Chapter 8

Estimating with Confidence

case study

Need Help? Give Us a Call!

If your cable television goes out, you phone the cable company to get it fixed. Does a real person answer your call? These days, probably not. It is far more likely that you will get an automated response. You will probably be offered several options, such as: to order cable service, press 1; for questions about your bill, press 2; to add new channels, press 3; (and finally) to speak with a customer service agent, press 4. Customers will get frustrated if they have to wait too long before speaking to a live person. So companies try hard to minimize the time required to connect to a customer service representative.

A large bank decided to study the call response times in its customer service department. The bank's goal was to have a representative answer an incoming call in less than 30 seconds. Figure 8.1 is a histogram of the response times in a random sample of 241 calls to the bank's customer service center in a given month. What does the graph suggest about how well the bank is meeting its goal?

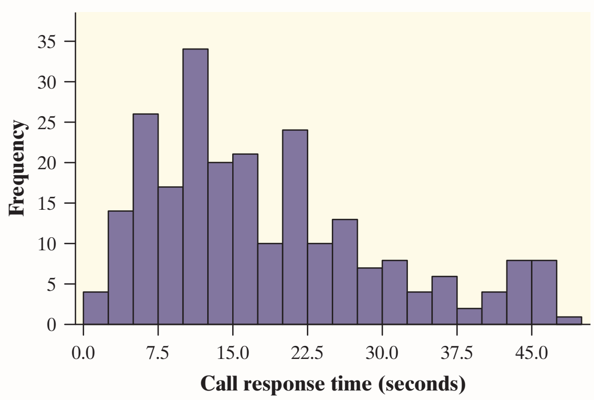


FIGURE 8.1 Histogram showing the response time (in seconds) at a bank's customer service center for a random sample of 241 calls.

By the end of this chapter, you will be able to use the sample data to make inferences about the proportion p of all calls to the customer service department that are answered within 30 seconds and the mean call response time μ.

Introduction

How long does a new model of laptop battery last? What proportion of college undergraduates have engaged in binge drinking? How much does the weight of a quarter-pound hamburger at a fast-food restaurant vary after cooking? These are the types of questions we would like to be able to answer.

It wouldn't be practical to determine the lifetime of *every* laptop battery, to ask *all* college undergraduates about their drinking habits, or to weigh *every* burger after cooking. Instead, we choose a sample of individuals (batteries, college students, burgers) to represent the population and collect data from those individuals. Our goal in each case is to use a sample statistic to estimate an unknown population parameter. From what we learned in Chapter 4, if we randomly select the sample, we should be able to generalize our results to the population of interest.

We cannot be certain that our conclusions are correct—a different sample would probably yield a different estimate. Statistical inference uses the language of probability to express the strength of our conclusions. Probability allows us to take chance variation due to random selection or random assignment into account. The following Activity gives you an idea of what lies ahead.

ACTIVITY | The Mystery Mean

MATERIALS:
TI-83/84 or TI-89 with display capability

In this Activity, each team of three to four students will try to estimate the mystery value of the population mean μ that your teacher entered before class.[1]

1. Before class, your teacher will store a value of μ (represented by M) in the display calculator. The teacher will then clear the home screen so you can't see the value of M.
2. With the class watching, the teacher will execute the following command: `mean(randNorm(M,20,16))`.

This tells the calculator to choose an SRS of 16 observations from a Normal population with mean M and standard deviation 20 and then compute the mean $\bar{x}$ of those 16 sample values. Is the sample mean shown likely to be equal to the mystery mean M? Why or why not?

3. Now for the challenge! Your group must determine an interval of *reasonable* values for the population mean μ. Use the result from Step 2 and what you learned about sampling distributions in the previous chapter.
4. Share your team's results with the class.

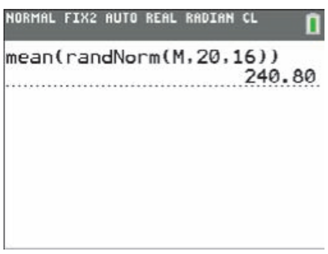
In this chapter and the next, we will meet the two most common types of formal statistical inference. Chapter 8 concerns *confidence intervals* for estimating the value of a parameter. Chapter 9 presents *significance tests*, which assess the evidence for a claim about a parameter. Both types of inference are based on the sampling distributions of statistics. That is, both report probabilities that state *what would happen if we used the inference method many times.*

Section 8.1 examines the idea of a confidence interval. We start by presenting the reasoning of confidence intervals in a general way that applies to estimating any unknown parameter. In Section 8.2, we show how to estimate a population proportion. Section 8.3 focuses on confidence intervals for a population mean.

8.1 Confidence Intervals: The Basics

WHAT YOU WILL LEARN By the end of the section, you should be able to:

- Determine the point estimate and margin of error from a confidence interval.
- Interpret a confidence interval in context.
- Interpret a confidence level in context.
- Describe how the sample size and confidence level affect the length of a confidence interval.
- Explain how practical issues like nonresponse, undercoverage, and response bias can affect the interpretation of a confidence interval.



Mr. Schiel's class did the mystery mean Activity from the Introduction. The TI screen shot displays the information that the students received about the unknown population mean μ. Here is a summary of what the class said about the calculator output:

- The population distribution is Normal and its standard deviation is $\sigma = 20$.
- A simple random sample of $n = 16$ observations was taken from this population.
- The sample mean is $\bar{x} = 240.80$.

If we had to give a single number to estimate the value of M that Mr. Schiel chose, what would it be? Such a value is known as a **point estimate.** How about 240.80? That makes sense, because the sample mean $\bar{x}$ is an unbiased estimator of the population mean μ. We are using the statistic $\bar{x}$ as a **point estimator** of the parameter μ.

DEFINITION: Point estimator and point estimate

A **point estimator** is a statistic that provides an estimate of a population parameter. The value of that statistic from a sample is called a **point estimate.**

As we saw in Chapter 7, the ideal point estimator will have no bias and low variability. Here's an example involving some of the more common point estimators.

EXAMPLE From Batteries to Smoking

Point estimators

PROBLEM: In each of the following settings, determine the point estimator you would use and calculate the value of the point estimate.

(a) Quality control inspectors want to estimate the mean lifetime μ of the AA batteries produced in an hour at a factory. They select a random sample of 50 batteries during each hour of production and then drain them under conditions that mimic normal use. Here are the lifetimes (in hours) of the batteries from one such sample:

16.73	15.60	16.31	17.57	16.14	17.28	16.67	17.28	17.27	17.50	15.46	16.50	16.19
15.59	17.54	16.46	15.63	16.82	17.16	16.62	16.71	16.69	17.98	16.36	17.80	16.61
15.99	15.64	17.20	17.24	16.68	16.55	17.48	15.58	17.61	15.98	16.99	16.93	16.01
17.54	17.41	16.91	16.60	16.78	15.75	17.31	16.50	16.72	17.55	16.46		

(b) What proportion p of U.S. high school students smoke? The 2011 Youth Risk Behavioral Survey questioned a random sample of 15,425 students in grades 9 to 12. Of these, 2792 said they had smoked cigarettes at least one day in the past month.

(c) The quality control inspectors in part (a) want to investigate the variability in battery lifetimes by estimating the population variance σ^2.

SOLUTION:

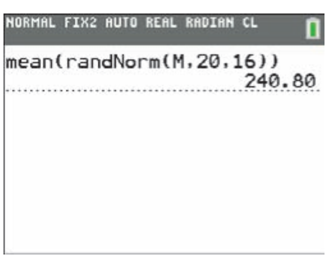
(a) Use the sample mean $\bar{x}$ as a point estimator for the population mean μ. For these data, our point estimate is $\bar{x} = 16.718$ hours.

(b) Use the sample proportion $\hat{p}$ as a point estimator for the population proportion p. For this survey, our point estimate is $\hat{p} = \dfrac{2792}{15{,}425} = 0.181$.

(c) Use the sample variance s_x^2 as a point estimator for the population variance σ^2. For the battery life data, our point estimate is $s_x^2 = 0.441$ hours^2.

For Practice *Try Exercises* **1** *and* **3**

The Idea of a Confidence Interval



Is the value of the population mean μ that Mr. Schiel entered in his calculator exactly 240.80? Probably not. Because $\bar{x} = 240.80$, we guess that μ is "somewhere around 240.80." How close to 240.80 is μ likely to be?

To answer this question, we ask another: *How would the sample mean $\bar{x}$ vary if we took many SRSs of size 16 from this same population?*

The sampling distribution of $\bar{x}$ describes how the values of $\bar{x}$ vary in repeated samples. Recall the facts about this sampling distribution from Chapter 7:

Shape: Because the population distribution is Normal, so is the sampling distribution of $\bar{x}$. Thus, the *Normal/Large Sample condition* is met.
Center: The mean of the sampling distribution of $\bar{x}$ is the same as the unknown mean μ of the entire population. That is, $\mu_{\bar{x}} = \mu$.
Spread: The standard deviation of the sampling distribution of $\bar{x}$ for samples of size $n = 16$ is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{16}} = 5$$

because the *10% condition* is met—we are sampling from an infinite population in this case.

Figure 8.2 summarizes these facts.

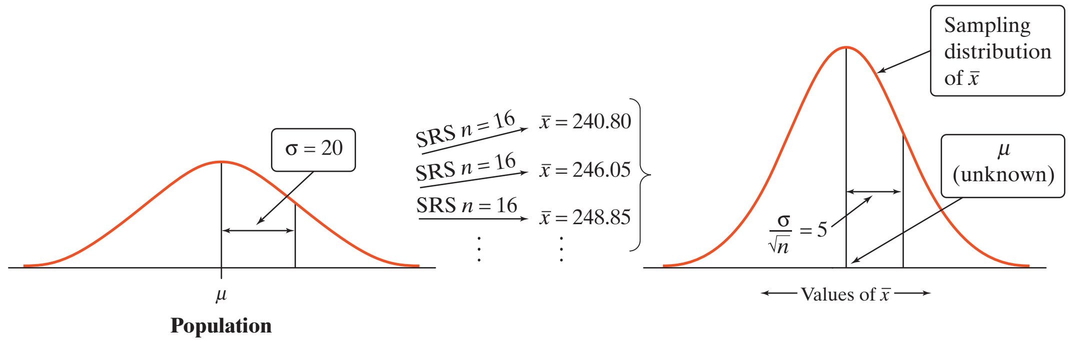
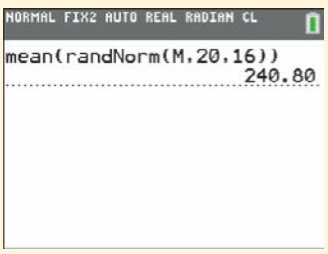


FIGURE 8.2 The sampling distribution of the mean score $\bar{x}$ for SRSs of 16 observations from a Normally distributed population with unknown mean μ and standard deviation $\sigma = 20$.

The next example gives the reasoning of statistical estimation in a nutshell.

EXAMPLE The Mystery Mean

Moving beyond a point estimate



When Mr. Schiel's class discussed the results of the mystery mean Activity, students used the following logic to come up with an "interval estimate" for the unknown population mean μ.

1. The sample mean $\bar{x} = 240.80$ is our point estimate for μ. We don't expect $\bar{x}$ to be exactly equal to μ, so we want to say how precise this estimate is.
2. In repeated samples, the values of $\bar{x}$ follow a Normal distribution with mean μ and standard deviation 5, as in Figure 8.2.
3. The 95 part of the 68–95–99.7 rule for Normal distributions says that $\bar{x}$ is within $2(5) = 10$ (that's 2 standard deviations) of the population mean μ in about 95% of all samples of size $n = 16$. See Figure 8.3.
4. Whenever $\bar{x}$ is within 10 points of μ, μ is within 10 points of $\bar{x}$. This happens in about 95% of all possible samples. So the interval from $\bar{x} - 10$ to $\bar{x} + 10$ "captures" the population mean μ in about 95% of all samples of size 16.
5. If we estimate that μ lies somewhere in the interval from

$$\bar{x} - 10 = 240.80 - 10 = 230.80$$

to

$$\bar{x} + 10 = 240.80 + 10 = 250.80$$

we'd be calculating this interval using a method that captures the true μ in about 95% of all possible samples of this size.

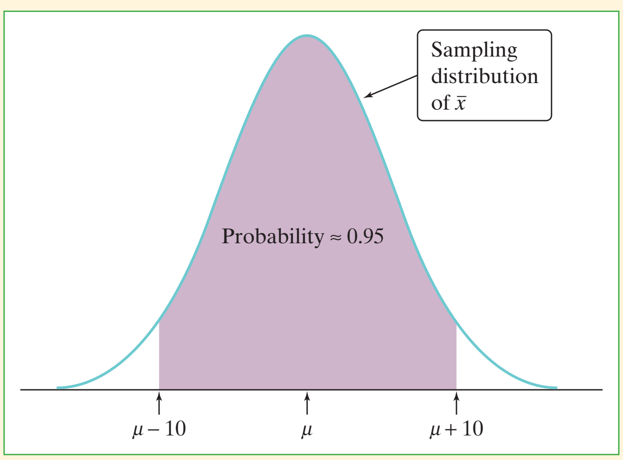


FIGURE 8.3 In about 95% of all samples, $\bar{x}$ lies within ± 10 of the unknown population mean μ. So μ also lies within ± 10 of $\bar{x}$ in those samples.

The big idea is that the sampling distribution of $\bar{x}$ tells us how close to μ the sample mean $\bar{x}$ is likely to be. Statistical estimation just turns that information around to say how close to $\bar{x}$ the unknown population mean μ is likely to be. In the mystery mean example, the value of μ is usually within $2(5) = 10$ of $\bar{x}$ for SRSs of size 16. Because the class's sample mean was $\bar{x} = 240.80$, the interval 240.80 ± 10 gives an approximate **95% confidence interval** for μ.

There are several ways to write a confidence interval. We can give the interval for Mr. Schiel's mystery mean as 240.80 ± 10, as 230.80 to 250.80, or as (230.80, 250.80).

All the confidence intervals we will meet have a form similar to this:

$$\text{point estimate} \pm \text{margin of error}$$

The point estimate ($\bar{x} = 240.80$ in our example) is our best guess for the value of the unknown parameter. The **margin of error,** 10, shows how close we believe our guess is, based on the variability of the estimate in repeated SRSs of size 16. We say that our **confidence level** is about 95% because the interval $\bar{x} \pm 10$ catches the unknown parameter in about 95% of all possible samples.

A confidence interval is sometimes referred to as an *interval estimate*. This is consistent with our earlier use of the term *point estimate*.

DEFINITION: Confidence interval, margin of error, confidence level

A ***C*% confidence interval** gives an interval of plausible values for a parameter. The interval is calculated from the data and has the form

$$\text{point estimate} \pm \text{margin of error}$$

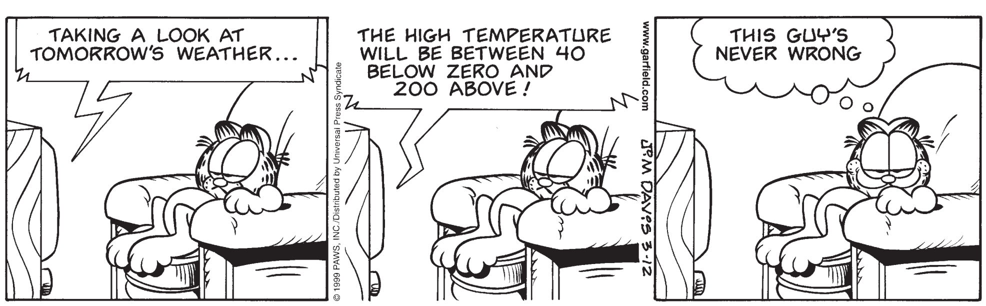
The difference between the point estimate and the true parameter value will be less than the **margin of error** in *C*% of all samples.

The **confidence level *C*** gives the overall success rate of the method for calculating the confidence interval. That is, in *C*% of all possible samples, the method would yield an interval that captures the true parameter value.

The interval from 230.80 to 250.80 gives the set of plausible values for Mr. Schiel's mystery mean μ with 95% confidence. We wouldn't be surprised if any of the values in this interval turned out to be the actual value of μ.

Plausible does not mean the same thing as possible. You could argue that just about any value of a parameter is *possible*. A *plausible* value of a parameter is a reasonable or believable value based on the data.

There is a trade-off between the confidence level and the amount of information provided by a confidence interval, as the cartoon below illustrates. We usually choose a confidence level of 90% or higher because we want to be quite sure of our conclusions. The most common confidence level is 95%.



Interpreting Confidence Intervals and Confidence Levels

Our 95% confidence interval for Mr. Schiel's mystery mean was (230.80, 250.80). How do we interpret this interval? We say, "We are 95% confident that the interval from 230.80 to 250.80 captures the mystery mean chosen by Mr. Schiel."

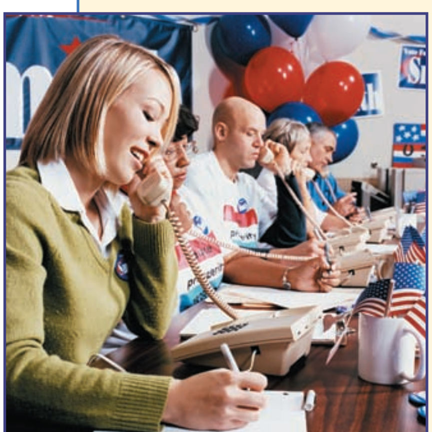
INTERPRETING CONFIDENCE INTERVALS

To interpret a $C\%$ confidence interval for an unknown parameter, say, "We are $C\%$ confident that the interval from _____ to _____ captures the [parameter in context]."

Here's an example that involves interpreting a confidence interval for a proportion.

EXAMPLE Who Will Win the Election?

Interpreting a confidence interval



Two weeks before a presidential election, a polling organization asked a random sample of registered voters the following question: "If the presidential election were held today, would you vote for candidate A or candidate B?" Based on this poll, the 95% confidence interval for the population proportion who favor candidate A is (0.48, 0.54).

PROBLEM:

(a) Interpret the confidence interval.

(b) What is the point estimate that was used to create the interval? What is the margin of error?

(c) Based on this poll, a political reporter claims that the majority of registered voters favor candidate A. Use the confidence interval to evaluate this claim.

SOLUTION:

(a) We are 95% confident that the interval from 0.48 to 0.54 captures the true proportion of all registered voters who favor candidate A in the election.

(b) A confidence interval has the form

$$\text{point estimate} \pm \text{margin of error}$$

The point estimate is at the midpoint of the interval. Here, the point estimate is $\hat{p} = 0.51$. The margin of error gives the distance from the point estimate to either end of the interval. So the margin of error for this interval is 0.03.

(c) Any value from 0.48 to 0.54 is a plausible value for the population proportion p that favors candidate A. Because there are plausible values of p less than 0.50, the confidence interval does not give convincing evidence to support the reporter's claim that a majority (more than 50%) of registered voters favor candidate A.

For Practice *Try Exercise* 9

The following Activity gives you a chance to explore the meaning of the confidence *level*.

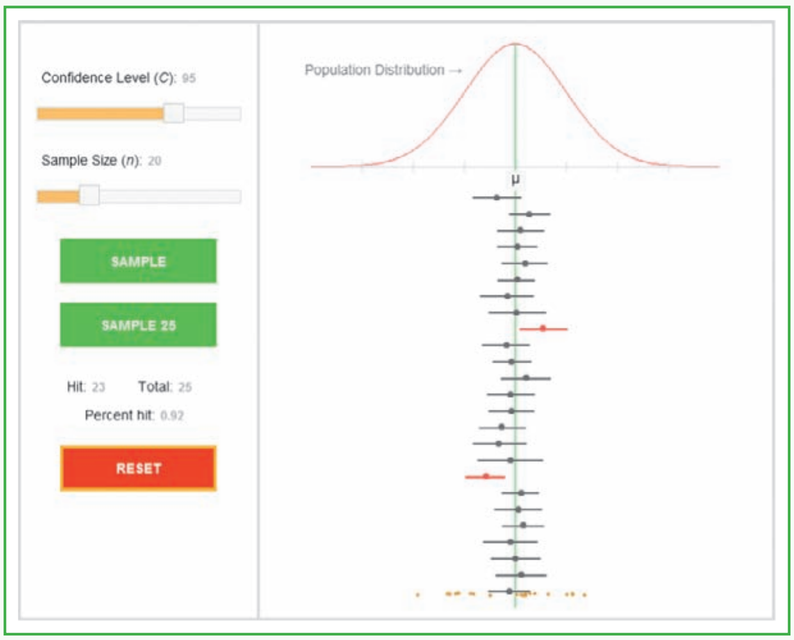
ACTIVITY | The *Confidence Intervals* Applet

MATERIALS:
Computer with Internet connection and display capability

APPLET

The *Confidence Intervals* applet at the book's Web site will quickly generate many confidence intervals. In this Activity, you will use the applet to investigate the idea of a confidence level.

1. Go to www.whfreeman.com/tps5e and launch the applet. Use the default settings: confidence level 95% and sample size $n = 20$.
2. Click "Sample" to choose an SRS and display the resulting confidence interval. Did the interval capture the population mean μ (what the applet calls a "hit")? Do this a total of 10 times. How many of the intervals captured the population mean μ? *Note*: So far, you have used the applet to take 10 SRSs, each of size $n = 20$. Be sure you understand the difference between sample size and the number of samples taken.



3. Reset the applet. Click "Sample 25" twice to choose 50 SRSs and display the confidence intervals based on those samples. How many captured the parameter μ? Keep clicking "Sample 25" and observe the value of "Percent hit." What do you notice?
4. Repeat Step 3 using a 90% confidence level.
5. Repeat Step 3 using an 80% confidence level.
6. Summarize what you have learned about the relationship between confidence level and "Percent hit" after taking many samples.

We will investigate the effect of changing the sample size later.

As the Activity confirms, *the confidence level is the overall capture rate if the method is used many times*. Figure 8.4 illustrates the behavior of the confidence interval $\bar{x} \pm 10$ for Mr. Schiel's mystery mean. Starting with the population, imagine taking many SRSs of 16 observations. The first sample has $\bar{x} = 240.80$, the second has $\bar{x} = 246.05$, the third has $\bar{x} = 248.85$, and so on. The sample mean varies from sample to sample, but when we use the formula $\bar{x} \pm 10$ to get an interval based on each sample, about 95% *of these intervals capture the unknown population mean* μ.

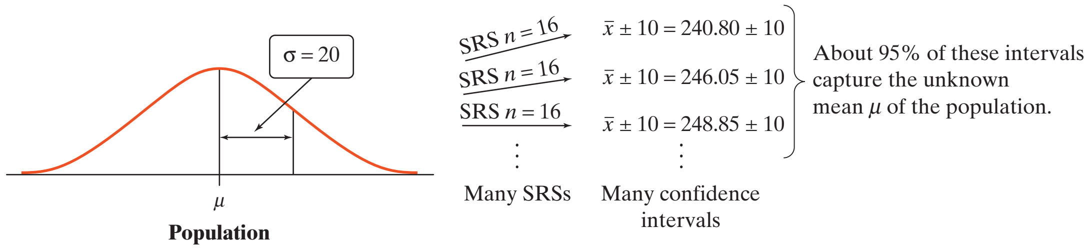


FIGURE 8.4 To say that $\bar{x} \pm 10$ is a 95% confidence interval for the population mean μ is to say that, in repeated samples, about 95% of these intervals capture μ.

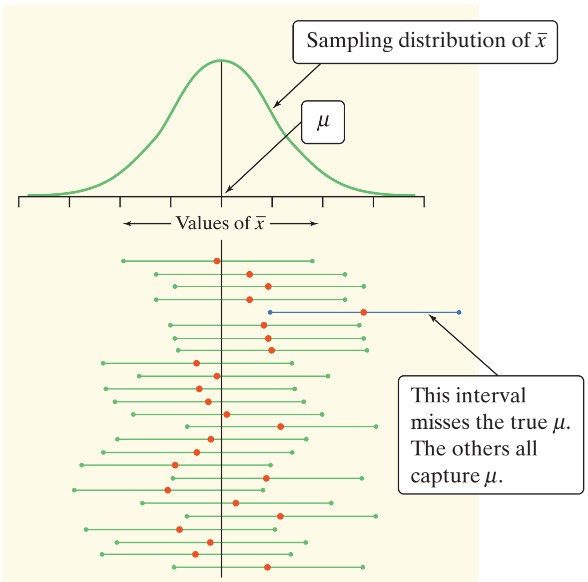


FIGURE 8.5 Twenty-five samples of the same size from the same population gave these 95% confidence intervals. In the long run, about 95% of samples give an interval that captures the population mean μ.

Figure 8.5 illustrates the idea of a confidence interval in a different form. It shows the result of drawing many SRSs from the same population and calculating a 95% confidence interval from each sample. The center of each interval is at $\bar{x}$ and therefore varies from sample to sample. The sampling distribution of $\bar{x}$ appears at the top of the figure to show the long-term pattern of this variation. The 95% confidence intervals from 25 SRSs appear below.

Here's what you should notice:

- The center $\bar{x}$ of each interval is marked by a dot.
- The distance from the dot to either endpoint of the interval is the margin of error.
- 24 of these 25 intervals (that's 96%) contain the true value of μ. If we took many samples, about 95% of the resulting confidence intervals would capture μ.

Figures 8.4 and 8.5 give us the insight we need to interpret a confidence level.

INTERPRETING CONFIDENCE LEVELS

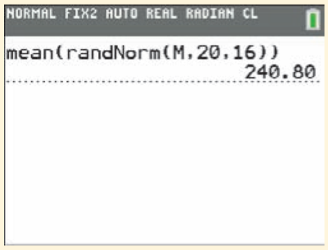
To say that we are 95% *confident* is shorthand for "If we take many samples of the same size from this population, about 95% of them will result in an interval that captures the actual parameter value."

The confidence level tells us how likely it is that the method we are using will produce an interval that captures the population parameter *if we use it many times*. However, in practice we tend to calculate only a single confidence interval for a given situation. *The confidence level does not tell us the chance that a particular confidence interval captures the population parameter.* Instead, the confidence interval gives us a set of plausible values for the parameter.

caution

EXAMPLE The Mystery Mean

Interpreting a confidence level



The confidence level in the mystery mean example—roughly 95%—tells us that if we take many SRSs of size 16 from Mr. Schiel's mystery population, the interval $\bar{x} \pm 10$ will capture the population mean μ for about 95% of those samples.

Be sure you understand the basis for our confidence. There are only two possibilities:

1. The interval from 230.80 to 250.80 contains the population mean μ.
2. The interval from 230.80 to 250.80 does *not* contain the population mean μ. Our SRS was one of the few samples for which $\bar{x}$ is not within 10 points of the true μ. Only about 5% of all samples result in a confidence interval that fails to capture μ.

We cannot know whether our sample is one of the 95% for which the interval $\bar{x} \pm 10$ catches μ or whether it is one of the unlucky 5%. The statement that we are "95% confident" that the unknown μ lies between 230.80 and 250.80 is shorthand for saying, "We got these numbers by a method that gives correct results 95% of the time."

THINK ABOUT IT

What's the probability that our 95% confidence interval captures the parameter? It's *not* 95%! Before we execute the command `mean(randNorm(M,20,16))`, we have a 95% chance of getting a sample mean that's within $2\sigma_{\bar{x}}$ of the mystery μ, which would lead to a confidence interval that captures μ. Once we have chosen a random sample, the sample mean $\bar{x}$ either is or isn't within $2\sigma_{\bar{x}}$ of μ. And the resulting confidence interval either does or doesn't contain μ. *After* we construct a confidence interval, the probability that it captures the population parameter is either *1 (it does)* or *0 (it doesn't)*.

We interpret confidence intervals and confidence levels in much the same way whether we are estimating a population mean, proportion, or some other parameter.

EXAMPLE Do You Use Twitter?

Interpreting a confidence interval and a confidence level

The Pew Internet and American Life Project asked a random sample of 2253 U.S. adults, "Do you ever . . . use Twitter or another service to share updates about yourself or to see updates about others?" Of the sample, 19% said "Yes." According to Pew, the resulting 95% confidence interval is (0.167, 0.213).[2]

PROBLEM: Interpret the confidence interval and the confidence level.

SOLUTION:

Confidence interval: We are 95% confident that the interval from 0.167 to 0.213 captures the true proportion p of all U.S. adults who use Twitter or another service for updates.

Confidence level: If many samples of 2253 U.S. adults were taken, the resulting confidence intervals would capture the true proportion of all U.S. adults who use Twitter or another service for updates for about 95% of those samples.

For Practice *Try Exercise* 15

AP® EXAM TIP On a given problem, you may be asked to interpret the confidence interval, the confidence level, or both. Be sure you understand the difference: the confidence interval gives a set of plausible values for the parameter and the confidence level describes the long-run capture rate of the method.

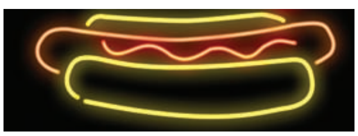
Confidence intervals are statements about *parameters*. In the previous example, it would be wrong to say, "We are 95% confident that the interval from 0.167 to 0.213 contains the *sample* proportion of U.S. adults who use Twitter." Why? Because we *know* that the *sample* proportion, $\hat{p} = 0.19$, is in the interval. Likewise, in the mystery mean example, it would be wrong to say that "95% of the values are between 230.80 and 250.80," whether we are referring to the sample or the population. All we can say is, "Based on Mr. Schiel's sample, we believe that the *population mean* is somewhere between 230.80 and 250.80."

caution

When interpreting a confidence interval, make it clear that you are describing a parameter and not a statistic. And be sure to include context.

CHECK YOUR UNDERSTANDING

How much does the fat content of Brand X hot dogs vary? To find out, researchers measured the fat content (in grams) of a random sample of 10 Brand X hot dogs. A 95% confidence interval for the population standard deviation σ is 2.84 to 7.55.



1. Interpret the confidence interval.
2. Interpret the confidence level.
3. True or false: The interval from 2.84 to 7.55 has a 95% chance of containing the actual population standard deviation σ. Justify your answer.

Constructing a Confidence Interval

Why settle for 95% confidence when estimating an unknown parameter? Do larger random samples yield "better" intervals? The *Confidence Intervals* applet might shed some light on these questions.

APPLET

ACTIVITY | The *Confidence Intervals* Applet

MATERIALS:
Computer with Internet connection and display capability

In this Activity, you will use the applet to explore the relationship between the confidence level, the sample size, and the confidence interval.

1. Go to www.whfreeman.com/tps5e and launch the *Confidence Intervals* applet. Use the default settings: confidence level 95% and sample size $n = 20$. Click "Sample 25."

2. Change the confidence level to 99%. What happens to the length of the confidence intervals?

3. Now change the confidence level to 90%. What happens to the length of the confidence intervals?

4. Finally, change the confidence level to 80%. What happens to the length of the confidence intervals?

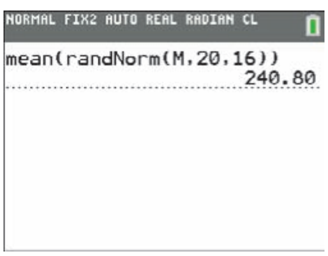
5. Summarize what you learned about the relationship between the confidence level and the length of the confidence intervals for a fixed sample size.

6. Reset the applet and change the confidence level to 95%. What happens to the length of the confidence intervals as you increase the sample size?

7. Does increasing the sample size increase the capture rate (percent hit)? Use the applet to investigate.

As the Activity illustrates, the price we pay for greater confidence is a wider interval. If we're satisfied with 80% confidence, then our interval of plausible values for the parameter will be much narrower than if we insist on 90%, 95%, or 99% confidence. But we'll also be much less confident in our estimate. Taking the idea of confidence to an extreme, what if we want to estimate with 100% confidence the proportion p of all U.S. adults who use Twitter? That's easy: we're 100% confident that the interval from 0 to 1 captures the true population proportion!

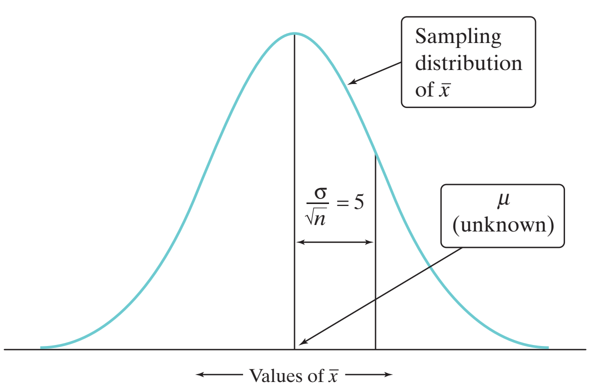
The activity also shows that we can get a more precise estimate of a parameter by increasing the sample size. Larger samples yield narrower confidence intervals. This result holds for any confidence level.



Let's look a bit more closely at the method we used earlier to calculate an approximate 95% confidence interval for Mr. Schiel's mystery mean. We started with

$$\text{point estimate} \pm \text{margin of error}$$

Our point estimate came from the sample statistic $\bar{x} = 240.80$. What about the margin of error? Because the population distribution is Normal, so is the sampling distribution of $\bar{x}$. About 95% of the values of $\bar{x}$ will lie within 2 standard deviations ($2\sigma_{\bar{x}}$) of the mystery mean μ. See the figure below. We could rewrite our interval as



$$240.80 \pm 2 \cdot 5 = \bar{x} \pm 2 \cdot \sigma_{\bar{x}}$$

This leads to the more general formula for a confidence interval:

$$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$

The **critical value** is a multiplier that makes the interval wide enough to have the stated capture rate. The critical value depends on both the confidence level C and the sampling distribution of the statistic.

CALCULATING A CONFIDENCE INTERVAL

The confidence interval for estimating a population parameter has the form

$$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$

where the statistic we use is the point estimator for the parameter.

The confidence interval for the mystery mean μ of Mr. Schiel's population illustrates several important properties that are shared by all confidence intervals in common use. The user chooses the confidence level, and the margin of error follows from this choice. We would like high confidence and also a small margin of error. High confidence says that our method almost always gives correct answers. A small margin of error says that we have pinned down the parameter quite precisely.

Our general formula for a confidence interval is

$$\text{Statistic} \pm \overbrace{(\text{critical value}) \cdot (\text{standard deviation of statistic})}^{\text{Margin of error}}$$

Recall that $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ and $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$. So as the sample size n increases, the standard deviation of the statistic decreases.

We can see that the margin of error depends on the critical value and the standard deviation of the statistic. The critical value is tied directly to the confidence level: greater confidence requires a larger critical value. The standard deviation of the statistic depends on the sample size n: larger samples give more precise estimates, which means less variability in the statistic.

So the margin of error gets smaller when:

- *The confidence level decreases*. There is a trade-off between the confidence level and the margin of error. To obtain a smaller margin of error from the same data, you must be willing to accept lower confidence. Earlier, we found that a 95% confidence interval for Mr. Schiel's mystery mean μ is 230.80 to 250.80. The 80% confidence interval for μ is 234.39 to 247.21. Figure 8.6 compares these two intervals.
- *The sample size n increases*. Increasing the sample size n reduces the margin of error for any fixed confidence level.

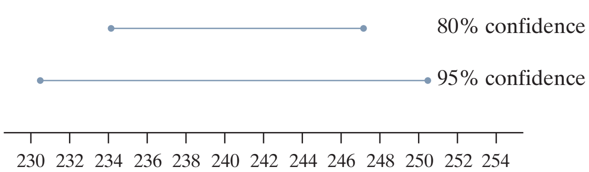


FIGURE 8.6 80% and 95% confidence intervals for Mr. Schiel's mystery mean. Higher confidence requires a longer interval.

Using Confidence Intervals Wisely

Our goal in this section has been to introduce you to the big ideas of confidence intervals without getting bogged down in details. You may have noticed that we only calculated intervals in a contrived setting: estimating an unknown population mean μ when we somehow knew the population standard deviation σ. In practice, when we don't know μ, we don't know σ either. We'll learn to construct confidence intervals for a population mean in this more realistic setting in Section 8.3. First, we will study confidence intervals for a population proportion p in Section 8.2. Although it is possible to estimate other parameters, confidence intervals for means and proportions are the most common tools in everyday use.

Here are two important cautions to keep in mind when constructing and interpreting confidence intervals.

- *Our method of calculation assumes that the data come from an SRS of size n from the population of interest.* Other types of random samples (stratified or cluster, say) might be preferable to an SRS in a given setting, but they require more complex calculations than the ones we'll use.
- *The margin of error in a confidence interval covers only chance variation due to random sampling or random assignment.* The margin of error is obtained from the sampling distribution. It indicates how close our estimate is likely to be to the unknown parameter if we repeat the random sampling or random assignment process many times. Practical difficulties, such as undercoverage and nonresponse in a sample survey, can lead to additional errors that may be larger than this chance variation. Remember this unpleasant fact when reading the results of an opinion poll or other sample survey. The way in which a survey or experiment is conducted influences the trustworthiness of its results in ways that are not included in the announced margin of error.

Section 8.1 Summary

- To estimate an unknown population parameter, start with a statistic that provides a reasonable guess. The chosen statistic is a **point estimator** for the parameter. The specific value of the point estimator that we use gives a **point estimate** for the parameter.
- A **C% confidence interval** uses sample data to estimate an unknown population parameter with an indication of how precise the estimate is and of how confident we are that the result is correct.
- A confidence interval gives an interval of plausible values for the parameter. The interval is computed from the data and has the form

$$\text{point estimate} \pm \text{margin of error}$$

When calculating a confidence interval, it is common to use the form

$$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$

- To interpret a C% confidence interval, say, "We are C% confident that the interval from ____ to ____ captures the [parameter in context]." Be sure that your interpretation describes a parameter and not a statistic.
- The **confidence level C** is the success rate of the method that produces the interval. If you use 95% confidence intervals often, in the long run about 95% of your intervals will contain the true parameter value. You don't know whether a 95% confidence interval calculated from a particular set of data actually captures the true parameter value.

- Other things being equal, the **margin of error** of a confidence interval gets smaller as
  - the confidence level C decreases.
  - the sample size n increases.
- Remember that the margin of error for a confidence interval includes only chance variation, not other sources of error like nonresponse and undercoverage.

Section 8.1 Exercises

In Exercises 1 to 4, determine the point estimator you would use and calculate the value of the point estimate.

1. **Got shoes?** How many pairs of shoes, on average, do female teens have? To find out, an AP® Statistics class conducted a survey. They selected an SRS of 20 female students from their school. Then they recorded the number of pairs of shoes that each student reported having. Here are the data:

pg 478

50	26	26	31	57	19	24	22	23	38
13	50	13	34	23	30	49	13	15	51

2. **Got shoes?** The class in Exercise 1 wants to estimate the variability in the number of pairs of shoes that female students have by estimating the population variance σ^2.

3. **Going to the prom** Tonya wants to estimate what proportion of the seniors in her school plan to attend the prom. She interviews an SRS of 50 of the 750 seniors in her school and finds that 36 plan to go to the prom.

pg 478

4. **Reporting cheating** What proportion of students are willing to report cheating by other students? A student project put this question to an SRS of 172 undergraduates at a large university: "You witness two students cheating on a quiz. Do you go to the professor?" Only 19 answered "Yes."[3]

5. **NAEP scores** Young people have a better chance of full-time employment and good wages if they are good with numbers. How strong are the quantitative skills of young Americans of working age? One source of data is the National Assessment of Educational Progress (NAEP) Young Adult Literacy Assessment Survey, which is based on a nationwide probability sample of households. The NAEP survey includes a short test of quantitative skills, covering mainly basic arithmetic and the ability to apply it to realistic problems. Scores on the test range from 0 to 500. For example, a person who scores 233 can add the amounts of two checks appearing on a bank deposit slip; someone scoring 325 can determine the price of a meal from a menu; a person scoring 375 can transform a price in cents per ounce into dollars per pound.[4]

   Suppose that you give the NAEP test to an SRS of 840 people from a large population in which the scores have mean 280 and standard deviation $\sigma = 60$. The mean $\bar{x}$ of the 840 scores will vary if you take repeated samples.

   (a) Describe the shape, center, and spread of the sampling distribution of $\bar{x}$.

   (b) Sketch the sampling distribution of $\bar{x}$. Mark its mean and the values 1, 2, and 3 standard deviations on either side of the mean.

   (c) According to the 68–95–99.7 rule, about 95% of all values of $\bar{x}$ lie within a distance m of the mean of the sampling distribution. What is m? Shade the region *on the axis* of your sketch that is within m of the mean.

   (d) Whenever $\bar{x}$ falls in the region you shaded, the population mean μ lies in the confidence interval $\bar{x} \pm m$. For what percent of all possible samples does the interval capture μ?

6. **Auto emissions** Oxides of nitrogen (called NOX for short) emitted by cars and trucks are important contributors to air pollution. The amount of NOX emitted by a particular model varies from vehicle to vehicle. For one light-truck model, NOX emissions vary with mean $\mu = 1.8$ grams per mile and standard deviation $\sigma = 0.4$ gram per mile. You test an SRS of 50 of these trucks. The sample mean NOX level $\bar{x}$ will vary if you take repeated samples.

(a) Describe the shape, center, and spread of the sampling distribution of $\bar{x}$.

(b) Sketch the sampling distribution of $\bar{x}$. Mark its mean and the values 1, 2, and 3 standard deviations on either side of the mean.

(c) According to the 68–95–99.7 rule, about 95% of all values of $\bar{x}$ lie within a distance m of the mean of the sampling distribution. What is m? Shade the region *on the axis* of your sketch that is within m of the mean.

(d) Whenever $\bar{x}$ falls in the region you shaded, the unknown population mean μ lies in the confidence interval $\bar{x} \pm m$. For what percent of all possible samples does the interval capture μ?

7. **NAEP scores** Refer to Exercise 5. Below your sketch, choose one value of $\bar{x}$ inside the shaded region and draw its corresponding confidence interval. Do the same for one value of $\bar{x}$ outside the shaded region. What is the most important difference between these intervals? (Use Figure 8.5, on page 483, as a model for your drawing.)

8. **Auto emissions** Refer to Exercise 6. Below your sketch, choose one value of $\bar{x}$ inside the shaded region and draw its corresponding confidence interval. Do the same for one value of $\bar{x}$ outside the shaded region. What is the most important difference between these intervals? (Use Figure 8.5, on page 483, as a model for your drawing.)

9. pg 481 **Prayer in school** A *New York Times*/CBS News Poll asked a random sample of U.S. adults the question, "Do you favor an amendment to the Constitution that would permit organized prayer in public schools?" Based on this poll, the 95% confidence interval for the population proportion who favor such an amendment is (0.63, 0.69).

(a) Interpret the confidence interval.

(b) What is the point estimate that was used to create the interval? What is the margin of error?

(c) Based on this poll, a reporter claims that more than two-thirds of U.S. adults favor such an amendment. Use the confidence interval to evaluate this claim.

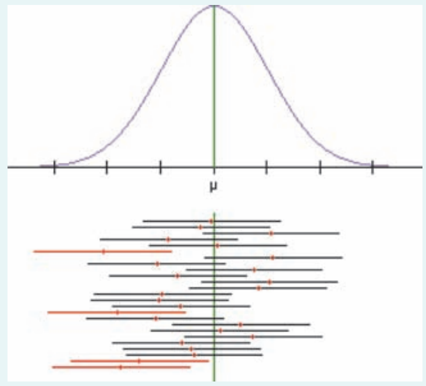
10. **Losing weight** A Gallup Poll asked a random sample of U.S. adults, "Would you like to lose weight?" Based on this poll, the 95% confidence interval for the population proportion who want to lose weight is (0.56, 0.62).[5]

(a) Interpret the confidence interval.

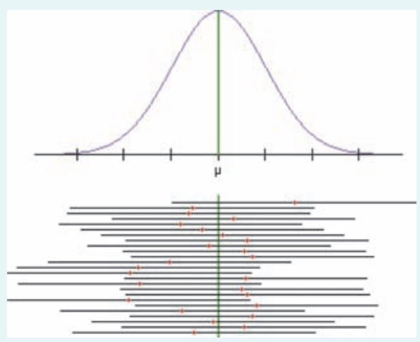
(b) What is the point estimate that was used to create the interval? What is the margin of error?

(c) Based on this poll, Gallup claims that more than half of U.S. adults want to lose weight. Use the confidence interval to evaluate this claim.

11. **How confident?** The figure below shows the result of taking 25 SRSs from a Normal population and constructing a confidence interval for each sample. Which confidence level—80%, 90%, 95%, or 99%—do you think was used? Explain.



12. **How confident?** The figure below shows the result of taking 25 SRSs from a Normal population and constructing a confidence interval for each sample. Which confidence level—80%, 90%, 95%, or 99%—do you think was used? Explain.



13. **Prayer in school** Refer to Exercise 9. The news article goes on to say: "The theoretical errors do not take into account · · · additional error resulting from the various practical difficulties in taking any survey of public opinion." List some of the "practical difficulties" that may cause errors which are not included in the ±3 percentage point margin of error.

14. **Losing weight** Refer to Exercise 10. As Gallup indicates, the 3 percentage point margin of error for this poll includes only sampling variability (what they call "sampling error"). What other potential sources of error (Gallup calls these "nonsampling errors") could affect the accuracy of the 95% confidence interval?

15. pg 484 **Shoes** The AP® Statistics class in Exercise 1 also asked an SRS of 20 boys at their school how many

pairs of shoes they have. A 95% confidence interval for the difference in the population means (girls – boys) is 10.9 to 26.5. Interpret the confidence interval and the confidence level.

16. **Lying online** Many teens have posted profiles on sites such as Facebook. A sample survey asked random samples of teens with online profiles if they included false information in their profiles. Of 170 younger teens (ages 12 to 14) polled, 117 said "Yes." Of 317 older teens (ages 15 to 17) polled, 152 said "Yes."[6] A 95% confidence interval for the difference in the population proportions (younger teens – older teens) is 0.120 to 0.297. Interpret the confidence interval and the confidence level.

17. **Shoes** Refer to Exercise 15. Does the confidence interval give convincing evidence of a difference in the population mean number of pairs of shoes for boys and girls at the school? Justify your answer.

18. **Lying online** Refer to Exercise 16. Does the confidence interval give convincing evidence of a difference in the population proportions of younger and older teens who include false information in their profiles? Justify your answer.

19. **Explaining confidence** A 95% confidence interval for the mean body mass index (BMI) of young American women is 26.8 ± 0.6. Discuss whether each of the following explanations is correct.

(a) We are confident that 95% of all young women have BMI between 26.2 and 27.4.

(b) We are 95% confident that future samples of young women will have mean BMI between 26.2 and 27.4.

(c) Any value from 26.2 to 27.4 is believable as the true mean BMI of young American women.

(d) If we take many samples, the population mean BMI will be between 26.2 and 27.4 in about 95% of those samples.

(e) The mean BMI of young American women cannot be 28.

20. **Explaining confidence** The admissions director from Big City University found that (107.8, 116.2) is a 95% confidence interval for the mean IQ score of all freshmen. Discuss whether each of the following explanations is correct.

(a) There is a 95% probability that the interval from 107.8 to 116.2 contains μ.

(b) There is a 95% chance that the interval (107.8, 116.2) contains $\bar{x}$.

(c) This interval was constructed using a method that produces intervals that capture the true mean in 95% of all possible samples.

(d) If we take many samples, about 95% of them will contain the interval (107.8, 116.2).

(e) The probability that the interval (107.8, 116.2) captures μ is either 0 or 1, but we don't know which.

Multiple choice: Select the best answer for Exercises 21 to 24.

Exercises 21 and 22 refer to the following setting. A researcher plans to use a random sample of families to estimate the mean monthly family income for a large population.

21. The researcher is deciding between a 95% confidence level and a 99% confidence level. Compared to a 95% confidence interval, a 99% confidence interval will be

(a) narrower and would involve a larger risk of being incorrect.

(b) wider and would involve a smaller risk of being incorrect.

(c) narrower and would involve a smaller risk of being incorrect.

(d) wider and would involve a larger risk of being incorrect.

(e) wider and would have the same risk of being incorrect.

22. The researcher is deciding between a sample of size $n = 500$ and a sample of size $n = 1000$. Compared to using a sample size of $n = 500$, a 95% confidence interval based on a sample size of $n = 1000$ will be

(a) narrower and would involve a larger risk of being incorrect.

(b) wider and would involve a smaller risk of being incorrect.

(c) narrower and would involve a smaller risk of being incorrect.

(d) wider and would involve a larger risk of being incorrect.

(e) narrower and would have the same risk of being incorrect.

23. In a poll,

I. Some people refused to answer questions.

II. People without telephones could not be in the sample.

III. Some people never answered the phone in several calls.

Which of these possible sources of bias is included in the $\pm 2\%$ margin of error announced for the poll?

(a) I only (c) III only (e) None of these

(b) II only (d) I, II, and III

24. You have measured the systolic blood pressure of an SRS of 25 company employees. A 95% confidence interval for the mean systolic blood pressure for the employees of this company is (122, 138). Which of the following statements is true?

(a) 95% of the sample of employees have a systolic blood pressure between 122 and 138.

(b) 95% of the population of employees have a systolic blood pressure between 122 and 138.

(c) If the procedure were repeated many times, 95% of the resulting confidence intervals would contain the population mean systolic blood pressure.

(d) If the procedure were repeated many times, 95% of the time the population mean systolic blood pressure would be between 122 and 138.

(e) If the procedure were repeated many times, 95% of the time the sample mean systolic blood pressure would be between 122 and 138.

25. **Power lines and cancer** (4.2, 4.3) Does living near power lines cause leukemia in children? The National Cancer Institute spent 5 years and $5 million gathering data on this question. The researchers compared 638 children who had leukemia with 620 who did not. They went into the homes and measured the magnetic fields in children's bedrooms, in other rooms, and at the front door. They recorded facts about power lines near the family home and also near the mother's residence when she was pregnant. Result: no connection between leukemia and exposure to magnetic fields of the kind produced by power lines was found.[7]

(a) Was this an observational study or an experiment? Justify your answer.

(b) Does this study show that living near power lines doesn't cause cancer? Explain.

26. **Sisters and brothers** (3.1, 3.2) How strongly do physical characteristics of sisters and brothers correlate? Here are data on the heights (in inches) of 11 adult pairs:[8]

Brother:	71	68	66	67	70	71	70	73	72	65	66
Sister:	69	64	65	63	65	62	65	64	66	59	62

(a) Construct a scatterplot using brother's height as the explanatory variable. Describe what you see.

(b) Use your calculator to compute the least-squares regression line for predicting sister's height from brother's height. Interpret the slope in context.

(c) Damien is 70 inches tall. Predict the height of his sister Tonya.

(d) Do you expect your prediction in (c) to be very accurate? Give appropriate evidence to support your answer.

8.2 Estimating a Population Proportion

WHAT YOU WILL LEARN **By the end of the section, you should be able to:**

- State and check the Random, 10%, and Large Counts conditions for constructing a confidence interval for a population proportion.
- Determine critical values for calculating a *C*% confidence interval for a population proportion using a table or technology.
- Construct and interpret a confidence interval for a population proportion.
- Determine the sample size required to obtain a *C*% confidence interval for a population proportion with a specified margin of error.

In Section 8.1, we saw that a confidence interval can be used to estimate an unknown population parameter. We are often interested in estimating the proportion p of some outcome in the population. Here are some examples:

- What proportion of U.S. adults are unemployed right now?
- What proportion of high school students have cheated on a test?

- What proportion of pine trees in a national park are infested with beetles?
- What proportion of college students pray daily?
- What proportion of a company's laptop batteries last as long as the company claims?

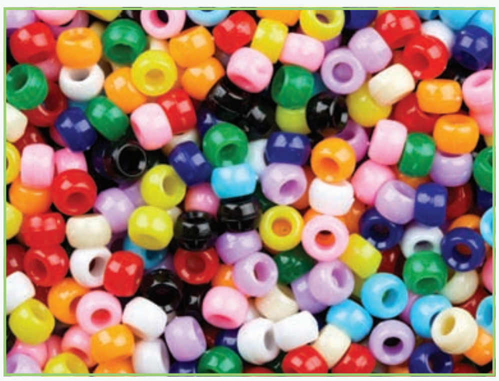
This section shows you how to construct and interpret a confidence interval for a population proportion. The following Activity gives you a taste of what lies ahead.

ACTIVITY | The Beads

MATERIALS:
Several thousand small plastic beads of at least two colors, container to hold all the beads, small cup for sampling, several small bowls

Before class, your teacher will prepare a large population of different-colored beads and put them into a container that you cannot see inside. Your goal is to estimate the actual proportion of beads in the population that have a particular color (say, red).

1. As a class, discuss how to use the cup provided to get a simple random sample of beads from the container. Think this through carefully, because you will get to take only one sample.



2. Have one student take an SRS of beads. Separate the beads into two groups: those that are red and those that aren't. Count the number of beads in each group.
3. Determine a point estimate for the unknown population proportion p of red beads in the container.
4. Now for the challenge: each team of three to four students will be given about 10 minutes to find a 95% confidence interval for the parameter p. Be sure to consider any conditions that are required for the methods you use.
5. Compare the results with those of the other teams in the class. Discuss any problems you encountered and how you dealt with them.

Conditions for Estimating p

Before constructing a confidence interval for p, you should check some important conditions:

- **Random:** The data should come from a well-designed random sample or randomized experiment. Otherwise, there's no scope for inference to a population (sampling) or inference about cause and effect (experiment). If we can't draw conclusions beyond the data at hand, there's not much point in constructing a confidence interval!

  Another important reason for random selection or random assignment is to introduce chance into the data-production process. We can model chance behavior with a probability distribution, like the sampling distributions of Chapter 7. The probability distribution helps us calculate a confidence interval.

  - **10%:** The procedure for calculating confidence intervals assumes that individual observations are independent. Well-conducted studies that use random sampling or random assignment can help ensure independent observations. However, our formula for the standard deviation of the sampling distribution of $\hat{p}$, $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$, acts as if we are sampling *with replacement* from a population. That's rarely the case. When sampling without replacement from a finite population, be sure to check that $n \leq \frac{1}{10}N$. Sampling more than 10% of the population would give a more precise estimate of the parameter p but would require us to use a different formula to calculate the standard deviation of the sampling distribution.

- **Large Counts:** The method that we use to construct a confidence interval for p depends on the fact that the sampling distribution of $\hat{p}$ is approximately Normal. From what we learned in Chapter 7, we can use the Normal approximation to the sampling distribution of $\hat{p}$ as long as $np \geq 10$ and $n(1-p) \geq 10$.

  In practice, of course, we don't know the value of p. If we did, we wouldn't need to construct a confidence interval for it! So we cannot check if np and $n(1-p)$ are at least 10. In large random samples, $\hat{p}$ will tend to be close to p. So we replace p by $\hat{p}$ in checking the Large Counts condition: $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$.

CONDITIONS FOR CONSTRUCTING A CONFIDENCE INTERVAL ABOUT A PROPORTION

- **Random:** The data come from a well-designed random sample or randomized experiment.
  - **10%:** When sampling without replacement, check that $n \leq \frac{1}{10}N$.
- **Large Counts:** Both $n\hat{p}$ and $n(1-\hat{p})$ are at least 10.

When Mr. Vignolini's class did the beads Activity, they got 107 red beads and 144 white beads. Their point estimate for the unknown proportion p of red beads in the population is

$$\hat{p} = \frac{107}{251} = 0.426$$

Let's see how the conditions play out for Mr. Vignolini's class.

EXAMPLE The Beads

Checking conditions

Mr. Vignolini's class wants to construct a confidence interval for the true proportion p of red beads in the container. Recall that the class's sample had 107 red beads and 144 white beads.

PROBLEM: **Check that the conditions for constructing a confidence interval for p are met.**

SOLUTION: **There are three conditions to check:**

- Random: The class took an SRS of 251 beads from the container.
  - 10%: Because the class sampled without replacement, they need to check that there are at least $10(251) = 2510$ beads in the population. Mr. Vignolini reveals that there are 3000 beads in the container.
- Large Counts: To use a Normal approximation for the sampling distribution of $\hat{p}$, we need both np and $n(1 - p)$ to be at least 10. Because we don't know p, we check

$$n\hat{p} = 251\left(\frac{107}{251}\right) = 107 \geq 10 \quad \text{and} \quad n(1 - \hat{p}) = 251\left(1 - \frac{107}{251}\right) = 251\left(\frac{144}{251}\right) = 144 \geq 10$$

That is, the counts of successes (red beads) and failures (non-red beads) are both at least 10.

All conditions are met, so it should be safe for the class to construct a confidence interval.

For Practice *Try Exercise* 27

Notice that $n\hat{p}$ and $n(1 - \hat{p})$ should be whole numbers. You don't really need to calculate these values since they are just the number of successes and failures in the sample. In the previous example, we could address the Large Counts condition simply by saying, "The numbers of successes (107) and failures (144) in the sample are both at least 10."

What happens if one of the conditions is violated? If the data come from a convenience sample or a poorly designed experiment, there's no point constructing a confidence interval for p. Violation of the Random condition severely limits our ability to make any inference beyond the data at hand.

The figure below shows a screen shot from the *Confidence Intervals for Proportions* applet at the book's Web site, www.whfreeman.com/tps5e. We set $n = 20$ and $p = 0.25$. The Large Counts condition is not met because $np = 20(0.25) = 5$. We used the applet to generate 1000 95% confidence intervals for p. Only 902 of those 1000 intervals contained $p = 0.25$, a capture rate of 90.2%. When the Large Counts condition is violated, the capture rate will be lower than the one advertised by the confidence level if the method is used many times.

That leaves just the 10% condition when sampling without replacement from a finite population. Large random samples give more precise estimates than small random samples. So randomly selecting more than 10% of a population should be a good thing! Unfortunately, the formula for the standard deviation of $\hat{p}$ that we developed in Chapter 7, $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$, is not correct when the 10% condition is violated. The formula gives a value that is too large. Confidence intervals based on this formula are longer than they need to be. If many 95% confidence intervals for a population proportion are constructed in this way, more than 95% of them will capture p. The actual capture rate is *greater* than the reported confidence level when the 10% condition is violated.

CHECK YOUR UNDERSTANDING

In each of the following settings, check whether the conditions for calculating a confidence interval for the population proportion p are met.

1. An AP® Statistics class at a large high school conducts a survey. They ask the first 100 students to arrive at school one morning whether or not they slept at least 8 hours the night before. Only 17 students say "Yes."
2. A quality control inspector takes a random sample of 25 bags of potato chips from the thousands of bags filled in an hour. Of the bags selected, 3 had too much salt.

Constructing a Confidence Interval for *p*

When the conditions are met, the sampling distribution of $\hat{p}$ will be approximately Normal with mean $\mu_{\hat{p}} = p$ and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$. Figure 8.7 displays this distribution. Inference about a population proportion p is based on the sampling distribution of $\hat{p}$.

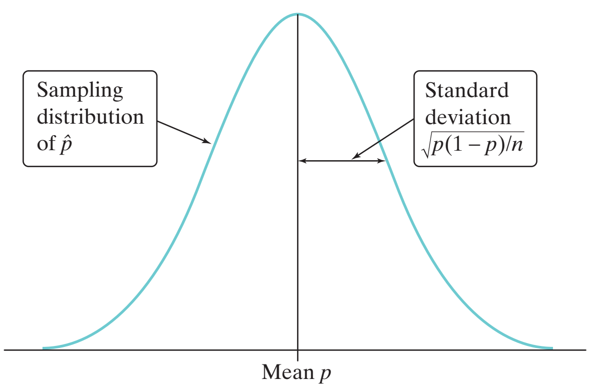


FIGURE 8.7 Suppose that a population has proportion p of successes. When the conditions for inference are met, the sampling distribution of the proportion of successes $\hat{p}$ in a sample is approximately Normal with mean p and standard deviation $\sqrt{\frac{p(1-p)}{n}}$.

We can use the general formula from Section 8.1 to construct a confidence interval for an unknown population proportion p:

statistic ± (critical value) · (standard deviation of statistic)

The sample proportion $\hat{p}$ is the statistic we use to estimate p. Doing so makes sense if the data came from a well-designed random sample or randomized experiment (the Random condition).

The standard deviation of the sampling distribution of $\hat{p}$ is

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

if the 10% condition is met. Because we don't know the value of p, we replace it with the sample proportion $\hat{p}$. The resulting quantity is called the **standard error** (**SE**) of the sample proportion $\hat{p}$.

Some books refer to $\sigma_{\hat{p}}$ as the "standard error" of $\hat{p}$ and to what we call the standard error as the "estimated standard error."

$$\text{SE}_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

It describes how close the sample proportion $\hat{p}$ will typically be to the population proportion p in repeated SRSs of size n.

DEFINITION: Standard error

When the standard deviation of a statistic is estimated from data, the result is called the **standard error** of the statistic.

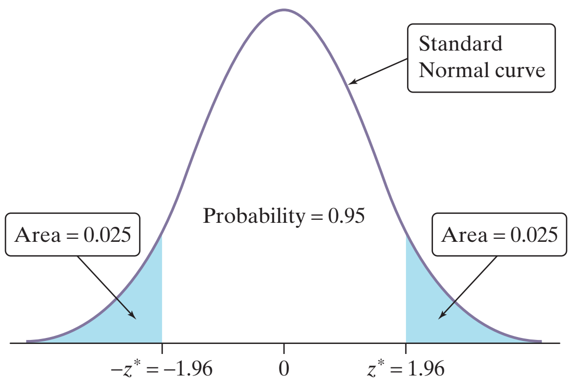


FIGURE 8.8 Finding the critical value for a 95% confidence interval. The correct value is $z^* = 1.96$, which is more precise than the value of 2 we had been using for 95% confidence.

How do we get the critical value for our confidence interval? If the Large Counts condition is met, we can use a Normal curve. For the approximate 95% confidence intervals of Section 8.1, we used a critical value of 2 based on the 68–95–99.7 rule for Normal distributions. We can get a more precise critical value from Table A or a calculator. As Figure 8.8 shows, the central 95% of the standard Normal distribution is marked off by 2 points, $z^* = 1.96$ and $-z^* = -1.96$. We use the * to remind you that this is a critical value, not a standardized score that has been calculated from data.

To find a level C confidence interval, we need to catch the central $C\%$ under the standard Normal curve. Here's an example that shows how to get the **critical value** z^* for a different confidence level.

EXAMPLE 80% Confidence

Finding a critical value

PROBLEM: Use Table A or technology to find the critical value z^* for an 80% confidence interval. Assume that the Large Counts condition is met.

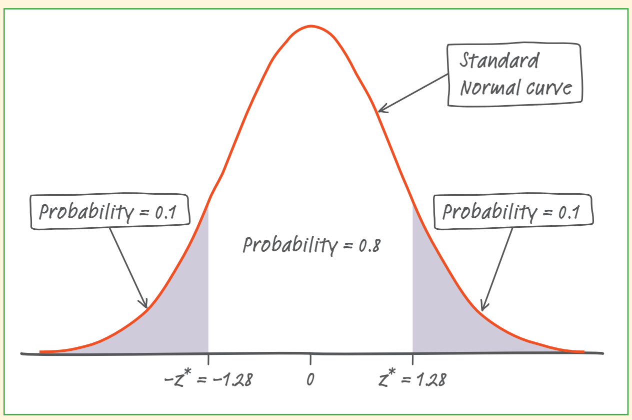


FIGURE 8.9 Finding the critical value for an 80% confidence interval.

SOLUTION: For an 80% confidence level, we need to capture the central 80% of the standard Normal distribution. In capturing the central 80%, we leave out 20%, or 10% in each tail. So the desired critical value z^* is the point with area 0.1 to its right under the standard Normal curve. Figure 8.9 shows the details in picture form.

Search the body of Table A to find the point $-z^*$ with area 0.1 to its left. The closest entry is $z = -1.28$. (See the excerpt from Table A below.) So the critical value we want is $z^* = 1.28$.

z	.07	.08	.09
−1.3	.0853	.0838	.0823
−1.2	.1020	.1003	.0985
−1.1	.1210	.1190	.1170

Using technology: The command `invNorm(area:0.1,μ:0, σ:1)` gives $z = -1.28$. The critical value is $z^* = 1.28$, which matches what we got from Table A.

For Practice *Try Exercise* **31**

Technically, the correct formula for a confidence interval is statistic ± (critical value) · (standard *error* of statistic). We are following the convention used on the AP® Statistics exam formula sheet.

Once we find the critical value z^*, our confidence interval for the population proportion p is

$$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$
$$= \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Notice that we replaced the standard deviation of $\hat{p}$ with the formula for its standard error. The resulting interval is sometimes called a **one-sample z interval for a population proportion.**

ONE-SAMPLE z INTERVAL FOR A POPULATION PROPORTION

When the conditions are met, a C% confidence interval for the unknown proportion p is

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where z^* is the critical value for the standard Normal curve with C% of its area between $-z^*$ and z^*.

Now we can get the desired confidence interval for Mr. Vignolini's class.

EXAMPLE The Beads

Calculating a confidence interval for p

PROBLEM: Mr. Vignolini's class took an SRS of beads from the container and got 107 red beads and 144 white beads.

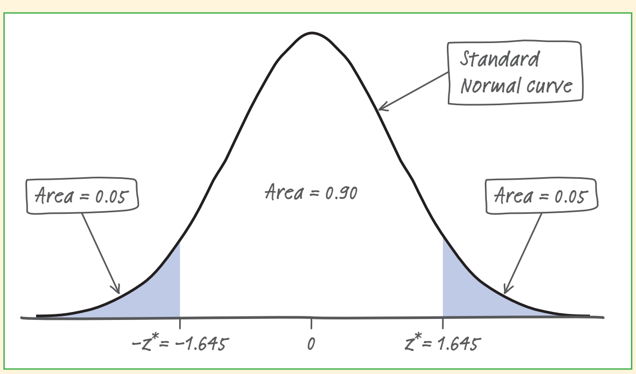
(a) Calculate and interpret a 90% confidence interval for p.

(b) Mr. Vignolini claims that exactly half of the beads in the container are red. Use your result from part (a) to comment on this claim.

SOLUTION: We checked conditions for calculating the interval earlier.

(a) Our confidence interval has the form

$$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$
$$= \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



The sample statistic is $\hat{p} = 107/251 = 0.426$. Now let's find the critical value. From Table A, we look for the point with area 0.05 to its left. As the excerpt from Table A shows, this point is between $z = -1.64$ and $z = -1.65$. The calculator's `invNorm(area:0.05,μ:0,σ:1)` gives $z = -1.645$. So we use $z^* = 1.645$ as our critical value.

z	.03	.04	.05
−1.7	.0418	.0409	.0401
−1.6	.0516	.0505	.0495
−1.5	.0630	.0618	.0606

The resulting 90% confidence interval is

$$= 0.426 \pm 1.645\sqrt{\frac{(0.426)(1-0.426)}{251}}$$
$$= 0.426 \pm 0.051$$
$$= (0.375, 0.477)$$

We are 90% confident that the interval from 0.375 to 0.477 captures the true proportion of red beads in Mr. Vignolini's container.

(b) The confidence interval in part (a) gives a set of plausible values for the population proportion of red beads. Because 0.5 is not contained in the interval, it is not a plausible value for *p*. We have good reason to doubt Mr. Vignolini's claim.

For Practice *Try Exercise* 35

CHECK YOUR UNDERSTANDING

Alcohol abuse has been described by college presidents as the number one problem on campus, and it is an important cause of death in young adults. How common is it? A survey of 10,904 randomly selected U.S. college students collected information on drinking behavior and alcohol-related problems.[9] The researchers defined "frequent binge drinking" as having five or more drinks in a row three or more times in the past two weeks. According to this definition, 2486 students were classified as frequent binge drinkers.

1. Identify the parameter of interest.
2. Check conditions for constructing a confidence interval for the parameter.
3. Find the critical value for a 99% confidence interval. Show your method. Then calculate the interval.
4. Interpret the interval in context.

Putting It All Together: The Four-Step Process

Taken together, the examples about Mr. Vignolini's class and the beads Activity show you how to get a confidence interval for an unknown population proportion *p*. We can use the familiar four-step process whenever a problem asks us to construct and interpret a confidence interval.

STEP 4

CONFIDENCE INTERVALS: A FOUR-STEP PROCESS

State: What *parameter* do you want to estimate, and at what *confidence level*?

Plan: Identify the appropriate inference *method*. Check *conditions*.

Do: If the conditions are met, perform *calculations*.

Conclude: *Interpret* your interval in the context of the problem.

The next example illustrates the four-step process in action.

EXAMPLE Teens Say Sex Can Wait

STEP 4

Confidence interval for *p*

The Gallup Youth Survey asked a random sample of 439 U.S. teens aged 13 to 17 whether they thought young people should wait to have sex until marriage.[10] Of the sample, 246 said "Yes." Construct and interpret a 95% confidence interval for the proportion of all teens who would say "Yes" if asked this question.

STATE: We want to estimate the true proportion p of all 13- to 17-year-olds in the United States who would say that young people should wait to have sex until they get married with 95% confidence.

PLAN: We should use a one-sample z interval for p if the conditions are met.

- Random: Gallup surveyed a random sample of U.S. teens.
  - 10%: Because Gallup is sampling without replacement, we need to check the 10% condition: there are at least 10(439) = 4390 U.S. teens aged 13 to 17.
- Large Counts: We check the counts of "successes" and "failures":

$$n\hat{p} = 246 \geq 10 \quad \text{and} \quad n(1-\hat{p}) = 193 \geq 10$$

DO: The sample statistic is $\hat{p} = 246/439 = 0.56$. A 95% confidence interval for p is given by

$$\begin{aligned}\hat{p} \pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0.56 \pm 1.96\sqrt{\frac{(0.56)(0.44)}{439}} \\ &= 0.56 \pm 0.046 \\ &= (0.514, 0.606)\end{aligned}$$

CONCLUDE: We are 95% confident that the interval from 0.514 to 0.606 captures the true proportion of 13- to 17-year-olds in the United States who would say that teens should wait until marriage to have sex.

AP® EXAM TIP If a free-response question asks you to construct and interpret a confidence interval, you are expected to do the entire four-step process. That includes clearly defining the parameter, identifying the procedure, and checking conditions.

For Practice *Try Exercise* 39

Simulation studies have shown that a variation of our method for calculating a 95% confidence interval for p can result in closer to a 95% capture rate in the long run, especially for small sample sizes. This simple adjustment, first suggested by Edwin Bidwell Wilson in 1927, is sometimes called the "plus four" estimate. Just pretend we have four additional observations, two of which are successes and two of which are failures. Then calculate the "plus four interval" using the plus four estimate in place of $\hat{p}$ in our usual formula.

caution

Remember that the margin of error in a confidence interval includes only sampling variability! There are other sources of error that are not taken into account. As is the case with many surveys, we are forced to assume that the teens answered truthfully. If they didn't, then our estimate may be biased. Other problems like nonresponse and question wording can also affect the results of this or any other poll. *Lesson:* Sampling beads is much easier than sampling people!

Your calculator will handle the "Do" part of the four-step process, as the following Technology Corner illustrates.

AP® EXAM TIP You may use your calculator to compute a confidence interval on the AP® exam. But there's a risk involved. If you just give the calculator answer with no work, you'll get either full credit for the "Do" step (if the interval is correct) or no credit (if it's wrong). If you opt for the calculator-only method, be sure to name the procedure (e.g., one-proportion z interval) and to give the interval (e.g., 0.514 to 0.607).

15. TECHNOLOGY CORNER — CONFIDENCE INTERVAL FOR A POPULATION PROPORTION

TI-Nspire instructions in Appendix B; HP Prime instructions on the book's Web site.

The TI-83/84 and TI-89 can be used to construct a confidence interval for an unknown population proportion. We'll demonstrate using the previous example. Of $n = 439$ teens surveyed, $X = 246$ said they thought that young people should wait to have sex until after marriage. To construct a confidence interval:

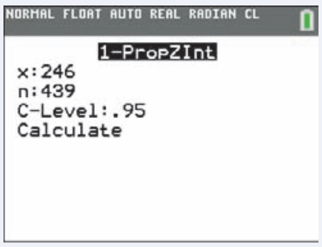
TI-83/84

- Press `STAT`, then choose TESTS and `1-PropZInt`.

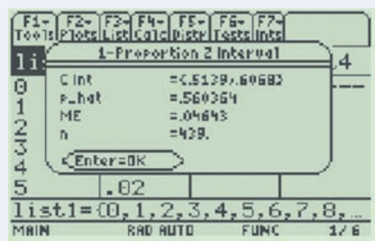
TI-89

- In the Statistics/List Editor, press `2nd` `F2` ([F7]) and choose `1-PropZInt`.

- When the 1-PropZInt screen appears, enter $x = 246$, $n = 439$, and confidence level 0.95.



- Highlight "Calculate" and press `ENTER`. The 95% confidence interval for p is reported, along with the sample proportion $\hat{p}$ and the sample size, as shown here.



Choosing the Sample Size

In planning a study, we may want to choose a sample size that allows us to estimate a population proportion within a given margin of error. National survey organizations like the Gallup Poll typically sample between 1000 and 1500 American adults, who are interviewed by telephone. Why do they choose such sample sizes?

The margin of error (ME) in the confidence interval for p is

$$ME = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Here, z^* is the standard Normal critical value for the confidence level we want. Because the margin of error involves the sample proportion of successes $\hat{p}$, we have to guess the value of $\hat{p}$ when choosing n. Here are two ways to do this:

1. Use a guess for $\hat{p}$ based on a pilot study or on past experience with similar studies. You should do several calculations that cover the range of $\hat{p}$-values you might get.
2. Use $\hat{p} = 0.5$ as the guess. The margin of error ME is largest when $\hat{p} = 0.5$, so this guess is conservative in the sense that if we get any other $\hat{p}$ when we do our study, we will get a margin of error smaller than planned.

Once you have a guess for $\hat{p}$, the formula for the margin of error can be solved to give the sample size n needed.

SAMPLE SIZE FOR DESIRED MARGIN OF ERROR

To determine the sample size n that will yield a $C\%$ confidence interval for a population proportion p with a maximum margin of error ME, solve the following inequality for n:

$$z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq ME$$

where $\hat{p}$ is a guessed value for the sample proportion. The margin of error will always be less than or equal to ME if you use $\hat{p} = 0.5$.

Here's an example that shows you how to determine the sample size.

EXAMPLE Customer Satisfaction

Determining sample size

A company has received complaints about its customer service. The managers intend to hire a consultant to carry out a survey of customers. Before contacting the consultant, the company president wants some idea of the sample size that she will be required to pay for. One critical question is the degree of satisfaction with the company's customer service, measured on a 5-point scale. The president wants to estimate the proportion p of customers who are satisfied (that is, who choose either "somewhat satisfied" or "very satisfied," the two highest levels on the 5-point scale). She decides that she wants the estimate to be within 3% (0.03) at a 95% confidence level. How large a sample is needed?

PROBLEM: Determine the sample size needed to estimate p within 0.03 with 95% confidence.

SOLUTION: The critical value for 95% confidence is $z^* = 1.96$. We have no idea about the true proportion p of satisfied customers, so we decide to use $\hat{p} = 0.5$ as our guess. Because the company president wants a margin of error of no more than 0.03, we need to solve the inequality

$$1.96\sqrt{\frac{0.5(1-0.5)}{n}} \leq 0.03$$

for n. Multiplying both sides by $\sqrt{n}$ and then dividing both sides by 0.03 yields

$$\frac{1.96}{0.03}\sqrt{0.5(1-0.5)} \leq \sqrt{n}$$

Squaring both sides gives

$$\left(\frac{1.96}{0.03}\right)^2 (0.5)(1-0.5) \leq n$$

$$1067.111 \leq n$$

We round up to 1068 respondents to ensure that the margin of error is no more than 3%.

For Practice *Try Exercise* 43

Why not round to the nearest whole number—in this case, 1067? Because a smaller sample size will result in a larger margin of error, possibly more than the desired 3% for the poll.

If you want a 2.5% margin of error rather than 3%, then

$$n \geq \left(\frac{1.96}{0.025}\right)^2 (0.5)(1-0.5) = 1536.64 \Rightarrow n = 1537$$

For a 2% margin of error, the sample size you need is

$$n \geq \left(\frac{1.96}{0.02}\right)^2 (0.5)(1-0.5) = 2401$$

As usual, smaller margins of error call for larger samples.

News reports frequently describe the results of surveys with sample sizes between 1000 and 1500 and a margin of error of about 3%. These surveys generally use sampling procedures more complicated than a simple random sample, so the calculation of confidence intervals is more involved than what we have studied in this section. The calculations of the previous example still give you a rough idea of how such surveys are planned.

CHECK YOUR UNDERSTANDING

Refer to the previous example about the company's customer satisfaction survey.

1. In the company's prior-year survey, 80% of customers surveyed said they were "somewhat satisfied" or "very satisfied." Using this value as a guess for $\hat{p}$, find the sample size needed for a margin of error of 3% at a 95% confidence level.

2. What if the company president demands 99% confidence instead? Determine how this would affect your answer to Question 1.

Section 8.2 Summary

- The conditions for constructing a confidence interval about a population proportion are
  - **Random:** The data were produced by a well-designed random sample or randomized experiment.
    - **10%:** When sampling without replacement, we check that the population is at least 10 times as large as the sample.
  - **Large Counts:** The sample is large enough that $n\hat{p}$ and $n(1-\hat{p})$, the counts of successes and failures in the sample, are both at least 10.
- Confidence intervals for a population proportion p are based on the sampling distribution of the sample proportion $\hat{p}$. When the conditions for inference are met, the sampling distribution of $\hat{p}$ is approximately Normal with mean p and standard deviation $\sqrt{p(1-p)/n}$.

- In practice, we use the sample proportion $\hat{p}$ to estimate the unknown parameter p. We therefore replace the standard deviation of $\hat{p}$ with its **standard error** when constructing a confidence interval. The $C\%$ **confidence interval for p** is

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where z^* is the standard Normal **critical value** with $C\%$ of its area between $-z^*$ and z^*.

- When constructing a confidence interval, follow the four-step process:

STEP 4

STATE: What *parameter* do you want to estimate, and at what *confidence level*?

PLAN: Identify the appropriate inference *method*. Check *conditions*.

DO: If the conditions are met, perform *calculations*.

CONCLUDE: *Interpret* your interval in the context of the problem.

- The sample size needed to obtain a confidence interval with approximate margin of error ME for a population proportion involves solving

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le ME$$

for n, where $\hat{p}$ is a guessed value for the sample proportion, and z^* is the critical value for the confidence level you want. Use $\hat{p} = 0.5$ if you don't have a good idea about the value of $\hat{p}$.

8.2 TECHNOLOGY CORNER

TI-Nspire Instructions in Appendix B; HP Prime instructions on the book's Web site.

15. Confidence interval for a population proportion page 501

Section 8.2 Exercises

For Exercises 27 to 30, check whether each of the conditions is met for calculating a confidence interval for the population proportion p.

27. **Rating school food** Latoya wants to estimate what proportion of the seniors at her boarding high school like the cafeteria food. She interviews an SRS of 50 of the 175 seniors living in the dormitory. She finds that 14 think the cafeteria food is good.

pg 494

28. **High tuition costs** Glenn wonders what proportion of the students at his school believe that tuition is too high. He interviews an SRS of 50 of the 2400 students at his college. Thirty-eight of those interviewed think tuition is too high.

29. **AIDS and risk factors** In the National AIDS Behavioral Surveys sample of 2673 adult heterosexuals, 0.2% had both received a blood transfusion and had a sexual partner from a group at high risk of AIDS. We want to estimate the proportion p in the population who share these two risk factors.

30. **Whelks and mussels** The small round holes you often see in sea shells were drilled by other sea creatures, who ate the former dwellers of the shells.

Whelks often drill into mussels, but this behavior appears to be more or less common in different locations. Researchers collected whelk eggs from the coast of Oregon, raised the whelks in the laboratory, then put each whelk in a container with some delicious mussels. Only 9 of 98 whelks drilled into a mussel.[11] The researchers want to estimate the proportion p of Oregon whelks that will spontaneously drill into mussels.

31. **98% confidence** Find z^* for a 98% confidence interval using Table A or your calculator. Show your method.

pg 497

32. **93% confidence** Find z^* for a 93% confidence interval using Table A or your calculator. Show your method.

33. **Going to the prom** Tonya wants to estimate what proportion of her school's seniors plan to attend the prom. She interviews an SRS of 50 of the 750 seniors in her school and finds that 36 plan to go to the prom.

(a) Identify the population and parameter of interest.

(b) Check conditions for constructing a confidence interval for the parameter.

(c) Construct a 90% confidence interval for p. Show your method.

(d) Interpret the interval in context.

34. **Reporting cheating** What proportion of students are willing to report cheating by other students? A student project put this question to an SRS of 172 undergraduates at a large university: "You witness two students cheating on a quiz. Do you go to the professor?" Only 19 answered "Yes."[12]

(a) Identify the population and parameter of interest.

(b) Check conditions for constructing a confidence interval for the parameter.

(c) Construct a 99% confidence interval for p. Show your method.

(d) Interpret the interval in context.

35. **Binge drinking** In a recent National Survey of Drug Use and Health, 2312 of 5914 randomly selected full-time U.S. college students were classified as binge drinkers.[13]

pg 498

STEP 4

(a) Calculate and interpret a 99% confidence interval for the population proportion p that are binge drinkers.

(b) A newspaper article claims that 45% of full-time U.S. college students are binge drinkers. Use your result from part (a) to comment on this claim.

36. **Teens' texting** A Pew Internet and American Life Project survey found that 392 of 799 randomly selected teens reported texting with their friends every day.

STEP 4

(a) Calculate and interpret a 95% confidence interval for the population proportion p that would report texting with their friends every day.

(b) Is it plausible that the true proportion of American teens who text with their friends every day is 0.45? Use your result from part (a) to support your answer.

37. **Binge drinking** Describe a possible source of error that is not included in the margin of error for the 99% confidence interval in Exercise 35.

38. **Teens' texting** Describe a possible source of error that is not included in the margin of error for the 95% confidence interval in Exercise 36.

39. **How common is SAT coaching?** A random sample of students who took the SAT college entrance examination twice found that 427 of the respondents had paid for coaching courses and that the remaining 2733 had not.[14] Construct and interpret a 99% confidence interval for the proportion of coaching among students who retake the SAT.

pg 500

STEP 4

40. **2010 begins** In January 2010 a Gallup Poll asked a random sample of adults, "In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?" In all, 256 said that they were satisfied and the remaining 769 said they were not. Construct and interpret a 90% confidence interval for the proportion of adults who are satisfied with how things are going.

STEP 4

41. **Equality for women?** Have efforts to promote equality for women gone far enough in the United States? A poll on this issue by the cable network MSNBC contacted 1019 adults. A newspaper article about the poll said, "Results have a margin of sampling error of plus or minus 3 percentage points."[15]

(a) The news article said that 65% of men, but only 43% of women, think that efforts to promote equality have gone far enough. Explain why we do not have enough information to give confidence intervals for men and women separately.

(b) Would a 95% confidence interval for women alone have a margin of error less than 0.03, about equal to 0.03, or greater than 0.03? Why? (You see that the news article's statement about the margin of error for poll results is a bit misleading.)

42. **A TV poll** A television news program conducts a call-in poll about a proposed city ban on handgun ownership. Of the 2372 calls, 1921 oppose the ban.

The station, following recommended practice, makes a confidence statement: "81% of the Channel 13 Pulse Poll sample opposed the ban. We can be 95% confident that the true proportion of citizens opposing a handgun ban is within 1.6% of the sample result." Is the station's conclusion justified? Explain.

43. **Can you taste PTC?** PTC is a substance that has a strong bitter taste for some people and is tasteless for others. The ability to taste PTC is inherited. About 75% of Italians can taste PTC, for example. You want to estimate the proportion of Americans who have at least one Italian grandparent and who can taste PTC.

pg 502

(a) How large a sample must you test to estimate the proportion of PTC tasters within 0.04 with 90% confidence? Answer this question using the 75% estimate as the guessed value for $\hat{p}$.

(b) Answer the question in part (a) again, but this time use the conservative guess $\hat{p} = 0.5$. By how much do the two sample sizes differ?

44. **School vouchers** A national opinion poll found that 44% of all American adults agree that parents should be given vouchers that are good for education at any public or private school of their choice. The result was based on a small sample.

(a) How large an SRS is required to obtain a margin of error of 0.03 (that is, ±3%) in a 99% confidence interval? Answer this question using the previous poll's result as the guessed value for $\hat{p}$.

(b) Answer the question in part (a) again, but this time use the conservative guess $\hat{p} = 0.5$. By how much do the two sample sizes differ?

45. **Election polling** Gloria Chavez and Ronald Flynn are the candidates for mayor in a large city. We want to estimate the proportion p of all registered voters in the city who plan to vote for Chavez with 95% confidence and a margin of error no greater than 0.03. How large a random sample do we need? Show your work.

46. **Starting a nightclub** A college student organization wants to start a nightclub for students under the age of 21. To assess support for this proposal, they will select an SRS of students and ask each respondent if he or she would patronize this type of establishment. What sample size is required to obtain a 90% confidence interval with an approximate margin of error of 0.04? Show your work.

47. **Teens and their TV sets** According to a Gallup Poll report, 64% of teens aged 13 to 17 have TVs in their rooms. Here is part of the footnote to this report:

These results are based on telephone interviews with a randomly selected national sample of 1028 teenagers in the Gallup Poll Panel of households, aged 13 to 17. For results based on this sample, one can say . . . that the maximum error attributable to sampling and other random effects is ±3 percentage points. 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.[16]

(a) We omitted the confidence level from the footnote. Use what you have learned to determine the confidence level, assuming that Gallup took an SRS.

(b) Give an example of a "practical difficulty" that could lead to biased results for this survey.

48. **Gambling and the NCAA** Gambling is an issue of great concern to those involved in college athletics. Because of this concern, the National Collegiate Athletic Association (NCAA) surveyed randomly selected student athletes concerning their gambling-related behaviors.[17] Of the 5594 Division I male athletes in the survey, 3547 reported participation in some gambling behavior. This includes playing cards, betting on games of skill, buying lottery tickets, betting on sports, and similar activities. A report of this study cited a 1% margin of error.

(a) The confidence level was not stated in the report. Use what you have learned to find the confidence level, assuming that the NCAA took an SRS.

(b) The study was designed to protect the anonymity of the student athletes who responded. As a result, it was not possible to calculate the number of students who were asked to respond but did not. How does this fact affect the way that you interpret the results?

Multiple choice: Select the best answer for Exercises 49 to 52.

49. A Gallup Poll found that only 28% of American adults expect to inherit money or valuable possessions from a relative. The poll's margin of error was ±3 percentage points at a 95% confidence level. This means that

(a) the poll used a method that gets an answer within 3% of the truth about the population 95% of the time.

(b) the percent of all adults who expect an inheritance is between 25% and 31%.

(c) if Gallup takes another poll on this issue, the results of the second poll will lie between 25% and 31%.

(d) there's a 95% chance that the percent of all adults who expect an inheritance is between 25% and 31%.

(e) Gallup can be 95% confident that between 25% and 31% of the sample expect an inheritance.

50. Most people can roll their tongues, but many can't. The ability to roll the tongue is genetically determined. Suppose we are interested in determining what proportion of students can roll their tongues. We test a simple random sample of 400 students and find that 317 can roll their tongues. The margin of error for a 95% confidence interval for the true proportion of tongue rollers among students is closest to

(a) 0.0008. (c) 0.03. (e) 0.05.

(b) 0.02. (d) 0.04.

51. You want to design a study to estimate the proportion of students at your school who agree with the statement, "The student government is an effective organization for expressing the needs of students to the administration." You will use a 95% confidence interval, and you would like the margin of error to be 0.05 or less. The minimum sample size required is

(a) 22. (b) 271. (c) 385. (d) 769. (e) 1795.

52. A newspaper reporter asked an SRS of 100 residents in a large city for their opinion about the mayor's job performance. Using the results from the sample, the C% confidence interval for the proportion of all residents in the city who approve of the mayor's job performance is 0.565 to 0.695. What is the value of C?

(a) 82 (b) 86 (c) 90 (d) 95 (e) 99

Exercises 53 and 54 refer to the following setting. The following table displays the number of accidents at a factory during each hour of a 24-hour shift (1 = 1:00 A.M.).

Hour	Number of accidents	Hour	Number of accidents
1	5	13	21
2	8	14	12
3	17	15	10
4	31	16	1
5	24	17	0
6	18	18	1
7	12	19	3
8	7	20	21
9	1	21	23
10	0	22	18
11	2	23	11
12	14	24	2

53. Accidents happen (1.2, 3.1)

(a) Construct a plot that displays the distribution of the number of accidents effectively.

(b) Construct a plot that shows the relationship between the number of accidents and the time when they occurred.

(c) Describe something that the plot in part (a) tells you about the data that the plot in part (b) does not.

(d) Describe something that the plot in part (b) tells you about the data that the plot in part (a) does not.

54. Accidents happen (1.3) Plant managers are concerned that the number of accidents may be significantly higher during the midnight to 8:00 A.M. shift than during the 4:00 P.M. to midnight shift. What would you tell them? Give appropriate statistical evidence to support your conclusion.

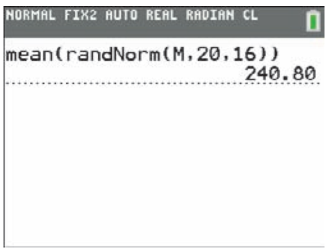
8.3 Estimating a Population Mean

WHAT YOU WILL LEARN By the end of the section, you should be able to:

- State and check the Random, 10%, and Normal/Large Sample conditions for constructing a confidence interval for a population mean.
- Explain how the t distributions are different from the standard Normal distribution and why it is necessary to use a t distribution when calculating a confidence interval for a population mean.
- Determine critical values for calculating a C% confidence interval for a population mean using a table or technology.
- Construct and interpret a confidence interval for a population mean.
- Determine the sample size required to obtain a C% confidence interval for a population mean with a specified margin of error.

Inference about a population proportion usually arises when we study *categorical* variables. We learned how to construct and interpret confidence intervals for a population proportion p in Section 8.2. To estimate a population mean, we have to record values of a *quantitative* variable for a sample of individuals. It makes sense to try to estimate the mean amount of sleep that students at a large high school got last night but not their mean eye color! In this section, we'll examine confidence intervals for a population mean μ.

The Problem of Unknown σ



Mr. Schiel's class did the mystery mean Activity (page 476) and got a value of $\bar{x} = 240.80$ from an SRS of size 16, as shown.

Their task was to estimate the unknown population mean μ. They knew that the population distribution was Normal and that its standard deviation was $\sigma = 20$. Their estimate was based on the sampling distribution of $\bar{x}$. Figure 8.10 shows this Normal sampling distribution once again.

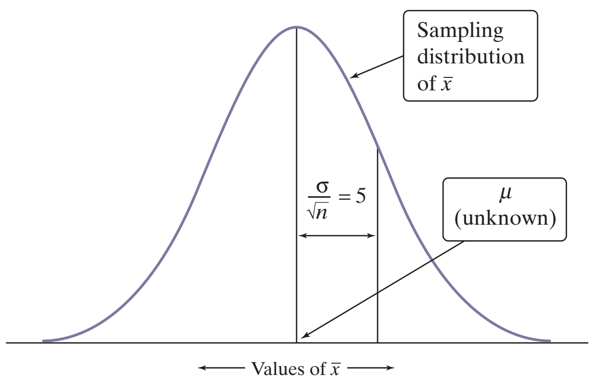


FIGURE 8.10 The Normal sampling distribution of $\bar{x}$ for the mystery mean Activity.

To calculate a 95% confidence interval for μ, we use our familiar formula:

$$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$

The critical value, $z^* = 1.96$, tells us how many standardized units we need to go out to catch the middle 95% of the sampling distribution. Our interval is

$$\bar{x} \pm z^* \cdot \frac{\sigma}{\sqrt{n}} = 240.80 \pm 1.96 \cdot \frac{20}{\sqrt{16}} = 240.80 \pm 9.80 = (231.00, 250.60)$$

We call such an interval a *one-sample z interval for a population mean.*

This method isn't very useful in practice, however. In most real-world settings, if we don't know the population mean μ, then we don't know the population standard deviation σ either.

How do we estimate μ when the population standard deviation σ is unknown? Our best guess for the value of σ is the sample standard deviation s_x. Maybe we could use the one-sample z interval for a population mean with s_x in place of σ:

$$\bar{x} \pm z^* \frac{s_x}{\sqrt{n}}$$

Let's try it.

ACTIVITY | Calculator BINGO!

MATERIALS:
TI-83/84 for each student

A farmer wants to estimate the mean weight (in grams) of all tomatoes grown on his farm. To do so, he will select a random sample of 4 tomatoes, calculate the mean weight (in grams), and use the sample mean $\bar{x}$ to create a 99% confidence interval for the population mean μ. Suppose that the weights of tomatoes on his farm are approximately Normally distributed with a mean of 100 grams and a standard deviation of 40 grams.

1. Use your calculator to simulate taking an SRS of size 4 from this population and creating a one-sample z interval for μ: $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = \bar{x} \pm 2.576 \frac{40}{\sqrt{4}}$. Enter the command shown below and press ENTER.

```
randNorm(100,40,4)→L1:ZInterval 40,mean(L1),4,99
```

Check to see whether the resulting interval captures $\mu = 100$. If it does not, shout "BINGO!"

To get the `randNorm` command, press MATH and arrow to PRB. The `ZInterval` command is in the Catalog. To get the `mean` and `stdDev` commands, press 2nd STAT (LIST) and arrow to the MATH menu.

Keep pressing ENTER to generate more 99% confidence intervals. Check each interval to see whether it captures $\mu = 100$. If it does not, shout "BINGO!" If this method of constructing confidence intervals is working properly, about what percent of the time should you get a BINGO? Does the method seem to be working?

The method in Step 1 works well if we know the population standard deviation σ. That's rarely the case in real life. What happens if we use the sample standard deviation s_x in place of σ when calculating a confidence interval for the population mean?

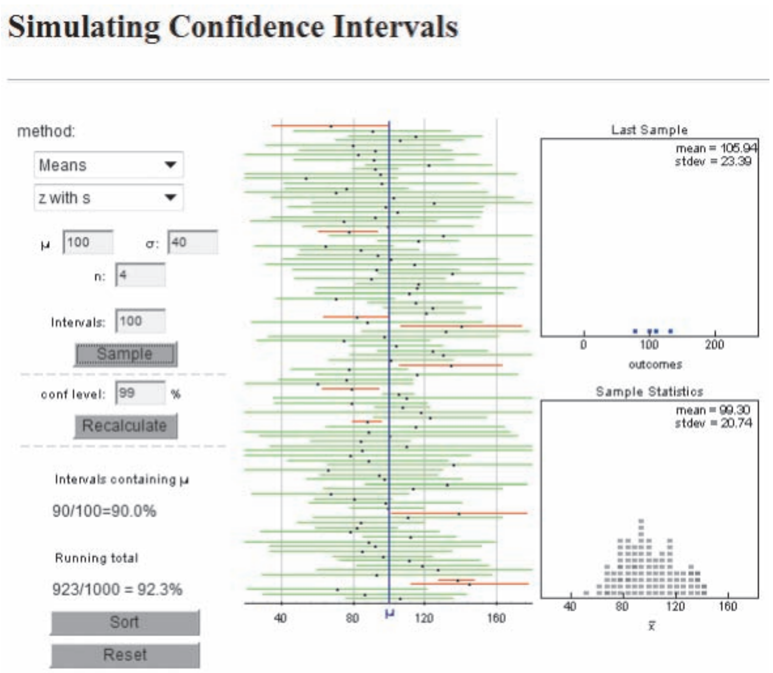
2. Use your calculator to simulate taking an SRS of size 4 from this population and creating a "modified" one-sample z interval for μ: $\bar{x} \pm z^* \frac{s_x}{\sqrt{n}} = \bar{x} \pm 2.576 \frac{s_x}{\sqrt{4}}$. Enter the command shown below and press ENTER.

```
randNorm(100,40,4)→L1:ZInterval stdDev(L1),mean(L1),4,99
```

Check to see whether the resulting interval captures $\mu = 100$. If it does not, shout "BINGO!"

Keep pressing ENTER to generate more 99% confidence intervals. Check each interval to see whether it captures $\mu = 100$. If it does not, shout "BINGO!" If this method of constructing confidence intervals is working properly, about what percent of the time should you get a BINGO? Does the method seem to be working?

The figure on the next page shows the results of using an applet from www.rossmanchance.com to repeatedly construct confidence intervals as described in Step 2 of the Activity. Of the 1000 intervals constructed, only 923 (that's 92.3%) captured the population mean. That's far below our desired 99% confidence level. What went wrong? The intervals that missed (those in red) came from samples with small standard deviations s_x and from samples in which $\bar{x}$ was far from the population mean μ. In those cases, multiplying $s_x/\sqrt{4}$ by $z^* = 2.576$ didn't produce long enough intervals to reach $\mu = 100$. We need to multiply by a larger critical value to achieve a 99% capture rate. But what critical value should we use?



When σ Is Unknown: The *t* Distributions

When the sampling distribution of $\bar{x}$ is close to Normal, we can find probabilities involving $\bar{x}$ by standardizing:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Recall that a statistic is a number computed from sample data. We know that the sample mean $\bar{x}$ is a statistic. So is the standardized value $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$. The sampling distribution of z shows the values this statistic takes in all possible SRSs of size n from the population.

Recall that the sampling distribution of $\bar{x}$ has mean μ and standard deviation $\sigma/\sqrt{n}$, as shown in Figure 8.11(a). What are the shape, center, and spread of the sampling distribution of the new statistic z?

From what we learned in Chapter 6, subtracting the constant μ from the values of the random variable $\bar{x}$ shifts the distribution left by μ units, making the mean 0. This transformation doesn't affect the shape or spread of the distribution. Dividing

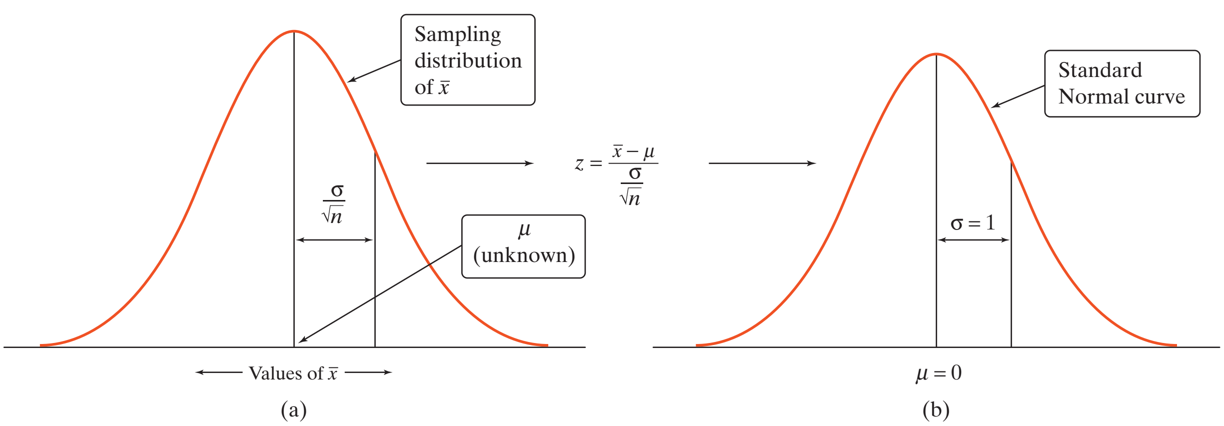


FIGURE 8.11 (a) Sampling distribution of $\bar{x}$ when the Normal/Large Sample condition is met. (b) Standardized values of $\bar{x}$ lead to the statistic z, which follows the standard Normal distribution.

by the constant $\sigma/\sqrt{n}$ keeps the mean at 0, makes the standard deviation 1, and leaves the shape unchanged. As shown in Figure 8.11(b), z has the standard Normal distribution $N(0, 1)$. Therefore, we can use Table A or a calculator to find the related probability involving z. That's how we have gotten the critical values for our confidence intervals so far.

When we don't know σ, we estimate it using the sample standard deviation s_x. What happens now when we standardize?

$$?? = \frac{\bar{x} - \mu}{s_x/\sqrt{n}}$$

To find out, let's start with a Normal population having mean $\mu = 100$ and standard deviation $\sigma = 5$. We'll simulate choosing an SRS of size $n = 4$ and calculating the sample mean $\bar{x}$. Then we will standardize the result in two ways:

$$z = \frac{\bar{x} - 100}{5/\sqrt{4}} \qquad \text{and} \qquad ?? = \frac{\bar{x} - 100}{s_x/\sqrt{4}}$$

Figure 8.12 shows the results of taking 500 SRSs of size $n = 4$ and standardizing the value of the sample mean $\bar{x}$ in both ways. The values of z follow a standard Normal distribution, as expected. The standardized values we get, using the sample standard deviation s_x in place of the population standard deviation σ, show much greater spread. In fact, in a few samples, the statistic

$$?? = \frac{\bar{x} - \mu}{s_x/\sqrt{n}}$$

took values below -6 or above 6.

This statistic has a distribution that is new to us, called a ***t* distribution.** It has a *different shape* than the standard Normal curve: still symmetric with a single peak at 0, but with much more area in the tails.

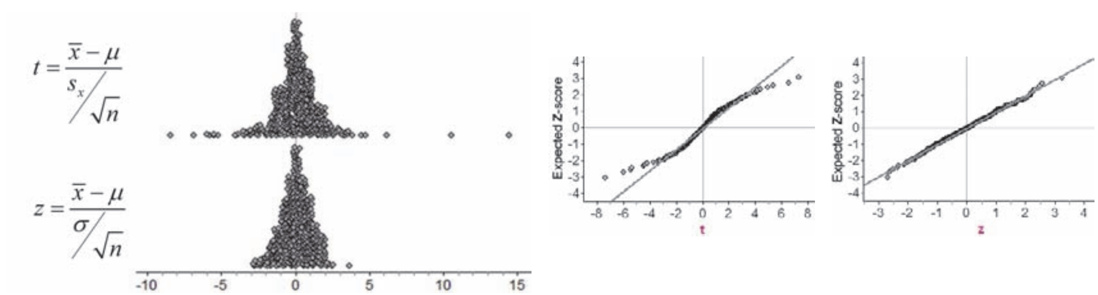


FIGURE 8.12 Fathom simulation showing standardized values of the sample mean $\bar{x}$ in 500 SRSs. The statistic z follows a standard Normal distribution. Replacing σ with s_x yields a statistic with much greater variability that doesn't follow the standard Normal curve. The Normal probability plot for the t statistic shows the departure from Normality in the tails of the t distribution.

The statistic t has the same interpretation as any standardized statistic: it says how far $\bar{x}$ is from its mean μ in standard deviation units. There is a different t distribution for each sample size. We specify a particular t distribution by giving its **degrees of freedom (df).** When we perform inference about a population mean μ using a t distribution, the appropriate degrees of freedom are found by subtracting 1 from the sample size n, making df $= n - 1$. We will write the t distribution with $n - 1$ degrees of freedom as t_{n-1} for short.

The *t* distribution and the *t* inference procedures were invented by William S. Gosset (1876–1937). Gosset worked for the Guinness brewery, and his goal in life was to make better beer. He used his new *t* procedures to find the best varieties of barley and hops. Gosset's statistical work helped him become head brewer. Because Gosset published under the pen name "Student," you will often see the *t* distribution called "Student's *t*" in his honor.

THE *t* DISTRIBUTIONS; DEGREES OF FREEDOM

Draw an SRS of size n from a large population that has a Normal distribution with mean μ and standard deviation σ. The statistic

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

has the **t distribution** with **degrees of freedom** df $= n - 1$. When the population distribution isn't Normal, this statistic will have approximately a t_{n-1} distribution if the sample size is large enough.

Figure 8.13 compares the density curves of the standard Normal distribution and the *t* distributions with 2 and 9 degrees of freedom. The figure illustrates these facts about the *t* distributions:

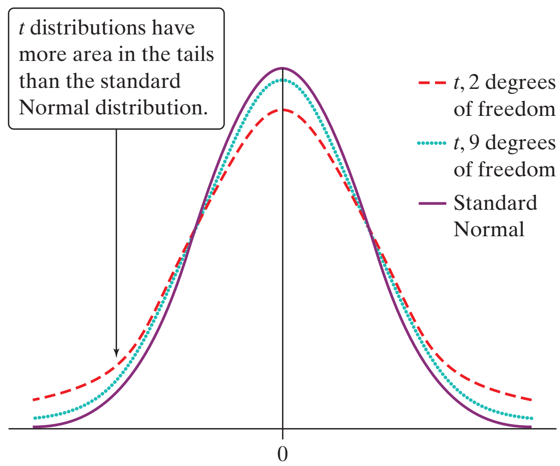


FIGURE 8.13 Density curves for the *t* distributions with 2 and 9 degrees of freedom and the standard Normal distribution. All are symmetric with center 0. The *t* distributions are somewhat more spread out.

- The density curves of the *t* distributions are similar in shape to the standard Normal curve. They are symmetric about 0, single-peaked, and bell-shaped.
- The spread of the *t* distributions is a bit greater than that of the standard Normal distribution. The *t* distributions in Figure 8.13 have more probability in the tails and less in the center than does the standard Normal. This is true because substituting the estimate s_x for the fixed parameter σ introduces more variation into the statistic.
- As the degrees of freedom increase, the *t* density curve approaches the standard Normal density curve ever more closely. This happens because s_x estimates σ more accurately as the sample size increases. So using s_x in place of σ causes little extra variation when the sample is large.

Table B in the back of the book gives critical values t^* for the *t* distributions. Each row in the table contains critical values for the *t* distribution whose degrees of freedom appear at the left of the row. For convenience, several of the more common confidence levels C are given at the bottom of the table. By looking down any column, you can check that the *t* critical values approach the Normal critical values z^* as the degrees of freedom increase.

When you use Table B to determine the correct value of t^* for a given confidence interval, all you need to know are the confidence level C and the degrees of freedom (df). Unfortunately, Table B does not include every possible sample size. *When the actual df does not appear in the table, use the greatest df available that is less than your desired df.* This guarantees a wider confidence interval than you need to justify a given confidence level. Better yet, use technology to find an accurate value of t^* for any df.

EXAMPLE Finding t^* Critical Values

Using Table B

PROBLEM: What critical value t^* from Table B should be used in constructing a confidence interval for the population mean in each of the following settings?

(a) A 95% confidence interval based on an SRS of size $n = 12$.

(b) A 90% confidence interval from a random sample of 48 observations.

SOLUTION:

(a) In Table B, we consult the row corresponding to df $= 12 - 1 = 11$. We move across that row to the entry that is directly above 95% confidence level on the bottom of the chart. The desired critical value is $t^* = 2.201$.

(b) With 48 observations, we want to find the t^* critical value for df $= 48 - 1 = 47$ and 90% confidence. There is no df $= 47$ row in Table B, so we use the more conservative df $= 40$. The corresponding critical value is $t^* = 1.684$.

The bottom row of Table B gives z^* critical values and is labeled as df $= \infty$. That's because the t distributions approach the standard Normal distribution as the degrees of freedom approach infinity.

	Upper-tail probability *p*			
df	.05	.025	.02	.01
10	1.812	2.228	2.359	2.764
11	1.796	2.201	2.328	2.718
12	1.782	2.179	2.303	2.681
∞	1.645	1.960	2.054	2.326
	90%	95%	96%	98%
	Confidence level *C*			

	Upper-tail probability *p*			
df	.10	.05	.025	.02
30	1.310	1.697	2.042	2.147
40	1.303	1.684	2.021	2.123
50	1.299	1.676	2.009	2.109
∞	1.282	1.645	1.960	2.054
	80%	90%	95%	96%
	Confidence level *C*			

For Practice *Try Exercise* **55**

For part (a) of the example, the corresponding standard Normal critical value for 95% confidence is $z^* = 1.96$. We have to go out farther than 1.96 standard deviations to capture the central 95% of the t distribution with 11 degrees of freedom.

Technology will quickly produce t^* critical values for any sample size.

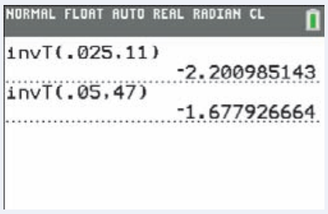
16. TECHNOLOGY CORNER INVERSE t ON THE CALCULATOR

TI-Nspire instructions in Appendix B; HP Prime instructions on the book's Web site.

Most newer TI-84 and TI-89 calculators allow you to find critical values t^* using the inverse t command. As with the calculator's inverse Normal command, you have to enter the *area to the left* of the desired critical value. Let's use the inverse t command to find the critical values in parts (a) and (b) of the example.

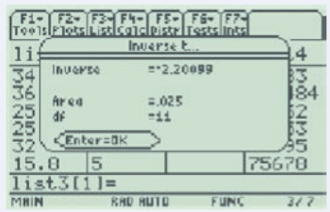
TI-83/84

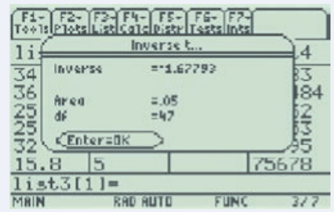
- Press [2nd] [VARS] (DISTR) and choose `invT(`.
- For part (a), **OS 2.55 or later:** In the dialog box, enter these values: `area:.025, df:11`, choose `Paste`, and then press [ENTER]. **Older OS:** Complete the command `invT(.025,11)` and press [ENTER].
- For part (b), use the command `invT(.05,47)`.



TI-89

- In the Statistics/List Editor, press [F5], choose `Inverse` and `Inverse t...`.
- For part (a), enter Area: .025 and Deg of Freedom, df: 11, and then press [ENTER].
- For part (b), use Area:.05 and df: 47.





Note that the t critical values are $t^* = 2.201$ and $t^* = 1.678$, respectively.

CHECK YOUR UNDERSTANDING

Use Table B to find the critical value t^* that you would use for a confidence interval for a population mean μ in each of the following settings. If possible, check your answer with technology.

1. A 96% confidence interval based on a random sample of 22 observations.
2. A 99% confidence interval from an SRS of 71 observations.

Conditions for Estimating μ

As with proportions, you should check some important conditions before constructing a confidence interval for a population mean. Two of the conditions should be familiar by now. The *Random condition* is crucial for doing inference. If the data don't come from a well-designed random sample or randomized experiment, you can't draw conclusions about a larger population or about cause and effect. When sampling without replacement, the *10% condition* ensures that our formula for the standard deviation of the statistic $\bar{x}$

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

is approximately correct.

The method we use to construct a confidence interval for μ depends on the fact that the sampling distribution of $\bar{x}$ is approximately Normal. From Chapter 7, we know that the sampling distribution of $\bar{x}$ is Normal if the population distribution is Normal. When the population distribution is not Normal, the central limit theorem tells us that the sampling distribution of $\bar{x}$ will be approximately Normal if the sample size is large enough ($n \geq 30$). Be sure to check this *Normal/Large Sample condition* before calculating a confidence interval.

CONDITIONS FOR CONSTRUCTING A CONFIDENCE INTERVAL ABOUT A MEAN

- **Random:** The data come from a well-designed random sample or randomized experiment.
  - **10%:** When sampling without replacement, check that $n \leq \frac{1}{10}N$.
- **Normal/Large Sample:** The population has a Normal distribution or the sample size is large ($n \geq 30$). If the population distribution has unknown shape and $n < 30$, use a graph of the sample data to assess the Normality of the population. Do not use t procedures if the graph shows strong skewness or outliers.

Larger samples improve the accuracy of critical values from the t distributions when the population is not Normal. This is true for two reasons:

1. The sampling distribution of $\bar{x}$ for large sample sizes is close to Normal.
2. As the sample size n grows, the sample standard deviation s_x will give a more accurate estimate of σ. This is important because we use $\frac{s_x}{\sqrt{n}}$ in place of $\frac{\sigma}{\sqrt{n}}$ when doing calculations.

The Normal/Large Sample condition is obviously met if we know that the population distribution is Normal or that the sample size is at least 30. What if we don't know the shape of the population distribution and $n < 30$? In that case, we have to graph the sample data. Our goal is to answer the question, "Is it reasonable to believe that these data came from a Normal population?"

How should graphs of data from small samples look if the population has a Normal distribution? The following Activity sheds some light on this question.

ACTIVITY | Sampling from a Normal Population

MATERIALS:
TI-83/84 or TI-89 for each student

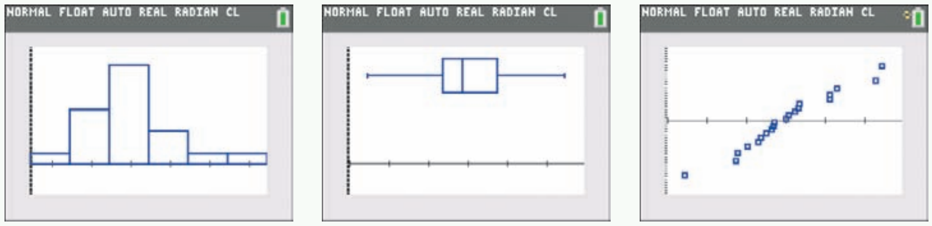
Let's use the calculator to simulate choosing random samples of size $n = 20$ from a Normal distribution with $\mu = 100$ and $\sigma = 15$ and then to plot the data.

1. Choose an SRS of 20 observations from this Normal population.

TI-83/84: Press `MATH`, arrow to `PRB` and choose `randNorm(`. Complete the command `randNorm(100,15,20)→L1` and press `ENTER`.

TI-89: Press `CATALOG` `F3` (Flash Apps), press `alpha` `2` (R) to jump to the r's, and choose `randNorm(`... Complete the command `tistat.randNorm(100,15,20)→list1` and press `ENTER`.

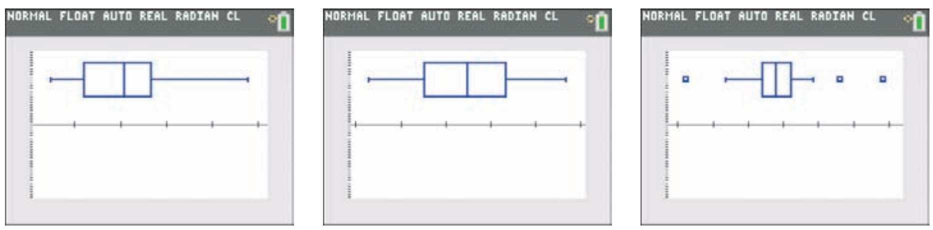
2. Make a histogram, a boxplot, and a Normal probability plot of the data in L1/list1. Do you see any obvious departures from Normality in the graphs of the sample data?



3. Repeat Steps 1 and 2 several times. Do the graphs of the sample data always look approximately Normal when the population distribution is Normal?

4. Compare the results with those of your classmates. How easy do you think it will be to use a graph of sample data to determine whether or not a population has a Normal distribution?

Did you expect that a random sample from a Normal population would yield a graph that looked Normal? Unfortunately, that's usually not the case. The figure below shows boxplots from three different SRSs of size 20 chosen in Step 3 of the Activity. The left-hand graph is skewed to the right. The right-hand graph shows three outliers in the sample. Only the middle graph looks symmetric and has no outliers.



As the Activity shows, it is very difficult to use a graph of sample data to assess the Normality of a population distribution. If the graph has a skewed shape or if there are outliers present, it could be because the population distribution isn't Normal. Skewness or outliers could also occur naturally in a random sample from a Normal population. To be safe, you should *only use a t distribution for small samples with no outliers or strong skewness*.

Caution

What constitutes strong skewness in a distribution? The following example gives you some idea.

EXAMPLE GPAs, Wood, and SATs

Can we use t?

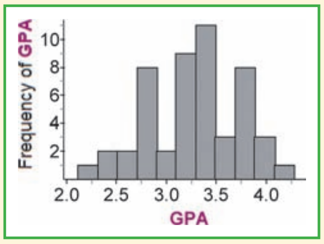
PROBLEM: **Determine if we can safely use a t^* critical value to calculate a confidence interval for the population mean in each of the following settings.**

(a) To estimate the average GPA of students at your school, you randomly select 50 students from classes you take. Figure 8.14(a) is a histogram of their GPAs.

(b) How much force does it take to pull wood apart? Figure 8.14(b) shows a stemplot of the force (in pounds) required to pull apart a random sample of 20 pieces of Douglas fir.

(c) Suppose you want to estimate the mean SAT Math score at a large high school. Figure 8.14(c) is a boxplot of the SAT Math scores for a random sample of 20 students at the school.

(a)



(b)

```
23 | 0
24 | 0
25 |
26 | 5
27 |          Key: 31/3 = 313 pounds
28 | 7
29 |
30 | 259
31 | 399
32 | 033677
33 | 0236
```

(c)

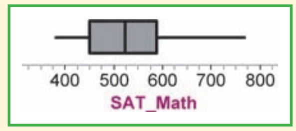


FIGURE 8.14 Can we use a *t* distribution for these data? (a) GPAs of 50 randomly selected students in your classes at school. (b) Force required to pull apart a random sample of 20 pieces of Douglas fir. (c) SAT Math scores for a random sample of 20 students at a large high school.

SOLUTION:

(a) No. Although the histogram is roughly symmetric with no outliers, the random sample of 50 students was only from *your* classes and not from all students at your school. So we should not use these data to calculate a confidence interval for the mean GPA of all students at the school.

(b) No. The graph is strongly skewed to the left with possible low outliers. We cannot trust a critical value from a *t* distribution with df = 19 in this case.

(c) Yes. The distribution is only moderately skewed to the right and there are no outliers present.

AP® EXAM TIP If a question on the AP® exam asks you to calculate a confidence interval, all the conditions should be met. However, you are still required to state the conditions and show evidence that they are met.

For Practice *Try Exercise* 59

THINK ABOUT IT

What's the difference between "strongly skewed" and "moderately skewed"? Look at the stemplot in Figure 8.14(b) and the boxplot in Figure 8.14(c). Compare the distance from the maximum to the median and from the median to the minimum in both graphs. In Figure 8.14(b), maximum − median = 336 − 319.5 = 16.5 and median − minimum = 319.5 − 230 = 89.5. The half of the stemplot with smaller values is more than five times as long as the half of the stemplot with larger values. In Figure 8.14(c), maximum − median ≈ 775 − 525 = 250 and median − minimum ≈ 525 − 375 = 150. The right half of the boxplot is not quite twice as long as the left half.

There is no accepted rule of thumb for identifying strong skewness. For that reason, we have chosen the data sets in examples and exercises to avoid borderline cases. You should be able to tell easily if strong skewness is present in a graph of data from a small sample.

Constructing a Confidence Interval for μ

When the conditions are met, the sampling distribution of $\bar{x}$ has roughly a Normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$. Because we don't know σ, we estimate it by the sample standard deviation s_x. We then estimate the standard deviation of the sampling distribution with $\mathrm{SE}_{\bar{x}} = \frac{s_x}{\sqrt{n}}$. This value is called the **standard error of the sample mean** $\bar{x}$, or just the standard error of the mean.

As with proportions, some books refer to the standard deviation of the sampling distribution of $\bar{x}$ as the "standard error" and what we call the standard error of the mean as the "estimated standard error." The standard error of the mean is often abbreviated SEM.

DEFINITION: Standard error of the sample mean

The **standard error of the sample mean** $\bar{x}$ is $\mathrm{SE}_{\bar{x}} = \frac{s_x}{\sqrt{n}}$, where s_x is the sample standard deviation. It describes how far $\bar{x}$ will typically be from μ in repeated SRSs of size n.

To construct a confidence interval for μ, replace the standard deviation $\sigma/\sqrt{n}$ of $\bar{x}$ by its standard error $s_x/\sqrt{n}$ in the formula for the one-sample z interval for a population mean. Use critical values from the t distribution with $n - 1$ degrees of freedom in place of the z critical values. That is,

$$\begin{aligned} &\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic}) \\ &= \bar{x} \pm t^* \frac{s_x}{\sqrt{n}} \end{aligned}$$

This **one-sample *t* interval for a population mean** is similar in both reasoning and computational detail to the one-sample z interval for a population proportion of Section 8.2. So we will now pay more attention to questions about using these methods in practice.

THE ONE-SAMPLE *t* INTERVAL FOR A POPULATION MEAN

When the conditions are met, a C% confidence interval for the unknown mean μ is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

where t^* is the critical value for the t_{n-1} distribution, with C% of the area between $-t^*$ and t^*.

The following example shows you how to construct a confidence interval for a population mean when σ is unknown. By now, you should recognize the four-step process.

EXAMPLE Auto Pollution

STEP 4

A one-sample t interval for μ

Environmentalists, government officials, and vehicle manufacturers are all interested in studying the auto exhaust emissions produced by motor vehicles. The major pollutants in auto exhaust from gasoline engines are hydrocarbons, carbon monoxide, and nitrogen oxides (NOX). Researchers collected data on the NOX levels (in grams/mile) for a random sample of 40 light-duty engines of the same type. The mean NOX reading was 1.2675 and the standard deviation was 0.3332.[18]

PROBLEM:

(a) Construct and interpret a 95% confidence interval for the mean amount of NOX emitted by light-duty engines of this type.

(b) The Environmental Protection Agency (EPA) sets a limit of 1.0 gram/mile for average NOX emissions. Are you convinced that this type of engine violates the EPA limit? Use your interval from (a) to support your answer.

SOLUTION:

(a) **STATE:** We want to estimate the true mean amount μ of NOX emitted by all light-duty engines of this type at a 95% confidence level.

PLAN: We should construct a one-sample t interval for μ if the conditions are met.

- Random: The data come from a random sample of 40 light-duty engines of this type.
  - 10%: We are sampling without replacement, so we need to assume that there are at least $10(40) = 400$ light-duty engines of this type.
- Normal/Large Sample: We don't know whether the population distribution of NOX emissions is Normal. Because the sample size is large ($n = 40 \geq 30$), we should be safe using a t distribution.

DO: The formula for the one-sample t interval is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

From the information given, $\bar{x} = 1.2675$ g/mi and $s_x = 0.3332$ g/mi. To find the critical value t^*, we use the t distribution with df $= 40 - 1 = 39$. Unfortunately, there is no row corresponding to 39 degrees of freedom in Table B. We can't pretend we have a larger sample size than we actually do, so we use the more conservative df $= 30$.

Upper-tail probability p			
df	.05	.025	.02
29	1.699	2.045	2.150
30	1.697	2.042	2.147
40	1.684	2.021	2.123
	90%	95%	96%
	Confidence level C		

At a 95% confidence level, the critical value is $t^* = 2.042$. So the 95% confidence interval for μ is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} = 1.2675 \pm 2.042 \frac{0.3332}{\sqrt{40}} = 1.2675 \pm 0.1076 = (1.1599, 1.3751)$$

Using technology: The command `invT(.025,39)` gives $t = -2.023$. Using the critical value $t^* = 2.023$ for the 95% confidence interval gives

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} = 1.2675 \pm 2.023 \frac{0.3332}{\sqrt{40}} = 1.2675 \pm 0.1066 = (1.1609, 1.3741)$$

This interval is slightly narrower than the one found using Table B.

CONCLUDE: We are 95% confident that the interval from 1.1609 to 1.3741 grams/mile captures the true mean level of nitrogen oxides emitted by this type of light-duty engine.

(b) The confidence interval from (a) tells us that any value from 1.1609 to 1.3741 g/mi is a plausible value of the mean NOX level μ for this type of engine. Because the entire interval exceeds 1.0, it appears that this type of engine violates EPA limits.

For Practice *Try Exercise* 65

Now that we've calculated our first confidence interval for a population mean μ, it's time to make a simple observation. Inference for *proportions* uses z; inference for *means* uses t. That's one reason why distinguishing categorical from quantitative variables is so important.

Here is another example, this time with a smaller sample size.

EXAMPLE Video Screen Tension

STEP 4

Constructing a confidence interval for μ

A manufacturer of high-resolution video terminals must control the tension on the mesh of fine wires that lies behind the surface of the viewing screen. Too much tension will tear the mesh, and too little will allow wrinkles. The tension is measured by an electrical device with output readings in millivolts (mV). Some variation is inherent in the production process. Here are the tension readings from a random sample of 20 screens from a single day's production:

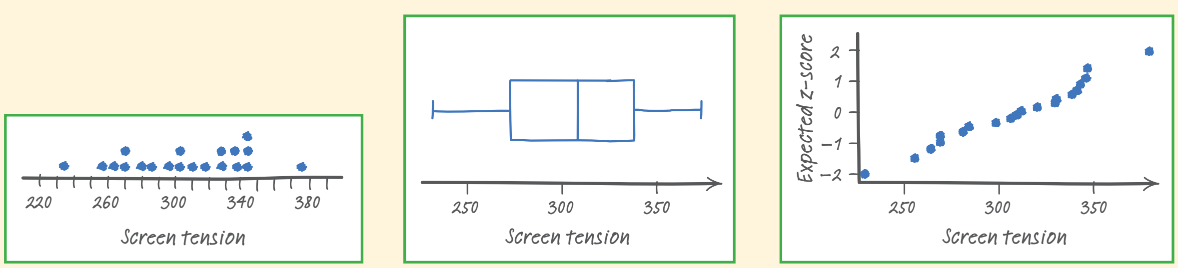
269.5	297.0	269.6	283.3	304.8	280.4	233.5	257.4	317.5	327.4
264.7	307.7	310.0	343.3	328.1	342.6	338.8	340.1	374.6	336.1

Construct and interpret a 90% confidence interval for the mean tension μ of all the screens produced on this day.

STATE: We want to estimate the true mean tension μ of all the video terminals produced this day with 90% confidence.

PLAN: If the conditions are met, we should use a one-sample *t* interval to estimate μ.

- Random: We are told that the data come from a random sample of 20 screens produced that day.
  - 10%: Because we are sampling without replacement, we must assume that at least 10(20) = 200 video terminals were produced this day.
- Normal/Large Sample: Because the sample size is small ($n = 20$), we must check whether it's reasonable to believe that the population distribution is Normal. So we examine the sample data. Figure 8.15 shows (a) a dotplot, (b) a boxplot, and (c) a Normal probability plot of the tension



(a) (b) (c)

FIGURE 8.15 (a) A dotplot, (b) boxplot, and (c) Normal probability plot of the video screen tension readings.

When the sample size is small ($n < 30$), as in this example, the Normal condition is about the shape of the *population* distribution. We look at a graph of the sample data to see if it's believable that the data came from a Normal population.

readings in the sample. Neither the dotplot nor the boxplot shows strong skewness or any outliers. The Normal probability plot looks roughly linear. These graphs give us no reason to doubt the Normality of the population.

DO: We used our calculator to find the mean and standard deviation of the tension readings for the 20 screens in the sample: $\bar{x} = 306.32$ mV and $s_x = 36.21$ mV. We use the t distribution with df = 19 to find the critical value. For a 90% confidence level, the critical value is $t^* = 1.729$. So the 90% confidence interval for μ is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} = 306.32 \pm 1.729 \frac{36.21}{\sqrt{20}} = 306.32 \pm 14.00 = (292.32, 320.32)$$

Upper-tail probability p			
df	.10	.05	.025
18	1.330	1.734	2.101
19	1.328	1.729	2.093
20	1.325	1.725	2.086
	80%	90%	95%
	Confidence level C		

Using technology: The calculator's `invT(.05,19)` gives $t = -1.729$, which matches the $t^* = 1.729$ critical value we got from the table.

CONCLUDE: We are 90% confident that the interval from 292.32 to 320.32 mV captures the true mean tension in the entire batch of video terminals produced that day.

For Practice *Try Exercise* 69

AP® EXAM TIP It is not enough just to make a graph of the data on your calculator when assessing Normality. You must *sketch* the graph on your paper to receive credit. You don't have to draw multiple graphs—any appropriate graph will do.

As you probably guessed, your calculator will compute a one-sample t interval for a population mean from sample data or summary statistics.

17. TECHNOLOGY CORNER — ONE-SAMPLE t INTERVALS FOR μ ON THE CALCULATOR

TI-Nspire instructions in Appendix B; HP Prime instructions on the book's Web site.

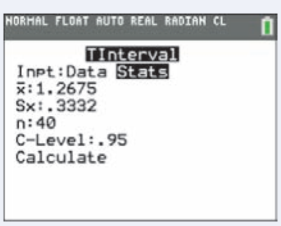
Confidence intervals for a population mean using t distributions can be constructed on the TI-83/84 and TI-89, thus avoiding the use of Table B. Here is a brief summary of the techniques when you have the actual data values and when you have only numerical summaries.

1. Using summary statistics (see auto pollution example, page 519)

TI-83/84

From the home screen,

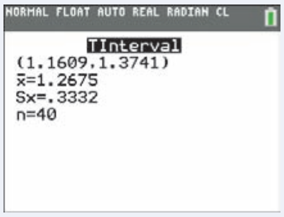
- Press STAT, arrow over to TESTS, and choose `TInterval....`
- On the TInterval screen, adjust your settings as shown and choose `Calculate`.

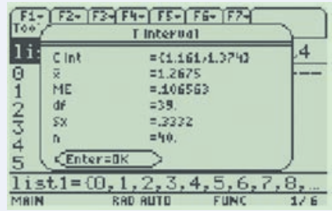


TI-89

From inside the Statistics/List Editor,

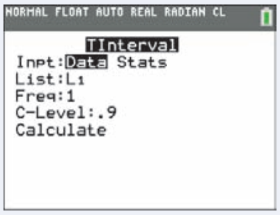
- Press 2nd F2 ([F7]) to go into the intervals (Ints) menu, then choose `TInterval....`
- Choose "Stats" as the Data Input Method.
- On the TInterval screen, adjust your settings as shown and press ENTER.

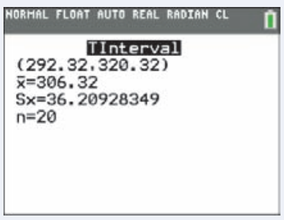




2. Using raw data (see video screen tension example, page 520)

Enter the 20 video screen tension readings data in L1/list1. Proceed to the TInterval screen as in Step 1, but choose Data as the input method. Then adjust your settings as shown and calculate the interval.





CHECK YOUR UNDERSTANDING

Biologists studying the healing of skin wounds measured the rate at which new cells closed a cut made in the skin of an anesthetized newt. Here are data from a random sample of 18 newts, measured in micrometers (millionths of a meter) per hour:[19]

29	27	34	40	22	28	14	35	26	35	12	30	23	18	11	22	23	33

Calculate and interpret a 95% confidence interval for the mean healing rate μ.

Choosing the Sample Size

A wise user of statistics never plans data collection without planning the inference at the same time. You can arrange to have both high confidence and a small margin of error by taking enough observations. When the population standard deviation σ is unknown and conditions are met, the $C\%$ confidence interval for μ is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

where t^* is the critical value for confidence level C and degrees of freedom df $= n - 1$. The margin of error (ME) of the confidence interval is

$$ME = t^* \frac{s_x}{\sqrt{n}}$$

To determine the sample size for a desired margin of error, it makes sense to set the expression for ME less than or equal to the specified value and solve the inequality for n. There are two problems with this approach:

1. We don't know the sample standard deviation s_x because we haven't produced the data yet.
2. The critical value t^* depends on the sample size n that we choose.

The second problem is more serious. To get the correct value of t^*, we need to know the sample size. But that's what we're trying to find! There is no easy solution to this problem.

One alternative (the one we recommend!) is to come up with a reasonable estimate for the *population* standard deviation σ from a similar study that was done in the past or from a small-scale pilot study. By pretending that σ is known, we can use the one-sample z interval for μ:

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

Using the appropriate standard Normal critical value z^* for confidence level C, we can solve

$$z^* \frac{\sigma}{\sqrt{n}} \leq ME$$

for n. Here is a summary of this strategy.

There are other methods of determining sample size that do not require us to use a known value of the population standard deviation σ. These methods are beyond the scope of this text. Our advice: consult with a statistician when planning your study!

CHOOSING SAMPLE SIZE FOR A DESIRED MARGIN OF ERROR WHEN ESTIMATING μ

To determine the sample size n that will yield a C% confidence interval for a population mean with a specified margin of error ME:

- Get a reasonable value for the population standard deviation σ from an earlier or pilot study.
- Find the critical value z^* from a standard Normal curve for confidence level C.
- Set the expression for the margin of error to be less than or equal to ME and solve for n:

$$z^* \frac{\sigma}{\sqrt{n}} \leq ME$$

The procedure is best illustrated with an example.

EXAMPLE How Many Monkeys?

Determining sample size from margin of error

Researchers would like to estimate the mean cholesterol level μ of a particular variety of monkey that is often used in laboratory experiments. They would like their estimate to be within 1 milligram per deciliter (mg/dl) of the true value of μ at a 95% confidence level. A previous study involving this variety of monkey suggests that the standard deviation of cholesterol level is about 5 mg/dl.

PROBLEM: Obtaining monkeys for research is time-consuming, expensive, and controversial. What is the minimum number of monkeys the researchers will need to get a satisfactory estimate?

SOLUTION: For 95% confidence, $z^* = 1.96$. We will use $\sigma = 5$ as our best guess for the standard deviation of the monkeys' cholesterol level. Set the expression for the margin of error to be at most 1 and solve for n:

$$1.96 \frac{5}{\sqrt{n}} \leq 1$$

$$\frac{(1.96)(5)}{1} \leq \sqrt{n}$$

$$96.04 \leq n$$

Remember: always round up to the next whole number when finding n.

Because 96 monkeys would give a slightly larger margin of error than desired, the researchers would need 97 monkeys to estimate the cholesterol levels to their satisfaction. (On learning the cost of getting this many monkeys, the researchers might want to consider studying rats instead!)

For Practice *Try Exercise* 73

Taking observations costs time and money. The required sample size may be impossibly expensive. *Notice that it is the size of the sample that determines the margin of error. The size of the population does not influence the sample size we need.* This is true as long as the population is much larger than the sample.

CHECK YOUR UNDERSTANDING

Administrators at your school want to estimate how much time students spend on homework, on average, during a typical week. They want to estimate μ at the 90% confidence level with a margin of error of at most 30 minutes. A pilot study indicated that the standard deviation of time spent on homework per week is about 154 minutes.

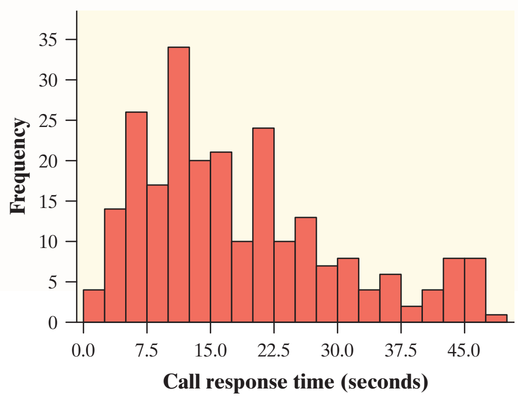
How many students need to be surveyed to meet the administrators' goal? Show your work.

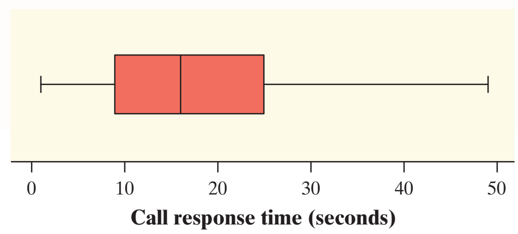
case closed

Need Help? Give Us a Call!

Refer to the chapter-opening Case Study on page 475. The bank manager wants to know whether or not the bank's customer service agents generally met the goal of answering incoming calls in less than 30 seconds. We can approach this question in two ways: by estimating the proportion p of all calls that were answered within 30 seconds or by estimating the mean response time μ.

Some graphs and numerical summaries of the data are provided below.





Descriptive Statistics: Call response time (sec)

Variable	N	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Call response time (sec)	241	18.353	0.758	11.761	1.000	9.000	16.000	25.000	49.000

1. Describe the distribution of call response times for the random sample of 241 calls.
2. About what proportion of the call response times in the sample were less than 30 seconds? Explain how you got your answer.
3. The bank's manager would like to estimate the true proportion p of calls to the bank's customer service center that are answered in less than 30 seconds.
   (a) What conditions must be met to calculate a 95% confidence interval for p? Show that the conditions are met in this case.
   (b) Explain the meaning of 95% confidence in this setting.
   (c) A 95% confidence interval for p is (0.783, 0.877). Give the margin of error and show how it was calculated.
   (d) Interpret the interval from part (c) in context.
4. Construct and interpret a 95% confidence interval for the true mean response time of calls to the bank's customer service center.
5. Is the customer service center meeting its goal of answering calls in less than 30 seconds? Give appropriate evidence to support your answer.

Section 8.3 Summary

- **Confidence intervals for the mean μ** of a Normal population are based on the sample mean $\bar{x}$ of an SRS. If we somehow know σ, we use the z critical value and the standard Normal distribution to help calculate confidence intervals.
- In practice, we usually don't know σ. Replace the standard deviation $\sigma/\sqrt{n}$ of the sampling distribution of $\bar{x}$ by the **standard error** $SE_{\bar{x}} = s_x/\sqrt{n}$ and use the t distribution with $n - 1$ **degrees of freedom (df).**
- There is a ***t* distribution** for every positive degrees of freedom. All t distributions are unimodal, symmetric, and centered at 0. The t distributions approach the standard Normal distribution as the number of degrees of freedom increases.
- The conditions for constructing a confidence interval about a population mean are
    - **Random:** The data were produced by a well-designed random sample or randomized experiment.
        - **10%:** When sampling without replacement, check that the population is at least 10 times as large as the sample.
    - **Normal/Large Sample:** The population distribution is Normal or the sample size is large ($n \geq 30$). When the sample size is small ($n < 30$), examine a graph of the sample data for any possible departures from Normality in the population. You should be safe using a t distribution as long as there is no strong skewness and no outliers are present.
- When conditions are met, a C% confidence interval for the mean μ is given by the **one-sample *t* interval**

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

The critical value t^* is chosen so that the t curve with $n - 1$ degrees of freedom has C% of the area between $-t^*$ and t^*.

STEP 4

- Follow the four-step process—State, Plan, Do, Conclude—whenever you are asked to construct and interpret a confidence interval for a population mean Remember: inference for proportions uses z; inference for means uses t.
- The sample size needed to obtain a confidence interval with approximate margin of error ME for a population mean involves solving

$$z^* \frac{\sigma}{\sqrt{n}} \leq ME$$

for n, where the standard deviation σ is a reasonable value from a previous or pilot study, and z^* is the critical value for the level of confidence we want.

8.3 TECHNOLOGY CORNERS

TI-Nspire Instructions in Appendix B; HP Prime instructions on the book's Web site.

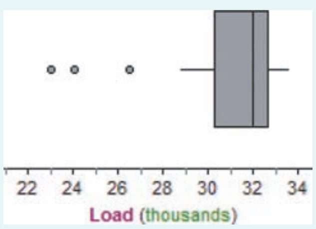
16. Inverse *t* on the calculator page 513
17. One-sample *t* intervals for μ on the calculator page 521

Section 8.3 Exercises

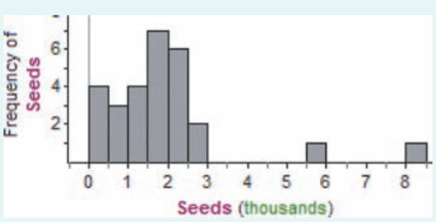
55. **Critical values** What critical value t^* from Table B would you use for a confidence interval for the population mean in each of the following situations?
pg 513
(a) A 95% confidence interval based on $n = 10$ randomly selected observations
(b) A 99% confidence interval from an SRS of 20 observations
(c) A 90% confidence interval based on a random sample of 77 individuals

56. **Critical values** What critical value t^* from Table B should be used for a confidence interval for the population mean in each of the following situations?
(a) A 90% confidence interval based on $n = 12$ randomly selected observations
(b) A 95% confidence interval from an SRS of 30 observations
(c) A 99% confidence interval based on a random sample of size 58

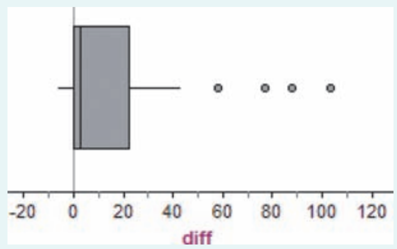
57. **Pulling wood apart** How heavy a load (pounds) is needed to pull apart pieces of Douglas fir 4 inches long and 1.5 inches square? A random sample of 20 similar pieces of Douglas fir from a large batch was selected for a science class. The Fathom boxplot below shows the class's data. Explain why it would not be wise to use a t critical value to construct a confidence interval for the population mean μ.



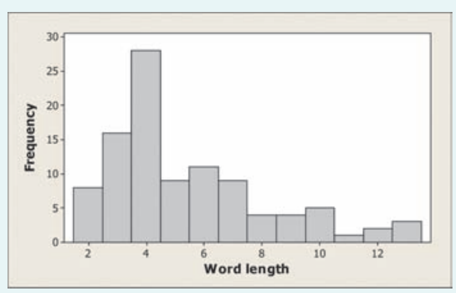
58. **Weeds among the corn** Velvetleaf is a particularly annoying weed in cornfields. It produces lots of seeds, and the seeds wait in the soil for years until conditions are right for sprouting. How many seeds do velvetleaf plants produce? The Fathom histogram below shows the counts from a random sample of 28 plants that came up in a cornfield when no herbicide was used.[20] Explain why it would not be wise to use a t critical value to construct a confidence interval for the mean number of seeds μ produced by velvetleaf plants.



59. **Should we use t?** Determine whether we can safely use a t^* critical value to calculate a confidence interval for the population mean in each of the following settings.
pg 516
(a) We collect data from a random sample of adult residents in a state. Our goal is to estimate the overall percent of adults in the state who are college graduates.
(b) The coach of a college men's basketball team records the resting heart rates of the 15 team members. We use these data to construct a confidence interval for the mean resting heart rate of all male students at this college.
(c) Do teens text more than they call? To find out, an AP® Statistics class at a large high school collected data on the number of text messages and calls sent or received by each of 25 randomly selected students. The Fathom boxplot below displays the difference (texts − calls) for each student.



60. **Should we use t?** Determine whether we can safely use a t^* critical value to calculate a confidence interval for the population mean in each of the following settings.
(a) We want to estimate the average age at which U.S. presidents have died. So we obtain a list of all U.S. presidents who have died and their ages at death.
(b) How much time do students spend on the Internet? We collect data from the 32 members of our AP® Statistics class and calculate the mean amount of time that each student spent on the Internet yesterday.
(c) Judy is interested in the reading level of a medical journal. She records the length of a random sample of 100 words. The Minitab histogram below displays the data.



61. **Blood pressure** A medical study finds that $\bar{x} = 114.9$ and $s_x = 9.3$ for the seated systolic blood pressure of the 27 members of one treatment group. What is the standard error of the mean? Interpret this value in context.

62. **Travel time to work** A study of commuting times reports the travel times to work of a random sample of 20 employed adults in New York State. The mean is $\bar{x} = 31.25$ minutes, and the standard deviation is $s_x = 21.88$ minutes. What is the standard error of the mean? Interpret this value in context.

63. **Willows in Yellowstone** Writers in some fields summarize data by giving $\bar{x}$ and its standard error rather than $\bar{x}$ and s_x. Biologists studying willow plants in Yellowstone National Park reported their results in a table with columns labeled $\bar{x} \pm$ SE. The table entry for the heights of willow plants (in centimeters) in one region of the park was 61.55 ± 19.03.[21] The researchers measured a total of 23 plants.

(a) Find the sample standard deviation s_x for these measurements. Show your work.

(b) A hasty reader believes that the interval given in the table is a 95% confidence interval for the mean height of willow plants in this region of the park. Find the actual confidence level for the given interval.

64. **Blink** When two lights close together blink alternately, we "see" one light moving back and forth if the time between blinks is short. What is the longest interval of time between blinks that preserves the illusion of motion? Ask subjects to turn a knob that slows the blinking until they "see" two lights rather than one light moving. A report gives the results in the form "mean plus or minus the standard error of the mean."[22] Data for 12 subjects are summarized as 251 ± 45 (in milliseconds).

(a) Find the sample standard deviation s_x for these measurements. Show your work.

(b) A hasty reader believes that the interval given in the report is a 95% confidence interval for the population mean. Find the actual confidence level for the given interval.

65. **Bone loss by nursing mothers** Breast-feeding mothers secrete calcium into their milk. Some of the calcium may come from their bones, so mothers may lose bone mineral. Researchers measured the percent change in bone mineral content (BMC) of the spines of 47 randomly selected mothers during three months of breast-feeding.[23] The mean change in BMC was −3.587% and the standard deviation was 2.506%.

pg 519

STEP 4

(a) Construct and interpret a 99% confidence interval to estimate the mean percent change in BMC in the population.

(b) Based on your interval from part (a), do these data give good evidence that on the average nursing mothers lose bone mineral? Explain.

66. **Reading scores in Atlanta** The Trial Urban District Assessment (TUDA) is a government-sponsored study of student achievement in large urban school districts. TUDA gives a reading test scored from 0 to 500. A score of 243 is a "basic" reading level and a score of 281 is "proficient." Scores for a random sample of 1470 eighth-graders in Atlanta had $\bar{x} = 240$ with standard deviation 42.17.[24]

STEP 4

(a) Calculate and interpret a 99% confidence interval for the mean score of all Atlanta eighth-graders.

(b) Based on your interval from part (a), is there good evidence that the mean for all Atlanta eighth-graders is less than the basic level? Explain.

67. **Men and muscle** Ask young men to estimate their own degree of body muscle by choosing from a set of 100 photos. Then ask them to choose what they believe women prefer. The researchers know the actual degree of muscle, measured as kilograms per square meter of fat-free mass, for each of the photos. They can therefore measure the difference between what a subject thinks women prefer and the subject's own self-image. Call this difference the "muscle gap." Here are summary statistics for the muscle gap from a random sample of 200 American and European young men: $\bar{x} = 2.35$ and $s_x = 2.5$.[25]

STEP 4

(a) Calculate and interpret a 95% confidence interval for the mean size of the muscle gap for the population of American and European young men.

(b) A graph of the sample data is strongly skewed to the right. Explain why this information does not invalidate the interval you calculated in part (a).

68. **A big-toe problem** A bunion on the big toe is fairly uncommon in youth and often requires surgery. Doctors used X-rays to measure the angle (in degrees) of deformity on the big toe in a random sample of 37 patients under the age of 21 who came to a medical center for surgery to correct a bunion. The angle is a measure of the seriousness of the deformity. For these 37 patients, the mean angle of deformity was 24.76 degrees and the standard deviation was 6.34 degrees. A dotplot of the data revealed no outliers or strong skewness.[26]

STEP 4

(a) Construct and interpret a 90% confidence interval for the mean angle of deformity in the population of all such patients.

(b) Researchers omitted one patient with a deformity angle of 50 degrees from the analysis due to a measurement issue. What effect would including this outlier have on the confidence interval in part (a)? Justify your answer without doing any calculations.

69. **Give it some gas!** Computers in some vehicles calculate various quantities related to performance. One of these is fuel efficiency, or gas mileage, usually expressed as miles per gallon (mpg). For one vehicle equipped in this way, the miles per gallon were recorded each time the gas tank was filled and the computer was then reset.[27] Here are the mpg values for a random sample of 20 of these records:

pg 520

STEP 4

15.8	13.6	15.6	19.1	22.4	15.6	22.5	17.2	19.4	22.6
19.4	18.0	14.6	18.7	21.0	14.8	22.6	21.5	14.3	20.9

Construct and interpret a 95% confidence interval for the mean fuel efficiency μ for this vehicle.

70. **Vitamin C content** Several years ago, the U.S. Agency for International Development provided 238,300 metric tons of corn-soy blend (CSB) for emergency relief in countries throughout the world. CSB is a highly nutritious, low-cost fortified food. As part of a study to evaluate appropriate vitamin C levels in this food, measurements were taken on samples of CSB produced in a factory.[28] The following data are the amounts of vitamin C, measured in milligrams per 100 grams (mg/100 g) of blend, for a random sample of size 8 from one production run:

26	31	23	22	11	22	14	31

Construct and interpret a 95% confidence interval for the mean amount of vitamin C μ in the CSB from this production run.

71. **Paired tires** Researchers were interested in comparing two methods for estimating tire wear. The first method used the amount of weight lost by a tire. The second method used the amount of wear in the grooves of the tire. A random sample of 16 tires was obtained. Both methods were used to estimate the total distance traveled by each tire. The table below provides the two estimates (in thousands of miles) for each tire.[29]

(a) Construct and interpret a 95% confidence interval for the mean difference μ in the estimates from these two methods in the population of tires.

(b) Does your interval in part (a) give convincing evidence of a difference in the two methods of estimating tire wear? Justify your answer.

Tire	Weight	Groove	Diff.	Tire	Weight	Groove	Diff.
1	45.9	35.7	10.2	9	30.4	23.1	7.3
2	41.9	39.2	2.7	10	27.3	23.7	3.6
3	37.5	31.1	6.4	11	20.4	20.9	−0.5
4	33.4	28.1	5.3	12	24.5	16.1	8.4
5	31.0	24.0	7.0	13	20.9	19.9	1.0
6	30.5	28.7	1.8	14	18.9	15.2	3.7
7	30.9	25.9	5.0	15	13.7	11.5	2.2
8	31.9	23.3	8.6	16	11.4	11.2	0.2

72. **Water** Trace metals found in wells affect the taste of drinking water, and high concentrations can pose a health risk. Researchers measured the concentration of zinc (in milligrams/liter) near the top and the bottom of 10 randomly selected wells in a large region. The data are provided in the table below.[30]

(a) Construct and interpret a 95% confidence interval for the mean difference μ in the zinc concentrations from these two locations in the wells.

(b) Does your interval in part (a) give convincing evidence of a difference in zinc concentrations at the top and bottom of wells in the region? Justify your answer.

Well:	1	2	3	4	5	6	7	8	9	10
Bottom:	0.430	0.266	0.567	0.531	0.707	0.716	0.651	0.589	0.469	0.723
Top:	0.415	0.238	0.390	0.410	0.605	0.609	0.632	0.523	0.411	0.612
Difference:	0.015	0.028	0.177	0.121	0.102	0.107	0.019	0.066	0.058	0.111

73. **Estimating BMI** pg 524 The body mass index (BMI) of all American young women is believed to follow a Normal distribution with a standard deviation of about 7.5. How large a sample would be needed to estimate the mean BMI μ in this population to within ± 1 with 99% confidence? Show your work.

74. **The SAT again** High school students who take the SAT Math exam a second time generally score higher than on their first try. Past data suggest that the score increase has a standard deviation of about 50 points. How large a sample of high school students would be needed to estimate the mean change in SAT score to within 2 points with 95% confidence? Show your work.

Multiple choice: Select the best answer for Exercises 75 to 78.

75. One reason for using a t distribution instead of the standard Normal curve to find critical values when calculating a level C confidence interval for a population mean is that

(a) z can be used only for large samples.

(b) z requires that you know the population standard deviation σ.

(c) z requires that you can regard your data as an SRS from the population.

(d) z requires that the sample size is at most 10% of the population size.

(e) a z critical value will lead to a wider interval than a t critical value.

76. You have an SRS of 23 observations from a large population. The distribution of sample values is roughly symmetric with no outliers. What critical value would you use to obtain a 98% confidence interval for the mean of the population?

(a) 2.177 (b) 2.183 (c) 2.326 (d) 2.500 (e) 2.508

77. A quality control inspector will measure the salt content (in milligrams) in a random sample of bags of potato chips from an hour of production. Which of the following would result in the smallest margin of error in estimating the mean salt content μ?

(a) 90% confidence; $n = 25$

(b) 90% confidence; $n = 50$

(c) 95% confidence; $n = 25$

(d) 95% confidence; $n = 50$

(e) $n = 100$ at any confidence level

78. Scientists collect data on the blood cholesterol levels (milligrams per deciliter of blood) of a random sample of 24 laboratory rats. A 95% confidence interval for the mean blood cholesterol level μ is 80.2 to 89.8. Which of the following would cause the most worry about the validity of this interval?

(a) There is a clear outlier in the data.

(b) A stemplot of the data shows a mild right skew.

(c) You do not know the population standard deviation σ.

(d) The population distribution is not exactly Normal.

(e) None of these are a problem when using a t interval.

79. **Watching TV** (6.1, 7.3) Choose a young person (aged 19 to 25) at random and ask, "In the past seven days, how many days did you watch television?" Call the response X for short. Here is the probability distribution for X:[31]

Days:	0	1	2	3	4	5	6	7
Probability:	0.04	0.03	0.06	0.08	0.09	0.08	0.05	???

(a) What is the probability that $X = 7$? Justify your answer.

(b) Calculate the mean of the random variable X. Interpret this value in context.

(c) Suppose that you asked 100 randomly selected young people (aged 19 to 25) to respond to the question and found that the mean $\bar{x}$ of their responses was 4.96. Would this result surprise you? Justify your answer.

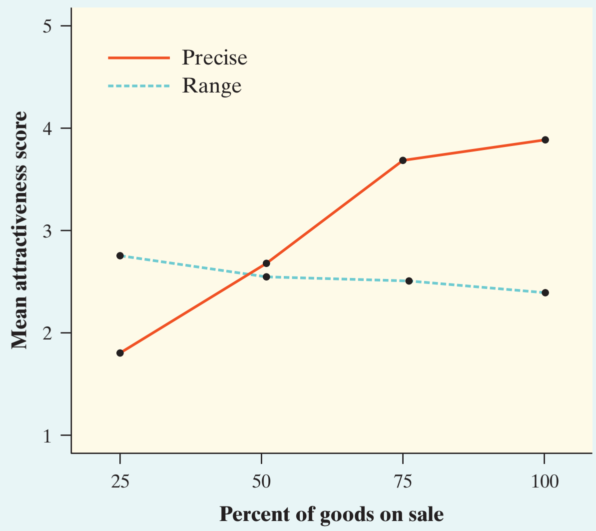
80. **Price cuts** (4.2) Stores advertise price reductions to attract customers. What type of price cut is most attractive? Experiments with more than one factor allow insight into interactions between the factors. A study of the attractiveness of advertised price discounts had two factors: percent of all foods on sale (25%, 50%, 75%, or 100%) and whether the discount was stated precisely (as in, for example, "60% off") or as a range (as in "40% to 70% off"). Subjects rated the attractiveness of the sale on a scale of 1 to 7.

(a) Describe a completely randomized design using 200 student subjects.

(b) Explain how you would use the partial table of random digits below to assign subjects to treatment groups. Then use your method to select the first 3 subjects for one of the treatment groups. Show your work clearly on your paper.

45740	41807	65561	33302	07051	93623	18132	09547
12975	13258	13048	45144	72321	81940	00360	02428

(c) The figure below shows the mean ratings for the eight treatments formed from the two factors.[32] Based on these results, write a careful description of how percent on sale and precise discount versus range of discounts influence the attractiveness of a sale.



FRAPPY! Free Response AP® Problem, Yay!

The following problem is modeled after actual AP® Statistics exam free response questions. Your task is to generate a complete, concise response in 15 minutes.

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

Members at a popular fitness club currently pay a \$40 per month membership fee. The owner of the club wants to raise the fee to \$50 but is concerned that some members will leave the gym if the fee increases. To investigate, the owner plans to survey a random sample of the club members and construct a 95% confidence interval for the proportion of all members who would quit if the fee was raised to \$50.

(a) Explain the meaning of "95% confidence" in the context of the study.

(b) After the owner conducted the survey, he calculated the confidence interval to be 0.18 ± 0.075. Interpret this interval in the context of the study.

(c) According to the club's accountant, the fee increase will be worthwhile if fewer than 20% of the members quit. According to the interval from part (b), can the owner be confident that the fee increase will be worthwhile? Explain.

(d) One of the conditions for calculating the confidence interval in part (b) is that $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$. Explain why it is necessary to check this condition.

After you finish, you can view two example solutions on the book's Web site (www.whfreeman.com/tps5e). Determine whether you think each solution is "complete," "substantial," "developing," or "minimal." If the solution is not complete, what improvements would you suggest to the student who wrote it? Finally, your teacher will provide you with a scoring rubric. Score your response and note what, if anything, you would do differently to improve your own score.

Chapter Review

Section 8.1: Confidence Intervals: The Basics

In this section, you learned that a point estimate is the single best guess for the value of a population parameter. You also learned that a confidence interval provides an interval of plausible values for a parameter. To interpret a confidence interval, say, "We are $C\%$ confident that the interval from ___ to ___ captures the [parameter in context]," where C is the confidence level of the interval.

The confidence level C describes the percentage of confidence intervals that we expect to capture the value of the parameter. To interpret a $C\%$ confidence level, we say, "If we took many samples of the same size and used them to construct $C\%$ confidence intervals, about $C\%$ of those intervals would capture the [parameter in context]."

Confidence intervals are formed by including a margin of error on either side of the point estimate. The size of the margin of error is determined by several factors, including the confidence level C and the sample size n. Increasing the sample size n makes the standard deviation of our estimate smaller, decreasing the margin of error. Increasing the confidence level C makes the margin of error larger, to ensure that the capture rate of the interval increases to $C\%$. Remember that the margin of error only accounts for sampling variability—it does not account for any bias in the data collection process.

Section 8.2: Estimating a Population Proportion

In this section, you learned how to construct and interpret confidence intervals for a population proportion. Several important conditions must be met for this type of confidence interval to be valid. First, the data used to calculate the interval must come from a well-designed random sample or randomized experiment (the Random condition). When the sample is taken without replacement from the population, the sample size should be no more than 10% of the population size (the 10% condition). Finally, the observed number of successes $n\hat{p}$ and observed number of failures $n(1-\hat{p})$ must both be at least 10 (the Large Counts condition).

The formula for calculating a confidence interval for a population proportion is

$$\hat{p} \pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where $\hat{p}$ is the sample proportion, z^* is the critical value, and n is the sample size. The value of z^* is based on the confidence level C. To find z^*, use Table A or technology to determine the values of z^* and $-z^*$ that capture the middle $C\%$ of the standard Normal distribution.

The four-step process (State, Plan, Do, Conclude) is perfectly suited for problems that ask you to construct and interpret a confidence interval. You should *state* the parameter you are estimating and at what confidence level, *plan* your work by naming the type of interval you will use and checking the appropriate conditions, *do* the calculations, and make a *conclusion* in the context of the problem. You can use technology for the Do step, but make sure that you identify the procedure you are using and type in the values correctly.

Finally, an important part of planning a study is determining the size of the sample to be selected. The necessary sample size is based on the confidence level, the proportion of successes, and the desired margin of error. To calculate the minimum sample size, solve the following inequality for n, where $\hat{p}$ is a guessed value for the sample proportion:

$$z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le ME$$

If you do not have an approximate value of $\hat{p}$ from a previous study or a pilot study, use $\hat{p} = 0.5$ to determine the sample size that will yield a value less than or equal to the desired margin of error.

Section 8.3: Estimating a Population Mean

In this section, you learned how to construct and interpret confidence intervals for a population mean. Remember that you have to check conditions before doing calculations. The Random and 10% conditions are the same as those for proportions. There's one new condition for means: the population must be Normally distributed or the sample size must be at least 30 (the Normal/Large Sample condition). If the population shape is unknown and the sample size is less than 30, graph the sample data and check for strong skewness or outliers. If there is no strong skewness or outliers, it is reasonable to assume that the population distribution is approximately Normal.

The formula for calculating a confidence interval for a population mean is

$$\bar{x} \pm t^*\frac{s_x}{\sqrt{n}}$$

where $\bar{x}$ is the sample mean, t^* is the critical value, s_x is the sample standard deviation, and n is the sample size. We use a t critical value instead of a z critical value when the population standard deviation is unknown—which is almost always the case. The value of t^* is based on the confidence level C and the degrees of freedom (df $= n - 1$). To find t^*, use Table B or technology to determine the values of t^* and $-t^*$ that capture the middle $C\%$ of the appropriate t distribution. The t distributions are bell-shaped, symmetric, and centered at 0. However, they are more variable and have a shape slightly different from that of the standard Normal distribution.

You also learned how to estimate the sample size when planning a study, as in Section 8.2. To calculate the minimum sample size, solve the following inequality for n, where σ is a guessed value for the population standard deviation:

$$z^*\frac{\sigma}{\sqrt{n}} \le ME$$

What Did You Learn?

Learning Objective	Section	Related Example on Page(s)	Relevant Chapter Review Exercise(s)
Determine the point estimate and margin of error from a confidence interval.	8.1	481	R8.2
Interpret a confidence interval in context.	8.1	481, 484	R8.3, R8.4, R8.6, R8.7
Interpret a confidence level in context.	8.1	484	R8.2
Describe how the sample size and confidence level affect the length of a confidence interval.	8.1	Discussion on 487	R8.9
Explain how practical issues like nonresponse, undercoverage, and response bias can affect the interpretation of a confidence interval.	8.1	Discussion on 488	R8.6
State and check the Random, 10%, and Large Counts conditions for constructing a confidence interval for a population proportion.	8.2	494	R8.3
Determine critical values for calculating a *C*% confidence interval for a population proportion using a table or technology.	8.2	497	R8.1
Construct and interpret a confidence interval for a population proportion.	8.2	498, 500	R8.3, R8.6
Determine the sample size required to obtain a *C*% confidence interval for a population proportion with a specified margin of error.	8.2	502	R8.5
State and check the Random, 10%, and Normal/Large Sample conditions for constructing a confidence interval for a population mean.	8.3	516	R8.4
Explain how the *t* distributions are different from the standard Normal distribution and why it is necessary to use a *t* distribution when calculating a confidence interval for a population mean.	8.3	Discussion on 511–512	R8.10
Determine critical values for calculating a *C*% confidence interval for a population mean using a table or technology.	8.3	513	R8.1
Construct and interpret a confidence interval for a population mean.	8.3	519–520	R8.4, R8.7
Determine the sample size required to obtain a *C*% confidence interval for a population mean with a specified margin of error.	8.3	524	R8.8

Chapter 8 Chapter Review Exercises

These exercises are designed to help you review the important ideas and methods of the chapter.

R8.1 **It's critical** Find the appropriate critical value for constructing a confidence interval in each of the following settings.

(a) Estimating a population proportion p at a 94% confidence level based on an SRS of size 125.

(b) Estimating a population mean μ at a 99% confidence level based on an SRS of size 58.

R8.2 **Batteries** A company that produces AA batteries tests the lifetime of a random sample of 30 batteries using a special device designed to imitate real-world use. Based on the testing, the company makes the following statement: "Our AA batteries last an average of 430 to 470 minutes, and our confidence in that interval is 95%."[33]

(a) Determine the point estimate, margin of error, standard error, and sample standard deviation.

(b) A reporter translates the statistical announcement into "plain English" as follows: "If you buy one of this company's AA batteries, there is a 95% chance that it will last between 430 and 470 minutes." Comment on this interpretation.

(c) Your friend, who has just started studying statistics, claims that if you select 40 more AA batteries at random from those manufactured by this company, there is a 95% probability that the mean lifetime will fall between 430 and 470 minutes. Do you agree? Explain.

(d) Give a statistically correct interpretation of the confidence level that could be published in a newspaper report.

R8.3 **We love football!** A recent Gallup Poll conducted telephone interviews with a random sample of adults aged 18 and older. Data were obtained for 1000 people. Of these, 37% said that football is their favorite sport to watch on television.

(a) Define the parameter p in this setting. Explain to someone who knows nothing about statistics why we can't just say that 37% of all adults would say that football is their favorite sport to watch on television.

(b) Check the conditions for constructing a confidence interval for p.

(c) Construct a 95% confidence interval for p.

(d) Interpret the interval in context.

R8.4 **Smart kids** A school counselor wants to know how smart the students in her school are. She gets funding from the principal to give an IQ test to an SRS of 60 of the over 1000 students in the school. The mean IQ score was 114.98 and the standard deviation was 14.80.[34]

(a) Define the parameter μ in this setting.

(b) Check the conditions for constructing a confidence interval for μ.

(c) Construct a 90% confidence interval for the mean IQ score of students at the school.

(d) Interpret your result from part (c) in context.

R8.5 **Do you go to church?** The Gallup Poll plans to ask a random sample of adults whether they attended a religious service in the last 7 days. How large a sample would be required to obtain a margin of error of at most 0.01 in a 99% confidence interval for the population proportion who would say that they attended a religious service? Show your work.

R8.6 **Running red lights** A random digit dialing telephone survey of 880 drivers asked, "Recalling the last ten traffic lights you drove through, how many of them were red when you entered the intersections?" Of the 880 respondents, 171 admitted that at least one light had been red.[35]

(a) Construct and interpret a 95% confidence interval for the population proportion.

(b) Nonresponse is a practical problem for this survey—only 21.6% of calls that reached a live person were completed. Another practical problem is that people may not give truthful answers. What is the likely direction of the bias: Do you think more or fewer than 171 of the 880 respondents really ran a red light? Why? Are these sources of bias included in the margin of error?

R8.7 **Engine parts** Here are measurements (in millimeters) of a critical dimension on an SRS of 16 of the more than 200 auto engine crankshafts produced in one day:

224.120	224.001	224.017	223.982	223.989	223.961	223.960	224.089
223.987	223.976	223.902	223.980	224.098	224.057	223.913	223.999

(a) Construct and interpret a 95% confidence interval for the process mean at the time these crankshafts were produced.

(b) The process mean is supposed to be $\mu = 224$ mm but can drift away from this target during production. Does your interval from part (a) suggest that the process mean has drifted? Explain.

R8.8 **Good wood?** A lab supply company sells pieces of Douglas fir 4 inches long and 1.5 inches square for force experiments in science classes. From experience, the strength of these pieces of wood follows a Normal distribution with standard deviation 3000 pounds. You want to estimate the mean load needed to pull apart these pieces of wood to within 1000 pounds with 95% confidence. How large a sample is needed? Show your work.

R8.9 **It's about *ME*** Explain how each of the following would affect the margin of error of a confidence interval, if all other things remained the same.

(a) Increasing the confidence level

(b) Quadrupling the sample size

R8.10 ***t* time** When constructing confidence intervals for a population mean, we almost always use critical values from a t distribution rather than the standard Normal distribution.

(a) When is it necessary to use a t critical value rather than a z critical value when constructing a confidence interval for a population mean?

(b) Describe two ways that the t distributions are different from the standard Normal distribution.

(c) Explain what happens to the t distributions as the degrees of freedom increase.

Chapter 8 AP® Statistics Practice Test

Section I: Multiple Choice *Select the best answer for each question.*

T8.1 The Gallup Poll interviews 1600 people. Of these, 18% say that they jog regularly. The news report adds: "The poll had a margin of error of plus or minus three percentage points at a 95% confidence level." You can safely conclude that

(a) 95% of all Gallup Poll samples like this one give answers within ±3% of the true population value.
(b) the percent of the population who jog is certain to be between 15% and 21%.
(c) 95% of the population jog between 15% and 21% of the time.
(d) we can be 95% confident that the sample proportion is captured by the confidence interval.
(e) if Gallup took many samples, 95% of them would find that 18% of the people in the sample jog.

T8.2 The weights (in pounds) of three adult males are 160, 215, and 195. The standard error of the mean of these three weights is

(a) 190. (b) 27.84. (c) 22.73. (d) 16.07. (e) 13.13.

T8.3 In preparing to construct a one-sample t interval for a population mean, suppose we are not sure if the population distribution is Normal. In which of the following circumstances would we *not* be safe constructing the interval based on an SRS of size 24 from the population?

(a) A stemplot of the data is roughly bell-shaped.
(b) A histogram of the data shows slight skewness.
(c) A boxplot of the data has a large outlier.
(d) The sample standard deviation is large.
(e) A Normal probability plot of the data is fairly linear.

T8.4 Many television viewers express doubts about the validity of certain commercials. In an attempt to answer their critics, Timex Group USA wishes to estimate the proportion of consumers who believe what is shown in Timex television commercials. Let p represent the true proportion of consumers who believe what is shown in Timex television commercials. What is the smallest number of consumers that Timex can survey to guarantee a margin of error of 0.05 or less at a 99% confidence level?

(a) 550 (b) 600 (c) 650 (d) 700 (e) 750

T8.5 You want to compute a 90% confidence interval for the mean of a population with unknown population standard deviation. The sample size is 30. The value of t^* you would use for this interval is

(a) 1.645. (b) 1.699. (c) 1.697. (d) 1.96. (e) 2.045.

T8.6 A radio talk show host with a large audience is interested in the proportion p of adults in his listening area who think the drinking age should be lowered to eighteen. To find this out, he poses the following question to his listeners: "Do you think that the drinking age should be reduced to eighteen in light of the fact that eighteen-year-olds are eligible for military service?" He asks listeners to phone in and vote "Yes" if they agree the drinking age should be lowered and "No" if not. Of the 100 people who phoned in, 70 answered "Yes." Which of the following conditions for inference about a proportion using a confidence interval are violated?

I. The data are a random sample from the population of interest.
II. The population is at least 10 times as large as the sample.
III. n is so large that both $n\hat{p}$ and $n(1-\hat{p})$ are at least 10.

(a) I only (b) II only (c) III only (d) I and II only (e) I, II, and III

T8.7 A 90% confidence interval for the mean μ of a population is computed from a random sample and is found to be 9 ± 3. Which of the following *could* be the 95% confidence interval based on the same data?

(a) 9 ± 2 (b) 9 ± 3 (c) 9 ± 4 (d) 9 ± 8
(e) Without knowing the sample size, any of the above answers could be the 95% confidence interval.

T8.8 Suppose we want a 90% confidence interval for the average amount spent on books by freshmen in their first year at a major university. The interval is to have a margin of error of \$2. Based on last year's book sales, we estimate that the standard deviation of the amount spent will be close to \$30. The number of observations required is closest to

(a) 25. (b) 30. (c) 608. (d) 609. (e) 865.

T8.9 A telephone poll of an SRS of 1234 adults found that 62% are generally satisfied with their lives. The announced margin of error for the poll was 3%. Does the margin of error account for the fact that some adults do not have telephones?

(a) Yes. The margin of error includes all sources of error in the poll.
(b) Yes. Taking an SRS eliminates any possible bias in estimating the population proportion.
(c) Yes. The margin of error includes undercoverage but not nonresponse.
(d) No. The margin of error includes nonresponse but not undercoverage.
(e) No. The margin of error only includes sampling variability.

T8.10 A Census Bureau report on the income of Americans says that with 90% confidence the median income of all U.S. households in a recent year was \$57,005 with a margin of error of ±\$742. This means that

(a) 90% of all households had incomes in the range \$57,005 ± \$742.
(b) we can be sure that the median income for all households in the country lies in the interval \$57,005 ± \$742.
(c) 90% of the households in the sample interviewed by the Census Bureau had incomes in the interval \$57,005 ± \$742.
(d) the Census Bureau got the result \$57,005 ± \$742 using a method that will capture the true median income 90% of the time when used repeatedly.
(e) 90% of all possible samples of this same size would result in a sample median that falls within \$742 of \$57,005.

Section II: Free Response *Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.*

T8.11 The U.S. Forest Service is considering additional restrictions on the number of vehicles allowed to enter Yellowstone National Park. To assess public reaction, the service asks a random sample of 150 visitors if they favor the proposal. Of these, 89 say "Yes."

(a) Construct and interpret a 99% confidence interval for the proportion of all visitors to Yellowstone who favor the restrictions.
(b) Based on your work in part (a), can the U.S. Forest Service conclude that more than half of visitors to Yellowstone National Park favor the proposal? Justify your answer.

T8.12 How many people live in South African households? To find out, we collected data from an SRS of 48 out of the over 700,000 South African students who took part in the CensusAtSchool survey project. The mean number of people living in a household was 6.208; the standard deviation was 2.576.

(a) Is the Normal/Large Sample condition met in this case? Justify your answer.
(b) Maurice claims that a 95% confidence interval for the population mean is $6.208 \pm 1.96\frac{2.576}{\sqrt{47}}$. Explain why this interval is wrong. Then give the correct interval.

T8.13 A milk processor monitors the number of bacteria per milliliter in raw milk received at the factory. A random sample of 10 one-milliliter specimens of milk supplied by one producer gives the following data:

5370	4890	5100	4500	5260	5150	4900	4760	4700	4870

Construct and interpret a 90% confidence interval for the population mean μ.