# Chapter 9 – Estimating the Value of a Parameter

## OUTLINE

1. Estimating a Population Proportion
2. Estimating a Population Mean
3. Putting It Together: Which Procedure Do I Use?
4. Estimating with Bootstrapping

## Putting It Together

Chapters 1 through 7 laid the groundwork for the remainder of the course. These chapters dealt with data collection (Chapter 1), descriptive statistics (Chapters 2 through 4), and probability (Chapters 5 through 7).

Chapter 8 formed a bridge between probability and statistical inference by giving us models that can be used to make probability statements about the sample mean and sample proportion.

We now discuss inferential statistics—the process of generalizing information obtained from a sample to a population. We will study two areas of inferential statistics:

Estimation: Sample data are used to estimate the value of unknown parameters such as  or *p*.

Hypothesis testing: Statements regarding a characteristic of one or more populations are tested using sample data.

In this chapter, we discuss estimation of an unknown parameter, and in the next chapter we discuss hypothesis testing.

## Section 9.1 Estimating a Population Proportion

### Objectives

1. Obtain a Point Estimate for the Population Proportion
2. Construct and Interpret a Confidence Interval for the Population Proportion
3. Determine the Sample Size Necessary for Estimating a Population Proportion within a Specified Margin of Error

#### Objective 1: Obtain a Point Estimate for the Population Proportion

Objective 1, Page 1

1. State the definition of a point estimate.

Objective 1, Page 2

**Example 1 *Obtaining a Point Estimate of a Population Proportion***

The Gallup Organization conducted a poll in April 2017 in which a simple random sample of 1019 Americans aged 18 and older were asked, “Do you regard the income tax that you will have to pay this year as fair?” Of the 1019 adult Americans surveyed, 620 said yes. Obtain a point estimate for the proportion of Americans aged 18 and older who believe that the amount of income tax they pay is fair.

#### Objective 2: Construct and Interpret a Confidence Interval for the Population Proportion

Objective 2, Page 1

Statistics such as  vary from sample to sample.

Due to variability in the sample proportion, we report a range (or *interval*) of values, including a measure of the likelihood that the interval includes the unknown population parameter.

Objective 2, Page 2

 *Watch the video, which discusses the logic behind the construction of confidence intervals for a population proportion. Answer the following as you watch.*

1. Give the definition for a confidence interval for an unknown parameter.
2. What does the level of confidence represent?
3. What is the form of confidence interval estimates for a population parameter?

**Note: Review of the Sampling Distribution of the Sample Proportion**

* The shape of the distribution of all possible sample proportions is approximately normal provided  and the sample size is no more than 5% of the population size. That is, .
* The mean of the distribution of the sample proportions equals the population proportion. That is, .
* The standard deviation of the distribution of the sample proportion (the standard error) is 

1. How is the margin of error computed for a 95% confidence interval for a population proportion?

Objective 2, Page 2(continued)

* For a 95% confidence interval, any sample proportion that lies within 1.96 standard errors of the population proportion will result in a confidence interval that includes *p*. This will happen in 95% of all possible samples.
* Any sample proportion that is more than 1.96 standard errors from the population proportion will result in a confidence interval that does not contain *p*. This will happen in 5% of all possible samples (those sample proportions in the tails of the distribution).
* Whether a confidence interval contains the population parameter depends solely on the value of the sample statistic.
* Any sample statistic that is in the tails of the sampling distribution will result in a confidence interval that does not include the population parameter.

Objective 2, Page 3

**Key Ideas Regarding Confidence Intervals**

* A confidence interval for an unknown parameter consists of an interval of numbers based on a point estimate.
* The level of confidence represents the expected proportion of intervals that will contain the parameter if a large number of different samples is obtained. The level of confidence is denoted 
* Whether a confidence interval contains the population parameter depends solely on the value of the sample statistic. Any sample statistic that is in the tails of the sampling distribution will result in a confidence interval that does not include the population parameter.

Objective 2, Page 6

 *Answer the following after Activity 1: Illustrating the Meaning of Level of Confidence Using Simulation*

1. What proportion of the 95% confidence intervals generated by simulation would you expect to contain the population parameter *p*? What proportion of the 99% confidence intervals would you expect to contain the population parameter *p*?

Objective 2, Page 7

 *Watch the video to reinforce the concepts learned from Activity 1 (Objective 2, Page 6).*

Objective 2, Page 10

 *Answer the following after watching the video.*

1. What does the “95%” in a 95% confidence interval represent?

Objective 2, Page 12

 of all sample proportions will result in confidence intervals that contain the population proportion. The sample proportions that are in the tails of the distribution, outside the interval , will result in confidence intervals that contain the population proportion.

Objective 2, Page 13

1. What does the critical value of a distribution represent?
2. As the level of confidence increases, what happens to the critical value?
3. List the critical value associated with the given level of confidence.
4. 90%
5. 95%
6. 99%

Objective 2, Page 14

1. State the interpretation of a confidence interval.

Objective 2, Page 15

**Example 2 *Interpreting a Confidence Interval***

The Gallup Organization conducted a poll in April 2017 in which a simple random sample of 1019 Americans aged 18 and older were asked, “Do you regard the income tax that you will have to pay this year as fair?” We learned from Example 1 that the proportion of those surveyed who responded yes was 0.608. Gallup reported its “survey methodology” as follows:

Results are based on telephone interviews with a random sample of 1019 national adults, aged 18 and older. For results based on the total sample of national adults, one can say with 95% confidence that the maximum margin of sampling error is 4 percentage points.

Determine and interpret the confidence interval for the proportion of Americans aged 18 and older who believe the amount of federal income tax they have to pay is fair.

Objective 2, Page 17

**Note:**

A 90% level of confidence does not tell us that there is a 90% probability that the parameter lies between the lower and upper bound.

It means that the interval includes the unknown parameter for 90% of all samples.

Objective 2, Page 18

1. Explain the method for constructing a confidence interval about the population proportion, *p*.

**Note:** It must be the case that  and  to construct this interval. Use  in place of *p* in the standard deviation. This is because *p* is unknown, and  is the best point estimate of *p*.

Objective 2, Page 19

**Example 3 *Constructing a Confidence Interval for a Population Proportion***

In the Parent-Teen Cell Phone Survey conducted by Princeton Survey Research Associates International, 800 randomly sampled 16- to 17-year-olds living in the United States were asked whether they have ever used their cell phone to text while driving. Of the 800 teenagers surveyed, 272 indicated that they text while driving. Obtain a 95% confidence interval for the proportion of 16- to 17-year-olds who text while driving.

Objective 2, Page 21

1. State the formula for the margin of error, *E*, for a  confidence interval for a population proportion.

Objective 2, Page 22

**Example 4 *The Role of the Level of Confidence in the Margin of Error***

In the Parent-Teen Cell Phone Survey conducted by Princeton Survey Research Associates International, 800 randomly sampled 16- to 17-year-olds living in the United States were asked whether they have ever used their cell phone to text while driving. Of the 800 teenagers surveyed, 272 indicated that they text while driving. From the last example, we concluded that we are 95% confident that the proportion of 16- to 17-year olds who text while driving is between 0.307 and 0.373. Determine the effect on the margin of error by increasing the level of confidence from 95% to 99%.

Objective 2, Page 23

1. As the sample size, *n*, increases, what happens to the margin of error?

Objective 2, Page 24

 *Answer the following after watching the video.*

1. If the normality condition is not satisfied, how does the proportion of intervals that capture the parameter compare to the level of confidence?

#### Objective 3: Determine the Sample Size Necessary for Estimating a Population Proportion within a Specified Margin of Error

Objective 3, Page 1

 *Watch the video for an explanation of where the formulas for determining sample size for estimating a population proportion within a specified margin of error come from.*

Objective 3, Page 2

1. List the formula for the sample size required to obtain a  confidence interval for *p* with a margin of error *E*, if  is a prior estimate of *p*.
2. List the formula for the sample size required to obtain a  confidence interval for *p* with a margin of error *E*, if a prior estimate of *p* is unavailable.

Objective 3, Page 3

**Example 5 *Determining Sample Size***

An economist wants to know if the proportion of the U.S. population who commutes to work via car-pooling is on the rise. What size sample should be obtained if the economist wants an estimate within 2 percentage points of the true proportion with 90% confidence?

1. Assume that the economist uses the estimate of 10% obtained from the American Community Survey.
2. Assume that the economist does not use any prior estimates.

## Section 9.2 Estimating a Population Mean

### Objectives

1. Obtain a Point Estimate for the Population Mean
2. State Properties of Student’s *t*-Distribution
3. Determine *t*-Values
4. Construct and Interpret a Confidence Interval for a Population Mean
5. Determine the Sample Size Necessary for Estimating a Population Mean within a Given Margin of Error

#### Objective 1: Obtain a Point Estimate for the Population Mean

Objective 1, Page 1

1. What is the point estimate for a population mean ?

Objective 1, Page 2

**Example 1 *Computing a Point Estimate of the Population Mean***

The website [fueleconomy.gov](https://fueleconomy.gov) allows drivers to report the miles per gallon of their vehicle. The data in Table 2 show the reported miles per gallon of 2011 Ford Focus automobiles for 16 different owners. Obtain a point estimate of the population mean miles per gallon of a 2011 Ford Focus.

**Table 2**

|  |  |  |  |
| --- | --- | --- | --- |
| 35.7 | 37.2 | 34.1 | 38.9 |
| 32.0 | 41.3 | 32.5 | 37.1 |
| 37.3 | 38.8 | 38.2 | 39.6 |
| 32.2 | 40.9 | 37.0 | 36.0 |

#### Objective 2: State Properties of Student’s t-Distribution

Objective 2, Page 1

Recall that the distribution of  is approximately normal if the population from which the sample is drawn is normal or the sample size is sufficiently large. In addition, the distribution of  has the same mean as the parent population, , and a standard deviation equal to the parent population’s standard deviation divided by the square root of the sample size, .

Objective 2, Page 2

Using  for a confidence interval for the mean presents a problem because it is not likely that we know the population standard deviation  but not know the population mean .

Objective 2, Page 3

Using *s* as an estimate for  also presents a problem because the sample standard deviation, *s*, is a statistic and therefore will vary from sample to sample. Using the normal model to determine the critical value, , in the margin of error does not take into account the additional variability introduced by using *s* in place of . A new model must be used to determine the margin of error in a confidence interval that accounts for the additional variability. This leads to the story of William Gosset.

Objective 2, Page 4

1. What was the name of the brewery that Gosset worked for? What pseudonym did he choose to publish his results about a model that accounts for the additional variability introduced by using *s* in place of  when determining margin of error?

Objective 2, Page 5

 *Watch the video that uses simulation to illustrate some of the work Gosset did by hand to develop his sampling distribution—Student’s t-distribution..*

Objective 2, Page 6

1. Suppose that a simple random sample of size *n* is taken from a population. If the population from which the sample is drawn follows a normal distribution, what does the distribution of  follow?

Objective 2, Page 7

1. State six properties of the *t*-distribution.

#### Objective 3: Determine t-Values

Objective 3, Page 1

1. What does  represent?

Objective 3, Page 2

**Example 2 *Finding t-Values***

Find the *t*-value such that the area under the *t*-distribution to the right of the *t*-value is 0.10, assuming 15 degrees of freedom (df). That is, find  with 15 degrees of freedom.

Objective 3, Page 4

If the degrees of freedom we desire are not listed in Table VII, choose the closest number in the “df” column.

In addition, the last row of Table VII lists the *z*-values from the standard normal distribution. Use these values when the degrees of freedom are more than 1000 because the *t*-distribution starts to behave like the standard normal distribution as *n* increases.

#### Objective 4: Construct and Interpret a Confidence Interval for a Population Mean

Objective 4, Page 1

1. List the three conditions required for constructing a confidence interval for a population mean .
2. List the formulas for the lower bound and upper bound for a  confidence interval for the population mean, .

Objective 4, Page 2

1. What does it mean when we say that the procedure for constructing a confidence interval is robust?
2. Because the sample mean and sample standard deviation are not resistant to outliers, sample data should always be inspected for serious departures from normality and for outliers. What tools can be used to check for serious departures from normality and for outliers?

Objective 4, Page 3

**Example 3 *Constructing a Confidence Interval about a Population Mean***

The website fueleconomy.gov allows drivers to report the miles per gallon of their vehicle. The data in Table 3 show the reported miles per gallon of 2011 Ford Focus automobiles for 16 different owners. Treat the sample as a simple random sample of all 2011 Ford Focus automobiles. Construct a 95% confidence interval for the mean miles per gallon of a 2011 Ford Focus. Interpret the interval.

**Table 3**

|  |  |  |  |
| --- | --- | --- | --- |
| 35.7 | 37.2 | 34.1 | 38.9 |
| 32.0 | 41.3 | 32.5 | 37.1 |
| 37.3 | 38.8 | 38.2 | 39.6 |
| 32.2 | 40.9 | 37.0 | 36.0 |

Objective 4, Page 4

The *t*-distribution gives a larger critical value than the *z*-distribution, so the width of the confidence interval is wider when it is constructed using Student’s *t*-distribution.

Objective 4, Page 7

*Answer the following after Activity 1: When Model Requirements Fail.*

1. What happens to the proportion of intervals that capture the population mean as the sample size increases?

Objective 4, Page 9

If the requirements to compute a *t*-interval are not met, one option is to use resampling methods, such as bootstrapping. Bootstrapping is presented later in this chapter.

#### Objective 5: Determine the Sample Size Necessary for Estimating a Population Mean within a Given Margin of Error

Objective 5, Page 1

1. State the formula for margin of error when constructing a confidence interval about the population mean.
2. List the formula for the sample size required to obtain a  confidence interval for  with a margin of error *E*.

Objective 5, Page 2

**Example 4 *Determining Sample Size***

We again consider the problem of estimating the miles per gallon of a 2011 Ford Focus. How large a sample is required to estimate the mean miles per gallon within 0.5 miles per gallon with 95% confidence? Note: The sample standard deviation is s = 2.92 mpg.

## Section 9.3 Putting It Together: Which Procedure Do I Use?

### Objectives

1. Determine the Appropriate Confidence Interval to Construct

#### Objective 1: Determine the Appropriate Confidence Interval to Construct

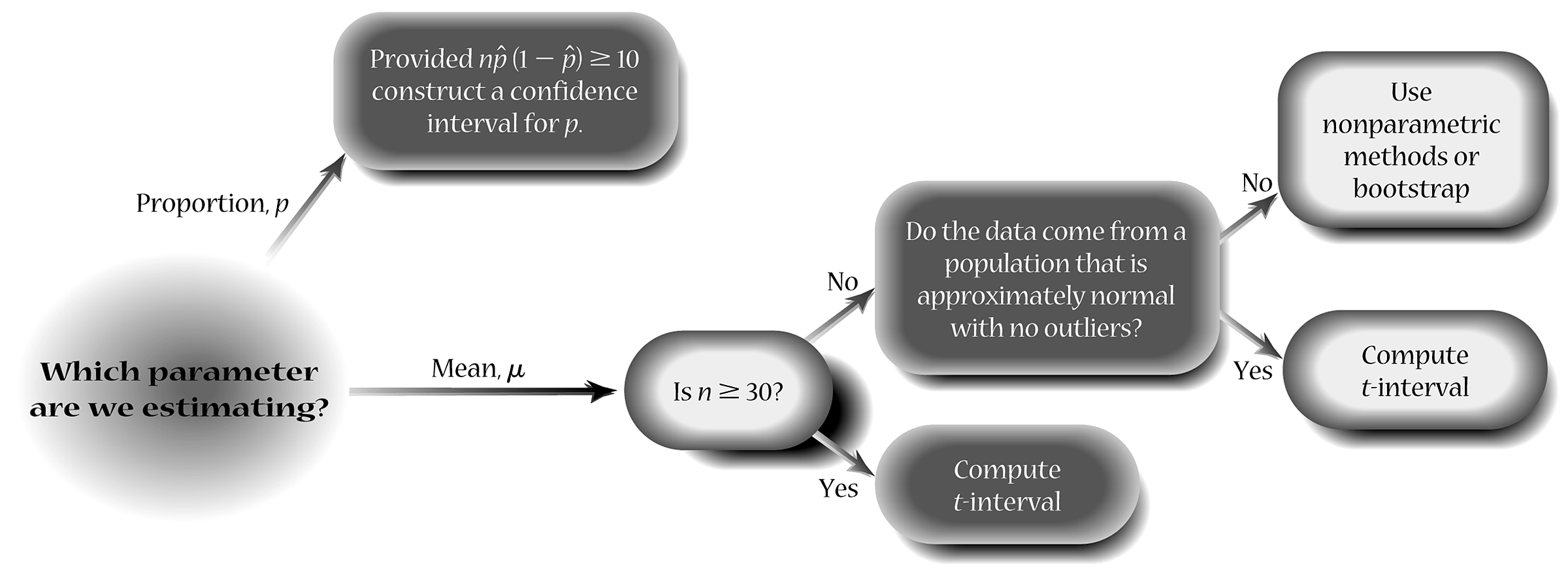
Objective 1, Page 1

 *Answer the following after watching the video.*

* 1. What type of data are needed to construct a confidence interval for a population proportion, *p*?
  2. Besides the fact that the sample must be obtained by simple random sampling or through a randomized experiment, list the two conditions that must be met when constructing a confidence interval for a population proportion, *p*.
  3. What type of data are needed to construct a confidence interval for a population mean, ?
  4. Besides the facts that the sample must be obtained by simple random sampling or through a randomized experiment and that the sample size must be small relative to the size of the population, what other condition must be satisfied?

Objective 1, Page 2

**Flowchart for Determining Which Type of Confidence to Construct**



## Section 9.4 Estimating with Bootstrapping

### Objectives

1. Estimate a Parameter Using the Bootstrap Method

#### Objective 1: Estimate a Parameter Using the Bootstrap Method

Objective 1, Page 1

1. State the definition for bootstrapping.

Objective 1, Page 2

 *Watch the video to learn about the logic behind the bootstrap method*.

Objective 1, Page 3

1. List the two basic requirements that must be satisfied to use bootstrapping.
2. What two percentiles are associated with a 95% confidence interval?

Objective 1, Page 4

1. List the three steps of the bootstrap algorithm.

Objective 1, Page 7

**Example 1 *Using the Bootstrap Method to* *Construct a 95% Confidence Interval***

The website fueleconomy.gov allows drivers to report the miles per gallon of their vehicle. The data in Table 4 show the reported miles per gallon of 2011 Ford Focus automobiles for 16 different owners. Treat the sample as a simple random sample of all 2011 Ford Focus automobiles. Construct a 95% confidence interval for the mean miles per gallon of a 2011 Ford Focus using a bootstrap sample. Interpret the interval.

**Table 4**

|  |  |  |  |
| --- | --- | --- | --- |
| 35.7 | 37.2 | 34.1 | 38.9 |
| 32.0 | 41.3 | 32.5 | 37.1 |
| 37.3 | 38.8 | 38.2 | 39.6 |
| 32.2 | 40.9 | 37.0 | 36.0 |

Objective 1, Page 10

The confidence interval constructed using the bootstrap method is slightly different than the confidence interval using Student’s *t*-distribution, but the two intervals are very similar.

Because bootstrapping relies on randomization, results will vary each time the method is used. So, the results you obtain by following Example 1 may differ from the results we obtained, but your results should be close.

Objective 1, Page 12

Many statistical spreadsheets (such as StatCrunch) have built-in algorithms that do bootstrapping.

Objective 1, Page 13

**Example 2 *Using StatCrunch’s Resample Command to Obtain a Bootstrap Confidence Interval***

Use StatCrunch to estimate a 95% confidence interval for the mean miles per gallon of a 2011 Ford Focus based on the sample data in Table 4.

**Table 4**

|  |  |  |  |
| --- | --- | --- | --- |
| 35.7 | 37.2 | 34.1 | 38.9 |
| 32.0 | 41.3 | 32.5 | 37.1 |
| 37.3 | 38.8 | 38.2 | 39.6 |
| 32.2 | 40.9 | 37.0 | 36.0 |

Objective 1, Page 15

To construct a bootstrap confidence interval for a proportion, we need raw data using 0 for a failure and 1 for a success. The sample proportion is then estimated using the mean of the 0s and 1s.

Objective 1, Page 16

1. List three items to be aware of when constructing confidence intervals using the bootstrap approach.