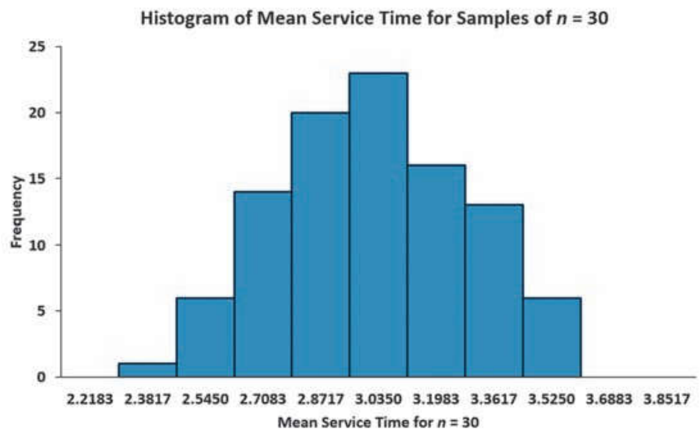
FIGURE 7.6
(continued)

Panel C: Histogram of the mean service time (in minutes) at the fast-food chain drive-through lane of 100 different random samples of $n = 30$

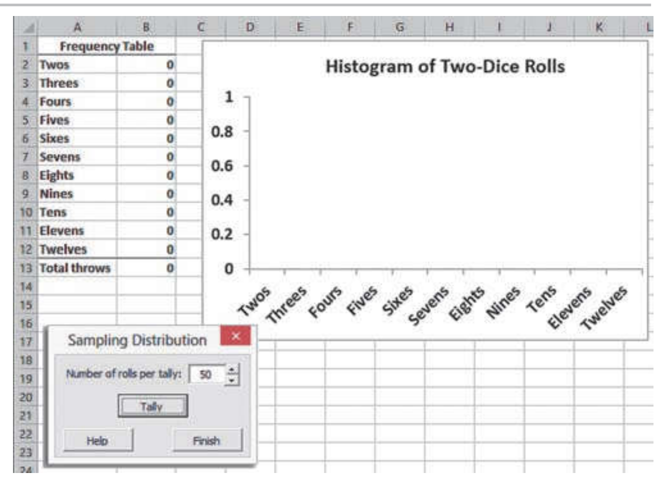


random samples of $n = 30$, shows a distribution that appears to be approximately bell-shaped with a concentration of values in the center of the distribution. The progression of the histograms from a skewed population towards a bell-shaped distribution as the sample size increases is consistent with the Central Limit Theorem.

VISUAL EXPLORATIONS Exploring Sampling Distributions

Open the **VE-Sampling Distribution add-in workbook** to observe the effects of simulated rolls on the frequency distribution of the sum of two dice. (For Excel technical requirements, review Appendix Section D.4) When this workbook opens properly, it adds a Sampling Distribution menu to the Add-ins tab (Apple menu in Excel 2011).

To observe the effects of simulated throws on the frequency distribution of the sum of the two dice, select **Sampling Distribution ➔ Two Dice Simulation**. In the Sampling Distribution dialog box, enter the **Number of rolls per tally** and click **Tally**. Click **Finish** when done.



Problems for Section 7.2

LEARNING THE BASICS

7.1 Given a normal distribution with $\mu = 100$ and $\sigma = 10$, if you select a sample of $n = 25$, what is the probability that $\overline{X}$ is

a. less than 95?
b. between 95 and 97.5?
c. above 102.2?
d. There is a 65% chance that $\overline{X}$ is above what value?

7.2 Given a normal distribution with $\mu = 50$ and $\sigma = 5$, if you select a sample of $n = 100$, what is the probability that $\overline{X}$ is

a. less than 47?
b. between 47 and 49.5?
c. above 51.1?
d. There is a 35% chance that $\overline{X}$ is above what value?

APPLYING THE CONCEPTS

7.3 For each of the following three populations, indicate what the sampling distribution for samples of 25 would consist of:

a. Customer receipts for a supermarket for a year.
b. Insurance payouts in a particular geographical area in a year.
c. Call center logs of inbound calls tracking handling time for a credit card company during the year.

7.4 The following data represent the number of days absent per year in a population of six employees of a small company:

1 3 6 7 9 10

a. Assuming that you sample without replacement, select all possible samples of $n = 2$ and construct the sampling distribution of the mean. Compute the mean of all the sample means and also compute the population mean. Are they equal? What is this property called?
b. Repeat (a) for all possible samples of $n = 3$.
c. Compare the shape of the sampling distribution of the mean in (a) and (b). Which sampling distribution has less variability? Why?
d. Assuming that you sample with replacement, repeat (a) through (c) and compare the results. Which sampling distributions have the least variability—those in (a) or (b)? Why?

7.5 The diameter of a brand of tennis balls is approximately normally distributed, with a mean of 2.63 inches and a standard deviation of 0.03 inch. If you select a random sample of 9 tennis balls,

a. what is the sampling distribution of the mean?
b. what is the probability that the sample mean is less than 2.61 inches?
c. what is the probability that the sample mean is between 2.62 and 2.64 inches?
d. The probability is 60% that the sample mean will be between what two values symmetrically distributed around the population mean?

7.6 The U.S. Census Bureau announced that the median sales price of new houses sold in 2012 was \$245,200, and the mean sales price was \$291,200 (**www.census.gov/newhomesales**, April 1, 2013). Assume that the standard deviation of the prices is \$90,000.

a. If you select samples of $n = 4$, describe the shape of the sampling distribution of $\overline{X}$.
b. If you select samples of $n = 100$, describe the shape of the sampling distribution of $\overline{X}$.
c. If you select a random sample of $n = 100$, what is the probability that the sample mean will be less than \$315,000?
d. If you select a random sample of $n = 100$, what is the probability that the sample mean will be between \$295,000 and \$310,000?

7.7 According to a mashable.com post, time spent on Tumblr, a microblogging platform and social networking website, has a mean of 14 minutes per visit. (Source: **on.mash.to/1757wfE**.) Assume that time spent on Tumblr per visit is normally distributed and that the standard deviation is 4 minutes. If you select a random sample of 25 visits,

a. what is the probability that the sample mean is between 13.6 and 14.4 minutes?
b. what is the probability that the sample mean is between 13 and 14 minutes?
c. If you select a random sample of 100 visits, what is the probability that the sample mean is between 13.6 and 14.4 minutes?
d. Explain the difference in the results of (a) and (c).

SELF Test **7.8** Today, full-time college students report spending a mean of 27 hours per week on academic activities, both inside and outside the classroom. (Source: "A Challenge to College Students for 2013: Don't Waste Your 6,570," *Huffington Post*, January 29, 2013, **huff.to/13dNtuT**.) Assume the standard deviation of time spent on academic activities is 4 hours. If you select a random sample of 16 full-time college students,

a. what is the probability that the mean time spent on academic activities is at least 26 hours per week?
b. there is an 85% chance that the sample mean is less than how many hours per week?
c. What assumption must you make in order to solve (a) and (b)?
d. If you select a random sample of 64 full-time college students, there is an 85% chance that the sample mean is less than how many hours per week?

7.3 Sampling Distribution of the Proportion

Student Tip
Do not confuse this use of the Greek letter pi, π, to represent the population proportion with the mathematical constant that is the ratio of the circumference to a diameter of a circle—approximately 3.14159—which is also known by the same Greek letter.

Consider a categorical variable that has only two categories, such as the customer prefers your brand or the customer prefers the competitor's brand. You are interested in the proportion of items belonging to one of the categories—for example, the proportion of customers that prefer your brand. The population proportion, represented by π, is the proportion of items in the entire population with the characteristic of interest. The sample proportion, represented by p, is the proportion of items in the sample with the characteristic of interest. The sample proportion, a statistic, is used to estimate the population proportion, a parameter. To calculate the sample proportion, you assign one of two possible values, 1 or 0, to represent the presence or absence of the characteristic. You then sum all the 1 and 0 values and divide by n, the sample size. For example, if, in a sample of five customers, three preferred your brand and two did not, you have three 1s and two 0s. Summing the three 1s and two 0s and dividing by the sample size of 5 results in a sample proportion of 0.60.

the survey percentage of 32% as the population proportion, you can calculate the probability that more than 40% of the vacationers are unable to stop thinking about work while on vacation by using Equation (7.8):

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

$$= \frac{0.40 - 0.32}{\sqrt{\frac{(0.32)(0.68)}{200}}} = \frac{0.08}{\sqrt{\frac{0.2176}{200}}} = \frac{0.08}{0.0330}$$

$$= 2.42$$

Using Table E.2, the area under the normal curve greater than 2.42 is 0.0078. Therefore, if the population proportion is 0.32, the probability is 0.78% that more than 40% of the 200 vacationers in the sample will be unable to stop thinking about work while on vacation.

Problems for Section 7.3

LEARNING THE BASICS

7.9 In a random sample of 64 people, 48 are classified as "successful."

a. Determine the sample proportion, p, of "successful" people.
b. If the population proportion is 0.70, determine the standard error of the proportion.

7.10 A random sample of 50 households was selected for a phone (landline and cellphone) survey. The key question asked was, "Do you or any member of your household own an Apple product (iPhone, iPod, iPad, or Mac computer)?" Of the 50 respondents, 20 said yes and 30 said no.

a. Determine the sample proportion, p, of households that own an Apple product.
b. If the population proportion is 0.45, determine the standard error of the proportion.

7.11 The following data represent the responses (Y for yes and N for no) from a sample of 40 college students to the question "Do you currently own shares in any stocks?"

N N Y N N Y N Y N Y N N Y N Y Y N N N Y
N Y N N N N Y N N Y Y N N N Y N N Y N N

a. Determine the sample proportion, p, of college students who own shares of stock.
b. If the population proportion is 0.30, determine the standard error of the proportion.

APPLYING THE CONCEPTS

SELF Test **7.12** A political pollster is conducting an analysis of sample results in order to make predictions on election night. Assuming a two-candidate election, if a specific candidate receives at least 55% of the vote in the sample, that candidate will be forecast as the winner of the election. If you select a random sample of 100 voters, what is the probability that a candidate will be forecast as the winner when

a. the population percentage of her vote is 50.1%?
b. the population percentage of her vote is 60%?
c. the population percentage of her vote is 49% (and she will actually lose the election)?
d. If the sample size is increased to 400, what are your answers to (a) through (c)? Discuss.

7.13 You plan to conduct a marketing experiment in which students are to taste one of two different brands of soft drink. Their task is to correctly identify the brand tasted. You select a random sample of 200 students and assume that the students have no ability to distinguish between the two brands. (Hint: If an individual has no ability to distinguish between the two soft drinks, then the two brands are equally likely to be selected.)

a. What is the probability that the sample will have between 50% and 60% of the identifications correct?
b. The probability is 90% that the sample percentage is contained within what symmetrical limits of the population percentage?
c. What is the probability that the sample percentage of correct identifications is greater than 65%?
d. Which is more likely to occur—more than 60% correct identifications in the sample of 200 or more than 55% correct identifications in a sample of 1,000? Explain.

7.14 Accenture's *Defining Success* global research study found that the majority of today's working women would prefer a better work–life balance to an increased salary. One of the most important contributors to work–life balance identified by the survey was "flexibility," with 80% of women saying that having a flexible work schedule is either very important or extremely important to their career success. (Source: **bit.ly/17IM8gq**.) Suppose you select a sample of 100 working women.

a. What is the probability that in the sample fewer than 85% say that having a flexible work schedule is either very important or extremely important to their career success?
b. What is the probability that in the sample between 75% and 85% say that having a flexible work schedule is either very important or extremely important to their career success?

c. What is the probability that in the sample more than 82% say that having a flexible work schedule is either very important or extremely important to their career success?
d. If a sample of 400 is taken, how does this change your answers to (a) through (c)?

7.15 The goal of corporate sustainability is to manage the environmental, economic, and social effects of a corporation's operations so it is profitable over the long term while acting in a responsible manner toward society. A Hill + Knowlton Strategies survey found that 57% of U.S. respondents are more likely to buy stock in a U.S. corporation, or shop at its stores, if it is making an effort to publicly talk about how it is becoming more sustainable. (Source: "Sustainability," **bit.ly/10A2Snl**.) Suppose you select a sample of 100 U.S. respondents.
a. What is the probability that in the sample, fewer than 57% are more likely to buy stock in a U.S. corporation, or shop at its stores, if it is making an effort to publicly talk about how it is becoming more sustainable?
b. What is the probability that in the sample, between 52% and 62% are more likely to buy stock in a U.S. corporation, or shop at its stores, if it is making an effort to publicly talk about how it is becoming more sustainable?
c. What is the probability that in the sample, more than 62% are more likely to buy stock in a U.S. corporation, or shop at its stores, if it is making an effort to publicly talk about how it is becoming more sustainable?
d. If a sample of 400 is taken, how does this change your answers to (a) through (c)?

7.16 According to *GMI Ratings' 2013 Women on Boards Report*, the percentage of women on U.S. boards has increased marginally in 2009–2012 and now stands at 14%, well below the values for Nordic countries and France. A number of initiatives are underway in an effort to increase the representation. For example, a network of investors, corporate leaders, and other advocates, known as the 30% coalition, is seeking to raise the proportion of female directors to that number (30%) by 2015. This study also reports that 15% of U.S. companies have three or more female board directors. (Data extracted from **bit.ly/13oSFem**.) If you select a random sample of 200 U.S. companies,
a. what is the probability that the sample will have between 12% and 18% U.S. companies that have three or more female board directors?
b. the probability is 90% that the sample percentage of U.S. companies having three or more female board directors will be contained within what symmetrical limits of the population percentage?
c. the probability is 95% that the sample percentage of U.S. companies having three or more female board directors will be contained within what symmetrical limits of the population percentage?

7.17 The Chartered Financial Analyst (CFA) institute reported that 51% of its U.S. members indicate that lack of ethical culture within financial firms has contributed the most to the lack of trust in the financial industry. (Source: Data extracted from *Global Market Sentiment Survey 2013*, **cfa.is/YqVCKB**.) Suppose that you select a sample of 100 CFA members.
a. What is the probability that the sample percentage indicating that lack of ethical culture within financial firms has contributed the most to the lack of trust in the financial industry will be between 50% and 55%?
b. The probability is 90% that the sample percentage will be contained within what symmetrical limits of the population percentage?
c. The probability is 95% that the sample percentage will be contained within what symmetrical limits of the population percentage?
d. Suppose you selected a sample of 400 CFA members. How does this change your answers in (a) through (c)?

7.18 A Pew Research Center project on the state of news media showed that the clearest pattern of news audience growth in 2012 came on digital platforms. According to Pew Research data, 39% of Americans get news online or from a mobile device in a typical day. (Data extracted from "Key Findings: State of the News Media," Pew Research Center, **bit.ly/10kKUTi**.)
a. Suppose that you take a sample of 100 Americans. If the population proportion of Americans who get news online or from a mobile device in a typical day is 0.39, what is the probability that fewer than 30% in your sample will get news online or from a mobile device in a typical day?
b. Suppose that you take a sample of 400 Americans. If the population proportion of Americans who get news online or from a mobile device in a typical day is 0.39, what is the probability that fewer than 30% in your sample will get news online or from a mobile device in a typical day?
c. Discuss the effect of sample size on the sampling distribution of the proportion in general and the effect on the probabilities in (a) and (b).

7.4 Sampling from Finite Populations

The Central Limit Theorem and the standard errors of the mean and of the proportion are based on samples selected with replacement. However, in nearly all survey research, you sample *without* replacement from populations that are of a finite size, N. The **Section 7.4 online topic** explains how you use a **finite population correction factor** to compute the standard error of the mean and the standard error of the proportion for such samples.

USING STATISTICS

Sampling Oxford Cereals, Revisited

Corbis

As the plant operations manager for Oxford Cereals, you were responsible for monitoring the amount of cereal placed in each box. To be consistent with package labeling, boxes should contain a mean of 368 grams of cereal. Thousands of boxes are produced during a shift, and weighing every single box was determined to be too time-consuming, costly, and inefficient. Instead, a sample of boxes was selected. Based on your analysis of the sample, you had to decide whether to maintain, alter, or shut down the process.

Using the concept of the sampling distribution of the mean, you were able to determine probabilities that such a sample mean could have been randomly selected from a population with a mean of 368 grams. Specifically, if a sample of size $n = 25$ is selected from a population with a mean of 368 and standard deviation of 15, you calculated the probability of selecting a sample with a mean of 365 grams or less to be 15.87%. If a larger sample size is selected, the sample mean should be closer to the population mean. This result was illustrated when you calculated the probability if the sample size were increased to $n = 100$. Using the larger sample size, you determined the probability of selecting a sample with a mean of 365 grams or less to be 2.28%.

SUMMARY

You studied the sampling distribution of the sample mean and the sampling distribution of the sample proportion and their relationship to the Central Limit Theorem. You learned that the sample mean is an unbiased estimator of the population mean, and the sample proportion is an unbiased estimator of the population proportion. In the next five chapters, the techniques of confidence intervals and tests of hypotheses commonly used for statistical inference are discussed.

REFERENCES

1. Cochran, W. G. *Sampling Techniques*, 3rd ed. New York: Wiley, 1977.
2. *Microsoft Excel 2013*. Redmond, WA: Microsoft Corp., 2012.
3. *Minitab Release 16*. State College, PA: Minitab, Inc., 2010.

KEY EQUATIONS

Population Mean

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} \tag{7.1}$$

Population Standard Deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}} \tag{7.2}$$

Standard Error of the Mean

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \tag{7.3}$$

Finding Z for the Sampling Distribution of the Mean

$$Z = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} \tag{7.4}$$

Finding $\overline{X}$ for the Sampling Distribution of the Mean

$$\overline{X} = \mu + Z\frac{\sigma}{\sqrt{n}} \tag{7.5}$$

Sample Proportion

$$p = \frac{X}{n} \tag{7.6}$$

Standard Error of the Proportion

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \tag{7.7}$$

Finding Z for the Sampling Distribution of the Proportion

$$Z = \frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \tag{7.8}$$

KEY TERMS

Central Limit Theorem 257
sampling distribution 251
sampling distribution of the mean 251
sampling distribution of the proportion 263
standard error of the mean 253
standard error of the proportion 263
unbiased 251

CHECKING YOUR UNDERSTANDING

7.19 Why is the sample mean an unbiased estimator of the population mean?

7.20 Why does the standard error of the mean decrease as the sample size, n, increases?

7.21 Why does the sampling distribution of the mean follow a normal distribution for a large enough sample size, even though the population may not be normally distributed?

7.22 What is the difference between a population distribution and a sampling distribution?

7.23 Under what circumstances does the sampling distribution of the proportion approximately follow the normal distribution?

CHAPTER REVIEW PROBLEMS

7.24 An industrial sewing machine uses ball bearings that are targeted to have a diameter of 0.75 inch. The lower and upper specification limits under which the ball bearing can operate are 0.74 inch (lower) and 0.76 inch (upper). Past experience has indicated that the actual diameter of the ball bearings is approximately normally distributed, with a mean of 0.753 inch and a standard deviation of 0.004 inch. If you select a random sample of 25 ball bearings, what is the probability that the sample mean is

a. between the target and the population mean of 0.753?
b. between the lower specification limit and the target?
c. greater than the upper specification limit?
d. less than the lower specification limit?
e. The probability is 93% that the sample mean diameter will be greater than what value?

7.25 The fill amount of bottles of a soft drink is normally distributed, with a mean of 2.0 liters and a standard deviation of 0.05 liter. If you select a random sample of 25 bottles, what is the probability that the sample mean will be

a. between 1.99 and 2.0 liters?
b. below 1.98 liters?
c. greater than 2.01 liters?
d. The probability is 99% that the sample mean amount of soft drink will be at least how much?
e. The probability is 99% that the sample mean amount of soft drink will be between which two values (symmetrically distributed around the mean)?

7.26 An orange juice producer buys oranges from a large orange grove that has one variety of orange. The amount of juice squeezed from these oranges is approximately normally distributed, with a mean of 4.70 ounces and a standard deviation of 0.40 ounce. Suppose that you select a sample of 25 oranges.

a. What is the probability that the sample mean amount of juice will be at least 4.60 ounces?
b. The probability is 70% that the sample mean amount of juice will be contained between what two values symmetrically distributed around the population mean?
c. The probability is 77% that the sample mean amount of juice will be greater than what value?

7.27 In Problem 7.26, suppose that the mean amount of juice squeezed is 5.0 ounces.

a. What is the probability that the sample mean amount of juice will be at least 4.60 ounces?
b. The probability is 70% that the sample mean amount of juice will be contained between what two values symmetrically distributed around the population mean?
c. The probability is 77% that the sample mean amount of juice will be greater than what value?

7.28 The stock market in Mexico reported strong returns in 2012. The population of stocks earned a mean return of 17.87% in 2012. (Data extracted from **asxiq.com/blog/stock-markets-in-the-world-returns-in-2012/**.) Assume that the returns for stocks on the Mexican stock market were distributed as a normal random variable, with a mean of 17.87 and a standard deviation of 20. If you selected a random sample of 16 stocks from this population, what is the probability that the sample would have a mean return

a. less than 0 (i.e., a loss)?
b. between -10 and 10?
c. greater than 10?

7.29 The article mentioned in Problem 7.28 reported that the stock market in China had a mean return of 1.54% in 2012. Assume that the returns for stocks on the Chinese stock market were distributed as a normal random variable, with a mean of 1.54 and a standard deviation of 10. If you select an individual stock from this population, what is the probability that it would have a return

a. less than 0 (i.e., a loss)?
b. between -10 and -20?
c. greater than -5?

If you selected a random sample of four stocks from this population, what is the probability that the sample would have a mean return

d. less than 0 (a loss)?
e. between -10 and -20?
f. greater than -5?
g. Compare your results in parts (d) through (f) to those in (a) through (c).

7.30 **(Class Project)** The table of random numbers is an example of a uniform distribution because each digit is equally likely to occur. Starting in the row corresponding to the day of the month in which you were born, use a table of random numbers (Table E.1) to take one digit at a time.

Select five different samples each of $n = 2, n = 5$, and $n = 10$. Compute the sample mean of each sample. Develop a frequency distribution of the sample means for the results of the entire class, based on samples of sizes $n = 2, n = 5$, and $n = 10$.

What can be said about the shape of the sampling distribution for each of these sample sizes?

7.31 **(Class Project)** Toss a coin 10 times and record the number of heads. If each student performs this experiment five times, a frequency distribution of the number of heads can be developed from the results of the entire class. Does this distribution seem to approximate the normal distribution?

7.32 **(Class Project)** The number of cars waiting in line at a car wash is distributed as follows:

Number of Cars	Probability
0	0.25
1	0.40
2	0.20
3	0.10
4	0.04
5	0.01

You can use a table of random numbers (Table E.1) to select samples from this distribution by assigning numbers as follows:

1. Start in the row corresponding to the day of the month in which you were born.
2. Select a two-digit random number.
3. If you select a random number from 00 to 24, record a length of 0; if from 25 to 64, record a length of 1; if from 65 to 84, record a length of 2; if from 85 to 94, record a length of 3; if from 95 to 98, record a length of 4; if 99, record a length of 5.

Select samples of $n = 2, n = 5$, and $n = 10$. Compute the mean for each sample. For example, if a sample of size 2 results in the random numbers 18 and 46, these would correspond to lengths 0 and 1, respectively, producing a sample mean of 0.5. If each student selects five different samples for each sample size, a frequency distribution of the sample means (for each sample size) can be developed from the results of the entire class. What conclusions can you reach concerning the sampling distribution of the mean as the sample size is increased?

7.33 **(Class Project)** Using a table of random numbers (Table E.1), simulate the selection of different-colored balls from a bowl, as follows:

1. Start in the row corresponding to the day of the month in which you were born.
2. Select one-digit numbers.
3. If a random digit between 0 and 6 is selected, consider the ball white; if a random digit is a 7, 8, or 9, consider the ball red.

Select samples of $n = 10, n = 25$, and $n = 50$ digits. In each sample, count the number of white balls and compute the proportion of white balls in the sample. If each student in the class selects five different samples for each sample size, a frequency distribution of the proportion of white balls (for each sample size) can be developed from the results of the entire class. What conclusions can you reach about the sampling distribution of the proportion as the sample size is increased?

7.34 **(Class Project)** Suppose that step 3 of Problem 7.33 uses the following rule: "If a random digit between 0 and 8 is selected, consider the ball to be white; if a random digit of 9 is selected, consider the ball to be red." Compare and contrast the results in this problem and those in Problem 7.33.