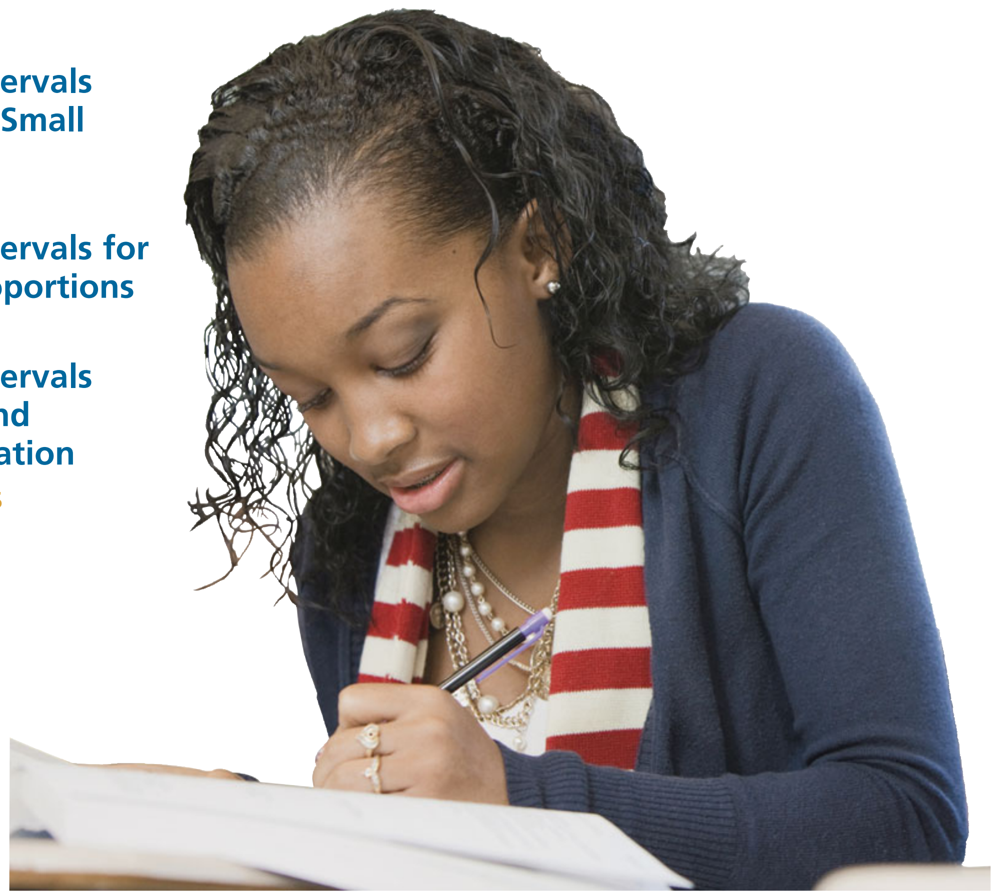
6

CHAPTER

CONFIDENCE INTERVALS



David Wechsler was one of the most influential psychologists of the 20th century. He is known for developing intelligence tests, such as the Wechsler Adult Intelligence Scale and the Wechsler Intelligence Scale for Children.

WHERE YOU'VE BEEN

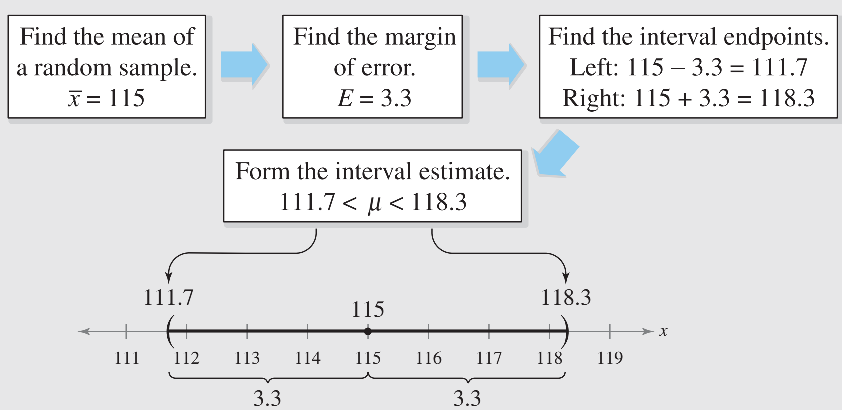
In Chapters 1 through 5, you studied descriptive statistics (how to collect and describe data) and probability (how to find probabilities and analyze discrete and continuous probability distributions). For instance, psychologists use descriptive statistics to analyze the data collected during experiments and trials.

One of the most commonly administered psychological tests is the Wechsler Adult Intelligence Scale. It is an intelligence quotient (IQ) test that is standardized to have a normal distribution with a mean of 100 and a standard deviation of 15.

WHERE YOU'RE GOING

In this chapter, you will begin your study of inferential statistics—the second major branch of statistics. For instance, a chess club wants to estimate the mean IQ of its members. The mean of a random sample of members is 115. Because this estimate consists of a single number represented by a point on a number line, it is called a point estimate. The problem with using a point estimate is that it is rarely equal to the exact parameter (mean, standard deviation, or proportion) of the population.

In this chapter, you will learn how to make a more meaningful estimate by specifying an interval of values on a number line, together with a statement of how confident you are that your interval contains the population parameter. Suppose the club wants to be 90% confident of its estimate for the mean IQ of its members. Here is an overview of how to construct an interval estimate.



So, the club can be 90% confident that the mean IQ of its members is between 111.7 and 118.3.

6.1 Confidence Intervals for the Mean (Large Samples)

WHAT YOU SHOULD LEARN

- How to find a point estimate and a margin of error
- How to construct and interpret confidence intervals for the population mean
- How to determine the minimum sample size required when estimating μ

Estimating Population Parameters ▸ Confidence Intervals for the Population Mean ▸ Sample Size

▸ ESTIMATING POPULATION PARAMETERS

In this chapter, you will learn an important technique of statistical inference—to use sample statistics to estimate the value of an unknown population parameter. In this section, you will learn how to use sample statistics to make an estimate of the population parameter μ when the sample size is at least 30 or when the population is normally distributed and the standard deviation σ is known. To make such an inference, begin by finding a *point estimate*.

DEFINITION

A **point estimate** is a single value estimate for a population parameter. The most unbiased point estimate of the population mean μ is the sample mean $\bar{x}$.

The validity of an estimation method is increased if a sample statistic is unbiased and has low variability. A statistic is unbiased if it does not overestimate or underestimate the population parameter. In Chapter 5, you learned that the mean of all possible sample means of the same size equals the population mean. As a result, $\bar{x}$ is an unbiased estimator of μ. When the standard error $\sigma/\sqrt{n}$ of a sample mean is decreased by increasing n, it becomes less variable.

EXAMPLE 1

▸ Finding a Point Estimate

A social networking website allows its users to add friends, send messages, and update their personal profiles. The following represents a random sample of the number of friends for 40 users of the website. Find a point estimate of the population mean μ. *(Adapted from Facebook)*

140	105	130	97	80	165	232	110	214	201	122
98	65	88	154	133	121	82	130	211	153	114
58	77	51	247	236	109	126	132	125	149	122
74	59	218	192	90	117	105				

▸ Solution

The sample mean of the data is

$$\bar{x} = \frac{\Sigma x}{n} = \frac{5232}{40} = 130.8.$$

So, the point estimate for the mean number of friends for all users of the website is 130.8 friends.

▸ Try It Yourself 1

Another random sample of the number of friends for 30 users of the website is shown at the left. Use this sample to find another point estimate for μ.

a. *Find* the sample mean.
b. *Estimate* the mean number of friends of the population. *Answer: Page A39*

Sample Data

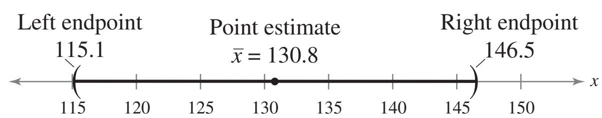
Number of Friends					
162	114	131	87	108	63
249	135	172	196	127	100
146	214	80	55	71	130
95	156	201	227	137	125
145	179	74	215	137	124

In Example 1, the probability that the population mean is exactly 130.8 is virtually zero. So, instead of estimating μ to be exactly 130.8 using a point estimate, you can estimate that μ lies in an interval. This is called making an *interval estimate*.

DEFINITION

An **interval estimate** is an interval, or range of values, used to estimate a population parameter.

Although you can assume that the point estimate in Example 1 is not equal to the actual population mean, it is probably close to it. To form an interval estimate, use the point estimate as the center of the interval, then add and subtract a margin of error. For instance, if the margin of error is 15.7, then an interval estimate would be given by 130.8 ± 15.7 or $115.1 < \mu < 146.5$. The point estimate and interval estimate are as follows.



Interval estimate

Before finding a margin of error for an interval estimate, you should first determine how confident you need to be that your interval estimate contains the population mean μ.

DEFINITION

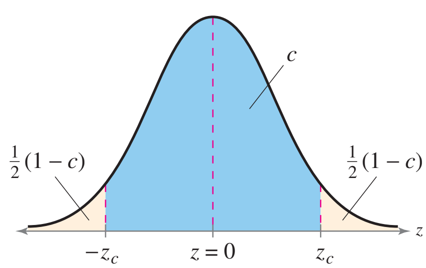
The **level of confidence c** is the probability that the interval estimate contains the population parameter.

STUDY TIP

In this course, you will usually use 90%, 95%, and 99% levels of confidence. The following z-scores correspond to these levels of confidence.

Level of Confidence	z_c
90%	1.645
95%	1.96
99%	2.575

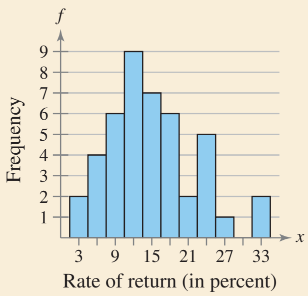
You know from the Central Limit Theorem that when $n \geq 30$, the sampling distribution of sample means is a normal distribution. The level of confidence c is the area under the standard normal curve between the *critical values*, $-z_c$ and z_c. **Critical values** are values that separate sample statistics that are probable from sample statistics that are improbable, or unusual. You can see from the graph that c is the percent of the area under the normal curve between $-z_c$ and z_c. The area remaining is $1 - c$, so the area in each tail is $\frac{1}{2}(1 - c)$. For instance, if $c = 90\%$, then 5% of the area lies to the left of $-z_c = -1.645$ and 5% lies to the right of $z_c = 1.645$.



If $c = 90\%$:	
$c = 0.90$	Area in blue region
$1 - c = 0.10$	Area in yellow regions
$\frac{1}{2}(1 - c) = 0.05$	Area in each tail
$-z_c = -1.645$	Critical value separating left tail
$z_c = 1.645$	Critical value separating right tail

PICTURING THE WORLD

Many investors choose mutual funds as a way to invest in the stock market. The mean annual rate of return for mutual funds in a recent year was estimated by taking a random sample of 44 mutual funds. The mean annual rate of return for the sample was 14.73%, with a standard deviation of 7.23%. (Source: Marketwatch, Inc.)



For a 95% confidence interval, what would be the margin of error for the population mean rate of return?

The difference between the point estimate and the actual parameter value is called the **sampling error.** When μ is estimated, the sampling error is the difference $\bar{x} - \mu$. In most cases, of course, μ is unknown, and $\bar{x}$ varies from sample to sample. However, you can calculate a maximum value for the error if you know the level of confidence and the sampling distribution.

DEFINITION

Given a level of confidence c, the **margin of error E** (sometimes also called the maximum error of estimate or error tolerance) is the greatest possible distance between the point estimate and the value of the parameter it is estimating.

$$E = z_c \sigma_{\bar{x}} = z_c \frac{\sigma}{\sqrt{n}}$$

In order to use this technique, it is assumed that the population standard deviation is known. This is rarely the case, but when $n \geq 30$, the sample standard deviation s can be used in place of σ.

EXAMPLE 2

Finding the Margin of Error

Use the data given in Example 1 and a 95% confidence level to find the margin of error for the mean number of friends for all users of the website. Assume that the sample standard deviation is about 53.0.

Solution

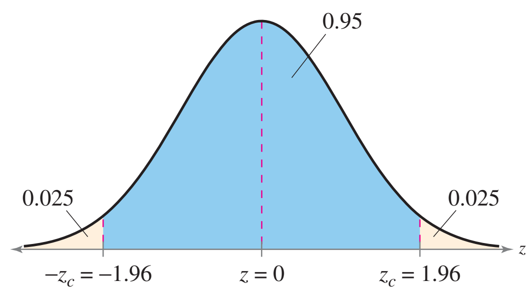
The z-score that corresponds to a 95% confidence level is 1.96. This implies that 95% of the area under the standard normal curve falls within 1.96 standard deviations of the mean. (You can approximate the distribution of the sample means with a normal curve by the Central Limit Theorem because $n = 40 \geq 30$.) You don't know the population standard deviation σ. But because $n \geq 30$, you can use s in place of σ.

Using the values $z_c = 1.96$, $\sigma \approx s \approx 53.0$, and $n = 40$,

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

$$\approx 1.96 \cdot \frac{53.0}{\sqrt{40}}$$

$$\approx 16.4.$$



Interpretation You are 95% confident that the margin of error for the population mean is about 16.4 friends.

Try It Yourself 2

Use the data given in Try It Yourself 1 and a 95% confidence level to find the margin of error for the mean number of friends for all users of the website.

a. *Identify* z_c, n, and s.
b. *Find* E using z_c, $\sigma \approx s$, and n.
c. *Interpret* the results.

Answer: Page A39

STUDY TIP

Remember that you can calculate the sample standard deviation s using the formula

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

or the shortcut formula

$$s = \sqrt{\frac{\Sigma x^2 - (\Sigma x)^2/n}{n - 1}}.$$

However, the most convenient way to find the sample standard deviation is to use the *1–Var Stats* feature of a graphing calculator.

CONFIDENCE INTERVALS FOR THE POPULATION MEAN

Using a point estimate and a margin of error, you can construct an interval estimate of a population parameter such as μ. This interval estimate is called a *confidence interval*.

STUDY TIP

When you compute a confidence interval for a population mean, the general *round-off rule* is to round off to the same number of decimal places given for the sample mean. Recall that rounding is done in the final step.

DEFINITION

A **c-confidence interval for the population mean μ** is

$$\bar{x} - E < \mu < \bar{x} + E.$$

The probability that the confidence interval contains μ is c.

GUIDELINES

Finding a Confidence Interval for a Population Mean ($n \geq 30$ or σ known with a normally distributed population)

IN WORDS	IN SYMBOLS
1. Find the sample statistics n and $\bar{x}$.	$\bar{x} = \dfrac{\Sigma x}{n}$
2. Specify σ, if known. Otherwise, if $n \geq 30$, find the sample standard deviation s and use it as an estimate for σ.	$s = \sqrt{\dfrac{\Sigma (x - \bar{x})^2}{n - 1}}$
3. Find the critical value z_c that corresponds to the given level of confidence.	Use the Standard Normal Table or technology.
4. Find the margin of error E.	$E = z_c \dfrac{\sigma}{\sqrt{n}}$
5. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\bar{x} - E$ Right endpoint: $\bar{x} + E$ Interval: $\bar{x} - E < \mu < \bar{x} + E$

EXAMPLE 3

SC Report 23

See MINITAB steps on page 352.

Constructing a Confidence Interval

Use the data given in Example 1 to construct a 95% confidence interval for the mean number of friends for all users of the website.

Solution In Examples 1 and 2, you found that $\bar{x} = 130.8$ and $E \approx 16.4$. The confidence interval is as follows.

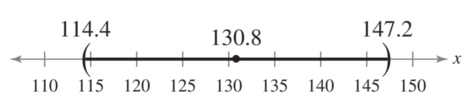
Left Endpoint

$$\bar{x} - E \approx 130.8 - 16.4 = 114.4$$

Right Endpoint

$$\bar{x} + E \approx 130.8 + 16.4 = 147.2$$

$$114.4 < \mu < 147.2$$



Interpretation With 95% confidence, you can say that the population mean number of friends is between 114.4 and 147.2.

STUDY TIP

Other ways to represent a confidence interval are $(\bar{x} - E, \bar{x} + E)$ and $\bar{x} \pm E$. For instance, in Example 3, you could write the confidence interval as (114.4, 147.2) or 130.8 ± 16.4.

INSIGHT

The width of a confidence interval is $2E$. Examine the formula for E to see why a larger sample size tends to give you a narrower confidence interval for the same level of confidence.

▸ Try It Yourself 3

Use the data given in Try It Yourself 1 to construct a 95% confidence interval for the mean number of friends for all users of the website. Compare your result with the interval found in Example 3.

a. *Find* $\overline{x}$ and E.
b. Find the *left* and *right endpoints* of the confidence interval.
c. *Interpret* the results and compare them with Example 3. *Answer: Page A39*

EXAMPLE 4

▸ Constructing a Confidence Interval Using Technology

Use a technology tool to construct a 99% confidence interval for the mean number of friends for all users of the website using the sample in Example 1.

▸ Solution

To use a technology tool to solve the problem, enter the data and recall that the sample standard deviation is $s \approx 53.0$. Then, use the confidence interval command to calculate the confidence interval (*1-Sample Z* for MINITAB). The display should look like the one shown below. To construct a confidence interval using a TI-83/84 Plus, follow the instructions in the margin.

STUDY TIP

Using a TI-83/84 Plus, you can either enter the original data into a list to construct the confidence interval or enter the descriptive statistics.

STAT

Choose the TESTS menu.

7: ZInterval…

Select the *Data* input option if you use the original data. Select the *Stats* input option if you use the descriptive statistics. In each case, enter the appropriate values, then select *Calculate*. Your results may differ slightly depending on the method you use. For Example 4, the original data values were entered.

```
ZInterval
 (109.21,152.39)
 x̄=130.8
 Sx=52.63234844
 n=40
```

MINITAB

One-Sample Z: Friends

The assumed standard deviation = 53

Variable	N	Mean	StDev	SE Mean	99% CI
Friends	40	130.80	52.63	8.38	(109.21, 152.39)

So, a 99% confidence interval for μ is $(109.2, 152.4)$.

Interpretation With 99% confidence, you can say that the population mean number of friends is between 109.2 and 152.4.

▸ Try It Yourself 4

Use the sample data in Example 1 and a technology tool to construct 75%, 85%, and 99% confidence intervals for the mean number of friends for all users of the website. How does the width of the confidence interval change as the level of confidence increases?

a. *Enter* the data.
b. *Use* the appropriate command to construct each confidence interval.
c. *Compare* the widths of the confidence intervals for $c = 0.75, 0.85$, and 0.99.
Answer: Page A39

In Example 4 and Try It Yourself 4, the same sample data were used to construct confidence intervals with different levels of confidence. Notice that as the level of confidence increases, the width of the confidence interval also increases. In other words, when the same sample data are used, *the greater the level of confidence, the wider the interval.*

If the population is normally distributed and the population standard deviation σ is known, you may use the normal sampling distribution for any sample size, as shown in Example 5.

STUDY TIP

Here are instructions for constructing a confidence interval in Excel. First, click *Insert* at the top of the screen and select *Function*. Select the category *Statistical* and select the *Confidence* function. In the dialog box, enter the values of alpha, the standard deviation, and the sample size. Then click OK. The value returned is the margin of error, which is used to construct the confidence interval.

	A	B
1	=CONFIDENCE(0.1,1.5,20)	
2		0.55170068

Alpha is the *level of significance*, which will be explained in Chapter 7. When using Excel in Chapter 6, you can think of alpha as the complement of the level of confidence. So, for a 90% confidence interval, alpha is equal to $1 - 0.90 = 0.10$.

EXAMPLE 5

See TI-83/84 Plus steps on page 353.

Constructing a Confidence Interval, σ Known

A college admissions director wishes to estimate the mean age of all students currently enrolled. In a random sample of 20 students, the mean age is found to be 22.9 years. From past studies, the standard deviation is known to be 1.5 years, and the population is normally distributed. Construct a 90% confidence interval of the population mean age.

Solution

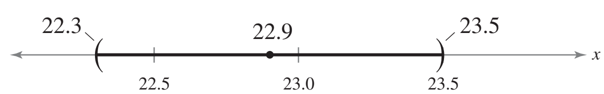
Using $n = 20$, $\bar{x} = 22.9$, $\sigma = 1.5$, and $z_c = 1.645$, the margin of error at the 90% confidence level is

$$E = z_c\frac{\sigma}{\sqrt{n}}$$
$$= 1.645 \cdot \frac{1.5}{\sqrt{20}} \approx 0.6.$$

The 90% confidence interval can be written as $\bar{x} \pm E \approx 22.9 \pm 0.6$ or as follows.

Left Endpoint	Right Endpoint
$\bar{x} - E \approx 22.9 - 0.6 = 22.3$	$\bar{x} + E \approx 22.9 + 0.6 = 23.5$

$$22.3 < \mu < 23.5$$

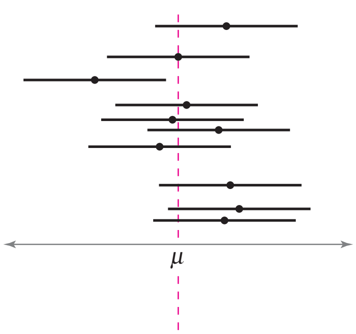


Interpretation With 90% confidence, you can say that the mean age of all the students is between 22.3 and 23.5 years.

Try It Yourself 5

Construct a 90% confidence interval of the population mean age for the college students in Example 5 with the sample size increased to 30 students. Compare your answer with Example 5.

a. *Identify* n, $\bar{x}$, σ, and z_c, and *find* E.
b. Find the *left* and *right endpoints* of the confidence interval.
c. *Interpret* the results and compare them with Example 5. *Answer: Page A39*



The horizontal segments represent 90% confidence intervals for different samples of the same size. In the long run, 9 of every 10 such intervals will contain μ.

After constructing a confidence interval, it is important that you interpret the results correctly. Consider the 90% confidence interval constructed in Example 5. Because μ is a fixed value predetermined by the population, it is either in the interval or not. It is *not* correct to say "There is a 90% probability that the actual mean will be in the interval $(22.3, 23.5)$." This statement is wrong because it suggests that the value of μ can vary, which is not true. The correct way to interpret your confidence interval is "If a large number of samples is collected and a confidence interval is created for each sample, approximately 90% of these intervals will contain μ."

SAMPLE SIZE

For the same sample statistics, as the level of confidence increases, the confidence interval widens. As the confidence interval widens, the precision of the estimate decreases. One way to improve the precision of an estimate without decreasing the level of confidence is to increase the sample size. But how large a sample size is needed to guarantee a certain level of confidence for a given margin of error?

INSIGHT

Using the formula for the margin of error,

$$E = z_c \frac{\sigma}{\sqrt{n}},$$

you can derive the minimum sample size n. (See Exercise 69.)

FIND A MINIMUM SAMPLE SIZE TO ESTIMATE μ

Given a c-confidence level and a margin of error E, the minimum sample size n needed to estimate the population mean μ is

$$n = \left(\frac{z_c \sigma}{E}\right)^2.$$

If σ is unknown, you can estimate it using s, provided you have a preliminary sample with at least 30 members.

EXAMPLE 6

Determining a Minimum Sample Size

You want to estimate the mean number of friends for all users of the website. How many users must be included in the sample if you want to be 95% confident that the sample mean is within seven friends of the population mean?

Solution

Using $c = 0.95$, $z_c = 1.96$, $\sigma \approx s \approx 53.0$ (from Example 2), and $E = 7$, you can solve for the minimum sample size n.

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$
$$\approx \left(\frac{1.96 \cdot 53.0}{7}\right)^2$$
$$\approx 220.23$$

STUDY TIP

When necessary, round up to obtain a whole number when determining a minimum sample size.

When necessary, round up to obtain a whole number. So, you should include at least 221 users in your sample.

Interpretation You already have 40, so you need 181 more. Note that 221 is the *minimum* number of users to include in the sample. You could include more, if desired.

Try It Yourself 6

How many users must be included in the sample if you want to be 95% confident that the sample mean is within 10 users of the population mean? Compare your answer with Example 6.

a. *Identify* z_c, E, and s.
b. Use z_c, E, and $\sigma \approx s$ to find the *minimum sample size* n.
c. *Interpret* the results and compare them with Example 6.

Answer: Page A40

6.1 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. When estimating a population mean, are you more likely to be correct if you use a point estimate or an interval estimate? Explain your reasoning.

2. A news reporter reports the results of a survey and states that 45% of those surveyed responded "yes" with a margin of error of "plus or minus 5%." Explain what this means.

3. Given the same sample statistics, which level of confidence would produce the widest confidence interval? Explain your reasoning.

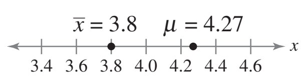
(a) 90% (b) 95% (c) 98% (d) 99%

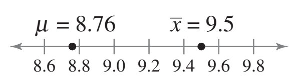
4. You construct a 95% confidence interval for a population mean using a random sample. The confidence interval is $24.9 < \mu < 31.5$. Is the probability that μ is in this interval 0.95? Explain.

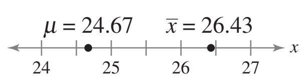
In Exercises 5–8, find the critical value z_c necessary to construct a confidence interval at the given level of confidence.

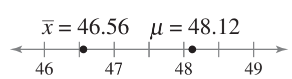
5. $c = 0.80$ **6.** $c = 0.85$ **7.** $c = 0.75$ **8.** $c = 0.97$

Graphical Analysis *In Exercises 9–12, use the values on the number line to find the sampling error.*

9.


10.


11.


12.


In Exercises 13–16, find the margin of error for the given values of c, s, and n.

13. $c = 0.95, s = 5.2, n = 30$

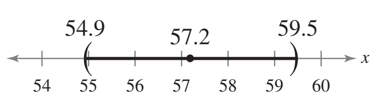
14. $c = 0.90, s = 2.9, n = 50$

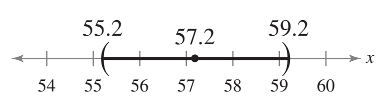
15. $c = 0.80, s = 1.3, n = 75$

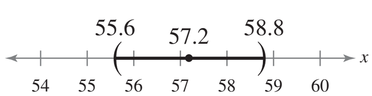
16. $c = 0.975, s = 4.6, n = 100$

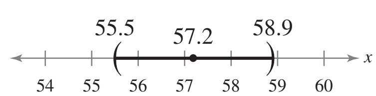
Matching *In Exercises 17–20, match the level of confidence c with its representation on the number line, given $\bar{x} = 57.2$, $s = 7.1$, and $n = 50$.*

17. $c = 0.88$ **18.** $c = 0.90$ **19.** $c = 0.95$ **20.** $c = 0.98$

(a)


(b)


(c)


(d)


In Exercises 21–24, construct the indicated confidence interval for the population mean μ. If convenient, use technology to construct the confidence interval.

21. $c = 0.90, \bar{x} = 12.3, s = 1.5, n = 50$

22. $c = 0.95, \bar{x} = 31.39, s = 0.8, n = 82$

23. $c = 0.99, \bar{x} = 10.5, s = 2.14, n = 45$

24. $c = 0.80, \bar{x} = 20.6, s = 4.7, n = 100$

In Exercises 25–28, use the given confidence interval to find the margin of error and the sample mean.

25. (12.0, 14.8)

26. (21.61, 30.15)

27. (1.71, 2.05)

28. (3.144, 3.176)

In Exercises 29–32, determine the minimum sample size n needed to estimate μ for the given values of c, s, and E.

29. $c = 0.90, s = 6.8, E = 1$

30. $c = 0.95, s = 2.5, E = 1$

31. $c = 0.80, s = 4.1, E = 2$

32. $c = 0.98, s = 10.1, E = 2$

USING AND INTERPRETING CONCEPTS

Finding the Margin of Error *In Exercises 33 and 34, use the given confidence interval to find the estimated margin of error. Then find the sample mean.*

33. Commute Times A government agency reports a confidence interval of (26.2, 30.1) when estimating the mean commute time (in minutes) for the population of workers in a city.

34. Book Prices A store manager reports a confidence interval of (44.07, 80.97) when estimating the mean price (in dollars) for the population of textbooks.

Constructing Confidence Intervals *In Exercises 35–38, you are given the sample mean and the sample standard deviation. Use this information to construct the 90% and 95% confidence intervals for the population mean. Interpret the results and compare the widths of the confidence intervals. If convenient, use technology to construct the confidence intervals.*

35. Home Theater Systems A random sample of 34 home theater systems has a mean price of \$452.80 and a standard deviation of \$85.50.

36. Gasoline Prices From a random sample of 48 days in a recent year, U.S. gasoline prices had a mean of \$2.34 and a standard deviation of \$0.32. *(Source: U.S. Energy Information Administration)*

37. Juice Drinks A random sample of 31 eight-ounce servings of different juice drinks has a mean of 99.3 calories and a standard deviation of 41.5 calories. *(Adapted from The Beverage Institute for Health and Wellness)*

38. Sodium Chloride Concentration In 36 randomly selected seawater samples, the mean sodium chloride concentration was 23 cubic centimeters per cubic meter and the standard deviation was 6.7 cubic centimeters per cubic meter. *(Adapted from Dorling Kindersley Visual Encyclopedia)*

39. Replacement Costs: Transmissions You work for a consumer advocate agency and want to estimate the population mean cost of replacing a car's transmission. As part of your study, you randomly select 50 replacement costs and find the mean to be \$2650.00. The sample standard deviation is \$425.00. Construct a 95% confidence interval for the population mean replacement cost. Interpret the results. *(Adapted from CostHelper)*

40. Repair Costs: Refrigerators In a random sample of 60 refrigerators, the mean repair cost was \$150.00 and the standard deviation was \$15.50. Construct a 99% confidence interval for the population mean repair cost. Interpret the results. *(Adapted from Consumer Reports)*

41. Repeat Exercise 39, changing the sample size to $n = 80$. Which confidence interval is wider? Explain.

42. Repeat Exercise 40, changing the sample size to $n = 40$. Which confidence interval is wider? Explain.

43. Swimming Times A random sample of forty-eight 200-meter swims has a mean time of 3.12 minutes and a standard deviation of 0.09 minute. Construct a 95% confidence interval for the population mean time. Interpret the results.

44. Hotels A random sample of 55 standard hotel rooms in the Philadelphia, PA area has a mean nightly cost of \$154.17 and a standard deviation of \$38.60. Construct a 99% confidence interval for the population mean cost. Interpret the results.

45. Repeat Exercise 43, using a standard deviation of $s = 0.06$ minute. Which confidence interval is wider? Explain.

46. Repeat Exercise 44, using a standard deviation of $s = \$42.50$. Which confidence interval is wider? Explain.

47. If all other quantities remain the same, how does the indicated change affect the width of a confidence interval?

(a) Increase in the level of confidence
(b) Increase in the sample size
(c) Increase in the standard deviation

48. Describe how you would construct a 90% confidence interval to estimate the population mean age for students at your school.

Constructing Confidence Intervals *In Exercises 49 and 50, use the given information to construct the 90% and 99% confidence intervals for the population mean. Interpret the results and compare the widths of the confidence intervals. If convenient, use technology to construct the confidence intervals.*

49. DVRs A research council wants to estimate the mean length of time (in minutes) the average U.S. adult spends watching TVs using digital video recorders (DVRs) each day. To determine this estimate, the research council takes a random sample of 20 U.S. adults and obtains the following results.

15, 18, 17, 20, 24, 12, 9, 15, 14, 25, 8, 6, 10, 14, 16, 20, 27, 10, 9, 13

From past studies, the research council assumes that σ is 1.3 minutes and that the population of times is normally distributed. *(Adapted from the Council for Research Excellence)*

50. Text Messaging A telecommunications company wants to estimate the mean length of time (in minutes) that 18- to 24-year-olds spend text messaging each day. In a random sample of twenty-seven 18- to 24-year-olds, the mean length of time spent text messaging was 29 minutes. From past studies, the company assumes that σ is 4.5 minutes and that the population of times is normally distributed. *(Adapted from the Council for Research Excellence)*

51. Minimum Sample Size Determine the minimum required sample size if you want to be 95% confident that the sample mean is within one unit of the population mean given $\sigma = 4.8$. Assume the population is normally distributed.

52. Minimum Sample Size Determine the minimum required sample size if you want to be 99% confident that the sample mean is within two units of the population mean given $\sigma = 1.4$. Assume the population is normally distributed.

53. Cholesterol Contents of Cheese A cheese processing company wants to estimate the mean cholesterol content of all one-ounce servings of cheese. The estimate must be within 0.5 milligram of the population mean.

(a) Determine the minimum required sample size to construct a 95% confidence interval for the population mean. Assume the population standard deviation is 2.8 milligrams.

(b) Repeat part (a) using a 99% confidence interval.

(c) Which level of confidence requires a larger sample size? Explain.

54. Ages of College Students An admissions director wants to estimate the mean age of all students enrolled at a college. The estimate must be within 1 year of the population mean. Assume the population of ages is normally distributed.

(a) Determine the minimum required sample size to construct a 90% confidence interval for the population mean. Assume the population standard deviation is 1.2 years.

(b) Repeat part (a) using a 99% confidence interval.

(c) Which level of confidence requires a larger sample size? Explain.

Error tolerance = 0.25 oz

FIGURE FOR EXERCISE 55

55. Paint Can Volumes A paint manufacturer uses a machine to fill gallon cans with paint (see figure).

(a) The manufacturer wants to estimate the mean volume of paint the machine is putting in the cans within 0.25 ounce. Determine the minimum sample size required to construct a 90% confidence interval for the population mean. Assume the population standard deviation is 0.85 ounce.

(b) Repeat part (a) using an error tolerance of 0.15 ounce. Which error tolerance requires a larger sample size? Explain.

Error tolerance = 1 mL

FIGURE FOR EXERCISE 56

56. Water Dispensing Machine A beverage company uses a machine to fill one-liter bottles with water (see figure). Assume that the population of volumes is normally distributed.

(a) The company wants to estimate the mean volume of water the machine is putting in the bottles within 1 milliliter. Determine the minimum sample size required to construct a 95% confidence interval for the population mean. Assume the population standard deviation is 3 milliliters.

(b) Repeat part (a) using an error tolerance of 2 milliliters. Which error tolerance requires a larger sample size? Explain.

57. Plastic Sheet Cutting A machine cuts plastic into sheets that are 50 feet (600 inches) long. Assume that the population of lengths is normally distributed.

(a) The company wants to estimate the mean length of the sheets within 0.125 inch. Determine the minimum sample size required to construct a 95% confidence interval for the population mean. Assume the population standard deviation is 0.25 inch.

(b) Repeat part (a) using an error tolerance of 0.0625 inch. Which error tolerance requires a larger sample size? Explain.

58. Paint Sprayer A company uses an automated sprayer to apply paint to metal furniture. The company sets the sprayer to apply the paint one mil (1/1000 of an inch) thick.

(a) The company wants to estimate the mean thickness of paint the sprayer is applying within 0.0425 mil. Determine the minimum sample size required to construct a 90% confidence interval for the population mean. Assume the population standard deviation is 0.15 mil.

(b) Repeat part (a) using an error tolerance of 0.02125 mil. Which error tolerance requires a larger sample size? Explain.

59. Soccer Balls A soccer ball manufacturer wants to estimate the mean circumference of soccer balls within 0.1 inch.

(a) Determine the minimum sample size required to construct a 99% confidence interval for the population mean. Assume the population standard deviation is 0.25 inch.

(b) Repeat part (a) using a standard deviation of 0.3 inch. Which standard deviation requires a larger sample size? Explain.

60. Mini-Soccer Balls A soccer ball manufacturer wants to estimate the mean circumference of mini-soccer balls within 0.15 inch. Assume that the population of circumferences is normally distributed.

(a) Determine the minimum sample size required to construct a 99% confidence interval for the population mean. Assume the population standard deviation is 0.20 inch.

(b) Repeat part (a) using a standard deviation of 0.10 inch. Which standard deviation requires a larger sample size? Explain.

61. If all other quantities remain the same, how does the indicated change affect the minimum sample size requirement?

(a) Increase in the level of confidence

(b) Increase in the error tolerance

(c) Increase in the standard deviation

62. When estimating the population mean, why not construct a 99% confidence interval every time?

Using Technology

In Exercises 63 and 64, you are given a data sample. Use a technology tool to construct a 95% confidence interval for the population mean. Interpret your answer.

63. Airfare The stem-and-leaf plot shows the results of a random sample of airfare prices (in dollars) for a one-way ticket from Boston, MA to Chicago, IL *(Adapted from Expedia, Inc.)*

18	3 3
19	7
20	9 9
21	2 2 2 3 3 3 3 3 3 3 6 6
22	2 2 2 2 3 6 6 8 8 8 8 8 9
23	8 8

Key: 18|3 = 183

FIGURE FOR EXERCISE 63

64. Stock Prices A random sample of the closing stock prices for the Oracle Corporation for a recent year *(Source: Yahoo! Inc.)*

18.41	16.91	16.83	17.72	15.54	15.56	18.01	19.11	19.79
18.32	18.65	20.71	20.66	21.04	21.74	22.13	21.96	22.16
22.86	20.86	20.74	22.05	21.42	22.34	22.83	24.34	17.97
14.47	19.06	18.42	20.85	21.43	21.97	21.81		

SC *In Exercises 65 and 66, use StatCrunch to construct the 80%, 90%, and 95% confidence intervals for the population mean. Interpret the results.*

65. Sodium A random sample of 30 sandwiches from a fast food restaurant has a mean of 1042.7 milligrams of sodium and a standard deviation of 344.9 milligrams of sodium. *(Source: McDonald's Corporation)*

66. Carbohydrates The following represents a random sample of the amounts of carbohydrates (in grams) for 30 sandwiches from a fast food restaurant. *(Source: McDonald's Corporation)*

31	33	34	33	37	40	40	45	37	38	63	61	59	38	40
44	51	59	52	60	54	62	39	33	26	34	27	35	28	26

EXTENDING CONCEPTS

Finite Population Correction Factor *In Exercises 67 and 68, use the following information.*

In this section, you studied the construction of a confidence interval to estimate a population mean when the population is large or infinite. When a population is finite, the formula that determines the standard error of the mean $\sigma_{\bar{x}}$ needs to be adjusted. If N is the size of the population and n is the size of the sample (where $n \geq 0.05N$), the standard error of the mean is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}.$$

The expression $\sqrt{(N-n)/(N-1)}$ is called the ***finite population correction factor.*** The margin of error is

$$E = z_c \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}.$$

67. Determine the finite population correction factor for each of the following.

(a) $N = 1000$ and $n = 500$

(b) $N = 1000$ and $n = 100$

(c) $N = 1000$ and $n = 75$

(d) $N = 1000$ and $n = 50$

(e) What happens to the finite population correction factor as the sample size n decreases but the population size N remains the same?

68. Determine the finite population correction factor for each of the following.

(a) $N = 100$ and $n = 50$

(b) $N = 400$ and $n = 50$

(c) $N = 700$ and $n = 50$

(d) $N = 1000$ and $n = 50$

(e) What happens to the finite population correction factor as the population size N increases but the sample size n remains the same?

69. Sample Size The equation for determining the sample size

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

can be obtained by solving the equation for the margin of error

$$E = \frac{z_c \sigma}{\sqrt{n}}$$

for n. Show that this is true and justify each step.

Marathon Training

A marathon is a foot race with a distance of 26.22 miles. It was one of the original events of the modern Olympics, where it was a men's-only event. The women's marathon did not become an Olympic event until 1984. The Olympic record for the men's marathon was set during the 2008 Olympics by Samuel Kamau Wanjiru of Kenya, with a time of 2 hours, 6 minutes, 32 seconds. The Olympic record for the women's marathon was set during the 2000 Olympics by Naoko Takahashi of Japan, with a time of 2 hours, 23 minutes, 14 seconds.

Training for a marathon typically lasts at least 6 months. The training is gradual, with increases in distance about every 2 weeks. About 1 to 3 weeks before the race, the distance run is decreased slightly. The stem-and-leaf plots below show the marathon training times (in minutes) for a sample of 30 male runners and 30 female runners.

Training Times (in minutes) of Male Runners

Key: $15|5 = 155$

15	5 8 9 9 9
16	0 0 0 0 1 2 3 4 4 5 8 9
17	0 1 1 3 5 6 6 7 7 9
18	0 1 5

Training Times (in minutes) of Female Runners

Key: $17|8 = 178$

17	8 9 9
18	0 0 0 0 1 2 3 4 6 6 7 9
19	0 0 0 1 3 4 5 5 6 6
20	0 0 1 2 3

EXERCISES

1. Use the sample to find a point estimate for the mean training time of the
(a) male runners.
(b) female runners.

2. Find the standard deviation of the training times for the
(a) male runners.
(b) female runners.

3. Use the sample to construct a 95% confidence interval for the population mean training time of the
(a) male runners.
(b) female runners.

4. Interpret the results of Exercise 3.

5. Use the sample to construct a 95% confidence interval for the population mean training time of all runners. How do your results differ from those in Exercise 3? Explain.

6. A trainer wants to estimate the population mean running times for both male and female runners within 2 minutes. Determine the minimum sample size required to construct a 99% confidence interval for the population mean training time of
(a) male runners. Assume the population standard deviation is 8.9 minutes.
(b) female runners. Assume the population standard deviation is 8.4 minutes.

6.2 Confidence Intervals for the Mean (Small Samples)

WHAT YOU SHOULD LEARN

- How to interpret the *t*-distribution and use a *t*-distribution table
- How to construct confidence intervals when $n < 30$, the population is normally distributed, and σ is unknown

The *t*-Distribution ▸ Confidence Intervals and *t*-Distributions

▸ THE *t*-DISTRIBUTION

In many real-life situations, the population standard deviation is unknown. Moreover, because of various constraints such as time and cost, it is often not practical to collect samples of size 30 or more. So, how can you construct a confidence interval for a population mean given such circumstances? If the random variable is normally distributed (or approximately normally distributed), you can use a *t-distribution*.

INSIGHT

The following example illustrates the concept of degrees of freedom. Suppose the number of chairs in a classroom equals the number of students: 25 chairs and 25 students. Each of the first 24 students to enter the classroom has a choice as to which chair he or she will sit in. There is no freedom of choice, however, for the 25th student who enters the room.

DEFINITION

If the distribution of a random variable x is approximately normal, then

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

follows a ***t*-distribution.**

Critical values of t are denoted by t_c. Several properties of the *t*-distribution are as follows.

1. The *t*-distribution is bell-shaped and symmetric about the mean.
2. The *t*-distribution is a family of curves, each determined by a parameter called the *degrees of freedom*. The **degrees of freedom** are the number of free choices left after a sample statistic such as $\bar{x}$ is calculated. When you use a *t*-distribution to estimate a population mean, the degrees of freedom are equal to one less than the sample size.

   d.f. $= n - 1$ Degrees of freedom

3. The total area under a *t*-curve is 1 or 100%.
4. The mean, median, and mode of the *t*-distribution are equal to 0.
5. As the degrees of freedom increase, the *t*-distribution approaches the normal distribution. After 30 d.f., the *t*-distribution is very close to the standard normal *z*-distribution.

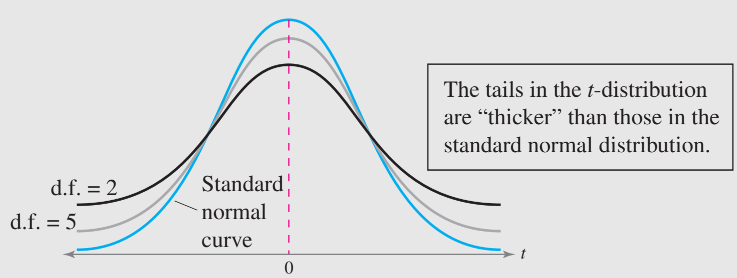


Table 5 in Appendix B lists critical values of t for selected confidence intervals and degrees of freedom.

EXAMPLE 1

Finding Critical Values of *t*

Find the critical value t_c for a 95% confidence level when the sample size is 15.

Solution

Because $n = 15$, the degrees of freedom are

$$\begin{aligned} \text{d.f.} &= n - 1 \\ &= 15 - 1 \\ &= 14. \end{aligned}$$

A portion of Table 5 is shown. Using d.f. $= 14$ and $c = 0.95$, you can find the critical value t_c, as shown by the highlighted areas in the table.

STUDY TIP

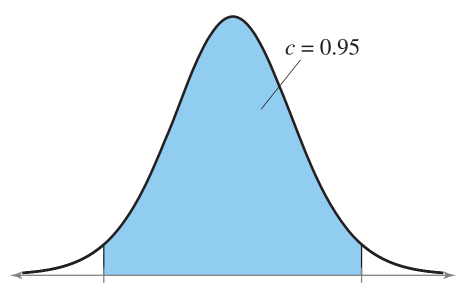
Unlike the *z*-table, critical values for a specific confidence interval can be found in the column headed by *c* in the appropriate d.f. row. (The symbol α will be explained in Chapter 7.)

d.f.	Level of confidence, *c*	0.50	0.80	0.90	0.95	0.98
	One tail, α	0.25	0.10	0.05	0.025	0.01
	Two tails, α	0.50	0.20	0.10	0.05	0.02
1		1.000	3.078	6.314	12.706	31.821
2		.816	1.886	2.920	4.303	6.965
3		.765	1.638	2.353	3.182	4.541
12		.695	1.356	1.782	2.179	2.681
13		.694	1.350	1.771	2.160	2.650
14		.692	1.345	1.761	2.145	2.624
15		.691	1.341	1.753	2.131	2.602
16		.690	1.337	1.746	2.120	2.583
28		.683	1.313	1.701	2.048	2.467
29		.683	1.311	1.699	2.045	2.462
∞		.674	1.282	1.645	1.960	2.326

INSIGHT

For 30 or more degrees of freedom, the critical values for the *t*-distribution are close to the corresponding critical values for the normal distribution. Moreover, the values in the last row of the table marked ∞ d.f. correspond *exactly* to the normal distribution values.

From the table, you can see that $t_c = 2.145$. The graph shows the *t*-distribution for 14 degrees of freedom, $c = 0.95$, and $t_c = 2.145$.



Interpretation So, 95% of the area under the *t*-distribution curve with 14 degrees of freedom lies between $t = \pm 2.145$.

Try It Yourself 1

Find the critical value t_c for a 90% confidence level when the sample size is 22.

a. Identify the *degrees of freedom*.
b. Identify the *level of confidence c*.
c. *Use* Table 5 in Appendix B to find t_c.

Answer: Page A40

CONFIDENCE INTERVALS AND *t*-DISTRIBUTIONS

Constructing a confidence interval using the t-distribution is similar to constructing a confidence interval using the normal distribution—both use a point estimate $\bar{x}$ and a margin of error E.

GUIDELINES

Constructing a Confidence Interval for the Mean: *t*-Distribution

IN WORDS	IN SYMBOLS
1. Find the sample statistics n, $\bar{x}$, and s.	$\bar{x} = \dfrac{\Sigma x}{n}$, $s = \sqrt{\dfrac{\Sigma(x-\bar{x})^2}{n-1}}$
2. Identify the degrees of freedom, the level of confidence c, and the critical value t_c.	d.f. $= n - 1$
3. Find the margin of error E.	$E = t_c \dfrac{s}{\sqrt{n}}$
4. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\bar{x} - E$ Right endpoint: $\bar{x} + E$ Interval: $\bar{x} - E < \mu < \bar{x} + E$

STUDY TIP

For a TI-83/84 Plus, constructing a confidence interval using the t-distribution is similar to constructing a confidence interval using the normal distribution.

STAT

Choose the TESTS menu.

8: TInterval...

Select the *Data* input option if you use the original data. Select the *Stats* input option if you use the descriptive statistics. In each case, enter the appropriate values, then select *Calculate*. Your results may vary slightly depending on the method you use. For Example 2, the descriptive statistics were entered.

```
TInterval
 (156.67,167.33)
 x̄=162
 Sx=10
 n=16
```

EXAMPLE 2

SC Report 24

See MINITAB steps on page 352.

Constructing a Confidence Interval

You randomly select 16 coffee shops and measure the temperature of the coffee sold at each. The sample mean temperature is 162.0°F with a sample standard deviation of 10.0°F. Construct a 95% confidence interval for the population mean temperature. Assume the temperatures are approximately normally distributed.

Solution

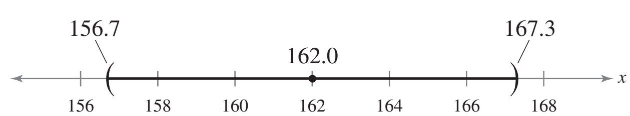
Because the sample size is less than 30, σ is unknown, and the temperatures are approximately normally distributed, you can use the t-distribution. Using $n = 16$, $\bar{x} = 162.0$, $s = 10.0$, $c = 0.95$, and d.f. $= 15$, you can use Table 5 to find that $t_c = 2.131$. The margin of error at the 95% confidence level is

$$E = t_c \frac{s}{\sqrt{n}} = 2.131 \cdot \frac{10.0}{\sqrt{16}} \approx 5.3.$$

The confidence interval is as follows.

Left Endpoint	Right Endpoint
$\bar{x} - E \approx 162 - 5.3 = 156.7$	$\bar{x} + E \approx 162 + 5.3 = 167.3$

$$156.7 < \mu < 167.3$$



Interpretation With 95% confidence, you can say that the population mean temperature of coffee sold is between 156.7°F and 167.3°F.

▸ Try It Yourself 2

Construct 90% and 99% confidence intervals for the population mean temperature.

a. *Find* t_c and E for each level of confidence.
b. Use $\bar{x}$ and E to find the *left* and *right endpoints* of the confidence interval.
c. *Interpret* the results.

Answer: Page A40

EXAMPLE 3

SC Report 25

See TI-83/84 Plus steps on page 353.

▸ Constructing a Confidence Interval

You randomly select 20 cars of the same model that were sold at a car dealership and determine the number of days each car sat on the dealership's lot before it was sold. The sample mean is 9.75 days, with a sample standard deviation of 2.39 days. Construct a 99% confidence interval for the population mean number of days the car model sits on the dealership's lot. Assume the days on the lot are normally distributed.

To explore this topic further, see Activity 6.2 on page 326.

▸ Solution

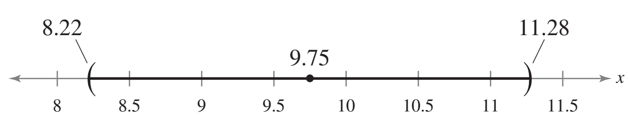
Because the sample size is less than 30, σ is unknown, and the days on the lot are normally distributed, you can use the t-distribution. Using $n = 20$, $\bar{x} = 9.75$, $s = 2.39$, $c = 0.99$, and d.f. $= 19$, you can use Table 5 to find that $t_c = 2.861$. The margin of error at the 99% confidence level is

$$E = t_c \frac{s}{\sqrt{n}} = 2.861 \cdot \frac{2.39}{\sqrt{20}} \approx 1.53.$$

The confidence interval is as follows.

Left Endpoint	Right Endpoint
$\bar{x} - E \approx 9.75 - 1.53 = 8.22$	$\bar{x} + E \approx 9.75 + 1.53 = 11.28$

$$8.22 < \mu < 11.28$$



Interpretation With 99% confidence, you can say that the population mean number of days the car model sits on the dealership's lot is between 8.22 and 11.28.

▸ Try It Yourself 3

Construct 90% and 95% confidence intervals for the population mean number of days the car model sits on the dealership's lot. Compare the widths of the confidence intervals.

a. *Find* t_c and E for each level of confidence.
b. Use $\bar{x}$ and E to find the *left* and *right endpoints* of the confidence interval.
c. *Interpret* the results and compare the widths of the confidence intervals.

Answer: Page A40

HISTORICAL REFERENCE

William S. Gosset (1876–1937)

Developed the t-distribution while employed by the Guinness Brewing Company in Dublin, Ireland. Gosset published his findings using the pseudonym Student. The t-distribution is sometimes referred to as Student's t-distribution. (See page 33 for others who were important in the history of statistics.)

PICTURING THE WORLD

Two footballs, one filled with air and the other filled with helium, were kicked on a windless day at Ohio State University. The footballs were alternated with each kick. After 10 practice kicks, each football was kicked 29 more times. The distances (in yards) are listed. (Source: The Columbus Dispatch)

Air Filled

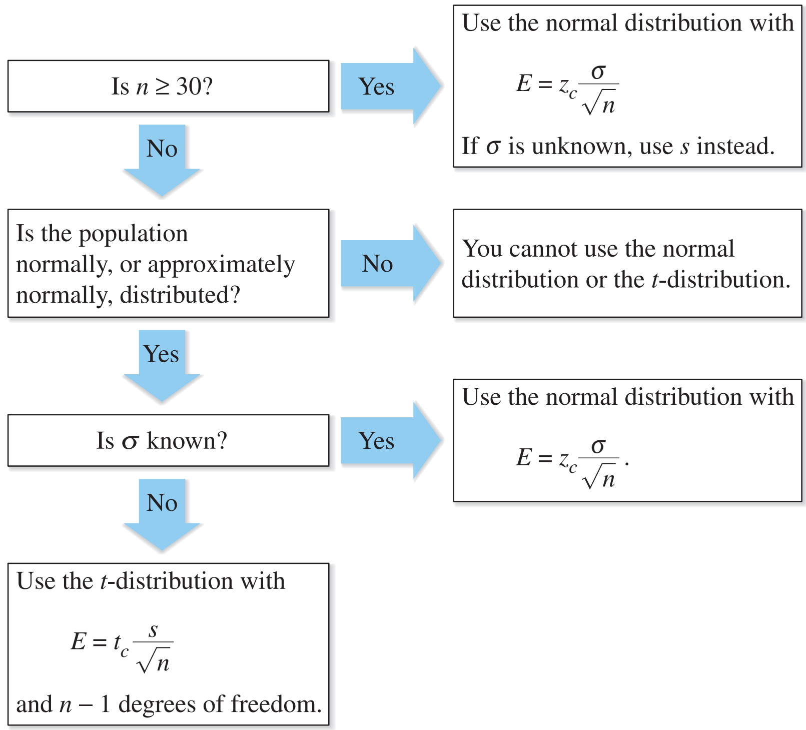
1	9
2	0 0 2 2 2
2	5 5 5 5 6 6
2	7 7 7 8 8 8 8 8 9 9 9
3	1 1 1 2
3	3 4 Key: 1|9 = 19

Helium Filled

1	1 2
1	4
1	
2	2
2	3 4 6 6 6
2	7 8 8 8 9 9 9 9
3	0 0 0 0 1 1 2 2
3	3 4 5
3	9 Key: 1|1 = 11

Assume that the distances are normally distributed for each football. Apply the flowchart at the right to each sample. Construct a 95% confidence interval for the population mean distance each football traveled. Do the confidence intervals overlap? What does this result tell you?

The following flowchart describes when to use the normal distribution and when to use a t-distribution to construct a confidence interval for the population mean.



EXAMPLE 4

Choosing the Normal Distribution or the t-Distribution

You randomly select 25 newly constructed houses. The sample mean construction cost is \$181,000 and the population standard deviation is \$28,000. Assuming construction costs are normally distributed, should you use the normal distribution, the t-distribution, or neither to construct a 95% confidence interval for the population mean construction cost? Explain your reasoning.

Solution

Because the population is normally distributed and the population standard deviation is known, you should use the normal distribution.

Try It Yourself 4

You randomly select 18 adult male athletes and measure the resting heart rate of each. The sample mean heart rate is 64 beats per minute, with a sample standard deviation of 2.5 beats per minute. Assuming the heart rates are normally distributed, should you use the normal distribution, the t-distribution, or neither to construct a 90% confidence interval for the population mean heart rate? Explain your reasoning.

Use the flowchart above to determine which distribution you should use to construct the 90% confidence interval for the population mean heart rate.

Answer: Page A40

6.2 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

In Exercises 1–4, find the critical value t_c for the given confidence level c and sample size n.

1. $c = 0.90,\ n = 10$

2. $c = 0.95,\ n = 12$

3. $c = 0.99,\ n = 16$

4. $c = 0.98,\ n = 20$

In Exercises 5–8, find the margin of error for the given values of c, s, and n.

5. $c = 0.95,\ s = 5,\ n = 16$

6. $c = 0.99,\ s = 3,\ n = 6$

7. $c = 0.90,\ s = 2.4,\ n = 12$

8. $c = 0.98,\ s = 4.7,\ n = 9$

In Exercises 9–12, (a) construct the indicated confidence interval for the population mean μ using a t-distribution. (b) If you had incorrectly used a normal distribution, which interval would be wider?

9. $c = 0.90,\ \bar{x} = 12.5,\ s = 2.0,\ n = 6$

10. $c = 0.95,\ \bar{x} = 13.4,\ s = 0.85,\ n = 8$

11. $c = 0.98,\ \bar{x} = 4.3,\ s = 0.34,\ n = 14$

12. $c = 0.99,\ \bar{x} = 24.7,\ s = 4.6,\ n = 10$

In Exercises 13–16, use the given confidence interval to find the margin of error and the sample mean.

13. (14.7, 22.1)

14. (6.17, 8.53)

15. (64.6, 83.6)

16. (16.2, 29.8)

USING AND INTERPRETING CONCEPTS

Constructing Confidence Intervals *In Exercises 17 and 18, you are given the sample mean and the sample standard deviation. Assume the random variable is normally distributed and use a t-distribution to find the margin of error and construct a 95% confidence interval for the population mean. Interpret the results. If convenient, use technology to construct the confidence interval.*

17. Commute Time to Work In a random sample of eight people, the mean commute time to work was 35.5 minutes and the sample standard deviation was 7.2 minutes.

18. Driving Distance to Work In a random sample of five people, the mean driving distance to work was 22.2 miles and the sample standard deviation was 5.8 miles.

19. You research commute times to work and find that the population standard deviation was 9.3 minutes. Repeat Exercise 17, using a normal distribution with the appropriate calculations for a standard deviation that is known. Compare the results.

20. You research driving distances to work and find that the population standard deviation was 5.2 miles. Repeat Exercise 18, using a normal distribution with the appropriate calculations for a standard deviation that is known. Compare the results.

Constructing Confidence Intervals *In Exercises 21 and 22, you are given the sample mean and the sample standard deviation. Assume the random variable is normally distributed and use a normal distribution or a t-distribution to construct a 90% confidence interval for the population mean. If convenient, use technology to construct the confidence interval.*

21. Waste Generated (a) In a random sample of 10 adults from the United States, the mean waste generated per person per day was 4.50 pounds and the standard deviation was 1.21 pounds. (b) Repeat part (a), assuming the same statistics came from a sample size of 500. Compare the results. *(Adapted from U.S. Environmental Protection Agency)*

22. Waste Recycled (a) In a random sample of 12 adults from the United States, the mean waste recycled per person per day was 1.50 pounds and the standard deviation was 0.28 pound. (b) Repeat part (a), assuming the same statistics came from a sample size of 600. Compare the results. *(Adapted from U.S. Environmental Protection Agency)*

Constructing Confidence Intervals *In Exercises 23–26, a data set is given. For each data set, (a) find the sample mean, (b) find the sample standard deviation, and (c) construct a 99% confidence interval for the population mean. Assume the population of each data set is normally distributed. If convenient, use a technology tool.*

23. Earnings The annual earnings of 16 randomly selected computer software engineers *(Adapted from U.S. Bureau of Labor Statistics)*

92,184	86,919	90,176	91,740	95,535	90,108	94,815	88,114
85,406	90,197	89,944	93,950	84,116	96,054	85,119	88,549

24. Earnings The annual earnings of 14 randomly selected physical therapists *(Adapted from U.S. Bureau of Labor Statistics)*

63,118	65,740	72,899	68,500	66,726	65,554	69,247
64,963	68,627	70,448	71,842	66,873	74,103	71,138

25. SAT Scores The SAT scores of 12 randomly selected high school seniors

1704	1940	1518	2005	1432	1872
1998	1658	1825	1670	2210	1380

26. GPA The grade point averages (GPA) of 15 randomly selected college students

2.3	3.3	2.6	1.8	0.2	3.1	4.0	0.7
2.3	2.0	3.1	3.4	1.3	2.6	2.6	

Choosing a Distribution *In Exercises 27–32, use a normal distribution or a t-distribution to construct a 95% confidence interval for the population mean. Justify your decision. If neither distribution can be used, explain why. Interpret the results. If convenient, use technology to construct the confidence interval.*

27. Body Mass Index In a random sample of 50 people, the mean body mass index (BMI) was 27.7 and the standard deviation was 6.12. Assume the body mass indexes are normally distributed. *(Adapted from Centers for Disease Control)*

28. Mortgages In a random sample of 15 mortgage institutions, the mean interest rate was 4.99% and the standard deviation was 0.36%. Assume the interest rates are normally distributed. *(Adapted from Federal Reserve)*

29. Sports Cars: Miles per Gallon You take a random survey of 25 sports cars and record the miles per gallon for each. The data are listed below. Assume the miles per gallon are normally distributed.

15 27 24 24 20 21 24 14 21 25 21 13 21
25 22 21 25 24 22 24 24 22 21 24 24

30. Yards Per Carry In a recent season, the standard deviation of the yards per carry for all running backs was 1.34. The yards per carry of 20 randomly selected running backs are listed below. Assume the yards per carry are normally distributed. *(Source: National Football League)*

5.6 4.4 3.8 4.5 3.3 5.0 3.6 3.7 4.8 3.5
5.6 3.0 6.8 4.7 2.2 3.3 5.7 3.0 5.0 4.5

31. Hospital Waiting Times In a random sample of 19 patients at a hospital's minor emergency department, the mean waiting time before seeing a medical professional was 23 minutes and the standard deviation was 11 minutes. Assume the waiting times are not normally distributed.

32. Hospital Length of Stay In a random sample of 13 people, the mean length of stay at a hospital was 6.3 days and the standard deviation was 1.7 days. Assume the lengths of stay are normally distributed. *(Adapted from American Hospital Association)*

SC *In Exercises 33 and 34, use StatCrunch to construct the 90%, 95%, and 99% confidence intervals for the population mean. Interpret the results and compare the widths of the confidence intervals. Assume the random variable is normally distributed.*

33. Homework The weekly time spent (in hours) on homework for 18 randomly selected high school students

12.0 11.3 13.5 11.7 12.0 13.0 15.5 10.8 12.5
12.3 14.0 9.5 8.8 10.0 12.8 15.0 11.8 13.0

34. Weight Lifting In a random sample of 11 college football players, the mean weekly time spent weight lifting was 7.2 hours and the standard deviation was 1.9 hours.

EXTENDING CONCEPTS

35. Tennis Ball Manufacturing A company manufactures tennis balls. When its tennis balls are dropped onto a concrete surface from a height of 100 inches, the company wants the mean height the balls bounce upward to be 55.5 inches. This average is maintained by periodically testing random samples of 25 tennis balls. If the t-value falls between $-t_{0.99}$ and $t_{0.99}$, the company will be satisfied that it is manufacturing acceptable tennis balls. A sample of 25 balls is randomly selected and tested. The mean bounce height of the sample is 56.0 inches and the standard deviation is 0.25 inch. Assume the bounce heights are approximately normally distributed. Is the company making acceptable tennis balls? Explain your reasoning.

36. Light Bulb Manufacturing A company manufactures light bulbs. The company wants the bulbs to have a mean life span of 1000 hours. This average is maintained by periodically testing random samples of 16 light bulbs. If the t-value falls between $-t_{0.99}$ and $t_{0.99}$, the company will be satisfied that it is manufacturing acceptable light bulbs. A sample of 16 light bulbs is randomly selected and tested. The mean life span of the sample is 1015 hours and the standard deviation is 25 hours. Assume the life spans are approximately normally distributed. Is the company making acceptable light bulbs? Explain your reasoning.

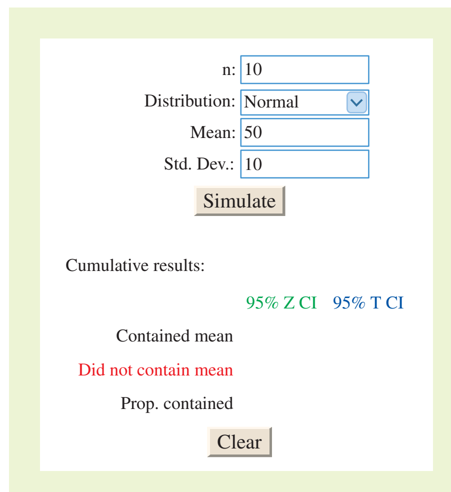
ACTIVITY 6.2 Confidence Intervals for a Mean (the impact of not knowing the standard deviation)

APPLET

The *confidence intervals for a mean (the impact of not knowing the standard deviation)* applet allows you to visually investigate confidence intervals for a population mean. You can specify the sample size n, the shape of the distribution (Normal or Right-skewed), the population mean (Mean), and the true population standard deviation (Std. Dev.). When you click SIMULATE, 100 separate samples of size n will be selected from a population with these population parameters. For each of the 100 samples, a 95% Z confidence interval (known standard deviation) and a 95% T confidence interval (unknown standard deviation) are displayed in the plot at the right. The 95% Z confidence interval is displayed in green and the 95% T confidence interval is displayed in blue. If an interval does not contain the population mean, it is displayed in red. Additional simulations can be carried out by clicking SIMULATE multiple times. The cumulative number of times that each type of interval contains the population mean is also shown. Press CLEAR to clear existing results and start a new simulation.

Explore

Step 1 Specify a value for n.
Step 2 Specify a distribution.
Step 3 Specify a value for the mean.
Step 4 Specify a value for the standard deviation.
Step 5 Click SIMULATE to generate the confidence intervals.



Draw Conclusions

APPLET

1. Set $n = 30$, Mean $= 25$, Std. Dev. $= 5$, and the distribution to Normal. Run the simulation so that at least 1000 confidence intervals are generated. Compare the proportion of the 95% Z confidence intervals and 95% T confidence intervals that contain the population mean. Is this what you would expect? Explain.

2. In a random sample of 24 high school students, the mean number of hours of sleep per night during the school week was 7.26 hours and the standard deviation was 1.19 hours. Assume the sleep times are normally distributed. Run the simulation for $n = 10$ so that at least 500 confidence intervals are generated. What proportion of the 95% Z confidence intervals and 95% T confidence intervals contain the population mean? Should you use a Z confidence interval or a T confidence interval for the mean number of hours of sleep? Explain.

6.3 Confidence Intervals for Population Proportions

WHAT YOU SHOULD LEARN

- How to find a point estimate for a population proportion
- How to construct a confidence interval for a population proportion
- How to determine the minimum sample size required when estimating a population proportion

Point Estimate for a Population Proportion ▸ Confidence Intervals for a Population Proportion ▸ Finding a Minimum Sample Size

POINT ESTIMATE FOR A POPULATION PROPORTION

Recall from Section 4.2 that the probability of success in a single trial of a binomial experiment is p. This probability is a **population proportion.** In this section, you will learn how to estimate a population proportion p using a confidence interval. As with confidence intervals for μ, you will start with a point estimate.

DEFINITION

The **point estimate for p,** the population proportion of successes, is given by the proportion of successes in a sample and is denoted by

$$\hat{p} = \frac{x}{n} \quad \text{Sample proportion}$$

where x is the number of successes in the sample and n is the sample size. The point estimate for the population proportion of failures is $\hat{q} = 1 - \hat{p}$. The symbols $\hat{p}$ and $\hat{q}$ are read as "p hat" and "q hat."

INSIGHT

In the first two sections, estimates were made for quantitative data. In this section, sample proportions are used to make estimates for qualitative data.

EXAMPLE 1

Finding a Point Estimate for p

In a survey of 1000 U.S. adults, 662 said that it is acceptable to check personal e-mail while at work. Find a point estimate for the population proportion of U.S. adults who say it is acceptable to check personal e-mail while at work. *(Adapted from Liberty Mutual)*

Solution

Using $n = 1000$ and $x = 662$,

$$\begin{aligned}\hat{p} &= \frac{x}{n}\\ &= \frac{662}{1000}\\ &= 0.662\\ &= 66.2\%\end{aligned}$$

So, the point estimate for the population proportion of U.S. adults who say it is acceptable to check personal e-mail while at work is 66.2%.

Try It Yourself 1

In a survey of 1006 U.S. adults, 181 said that Abraham Lincoln was the greatest president. Find a point estimate for the population proportion of U.S. adults who say Abraham Lincoln was the greatest president. *(Adapted from The Gallup Poll)*

a. *Identify* x and n.
b. Use x and n to find $\hat{p}$.

Answer: Page A40

PICTURING THE WORLD

In a recent year, there were about 9600 bird-aircraft collisions reported. A poll surveyed 2138 people about bird-aircraft collisions. Of those surveyed, 667 said that they are worried about bird-aircraft collisions. (Adapted from TripAdvisor)

Find a 90% confidence interval for the population proportion of people that are worried about bird-aircraft collisions.

▸ CONFIDENCE INTERVALS FOR A POPULATION PROPORTION

Constructing a confidence interval for a population proportion p is similar to constructing a confidence interval for a population mean. You start with a point estimate and calculate a margin of error.

DEFINITION

A ***c*-confidence interval for a population proportion *p*** is

$$\hat{p} - E < p < \hat{p} + E$$

where

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

The probability that the confidence interval contains p is c.

In Section 5.5, you learned that a binomial distribution can be approximated by a normal distribution if $np \geq 5$ and $nq \geq 5$. When $n\hat{p} \geq 5$ and $n\hat{q} \geq 5$, the sampling distribution of $\hat{p}$ is approximately normal with a mean of

$$\mu_{\hat{p}} = p$$

and a standard error of

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}.$$

GUIDELINES

Constructing a Confidence Interval for a Population Proportion

IN WORDS	IN SYMBOLS
1. Identify the sample statistics n and x.	
2. Find the point estimate $\hat{p}$.	$\hat{p} = \frac{x}{n}$
3. Verify that the sampling distribution of $\hat{p}$ can be approximated by a normal distribution.	$n\hat{p} \geq 5,\ n\hat{q} \geq 5$
4. Find the critical value z_c that corresponds to the given level of confidence c.	Use the Standard Normal Table or technology.
5. Find the margin of error E.	$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}$
6. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\hat{p} - E$ Right endpoint: $\hat{p} + E$ Interval: $\hat{p} - E < p < \hat{p} + E$

STUDY TIP

Here are instructions for constructing a confidence interval for a population proportion on a TI-83/84 Plus.

STAT

Choose the TESTS menu.

A: 1–PropZInt

Enter the values of *x*, *n*, and the level of confidence *c* (C-Level). Then select *Calculate*.

MINITAB and TI-83/84 Plus steps are shown on pages 352 and 353.

EXAMPLE 2

SC Report 26

Constructing a Confidence Interval for *p*

Use the data given in Example 1 to construct a 95% confidence interval for the population proportion of U.S. adults who say that it is acceptable to check personal e-mail while at work.

Solution

From Example 1, $\hat{p} = 0.662$. So,

$$\hat{q} = 1 - 0.662 = 0.338.$$

Using $n = 1000$, you can verify that the sampling distribution of $\hat{p}$ can be approximated by a normal distribution.

$$n\hat{p} = 1000 \cdot 0.662 = 662 > 5$$

$$n\hat{q} = 1000 \cdot 0.338 = 338 > 5$$

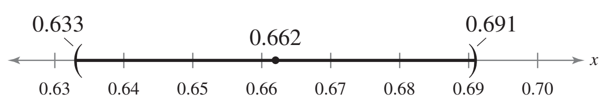
Using $z_c = 1.96$, the margin of error is

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{(0.662)(0.338)}{1000}} \approx 0.029.$$

The 95% confidence interval is as follows.

Left Endpoint	Right Endpoint
$\hat{p} - E \approx 0.662 - 0.029 = 0.633$	$\hat{p} + E \approx 0.662 + 0.029 = 0.691$

$$0.633 < p < 0.691$$



Interpretation With 95% confidence, you can say that the population proportion of U.S. adults who say that it is acceptable to check personal e-mail while at work is between 63.3% and 69.1%.

Try It Yourself 2

Use the data given in Try It Yourself 1 to construct a 90% confidence interval for the population proportion of U.S. adults who say that Abraham Lincoln was the greatest president.

a. *Find* $\hat{p}$ and $\hat{q}$.
b. *Verify* that the sampling distribution of $\hat{p}$ can be approximated by a normal distribution.
c. *Find* z_c and E.
d. Use $\hat{p}$ and E to find the *left* and *right endpoints* of the confidence interval.
e. *Interpret* the results.

Answer: Page A40

STUDY TIP

Notice in Example 2 that the confidence interval for the population proportion *p* is rounded to three decimal places. This *round-off rule* will be used throughout the text.

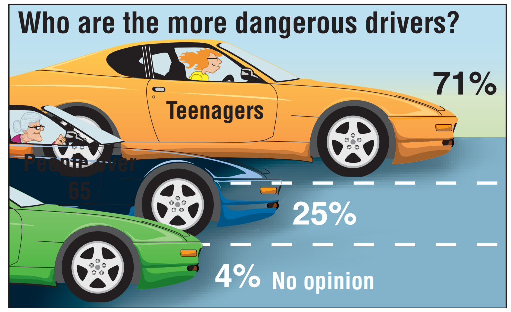
The confidence level of 95% used in Example 2 is typical of opinion polls. The result, however, is usually not stated as a confidence interval. Instead, the result of Example 2 would be stated as "66.2% with a margin of error of ±2.9%."

EXAMPLE 3

SC Report 27

Constructing a Confidence Interval for p

The graph shown at the right is from a survey of 498 U.S. adults. Construct a 99% confidence interval for the population proportion of U.S. adults who think that teenagers are the more dangerous drivers. *(Source: The Gallup Poll)*



INSIGHT

In Example 3, note that $n\hat{p} \geq 5$ and $n\hat{q} \geq 5$. So, the sampling distribution of $\hat{p}$ is approximately normal.

Solution

From the graph, $\hat{p} = 0.71$. So,

$$\hat{q} = 1 - 0.71$$
$$= 0.29.$$

Using these values and the values $n = 498$ and $z_c = 2.575$, the margin of error is

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\approx 2.575\sqrt{\frac{(0.71)(0.29)}{498}}$$

Use Table 4 in Appendix B to estimate that z_c is halfway between 2.57 and 2.58.

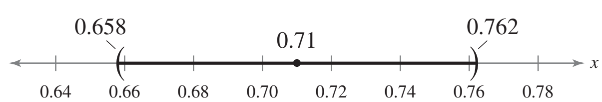
$$\approx 0.052.$$

To explore this topic further, see Activity 6.3 on page 336.

The 99% confidence interval is as follows.

Left Endpoint	Right Endpoint
$\hat{p} - E \approx 0.71 - 0.052 = 0.658$	$\hat{p} + E \approx 0.71 + 0.052 = 0.762$

$$0.658 < p < 0.762$$



Interpretation With 99% confidence, you can say that the population proportion of U.S. adults who think that teenagers are the more dangerous drivers is between 65.8% and 76.2%.

Try It Yourself 3

Use the data given in Example 3 to construct a 99% confidence interval for the population proportion of adults who think that people over 65 are the more dangerous drivers.

a. *Find* $\hat{p}$ and $\hat{q}$.
b. *Verify* that the sampling distribution of $\hat{p}$ can be approximated by a normal distribution.
c. *Find* z_c and E.
d. Use $\hat{p}$ and E to find the *left* and *right endpoints* of the confidence interval.
e. *Interpret* the results.

Answer: Page A40

FINDING A MINIMUM SAMPLE SIZE

One way to increase the precision of a confidence interval without decreasing the level of confidence is to increase the sample size.

INSIGHT

The reason for using 0.5 as the values of $\hat{p}$ and $\hat{q}$ when no preliminary estimate is available is that these values yield a maximum value of the product $\hat{p}\hat{q} = \hat{p}(1 - \hat{p})$. In other words, if you don't estimate the values of $\hat{p}$ and $\hat{q}$, you must pay the penalty of using a larger sample.

FINDING A MINIMUM SAMPLE SIZE TO ESTIMATE p

Given a c-confidence level and a margin of error E, the minimum sample size n needed to estimate p is

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2.$$

This formula assumes that you have preliminary estimates of $\hat{p}$ and $\hat{q}$. If not, use $\hat{p} = 0.5$ and $\hat{q} = 0.5$.

EXAMPLE 4

Determining a Minimum Sample Size

You are running a political campaign and wish to estimate, with 95% confidence, the population proportion of registered voters who will vote for your candidate. Your estimate must be accurate within 3% of the population proportion. Find the minimum sample size needed if (1) no preliminary estimate is available and (2) a preliminary estimate gives $\hat{p} = 0.31$. Compare your results.

Solution

1. Because you do not have a preliminary estimate of $\hat{p}$, use $\hat{p} = 0.5$ and $\hat{q} = 0.5$. Using $z_c = 1.96$ and $E = 0.03$, you can solve for n.

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2 = (0.5)(0.5)\left(\frac{1.96}{0.03}\right)^2 \approx 1067.11$$

Because n is a decimal, round up to the nearest whole number, 1068.

2. You have a preliminary estimate of $\hat{p} = 0.31$. So, $\hat{q} = 0.69$. Using $z_c = 1.96$ and $E = 0.03$, you can solve for n.

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2 = (0.31)(0.69)\left(\frac{1.96}{0.03}\right)^2 \approx 913.02$$

Because n is a decimal, round up to the nearest whole number, 914.

Interpretation With no preliminary estimate, the minimum sample size should be at least 1068 registered voters. With a preliminary estimate of $\hat{p} = 0.31$, the sample size should be at least 914 registered voters. So, you will need a larger sample size if no preliminary estimate is available.

Try It Yourself 4

You wish to estimate, with 90% confidence, the population proportion of females who refuse to eat leftovers. Your estimate must be accurate within 2% of the population proportion. Find the minimum sample size needed if (1) no preliminary estimate is available and (2) a previous survey found that 11% of females refuse to eat leftovers. *(Source: Consumer Reports National Research Center)*

a. *Identify* $\hat{p}$, $\hat{q}$, z_c, and E. If $\hat{p}$ is unknown, use 0.5.
b. Use $\hat{p}$, $\hat{q}$, z_c, and E to find the *minimum sample size* n.
c. *Determine* how many females should be included in the sample.

Answer: Page A40

6.3 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

True or False? *In Exercises 1 and 2, determine whether the statement is true or false. If it is false, rewrite it as a true statement.*

1. To estimate the value of p, the population proportion of successes, use the point estimate x.

2. The point estimate for the proportion of failures is $1 - \hat{p}$.

Finding $\hat{p}$ and $\hat{q}$ *In Exercises 3–6, let p be the population proportion for the given condition. Find point estimates of p and q.*

3. **Recycling** In a survey of 1002 U.S. adults, 752 say they recycle. *(Adapted from ABC News Poll)*

4. **Charity** In a survey of 2939 U.S. adults, 2439 say they have contributed to a charity in the past 12 months. *(Adapted from Harris Interactive)*

5. **Computers** In a survey of 11,605 parents, 4912 think that the government should subsidize the costs of computers for lower-income families. *(Adapted from DisneyFamily.com)*

6. **Vacation** In a survey of 1003 U.S. adults, 110 say they would go on vacation to Europe if cost did not matter. *(Adapted from The Gallup Poll)*

In Exercises 7–10, use the given confidence interval to find the margin of error and the sample proportion.

7. (0.905, 0.933)

8. (0.245, 0.475)

9. (0.512, 0.596)

10. (0.087, 0.263)

USING AND INTERPRETING CONCEPTS

Constructing Confidence Intervals *In Exercises 11 and 12, construct 90% and 95% confidence intervals for the population proportion. Interpret the results and compare the widths of the confidence intervals. If convenient, use technology to construct the confidence intervals.*

11. **Dental Visits** In a survey of 674 U.S. males ages 18–64, 396 say they have gone to the dentist in the past year. *(Adapted from National Center for Health Statistics)*

12. **Dental Visits** In a survey of 420 U.S. females ages 18–64, 279 say they have gone to the dentist in the past year. *(Adapted from National Center for Health Statistics)*

Constructing Confidence Intervals *In Exercises 13 and 14, construct a 99% confidence interval for the population proportion. Interpret the results.*

13. **Going Green** In a survey of 3110 U.S. adults, 1435 say they have started paying bills online in the last year. *(Adapted from Harris Interactive)*

14. **Seen a Ghost** In a survey of 4013 U.S. adults, 722 say they have seen a ghost. *(Adapted from Pew Research Center)*

15. Nail Polish In a survey of 7000 women, 4431 say they change their nail polish once a week. Construct a 95% confidence interval for the population proportion of women who change their nail polish once a week. *(Adapted from Essie Cosmetics)*

16. World Series In a survey of 891 U.S. adults who follow baseball in a recent year, 184 said that the Boston Red Sox would win the World Series. Construct a 90% confidence interval for the population proportion of U.S. adults who follow baseball who in a recent year said that the Boston Red Sox would win the World Series. *(Adapted from Harris Interactive)*

17. Alternative Energy You wish to estimate, with 95% confidence, the population proportion of U.S. adults who want more funding for alternative energy. Your estimate must be accurate within 4% of the population proportion.

(a) No preliminary estimate is available. Find the minimum sample size needed.

(b) Find the minimum sample size needed, using a prior study that found that 78% of U.S. adults want more funding for alternative energy. *(Source: Pew Research Center)*

(c) Compare the results from parts (a) and (b).

18. Reading Fiction You wish to estimate, with 99% confidence, the population proportion of U.S. adults who read fiction books. Your estimate must be accurate within 2% of the population proportion.

(a) No preliminary estimate is available. Find the minimum sample size needed.

(b) Find the minimum sample size needed, using a prior study that found that 47% of U.S. adults read fiction books. *(Source: National Endowment for the Arts)*

(c) Compare the results from parts (a) and (b).

19. Emergency Room Visits You wish to estimate, with 90% confidence, the population proportion of U.S. adults who made one or more emergency room visits in the past year. Your estimate must be accurate within 3% of the population proportion.

(a) No preliminary estimate is available. Find the minimum sample size needed.

(b) Find the minimum sample size needed, using a prior study that found that 20.1% of U.S. adults made one or more emergency room visits in the past year. *(Source: National Center for Health Statistics)*

(c) Compare the results from parts (a) and (b).

20. Ice Cream You wish to estimate, with 95% confidence, the population proportion of U.S. adults who say chocolate is their favorite ice cream flavor. Your estimate must be accurate within 5% of the population proportion.

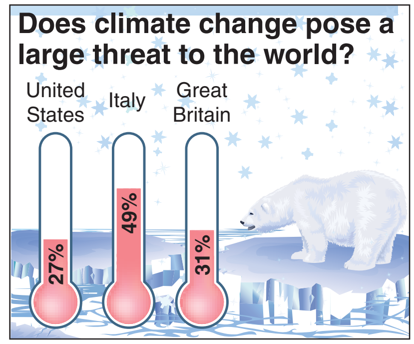
(a) No preliminary estimate is available. Find the minimum sample size needed.

(b) Find the minimum sample size needed, using a prior study that found that 27% of U.S. adults say that chocolate is their favorite ice cream flavor. *(Source: Harris Interactive)*

(c) Compare the results from parts (a) and (b).

Constructing Confidence Intervals

In Exercises 21 and 22, use the following information. The graph shows the results of a survey in which 1017 adults from the United States, 1060 adults from Italy, and 1126 adults from Great Britain were asked if they believe climate change poses a large threat to the world. (Source: Harris Interactive)



21. Global Warming Construct a 99% confidence interval for

(a) the population proportion of adults from the United States who say that climate change poses a large threat to the world.

(b) the population proportion of adults from Italy who say that climate change poses a large threat to the world.

(c) the population proportion of adults from Great Britain who say that climate change poses a large threat to the world.

22. Global Warming Determine whether it is possible that the following proportions are equal and explain your reasoning.

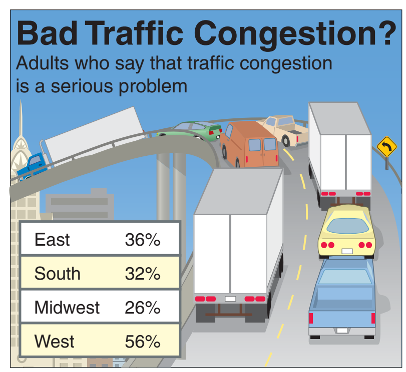
(a) The proportion of adults from Exercise 21(a) and the proportion of adults from Exercise 21(b).

(b) The proportion of adults from Exercise 21(b) and the proportion of adults from Exercise 21(c).

(c) The proportion of adults from Exercise 21(a) and the proportion of adults from Exercise 21(c).

Constructing Confidence Intervals

In Exercises 23 and 24, use the following information. The table shows the results of a survey in which separate samples of 400 adults each from the East, South, Midwest, and West were asked if traffic congestion is a serious problem in their community. (Adapted from Harris Interactive)



East	36%
South	32%
Midwest	26%
West	56%

23. South and West Construct a 95% confidence interval for the population proportion of adults

(a) from the South who say traffic congestion is a serious problem.

(b) from the West who say traffic congestion is a serious problem.

24. East and Midwest Construct a 95% confidence interval for the population proportion of adults

(a) from the East who say traffic congestion is a serious problem.

(b) from the Midwest who say traffic congestion is a serious problem.

25. Writing Is it possible that the proportions in Exercise 23 are equal? What if you used a 99% confidence interval? Explain your reasoning.

26. Writing Is it possible that the proportions in Exercise 24 are equal? What if you used a 99% confidence interval? Explain your reasoning.

SC *In Exercises 27 and 28, use StatCrunch to construct 90%, 95%, and 99% confidence intervals for the population proportion. Interpret the results and compare the widths of the confidence intervals.*

27. Congress In a survey of 1025 U.S. adults, 802 disapprove of the job Congress is doing. *(Adapted from The Gallup Poll)*

28. UFOs In a survey of 2303 U.S. adults, 734 believe in UFOs. *(Adapted from Harris Interactive)*

EXTENDING CONCEPTS

Newspaper Surveys *In Exercises 29 and 30, translate the newspaper excerpt into a confidence interval for p. Approximate the level of confidence.*

29. In a survey of 8451 U.S. adults, 31.4% said they were taking vitamin E as a supplement. The survey's margin of error is plus or minus 1%. *(Source: Decision Analyst, Inc.)*

30. In a survey of 1000 U.S. adults, 19% are concerned that their taxes will be audited by the Internal Revenue Service. The survey's margin of error is plus or minus 3%. *(Source: Rasmussen Reports)*

31. Why Check It? Why is it necessary to check that $n\hat{p} \geq 5$ and $n\hat{q} \geq 5$?

32. Sample Size The equation for determining the sample size

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2$$

can be obtained by solving the equation for the margin of error

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

for n. Show that this is true and justify each step.

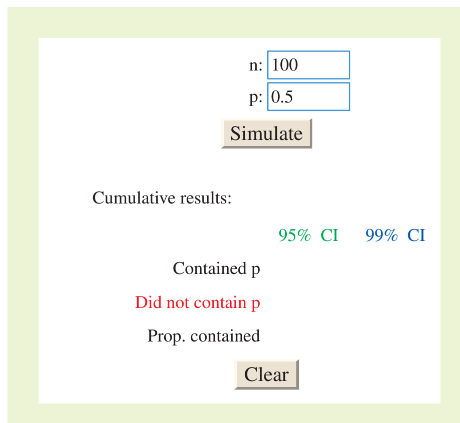
33. Maximum Value of $\hat{p}\hat{q}$ Complete the tables for different values of $\hat{p}$ and $\hat{q} = 1 - \hat{p}$. From the tables, which value of $\hat{p}$ appears to give the maximum value of the product $\hat{p}\hat{q}$?

$\hat{p}$	$\hat{q} = 1 - \hat{p}$	$\hat{p}\hat{q}$
0.0	1.0	0.00
0.1	0.9	0.09
0.2	0.8	
0.3		
0.4		
0.5		
0.6		
0.7		
0.8		
0.9		
1.0		

$\hat{p}$	$\hat{q} = 1 - \hat{p}$	$\hat{p}\hat{q}$
0.45		
0.46		
0.47		
0.48		
0.49		
0.50		
0.51		
0.52		
0.53		
0.54		
0.55		

ACTIVITY 6.3 Confidence Intervals for a Proportion

The *confidence intervals for a proportion* applet allows you to visually investigate confidence intervals for a population proportion. You can specify the sample size n and the population proportion p. When you click SIMULATE, 100 separate samples of size n will be selected from a population with a proportion of successes equal to p. For each of the 100 samples, a 95% confidence interval (in green) and a 99% confidence interval (in blue) are displayed in the plot at the right. Each of these intervals is computed using the standard normal approximation. If an interval does not contain the population proportion, it is displayed in red. Note that the 99% confidence interval is always wider than the 95% confidence interval. Additional simulations can be carried out by clicking SIMULATE multiple times. The cumulative number of times that each type of interval contains the population proportion is also shown. Press CLEAR to clear existing results and start a new simulation.



Explore

Step 1 Specify a value for n.
Step 2 Specify a value for p.
Step 3 Click SIMULATE to generate the confidence intervals.

Draw Conclusions

1. Run the simulation for $p = 0.6$ and $n = 10$, 20, 40, and 100. Clear the results after each trial. What proportion of the confidence intervals for each confidence level contains the population proportion? What happens to the proportion of confidence intervals that contains the population proportion for each confidence level as the sample size increases?

2. Run the simulation for $p = 0.4$ and $n = 100$ so that at least 1000 confidence intervals are generated. Compare the proportion of confidence intervals that contains the population proportion for each confidence level. Is this what you would expect? Explain.

6.4 Confidence Intervals for Variance and Standard Deviation

WHAT YOU SHOULD LEARN

- How to interpret the chi-square distribution and use a chi-square distribution table
- How to use the chi-square distribution to construct a confidence interval for the variance and standard deviation

The Chi-Square Distribution ▸ Confidence Intervals for σ^2 and σ

THE CHI-SQUARE DISTRIBUTION

In manufacturing, it is necessary to control the amount that a process varies. For instance, an automobile part manufacturer must produce thousands of parts to be used in the manufacturing process. It is important that the parts vary little or not at all. How can you measure, and consequently control, the amount of variation in the parts? You can start with a point estimate.

DEFINITION

The **point estimate for σ^2** is s^2 and the **point estimate for σ** is s. The most unbiased estimate for σ^2 is s^2.

You can use a *chi-square distribution* to construct a confidence interval for the variance and standard deviation.

STUDY TIP

The Greek letter χ is pronounced "$k\bar{\imath}$," which rhymes with the more familiar Greek letter π.

DEFINITION

If a random variable x has a normal distribution, then the distribution of

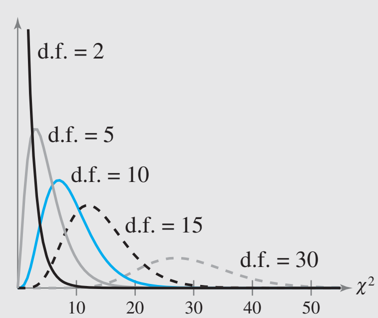
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

forms a **chi-square distribution** for samples of any size $n > 1$. Four properties of the chi-square distribution are as follows.

1. All chi-square values χ^2 are greater than or equal to 0.
2. The chi-square distribution is a family of curves, each determined by the degrees of freedom. To form a confidence interval for σ^2, use the χ^2-distribution with degrees of freedom equal to one less than the sample size.

   d.f. $= n - 1$ Degrees of freedom

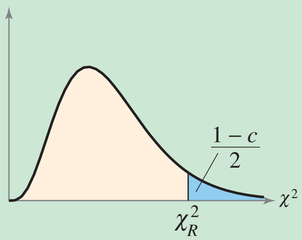
3. The area under each curve of the chi-square distribution equals 1.
4. Chi-square distributions are positively skewed.



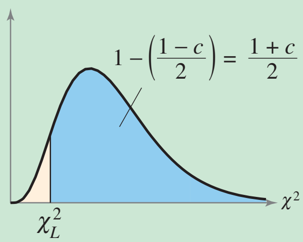
Chi-Square Distributions

STUDY TIP

For chi-square critical values with a *c*-confidence level, the following values are what you look up in Table 6 in Appendix B.

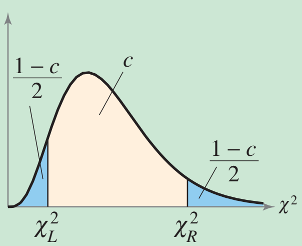


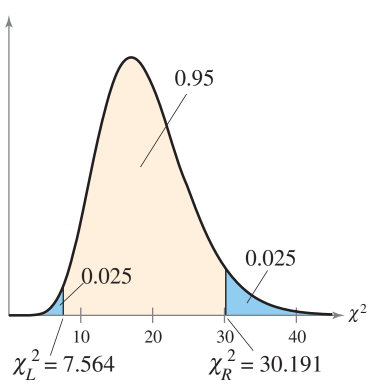
Area to the right of χ^2_R



Area to the right of χ^2_L

The result is that you can conclude that the area between the left and right critical values is *c*.





There are two critical values for each level of confidence. The value χ^2_R represents the right-tail critical value and χ^2_L represents the left-tail critical value. Table 6 in Appendix B lists critical values of χ^2 for various degrees of freedom and areas. Each area in the table represents the region under the chi-square curve to the *right* of the critical value.

EXAMPLE 1

Finding Critical Values for χ^2

Find the critical values χ^2_R and χ^2_L for a 95% confidence interval when the sample size is 18.

Solution

Because the sample size is 18, there are

$$\text{d.f.} = n - 1 = 18 - 1 = 17 \text{ degrees of freedom.}$$

The areas to the right of χ^2_R and χ^2_L are

$$\text{Area to right of } \chi^2_R = \frac{1-c}{2} = \frac{1-0.95}{2} = 0.025$$

and

$$\text{Area to right of } \chi^2_L = \frac{1+c}{2} = \frac{1+0.95}{2} = 0.975.$$

Part of Table 6 is shown. Using d.f. $= 17$ and the areas 0.975 and 0.025, you can find the critical values, as shown by the highlighted areas in the table.

Degrees of freedom	α							
	0.995	**0.99**	**0.975**	**0.95**	**0.90**	**0.10**	**0.05**	**0.025**
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170

χ^2_L χ^2_R

From the table, you can see that $\chi^2_R = 30.191$ and $\chi^2_L = 7.564$.

Interpretation So, 95% of the area under the curve lies between 7.564 and 30.191.

Try It Yourself 1

Find the critical values χ^2_R and χ^2_L for a 90% confidence interval when the sample size is 30.

a. Identify the *degrees of freedom* and the *level of confidence*.
b. *Find the areas* to the right of χ^2_R and χ^2_L.
c. *Use* Table 6 in Appendix B to find χ^2_R and χ^2_L.
d. *Interpret* the results.

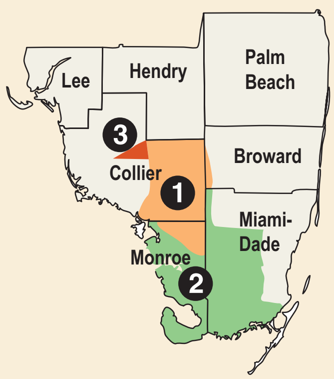
Answer: Page A40

CONFIDENCE INTERVALS FOR σ^2 AND σ

You can use the critical values χ^2_R and χ^2_L to construct confidence intervals for a population variance and standard deviation. The best point estimate for the variance is s^2 and the best point estimate for the standard deviation is s.

PICTURING THE WORLD

The Florida panther is one of the most endangered mammals on Earth. In the southeastern United States, the only breeding population (about 100) can be found on the southern tip of Florida. Most of the panthers live in (1) the Big Cypress National Preserve, (2) Everglades National Park, and (3) the Florida Panther National Wildlife Refuge, as shown on the map. In a recent study of 19 female panthers, it was found that the mean litter size was 2.4 kittens, with a standard deviation of 0.9. (Source: U.S. Fish & Wildlife Service)



Construct a 90% confidence interval for the standard deviation of the litter size for female Florida panthers. Assume the litter sizes are normally distributed.

DEFINITION

The c-confidence intervals for the population variance and standard deviation are as follows.

Confidence Interval for σ^2:

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

Confidence Interval for σ:

$$\sqrt{\frac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_L}}$$

The probability that the confidence intervals contain σ^2 or σ is c.

GUIDELINES

Constructing a Confidence Interval for a Variance and Standard Deviation

IN WORDS	IN SYMBOLS
1. Verify that the population has a normal distribution.	
2. Identify the sample statistic n and the degrees of freedom.	d.f. $= n - 1$
3. Find the point estimate s^2.	$s^2 = \dfrac{\Sigma(x-\bar{x})^2}{n-1}$
4. Find the critical values χ^2_R and χ^2_L that correspond to the given level of confidence c.	Use Table 6 in Appendix B.
5. Find the left and right endpoints and form the confidence interval for the population variance.	Left Endpoint, Right Endpoint: $\dfrac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \dfrac{(n-1)s^2}{\chi^2_L}$
6. Find the confidence interval for the population standard deviation by taking the square root of each endpoint.	$\sqrt{\dfrac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\dfrac{(n-1)s^2}{\chi^2_L}}$

EXAMPLE 2

SC Report 28

Constructing a Confidence Interval

You randomly select and weigh 30 samples of an allergy medicine. The sample standard deviation is 1.20 milligrams. Assuming the weights are normally distributed, construct 99% confidence intervals for the population variance and standard deviation.

Solution

The areas to the right of χ_R^2 and χ_L^2 are

$$\text{Area to right of } \chi_R^2 = \frac{1-c}{2} = \frac{1-0.99}{2} = 0.005$$

and

$$\text{Area to right of } \chi_L^2 = \frac{1+c}{2} = \frac{1+0.99}{2} = 0.995.$$

Using the values $n = 30$, d.f. $= 29$, and $c = 0.99$, the critical values χ_R^2 and χ_L^2 are

$$\chi_R^2 = 52.336 \qquad \text{and} \qquad \chi_L^2 = 13.121.$$

Using these critical values and $s = 1.20$, the confidence interval for σ^2 is as follows.

Left Endpoint

$$\frac{(n-1)s^2}{\chi_R^2} = \frac{(30-1)(1.20)^2}{52.336} \approx 0.80$$

Right Endpoint

$$\frac{(n-1)s^2}{\chi_L^2} = \frac{(30-1)(1.20)^2}{13.121} \approx 3.18$$

$$0.80 < \sigma^2 < 3.18$$

The confidence interval for σ is

$$\sqrt{\frac{(30-1)(1.20)^2}{52.336}} < \sigma < \sqrt{\frac{(30-1)(1.20)^2}{13.121}}$$

$$0.89 < \sigma < 1.78.$$

Interpretation With 99% confidence, you can say that the population variance is between 0.80 and 3.18, and the population standard deviation is between 0.89 and 1.78 milligrams.

Try It Yourself 2

Find the 90% and 95% confidence intervals for the population variance and standard deviation of the medicine weights.

a. Find the *critical values* χ_R^2 and χ_L^2 for each confidence interval.

b. Use n, s, χ_R^2, and χ_L^2 to find the *left* and *right endpoints* for each confidence interval for the population variance.

c. Find the *square roots* of the endpoints of each confidence interval.

d. *Specify* the 90% and 95% confidence intervals for the population variance and standard deviation.

Answer: Page A40

STUDY TIP

When a confidence interval for a population variance or standard deviation is computed, the general *round-off rule* is to round off to the same number of decimal places given for the sample variance or standard deviation.

6.4 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. Does a population have to be normally distributed in order to use the chi-square distribution?

2. What happens to the shape of the chi-square distribution as the degrees of freedom increase?

In Exercises 3–8, find the critical values χ^2_R and χ^2_L for the given confidence level c and sample size n.

3. $c = 0.90,\ n = 8$

4. $c = 0.99,\ n = 15$

5. $c = 0.95,\ n = 20$

6. $c = 0.98,\ n = 26$

7. $c = 0.99,\ n = 30$

8. $c = 0.80,\ n = 51$

USING AND INTERPRETING CONCEPTS

Constructing Confidence Intervals *In Exercises 9–24, assume each sample is taken from a normally distributed population and construct the indicated confidence intervals for (a) the population variance σ^2 and (b) the population standard deviation σ. Interpret the results.*

9. Vitamins To analyze the variation in weights of vitamin supplement tablets, you randomly select and weigh 14 tablets. The results (in milligrams) are shown. Use a 90% level of confidence.

500.000	499.995	500.010	499.997	500.015
499.988	500.000	499.996	500.020	500.002
499.998	499.996	500.003	500.000	

10. Cough Syrup You randomly select and measure the volumes of the contents of 15 bottles of cough syrup. The results (in fluid ounces) are shown. Use a 90% level of confidence.

4.211	4.246	4.269	4.241	4.260
4.293	4.189	4.248	4.220	4.239
4.253	4.209	4.300	4.256	4.290

11. Car Batteries The reserve capacities (in hours) of 18 randomly selected automotive batteries are shown. Use a 99% level of confidence. *(Adapted from Consumer Reports)*

1.70	1.60	1.94	1.58	1.74	1.60
1.86	1.72	1.38	1.46	1.64	1.49
1.55	1.70	1.75	0.88	1.77	2.07

12. Bolts You randomly select and measure the lengths of 17 bolts. The results (in inches) are shown. Use a 95% level of confidence.

1.286	1.138	1.240	1.132	1.381	1.137
1.300	1.167	1.240	1.401	1.241	1.171
1.217	1.360	1.302	1.331	1.383	

13. LCD TVs A magazine includes a report on the energy costs per year for 32-inch liquid crystal display (LCD) televisions. The article states that 14 randomly selected 32-inch LCD televisions have a sample standard deviation of \$3.90. Use a 99% level of confidence. *(Adapted from Consumer Reports)*

14. Digital Cameras A magazine includes a report on the prices of subcompact digital cameras. The article states that 11 randomly selected subcompact digital cameras have a sample standard deviation of \$109. Use an 80% level of confidence. *(Adapted from Consumer Reports)*

15. Spring Break As part of your spring break planning, you randomly select 10 hotels in Cancun, Mexico, and record the room rate for each hotel. The results are shown in the stem-and-leaf plot. Use a 98% level of confidence. *(Source: Expedia, Inc.)*

6	9
7	4
8	
9	0 9 9
10	
11	2
12	
13	6 9
14	9
15	0

Key: $7|4 = 74$

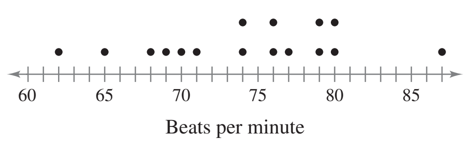
16. Cordless Drills The weights (in pounds) of a random sample of 14 cordless drills are shown in the stem-and-leaf plot. Use a 99% level of confidence. *(Adapted from Consumer Reports)*

3	4 6 9
4	6 8 9
5	1 3 4 5 7 9
6	0 1

Key: $3|4 = 3.4$

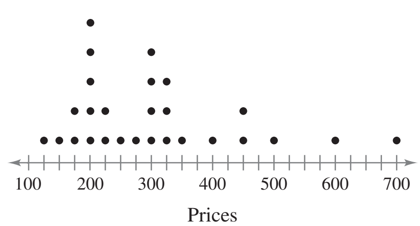
17. Pulse Rates The pulse rates of a random sample of 16 adults are shown in the dot plot. Use a 95% level of confidence.

Pulse Rates

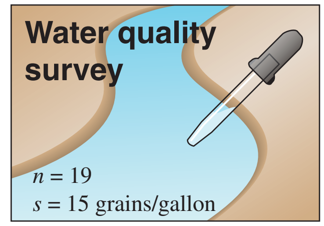


18. Blu-Ray™ Players The prices of a random sample of 27 Blu-ray™ players are shown in the dot plot. Use a 98% level of confidence. *(Adapted from Consumer Reports)*

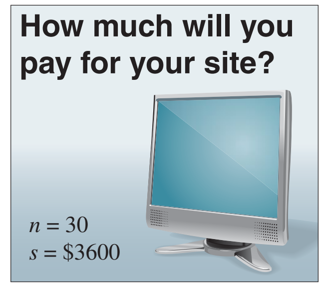
Blu-Ray™ Players



19. Water Quality As part of a water quality survey, you test the water hardness in several randomly selected streams. The results are shown in the figure. Use a 95% level of confidence.



20. Website Costs As part of a survey, you ask a random sample of business owners how much they would be willing to pay for a website for their company. The results are shown in the figure. Use a 90% level of confidence.



21. Annual Earnings The annual earnings of 14 randomly selected computer software engineers have a sample standard deviation of $3725. Use an 80% level of confidence. *(Adapted from U.S. Bureau of Labor Statistics)*

22. Annual Precipitation The average annual precipitations (in inches) of a random sample of 30 years in San Francisco, California have a sample standard deviation of 8.18 inches. Use a 98% level of confidence. *(Source: Golden Gate Weather Services)*

23. Waiting Times The waiting times (in minutes) of a random sample of 22 people at a bank have a sample standard deviation of 3.6 minutes. Use a 98% level of confidence.

24. Motorcycles The prices of a random sample of 20 new motorcycles have a sample standard deviation of $3900. Use a 90% level of confidence.

SC *In Exercises 25–28, use StatCrunch to help you construct the indicated confidence intervals for the population variance σ^2 and the population standard deviation σ. Assume each sample is taken from a normally distributed population.*

25. $c = 0.95, s^2 = 11.56, n = 30$

26. $c = 0.99, s^2 = 0.64, n = 7$

27. $c = 0.90, s = 35, n = 18$

28. $c = 0.97, s = 278.1, n = 45$

EXTENDING CONCEPTS

29. Vitamin Tablet Weights You are analyzing the sample of vitamin supplement tablets in Exercise 9. The population standard deviation of the tablets' weights should be less than 0.015 milligram. Does the confidence interval you constructed for σ suggest that the variation in the tablets' weights is at an acceptable level? Explain your reasoning.

30. Cough Syrup Bottle Contents You are analyzing the sample of cough syrup bottles in Exercise 10. The population standard deviation of the volumes of the bottles' contents should be less than 0.025 fluid ounce. Does the confidence interval you constructed for σ suggest that the variation in the volumes of the bottles' contents is at an acceptable level? Explain your reasoning.

31. In your own words, explain how finding a confidence interval for a population variance is different from finding a confidence interval for a population mean or proportion.

Statistics in the Real World

USES AND ABUSES

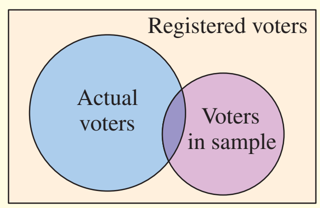
Uses

By now, you know that complete information about population parameters is often not available. The techniques of this chapter can be used to make interval estimates of these parameters so that you can make informed decisions.

From what you learned in this chapter, you know that point estimates (sample statistics) of population parameters are usually close but rarely equal to the actual values of the parameters they are estimating. Remembering this can help you make good decisions in your career and in everyday life. For instance, suppose the results of a survey tell you that 52% of the population plans to vote in favor of the rezoning of a portion of a town from residential to commercial use. You know that this is only a point estimate of the actual proportion that will vote in favor of rezoning. If the interval estimate is $0.49 < p < 0.55$, then you know this means it is possible that the item will not receive a majority vote.

Abuses

Unrepresentative Samples There are many ways that surveys can result in incorrect predictions. When you read the results of a survey, remember to question the sample size, the sampling technique, and the questions asked. For instance, suppose you want to know the proportion of people who will vote in favor of rezoning. From the diagram below, you can see that even if your sample is large enough, it may not consist of actual voters.



Using a small sample might be the only way to make an estimate, but be aware that a change in one data value may completely change the results. Generally, the larger the sample size, the more accurate the results will be.

Biased Survey Questions In surveys, it is also important to analyze the wording of the questions. For instance, the question about rezoning might be presented as: "Knowing that rezoning will result in more businesses contributing to school taxes, would you support the rezoning?"

EXERCISES

1. ***Unrepresentative Samples*** Find an example of a survey that is reported in a newspaper, magazine, or on a website. Describe different ways that the sample could have been unrepresentative of the population.

2. ***Biased Survey Questions*** Find an example of a survey that is reported in a newspaper, magazine, or on a website. Describe different ways that the survey questions could have been biased.

6 CHAPTER SUMMARY

What did you learn?	EXAMPLE(S)	REVIEW EXERCISES
Section 6.1		
■ How to find a point estimate and a margin of error $E = z_c \dfrac{\sigma}{\sqrt{n}}$ Margin of error	*1, 2*	*1, 2*
■ How to construct and interpret confidence intervals for the population mean $\bar{x} - E < \mu < \bar{x} + E$	*3–5*	*3–6*
■ How to determine the minimum sample size required when estimating μ	*6*	*7–10*
Section 6.2		
■ How to interpret the t-distribution and use a t-distribution table $t = \dfrac{(\bar{x} - \mu)}{(s/\sqrt{n})}$	*1*	*11–16*
■ How to construct confidence intervals when $n < 30$, the population is normally distributed, and σ is unknown $\bar{x} - E < \mu < \bar{x} + E, \quad E = t_c \dfrac{s}{\sqrt{n}}$	*2–4*	*17–26*
Section 6.3		
■ How to find a point estimate for a population proportion $\hat{p} = \dfrac{x}{n}$	*1*	*27–34*
■ How to construct a confidence interval for a population proportion $\hat{p} - E < p < \hat{p} + E, \quad E = z_c \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$	*2, 3*	*35–42*
■ How to determine the minimum sample size required when estimating a population proportion	*4*	*43, 44*
Section 6.4		
■ How to interpret the chi-square distribution and use a chi-square distribution table $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$	*1*	*45–48*
■ How to use the chi-square distribution to construct a confidence interval for the variance and standard deviation $\dfrac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \dfrac{(n-1)s^2}{\chi_L^2}, \quad \sqrt{\dfrac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\dfrac{(n-1)s^2}{\chi_L^2}}$	*2*	*49–52*

6 REVIEW EXERCISES

SECTION 6.1

In Exercises 1 and 2, find (a) the point estimate of the population mean μ and (b) the margin of error for a 90% confidence interval.

1. Waking times of 40 people who start work at 8:00 A.M. (in minutes past 5:00 A.M.)

135	145	95	140	135	95	110	50
90	165	110	125	80	125	130	110
25	75	65	100	60	125	115	135
95	90	140	40	75	50	130	85
100	160	135	45	135	115	75	130

2. Lengths of commutes to work of 32 people (in miles)

12	9	7	2	8	7	3	27
21	10	13	3	7	2	30	7
6	13	6	14	4	1	10	3
13	6	2	9	2	12	16	18

In Exercises 3 and 4, construct the indicated confidence interval for the population mean μ. If convenient, use technology to construct the confidence interval.

3. $c = 0.99, \bar{x} = 15.8, s = 0.85, n = 80$

4. $c = 0.95, \bar{x} = 7.675, s = 0.105, n = 55$

In Exercises 5 and 6, use the given confidence interval to find the margin of error and the sample mean.

5. $(20.75, 24.10)$

6. $(7.428, 7.562)$

In Exercises 7–10, determine the minimum sample size n needed to estimate μ.

7. Use the results of Exercise 1. Determine the minimum survey size that is necessary to be 95% confident that the sample mean waking time is within 10 minutes of the actual mean waking time.

8. Use the results of Exercise 1. Now suppose you want 99% confidence with a margin of error of 2 minutes. How many people would you need to survey?

9. Use the results of Exercise 2. Determine the minimum survey size that is necessary to be 95% confident that the sample mean length of commutes to work is within 2 miles of the actual mean length of commutes to work.

10. Use the results of Exercise 2. Now suppose you want 98% confidence with a margin of error of 0.5 mile. How many people would you need to survey?

SECTION 6.2

In Exercises 11–14, find the critical value t_c for the given confidence level c and sample size n.

11. $c = 0.80, n = 10$

12. $c = 0.95, n = 24$

13. $c = 0.98, n = 15$

14. $c = 0.99, n = 30$

15. Consider a 90% confidence interval for μ. Assume σ is not known. For which sample size, $n = 20$ or $n = 30$, is the critical value t_c larger?

16. Consider a 90% confidence interval for μ. Assume σ is not known. For which sample size, $n = 20$ or $n = 30$, is the confidence interval wider?

In Exercises 17–20, find the margin of error for μ.

17. $c = 0.90, s = 25.6, n = 16, \bar{x} = 72.1$

18. $c = 0.95, s = 1.1, n = 25, \bar{x} = 3.5$

19. $c = 0.98, s = 0.9, n = 12, \bar{x} = 6.8$

20. $c = 0.99, s = 16.5, n = 20, \bar{x} = 25.2$

In Exercises 21–24, construct the confidence interval for μ using the statistics from the given exercise. If convenient, use technology to construct the confidence interval.

21. Exercise 17

22. Exercise 18

23. Exercise 19

24. Exercise 20

25. In a random sample of 28 sports cars, the average annual fuel cost was $2218 and the standard deviation was $523. Construct a 90% confidence interval for μ. Assume the annual fuel costs are normally distributed. *(Adapted from U.S. Department of Energy)*

26. Repeat Exercise 25 using a 99% confidence interval.

SECTION 6.3

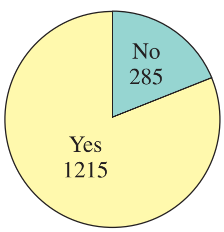


FIGURE FOR EXERCISE 27

In Exercises 27–34, let p be the proportion of the population who respond yes. Use the given information to find $\hat{p}$ and $\hat{q}$.

27. A survey asks 1500 U.S. adults if they will participate in the 2010 Census. The results are shown in the pie chart. *(Adapted from Pew Research Center)*

28. In a survey of 500 U.S. adults, 425 say they would trust doctors to tell the truth. *(Adapted from Harris Interactive)*

29. In a survey of 1023 U.S. adults, 552 say they have worked the night shift at some point in their lives. *(Adapted from CNN/Opinion Research)*

30. In a survey of 800 U.S. adults, 90 are making the minimum payment(s) on their credit card(s). *(Adapted from Cambridge Consumer Credit Index)*

31. In a survey of 1008 U.S. adults, 141 say the cost of health care is the most important financial problem facing their family today. *(Adapted from Gallup, Inc.)*

32. In a survey of 938 U.S. adults, 235 say the phrase "you know" is the most annoying conversational phrase. *(Adapted from Marist Poll)*

33. In a survey of 706 parents with kids 4 to 8 years old, 346 say that they know their state booster seat law. *(Adapted from Knowledge Networks, Inc.)*

34. In a survey of 2365 U.S. adults, 1230 say they worry most about missing deductions when filing their taxes. *(Adapted from USA TODAY)*

In Exercises 35–42, construct the indicated confidence interval for the population proportion p. If convenient, use technology to construct the confidence interval. Interpret the results.

35. Use the sample in Exercise 27 with $c = 0.95$.

36. Use the sample in Exercise 28 with $c = 0.99$.

37. Use the sample in Exercise 29 with $c = 0.90$.

38. Use the sample in Exercise 30 with $c = 0.98$.

39. Use the sample in Exercise 31 with $c = 0.99$.

40. Use the sample in Exercise 32 with $c = 0.90$.

41. Use the sample in Exercise 33 with $c = 0.80$.

42. Use the sample in Exercise 34 with $c = 0.98$.

43. You wish to estimate, with 95% confidence, the population proportion of U.S. adults who think they should be saving more money. Your estimate must be accurate within 5% of the population proportion.

(a) No preliminary estimate is available. Find the minimum sample size needed.

(b) Find the minimum sample size needed, using a prior study that found that 63% of U.S. adults think that they should be saving more money. *(Source: Pew Research Center)*

(c) Compare the results from parts (a) and (b).

44. Repeat Exercise 43 part (b), using a 99% confidence level and a margin of error of 2.5%. How does this sample size compare with your answer from Exercise 43 part (b)?

SECTION 6.4

In Exercises 45–48, find the critical values χ^2_R and χ^2_L for the given confidence level c and sample size n.

45. $c = 0.95, n = 13$

46. $c = 0.98, n = 25$

47. $c = 0.90, n = 8$

48. $c = 0.99, n = 10$

In Exercises 49–52, construct the indicated confidence intervals for the population variance σ^2 and the population standard deviation σ. Assume each sample is taken from a normally distributed population.

49. A random sample of the weights (in ounces) of 17 superzoom digital cameras is shown in the stem-and-leaf plot. Use a 95% level of confidence. *(Adapted from Consumer Reports)*

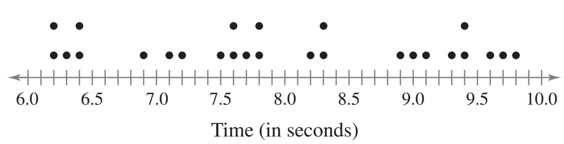
0	7 8 8 9
1	0 1 3 4 5 5 5 7 7 9
2	1 4
3	5

Key: $1|3 = 13$

50. Repeat Exercise 49 using a 99% level of confidence. Interpret the results and compare with Exercise 49.

51. A random sample of the acceleration times (in seconds) from 0 to 60 miles per hour for 26 sedans is shown in the dot plot. Use a 98% level of confidence. *(Adapted from Consumer Reports)*

Acceleration Times From 0–60 Miles Per Hour for Sedans



52. Repeat Exercise 51 using a 90% level of confidence. Interpret the results and compare with Exercise 51.

6 CHAPTER QUIZ

Take this quiz as you would take a quiz in class. After you are done, check your work against the answers given in the back of the book.

1. The following data set represents the amounts of time (in minutes) spent watching online videos each day for a random sample of 30 college students. *(Adapted from the Council for Research Excellence)*

5.0	6.25	8.0	5.5	4.75	4.5	7.2	6.6	5.8	5.5
4.2	5.4	6.75	9.8	8.2	6.4	7.8	6.5	5.5	6.0
3.8	6.75	9.25	10.0	9.6	7.2	6.4	6.8	9.8	10.2

(a) Find the point estimate of the population mean.

(b) Find the margin of error for a 95% level of confidence. Interpret the result.

(c) Construct a 95% confidence interval for the population mean. Interpret the results.

2. You want to estimate the mean time college students spend watching online videos each day. The estimate must be within 1 minute of the population mean. Determine the required sample size to construct a 99% confidence interval for the population mean. Assume the population standard deviation is 2.4 minutes.

3. The following data set represents the average number of minutes played for a random sample of professional basketball players in a recent season. *(Source: ESPN)*

35.9	33.8	34.7	31.5	33.2	29.1	30.7	31.2	36.1	34.9

(a) Find the sample mean and the sample standard deviation.

(b) Construct a 90% confidence interval for the population mean and interpret the results. Assume the population of the data set is normally distributed.

(c) Repeat part (b), assuming $\sigma = 5.25$ minutes per game. Interpret and compare the results.

4. In a random sample of seven aerospace engineers, the mean monthly income was \$6824 and the standard deviation was \$340. Assume the monthly incomes are normally distributed and construct a 95% confidence interval for the population mean monthly income for aerospace engineers. *(Adapted from U.S. Bureau of Labor Statistics)*

5. In a survey of 1383 U.S. adults, 1079 favor increasing federal funding for research on wind, solar, and hydrogen technology. *(Adapted from Pew Research Center)*

(a) Find a point estimate for the population proportion p of those in favor of increasing federal funding for research on wind, solar, and hydrogen technology.

(b) Construct a 90% confidence interval for the population proportion.

(c) Find the minimum sample size needed to estimate the population proportion at the 99% confidence level in order to ensure that the estimate is accurate within 4% of the population proportion.

6. Refer to the data set in Exercise 1. Assume the population of times spent watching online videos each day is normally distributed.

(a) Construct a 95% confidence interval for the population variance.

(b) Construct a 95% confidence interval for the population standard deviation.

PUTTING IT ALL TOGETHER

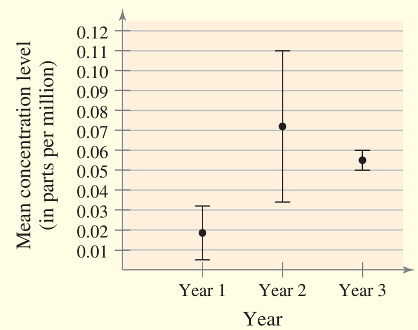
Real Statistics — Real Decisions

In 1974, the Safe Drinking Water Act was passed "to protect public health by regulating the nation's public drinking water supply." In accordance with the act, the Environmental Protection Agency (EPA) has regulations that limit the levels of contaminants in drinking water supplied by water utilities. These utilities are required to supply water quality reports to their customers annually. These reports discuss the source of the water, its treatment, and the results of water quality monitoring that is performed daily. The results of this monitoring indicate whether or not drinking water is healthy enough for consumption.

A water department tests for contaminants at water treatment plants and at customers' taps. These regulated parameters include microorganisms, organic chemicals, and inorganic chemicals. For instance, cyanide is an inorganic chemical that is regulated. Its presence in drinking water is the result of discharges from steel, plastics, and fertilizer factories. The maximum contaminant level for cyanide is set at 0.2 part per million.

You work for a city's water department and are interpreting the results shown in the graph at the right. The graph shows the point estimates for the population mean concentration and the 95% confidence intervals for μ for cyanide over a three-year period. The data are based on random water samples taken by the city's three water treatment plants.

Cyanide



EXERCISES

1. *Interpreting the Results*

Use the graph to decide if there has been a change in the mean concentration level of cyanide for the given years. Explain your reasoning.

(a) From Year 1 to Year 2

(b) From Year 2 to Year 3

(c) From Year 1 to Year 3

2. *What Can You Conclude?*

Using the results of Exercise 1, what can you conclude about the concentrations of cyanide in the drinking water?

3. *What Do You Think?*

The confidence interval for Year 2 is much larger than the other years. What do you think may have caused this larger confidence level?

4. *How Do You Think They Did It?*

How do you think the water department constructed the 95% confidence intervals for the population mean concentration of cyanide in the water? Do the following to answer the question. (You do not need to make any calculations.)

(a) What sampling distribution do you think they used? Why?

(b) Do you think they used the population standard deviation in calculating the margin of error? Why or why not? If not, what could they have used?

TECHNOLOGY

MINITAB | EXCEL | TI-83/84 PLUS

THE GALLUP ORGANIZATION

WWW.GALLUP.COM

MOST ADMIRED POLLS

Since 1946, the Gallup Organization has conducted a "most admired" poll. The methodology for the 2009 poll is described at the right.

> ***Survey Question***
> *What man* that you have heard or read about, living today in any part of the world, do you admire most? And who is your second choice?*

Reprinted with permission from GALLUP.

*Survey respondents are asked an identical question about most admired woman.

"Results are based on telephone interviews with 1,025 national adults, aged 18 and older, conducted Dec. 11–13, 2009. For results based on the total sample of national adults, one can say with 95% confidence that the maximum margin of sampling error is ±4 percentage points. Interviews are conducted with respondents on land-line telephones (for respondents with a land-line telephone) and cellular phones (for respondents who are cell-phone only). In addition to sampling error, question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of public opinion polls."

EXERCISES

1. In 2009, the most named man was Barack Obama at 30%. Use a technology tool to find a 95% confidence interval for the population proportion that would have chosen Barack Obama.

2. In 2009, the most named woman was Hillary Clinton at 16%. Use a technology tool to find a 95% confidence interval for the population proportion that would have chosen Hillary Clinton.

3. Do the confidence intervals you obtained in Exercises 1 and 2 agree with the statement issued by the Gallup Organization that the margin of error is ±4%? Explain.

4. The second most named woman was Sarah Palin, who was named by 15% of the people in the sample. Use a technology tool to find a 95% confidence interval for the population proportion that would have chosen Sarah Palin.

5. Use a technology tool to simulate a most admired poll. Assume that the actual population proportion who most admire Sarah Palin is 18%. Run the simulation several times using $n = 1025$.

(a) What was the least value you obtained for $\hat{p}$?

(b) What was the greatest value you obtained for $\hat{p}$?

MINITAB

Number of rows of data to generate: 200

Store in column(s): C1

Number of trials: 1025

Event probability: 0.18

6. Is it probable that the population proportion who most admire Sarah Palin is 18% or greater? Explain your reasoning.

Extended solutions are given in the *Technology Supplement.*
Technical instruction is provided for MINITAB, Excel, and the TI-83/84 Plus.

6 USING TECHNOLOGY TO CONSTRUCT CONFIDENCE INTERVALS

Here are some MINITAB and TI-83/84 Plus printouts for some examples in this chapter. Answers may be slightly different because of rounding.

(See Example 3, page 307.)

140	105	130	97	80	165	232	110	214	201	122	98	65	88
154	133	121	82	130	211	153	114	58	77	51	247	236	109
126	132	125	149	122	74	59	218	192	90	117	105		

Display Descriptive Statistics...
Store Descriptive Statistics...
Graphical Summary...
1-Sample Z...
1-Sample t...
2-Sample t...
Paired t...
1 Proportion...
2 Proportions...

MINITAB

One-Sample Z: Friends

The assumed standard deviation = 53

Variable	N	Mean	StDev	SE Mean	95% CI
Friends	40	130.80	52.63	8.38	(114.38, 147.22)

(See Example 2, page 320.)

Display Descriptive Statistics...
Store Descriptive Statistics...
Graphical Summary...
1-Sample Z...
1-Sample t...
2-Sample t...
Paired t...
1 Proportion...
2 Proportions...

MINITAB

One-Sample T

N	Mean	StDev	SE Mean	95% CI
16	162.00	10.00	2.50	(156.67, 167.33)

(See Example 2, page 329.)

Display Descriptive Statistics...
Store Descriptive Statistics...
Graphical Summary...
1-Sample Z...
1-Sample t...
2-Sample t...
Paired t...
1 Proportion...
2 Proportions...

MINITAB

Test and CI for One Proportion

Sample	X	N	Sample p	95% CI
1	662	1000	0.662000	(0.631738, 0.691305)

(See Example 5, page 309.)

TI-83/84 PLUS

```
EDIT CALC TESTS
1: Z-Test...
2: T-Test...
3: 2-SampZTest...
4: 2-SampTTest...
5: 1-PropZTest...
6: 2-PropZTest...
7↓ ZInterval...
```

TI-83/84 PLUS

```
ZInterval
  Inpt: Data Stats
  s: 1.5
  x̄: 22.9
  n: 20
  C-Level: .9
  Calculate
```

TI-83/84 PLUS

```
ZInterval
  (22.348, 23.452)
  x̄= 22.9
  n= 20
```

(See Example 3, page 321.)

TI-83/84 PLUS

```
EDIT CALC TESTS
2↑T-Test...
3: 2-SampZTest...
4: 2-SampTTest...
5: 1-PropZTest...
6: 2-PropZTest...
7: ZInterval...
8↓ TInterval...
```

TI-83/84 PLUS

```
TInterval
  Inpt: Data Stats
  x̄: 9.75
  Sx: 2.39
  n: 20
  C-Level: .99
  Calculate
```

TI-83/84 PLUS

```
TInterval
  (8.2211, 11.279)
  x̄= 9.75
  Sx= 2.39
  n= 20
```

(See Example 2, page 329.)

TI-83/84 PLUS

```
EDIT CALC TESTS
5↑1-PropZTest...
6: 2-PropZTest...
7: ZInterval...
8: TInterval...
9: 2-SampZInt...
0: 2-SampTInt...
A↓ 1-PropZInt...
```

TI-83/84 PLUS

```
1-PropZInt
  x: 662
  n: 1000
  C-Level: .95
  Calculate
```

TI-83/84 PLUS

```
1-PropZInt
  (.63268, .69132)
  p̂= 0.662
  n= 1000
```