad \textbf{(3.9)}$$

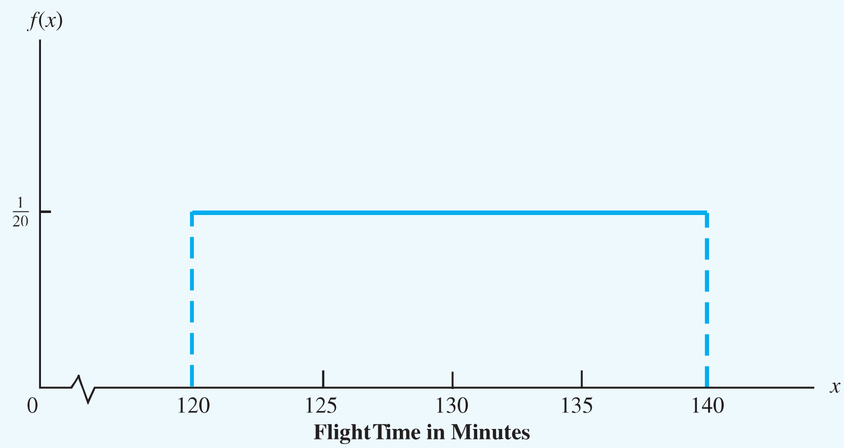
Figure 3.3 shows a graph of this probability density function. In general, the uniform probability density function for a random variable x is

$$f(x) = \begin{cases} \dfrac{1}{b - a} & \text{for } a \le x \le b \\ 0 & \text{elsewhere} \end{cases} \quad \textbf{(3.10)}$$

In the flight time example, $a = 120$ and $b = 140$.

In the graph of a probability density function, $f(x)$ shows the height or value of the function at any particular value of x. Because the probability density function for flight time is

FIGURE 3.3 UNIFORM PROBABILITY DENSITY FUNCTION FOR FLIGHT TIME



uniform, the height or value of the function is the same for each value of x between 120 and 140. That is, $f(x) = {}^1/_{20}$ for all values of x between 120 and 140. The probability density function $f(x)$, unlike the probability function for a discrete random variable, represents the height of the function at any particular value of x and *not* probability. Recall that, for each value of a discrete random variable (say, $x = 2$), the probability function yielded the probability of x having *exactly* that value [that is, $f(2)$]. However, a continuous random variable has infinitely many values, so we can no longer identify the probability for each specific value of x. Rather, we must consider probability in terms of the likelihood that a random variable takes on a value within a *specified interval.* For instance, in the flight time example an acceptable probability question is: What is the probability that the flight time is between 120 and 130 minutes? That is, what is $P(120 \leq x \leq 130)$? As the flight time must be between 120 and 140 minutes and as the probability is uniformly distributed over this interval, we feel comfortable saying that $P(120 \leq x \leq 130) = 0.50$. Indeed, as we will show, this is correct.

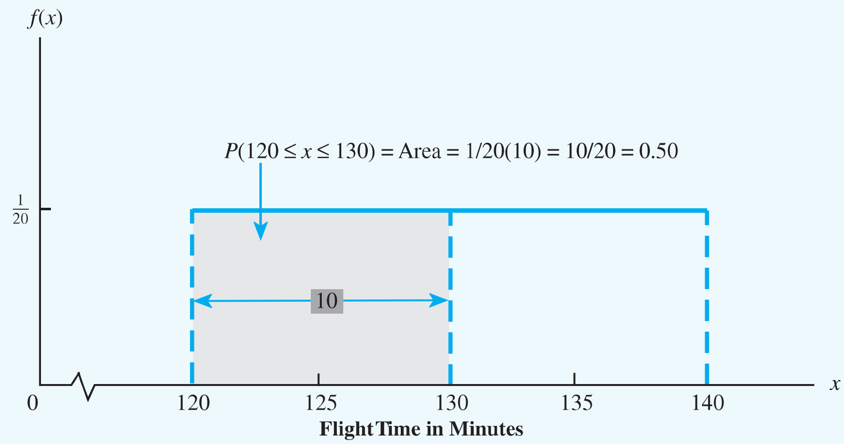
Area as a Measure of Probability

Refer to Figure 3.4 and consider the *area under the graph of* $f(x)$ over the interval from 120 to 130. Note that the region is rectangular in shape and that the area of a rectangle is simply the width times the height. With the width of the interval equal to $130 - 120 = 10$ and the height of the graph $f(x) = {}^1/_{20}$, the area $=$ width $\times$ height $= 10({}^1/_{20}) = {}^{10}/_{20} = 0.50$.

What observation can you make about the area under the graph of $f(x)$ and probability? They are identical! Indeed, that is true for all continuous random variables. In other words, once you have identified a probability density function $f(x)$ for a continuous random variable, you can obtain the probability that x takes on a value between some lower value a and some higher value b by computing the *area* under the graph of $f(x)$ over the interval a to b.

With the appropriate probability distribution and the interpretation of area as probability, we can answer any number of probability questions. For example, what is the probability of a flight time between 128 and 136 minutes? The width of the interval is $136 - 128 = 8$. With the uniform height of ${}^1/_{20}$, $P(128 \leq x \leq 136) = {}^8/_{20} = 0.40$.

FIGURE 3.4 AREA PROVIDES PROBABILITY OF FLIGHT TIME



Note that $P(120 \leq x \leq 140) = 20(^1/_{20}) = 1$; the total area under the $f(x)$ graph is equal to 1. This property holds for all continuous probability distributions and is the analog of the requirement that the sum of the probabilities must equal 1 for a discrete probability distribution. A continuous probability distribution also requires that $f(x) \geq 0$ for all values of x. It is the analog of the requirement that $f(x) \geq 0$ for discrete probability distributions.

Two principal differences between continuous random variables and probability distributions and their discrete counterparts stand out.

Try Problem 18 to practice computing probabilities using the uniform probability distribution.

1. We no longer talk about the probability of the random variable taking on a particular value. Instead we talk about the probability of the random variable taking on a value within some given interval.
2. The probability of the random variable taking on a value within some given interval is defined to be the area under the graph of the probability density function over the interval. This definition implies that the probability that a continuous random variable takes on any particular value is zero because the area under the graph of $f(x)$ at a single point is zero.

NOTES AND COMMENTS

1. For any continuous random variable the probability of any particular value is zero, so $P(a \leq x \leq b) = P(a < x < b)$. Thus, the probability of a random variable assuming a value in any interval is the same whether the endpoints are included or not.
2. To see more clearly why the height of a probability density function is not a probability, think about a random variable with a uniform probability distribution of

$$f(x) = \begin{cases} 2 & \text{for } 0 \leq x \leq 0.5 \\ 0 & \text{elsewhere} \end{cases}$$

The height of the probability density function is 2 for values of x between 0 and 0.5. But we know that probabilities can never be greater than 1.

3.6 Normal Probability Distribution

Perhaps the most important probability distribution used to describe a continuous random variable is the **normal probability distribution**. It is applicable in a great many practical problem situations, and its probability density function has the form of the bell-shaped curve shown in Figure 3.5. The mathematical function that provides the bell-shaped curve of the normal probability density function follows:

The normal distribution was first observed by Abraham de Moivre, a French mathematician, in the early 1700s. De Moivre's work was motivated by the study of probability associated with gambling and games of chance.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad \text{for } -\infty < x < \infty \tag{3.11}$$

where

μ = mean or expected value of the random variable x
σ^2 = variance of the random variable x
σ = standard deviation of the random variable x
π = 3.14159
e = 2.71828

Recall from the previous discussion of continuous random variables that $f(x)$ is the height of the curve at a particular value of x. Thus, once we specify the mean (μ) and either the standard deviation (σ) or variance (σ^2), we can use equation (3.11) to determine the graph for the corresponding normal distribution. Figure 3.6 shows two normal distributions, one with $\mu = 50$ and $\sigma = 15$ and another with $\mu = 50$ and $\sigma = 7.5$. Note in particular the effect that the standard deviation σ has on the general shape of the normal curve. A larger standard deviation tends to flatten and broaden the curve because larger values of σ indicate greater variability in the values of the random variable.

Fortunately, whenever we use the normal distribution to answer probability questions, we do not have to use the probability density function of equation (3.11). In fact, when we use the normal distribution, we will have tables of probability values [areas under the $f(x)$ curve] that can provide the desired probability information. To show how to use the tables of areas or probabilities for the normal distribution, we must first introduce the standard normal distribution.

FIGURE 3.5 NORMAL PROBABILITY DISTRIBUTION

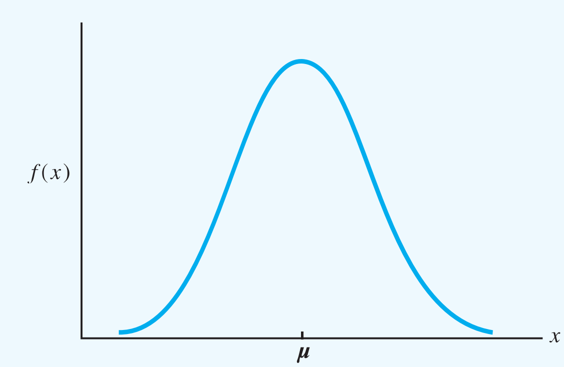
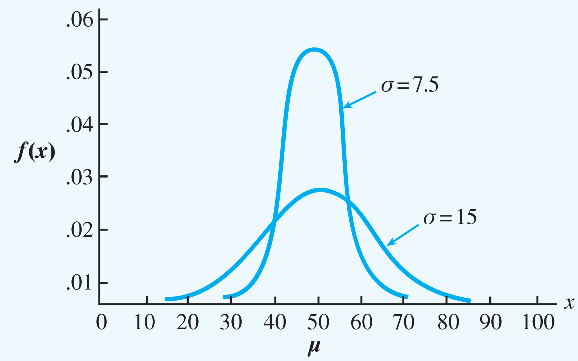


FIGURE 3.6 TWO NORMAL DISTRIBUTIONS WITH $\mu = 50$



Standard Normal Distribution

A random variable that has a normal distribution with a mean of 0 and a standard deviation of 1 is said to have a **standard normal distribution**. We use the letter z to designate this particular normal random variable. Figure 3.7 shows the graph of the standard normal distribution. Note that it has the same general appearance as other normal distributions, but with the special properties of $\mu = 0$ and $\sigma = 1$. The units on the horizontal axis (z) measure the number of standard deviations from the mean.

Recall the procedure for finding probabilities associated with a continuous random variable. We want to determine the probability of the random variable having a value in a specified interval from a to b. Thus we have to find the area under the curve in the interval from a to b. In the preceding section we showed that finding probabilities, or areas under the curve, for a uniform distribution was relatively easy. All we had to do was multiply the width of the interval by the height of the graph. However, finding areas under the normal distribution curve appears at first glance to be much more difficult because the height of the curve varies. The mathematical technique for obtaining these areas is beyond the scope of the text, but fortunately tables are available that provide the areas or probability values for the standard normal distribution. Table 3.9 is such a table of areas. This table is also available as Appendix D.

FIGURE 3.7 STANDARD NORMAL DISTRIBUTION

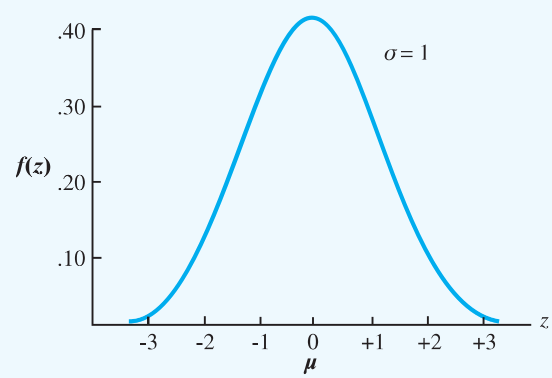
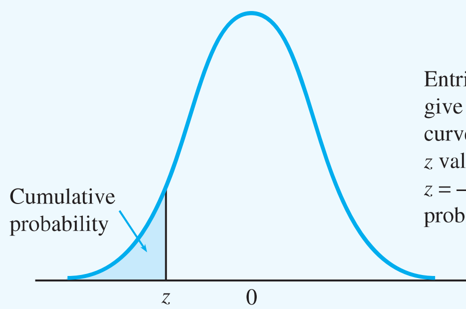


TABLE 3.9 CUMULATIVE PROBABILITIES FOR THE STANDARD NORMAL DISTRIBUTION

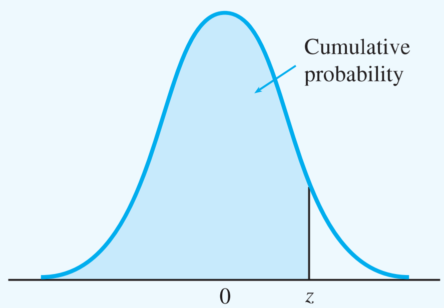


Entries in the table give the area under the curve to the left of the z value. For example, for $z = -0.85$, the cumulative probability is 0.1977.

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	**0.1977**	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

(Continued)

TABLE 3.9 CUMULATIVE PROBABILITIES FOR THE STANDARD NORMAL DISTRIBUTION (*Continued*)



Entries in the table give the area under the curve to the left of the z value. For example, for $z = 1.25$, the cumulative probability is 0.8944.

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	**0.8944**	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9913
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

The graphs at the top of Table 3.9 show that the area in the table is the probability of a standard normal random variable being less than or equal to a specific value of z. Such probabilities are referred to as **cumulative probabilities**. In using Table 3.9 to determine a cumulative probability, note that the value of z with one decimal appears in the left-hand column of the table, with the second decimal appearing in the top row. Negative values of z are provided on the first page of the table, whereas positive values of z are provided on the second page. For example, for $z = -0.85$, we find -0.8 in the left-hand column and the second decimal 0.05 in the top row of the first page of the table. Then, by looking in the body of the table, we find an area or probability of 0.1977. This is the cumulative probability that the standard normal random variable is less than or equal to $z = -0.85$. This area is shown graphically at the top of Table 3.9. As another example, we can use the second page of the table to determine that the cumulative probability that the standard normal random variable is less than or equal to $z = 1.25$. We find 1.2 in the left-hand column and the second decimal 0.05 in the top row of the second page of the table. In the body of the table we find an area or probability of 0.8944. This is the cumulative probability that the standard normal random variable is less than or equal to $z = 1.25$. This area is also shown graphically at the top of Table 3.9.

Suppose that we wanted to use the cumulative standard normal distribution table to determine the probability that the standard normal random variable z will be between -1.00 and $+1.00$. Table 3.9 shows the cumulative probability that z is less than or equal to $+1.00$ is 0.8413 and the cumulative probability that z is less than or equal to -1.00 is 0.1587. Thus, the probability that z will be between -1.00 and $+1.00$ must be the difference between these two cumulative probabilities: $0.8413 - 0.1587 = 0.6826$. This is shown graphically in Figure 3.8.

The probability that a standard normal random variable z is between a and b is always the difference between two cumulative probabilities: one for $z = b$ and one for $z = a$.

Similarly, we can find the probability that the standard normal random variable z will be between -2.00 and $+2.00$. Using the cumulative probabilities at $z = +2.00$ and $z = -2.00$, the probability that z will be between -2.00 and $+2.00$ is $0.9772 - 0.0228 = 0.9544$. In addition, we can use the cumulative probabilities at $z = +3.00$ and $z = -3.00$ to conclude that the probability z will be between -3.00 and $+3.00$ is $0.9986 - 0.0013 = 0.9973$. Since the total probability or total area under the curve is equal to 1.0000, the probability of 0.9973 tells us that z will almost always fall between -3.00 and $+3.00$.

As a final example, what is the probability that the normal random variable z is greater than 2.00? From Table 3.9, we find that the cumulative probability that z is less than or equal

FIGURE 3.8 PROBABILITY OF z BETWEEN -1.00 AND $+1.00$

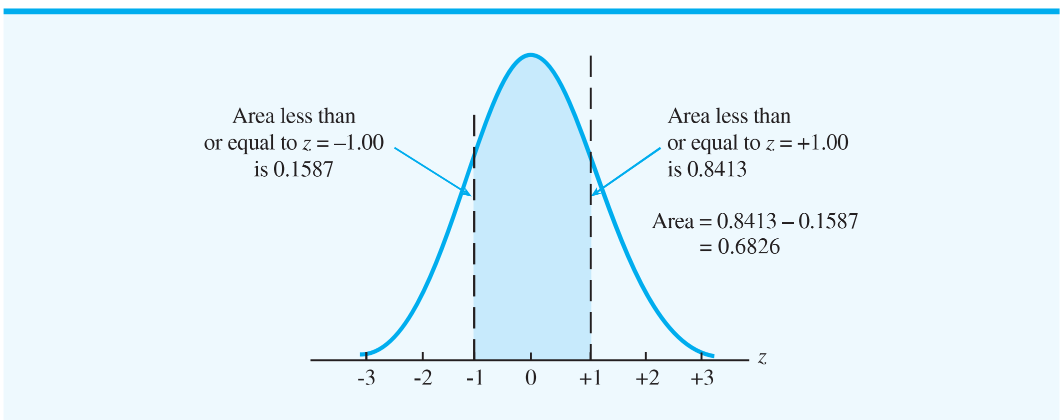
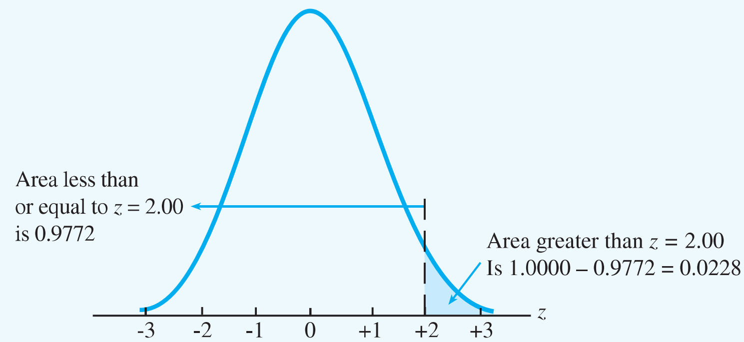


FIGURE 3.9 PROBABILITY OF z GREATER THAN 2.00



to 2.00 is 0.9772. Since the total area under the curve is equal to 1.0000, the probability that z will be greater than 2.00 must be $1.0000 - 0.9772 = 0.0228$. This is shown graphically in Figure 3.9. As the examples in this section have shown, you should be able to use the cumulative probabilities in Table 3.9 to answer a variety of probability questions about the standard normal random variable z.

Computing Probabilities for Any Normal Distribution

The reason we discuss the standard normal distribution so extensively is that we can compute probabilities for any normal distribution by first converting to the standard normal distribution. Thus, when we have a normal distribution with any mean μ and any standard deviation σ, we can answer probability questions about this distribution by converting to the standard normal distribution. We then use Table 3.9 and the appropriate z values to find the probability. The formula used to convert any normal random variable x with mean μ and standard deviation σ to the standard normal distribution is

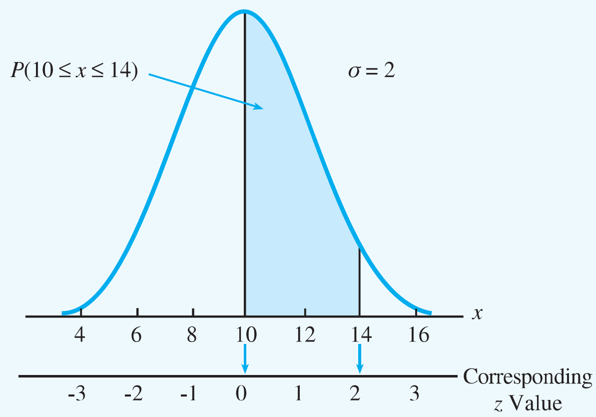
$$z = \frac{x - \mu}{\sigma} \quad (3.12)$$

When used in this way, z is a measure of the number of standard deviations that x is from μ.

We can use an example to show most easily how the conversion to the z value allows us to use the standard normal distribution to compute probabilities for any normal distribution. Suppose that we have a normal distribution with $\mu = 10$ and $\sigma = 2$, as shown graphically in Figure 3.10. Note that, in addition to the values of the random variable shown on the x axis, we have included a second axis (the z axis) to show that for each value of x there is a corresponding value of z. For example, when $x = 10$, the corresponding z value (the number of standard deviations away from the mean) is $z = (x - \mu)/\sigma = (10 - 10)/2 = 0$. Similarly, for $x = 14$ we have $z = (x - \mu)/\sigma = (14 - 10)/2 = 2$.

Now suppose that we want to know the probability that the random variable x is between 10 and 14; that is, $P(10 \leq x \leq 14)$. We do not have tables that provide this probability directly. However, note that in Figure 3.10 the area under the curve (probability) for x between 10 and 14 is the same as the area under the curve for z between 0 and 2. Using

FIGURE 3.10 NORMAL DISTRIBUTION WITH $\mu = 10$ AND $\sigma = 2$



$z = 2.00$ and Table 3.9, we find that the cumulative probability of z being less than or equal to 2.00 is 0.9772. Similarly, Table 3.9 shows that the cumulative probability of z being less than or equal to 0.00 is 0.5000. Thus, the probability that the standard normal random variable z is between 0.00 and 2.00 is $0.9772 - 0.5000 = 0.4772$. Thus, we conclude that the probability of x being between 10 and 14 is also 0.4772.

This procedure applies to any normal distribution problem. That is, for any x value a corresponding z value is given by equation (3.12). To find the probability that x is in a specified interval, simply convert the x interval to its corresponding z interval. Then use the table for the standard normal distribution to answer the probability question.

Grear Tire Company Problem

Suppose that Grear Tire Company just developed a new steel-belted radial tire that will be sold through a national chain of discount stores. Because the tire is a new product, Grear's management believes that the mileage guarantee offered with the tire will be an important factor in consumer acceptance of the product. Before finalizing the tire mileage guarantee policy, Grear's management wants some probability information concerning the number of miles the tires will last.

From actual road tests with the tires, Grear's engineering group estimates the mean tire mileage at $\mu = 36{,}500$ miles and the standard deviation at $\sigma = 5000$ miles. In addition, the data collected indicate that a normal distribution is a reasonable assumption.

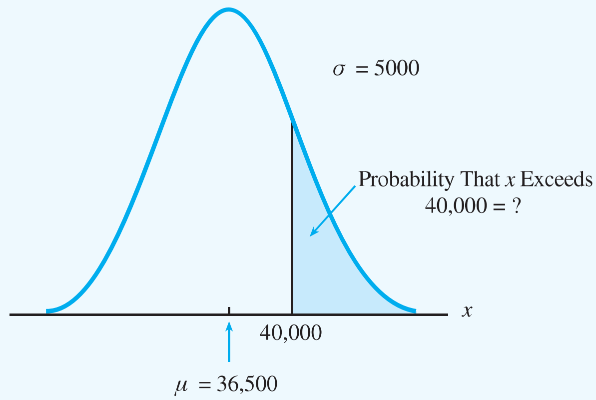
What percentage of the tires, then, can be expected to last more than 40,000 miles? In other words, what is the probability that the tire mileage will exceed 40,000? To compute this probability, we need to find the area of the shaded region in Figure 3.11.

At $x = 40{,}000$ we have

$$z = \frac{x - \mu}{\sigma} = \frac{40{,}000 - 36{,}500}{5000} = \frac{3500}{5000} = 0.70$$

Thus the probability that the normal distribution for tire mileage will have an x value greater than 40,000 is the same as the probability that the standard normal distribution will have a

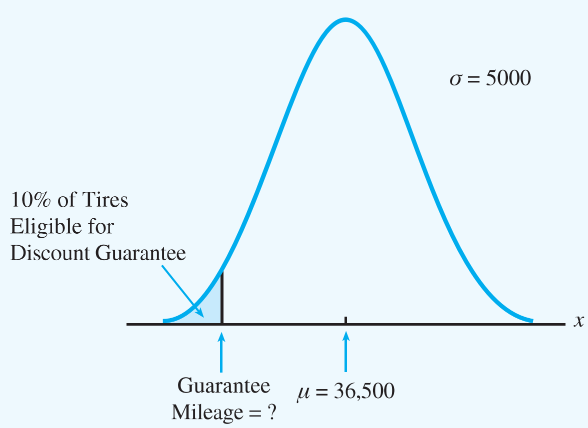
FIGURE 3.11 GREAR TIRE COMPANY TIRE MILEAGE



z value greater than 0.70. Using Table 3.9, we find that the cumulative probability that z is less than or equal to 0.70 is 0.7580. Thus, the probability that z will be greater than 0.70 must be $1.0000 - 0.7580 = 0.2420$. In terms of the tire mileage x, we can conclude that there is a 0.2420 probability that x will be greater than 40,000 miles. Thus, we can anticipate about 24.2% of the tires manufactured by Grear will last more than 40,000 miles.

Let us now assume that Grear is considering a guarantee that will provide a discount on a new set of tires if the mileage on the original tires doesn't exceed the mileage stated on the guarantee. What should the guarantee mileage be if Grear wants no more than 10% of the tires to be eligible for the discount? This question is interpreted graphically in Figure 3.12. Note that 10% of the area is below the unknown guarantee mileage. Because this 10% is the lower tail of the normal probability distribution, 0.1000 is the cumulative probability that the tire mileage will be less than or equal to the unknown guarantee mileage. The question is

FIGURE 3.12 GREAR'S DISCOUNT GUARANTEE



now, How many standard deviations (z value) do we have to be below the mean to a achieve a 0.1000 cumulative probability? This time we will look into the body of Table 3.9 and try to find the cumulative probability 0.1000. We cannot find 0.1000 exactly, but a cumulative probability 0.1003 is close. Here we find the corresponding $z = -1.28$. This tells us that we must be 1.28 below the mean to find the desired tire guarantee mileage. This mileage is

$$\begin{aligned}\text{Guarantee mileage} &= \mu - 1.28\sigma \\ &= 36{,}500 - 1.28(5000) = 30{,}100\end{aligned}$$

Therefore, a guarantee of 30,100 miles will meet the requirement that approximately 10% of the tires will be eligible for the discount. With this information the firm might confidently set its tire mileage guarantee at 30,000 miles.

Try Problem 23 for practice finding a z value that cuts off a particular probability.

Again we see the important role that probability distributions play in providing decision-making information. Once a probability distribution is established for a particular problem, it can be used rather quickly and easily to provide information about the likelihood of various scenarios. Although this information does not make a decision recommendation directly, it does provide information that helps the decision maker understand the problem better. Ultimately, this information may assist the decision maker in reaching a good decision.

3.7 Exponential Probability Distribution

A continuous probability distribution that is often useful in describing the time needed to complete a task is the **exponential probability distribution**. The exponential random variable can be used to describe the time between arrivals at a car wash, the time required to load a truck, the distance between major defects in a highway, and so on. The exponential probability density function is

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \qquad \text{for } x \geq 0, \mu > 0 \tag{3.13}$$

To provide an example of the exponential probability distribution, suppose that the loading time for a truck at a factory dock follows an exponential probability distribution. If the mean, or average, loading time is 15 minutes ($\mu = 15$), the appropriate probability density function is

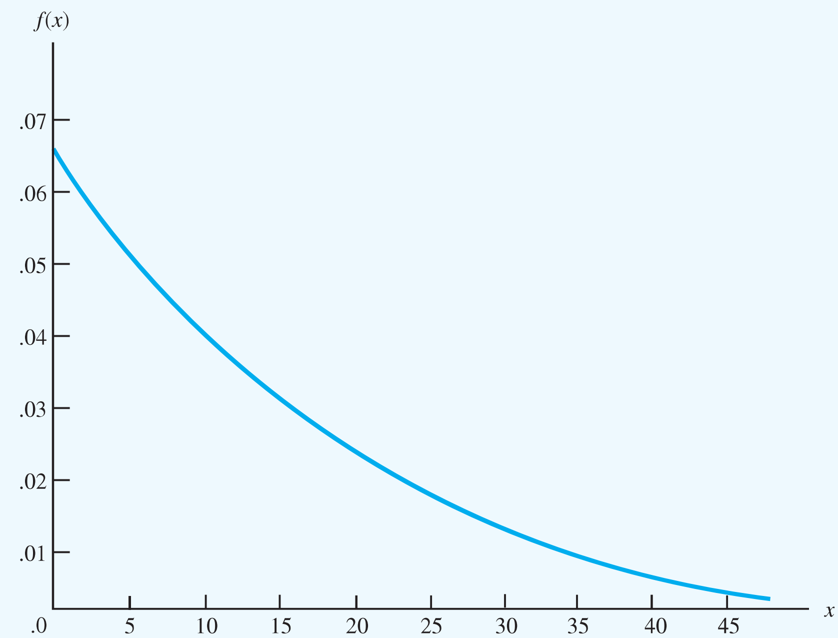
$$f(x) = \frac{1}{15} e^{-x/15}$$

Figure 3.13 shows the graph of this density function.

Computing Probabilities for the Exponential Distribution

As with any continuous probability distribution, the area under the curve corresponding to some interval provides the probability that the random variable takes on a value in that interval. For instance, for the factory loading dock example, the probability that 6 minutes

FIGURE 3.13 EXPONENTIAL DISTRIBUTION FOR LOADING TIME AT THE LOADING DOCK WITH $\mu = 15$



or less ($x \le 6$) are needed to load a truck is defined to be the area under the curve from $x = 0$ to $x = 6$. Similarly, the probability that a loading time is 18 minutes or less ($x \le 18$) is the area under the curve from $x = 0$ to $x = 18$. Note also that the probability of loading a truck in between 6 and 18 minutes ($6 \le x \le 18$) is the area under the curve from $x = 6$ to $x = 18$.

To compute exponential probabilities such as those previously described, the following formula provides the probability of obtaining a value for the exponential random variable of less than or equal to some specific value of x, denoted by x_0:

$$P(x \le x_0) = 1 - e^{-x_0/\mu} \tag{3.14}$$

Thus, for the factory loading dock example, equation (3.14) becomes

$$P(\text{loading time} \le x_0) = 1 - e^{-x_0/15}$$

Hence, the probability that 6 minutes or less ($x \le 6$) are needed to load a truck is

$$P(\text{loading time} \le 6) = 1 - e^{-6/15} = 0.3297$$

Note also that the probability that 18 minutes or less ($x \le 18$) are needed to load a truck is

$$P(\text{loading time} \le 18) = 1 - e^{-18/15} = 0.6988$$

Try Problem 29 for practice finding probabilities with the exponential probability distribution.

Thus, the probability that 6 to 18 minutes are required to load a truck is 0.6988 − 0.3297 = 0.3691. Probabilities for any other interval can be computed in a similar manner.

Relationship Between the Poisson and Exponential Distributions

In Section 3.4 we introduced the Poisson distribution as a discrete probability distribution that often is useful when we are dealing with the number of occurrences over a specified interval of time or space. Recall that the Poisson probability function is

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

where

$$\lambda = \text{expected value or mean number of occurrences in an interval}$$

The Poisson and exponential probability distributions are used in the chapter on waiting line models (Chapter 15). In these models the Poisson distribution is used as the probability distribution for the number of arrivals, while the exponential probability distribution is used as the probability distribution for the service time.

The continuous exponential probability distribution is related to the discrete Poisson distribution. The Poisson distribution provides an appropriate description of the number of occurrences per interval, and the exponential distribution provides a description of the length of the interval between occurrences.

To illustrate this relationship, let us suppose that the number of cars that arrive at a car wash during 1 hour is described by a Poisson probability distribution with a mean of 10 cars per hour. Thus the Poisson probability function that provides the probability of x arrivals per hour is

$$f(x) = \frac{10^x e^{-10}}{x!}$$

The average number of arrivals is 10 cars per hour, so the average time between cars arriving is

$$\frac{1 \text{ hour}}{10 \text{ cars}} = 0.1 \text{ hour/car}$$

Try Problems 30 and 31 for applications of the exponential probability distribution.

Thus, the corresponding exponential distribution that describes the time between the arrival of cars has a mean of $\mu = 0.1$ hours per car. The appropriate exponential probability density function is

$$f(x) = \frac{1}{0.1} e^{-x/0.1} = 10e^{-10x}$$

Summary

In this chapter we continued the discussion of probability by introducing the important concepts of random variables and probability distributions. Random variables provide numeric descriptions of the outcomes of experiments. When random variables are used, computations of the expected value, variance, and standard deviation can help the decision maker understand characteristics of the problem under study. We discussed the probability distributions for both discrete and continuous random variables.

Of particular interest are special probability distributions such as the binomial, Poisson, uniform, normal, and exponential distributions. These distributions provide wide applicability, and special formulas and/or tables make the probability information easily available.

Through a variety of problems and applications, we illustrated the role that probability distributions play in providing decision-making information. Although the probability values generated by the techniques and methods of this chapter do not by themselves make decision recommendations, they do provide assistance to the decision maker in terms of understanding the uncertainties inherent in the problem. Ultimately, this better understanding may lead to new and better decisions.

Glossary

Random variable A numeric description of the outcome of an experiment.

Discrete random variable A random variable that may assume only a finite or infinite sequence of values.

Continuous random variable A random variable that may assume any value in an interval or collection of intervals.

Probability function A function, denoted $f(x)$, that provides the probability that a discrete random variable x takes on some specific value.

Discrete probability distribution A table, graph, or equation describing the values of the random variable and the associated probabilities.

Expected value A weighted average of the values of the random variable, for which the probability function provides the weights. If an experiment can be repeated a large number of times, the expected value can be interpreted as the "long-run average."

Variance A measure of the dispersion or variability in the random variable. It is a weighted average of the squared deviations from the mean, μ.

Standard deviation The positive square root of the variance.

Binomial probability distribution The probability distribution for a discrete random variable, used to compute the probability of x successes in n trials.

Poisson probability distribution The probability distribution for a discrete random variable, used to compute the probability of x occurrences over a specified interval.

Uniform probability distribution A continuous probability distribution in which the probability that the random variable will assume a value in any interval of equal length is the same for each interval.

Probability density function The function that describes the probability distribution of a continuous random variable.

Normal probability distribution A continuous probability distribution whose probability density function is bell shaped and determined by the mean, μ, and standard deviation, σ.

Standard normal distribution A normal distribution with a mean of 0 and a standard deviation of 1.

Cumulative probability The probability that a random variable takes on a value less than or equal to stated value.

Exponential probability distribution A continuous probability distribution that is useful in describing the time to complete a task or the time between occurrences of an event.

Problems

SELF test

1. The following examples are experiments and their associated random variables. In each case identify the values the random variable can assume and state whether the random variable is discrete or continuous:

Experiment	Random Variable (x)
a. Take a 20-question examination	Number of questions answered correctly
b. Observe cars arriving at a tollbooth for 1 hour	Number of cars arriving at the tollbooth
c. Audit 50 tax returns	Number of returns containing errors
d. Observe an employee's work for 8 hours	Number of nonproductive hours
e. Weigh a shipment of goods	Number of pounds

2. The following table shows a partial probability distribution for the MRA Company's projected profits (in thousands of dollars) for the first year of operation (the negative value denotes a loss):

x	$f(x)$
−100	0.10
0	0.20
50	0.30
100	0.25
150	0.10
200	

   a. Find the missing value of $f(200)$. What is your interpretation of this value?
   b. What is the probability that MRA will be profitable?
   c. What is the probability that MRA will make at least $100,000?

SELF test

3. Data were collected on the number of operating rooms in use at Tampa General Hospital over a 20-day period. On 3 of the days only one operating room was used; on 5 days, two were used; on 8 days, three were used; and on 4 days all four rooms were used.
   a. Use the relative frequency approach to construct a probability distribution for the number of operating rooms in use on any given day.
   b. Draw a graph of the probability distribution.
   c. Show that your probability distribution satisfies the requirements for a valid discrete probability distribution.

SELF test

4. Shown is a probability distribution for the random variable x.

x	$f(x)$
3	0.25
6	0.50
9	0.25
Total	1.00

a. Compute $E(x)$, the expected value of x.
b. Compute σ^2, the variance of x.
c. Compute σ, the standard deviation of x.

5. Brandon Lang is a creative entrepreneur who has developed a novelty soap item called Jackpot to target consumers with a gambling habit. Inside each bar of Jackpot shower soap is a single rolled-up bill of U.S. currency. The currency (rolled up and sealed in shrink-wrap) is appropriately inserted into the soap mixture prior to the cutting and stamping procedure. The distribution of paper currency (per 1000 bars of soap) is given in the following table.

Distribution of Paper Currency Prizes

Bill Denomination	Number of Bills
\$1	520
\$5	260
\$10	120
\$20	70
\$50	29
\$100	1
Total	1000

a. What is the expected amount of money in a single bar of Jackpot soap?
b. What is the standard deviation of the money in a single bar of Jackpot soap?
c. How many bars of soap would a customer have to buy so that, on average, he or she has purchased three bars containing a \$50 or \$100 bill?
d. If a customer buys 8 bars of soap, what is the probability that at least one of these bars contains a bill of \$20 or larger?

6 The National Center for Health Statistics reported the following data on the number of children born in individual pregnancies in 1996 and 2006 (*The World Almanac*, 2010):

Number of Children	1996 Frequency	2006 Frequency
one child	3,671,455	3,971,276
twins	100,750	137,085
triplets	5,298	6,118
quadruplets	560	355
quintuplets or more	81	67

a. Define a random variable x = number of children born in a single pregnancy in 1996 and develop a probability distribution for the random variable. Let $x = 5$ represent quintuplets or more.
b. Compute the expected value and variance for the number of children born in a single pregnancy in 1996.
c. Define a random variable y = number of children born in a single pregnancy in 2006 and develop a probability distribution for the random variable. Let $y = 5$ represent quintuplets or more.
d. Compute the expected value and variance for the number of children born in a single pregnancy in 2006.
e. Do these data support the conclusion that the increased use of fertility drugs by older women has generated an upward trend in multiple births?

7. The demand for Carolina Industries' product varies greatly from month to month. Based on the past two years of data, the following probability distribution shows the company's monthly demand:

Unit Demand	Probability
300	0.20
400	0.30
500	0.35
600	0.15

a. If the company places monthly orders equal to the expected value of the monthly demand, what should Carolina's monthly order quantity be for this product?
b. Assume that each unit demanded generates \$70 in revenue and that each unit ordered costs \$50. How much will the company gain or lose in a month if it places an order based on your answer to part (a) and the actual demand for the item is 300 units?
c. What are the variance and standard deviation for the number of units demanded?

8. The J. R. Ryland Computer Company is considering a plant expansion that will enable the company to begin production of a new computer product. The company's president must determine whether to make the expansion a medium- or large-scale project. The demand for the new product involves an uncertainty, which for planning purposes may be low demand, medium demand, or high demand. The probability estimates for the demands are 0.20, 0.50, and 0.30, respectively. Letting x indicate the annual profit in \$1000s, the firm's planners developed profit forecasts for the medium- and large-scale expansion projects.

		Medium-Scale Expansion Profits		Large-Scale Expansion Profits	
		x	$f(x)$	y	$f(y)$
	Low	50	0.20	0	0.20
Demand	**Medium**	150	0.50	100	0.50
	High	200	0.30	300	0.30

a. Compute the expected value for the profit associated with the two expansion alternatives. Which decision is preferred for the objective of maximizing the expected profit?
b. Compute the variance for the profit associated with the two expansion alternatives. Which decision is preferred for the objective of minimizing the risk or uncertainty?

SELF test

9. Consider a binomial experiment with 2 trials and $p = 0.4$.
a. Compute the probability of 1 success, $f(1)$.
b. Compute $f(0)$.
c. Compute $f(2)$.
d. Find the probability of at least one success.
e. Find the expected value, variance, and standard deviation.

10. Consider a binomial experiment with $n = 10$ and $p = 0.10$. Use the binomial tables (Appendix B) to answer parts (a) through (d).
a. Find $f(0)$.
b. Find $f(2)$.
c. Find $P(x \leq 2)$.

d. Find $P(x \geq 1)$.
e. Find $E(x)$.
f. Find Var(x) and σ.

11. A survey on British Social Attitudes asked respondents if they had ever boycotted goods for ethical reasons (*Statesman*, January 28, 2008). The survey found that 23% of the respondents have boycotted goods for ethical reasons.
    a. In a sample of six British citizens, what is the probability that two have ever boycotted goods for ethical reasons?
    b. In a sample of six British citizens, what is the probability that at least two respondents have boycotted goods for ethical reasons?
    c. In a sample of ten British citizens, what is the probability that none have boycotted goods for ethical reasons?

SELF test

12. When a new machine is functioning properly, only 3% of the items produced are defective. Assume that we will randomly select two parts produced on the machine and that we are interested in the number of defective parts found.
    a. Describe the conditions under which this situation would be a binomial experiment.
    b. How many experimental outcomes yield one defect?
    c. Compute the probabilities associated with finding no defects, one defect, and two defects.

13. Military radar and missile detection systems are designed to warn a country of enemy attacks. A reliability question deals with the ability of the detection system to identify an attack and issue the warning. Assume that a particular detection system has a 0.90 probability of detecting a missile attack. Answer the following questions using the binomial probability distribution:
    a. What is the probability that one detection system will detect an attack?
    b. If two detection systems are installed in the same area and operate independently, what is the probability that at least one of the systems will detect the attack?
    c. If three systems are installed, what is the probability that at least one of the systems will detect the attack?
    d. Would you recommend that multiple detection systems be operated? Explain.

SELF test

14. Consider a Poisson probability distribution with 2 as the average number of occurrences per time period.
    a. Write the appropriate Poisson probability function.
    b. What is the average number of occurrences in three time periods?
    c. Write the appropriate Poisson probability function to determine the probability of x occurrences in three time periods.
    d. Find the probability of two occurrences in one time period.
    e. Find the probability of six occurrences in three time periods.
    f. Find the probability of five occurrences in two time periods.

15. Telephone calls arrive at the rate of 48 per hour at the reservation desk for Regional Airways.
    a. Find the probability of receiving 3 calls in a 5-minute interval.
    b. Find the probability of receiving 10 calls in 15 minutes.
    c. Suppose that no calls are currently on hold. If the agent takes 5 minutes to complete processing the current call, how many callers do you expect to be waiting by that time? What is the probability that no one will be waiting?
    d. If no calls are currently being processed, what is the probability that the agent can take 3 minutes for personal time without being interrupted?

16. More than 50 million guests stayed at bed and breakfasts (B&Bs) last year. The website for the Bed and Breakfast Inns of North America, which averages approximately seven visitors

per minute, enables many B&Bs to attract guests without waiting years to be mentioned in guidebooks (*Time,* September 2001).

a. What is the probability of no website visitors in a 1-minute period?
b. What is the probability of two or more website visitors in a 1-minute period?
c. What is the probability of one or more website visitors in a 30-second period?
d. What is the probability of five or more website visitors in a 1-minute period?

17. Airline passengers arrive randomly and independently at the passenger screening facility at a major international airport. The mean arrival rate is 10 passengers per minute.
a. What is the probability of no arrivals in a 1-minute period?
b. What is the probability of 3 or fewer arrivals in a 1-minute period?
c. What is the probability of no arrivals in a 15-second period?
d. What is the probability of at least 1 arrival in a 15-second period?

SELF test

18. A random variable x is uniformly distributed between 1.0 and 1.5.
a. Show the graph of the probability density function.
b. Find $P(x = 1.25)$.
c. Find $P(1.00 \leq x \leq 1.25)$.
d. Find $P(1.20 < x < 1.50)$.

19. Delta Airlines quotes a flight time of 2 hours, 5 minutes for its flights from Cincinnati to Tampa. Suppose we believe that actual flight times are uniformly distributed between 2 hours and 2 hours, 20 minutes.
a. Show the graph of the probability density function for flight times.
b. What is the probability that the flight will be no more than 5 minutes late?
c. What is the probability that the flight will be more than 10 minutes late?
d. What is the expected flight time?

20. Most computer languages have a function that can be used to generate random numbers. In Microsoft's Excel, the RAND function can be used to generate random numbers between 0 and 1. If we let x denote the random number generated, then x is a continuous random variable with the probability density function:

$$f(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

a. Graph the probability density function.
b. What is the probability of generating a random number between 0.25 and 0.75?
c. What is the probability of generating a random number with a value less than or equal to 0.30?
d. What is the probability of generating a random number with a value greater than 0.60?

SELF test

21. For the standard normal random variable z, compute the following probabilities:
a. $P(0 \leq z \leq 0.83)$
b. $P(-1.57 \leq z \leq 0)$
c. $P(z > 0.44)$
d. $P(z \geq 20.23)$
e. $P(z < 1.20)$
f. $P(z \leq 20.71)$

22. For the standard normal random variable z, find z for each situation.
a. The area to the left of z is 0.9750.
b. The area between 0 and z is 0.4750.
c. The area to the left of z is 0.7291.
d. The area to the right of z is 0.1314.
e. The area to the left of z is 0.6700.
f. The area to the right of z is 0.3300.

SELF test

23. For the standard normal random variable z, find z for each situation.
   a. The area to the left of z is 0.2119.
   b. The area between $-z$ and z is 0.9030.
   c. The area between $-z$ and z is 0.2052.
   d. The area to the left of z is 0.9948.
   e. The area to the right of z is 0.6915.
24. The demand for a new product is estimated to be normally distributed with $\mu = 200$ and $\sigma = 40$. Let x be the number of units demanded, and find the following probabilities:
   a. $P(180 \leq x \leq 220)$
   b. $P(x \geq 250)$
   c. $P(x \leq 100)$
   d. $P(225 \leq x \leq 250)$
25. In 2003, the average stock price for companies making up the S&P 500 was $30, and the standard deviation was $8.20 (*BusinessWeek*, Special Annual Issue, Spring 2003). Assume the stock prices are normally distributed.
   a. What is the probability that a company will have a stock price of at least $40?
   b. What is the probability that a company will have a stock price no higher than $20?
   c. How high does a stock price have to be to put a company in the top 10%?
26. General Hospital's patient account division has compiled data on the age of accounts receivable. The data collected indicate that the age of the accounts follows a normal distribution with $\mu = 28$ days and $\sigma = 8$ days.
   a. What portion of the accounts is between 20 and 40 days old—that is, $P(20 \leq x \leq 40)$?
   b. The hospital administrator is interested in sending reminder letters to the oldest 15% of accounts. How many days old should an account be before a reminder letter is sent?
   c. The hospital administrator wants to give a discount to those accounts that pay their balance by the twenty-first day. What percentage of the accounts will receive the discount?
27. To boost holiday sales, a jewelry store in Bismarck, North Dakota, is advertising the following promotion: "If more than seven inches of cumulative snow fall on December 24, 25, 26, 27, and 28, you get your money back on all purchase made on December 17." To analyze this promotion, the store manager has collected data and determined that snowfall over this 5-day period in December is normally distributed with an average of 6 inches and standard deviation of 0.559 inches. What is the probability that the store will have to refund the money to its December 17 customers?
28. A machine fills containers with a particular product. The standard deviation of filling weights computed from past data is 0.6 ounces. If only 2% of the containers hold less than 18 ounces, what is the mean filling weight for the machine? That is, what must μ equal? Assume that the filling weights have a normal distribution.

SELF test

29. Consider the exponential probability density function:

$$f(x) = \frac{1}{3} e^{-x/3} \qquad \text{for } x \geq 0$$

   a. Write the formula for $P(x \leq x_0)$.
   b. Find $P(x \leq 2)$.
   c. Find $P(x \geq 3)$.
   d. Find $P(x \leq 5)$.
   e. Find $P(2 \leq x \leq 5)$.

SELF test

30. *Internet Magazine* monitors Internet service providers (ISPs) and provides statistics on their performance. The average time to download a web page for free ISPs is approximately 20 seconds for European web pages (*Internet Magazine,* January 2000). Assume the time to download a web page follows an exponential distribution.
    a. What is the probability that it will take less than 10 seconds to download a web page?
    b. What is the probability that it will take more than 30 seconds to download a web page?
    c. What is the probability that it will take between 10 and 30 seconds to download a web page?

SELF test

31. The time between arrivals of vehicles at a particular intersection follows an exponential probability distribution with a mean of 12 seconds.
    a. Sketch this exponential probability distribution.
    b. What is the probability that the time between vehicle arrivals is 12 seconds or less?
    c. What is the probability that the time between vehicle arrivals is 6 seconds or less?
    d. What is the probability that there will be 30 or more seconds between arriving vehicles?
32. The lifetime (hours) of an electronic device is a random variable with the exponential probability density function:

$$f(x) = \frac{1}{50} e^{-x/50} \qquad \text{for } x \geq 0$$

    a. What is the mean lifetime of the device?
    b. What is the probability that the device fails in the first 25 hours of operation?
    c. What is the probability that the device operates 100 or more hours before failure?
33. The time (in minutes) between telephone calls at an insurance claims office has the exponential probability distribution:

$$f(x) = 0.50e^{-0.50x} \qquad \text{for } x \geq 0$$

    a. What is the mean time between telephone calls?
    b. What is the probability of 30 seconds or less between telephone calls?
    c. What is the probability of 1 minute or less between telephone calls?
    d. What is the probability of 5 or more minutes without a telephone call?
34. Sparagowski & Associates conducted a study of service times at the drive-up window of fast-food restaurants. The average time between placing an order and receiving the order at McDonald's restaurants was 2.78 minutes (*The Cincinnati Enquirer*, July 9, 2000). Waiting times, such as these, frequently follow an exponential distribution.
    a. What is the probability that a customer's service time is less than 2 minutes?
    b. What is the probability that a customer's service time is more than 5 minutes?
    c. What is the probability that a customer's service time is more than 2.78 minutes?

Case Problem Specialty Toys

Specialty Toys, Inc., sells a variety of new and innovative children's toys and believes that the preholiday season is the best time to introduce a new toy. Many families use this time to look for new ideas for December holiday gifts. When Specialty has a new toy with good market potential, it chooses an October market entry date.

In order to get toys in its stores by October, Specialty places one-time orders with its manufacturers in June or July of each year. Demand for children's toys can be highly volatile. If a new toy catches on, a sense of shortage in the marketplace often increases the demand to very high levels and large profits can be realized. On the other hand, new toys can also flop, leaving Specialty stuck with high levels of inventory that must be sold at reduced prices. The most important question the company faces is deciding how many units of a new toy should be purchased to meet expected sales demand. If too few are purchased, sales will be lost; if too many are purchased, profits will be reduced because of low prices realized in clearance sales.

For the coming season, Specialty plans to introduce a new product called Weather Teddy. This variation of a talking teddy bear is made by a company in Taiwan. When a child presses Teddy's hand, the bear begins to talk. With the aid of a built-in barometer, Teddy says one of five responses that predict the weather conditions. The responses range from "It looks to be a very nice day! Have fun" to "I think it may rain today. Don't forget your umbrella." Tests with the product show that even though it is not a perfect weather predictor, its predictions are surprisingly good. Several of Specialty's managers claimed Teddy gave predictions of the weather that were as good as the local television weather forecasters.

Specialty faces the decision of how many Weather Teddy units to order for the coming holiday season. Members of the management team recommended order quantities of 15,000, 18,000, 24,000, and 28,000. Considerable disagreement concerning the market potential is evidenced by the different order quantities suggested. The product management team has asked you for an analysis of the stock-out probabilities for various order quantities, an estimate of the profit potential, and help in making an order quantity recommendation. Specialty expects to sell Weather Teddy for $24, and the cost is $16 per unit. If inventory remains after the holiday season, Specialty will sell all surplus inventory for $5 per unit. After reviewing the sales history of similar products, Specialty's senior sales forecaster predicted an expected demand of 20,000 units with a 0.95 probability that demand would be between 10,000 units and 30,000 units.

Managerial Report

Prepare a managerial report that addresses the following issues and recommends an order quantity for the Weather Teddy product:

1. Use the sales forecaster's prediction to describe a normal probability distribution that can be used to approximate the demand distribution. Sketch the distribution and show its mean and standard deviation.
2. Compute the probability of a stock-out for the order quantities suggested by members of the management team.
3. Compute the projected profit for the order quantities suggested by the management team under three scenarios. Worst case: sales = 10,000 units; most likely case: sales = 20,000 units; and best case: sales = 30,000 units.
4. One of Specialty's managers felt that the profit potential was so great that the order quantity should have a 70% chance of meeting demand and only a 30% chance of any stock-outs. What quantity would be ordered under this policy, and what is the projected profit under the three scenarios in part 3?
5. Provide your own recommendation for an order quantity and note the associated profit projections. Provide a rationale for your recommendation.

Appendix 3.1 Computing Discrete Probabilities with Excel

Excel has the capability of computing probabilities for several discrete probability distributions including the binomial and Poisson. In this appendix, we describe how Excel can be used to compute the probabilities for any binomial probability distribution. The procedures for the Poisson probability distributions are similar to the one we describe for the binomial probability distribution.

Let us return to the Nastke Clothing Store problem, where the binomial probabilities of interest are based on a binomial experiment with $n = 10$ and $p = 0.30$. We assume that the user is interested in the probability of $x = 4$ successes in the 10 trials. The following steps describe how to use Excel to produce the desired binomial probability:

Step 1. Select a cell in the worksheet where you want the binomial probability to appear
Step 2. Select the **Formulas** tab (see Appendix A)
Step 3. Choose the **Insert Function** option
Step 4. When the **Insert Function** dialog box appears:
Choose **Statistical** from the **Or select a category** box
Choose **BINOM.DIST** from the **Select a function** box
Click **OK**
Step 5. When the **Function Arguments** dialog box appears:
Enter 4 in the **Number_s** box (the value of x)
Enter 10 in the **Trials** box (the value of n)
Enter 0.30 in the **Probability_s** box (the value of p)
Enter false in the **Cumulative** box[2]
Note: At this point the desired binomial probability of 0.2001 is automatically computed and appears near the bottom of the dialog box.
Click **OK** and the binomial probability will appear in the worksheet cell requested in Step 1.

A user who wants other binomial probabilities may obtain the information without repeating the steps for each probability desired. Perhaps the easiest alternative is to stay in step 5. After the four entries have been made and the first probability appears, simply return to the **Number_s** box and insert a new value of x. The new probability will appear. Repeated changes can be made in the dialog box, including changes to the trials, probability, and/or cumulative boxes. For each change, the desired probability will appear. When **OK** is selected, only the last binomial probability will be placed in the worksheet.

The Excel functions BINOM.DIST and POISSON.DIST are only recognized by Excel 2010. Earlier versions of Excel will use the function names BINOMDIST and POISSON, respectively, to compute the same values using the same steps.

If the user wants to insert multiple binomial probabilities into the worksheet, the desired values of x are entered into the worksheet first. Then, in step 5, the user enters the cell location of one of the values of x in the numbers box. After completing the steps for one binomial probability, individuals experienced with Excel can use Excel's Copy command to copy the binomial function into the cells where the other binomial probabilities are to appear.

The Excel procedure for generating Poisson probabilities is similar to the procedure just described. Step 4 can be used to select the **POISSON.DIST** function name. The dialog box in step 5 will guide the user through the input values required to compute the desired probabilities.

Appendix 3.2 Computing Probabilities for Continuous Distributions with Excel

Excel has the capability of computing probabilities for several continuous probability distributions, including the normal and exponential probability distributions. In this appendix, we describe how Excel can be used to compute probabilities for any normal probability

[2]Placing false in the cumulative box provides the probability of exactly four successes. Placing true in this box provides the cumulative probability of four or fewer successes.

distribution. The procedures for the exponential and other continuous probability distributions are similar to the one we describe for the normal probability distribution.

Let us return to the Grear Tire Company problem, where the tire mileage was described by a normal probability distribution with $\mu = 36{,}500$ and $\sigma = 5000$. Assume we are interested in the probability that tire mileage will exceed 40,000 miles. The following steps describe how to use Excel to compute the desired normal probability:

Step 1. Select a cell in the worksheet where you want the normal probability to appear
Step 2. Select the **Formulas** tab (see Appendix A)
Step 3. Choose the **Insert Function** option
Step 4. When the **Insert Function** dialog box appears:
Choose **Statistical** from the **Or select a category** box
Choose **NORM.DIST** from the **Select a function** box
Click **OK**
Step 5. When the **Function Arguments** dialog box appears:
Enter 40000 in the **X** box
Enter 36500 in the **Mean** box
Enter 5000 in the **Standard_dev** box
Enter true in the **Cumulative** box
Click **OK**

At this point, 0.7580 will appear in the cell selected in step 1, indicating that the cumulative probability that the tire mileage is less than or equal to 40,000 miles is 0.7580. Therefore, the probability that tire mileage will exceed 40,000 miles is $1 - 0.7580 = 0.2420$.

Excel uses an inverse computation to convert a given cumulative normal probability into a value for the random variable. For example, what mileage guarantee should Grear offer if the company wants no more than 10% of the tires to be eligible for the guarantee? To compute the mileage guarantee by using Excel, follow the procedure just described. However, two changes are necessary: In step 4, choose **NORM.INV** from the **Select a function** box; in step 5, enter the cumulative probability of 0.10 in the **Probability** box and then enter the mean and the standard deviation. When **OK** is selected in step 5, the tire mileage guarantee of 30,092, or approximately 30,100, miles appears in the worksheet.

The Excel functions NORM.DIST, NORM.INV and EXPON.DIST are only recognized by Excel 2010. Earlier versions of Excel will use the function names NORMDIST, NORMINV and EXPONDIST, respectively, to compute the same values using the same steps.

The Excel procedure for generating exponential probabilities is similar to the procedure just described. Step 4 can be used to choose the **EXPON.DIST** function name. The dialog box in step 5 will guide the user through the input values required to compute the desired probability. Note that the value entered in the **Lambda** box is $1/\mu$. When **OK** is selected in step 5, the cumulative exponential probability appears in the worksheet.