8

Sampling Distributions

Outline

Making an Informed Decision

The American Time Use Survey, conducted by the Bureau of Labor Statistics, investigates how adult Americans allocate their time during a day. As a reporter for the school newspaper, you wish to file a report that compares the typical student at your school to other Americans. See the Decisions project on page 446.

PUTTING IT TOGETHER

In Chapters 6 and 7, we learned about random variables and their probability distributions. A random variable is a numerical measure of the outcome to a probability experiment. A probability distribution provides a way to assign probabilities to the possible values of the random variable. For discrete random variables, we discussed the binomial probability distribution and the Poisson probability distribution. We assigned probabilities using a formula. For continuous random variables, we discussed the normal probability distribution. To compute probabilities for a normal random variable, we found the area under a normal density curve.

In this chapter, we continue our discussion of probability distributions where statistics, such as $\bar{x}$, will be the random variable. Statistics are random variables because the value of a statistic varies from sample to sample. For this reason, statistics have probability distributions associated with them. For example, there is a probability distribution for the sample mean, sample proportion, and so on. We use probability distributions to make probability statements regarding the statistic. So this chapter discusses the shape, center, and spread of statistics such as $\bar{x}$.

Preparing for This Section Before getting started, review the following:

- Simple random sampling (Section 1.3, pp. 50–53)
- The mean (Section 3.1, pp. 146–148)
- The standard deviation (Section 3.2, pp. 161–165)
- Applications of the normal distribution (Section 7.2, pp. 394–400)

Objectives

1. Describe the distribution of the sample mean: normal population
2. Describe the distribution of the sample mean: nonnormal population

Suppose the government wanted to determine the mean income of all U.S. households. One approach the government could take is to survey every U.S. household to determine the population mean, μ. This would be a very expensive and time-consuming survey!

A second approach the government could (and does) take is to survey a random sample of U.S. households and use the results to estimate the mean household income. The Current Population Survey is administered to approximately 250,000 randomly selected households each month. Among the many questions on the survey, respondents are asked to report the income of each individual in the household. From this information, the federal government obtains a sample mean household income for U.S. households. For example, in 2013 the mean annual household income in the United States was estimated to be $\bar{x} = \$72{,}641$. The government might infer from this survey that the mean annual household income of *all* U.S. households in 2013 was $\mu = \$72{,}641$.

The households in the Current Population Survey were determined by chance (random sampling). A second random sample of households would likely lead to a different sample mean, such as $\bar{x} = \$71{,}849$, and a third random sample of households would likely lead to a third sample mean, such as $\bar{x} = \$72{,}978$. Because the households selected will vary from sample to sample, the sample mean of household income will also vary from sample to sample. For this reason, the sample mean $\bar{x}$ is a random variable, so it has a probability distribution. Our goal in this section is to describe the distribution of the sample mean. Remember, when we describe a distribution, we do so in terms of its shape, center, and spread.

Definitions

The **sampling distribution** of a statistic is a probability distribution for all possible values of the statistic computed from a sample of size n.

The **sampling distribution of the sample mean** $\bar{x}$ is the probability distribution of all possible values of the random variable $\bar{x}$ computed from a sample of size n from a population with mean μ and standard deviation σ.

The idea behind obtaining the sampling distribution of the sample mean is as follows:

Step 1: Obtain a simple random sample of size n.

Step 2: Compute the sample mean.

Step 3: Assuming that we are sampling from a finite population, repeat Steps 1 and 2 until all distinct simple random samples of size n have been obtained.

Note: Once a particular sample is obtained, it cannot be obtained a second time.

In Other Words
If the number of individuals in a population is a positive integer, we say the population is finite. Otherwise, the population is infinite.

1 Describe the Distribution of the Sample Mean: Normal Population

The probability distribution of the sample mean is determined from statistical theory. We will use simulation to help justify the result that statistical theory provides. We consider two possibilities. In the first case (Examples 1, 2, and 3), we sample from a population

that is normally distributed. In the second case (Examples 4 and 5), we sample from a population that is not normally distributed.

EXAMPLE 1 Sampling Distribution of the Sample Mean: Normal Population

Problem An intelligence quotient, or IQ, is a measurement of intelligence derived from a standardized test, such as the Stanford Binet IQ test. Scores on this test are approximately normally distributed with a mean score of 100 and a standard deviation of 15. What is the sampling distribution of the sample mean for a sample of size $n = 9$?

Approach The problem asks us to determine the shape, center, and spread of the distribution of the sample mean. Remember, the sampling distribution of the sample mean would be the distribution of *all* possible sample means of size $n = 9$. To get a sense of this distribution, use Minitab to simulate obtaining 1000 samples of size $n = 9$ by randomly generating 1000 rows of IQs over 9 columns. Each row represents a random sample of size 9. For each of the 1000 samples (the 1000 rows), we determine the mean IQ score. Draw a histogram to gauge the shape of the distribution of the sample mean, determine the mean of the 1000 sample means to approximate the mean of the sampling distribution, and determine the standard deviation of the 1000 sample means to approximate the standard deviation of the sampling distribution.

Using Technology

We are using Minitab's Random Data command under the Calc menu to generate these data. Select Normal ... from the Random Data menu.

To obtain normal random data using StatCrunch, select

Data > Simulate > Normal

Solution Figure 1 shows random samples from Minitab. Row 1 contains the first sample, where the IQ scores of the nine individuals are 90, 74, 95, 88, 91, 91, 102, 91, and 96. The mean of these nine IQ scores is 90.9. Row 2 represents a second sample with nine different IQ scores; row 3 represents a third sample, and so on. Column C10 (xbar) lists the sample means for each of the different samples.

Figure 1

Worksheet 1 ***

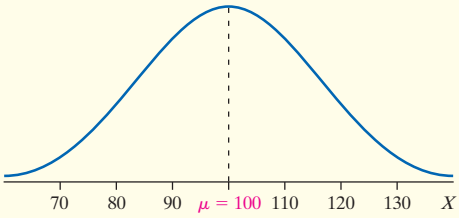
+	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
										xbar
1	90	74	95	88	91	91	102	91	96	90.9
2	114	86	96	82	80	68	93	136	111	96.2
3	89	89	86	98	96	96	89	99	107	94.3
4	87	94	89	116	92	124	115	83	111	101.3
5	107	103	86	86	109	104	94	82	110	97.8
6	113	84	101	89	92	71	89	86	108	92.7
7	75	116	107	118	112	86	104	97	106	102.4
8	118	117	81	96	86	94	109	96	104	100.1
9	91	91	79	80	109	89	97	81	99	90.6
10	112	98	115	73	95	104	76	95	87	95.0
11	103	92	100	75	95	108	105	125	62	96.1
12	75	124	101	113	91	85	121	115	85	101.2
13	104	103	81	134	111	108	101	88	115	105.1
14	121	86	118	88	96	84	103	77	102	97.1
15	119	84	119	80	99	98	88	88	89	95.9
16	71	86	94	101	95	84	124	105	83	93.7

Sample 1 — $\bar{x} = 90.9$

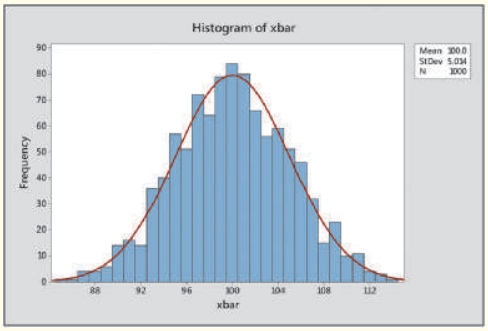
Sample 2 — $\bar{x} = 96.2$

Figure 2(a) shows the distribution of the population, and Figure 2(b) shows the distribution of the sample means from column C10 (using Minitab). The shape of the distribution of the population is normal. The histogram in Figure 2(b) shows that the shape of the distribution of the sample means is also normal. In addition, we notice that the center of the distribution of the sample means is the same as the center of the distribution of the population, but the spread of the distribution of the sample means is smaller than the spread of the distribution of the population. In fact, the mean of the 1000 sample means is 100.02, which is close to the population mean, 100; the standard deviation of the sample means is 5.01, which is less than the population standard deviation, 15.

Figure 2



(a)



(b)

We draw the following conclusions:

- **Shape:** The shape of the distribution of the sample mean is normal.
- **Center:** The mean of the distribution of the sample mean equals the mean of the population, 100.
- **Spread:** The standard deviation of the sample mean is less than the standard deviation of the population. ●

Why is the standard deviation of the sample mean less than the standard deviation of the population? Consider that, if we randomly select any one individual, according to the Empirical Rule, there is about a 68% chance that the individual's IQ score is between 85 and 115 (that is, within 1 standard deviation of the mean). If we had a sample of 9 individuals, we would not expect as much spread in the sample mean as there is for a single individual, since individuals with lower IQs will offset individuals in the sample with higher IQs, resulting in a sample mean closer to the expected value of 100. Look back at Figure 1. In the first sample (row 1), the low-IQ individual $(\text{IQ} = 74)$ is offset by the higher-IQ individual $(\text{IQ} = 102)$, which is why the sample mean is closer to 100. In the second sample (row 2), the low-IQ individual $(\text{IQ} = 68)$ is offset by the higher-IQ individual $(\text{IQ} = 136)$, so the sample mean of the second sample is closer to 100. Therefore, the spread in the distribution of sample means should be less than the spread in the population from which the sample is drawn.

Based on this, what role do you think n, the sample size, plays in the standard deviation of the distribution of the sample mean?

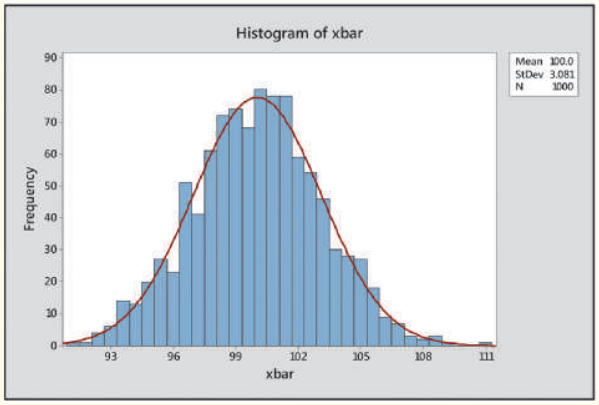
EXAMPLE 2 The Impact of Sample Size on Sampling Variability

Problem Repeat the problem in Example 1 with a sample of size $n = 25$.

Approach Use the approach presented in Example 1, but let $n = 25$ instead of $n = 9$.

Solution Figure 3 shows the histogram of the sample means. Notice that the sample means appear to be normally distributed with the center at 100. The histogram in Figure 3 shows less dispersion than the histogram in Figure 2(b). This implies that the distribution of $\bar{x}$ with $n = 25$ has less variability than the distribution of $\bar{x}$ with $n = 9$. In fact, the mean of the 1000 sample means is 100.05, and the standard deviation is 3.08.

Figure 3



●

From the results of Examples 1 and 2, we conclude that, as the sample size n increases, the standard deviation of the distribution of $\bar{x}$ decreases. Although the proof is beyond the scope of this text, we should be convinced that the following result is reasonable.

In Other Words
Regardless of the distribution of the population, the sampling distribution of $\bar{x}$ will have a mean equal to the mean of the population and a standard deviation equal to the standard deviation of the population divided by the square root of the sample size!

The Mean and Standard Deviation of the Sampling Distribution of $\bar{x}$

Suppose that a simple random sample of size n is drawn from a population* with mean μ and standard deviation σ. The sampling distribution of $\bar{x}$ has mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$. The standard deviation of the sampling distribution of $\bar{x}$, $\sigma_{\bar{x}}$, is called the **standard error of the mean**.

CAUTION!
It is important that two assumptions are satisfied with regard to sampling from a population.
1. The sample must be a random sample.
2. The sampled values must be independent. When sampling without replacement (which is the case when obtaining simple random samples), we shall verify this assumption by checking that the size is less than 5% of the population size ($n < 0.05N$).

For the population presented in Example 1, if we draw a simple random sample of size $n = 9$, the sampling distribution $\bar{x}$ will have mean $\mu_{\bar{x}} = 100$ and standard deviation

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{9}} = 5$$

This standard error of the mean is close to the approximate standard error of 5.01 found in our simulation in Example 1.

In Example 2, where the simple random sample was of size $n = 25$, the sampling distribution of $\bar{x}$ will have mean $\mu_{\bar{x}} = 100$ and standard deviation

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{25}} = 3$$

This standard error of the mean is close to the approximate standard error of 3.08 found in our simulation in Example 2.

• Now Work Problem 9

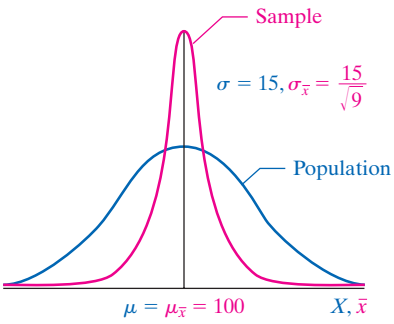
Now that we can find the mean and standard deviation for any sampling distribution of $\bar{x}$, we can concentrate on the shape of the distribution. Refer back to Figures 2(b) and 3 from Examples 1 and 2. Both histograms appear to be normal. Recall that the population from which the sample was drawn was normal. This leads us to believe that, if the population is normal, then the distribution of the sample mean is also normal.

The Shape of the Sampling Distribution of $\bar{x}$ If X Is Normal

If a random variable X is normally distributed, the sampling distribution of the sample mean, $\bar{x}$, is normally distributed.

For example, the IQ scores of individuals are modeled by a normal random variable with mean $\mu = 100$ and standard deviation $\sigma = 15$. The sampling distribution of the sample mean, $\bar{x}$, the mean IQ of a simple random sample of $n = 9$ individuals, is normal, with mean $\mu_{\bar{x}} = 100$ and standard deviation $\sigma_{\bar{x}} = \dfrac{15}{\sqrt{9}}$. See Figure 4.

Figure 4



EXAMPLE 3 Describing the Distribution of the Sample Mean

Problem The IQ, X, of humans is approximately normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$. Compute the probability that a simple random sample of size $n = 10$ results in a sample mean greater than 110. That is, compute $P(\bar{x} > 110)$.

*Technically, we assume that we are drawing a simple random sample from an infinite population. For populations of finite size N, $\sigma_{\bar{x}} = \sqrt{\dfrac{N-n}{N-1}} \cdot \dfrac{\sigma}{\sqrt{n}}$. However, if the sample size is less than 5% of the population size ($n < 0.05N$), the effect of $\sqrt{\dfrac{N-n}{N-1}}$ (the finite population correction factor) can be ignored without significantly affecting the results.

Approach The random variable X is approximately normally distributed, so the sampling distribution of $\bar{x}$ will be normally distributed. Verify the independence requirement. The mean of the sampling distribution is $\mu_{\bar{x}} = \mu$, and its standard deviation is $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$. Convert the sample mean $\bar{x} = 110$ to a z-score and then find the area under the standard normal curve to the right of this z-score.

Solution The sample mean is normally distributed, with mean $\mu_{\bar{x}} = 100$ and standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{15}{\sqrt{10}} = 4.743$. The sample size is definitely less than 5% of the population size.

Figure 5

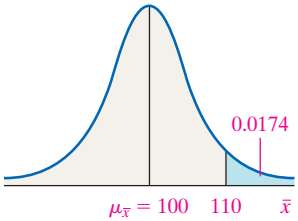


Figure 5 displays the normal curve with the area we want to compute shaded. To find the area by hand, convert $\bar{x} = 110$ to a z-score and obtain

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_{\bar{x}}}{\dfrac{\sigma}{\sqrt{n}}} = \frac{110 - 100}{\dfrac{15}{\sqrt{10}}} = 2.11$$

The area to the right of $z = 2.11$ is $1 - 0.9826 = 0.0174$.

Using technology, the area to the right of $\bar{x} = 110$ is 0.0175.

Interpretation The probability of obtaining a sample mean IQ greater than 110 from a population whose mean is 100 is approximately 0.02. That is, $P(\bar{x} > 110) = 0.0174$ (or 0.0175 using technology). If we take 100 simple random samples of $n = 10$ individuals from this population and if the population mean is 100, about 2 of the samples will result in a mean IQ that is greater than 110. ●

● **Now Work Problem 19**

2 Describe the Distribution of the Sample Mean: Nonnormal Population

Now we explore the distribution of the sample mean assuming the population from which the sample is drawn is not normal. Again we use simulation.

EXAMPLE 4 Sampling from a Population That Is Not Normal

Problem The data in Table 1 represent the probability distribution of the number of people living in households in the United States. Figure 6 shows a graph of the probability distribution. From the data in Table 1, we determine the mean and standard deviation number of people living in households in the United States to be $\mu = 2.9$ and $\sigma = 1.48$.

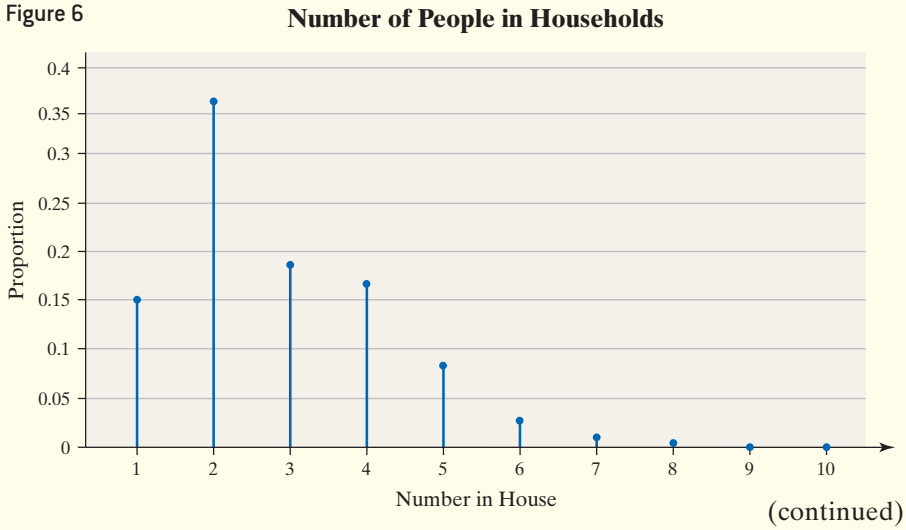
Clearly, the distribution is not normal. In fact, the random variable is discrete! Approximate the sampling distribution of the sample mean $\bar{x}$ by obtaining, through simulation, 1000 samples of size (a) $n = 4$, (b) $n = 10$, and (c) $n = 30$ from the population.

Table 1

Number in Household	Proportion
1	0.147
2	0.361
3	0.187
4	0.168
5	0.083
6	0.034
7	0.012
8	0.004
9	0.002
10	0.002

Source: General Social Survey

Figure 6



(continued)

Approach Use Minitab to obtain 1000 random samples of size $n = 4$ from the population. This simulates going to 4 households 1000 times and determining the number of people living in the household. Next, compute the mean of each of the 1000 random samples. Finally, draw a histogram, determine the mean, and determine the standard deviation of the 1000 sample means. Repeat this for samples of size $n = 10$ and $n = 30$.

Solution Figure 7 shows partial output from Minitab for random samples of size $n = 4$. Columns 1 and 2 represent the probability distribution. Each row in Columns 3 through 6 lists the number of individuals in the household for each sample. Column 7 (xbar) lists the sample mean for each sample (each row). For example, in the first sample (row 1), there are 4 individuals in the first house surveyed, and 2 individuals in the second, third, and fourth houses surveyed. The mean number of individuals in the household for the first sample is 2.5.

Figure 7

Worksheet 2 ***

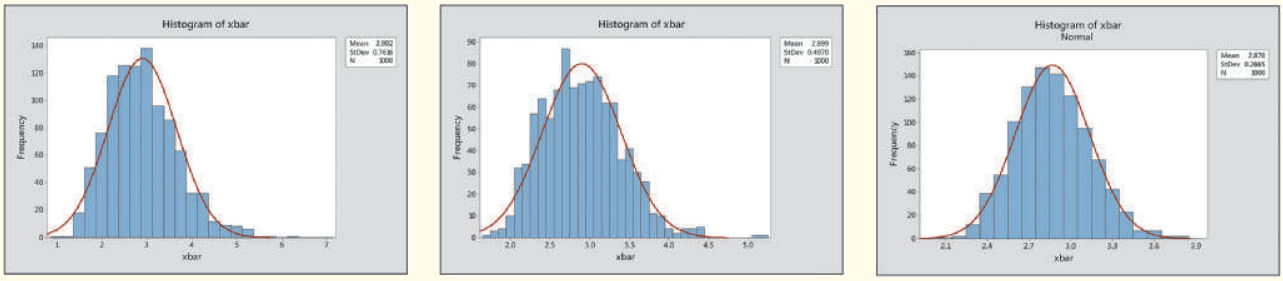
	C1	C2	C3	C4	C5	C6	C7
	Number	Proportion					xbar
1	1	0.147	4	2	2	2	2.50
2	2	0.361	1	3	4	2	2.50
3	3	0.187	4	4	2	4	3.50
4	4	0.168	5	4	2	6	4.25
5	5	0.083	7	2	4	4	4.25
6	6	0.034	2	2	6	2	3.00
7	7	0.012	4	8	3	1	4.00
8	8	0.004	3	2	2	2	2.25
9	9	0.002	2	2	4	2	2.50
10	10	0.002	2	2	4	1	2.25
11			2	1	7	2	3.00
12			3	7	2	1	3.25
13			5	3	2	3	3.25
14			1	5	2	2	2.50
15			1	1	4	7	3.25

Figure 8(a) shows the histogram of the 1000 sample means for a sample of size $n = 4$. The distribution of sample means is skewed right (just like the parent population, but not as strongly). The mean of the 1000 samples is 2.9, and the standard deviation is 0.76. So, the mean of the 1000 samples, $\mu_{\bar{x}}$, equals the population mean μ, and the standard deviation of the 1000 samples, $\sigma_{\bar{x}}$, is close to $\frac{\sigma}{\sqrt{n}} = \frac{1.48}{\sqrt{4}} = 0.74$.

Figure 8(b) shows the histogram of the 1000 sample means for a sample of size $n = 10$. The distribution of these sample means is also skewed right, but not as skewed as the distribution in Figure 8(a). The mean of the 1000 samples is 2.9, and the standard deviation is 0.50. So, the mean of the 1000 samples, $\mu_{\bar{x}}$, equals the population mean, μ, and the standard deviation of the 1000 samples, $\sigma_{\bar{x}}$, is close to $\frac{\sigma}{\sqrt{n}} = \frac{1.48}{\sqrt{10}} = 0.47$.

Figure 8(c) shows the histogram of the 1000 sample means for a sample of size $n = 30$. The distribution of sample means is approximately normal! The mean of the 1000 samples is 2.9, and the standard deviation is 0.27. So, the mean of the 1000 samples, $\mu_{\bar{x}}$, equals the population mean, μ, and the standard deviation of the 1000 samples, $\sigma_{\bar{x}}$, equals $\frac{\sigma}{\sqrt{n}} = \frac{1.48}{\sqrt{30}} = 0.27$.

Figure 8



(a) (b) (c)

There are two key concepts to understand in Example 4.

1. The mean of the sampling distribution of the sample mean is equal to the mean of the underlying population, and the standard deviation of the sampling distribution of the sample mean is $\frac{\sigma}{\sqrt{n}}$, regardless of the size of the sample.
2. The shape of the distribution of the sample mean becomes approximately normal as the sample size n increases, regardless of the shape of the underlying population.

We formally state point 2 as the *Central Limit Theorem*.

In Other Words
For any population, regardless of its shape, as the sample size increases, the shape of the distribution of the sample mean becomes more "normal."

CAUTION!
The Central Limit Theorem only has to do with the shape of the distribution of $\bar{x}$, not the center or spread. Regardless of the size of the sample, $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

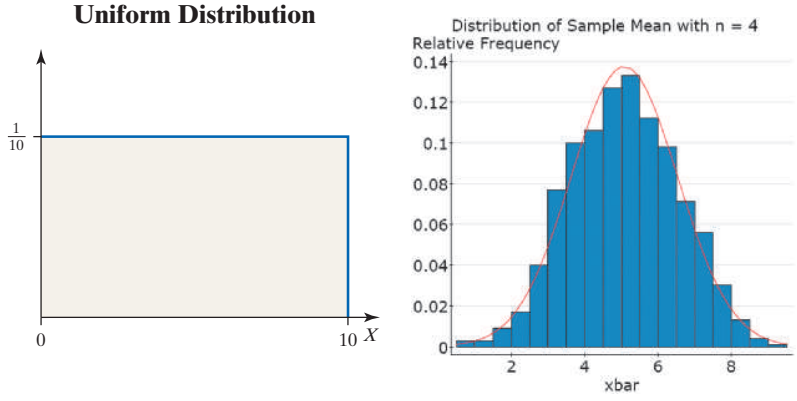
The Central Limit Theorem

Regardless of the shape of the underlying population, the sampling distribution of $\bar{x}$ becomes approximately normal as the sample size, n, increases.

How large does the sample size need to be before we can say that the sampling distribution of $\bar{x}$ is approximately normal? The answer depends on the shape of the distribution of the underlying population. Distributions that are highly skewed will require a larger sample size for the distribution of $\bar{x}$ to become approximately normal.

For example, the right-skewed distribution in Example 4 required a sample size of about 30 before the distribution of the sample mean became approximately normal. However, Figure 9(a) shows a uniform distribution for $0 \le X \le 10$. Figure 9(b) shows the distribution of the sample mean obtained via simulation using StatCrunch for $n = 4$. Even for samples as small as $n = 4$, the distribution of the sample mean is approximately normal.

Figure 9



(a) (b)

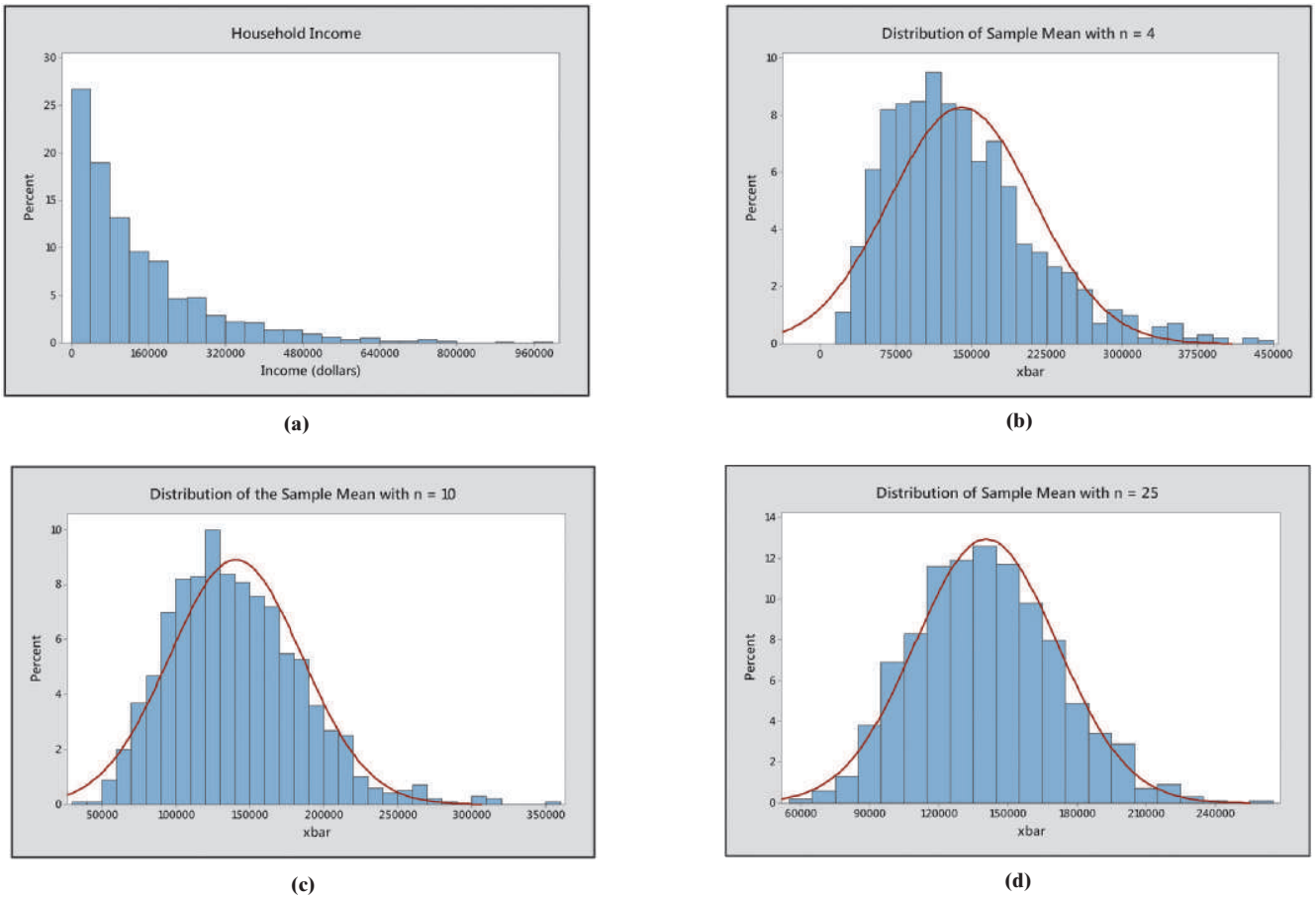
Figure 10(a) on the next page shows a distribution of household incomes for a town. Figure 10(b) shows the distribution of the sample mean for a random sample of $n = 4$ households from Minitab. Figure 10(c) shows the distribution of the sample mean for a random sample of $n = 10$ households, and Figure 10(d) shows the distribution of the

Historical Note

Pierre-Simon Laplace was born on March 23, 1749, in Normandy, France. At age 16, Laplace attended Caen University, where he studied theology. While there, his mathematical talents were discovered, which led him to Paris, where he obtained a job as professor of mathematics at the École Militaire. In 1773, Laplace was elected to the Académie des Sciences. Laplace was not humble. It is reported that, in 1780, he stated that he was the best mathematician in Paris. In 1799, Laplace published the first two volumes of *Mécanique céleste*, in which he discussed methods for calculating the motion of the planets. On April 9, 1810, Laplace presented the Central Limit Theorem to the Academy.

sample mean for a random sample of $n = 25$ households, also from Minitab. Notice the distribution of the sample mean is approximately normal for $n = 25$.

Figure 10



The results of Example 4 and Figures 9 and 10 confirm that the shape of the distribution of the population dictates the size of the sample required for the distribution of the sample mean to be normal. The more skewed the distribution of the population is, the larger the sample size needed to invoke the Central Limit Theorem. We will err on the side of caution and use the following rule of thumb:

> If the distribution of the population is unknown or not normal, then the distribution of the sample mean is approximately normal provided that the sample size is greater than or equal to 30.

EXAMPLE 5 Weight Gain During Pregnancy

Problem The mean weight gain during pregnancy is 30 pounds, with a standard deviation of 12.9 pounds. Weight gain during pregnancy is skewed right. An obstetrician obtains a random sample of 35 low-income patients and determines their mean weight gain during pregnancy was 36.2 pounds. Does this result suggest anything unusual?

Approach We want to know whether the sample mean obtained is unusual. Therefore, determine the likelihood of obtaining a sample mean of 36.2 pounds or higher (if a 36.2-pound weight gain is unusual, certainly any weight gain above 36.2 pounds is also

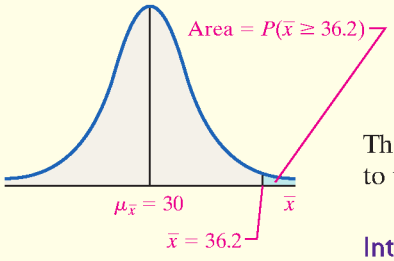
unusual). Assume that the patients come from the population whose mean weight gain is 30 pounds. Verify the independence assumption. Use the normal model to obtain the probability since the sample size is large enough to use the Central Limit Theorem. Determine the area under the normal curve to the right of 36.2 pounds with $\mu_{\bar{x}} = \mu = 30$ and $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{12.9}{\sqrt{35}}$.

Solution It seems reasonable there are at least 700 low-income pregnant women in the population. So the sample size is less than 5% of the population size. The probability is represented by the area under the normal curve to the right of 36.2. See Figure 11.

To find $P(\bar{x} \geq 36.2)$ by hand, convert the sample mean $\bar{x} = 36.2$ to a z-score.

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{36.2 - 30}{\frac{12.9}{\sqrt{35}}} = 2.84$$

Figure 11



The area under the standard normal curve to the left of $z = 2.84$ is 0.9977. So the area to the right is 0.0023. Therefore, $P(\bar{x} \geq 36.2) = 0.0023$.

If we use technology to find the area to the right of $\bar{x} = 36.2$, we obtain 0.0022.

Interpretation If the population from which this sample is drawn has a mean weight gain of 30 pounds, the probability that a random sample of 35 women has a sample mean weight gain of 36.2 pounds (or more) is approximately 0.002. This means that about 2 samples in 1000 will result in a sample mean of 36.2 pounds or higher if the population mean is 30 pounds. We can conclude one of two things based on this result:

1. The mean weight gain for low-income patients is 30 pounds, and we happened to select women who, on average, gained more weight.
2. The mean weight gain for low-income patients is more than 30 pounds.

We are inclined to accept the second explanation over the first since our sample was obtained randomly. Therefore, the obstetrician should be concerned. Perhaps she should look at the diets and/or lifestyles of low-income patients while they are pregnant. ●

● Now Work Problem 25

Summary: Shape, Center, and Spread of the Sampling Distribution of $\bar{x}$

Shape, Center, and Spread of the Population	Distribution of the Sample Mean: Shape	Center	Spread
Population is normal with mean μ and standard deviation σ	Regardless of the sample size n, the shape of the distribution of the sample mean is normal	$\mu_{\bar{x}} = \mu$	$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$
Population is not normal with mean μ and standard deviation σ	As the sample size n increases, the distribution of the sample mean becomes approximately normal	$\mu_{\bar{x}} = \mu$	$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

8.1 Assess Your Understanding

Vocabulary and Skill Building

1. The _____ _____ of the sample mean, $\bar{x}$, is the probability distribution of all possible values of the random variable $\bar{x}$ computed from a sample of size n from a population with mean μ and standard deviation σ.

2. Suppose a simple random sample of size n is drawn from a large population with mean μ and standard deviation σ. The sampling distribution of $\bar{x}$ has mean $\mu_{\bar{x}} =$ _____ and standard deviation $\sigma_{\bar{x}} =$ _____ .

3. The standard deviation of the sampling distribution of $\bar{x}$, $\sigma_{\bar{x}}$, is called the _____ _____ of the_____.