7

Continuous Probability Distributions

GOALS

When you have completed this chapter you will be able to:

1 Understand the difference between discrete and continuous distributions.

2 Compute the mean and the standard deviation for a *uniform distribution*.

3 Compute probabilities by using the uniform distribution.

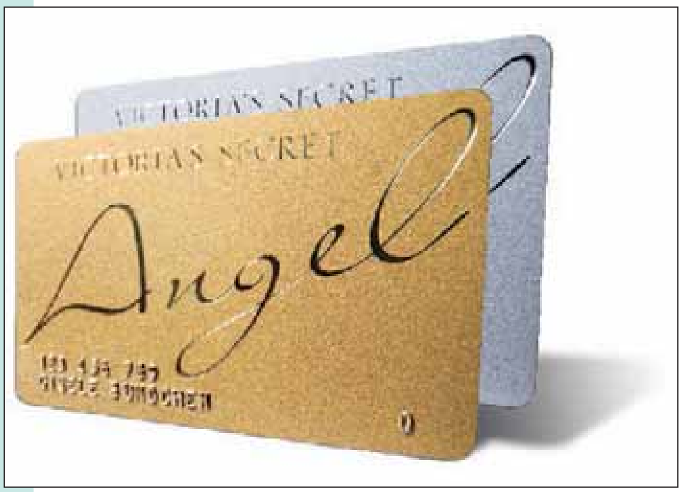
4 List the characteristics of the *normal probability distribution*.

5 Define and calculate *z* values.

6 Determine the probability an observation is between two points on a normal probability distribution.

7 Determine the probability an observation is above (or below) a point on a normal probability distribution.

8 Use the normal probability distribution to approximate the binomial distribution.



Most retail stores offer their own credit cards. At the time of the credit application the customer is given a 10% discount on his or her purchase. The time it takes to fill out the credit application form follows a uniform distribution with the times ranging from 4 to 10 minutes. What is the standard deviation for the process time? (See Goal 2 and Exercise 39.)

Introduction

Chapter 6 began our study of probability distributions. We consider three *discrete* probability distributions: binomial, hypergeometric, and Poisson. These distributions are based on discrete random variables, which can assume only clearly separated values. For example, we select for study 10 small businesses that began operations during the year 2000. The number still operating in 2006 can be 0, 1, 2, . . . , 10. There cannot be 3.7, 12, or −7 still operating in 2006. In this example, only certain outcomes are possible and these outcomes are represented by clearly separated values. In addition, the result is usually found by counting the number of successes. We count the number of the businesses in the study that are still in operation in 2006.

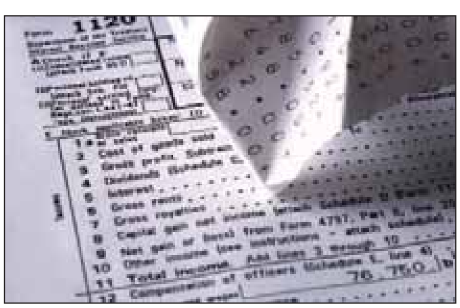
In this chapter, we continue our study of probability distributions by examining *continuous* probability distributions. A continuous probability distribution usually results from measuring something, such as the distance from the dormitory to the classroom, the weight of an individual, or the amount of bonus earned by CEOs. Suppose we select five students and find the distance, in miles, they travel to attend class as 12.2, 8.9, 6.7, 3.6, and 14.6. When examining a continuous distribution we are usually interested in information such as the percent of students who travel less than 10 miles or the percent who travel more than 8 miles. In other words, for a continuous distribution we may wish to know the percent of observations that occur within a certain range. It is important to realize that a continuous random variable has an infinite number of values within a particular range. So you think of the probability a variable will have a value within a specified range, rather than the probability for a specific value.

We consider two families of continuous probability distributions, the **uniform probability distribution** and the **normal probability distribution.** These distributions describe the likelihood that a continuous random variable that has an infinite number of possible values will fall within a specified range. For example, suppose the time to access the McGraw-Hill web page (www.mhhe.com) is uniformly distributed with a minimum time of 20 milliseconds and a maximum time of 60 milliseconds. Then we can determine the probability the page can be accessed in 30 milliseconds or less. The access time is measured on a continuous scale.

The second continuous distribution discussed in this chapter is the normal probability distribution. The normal distribution is described by its mean and standard deviation. For example, assume the life of an Energizer C-size battery follows a normal distribution with a mean of 45 hours and a standard deviation of 10 hours when used in a particular toy. We can determine the likelihood the battery will last more than 50 hours, between 35 and 62 hours, or less than 39 hours. The life of the battery is measured on a continuous scale.

The Family of Uniform Probability Distributions

The uniform probability distribution is perhaps the simplest distribution for a continuous random variable. This distribution is rectangular in shape and is defined by minimum and maximum values. Here are some examples that follow a uniform distribution.



- The time to fly via a commercial airliner from Orlando, Florida, to Atlanta, Georgia, ranges from 60 minutes to 120 minutes. The random variable is the flight time within this interval. Note the variable of interest, flight time in minutes, is continuous in the interval from 60 minutes to 120 minutes.
- Volunteers at the Grand Strand Public Library prepare federal income tax forms. The time to prepare form 1040-EZ follows a uniform distribution over the interval between 10 minutes and 30 minutes. The random variable is the number of minutes to complete the form, and it can assume any value between 10 and 30.

A uniform distribution is shown in Chart 7–1. The distribution's shape is rectangular and has a minimum value of *a* and a maximum of *b*. Also notice in Chart 7–1 the height of the distribution is constant or uniform for all values between *a* and *b*.

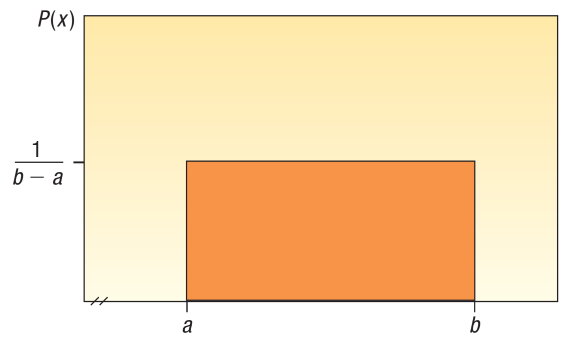


CHART 7–1 A Continuous Uniform Distribution

The mean of a uniform distribution is located in the middle of the interval between the minimum and maximum values. It is computed as:

MEAN OF THE UNIFORM DISTRIBUTION $$\mu = \frac{a + b}{2}$$ **[7–1]**

The standard deviation describes the dispersion of a distribution. In the uniform distribution, the standard deviation is also related to the interval between the maximum and minimum values.

STANDARD DEVIATION OF THE UNIFORM DISTRIBUTION $$\sigma = \sqrt{\frac{(b - a)^2}{12}}$$ **[7–2]**

The equation for the uniform probability distribution is:

UNIFORM DISTRIBUTION $$P(x) = \frac{1}{b - a} \quad \text{if } a \le x \le b \text{ and 0 elsewhere}$$ **[7–3]**

As shown in Chapter 6, probability distributions are useful for making probability statements concerning the values of a random variable. For distributions describing a continuous random variable, areas within the distribution represent probabilities. In the uniform distribution, its rectangular shape allows us to apply the area formula for a rectangle. Recall we find the area of a rectangle by multiplying its length by its height. For the uniform distribution the height of the rectangle is $P(x)$, which is $1/(b - a)$. The length or base of the distribution is $b - a$. Notice that if we multiply the height of the distribution by its entire range to find the area, the result is always 1.00. To put it another way, the total area within a continuous probability distribution is equal to 1.00. In general

$$\text{Area} = (\text{height})(\text{base}) = \frac{1}{(b - a)}(b - a) = 1.00$$

So if a uniform distribution ranges from 10 to 15, the height is 0.20, found by $1/(15 - 10)$. The base is 5, found by $15 - 10$. The total area is:

$$\text{Area} = (\text{height})(\text{base}) = \frac{1}{(15 - 10)}(15 - 10) = 1.00$$

An example will illustrate the features of a uniform distribution and how we calculate probabilities using it.

Example

Southwest Arizona State University provides bus service to students while they are on campus. A bus arrives at the North Main Street and College Drive stop every 30 minutes between 6 A.M. and 11 P.M. during weekdays. Students arrive at the bus stop at random times. The time that a student waits is uniformly distributed from 0 to 30 minutes.

1. Draw a graph of this distribution.
2. Show that the area of this uniform distribution is 1.00.
3. How long will a student "typically" have to wait for a bus? In other words what is the mean waiting time? What is the standard deviation of the waiting times?
4. What is the probability a student will wait more than 25 minutes?
5. What is the probability a student will wait between 10 and 20 minutes?

Solution

In this case the random variable is the length of time a student must wait. Time is measured on a continuous scale, and the wait times may range from 0 minutes up to 30 minutes.

1. The graph of the uniform distribution is shown in Chart 7–2. The horizontal line is drawn at a height of .0333, found by $1/(30 - 0)$. The range of this distribution is 30 minutes.

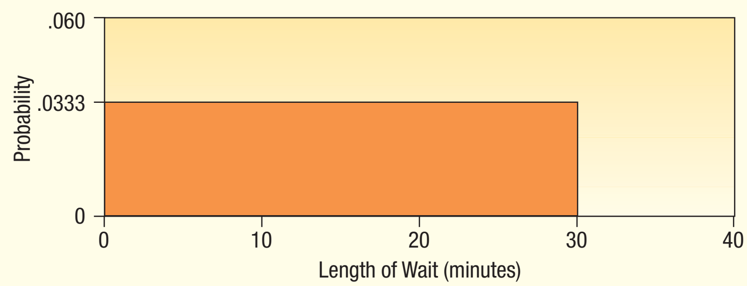


CHART 7–2 Uniform Probability Distribution of Student Waiting Times

2. The times students must wait for the bus is uniform over the interval from 0 minutes to 30 minutes, so in this case a is 0 and b is 30.

$$\text{Area} = (\text{height})(\text{base}) = \frac{1}{(30 - 0)}(30 - 0) = 1.00$$

3. To find the mean, we use formula (7–1).

$$\mu = \frac{a + b}{2} = \frac{0 + 30}{2} = 15$$

The mean of the distribution is 15 minutes, so the typical wait time for bus service is 15 minutes.

To find the standard deviation of the wait times, we use formula (7–2).

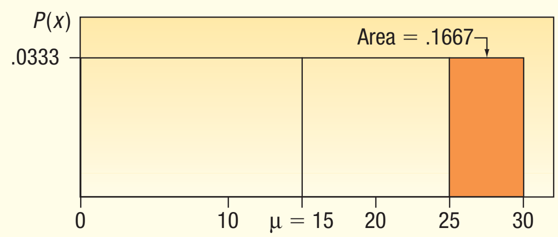
$$\sigma = \sqrt{\frac{(b - a)^2}{12}} = \sqrt{\frac{(30 - 0)^2}{12}} = 8.66$$

The standard deviation of the distribution is 8.66 minutes. This measures the variation in the student wait times.

4. The area within the distribution for the interval 25 to 30 represents this particular probability. From the area formula:

$$P(25 < \text{wait time} < 30) = (\text{height})(\text{base}) = \frac{1}{(30 - 0)}(5) = .1667$$

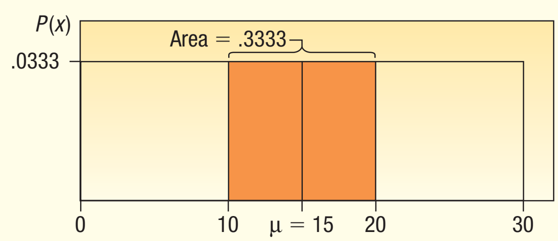
So the probability a student waits between 25 and 30 minutes is .1667. This conclusion is illustrated by the following graph.



5. The area within the distribution for the interval 10 to 20 represents the probability.

$$P(10 < \text{wait time} < 20) = (\text{height})(\text{base}) = \frac{1}{(30 - 0)}(10) = .3333$$

We can illustrate this probability as follows.



Self-Review 7–1

Australian sheepdogs have a relatively short life. The length of their life follows a uniform distribution between 8 and 14 years.
(a) Draw this uniform distribution. What are the height and base values?
(b) Show the total area under the curve is 1.00.
(c) Calculate the mean and the standard deviation of this distribution.
(d) What is the probability a particular dog lives between 10 and 14 years?
(e) What is the probability a dog will live less than 9 years?

Exercises

1. A uniform distribution is defined over the interval from 6 to 10.
  a. What are the values for *a* and *b*?
  b. What is the mean of this uniform distribution?
  c. What is the standard deviation?
  d. Show that the total area is 1.00.
  e. Find the probability of a value more than 7.
  f. Find the probability of a value between 7 and 9.

2. A uniform distribution is defined over the interval from 2 to 5.
  a. What are the values for *a* and *b*?
  b. What is the mean of this uniform distribution?
  c. What is the standard deviation?
  d. Show that the total area is 1.00.
  e. Find the probability of a value more than 2.6.
  f. Find the probability of a value between 2.9 and 3.7.

3. America West Airlines reports the flight time from Los Angeles International Airport to Las Vegas is 1 hour and 5 minutes, or 65 minutes. Suppose the actual flying time is uniformly distributed between 60 and 70 minutes.
  a. Show a graph of the continuous probability distribution.
  b. What is the mean flight time? What is the variance of the flight times?

c. What is the probability the flight time is less than 68 minutes?
d. What is the probability the flight takes more than 64 minutes?

4. According to the Insurance Institute of America, a family of four spends between $400 and $3,800 per year on all types of insurance. Suppose the money spent is uniformly distributed between these amounts.
   a. What is the mean amount spent on insurance?
   b. What is the standard deviation of the amount spent?
   c. If we select a family at random, what is the probability they spend less than $2,000 per year on insurance per year?
   d. What is the probability a family spends more than $3,000 per year?
5. The April rainfall in Flagstaff, Arizona, follows a uniform distribution between 0.5 and 3.00 inches.
   a. What are the values for a and b?
   b. What is the mean amount of rainfall for the month? What is the standard deviation?
   c. What is the probability of less than an inch of rain for the month?
   d. What is the probability of *exactly* 1.00 inch of rain?
   e. What is the probability of more than 1.50 inches of rain for the month?
6. Customers experiencing technical difficulty with their internet cable hookup may call an 800 number for technical support. It takes the technician between 30 seconds to 10 minutes to resolve the problem. The distribution of this support time follows the uniform distribution.
   a. What are the values for a and b in minutes?
   b. What is the mean time to resolve the problem? What is the standard deviation of the time?
   c. What percent of the problems take more than 5 minutes to resolve?
   d. Suppose we wish to find the middle 50 percent of the problem-solving times. What are the end points of these two times?

The Family of Normal Probability Distributions

Next we consider the normal probability distribution. Unlike the uniform distribution [see formula (7–3)] the normal probability distribution has a very complex formula.

NORMAL PROBABILITY DISTRIBUTION

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(X-\mu)^2}{2\sigma^2}\right]} \qquad [7\text{–}4]$$

However, do not be bothered by how complex this formula looks. You are already familiar with many of the values. The symbols μ and σ refer to the mean and the standard deviation, as usual. The Greek symbol π is a natural mathematical constant and its value is approximately 22/7 or 3.1416. The letter e is also a mathematical constant. It is the base of the natural log system and is equal to 2.718. X is the value of a continuous random variable. So a normal distribution is based on—that is, it is defined by—its mean and standard deviation.

You will not need to make any calculations from formula (7–4). Instead you will be using a table, which is given as Appendix B.1, to look up the various probabilities.

The normal probability distribution has the following major characteristics.

1. It is **bell-shaped** and has a single peak at the center of the distribution. The arithmetic mean, median, and mode are equal and located in the center of the distribution. The total area under the curve is 1.00. Half the area under the normal curve is to the right of this center point and the other half to the left of it.
2. It is **symmetrical** about the mean. If we cut the normal curve vertically at the center value, the two halves will be mirror images.
3. It falls off smoothly in either direction from the central value. That is, the distribution is **asymptotic:** The curve gets closer and closer to the X-axis but never actually touches it. To put it another way, the tails of the curve extend indefinitely in both directions.
4. The location of a normal distribution is determined by the mean, μ. The dispersion or spread of the distribution is determined by the standard deviation, σ.

These characteristics are shown graphically in Chart 7–3.

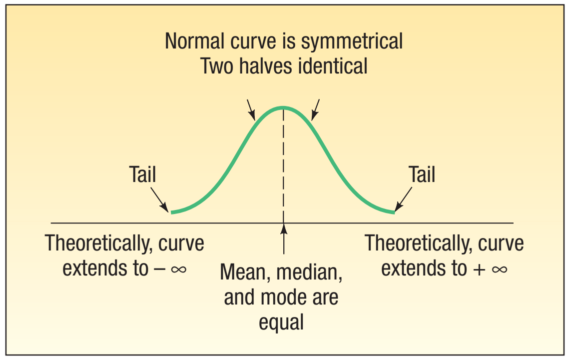


CHART 7–3 Characteristics of a Normal Distribution

There is not just one normal probability distribution, but rather a "family" of them. For example, in Chart 7–4 the probability distributions of length of employee service in three different plants can be compared. In the Camden plant, the mean is 20 years and the standard deviation is 3.1 years. There is another normal probability distribution for the length of service in the Dunkirk plant, where $\mu = 20$ years and $\sigma = 3.9$ years. In the Elmira plant, $\mu = 20$ years and $\sigma = 5.0$ years. Note that the means are the same but the standard deviations are different.

Equal means, unequal standard deviations

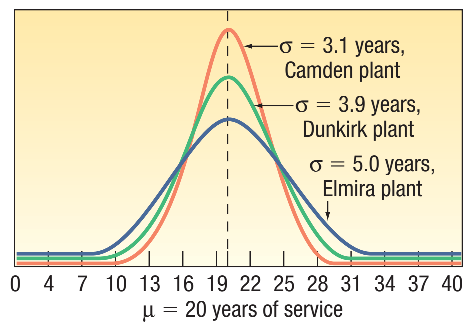


CHART 7–4 Normal Probability Distributions with Equal Means but Different Standard Deviations

Chart 7–5 shows the distribution of box weights of three different cereals. The weights follow a normal distribution with different means but identical standard deviations.

Unequal means, equal standard deviations

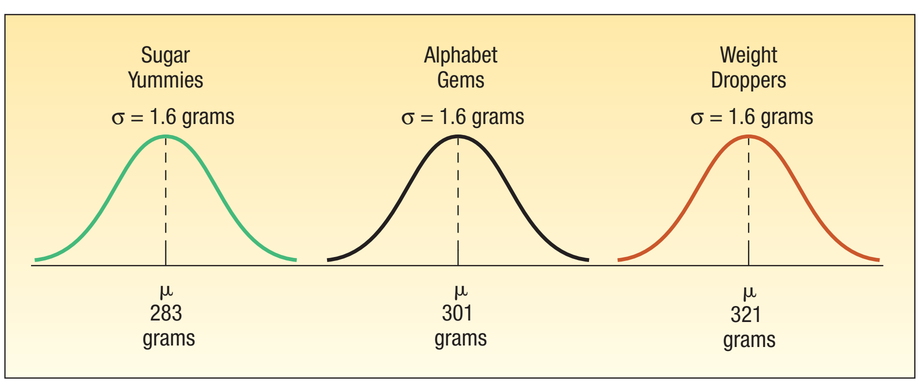


CHART 7–5 Normal Probability Distributions Having Different Means but Equal Standard Deviations

Finally, Chart 7–6 shows three normal distributions having different means and standard deviations. They show the distribution of tensile strengths, measured in pounds per square inch (psi), for three types of cables.

Unequal means, unequal standard deviations

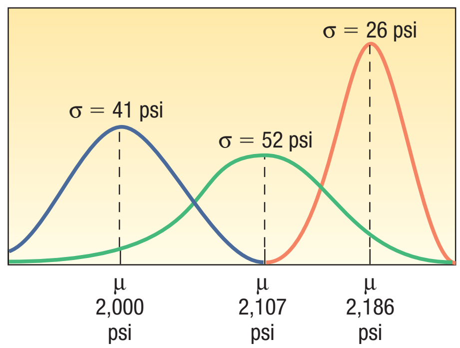


CHART 7–6 Normal Probability Distributions with Different Means and Standard Deviations

In Chapter 6, recall that discrete probability distributions show the specific likelihood a discrete value will occur. For example, on page 190, the binomial distribution is used to calculate the probability that none of the five flights arriving at the Bradford Pennsylvania Regional Airport would be late.

With a continuous probability distribution, areas below the curve define probabilities. The total area under the normal curve is 1.0. This accounts for all possible outcomes. Since a normal probability distribution is symmetric, the area under the curve to the left of the mean is 0.5, and the area under the curve to the right of the mean is 0.5. Apply this to the distribution of Sugar Yummies in Chart 7–5. It is normally distributed with a mean of 283 grams. Therefore, the probability of filling a box with more than 283 grams is 0.5 and the probability of filling a box with less than 283 grams is 0.5. We can also determine the probability that a box weighs between 280 and 286 grams. However, to determine this probability we need to know about the standard normal probability distribution.

The Standard Normal Probability Distribution

The number of normal distributions is unlimited, each having a different mean (μ), standard deviation (σ), or both. While it is possible to provide probability tables for discrete distributions such as the binomial and the Poisson, providing tables for the infinite number of normal distributions is impossible. Fortunately, one member of the family can be used to determine the probabilities for all normal probability distributions. It is called the **standard normal probability distribution,** and it is unique because it has a mean of 0 and a standard deviation of 1.

Any *normal probability distribution* can be converted into a *standard normal probability distribution* by subtracting the mean from each observation and dividing this difference by the standard deviation. The results are called ***z* values** or ***z* scores.**

> ***z* VALUE** The signed distance between a selected value, designated *X*, and the mean, μ, divided by the standard deviation, σ.

So, a *z* value is the distance from the mean, measured in units of the standard deviation.

In terms of a formula:

STANDARD NORMAL VALUE

$$z = \frac{X - \mu}{\sigma} \qquad [7\text{–}5]$$

where:

X is the value of any particular observation or measurement.
μ is the mean of the distribution.
σ is the standard deviation of the distribution.

As noted in the above definition, a z value expresses the distance or difference between a particular value of X and the arithmetic mean in units of the standard deviation. Once the normally distributed observations are standardized, the z values are normally distributed with a mean of 0 and a standard deviation of 1. So the z distribution has all the characteristics of any normal probability distribution. These characteristics are listed on page 227. The table in Appendix B.1 (also on the inside back cover) lists the probabilities for the standard normal probability distribution.

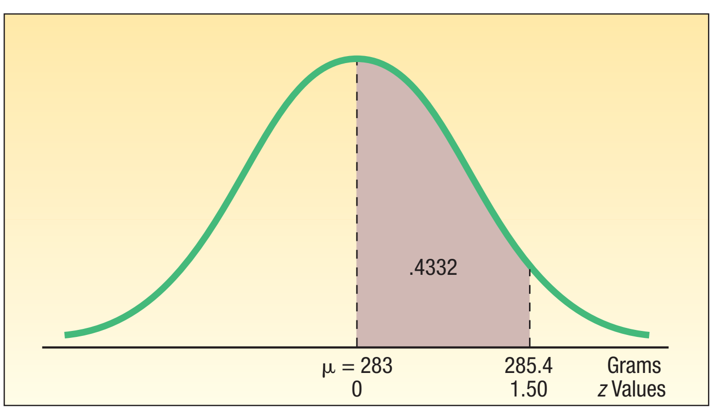
Statistics in Action

An individual's skills depend on a combination of many hereditary and environmental factors, each having about the same amount of weight or influence on the skills. Thus, much like a binomial distribution with a large number of trials, many skills and attributes follow the normal distribution. For example, scores on the Scholastic Aptitude Test (SAT) are normally distributed with a mean of 1000 and a standard deviation of 140.

TABLE 7–1 Areas under the Normal Curve

z	0.00	0.01	0.02	0.03	0.04	0.05	. . .
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	
.							
.							
.							

To explain, suppose we wish to compute the probability that boxes of Sugar Yummies weigh between 283 and 285.4 grams. From Chart 7–5, we know that the box weight of Sugar Yummies follows the normal distribution with a mean of 283 grams and a standard deviation of 1.6 grams. We want to know the probability or area under the curve between the mean, 283 grams, and 285.4 grams. We can also express this problem using probability notation, similar to the style used in the previous chapter: $P(283 < \text{weight} < 285.4)$. To find the probability, it is necessary to convert both 283 grams and 285.4 grams to z values using formula (7–5). The z value corresponding to 283 is 0, found by $(283 - 283)/1.6$. The z value corresponding to 285.4 is 1.50 found by $(285.4 - 283)/1.6$. Next we go to the table in Appendix B.1. A portion of the table is repeated as Table 7–1. Go down the column of the table headed by the letter z to 1.5. Then move horizontally to the right and read the probability under the column headed 0.00. It is 0.4332. This means the area under the curve between 0.00 and 1.50 is 0.4332. This is the probability that a randomly selected box of Sugar Yummies will weigh between 283 and 285.4 grams. This is illustrated in the following graph.



Applications of the Standard Normal Distribution

What is the area under the curve between the mean and X for the following z values? Check your answers against those given. Not all the values are available in Table 7–5. You will need to use Appendix B.1 or the table located on the inside back cover of the text.

Computed z Value	Area under Curve
2.84	.4977
1.00	.3413
0.49	.1879

Now we will compute the z value given the population mean, μ, the population standard deviation, σ, and a selected X.

Example

The weekly incomes of shift foremen in the glass industry follow the normal probability distribution with a mean of $1,000 and a standard deviation of $100. What is the z value for the income, let's call it X, of a foreman who earns $1,100 per week? For a foreman who earns $900 per week?

Solution

Using formula (7–5), the z values for the two X values ($1,100 and $900) are:

For X = $1,100:

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}100 - \$1{,}000}{\$100} = 1.00$$

For X = $900:

$$z = \frac{X - \mu}{\sigma} = \frac{\$900 - \$1{,}000}{\$100} = -1.00$$

The z of 1.00 indicates that a weekly income of $1,100 is one standard deviation above the mean, and a z of -1.00 shows that a $900 income is one standard deviation below the mean. Note that both incomes ($1,100 and $900) are the same distance ($100) from the mean.

Self-Review 7–2

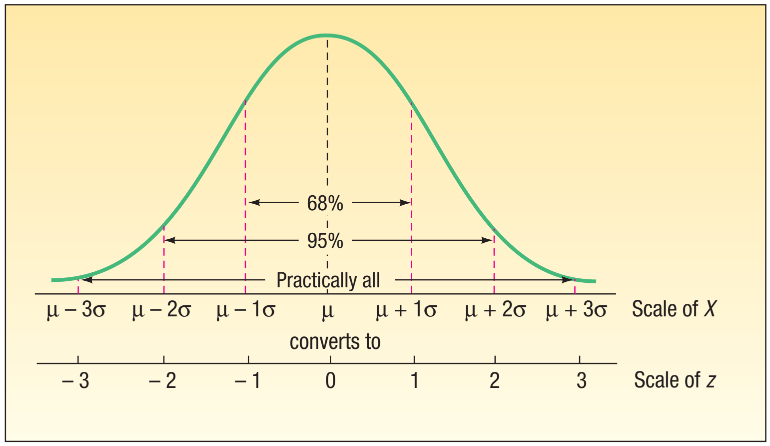
Using the information in the preceding example (μ = $1,000, σ = $100), convert:
(a) The weekly income of $1,225 to a z value.
(b) The weekly income of $775 to a z value.

The Empirical Rule

Before examining further applications of the standard normal probability distribution, we will consider three areas under the normal curve that will be used extensively in the following chapters. These facts were called the Empirical Rule in Chapter 3, see page 82.

1. About 68 percent of the area under the normal curve is within one standard deviation of the mean. This can be written as $\mu \pm 1\sigma$.
2. About 95 percent of the area under the normal curve is within two standard deviations of the mean, written $\mu \pm 2\sigma$.
3. Practically all of the area under the normal curve is within three standard deviations of the mean, written $\mu \pm 3\sigma$.

This information is summarized in the following graph.



Transforming measurements to standard normal deviates changes the scale. The conversions are also shown in the graph. For example, $\mu + 1\sigma$ is converted to a z value of 1.00. Likewise, $\mu - 2\sigma$ is transformed to a z value of -2.00. Note that the center of the z distribution is zero, indicating no deviation from the mean, μ.

Example

As part of its quality assurance program, the Autolite Battery Company conducts tests on battery life. For a particular D-cell alkaline battery, the mean life is 19 hours. The useful life of the battery follows a normal distribution with a standard deviation of 1.2 hours. Answer the following questions.

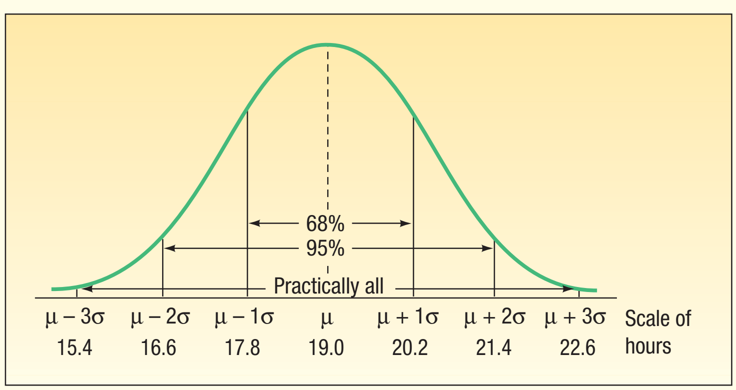
1. About 68 percent of the batteries failed between what two values?
2. About 95 percent of the batteries failed between what two values?
3. Virtually all of the batteries failed between what two values?

Solution

We can use the results of the Empirical Rule to answer these questions.

1. About 68 percent of the batteries will fail between 17.8 and 20.2 hours, found by $19.0 \pm 1(1.2)$ hours.
2. About 95 percent of the batteries will fail between 16.6 and 21.4 hours, found by $19.0 \pm 2(1.2)$ hours.
3. Virtually all failed between 15.4 and 22.6 hours, found by $19.0 \pm 3(1.2)$ hours.

This information is summarized on the following chart.



Self-Review 7–3

The distribution of the annual incomes of a group of middle-management employees at Compton Plastics approximates a normal distribution with a mean of \$47,200 and a standard deviation of \$800.

(a) About 68 percent of the incomes lie between what two amounts?
(b) About 95 percent of the incomes lie between what two amounts?
(c) Virtually all of the incomes lie between what two amounts?
(d) What are the median and the modal incomes?
(e) Is the distribution of incomes symmetrical?

Exercises

7. Explain what is meant by this statement: "There is not just one normal probability distribution but a 'family' of them."

8. List the major characteristics of a normal probability distribution.

9. The mean of a normal probability distribution is 500; the standard deviation is 10.
   a. About 68 percent of the observations lie between what two values?
   b. About 95 percent of the observations lie between what two values?
   c. Practically all of the observations lie between what two values?

10. The mean of a normal probability distribution is 60; the standard deviation is 5.
   a. About what percent of the observations lie between 55 and 65?
   b. About what percent of the observations lie between 50 and 70?
   c. About what percent of the observations lie between 45 and 75?

11. The Kamp family has twins, Rob and Rachel. Both Rob and Rachel graduated from college 2 years ago, and each is now earning \$50,000 per year. Rachel works in the retail industry, where the mean salary for executives with less than 5 years' experience is \$35,000 with a standard deviation of \$8,000. Rob is an engineer. The mean salary for engineers with less than 5 years' experience is \$60,000 with a standard deviation of \$5,000. Compute the z values for both Rob and Rachel and comment on your findings.

12. A recent article in the *Cincinnati Enquirer* reported that the mean labor cost to repair a heat pump is \$90 with a standard deviation of \$22. Monte's Plumbing and Heating Service completed repairs on two heat pumps this morning. The labor cost for the first was \$75 and it was \$100 for the second. Assume the distribution of labor costs follows the normal probability distribution. Compute z values for each and comment on your findings.

Finding Areas under the Normal Curve

The next application of the standard normal distribution involves finding the area in a normal distribution between the mean and a selected value, which we identify as X. The following example will illustrate the details.

Example

Recall in an earlier example (see page 231) we reported that the mean weekly income of a shift foreman in the glass industry is normally distributed with a mean of \$1,000 and a standard deviation of \$100. That is, $\mu = \$1,000$ and $\sigma = \$100$. What is the likelihood of selecting a foreman whose weekly income is between \$1,000 and \$1,100? We write this question in probability notation as: $P(\$1,000 < \text{weekly income} < \$1,100)$.

Solution

We have already converted \$1,100 to a z value of 1.00 using formula (7–5). To repeat:

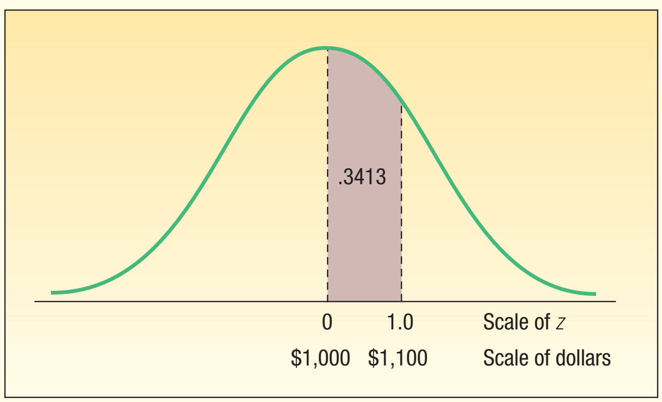
$$z = \frac{X - \mu}{\sigma} = \frac{\$1,100 - \$1,000}{\$100} = 1.00$$

The probability associated with a *z* of 1.00 is available in Appendix B.1. A portion of Appendix B.1 follows. To locate the probability, go down the left column to 1.0, and then move horizontally to the column headed .00. The value is .3413.

z	0.00	0.01	0.02
.	.	.	.
.	.	.	.
.	.	.	.
0.7	.2580	.2611	.2642
0.8	.2881	.2910	.2939
0.9	.3159	.3186	.3212
1.0	.3413	.3438	.3461
1.1	.3643	.3665	.3686
.	.	.	.
.	.	.	.
.	.	.	.

The area under the normal curve between $1,000 and $1,100 is .3413. We could also say 34.13 percent of the shift foremen in the glass industry earn between $1,000 and $1,100 weekly, or the likelihood of selecting a foreman and finding his or her income is between $1,000 and $1,100 is .3413.

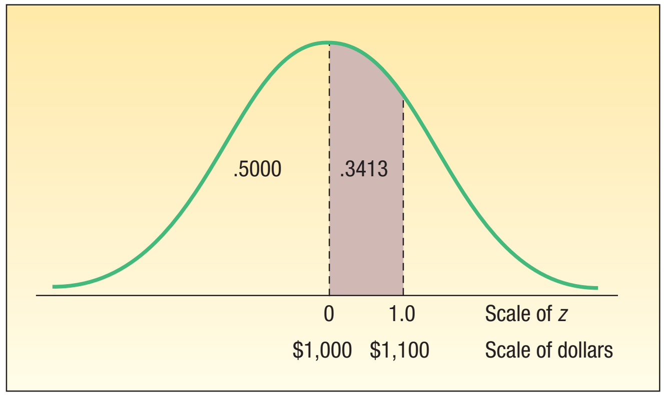
This information is summarized in the following diagram.



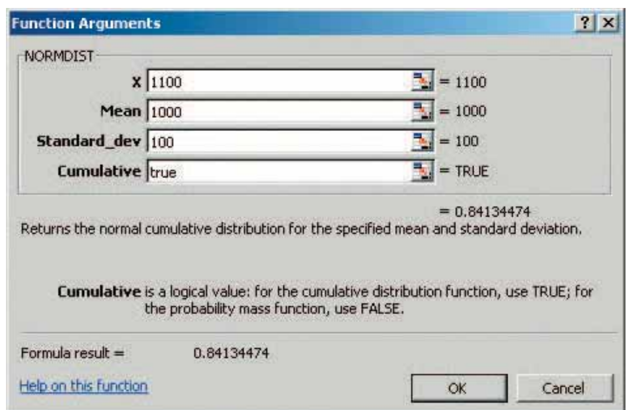
In the example just completed, we are interested in the probability between the mean and a given value. Let's change the question. Instead of wanting to know the probability of selecting at random a foreman who earned between $1,000 and $1,100, suppose we wanted the probability of selecting a foreman who earned less than $1,100. In probability notation we write this statement as *P*(weekly income $<$ $1,100). The method of solution is the same. We find the probability of selecting a foreman who earns between $1,000, the mean, and $1,100. This probability is .3413. Next, recall that half the area, or probability, is above the mean and half is below. So the probability of selecting a foreman earning less than $1,000 is .5000. Finally, we add the two probabilities, so .3413 + .5000 = .8413. About 84 percent of the foremen in the glass industry earn less than $1,100 per month. See the following diagram.

Statistics in Action

Many processes, such as filling soda bottles and canning fruit, are normally distributed. Manufacturers must guard against both over- and underfilling. If they put too much in the can or bottle, they are giving away their product. If they put too little in, the customer may feel cheated and the government may question the label description. "Control charts," with limits drawn three standard deviations above and below the mean, are routinely used to monitor this type of production process.



Excel will calculate this probability. The necessary commands are in the **Software Commands** section at the end of the chapter. The answer is .8413, the same as we calculated.



Example

Refer to the information regarding the weekly income of shift foremen in the glass industry. The distribution of weekly incomes follows the normal probability distribution, with a mean of $1,000 and a standard deviation of $100. What is the probability of selecting a shift foreman in the glass industry whose income is:

1. Between $790 and $1,000?
2. Less than $790?

Solution

We begin by finding the *z* value corresponding to a weekly income of $790. From formula (7–5):

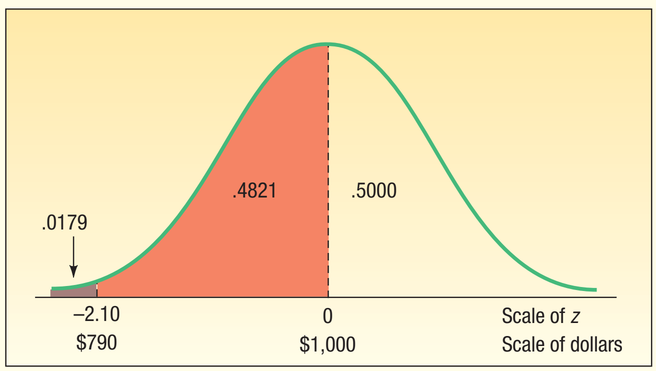
$$z = \frac{X - \mu}{s} = \frac{\$790 - \$1{,}000}{\$100} = -2.10$$

See Appendix B.1. Move down the left margin to the row 2.1 and across that row to the column headed 0.00. The value is .4821. So the area under the standard normal curve corresponding to a *z* value of 2.10 is .4821. However, because the normal distribution is symmetric, the area between 0 and a negative *z* value is the same as that between 0 and the corresponding positive *z* value. The likelihood of finding a foreman earning between $790 and $1,000 is .4821. In probability notation we write $P(\$790 < \text{weekly income} < \$1000) = .4821$.

z	0.00	0.01	0.02
.	.	.	.
.	.	.	.
.	.	.	.
2.0	.4772	.4778	.4783
2.1	.4821	.4826	.4830
2.2	.4861	.4864	.4868
2.3	.4893	.4896	.4898
.	.	.	.
.	.	.	.
.	.	.	.

The mean divides the normal curve into two identical halves. The area under the half to the left of the mean is .5000, and the area to the right is also .5000. Because the area under the curve between \$790 and \$1,000 is .4821, the area below \$790 is .0179, found by .5000 − .4821. In probability notation we write P(weekly income < \$790) = .0179.

This means that 48.21 percent of the foremen have weekly incomes between \$790 and \$1,000. Further, we can anticipate that 1.79 percent earn less than \$790 per week. This information is summarized in the following diagram.



Self-Review 7–4

The employees of Cartwright Manufacturing are awarded efficiency ratings based on such factors as monthly output, attitude, and attendance. The distribution of the ratings follows the normal probability distribution. The mean is 400, the standard deviation 50.

(a) What is the area under the normal curve between 400 and 482? Write this area in probability notation.
(b) What is the area under the normal curve for ratings greater than 482? Write this area in probability notation.
(c) Show the facets of this problem in a chart.

Exercises

13. A normal population has a mean of 20.0 and a standard deviation of 4.0.
  a. Compute the z value associated with 25.0.
  b. What proportion of the population is between 20.0 and 25.0?
  c. What proportion of the population is less than 18.0?

14. A normal population has a mean of 12.2 and a standard deviation of 2.5.
  a. Compute the z value associated with 14.3.
  b. What proportion of the population is between 12.2 and 14.3?
  c. What proportion of the population is less than 10.0?

15. A recent study of the hourly wages of maintenance crew members for major airlines showed that the mean hourly salary was \$20.50, with a standard deviation of \$3.50. Assume the distribution of hourly wages follows the normal probability distribution. If we select a crew member at random, what is the probability the crew member earns:
- **a.** Between \$20.50 and \$24.00 per hour?
- **b.** More than \$24.00 per hour?
- **c.** Less than \$19.00 per hour?

16. The mean of a normal probability distribution is 400 pounds. The standard deviation is 10 pounds.
- **a.** What is the area between 415 pounds and the mean of 400 pounds?
- **b.** What is the area between the mean and 395 pounds?
- **c.** What is the probability of selecting a value at random and discovering that it has a value of less than 395 pounds?

Another application of the normal distribution involves combining two areas, or probabilities. One of the areas is to the right of the mean and the other to the left.

Example

Recall the distribution of weekly incomes of shift foremen in the glass industry. The weekly incomes follow the normal probability distribution, with a mean of \$1,000 and a standard deviation of \$100. What is the area under this normal curve between \$840 and \$1,200?

Solution

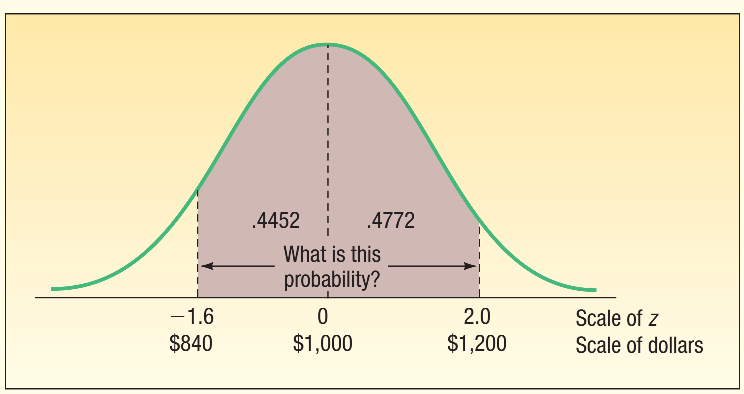
The problem can be divided into two parts. For the area between \$840 and the mean of \$1,000:

$$z = \frac{\$840 - \$1{,}000}{\$100} = \frac{-\$160}{\$100} = -1.60$$

For the area between the mean of \$1,000 and \$1,200:

$$z = \frac{\$1{,}200 - \$1{,}000}{\$100} = \frac{\$200}{\$100} = 2.00$$

The area under the curve for a z of -1.60 is .4452 (from Appendix B.1). The area under the curve for a z of 2.00 is .4772. Adding the two areas: $.4452 + .4772 = .9224$. Thus, the probability of selecting an income between \$840 and \$1,200 is .9224. In probability notation we write $P(\$840 < \text{weekly income} < \$1{,}200) = .4452 + .4772 = .9224$. To summarize, 92.24 percent of the foremen have weekly incomes between \$840 and \$1,200. This is shown in a diagram:



Another application of the normal distribution involves determining the area between values on the *same* side of the mean.

Example

Returning to the weekly income distribution of shift foremen in the glass industry ($\mu =$ \$1,000, $\sigma =$ \$100), what is the area under the normal curve between \$1,150 and \$1,250?

Solution

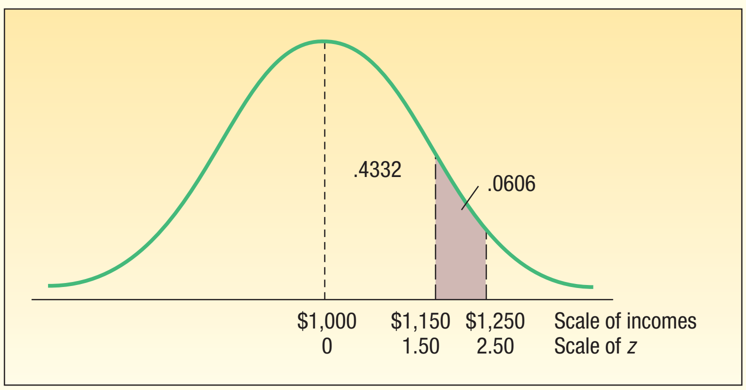
The situation is again separated into two parts, and formula (7–5) is used. First, we find the z value associated with a weekly salary of \$1,250:

$$z = \frac{\$1{,}250 - \$1{,}000}{\$100} = 2.50$$

Next we find the z value for a weekly salary of \$1,150:

$$z = \frac{\$1{,}150 - \$1{,}000}{\$100} = 1.50$$

From Appendix B.1 the area associated with a z value of 2.50 is .4938. So the probability of a weekly salary between \$1,000 and \$1,250 is .4938. Similarly, the area associated with a z value of 1.50 is .4332, so the probability of a weekly salary between \$1,000 and \$1,150 is .4332. The probability of a weekly salary between \$1,150 and \$1,250 is found by subtracting the area associated with a z value of 1.50 (.4332) from that associated with a z of 2.50 (.4938). Thus, the probability of a weekly salary between \$1,150 and \$1,250 is .0606. In probability notation we write $P(\$1{,}150 < \text{weekly income} < \$1{,}250) = .4938 - .4332 = .0606$.



In brief, there are four situations for finding the area under the standard normal probability distribution.

1. To find the area between 0 and z (or $-z$), look up the probability directly in the table.
2. To find the area beyond z or $(-z)$, locate the probability of z in the table and subtract that probability from .5000.
3. To find the area between two points on different sides of the mean, determine the z values and add the corresponding probabilities.
4. To find the area between two points on the same side of the mean, determine the z values and subtract the smaller probability from the larger.

Self-Review 7–5

Refer to the previous example, where the distribution of weekly incomes follows the normal distribution with a mean of \$1,000 and the standard deviation is \$100.

(a) What fraction of the shift foremen earn a weekly income between \$750 and \$1,225? Draw a normal curve and shade the desired area on your diagram.

(b) What fraction of the shift foremen earn a weekly income between \$1,100 and \$1,225? Draw a normal curve and shade the desired area on your diagram.

Exercises

17. A normal distribution has a mean of 50 and a standard deviation of 4.
   a. Compute the probability of a value between 44.0 and 55.0.
   b. Compute the probability of a value greater than 55.0.
   c. Compute the probability of a value between 52.0 and 55.0.

18. A normal population has a mean of 80.0 and a standard deviation of 14.0.
   a. Compute the probability of a value between 75.0 and 90.0.
   b. Compute the probability of a value 75.0 or less.
   c. Compute the probability of a value between 55.0 and 70.0.

19. According to the Internal Revenue Service, the mean tax refund for the year 2004 was $2,454. Assume the standard deviation is $650 and that the amounts refunded follow a normal probability distribution.
   a. What percent of the refunds are more than $3,000?
   b. What percent of the refunds are more than $3,000 but less than $3,500?
   c. What percent of the refunds are more than $2,500 but less than $3,500?

20. The amounts of money requested on home loan applications at Down River Federal Savings follow the normal distribution, with a mean of $70,000 and a standard deviation of $20,000. A loan application is received this morning. What is the probability:
   a. The amount requested is $80,000 or more?
   b. The amount requested is between $65,000 and $80,000?
   c. The amount requested is $65,000 or more?

21. WNAE, an all-news AM station, finds that the distribution of the lengths of time listeners are tuned to the station follows the normal distribution. The mean of the distribution is 15.0 minutes and the standard deviation is 3.5 minutes. What is the probability that a particular listener will tune in:
   a. More than 20 minutes?
   b. For 20 minutes or less?
   c. Between 10 and 12 minutes?

22. Among U.S. cities with a population of more than 250,000 the mean one-way commute to work is 24.3 minutes. The longest one-way travel time is New York City, where the mean time is 38.3 minutes. Assume the distribution of travel times in New York City follows the normal probability distribution and the standard deviation is 7.5 minutes.
   a. What percent of the New York City commutes are for less than 30 minutes?
   b. What percent are between 30 and 35 minutes?
   c. What percent are between 30 and 40 minutes?

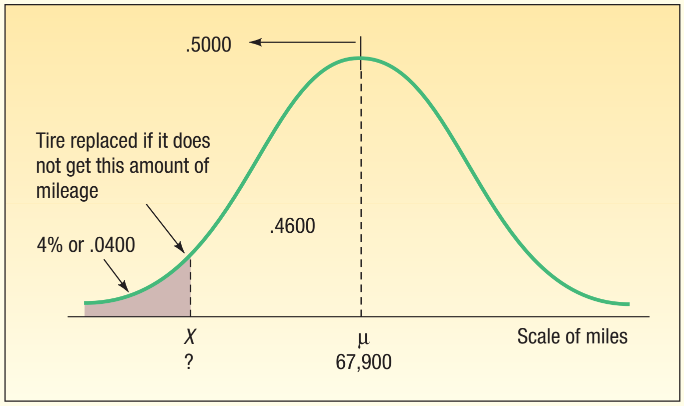
Previous examples require finding the percent of the observations located between two observations or the percent of the observations above, or below, a particular observation X. A further application of the normal distribution involves finding the value of the observation X when the percent above or below the observation is given.

Example

Layton Tire and Rubber Company wishes to set a minimum mileage guarantee on its new MX100 tire. Tests reveal the mean mileage is 67,900 with a standard deviation of 2,050 miles and that the distribution of miles follows the normal probability distribution. Layton wants to set the minimum guaranteed mileage so that no more than 4 percent of the tires will have to be replaced. What minimum guaranteed mileage should Layton announce?

Solution

The facets of this case are shown in the following diagram, where X represents the minimum guaranteed mileage.



Inserting these values in formula (7–5) for z gives:

$$z = \frac{X - \mu}{\sigma} = \frac{X - 67{,}900}{2{,}050}$$

Notice that there are two unknowns, z and X. To find X, we first find z, and then solve for X. Notice the area under the normal curve to the left of μ is .5000. The area between μ and X is .4600, found by .5000 − .0400. Now refer to Appendix B.1. Search the body of the table for the area closest to .4600. The closest area is .4599. Move to the margins from this value and read the z value of 1.75. Because the value is to the left of the mean, it is actually −1.75. These steps are illustrated in Table 7–2.

TABLE 7–2 Selected Areas under the Normal Curve

z ...	.03	.04	.05	.06
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
1.5	.4370	.4382	.4394	.4406
1.6	.4484	.4495	.4505	.4515
1.7	.4582	.4591	.4599	.4608
1.8	.4664	.4671	.4678	.4686

Knowing that the distance between μ and X is -1.75σ or $z = -1.75$, we can now solve for X (the minimum guaranteed mileage):

$$z = \frac{X - 67{,}900}{2{,}050}$$

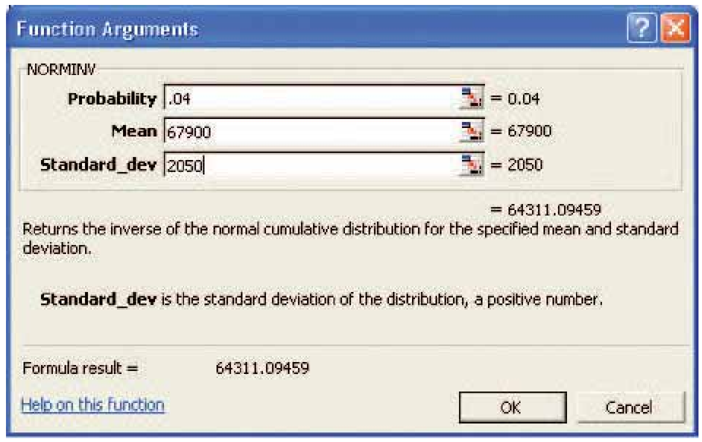
$$-1.75 = \frac{X - 67{,}900}{2{,}050}$$

$$-1.75(2{,}050) = X - 67{,}900$$

$$X = 67{,}900 - 1.75(2{,}050) = 64{,}312$$

So Layton can advertise that it will replace for free any tire that wears out before it reaches 64,312 miles, and the company will know that only 4 percent of the tires will be replaced under this plan.

Excel will also find the mileage value. See the following output. The necessary commands are given in the **Software Commands** section at the end of the chapter.



Self-Review 7–6

An analysis of the final test scores for Introduction to Business reveals the scores follow the normal probability distribution. The mean of the distribution is 75 and the standard deviation is 8. The professor wants to award an A to students whose score is in the highest 10 percent. What is the dividing point for those students who earn an A and those earning a B?

Exercises

23. A normal distribution has a mean of 50 and a standard deviation of 4. Determine the value below which 95 percent of the observations will occur.

24. A normal distribution has a mean of 80 and a standard deviation of 14. Determine the value above which 80 percent of the values will occur.

25. Assume that the mean hourly cost to operate a commercial airplane follows the normal distribution with a mean of $2,100 per hour and a standard deviation of $250. What is the operating cost for the lowest 3 percent of the airplanes?

26. The monthly sales of mufflers in the Richmond, Virginia, area follow the normal distribution with a mean of 1,200 and a standard deviation of 225. The manufacturer would like to establish inventory levels such that there is only a 5 percent chance of running out of stock. Where should the manufacturer set the inventory levels?

27. According to media research, the typical American listened to 195 hours of music in the year 2004. This is down from 290 hours in 1999. Dick Trythall is a big country and western music fan. He listens to music while working around the house, reading, and riding in his truck. Assume the number of hours spent listening to music follows a normal probability distribution with a standard deviation of 8.5 hours.

a. If Dick is in the top 1 percent in terms of listening time, how many hours does he listen per year?

b. Assume that the distribution of times for 1999 also follows the normal probability distribution with a standard deviation of 8.5 hours. How many hours did the 1 percent who listen to the *least* music actually listen?

28. In 2004–2005 the mean annual cost to attend a private university in the United States was $20,082. Assume the distribution of annual costs follows a normal probability distribution and the standard deviation is $4,500. Ninety-five percent of all students at private universities pay less than what amount?

29. The newsstand at the corner of East 9th Street and Euclid Avenue in downtown Cleveland sells the daily edition of the *Cleveland Plain Dealer*. The number of papers sold each day follows a normal probability distribution with a mean of 200 copies and a standard deviation of 17 copies. How many copies should the owner of the newsstand order, so that he only runs out of papers on 20 percent of the days?

30. The manufacturer of a laser printer reports the mean number of pages a cartridge will print before it needs replacing is 12,200. The distribution of pages printed per cartridge closely follows the normal probability distribution and the standard deviation is 820 pages. The manufacturer wants to provide guidelines to potential customers as to how long they can expect a cartridge to last. How many pages should the manufacturer advertise for each cartridge if it wants to be correct 99 percent of the time?

The Normal Approximation to the Binomial

Chapter 6 describes the binomial probability distribution, which is a discrete distribution. The table of binomial probabilities in Appendix B.9 goes successively from an n of 1 to an n of 15. If a problem involved taking a sample of 60, generating a binomial distribution for that large a number would be very time consuming. A more efficient approach is to apply the *normal approximation to the binomial*.

Using the normal distribution (a continuous distribution) as a substitute for a binomial distribution (a discrete distribution) for large values of n seems reasonable because, as n increases, a binomial distribution gets closer and closer to a normal distribution. Chart 7–7 depicts the change in the shape of a binomial distribution with $\pi = .50$ from an n of 1, to an n of 3, to an n of 20. Notice how the case where $n = 20$ approximates the shape of the normal distribution. That is, compare the case where $n = 20$ to the normal curve in Chart 7–3 on page 228.

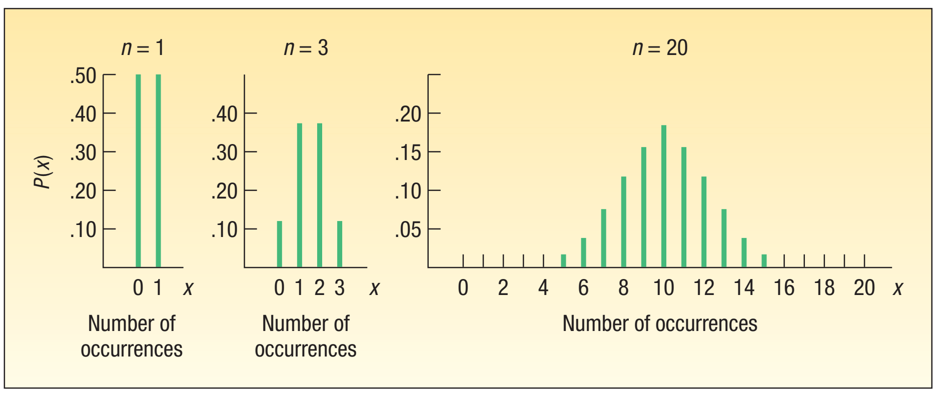


CHART 7–7 Binomial Distributions for an n of 1, 3, and 20, Where $\pi = .50$

When to use the normal approximation

When can we use the normal approximation to the binomial? The normal probability distribution is a good approximation to the binomial probability distribution when $n\pi$ and $n(1 - \pi)$ are both at least 5. However, before we apply the normal approximation, we must make sure that our distribution of interest is in fact a binomial distribution. Recall from Chapter 6 that four criteria must be met:

1. There are only two mutually exclusive outcomes to an experiment: a "success" and a "failure."
2. The distribution results from counting the number of successes in a fixed number of trials.
3. The probability of a success, π, remains the same from trial to trial.
4. Each trial is independent.

Continuity Correction Factor

To show the application of the normal approximation to the binomial and the need for a correction factor, suppose the management of the Santoni Pizza Restaurant found that 70 percent of its new customers return for another meal. For a week

in which 80 new (first-time) customers dined at Santoni's, what is the probability that 60 or more will return for another meal?

Notice the binomial conditions are met: (1) There are only two possible outcomes—a customer either returns for another meal or does not return. (2) We can count the number of successes, meaning, for example, that 57 of the 80 customers return. (3) The trials are independent, meaning that if the 34th person returns for a second meal, that does not affect whether the 58th person returns. (4) The probability of a customer returning remains at .70 for all 80 customers.

Therefore, we could use the binomial formula (6–3) described on page 190.

$$P(x) = {}_nC_x\,(\pi)^x\,(1 - \pi)^{n-x}$$

To find the probability 60 or more customers return for another pizza, we need to first find the probability exactly 60 customers return. That is:

$$P(x = 60) = {}_{80}C_{60}\,(.70)^{60}\,(1 - .70)^{20} = .063$$

Next we find the probability that exactly 61 customers return. It is:

$$P(x = 61) = {}_{80}C_{61}\,(.70)^{61}\,(1 - .70)^{19} = .048$$

We continue this process until we have the probability that all 80 customers return. Finally, we add the probabilities from 60 to 80. Solving the above problem in this manner is tedious. We can also use a computer software package such as MINITAB or Excel to find the various probabilities. Listed below are the binomial probabilities for $n = 80$, $\pi = .70$, and x, the number of customers returning, ranging from 43 to 68. The probability of any number of customers less than 43 or more than 68 returning is less than .001. We can assume these probabilities are 0.000.

Number Returning	Probability	Number Returning	Probability
43	.001	56	.097
44	.002	57	.095
45	.003	58	.088
46	.006	59	.077
47	.009	60	.063
48	.015	61	.048
49	.023	62	.034
50	.033	63	.023
51	.045	64	.014
52	.059	65	.008
53	.072	66	.004
54	.084	67	.002
55	.093	68	.001

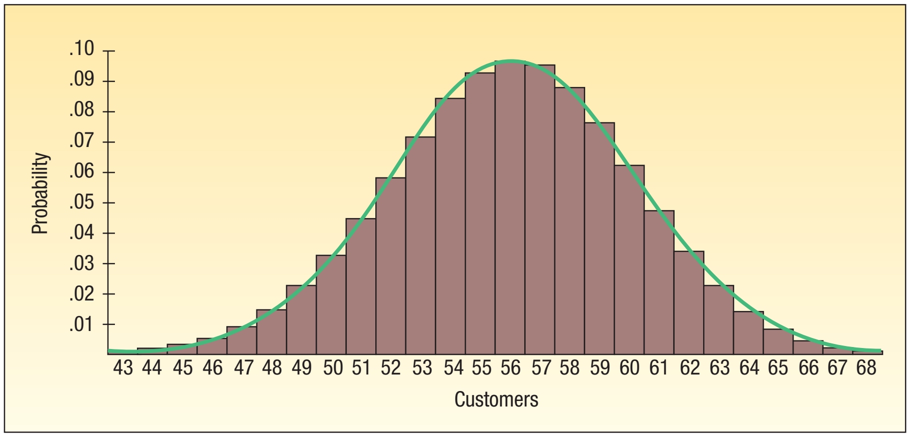
We can find the probability of 60 or more returning by summing $.063 + .048 + \cdots + .001$, which is .197. However, a look at the plot on page 244 shows the similarity of this distribution to a normal distribution. All we need do is "smooth out" the discrete probabilities into a continuous distribution. Furthermore, working with a normal distribution will involve far fewer calculations than working with the binomial.

The trick is to let the discrete probability for 56 customers be represented by an area under the continuous curve between 55.5 and 56.5. Then let the probability for 57 customers be represented by an area between 56.5 and 57.5 and so on. This is just the opposite of rounding off the numbers to a whole number.

Statistics in Action

Many variables are approximately, normally distributed, such as IQ scores, life expectancies, and adult height. This implies that nearly all observations occur within 3 standard deviations of the mean. On the other hand, observations that occur beyond 3 standard deviations from the mean are extremely rare. For example, the mean adult male height is 68.2 inches (about 5 feet 8 inches) with a standard deviation of 2.74. This means that almost all males are between 60.0 inches (5 feet) and 76.4 inches (6 feet 4 inches). Shaquille O'Neal, a professional basketball player with the Miami Heat, is 86 inches or 7 feet 2 inches, which is clearly beyond 3 standard deviations from the mean. The height of a standard doorway is 6 feet 8 inches, and should be high enough for almost all adult males, except for a rare person like Shaquille O'Neal.

(continued)



Because we use the normal distribution to determine the binomial probability of 60 or more successes, we must subtract, in this case, .5 from 60. The value .5 is called the **continuity correction factor.** This small adjustment must be made because a continuous distribution (the normal distribution) is being used to approximate a discrete distribution (the binomial distribution). Subtracting, $60 - .5 = 59.5$.

CONTINUITY CORRECTION FACTOR The value .5 subtracted or added, depending on the question, to a selected value when a discrete probability distribution is approximated by a continuous probability distribution.

How to Apply the Correction Factor

Only four cases may arise. These cases are:

1. For the probability *at least* X occurs, use the area *above* $(X - .5)$.
2. For the probability that *more than* X occurs, use the area *above* $(X + .5)$.
3. For the probability that X *or fewer* occurs, use the area *below* $(X + .5)$.
4. For the probability that *fewer than* X occurs, use the area *below* $(X - .5)$.

To use the normal distribution to approximate the probability that 60 or more first-time Santoni customers out of 80 will return, follow the procedure shown below.

Step 1. Find the z corresponding to an X of 59.5 using formula (7–5), and formulas (6–4) and (6–5) for the mean and the variance of a binomial distribution:

$$\mu = n\pi = 80(.70) = 56$$

$$\sigma^2 = n\pi(1 - \pi) = 80(.70)(1 - .70) = 16.8$$

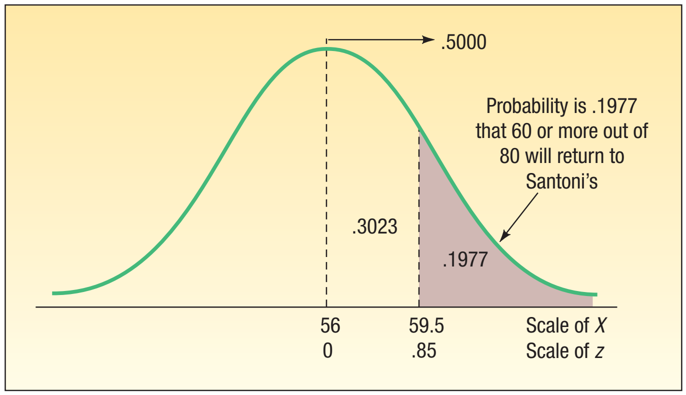
$$\sigma = \sqrt{16.8} = 4.10$$

$$z = \frac{X - \mu}{\sigma} = \frac{59.5 - 56}{4.10} = 0.85$$

Step 2. Determine the area under the normal curve between a μ of 56 and an X of 59.5. From step 1, we know that the z value corresponding to 59.5 is 0.85. So we go to Appendix B.1 and read down the left margin to 0.8, and then we go horizontally to the area under the column headed by .05. That area is .3023.

As another example, the driver's seat in most vehicles is set to comfortably fit a person who is at least 159 cm (62.5 inches) tall. The distribution of heights of adult women is approximately a normal distribution with a mean of 161.5 cm and a standard deviation of 6.3 cm. Thus about 35 percent of adult women will not fit comfortably in the driver's seat.

Step 3. Calculate the area beyond 59.5 by subtracting .3023 from .5000 (.5000 − .3023 = .1977). Thus, .1977 is the probability that 60 or more first-time Santoni customers out of 80 will return for another meal. In probability notation $P(\text{customers} > 59.5) = .5000 - .3023 = .1977$. The facets of this problem are shown graphically:



No doubt you will agree that using the normal approximation to the binomial is a more efficient method of estimating the probability of 60 or more first-time customers returning. The result compares favorably with that computed on page 243, using the binomial distribution. The probability using the binomial distribution is .197, whereas the probability using the normal approximation is .1977.

Self-Review 7–7

A study by Great Southern Home Insurance revealed that none of the stolen goods were recovered by the homeowners in 80 percent of reported thefts.

(a) During a period in which 200 thefts occurred, what is the probability that no stolen goods were recovered in 170 or more of the robberies?

(b) During a period in which 200 thefts occurred, what is the probability that no stolen goods were recovered in 150 or more robberies?

Exercises

31. Assume a binomial probability distribution with $n = 50$ and $\pi = .25$. Compute the following:
- **a.** The mean and standard deviation of the random variable.
- **b.** The probability that X is 15 or more.
- **c.** The probability that X is 10 or less.

32. Assume a binomial probability distribution with $n = 40$ and $\pi = .55$. Compute the following:
- **a.** The mean and standard deviation of the random variable.
- **b.** The probability that X is 25 or greater.
- **c.** The probability that X is 15 or less.
- **d.** The probability that X is between 15 and 25, inclusive.

33. Dottie's Tax Service specializes in federal tax returns for professional clients, such as physicians, dentists, accountants, and lawyers. A recent audit by the IRS of the returns she prepared indicated that an error was made on 5 percent of the returns she prepared last year. Assuming this rate continues into this year and she prepares 60 returns, what is the probability that she makes errors on:
- **a.** More than six returns?
- **b.** At least six returns?
- **c.** Exactly six returns?

34. Shorty's Muffler advertises it can install a new muffler in 30 minutes or less. However, the work standards department at corporate headquarters recently conducted a study and found that 20 percent of the mufflers were not installed in 30 minutes or less. The Maumee branch installed 50 mufflers last month. If the corporate report is correct:
   a. How many of the installations at the Maumee branch would you expect to take more than 30 minutes?
   b. What is the likelihood that fewer than eight installations took more than 30 minutes?
   c. What is the likelihood that eight or fewer installations took more than 30 minutes?
   d. What is the likelihood that exactly 8 of the 50 installations took more than 30 minutes?

35. A study conducted by the nationally known Taurus Health Club revealed that 30 percent of its new members are significantly overweight. A membership drive in a metropolitan area resulted in 500 new members.
   a. It has been suggested that the normal approximation to the binomial be used to determine the probability that 175 or more of the new members are significantly overweight. Does this problem qualify as a binomial problem? Explain.
   b. What is the probability that 175 or more of the new members are significantly overweight?
   c. What is the probability that 140 or more new members are significantly overweight?

36. A recent issue of *Bride Magazine* suggested that couples planning their wedding should expect two-thirds of those who are sent an invitation to respond that they will attend. Rich and Stacy are planning to be married later this year. They plan to send 197 invitations.
   a. How many guests would you expect to accept the invitation?
   b. What is the standard deviation?
   c. What is the probability 140 or more will accept the invitation?
   d. What is the probability exactly 140 will accept the invitation?

Chapter Summary

I. The uniform distribution is a continuous probability distribution with the following characteristics.
   A. It is rectangular in shape.
   B. The mean and the median are equal.
   C. It is completely described by its minimum value *a* and its maximum value *b*.
   D. It is also described by the following equation for the region from *a* to *b*:

$$P(x) = \frac{1}{b - a} \qquad \textbf{[7–3]}$$

   E. The mean and standard deviation of a uniform distribution are computed as follows:

$$\mu = \frac{(a + b)}{2} \qquad \textbf{[7–1]}$$

$$\sigma = \sqrt{\frac{(b - a)^2}{12}} \qquad \textbf{[7–2]}$$

II. The normal probability distribution is a continuous distribution with the following characteristics.
   A. It is bell-shaped and has a single peak at the center of the distribution.
   B. The distribution is symmetric.
   C. It is asymptotic, meaning the curve approaches but never touches the *X*-axis.
   D. It is completely described by its mean and standard deviation.
   E. There is a family of normal probability distributions.
      1. Another normal probability distribution is created when either the mean or the standard deviation changes.
      2. The normal probability distribution is described by the following formula:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x - \mu)^2}{2\sigma^2}\right]} \qquad \textbf{[7–4]}$$

III. The standard normal probability distribution is a particular normal distribution.
- **A.** It has a mean of 0 and a standard deviation of 1.
- **B.** Any normal probability distribution can be converted to the standard normal probability distribution by the following formula.

$$z = \frac{X - \mu}{\sigma} \qquad \textbf{[7–5]}$$

- **C.** By standardizing a normal probability distribution, we can report the distance of a value from the mean in units of the standard deviation.

IV. The normal probability distribution can approximate a binomial distribution under certain conditions.
- **A.** $n\pi$ and $n(1 - \pi)$ must both be at least 5.
  - **1.** n is the number of observations.
  - **2.** π is the probability of a success.
- **B.** The four conditions for a binomial probability distribution are:
  - **1.** There are only two possible outcomes.
  - **2.** π remains the same from trial to trial.
  - **3.** The trials are independent.
  - **4.** The distribution results from a count of the number of successes in a fixed number of trials.
- **C.** The mean and variance of a binomial distribution are computed as follows:

$$\mu = n\pi$$
$$\sigma^2 = n\pi(1 - \pi)$$

- **D.** The continuity correction factor of .5 is used to extend the continuous value of X one-half unit in either direction. This correction compensates for approximating a discrete distribution by a continuous distribution.

Chapter Exercises

37. The amount of cola in a 12-ounce can is uniformly distributed between 11.96 ounces and 12.05 ounces.
- **a.** What is the mean amount per can?
- **b.** What is the standard deviation amount per can?
- **c.** What is the probability of selecting a can of cola and finding it has less than 12 ounces?
- **d.** What is the probability of selecting a can of cola and finding it has more than 11.98 ounces?
- **e.** What is the probability of selecting a can of cola and finding it has more than 11.00 ounces?

38. A tube of Listerine Tartar Control toothpaste contains 4.2 ounces. As people use the toothpaste, the amount remaining in any tube is random. Assume the amount of toothpaste left in the tube follows a uniform distribution. From this information, we can determine the following information about the amount remaining in a toothpaste tube without invading anyone's privacy.
- **a.** How much toothpaste would you expect to be remaining in the tube?
- **b.** What is the standard deviation of the amount remaining in the tube?
- **c.** What is the likelihood there is less than 3.0 ounces remaining in the tube?
- **d.** What is the probability there is more than 1.5 ounces remaining in the tube?

39. Many retail stores offer their own credit cards. At the time of the credit application the customer is given a 10 percent discount on the purchase. The time required for the credit application process follows a uniform distribution with the times ranging from 4 minutes to 10 minutes.
- **a.** What is the mean time for the application process?
- **b.** What is the standard deviation of the process time?
- **c.** What is the likelihood a particular application will take less than 6 minutes?
- **d.** What is the likelihood an application will take more than 5 minutes?

40. The time patrons at the Grande Dunes Hotel in the Bahamas spend waiting for an elevator follows a uniform distribution between 0 and 3.5 minutes.
- **a.** Show that the area under the curve is 1.00.
- **b.** How long does the typical patron wait for elevator service?

c. What is the standard deviation of the waiting time?
d. What percent of the patrons wait for less than a minute?
e. What percent of the patrons wait more than 2 minutes?

41. The net sales and the number of employees for aluminum fabricators with similar characteristics are organized into frequency distributions. Both are normally distributed. For the net sales, the mean is $180 million and the standard deviation is $25 million. For the number of employees, the mean is 1,500 and the standard deviation is 120. Clarion Fabricators had sales of $170 million and 1,850 employees.
a. Convert Clarion's sales and number of employees to *z* values.
b. Locate the two *z* values.
c. Compare Clarion's sales and number of employees with those of the other fabricators.

42. The accounting department at Weston Materials, Inc., a national manufacturer of unattached garages, reports that it takes two construction workers a mean of 32 hours and a standard deviation of 2 hours to erect the Red Barn model. Assume the assembly times follow the normal distribution.
a. Determine the *z* values for 29 and 34 hours. What percent of the garages take between 32 hours and 34 hours to erect?
b. What percent of the garages take between 29 hours and 34 hours to erect?
c. What percent of the garages take 28.7 hours or less to erect?
d. Of the garages, 5 percent take how many hours or more to erect?

43. A recent report in *USA Today* indicated a typical family of four spends $490 per month on food. Assume the distribution of food expenditures for a family of four follows the normal distribution, with a mean of $490 and a standard deviation of $90.
a. What percent of the families spend more than $30 but less than $490 per month on food?
b. What percent of the families spend less than $430 per month on food?
c. What percent spend between $430 and $600 per month on food?
d. What percent spend between $500 and $600 per month on food?

44. A study of long-distance phone calls made from the corporate offices of Pepsi Bottling Group, Inc., in Somers, New York, revealed the length of the calls, in minutes, follows the normal probability distribution. The mean length of time per call was 4.2 minutes and the standard deviation was 0.60 minutes.
a. What fraction of the calls last between 4.2 and 5 minutes?
b. What fraction of the calls last more than 5 minutes?
c. What fraction of the calls last between 5 and 6 minutes?
d. What fraction of the calls last between 4 and 6 minutes?
e. As part of her report to the president, the director of communications would like to report the length of the longest (in duration) 4 percent of the calls. What is this time?

45. Shaver Manufacturing, Inc., offers dental insurance to its employees. A recent study by the human resource director shows the annual cost per employee per year followed the normal probability distribution, with a mean of $1,280 and a standard deviation of $420 per year.
a. What fraction of the employees cost more than $1,500 per year for dental expenses?
b. What fraction of the employees cost between $1,500 and $2,000 per year?
c. Estimate the percent that did not have any dental expense.
d. What was the cost for the 10 percent of employees who incurred the highest dental expense?

46. The annual commissions earned by sales representatives of Machine Products, Inc., a manufacturer of light machinery, follow the normal probability distribution. The mean yearly amount earned is $40,000 and the standard deviation is $5,000.
a. What percent of the sales representatives earn more than $42,000 per year?
b. What percent of the sales representatives earn between $32,000 and $42,000?
c. What percent of the sales representatives earn between $32,000 and $35,000?
d. The sales manager wants to award the sales representatives who earn the largest commissions a bonus of $1,000. He can award a bonus to 20 percent of the representatives. What is the cutoff point between those who earn a bonus and those who do not?

47. According to the South Dakota Department of Health, the mean number of hours of TV viewing per week is higher among adult women than men. A recent study showed women spent an average of 34 hours per week watching TV and men 29 hours per week (www.state.sd.us/DOH/Nutriton/TV.pdf). Assume that the distribution of hours watched follows the normal distribution for both groups, and that the standard deviation among the women is 4.5 hours and is 5.1 hours for the men.

a. What percent of the women watch TV less than 40 hours per week?
b. What percent of the men watch TV more than 25 hours per week?
c. How many hours of TV do the one percent of women who watch the most TV per week watch? Find the comparable value for the men.

48. According to a government study among adults in the 25- to 34-year age group, the mean amount spent per year on reading and entertainment is $1,994 (www.infoplease.com/ipa/A0908759.html). Assume that the distribution of the amounts spent follows the normal distribution with a standard deviation of $450.
a. What percent of the adults spend more than $2,500 per year on reading and entertainment?
b. What percent spend between $2,500 and $3,000 per year on reading and entertainment?
c. What percent spend less than $1,000 per year on reading and entertainment?

49. Management at Gordon Electronics is considering adopting a bonus system to increase production. One suggestion is to pay a bonus on the highest 5 percent of production based on past experience. Past records indicate weekly production follows the normal distribution. The mean of this distribution is 4,000 units per week and the standard deviation is 60 units per week. If the bonus is paid on the upper 5 percent of production, the bonus will be paid on how many units or more?

50. Fast Service Truck Lines uses the Ford Super Duty F-750 exclusively. Management made a study of the maintenance costs and determined the number of miles traveled during the year followed the normal distribution. The mean of the distribution was 60,000 miles and the standard deviation 2,000 miles.
a. What percent of the Ford Super Duty F-750s logged 65,200 miles or more?
b. What percent of the trucks logged more than 57,060 but less than 58,280 miles?
c. What percent of the Fords traveled 62,000 miles or less during the year?
d. Is it reasonable to conclude that any of the trucks were driven more than 70,000 miles? Explain.

51. Best Electronics, Inc., offers a "no hassle" returns policy. The number of items returned per day follows the normal distribution. The mean number of customer returns is 10.3 per day and the standard deviation is 2.25 per day.
a. In what percent of the days are there 8 or fewer customers returning items?
b. In what percent of the days are between 12 and 14 customers returning items?
c. Is there any chance of a day with no returns?

52. A recent report in *BusinessWeek* indicated that 20 percent of all employees steal from their company each year. If a company employs 50 people, what is the probability that:
a. Fewer than 5 employees steal?
b. More than 5 employees steal?
c. Exactly 5 employees steal?
d. More than 5 but fewer than 15 employees steal?

53. The *Orange County Register*, as part of its Sunday health supplement, reported that 64 percent of American men over the age of 18 consider nutrition a top priority in their lives. Suppose we select a sample of 60 men. What is the likelihood that:
a. 32 or more consider nutrition important?
b. 44 or more consider nutrition important?
c. More than 32 but fewer than 43 consider nutrition important?
d. Exactly 44 consider diet important?

54. It is estimated that 10 percent of those taking the quantitative methods portion of the CPA examination fail that section. Sixty students are taking the exam this Saturday.
a. How many would you expect to fail? What is the standard deviation?
b. What is the probability that exactly two students will fail?
c. What is the probability at least two students will fail?

55. The Georgetown, South Carolina, Traffic Division reported 40 percent of high-speed chases involving automobiles result in a minor or major accident. During a month in which 50 high-speed chases occur, what is the probability that 25 or more will result in a minor or major accident?

56. Cruise ships of the Royal Viking line report that 80 percent of their rooms are occupied during September. For a cruise ship having 800 rooms, what is the probability that 665 or more are occupied in September?

57. The goal at U.S. airports handling international flights is to clear these flights within 45 minutes. Let's interpret this to mean that 95 percent of the flights are cleared in 45 minutes, so 5 percent of the flights take longer to clear. Let's also assume that the distribution is approximately normal.

a. If the standard deviation of the time to clear an international flight is 5 minutes, what is the mean time to clear a flight?
b. Suppose the standard deviation is 10 minutes, not the 5 minutes suggested in part (a). What is the new mean?
c. A customer has 30 minutes from the time her flight lands to catch her limousine. Assuming a standard deviation of 10 minutes, what is the likelihood that she will be cleared in time?

58. The funds dispensed at the ATM machine located near the checkout line at the Kroger's in Union, Kentucky, follows a normal probability distribution with a mean of $4,200 per day and a standard deviation of $720 per day. The machine is programmed to notify the nearby bank if the amount dispensed is very low (less than $2,500) or very high (more than $6,000).
a. What percent of the days will the bank be notified because the amount dispensed is very low?
b. What percent of the time will the bank be notified because the amount dispensed is high?
c. What percent of the time will the bank not be notified regarding the amount of funds dispersed?

59. The weights of canned hams processed at Henline Ham Company follow the normal distribution, with a mean of 9.20 pounds and a standard deviation of 0.25 pounds. The label weight is given as 9.00 pounds.
a. What proportion of the hams actually weigh less than the amount claimed on the label?
b. The owner, Glen Henline, is considering two proposals to reduce the proportion of hams below label weight. He can increase the mean weight to 9.25 and leave the standard deviation the same, or he can leave the mean weight at 9.20 and reduce the standard deviation from 0.25 pounds to 0.15. Which change would you recommend?

60. The *Cincinnati Enquirer*, in its Sunday business supplement, reported that the mean number of hours worked per week by those employed full time is 43.9. The article further indicated that about one-third of those employed full time work less than 40 hours per week.
a. Given this information and assuming that number of hours worked follows the normal distribution, what is the standard deviation of the number of hours worked?
b. The article also indicated that 20 percent of those working full time work more than 49 hours per week. Determine the standard deviation with this information. Are the two estimates of the standard deviation similar? What would you conclude?

61. Most four-year automobile leases allow up to 60,000 miles. If the lessee goes beyond this amount, a penalty of 20 cents per mile is added to the lease cost. Suppose the distribution of miles driven on four-year leases follows the normal distribution. The mean is 52,000 miles and the standard deviation is 5,000 miles.
a. What percent of the leases will yield a penalty because of excess mileage?
b. If the automobile company wanted to change the terms of the lease so that 25 percent of the leases went over the limit, where should the new upper limit be set?
c. One definition of a low-mileage car is one that is 4 years old and has been driven less than 45,000 miles. What percent of the cars returned are considered low-mileage?

62. The price of shares of Bank of Florida at the end of trading each day for the last year followed the normal distribution. Assume there were 240 trading days in the year. The mean price was $42.00 per share and the standard deviation was $2.25 per share.
a. What percent of the days was the price over $45.00? How many days would you estimate?
b. What percent of the days was the price between $38.00 and $40.00?
c. What was the stock's price on the *highest* 15 percent of days?

63. The annual sales of romance novels follow the normal distribution. However, the mean and the standard deviation are unknown. Forty percent of the time sales are more than 470,000, and 10 percent of the time sales are more than 500,000. What are the mean and the standard deviation?

64. In establishing warranties on HDTV sets, the manufacturer wants to set the limits so that few will need repair at the manufacturer's expense. On the other hand, the warranty period must be long enough to make the purchase attractive to the buyer. For a new HDTV the mean number of months until repairs are needed is 36.84 with a standard deviation of 3.34 months. Where should the warranty limits be set so that only 10 percent of the HDTVs need repairs at the manufacturer's expense?

65. DeKorte Tele-Marketing, Inc., is considering purchasing a machine that randomly selects and automatically dials telephone numbers. DeKorte Tele-Marketing makes most of its calls during the evening, so calls to business phones are wasted. The manufacturer of the machine claims that its programming reduces the calling to business phones to 15 percent of all calls. To test this claim the director of purchasing at DeKorte programmed the machine to select a sample of 150 phone numbers. What is the likelihood that more than 30 of the phone numbers selected are those of businesses, assuming the manufacturer's claim is correct?

Data Set Exercises

66. Refer to the Real Estate data, which report information on homes sold in the Denver, Colorado, area during the last year.
- **a.** The mean selling price (in $ thousands) of the homes was computed earlier to be $221.10, with a standard deviation of $47.11. Use the normal distribution to estimate the percent of homes selling for more than $280.0. Compare this to the actual results. Does the normal distribution yield a good approximation of the actual results?
- **b.** The mean distance from the center of the city is 14.629 miles with a standard deviation of 4.874 miles. Use the normal distribution to estimate the number of homes 18 or more miles but less than 22 miles from the center of the city. Compare this to the actual results. Does the normal distribution yield a good approximation of the actual results?

67. Refer to the Baseball 2005 data, which report information on the 30 Major League Baseball teams for the 2005 season.
- **a.** The mean attendance per team for the season was 2,496,458 with a standard deviation of 672,879. Use the normal distribution to estimate the number of teams with attendance of more than 3.5 million. Compare that estimate with the actual number. Comment on the accuracy of your estimate.
- **b.** The mean team salary was 73.06 million with a standard deviation of 34.23 million. Use the normal distribution to estimate the number of teams with a team salary of more than $50 million. Compare that estimate with the actual number. Comment on the accuracy of the estimate.

68. Refer to the CIA data, which report demographic and economic information on 46 countries.
- **a.** The mean of the GDP/capita variable is 16.58 with a standard deviation of 9.27. Use the normal distribution to estimate the percentage of countries with exports above 24. Compare this estimate with the actual proportion. Does the normal distribution appear accurate in this case? Explain.
- **b.** The mean of the exports is 116.3 with a standard deviation of 157.4. Use the normal distribution to estimate the percentage of countries with exports above 170. Compare this estimate with the actual proportion. Does the normal distribution appear accurate in this case? Explain.

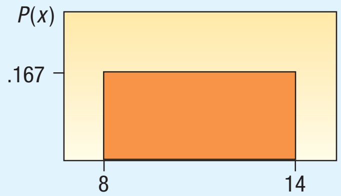
Software Commands

1. The Excel commands necessary to produce the output on page 235 are:
- **a.** Select **Insert** and **Function,** then from the category box select **Statistical** and below that **NORMDIST** and click **OK.**
- **b.** In the dialog box put *1100* in the box for **X**, *1000* for the **Mean,** *100* for the **Standard_dev,** *True* in the **Cumulative** box, and click **OK.**
- **c.** The result will appear in the dialog box. If you click **OK,** the answer appears in your spreadsheet.

2. The Excel Commands necessary to produce the output on page 241 are:
- **a.** Select **Insert** and **Function,** then from the category box select **Statistical** and then below **NORMINV** and click **OK.**
- **b.** In the dialog box, set the **Probability** to *.04*, the **Mean** to *67900,* and the **Standard_dev** to 2050.
- **c.** The results will appear in the dialog box. Note that the answer is different from page 240 because of rounding error. If you click **OK,** the answer also appears in your spreadsheet.
- **d.** Try entering a **Probability** of *.04*, a **Mean** of *0*, and a **Standard_dev** of *1*. The z value will be computed.

Chapter 7 Answers to Self-Review

7–1 a.



b. $P(x) = \text{(height)(base)}$

$$= \left(\frac{1}{14-8}\right)(14-8)$$

$$= \left(\frac{1}{6}\right)(6) = 1.00$$

c. $\mu = \frac{a+b}{2} = \frac{14+8}{2} = \frac{22}{2} = 11$

$$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(14-8)^2}{12}} = \sqrt{\frac{36}{12}} = \sqrt{3}$$

$= 1.73$

d. $P(10 < x < 14) = \text{(height)(base)}$

$$= \left(\frac{1}{14-8}\right)(14-10)$$

$$= \frac{1}{6}(4)$$

$= .667$

e. $P(x < 9) = \text{(height)(base)}$

$$= \left(\frac{1}{14-8}\right)(9-8)$$

$= 0.167$

7–2 a. 2.25, found by:

$$z = \frac{\$1{,}225 - \$1{,}000}{\$100} = \frac{\$225}{\$100} = 2.25$$

b. −2.25, found by:

$$z = \frac{\$775 - \$1{,}000}{\$100} = \frac{-\$225}{\$100} = -2.25$$

7–3 a. \$46,400 and \$48,000, found by \$47,200 ± 1(\$800).

b. \$45,600 and \$48,800, found by \$47,200 ± 2(\$800).

c. \$44,800 and \$49,600, found by \$47,200 ± 3(\$800).

d. \$47,200. The mean, median, and mode are equal for a normal distribution.

e. Yes, a normal distribution is symmetrical.

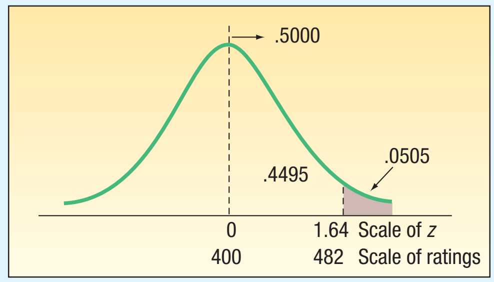
7–4 a. Computing z:

$$z = \frac{482 - 400}{50} = +1.64$$

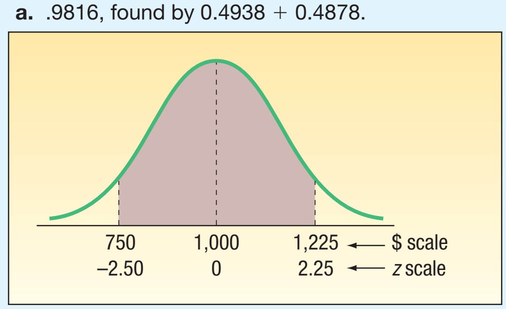
Referring to Appendix B.1, the area is .4495.
$P(400 < \text{rating} < 482) = .4495$

b. .0505, found by .5000 − .4495
$P(\text{rating} > 482) = .5000 - .4495 = .0505$

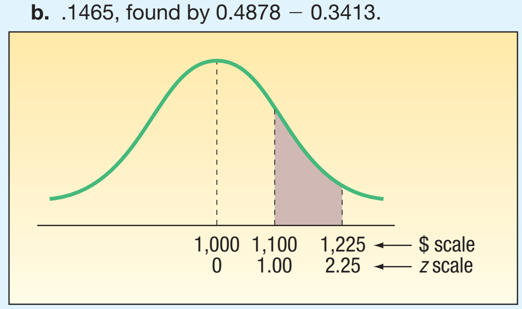
c.



7–5 a. .9816, found by 0.4938 + 0.4878.



b. .1465, found by 0.4878 − 0.3413.



7–6 85.24 (instructor would no doubt make it 85). The closest area to .4000 is .3997; z is 1.28. Then:

$$1.28 = \frac{X - 75}{8}$$

$$10.24 = X - 75$$

$$X = 85.24$$

7–7 a. .0465, found by $\mu = n\pi = 200(.80) = 160$, and $\sigma^2 = n\pi(1-\pi) = 200(.80)(1 - .80) = 32$. Then,

$$\sigma = \sqrt{32} = 5.66$$

$$z = \frac{169.5 - 160}{5.66} = 1.68$$

Area from Appendix B.1 is .4535. Subtracting from .5000 gives .0465.

b. .9686, found by .4686 + .5000. First calculate z:

$$z = \frac{149.5 - 160}{5.66} = -1.86$$

Area from Appendix B.1 is .4686.

A Review of Chapters 5–7

This section is a review of the major concepts, terms, symbols, and equations introduced in Chapters 5, 6, and 7. These three chapters are concerned with methods of dealing with uncertainty. As an example of the uncertainty in business, consider the role of the quality assurance department in most mass-production firms. Usually, the department has neither the personnel nor the time to check, say, all 200 plug-in modules produced during a two-hour period. Standard operating procedure may call for selecting a sample of 5 modules and shipping all 200 modules if the 5 operate correctly. However, if 1 or more in the sample are defective, all 200 are checked. Assuming that all 5 function correctly, quality assurance personnel cannot be absolutely certain that their action (allowing shipment of the modules) will prove to be correct. The study of probability allows us to measure the uncertainty of shipping defective modules. Also, probability as a measurement of uncertainty comes into play when Gallup, Harris, and other pollsters predict that Jim Barstow will win the vacant senatorial seat in Georgia.

Chapter 5 notes that a *probability* is a value between 0 and 1, inclusive, that expresses one's belief that a particular event will occur. A weather forecaster might state that the probability of rain tomorrow is .20. The project director of a firm bidding on a subway station in Bangkok might assess the firm's chance of being awarded the contract at .50. We look at the ways probabilities can be combined using rules of addition and multiplication, some principles of counting, and the importance of Bayes' theorem.

Chapter 6 presents *discrete* probability distributions—the *binomial distribution*, the *hypergeometric distribution*, and the *Poisson distribution*. Other probability distributions will be discussed in forthcoming chapters (*t* distribution, chi-square distribution, etc.). Probability distributions are listings of all the possible outcomes of an experiment and the probability associated with each outcome. A probability distribution allows us to evaluate sample results.

Chapter 7 describes two continuous probability distributions—the *uniform probability distribution* and the *normal probability distribution*. The uniform probability distribution is rectangular in shape and is defined by a minimum and maximum value. The mean and the median of a uniform probability distribution are equal, and it does not have a mode.

A normal probability distribution is used to describe phenomena that follow a normal bell-shaped distribution, such as the tensile strength of wires and the weights or volumes of cans and bottles. Actually, there is a family of normal probability distributions—each with its own mean and standard deviation. There is a normal probability distribution, for example, for a mean of $100 and a standard deviation of $5, another for a mean of $149 and a standard deviation of $5.26, and so on. A normal probability distribution is symmetrical about its mean, and the tails of the normal curve extend in either direction infinitely.

Since there are an unlimited number of normal probability distributions, it is difficult to assign probabilities. Instead, any normal distribution can be changed to a *standard normal probability distribution* by computing *z scores*. The standard normal probability distribution has a mean of 0 and a standard deviation of 1. It is useful because the probability for any event from a normal probability distribution can be computed using standard normal probability tables.

Glossary

Chapter 5

Bayes' theorem Developed by Reverend Bayes in the 1700s, it is designed to find the probability of one event, *A*, occurring, given that another event, *B*, has already occurred.

Classical probability Probability based on the assumption that each of the outcomes is equally likely. According to this concept of probability, if there are *n* possible outcomes, the probability of a particular outcome is $1/n$. Thus, on the toss of a coin, the probability of a head is $1/n = 1/2$.

Combination formula A formula to count the number of possible outcomes. If the order *a*, *b*, *c* is considered the same as *b*, *a*, *c*, or *c*, *b*, *a*, and so on, the number of arrangements is found by:

$$_nC_r = \frac{n!}{r!(n - r)!}$$

Conditional probability The likelihood that an event will occur given that another event has already occurred.

Empirical probability A concept of probability based on past experience. For example, Metropolitan Life Insurance Company reported that, during the year, 100.2 of every 100,000 persons in Wyoming died of accidental causes (motor vehicle accidents, falls, drowning, firearms, etc.). On the basis of this experience, Metropolitan can estimate the probability of accidental death for a particular person in Wyoming: 100.2/100,000 = .001002.

Event A collection of one or more outcomes of an experiment. For example, an event is the collection of even numbers in the roll of a fair die.

Experiment An activity that is either observed or measured. An experiment may be counting the number of correct responses to a question, for example.

General rule of addition Used to find the probabilities of complex events made up of *A* or *B.*

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

General rule of multiplication Used to find the probabilities of events *A* and *B* both happening. Example: It is known that there are 3 defective radios in a box containing 10 radios. What is the probability of selecting 2 defective radios on the first two selections from the box?

$$P(A \text{ and } B) = P(A)P(B \mid A) = \frac{3}{10} \times \frac{2}{9} = \frac{6}{90} = .067$$

where $P(B|A)$ is the conditional probability and means "the probability of *B* occurring given that *A* has already occurred."

Independent The occurrence of one event has no effect on the probability of the occurrence of another event.

Multiplication formula One of the formulas used to count the number of possible outcomes of an experiment. It states that if there are *m* ways of doing one thing and *n* ways of doing another, there are $m \times n$ ways of doing both. Example: A men's clothier offers two sport coats and three contrasting pants for $400. How many different outfits can there be? Answer: $m \times n = 2 \times 3 = 6$.

Outcome A particular observation or measurement of an experiment.

Permutation formula A formula to count the number of possible outcomes. If *a*, *b*, *c* is one arrangement, *b*, *a*, *c* another, *c*, *a*, *b* another, and so on, the total number of arrangements is determined by

$$_nP_r = \frac{n!}{(n-r)!}$$

Probability A value between 0 and 1, inclusive, that reports the likelihood that a specific event will occur.

Special rule of addition For this rule to apply, the events must be mutually exclusive. For two events, the probability of *A* or *B* occurring is found by:

$$P(A \text{ or } B) = P(A) + P(B)$$

Example: The probability of a one-spot or a two-spot occurring on the toss of one die.

$$P(A \text{ or } B) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

Special rule of multiplication If two events are not related—that is, they are independent—this rule is applied to determine the probability of their joint occurrence.

$$P(A \text{ and } B) = P(A)P(B)$$

Example: The probability of two heads on two tosses of a coin is:

$$P(A \text{ and } B) = P(A)P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Subjective probability The chance of an event happening based on whatever information is available—hunches, personal opinion, opinions of others, rumors, and so on.

Chapter 6

Binomial probability distribution A probability distribution based on a discrete random variable. Its major characteristics are:

1. Each outcome can be classified into one of two mutually exclusive categories.
2. The distribution is the result of counting the number of successes.
3. Each trial is independent, meaning that the answer to trial 1 (correct or wrong) in no way affects the answer to trial 2.
4. The probability of a success stays the same from trial to trial.

Continuous random variable A random variable that may assume an infinite number of values within a given range.

Discrete random variable A random variable that can assume only certain separate values.

Hypergeometric probability distribution A probability distribution based on a discrete random variable. Its major characteristics are:

1. There is a fixed number of trials.
2. The probability of success is not the same from trial to trial.
3. There are only two possible outcomes.

Poisson distribution A distribution often used to approximate binomial probabilities when *n* is large and π is small. What is considered "large" or "small" is not precisely defined, but a general rule is that *n* should be equal to or greater than 20 and π equal to or less than .05.

Probability distribution A listing of the possible outcomes of an experiment and the probability associated with each outcome.

Random variable A quantity obtained from an experiment that may, by chance, result in different values. For example, a count of the number of accidents (the experiment) on I-75 during a week might be 10, or 11, or 12, or some other number.

Chapter 7

Continuity correction factor Used to improve the accuracy of the approximation of a discrete distribution by a continuous distribution.

Normal probability distribution A continuous distribution that is bell-shaped, with the mean dividing the distribution into two equal parts. Further, the normal curve extends indefinitely in either direction, and it never touches the *X*-axis. The distribution is defined by its mean and standard deviation.

Uniform probability distribution A continuous probability distribution that is rectangular in shape. It is completely described by using the minimum and maximum values of the distribution to compute the mean and standard deviation. Also, minimum and maximum values are used to compute the probability for any event.

***z* value** The distance between a selected value and the population mean measured in units of the standard deviation.

Exercises

Part I—Multiple Choice

1. Which of the following is *not* a correct statement about a probability?
   - **a.** It must have a value between 0 and 1.
   - **b.** It can be reported as a decimal or a fraction.
   - **c.** A value near 0 means the event is not likely to happen.
   - **d.** It is the collection of several experiments.
2. The collection of one or more outcomes from an experiment is called a/an
   - **a.** Event.
   - **b.** Probability.
   - **c.** Random variable.
   - **d.** *z* value.
3. If the occurrence of one event means that another cannot happen, the events are
   - **a.** Independent.
   - **b.** Mutually exclusive.
   - **c.** Bayesian.
   - **d.** Empirical.
4. Under which approach to probability are the outcomes equally likely?
   - **a.** Classical.
   - **b.** Subjective.
   - **c.** Relative frequency.
   - **d.** Independent.
5. To apply the special rule of addition, the events are always
   - **a.** Independent.
   - **b.** Mutually exclusive.
   - **c.** Bayesian.
   - **d.** Empirical.
6. A joint probability is
   - **a.** The likelihood of two events happening.
   - **b.** The likelihood of an event happening given another event has already happened.
   - **c.** Based on two mutually exclusive events.
   - **d.** Also called a prior probability.
7. To apply the special rule of multiplication, the events are always
   - **a.** Independent.
   - **b.** Mutually exclusive.
   - **c.** Bayesian.
   - **d.** Empirical.
8. A table used to classify sample observations according to two criteria is called a:
   - **a.** Probability table.
   - **b.** Contingency table.
   - **c.** Bayesian table.
   - **d.** Scatter diagram
9. A listing of the possible outcomes of an experiment and the corresponding probability is called a
   - **a.** Random variable.
   - **b.** Contingency table.
   - **c.** Probability distribution.
   - **d.** Frequency distribution.
10. Which of the following is *not* an example of a discrete probability distribution?
    - **a.** The purchase price of a house.
    - **b.** The number of bedrooms in a house.
    - **c.** The number of bathrooms in a house.
    - **d.** Whether or not a home has a swimming pool.
11. Which of the following is *not* a condition of the binomial distribution?
    - **a.** Only 2 possible outcomes.
    - **b.** Constant probability of success.
    - **c.** Must have at least 3 trials.
    - **d.** Independent trials.
12. In a Poisson probability distribution
    - **a.** The mean and variance of the distribution are equal.
    - **b.** The probability of a success is always greater than .5.

c. The number of trials is always less than 5.
d. It always contains a contingency table.

13. Which of the following statements is *not* correct regarding the normal probability distribution?
a. It is defined by its mean and standard deviation.
b. The mean and median are equal.
c. It is symmetric.
d. It is based on only two observations.

14. To employ the normal approximation to the binomial
a. The probability of a success should be at least .5.
b. The sample size or the number of trials should be at least 30.
c. The value of $n\pi$ is greater than 5.
d. The outcomes should be mutually exclusive.

15. Using the standard normal probability distribution, what is the likelihood of finding a z value greater than 1.66?
a. .4515 b. .9515 c. .5000 d. .0485

Part II—Problems

16. It is claimed that Proactine, a new medicine for acne, is 80 percent effective—that is, of every 100 persons who apply it, 80 show significant improvement. It is applied to the affected area of a group of 15 people. What is the probability that:
a. All 15 will show significant improvement?
b. Fewer than 9 of 15 will show significant improvement?
c. That 12 or more people will show significant improvement?

17. First National Bank thoroughly investigates its applicants for small home-improvement loans. Its default record is very impressive: The probability that a homeowner will default is only .005. The bank has approved 400 small home-improvement loans. Assuming the Poisson probability distribution applies to this problem:
a. What is the probability that no homeowners out of the 400 will default?
b. How many of the 400 are expected not to default?
c. What is the probability that 3 or more homeowners will default on their small home-improvement loans?

18. A study of the attendance at the University of Alabama's basketball games revealed that the distribution of attendance is normally distributed with a mean of 10,000 and a standard deviation of 2,000.
a. What is the probability a particular game has an attendance of 13,500 or more?
b. What percent of the games have an attendance between 8,000 and 11,500?
c. Ten percent of the games have an attendance of how many or less?

19. A study by the Human Resources Department at the North Ocean Medical Center revealed the following information on the number of absences last month for housekeeping employees.

Days Absent	Number of Employees
0	20
1	35
2	90
3	40
4	10
5 or more	5

What is the probability that an employee selected at random:
a. Was not absent during the month?
b. Was absent less than 3 days?
c. Was absent 4 or more days?

20. The Internal Revenue Service has set aside 200 tax returns where the amount of charitable contributions seemed excessive. A sample of six returns is selected from the group. If two or more of this sampled group have "excessive" amounts deducted for charitable contributions, the entire group will be audited. What is the probability the entire group will be audited, if the true proportion of "excessive" deductions is 20 percent? What if the true proportion is 30 percent?

21. Daniel-James Insurance Company will insure an offshore Mobil Oil production platform against weather losses for one year. The president of Daniel-James estimates the following losses for that platform (in millions of dollars) with the accompanying probabilities:

Amount of Loss ($ millions)	Probability of Loss
0	.98
40	.016
300	.004

a. What is the expected amount Daniel-James will have to pay to Mobil in claims?
b. What is the likelihood that Daniel-James will actually lose less than the expected amount?
c. Given that Daniel-James suffers a loss, what is the likelihood that it is for $300 million?
d. Daniel-James has set the annual premium at $2.0 million. Does that seem like a fair premium?
Will it cover its risk?

22. The distribution of the number of school-age children per family in the Whitehall Estates area of Boise, Idaho, is:

Number of children	0	1	2	3	4
Percent of families	40	30	15	10	5

a. Determine the mean and standard deviation of the number of school-age children per family in Whitehall Estates.
b. A new school is planned in Whitehall Estates. An estimate of the number of school-age children is needed. There are 500 family units. How many children would you estimate?
c. Some additional information is needed about only the families having children. Convert the preceding distribution to one for families with children. What is the mean number of children among families that have children?

23. The following table shows a breakdown of the 108th U.S. Congress by party affiliation.

	Party		
	Democrats	Republicans	Others
House	205	229	1
Senate	48	51	1

a. A member of Congress is selected at random. What is the probability of selecting a Republican?
b. Given that the person selected is a member of the House of Representatives, what is the probability he or she is a Republican?
c. What is the probability of selecting a member of the House of Representatives or a Democrat?

Cases

A. Century National Bank

Refer to the Century National Bank data. Is it reasonable that the distribution of checking account balances approximates a normal probability distribution? Determine the mean and the standard deviation for the sample of 60 customers. Compare the actual distribution with the theoretical distribution. Cite some specific examples and comment on your findings.

Divide the account balances into three groups, of about 20 each, with the smallest third of the balances in the first group, the middle third in the second group, and those with the largest balances in the third group. Next, develop a table that shows the number in each of the categories of the account balances by branch. Does it appear that account balances are related to the branch? Cite some examples and comment on your findings.

B. Elections Auditor

An item such as an increase in taxes, recall of elected officials, or an expansion of public services can be placed on the ballot if a required number of valid signatures are collected on the petition. Unfortunately, many people will sign the petition even though they are not registered to vote in that particular district, or they will sign the petition more than once.

Sara Ferguson, the elections auditor in Venango County, must certify the validity of these signatures after the petition is officially presented. Not surprisingly, her staff is overloaded, so she is considering using statistical methods to validate the pages of 200 signatures, instead of validating each individual signature. At a recent professional meeting, she found that, in some communities in the state, election officials were checking only five signatures on each page and rejecting the entire page if two or more signatures were invalid. Some people are concerned that five may not be enough to make a good decision. They suggest that you should check 10 signatures and reject the page if three or more are invalid.

In order to investigate these methods, Sara asks her staff to pull the results from the last election and sample 30 pages. It happens that the staff selected 14 pages from the Avondale district, 9 pages from the Midway district, and 7 pages from the Kingston district. Each page had 200 signatures, and the data below show the number of invalid signatures on each.

Use the data to evaluate Sara's two proposals. Calculate the probability of rejecting a page under each of the approaches. Would you get about the same results by examining every single signature? Offer a plan of your own, and discuss how it might be better or worse than the two plans proposed by Sara.

Avondale	Midway	Kingston
9	19	38
14	22	39
11	23	41
8	14	39
14	22	41
6	17	39
10	15	39
13	20	
8	18	
8		
9		
12		
7		
13		

C. Geoff "Applies" His Education

Geoff Brown is the manager for a small telemarketing firm and is evaluating the sales rate of experienced workers in order to set minimum standards for new hires. During the past few weeks, he has recorded the number of successful calls per hour for the staff. These data appear below along with some summary statistics he worked out with a statistical software package. Geoff has been a student at the local community college and has heard of many different kinds of probability distributions (binomial, normal, hypergeometric, Poisson, etc.). Could you give Geoff some advice on which distribution to use to fit these data as well as possible and how to decide when a probationary employee should be accepted as having reached full production status? This is important because it means a pay raise for the employee, and there have been some probationary employees in the past who have quit because of discouragement that they would never meet the standard.

Successful sales calls per hour during the week of August 14:

4	2	3	1	4	5	5	2	3	2	2	4	5	2	5	3	3	0
1	3	2	8	4	5	2	2	4	1	5	5	4	5	1	2	4	

Descriptive statistics:

N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
35	3.229	3.000	3.194	1.682	0.284
MIN	MAX	Q1	Q3		
0.0	8.000	2.000	5.000		

Which distribution do you think Geoff should use for his analysis?

D. CNP Bank Card

Before banks issue a credit card, they usually rate or score the customer in terms of his or her projected probability of being a profitable customer. A typical scoring table appears below.

Age	**Under 25** (12 pts.)	**25–29** (5 pts.)	**30–34** (0 pts.)	**35+** (18 pts.)
Time at same address	**<1 yr.** (9 pts.)	**1–2 yrs.** (0 pts.)	**3–4 yrs.** (13 pts.)	**5+ yrs.** (20 pts.)
Auto age	**None** (18 pts.)	**0–1yr.** (12 pts.)	**2–4 yrs.** (13 pts.)	**5+ yrs.** (3 pts.)
Monthly car payment	**None** (15 pts.)	**\$1–\$99** (6 pts.)	**\$100–\$299** (4 pts.)	**\$300+** (0 pts.)
Housing cost	**\$1–\$199** (0 pts.)	**\$200–\$399** (10 pts.)	**Owns** (12 pts.)	**Lives with relatives** (24 pts.)
Checking/ savings accounts	**Both** (15 pts.)	**Checking only** (3 pts.)	**Savings only** (2 pts.)	**Neither** (0 pts.)

The score is the sum of the points on the six items. For example, Sushi Brown is under 25 years old (12 pts.), has lived at the same address for 2 years (0 pts.), owns a 4-year-old car (13 pts.), with car payments of \$75 (6 pts.), housing cost of $200 (10 pts.), and a checking account (3 pts.). She would score 44.

A second chart is then used to convert scores into the probability of being a profitable customer. A sample chart of this type appears below.

Score	30	40	50	60	70	80	90
Probability	.70	.78	.85	.90	.94	.95	.96

Sushi's score of 44 would translate into a probability of being profitable of approximately .81. In other words 81 percent of customers like Sushi will make money for the bank card operations.

Here are the interview results for three potential customers.

Name	David Born	Edward Brendan	Ann McLaughlin
Age	42	23	33
Time at same address	9	2	5
Auto age	2	3	7
Monthly car payment	$140	$99	$175
Housing cost	$300	$200	Owns clear
Checking/savings accounts	Both	Checking only	Neither

1. Score each of these customers and estimate their probability of being profitable.
2. What is the probability that all three are profitable?
3. What is the probability that none of them are profitable?
4. Find the entire probability distribution for the number of profitable customers among this group of three.