

## 8.S: Confidence Intervals (Summary)

### Chapter Review

In this module, we learned how to calculate the confidence interval for a single population mean where the population standard deviation is known. When estimating a population mean, the margin of error is called the error bound for a population mean ( $E$ ). A confidence interval has the general form:

$$(\text{lower bound}, \text{upper bound}) = (\text{point estimate} - E, \text{point estimate} + E) \quad (8.S.1)$$

The calculation of  $E$  depends on the size of the sample and the level of confidence desired. The confidence level is the percent of all possible samples that can be expected to include the true population parameter. As the confidence level increases, the corresponding  $E$  increases as well. As the sample size increases, the  $E$  decreases. By the central limit theorem,

$$E = z \frac{\sigma}{\sqrt{n}} \quad (8.S.2)$$

Given a confidence interval, you can work backwards to find the error bound ( $E$ ) or the sample mean. To find the error bound, find the difference of the upper bound of the interval and the mean. If you do not know the sample mean, you can find the error bound by calculating half the difference of the upper and lower bounds. To find the sample mean given a confidence interval, find the difference of the upper bound and the error bound. If the error bound is unknown, then average the upper and lower bounds of the confidence interval to find the sample mean.

Sometimes researchers know in advance that they want to estimate a population mean within a specific margin of error for a given level of confidence. In that case, solve the  $E$  formula for  $n$  to discover the size of the sample that is needed to achieve this goal:

$$n = \frac{z^2 \sigma^2}{E^2} \quad (8.S.3)$$

### Formula Review

$$\bar{X} \sim N\left(\mu_x, \frac{\sigma}{\sqrt{n}}\right) \quad (8.S.4)$$

The distribution of sample means is normally distributed with mean equal to the population mean and standard deviation given by the population standard deviation divided by the square root of the sample size.

The general form for a confidence interval for a single population mean, known standard deviation, normal distribution is given by

$$\begin{aligned} (\text{lower bound}, \text{upper bound}) &= (\text{point estimate} - E, \text{point estimate} + E) \\ &= \bar{x} - E, \bar{x} + E \\ &= \left( \bar{x} - z \frac{\sigma}{\sqrt{n}}, \bar{x} + z \frac{\sigma}{\sqrt{n}} \right) \end{aligned}$$

$E = z \frac{\sigma}{\sqrt{n}}$  = the error bound for the mean, or the margin of error for a single population mean; this formula is used when the population standard deviation is known.

$CL$  = confidence level, or the proportion of confidence intervals created that are expected to contain the true population parameter

$\alpha = 1 - CL$  = the proportion of confidence intervals that will not contain the population parameter

$z_{\frac{\alpha}{2}}$  = the  $z$ -score with the property that the area to the right of the  $z$ -score is  $\frac{\alpha}{2}$  this is the  $z$ -score used in the calculation of " $E$ " where  $\alpha = 1 - CL$ .

$n = \frac{z^2 \sigma^2}{E^2}$  the formula used to determine the sample size ( $n$ ) needed to achieve a desired margin of error at a given level of confidence

General form of a confidence interval

$$(\text{lower value}, \text{upper value}) = (\text{point estimate} - \text{error bound}, \text{point estimate} + \text{error bound}) \quad (8.S.5)$$

To find the error bound when you know the confidence interval

$$\text{error bound} = \text{upper value} - \text{point estimate} \quad (8.S.6)$$

OR

$$\text{error bound} = \frac{\text{upper value} - \text{lower value}}{2} \quad (8.S.7)$$

Single Population Mean, Known Standard Deviation, Normal Distribution

Use the Normal Distribution for Means, Population Standard Deviation is Known  $E = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

The confidence interval has the format  $(\bar{x} - E, \bar{x} + E)$ .

Use the following information to answer the next five exercises: The standard deviation of the weights of elephants is known to be approximately 15 pounds. We wish to construct a 95% confidence interval for the mean weight of newborn elephant calves. Fifty newborn elephants are weighed. The sample mean is 244 pounds. The sample standard deviation is 11 pounds.

### Exercise 8.2.8

Identify the following:

- a.  $\bar{x}$  = \_\_\_\_\_
- b.  $\sigma$  = \_\_\_\_\_
- c.  $n$  = \_\_\_\_\_

**Answer**

- a. 244
- b. 15
- c. 50

### Exercise 8.2.9

In words, define the random variables  $X$  and  $\bar{X}$ .

### Exercise 8.2.10

Which distribution should you use for this problem?

**Answer**

$$N\left(244, \frac{15}{\sqrt{50}}\right)$$

### Exercise 8.2.11

Construct a 95% confidence interval for the population mean weight of newborn elephants. State the confidence interval, sketch the graph, and calculate the error bound.

### Exercise 8.2.12

What will happen to the confidence interval obtained, if 500 newborn elephants are weighed instead of 50? Why?

**Answer**

As the sample size increases, there will be less variability in the mean, so the interval size decreases.

Use the following information to answer the next seven exercises: The U.S. Census Bureau conducts a study to determine the time needed to complete the short form. The Bureau surveys 200 people. The sample mean is 8.2 minutes. There is a known standard deviation of 2.2 minutes. The population distribution is assumed to be normal.

### Exercise 8.2.13

Identify the following:

- a.  $\bar{x}$  = \_\_\_\_\_
- b.  $\sigma$  = \_\_\_\_\_

c.  $n =$  \_\_\_\_\_

### Exercise 8.2.14

In words, define the random variables  $X$  and  $\bar{X}$ .

#### Answer

$X$  is the time in minutes it takes to complete the U.S. Census short form.  $\bar{X}$  is the mean time it took a sample of 200 people to complete the U.S. Census short form.

### Exercise 8.2.15

Which distribution should you use for this problem?

### Exercise 8.2.16

Construct a 90% confidence interval for the population mean time to complete the forms. State the confidence interval, sketch the graph, and calculate the error bound.

#### Answer

$CI : (7.9441, 8.4559)$

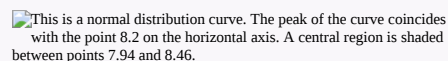
 This is a normal distribution curve. The peak of the curve coincides with the point 8.2 on the horizontal axis. A central region is shaded between points 7.94 and 8.46.

Figure 8.2.3.

$E = 0.26$

### Exercise 8.2.17

If the Census wants to increase its level of confidence and keep the error bound the same by taking another survey, what changes should it make?

### Exercise 8.2.18

If the Census did another survey, kept the error bound the same, and surveyed only 50 people instead of 200, what would happen to the level of confidence? Why?

#### Answer

The level of confidence would decrease because decreasing  $n$  makes the confidence interval wider, so at the same error bound, the confidence level decreases.

### Exercise 8.2.19

Suppose the Census needed to be 98% confident of the population mean length of time. Would the Census have to survey more people? Why or why not?

Use the following information to answer the next ten exercises: A sample of 20 heads of lettuce was selected. Assume that the population distribution of head weight is normal. The weight of each head of lettuce was then recorded. The mean weight was 2.2 pounds with a standard deviation of 0.1 pounds. The population standard deviation is known to be 0.2 pounds.

### Exercise 8.2.20

Identify the following:

- a.  $\bar{x} =$  \_\_\_\_\_
- b.  $\sigma =$  \_\_\_\_\_
- c.  $n =$  \_\_\_\_\_

#### Answer

- a.  $\bar{x} = 2.2$
- b.  $\sigma = 0.2$
- c.  $n = 20$

### Exercise 8.2.21

In words, define the random variable  $X$ .

### Exercise 8.2.22

In words, define the random variable  $\bar{X}$ .

#### Answer

$\bar{X}$  is the mean weight of a sample of 20 heads of lettuce.

### Exercise 8.2.23

Which distribution should you use for this problem?

### Exercise 8.2.24

Construct a 90% confidence interval for the population mean weight of the heads of lettuce. State the confidence interval, sketch the graph, and calculate the error bound.

#### Answer

$$E = 0.07$$

$$CI : (2.1264, 2.2736)$$


 This is a normal distribution curve. The peak of the curve coincides with the point 2.2 on the horizontal axis. A central region is shaded between points 2.13 and 2.27.

Figure 8.2.4.

### Exercise 8.2.25

Construct a 95% confidence interval for the population mean weight of the heads of lettuce. State the confidence interval, sketch the graph, and calculate the error bound.

### Exercise 8.2.26

In complete sentences, explain why the confidence interval in Exercise is larger than in Exercise.

#### Answer

The interval is greater because the level of confidence increased. If the only change made in the analysis is a change in confidence level, then all we are doing is changing how much area is being calculated for the normal distribution. Therefore, a larger confidence level results in larger areas and larger intervals.

### Exercise 8.2.27

In complete sentences, give an interpretation of what the interval in Exercise means.

### Exercise 8.2.28

What would happen if 40 heads of lettuce were sampled instead of 20, and the error bound remained the same?

#### Answer

The confidence level would increase.

### Exercise 8.2.29

What would happen if 40 heads of lettuce were sampled instead of 20, and the confidence level remained the same?

Use the following information to answer the next 14 exercises: The mean age for all Foothill College students for a recent Fall term was 33.2. The population standard deviation has been pretty consistent at 15. Suppose that twenty-five Winter students were randomly selected. The mean age for the sample was 30.4. We are interested in the true mean age for Winter Foothill College students. Let  $X$  = the age of a Winter Foothill College student.

### Exercise 8.2.30

$$\bar{x} = \underline{\hspace{2cm}}$$

**Answer**

30.4

Exercise 8.2.31

$n =$  \_\_\_\_\_

Exercise 8.2.32

\_\_\_\_\_ = 15

**Answer**

$\sigma$

Exercise 8.2.33

In words, define the random variable  $\bar{X}$ .

Exercise 8.2.34

What is  $\bar{x}$  estimating?

**Answer**

$\mu$

Exercise 8.2.35

Is  $\sigma_x$  known?

Exercise 8.2.36

As a result of your answer to Exercise, state the exact distribution to use when calculating the confidence interval.

**Answer**

normal

*Construct a 95% Confidence Interval for the true mean age of Winter Foothill College students by working out then answering the next seven exercises.*

Exercise 8.2.37

How much area is in both tails (combined)?  $\alpha =$  \_\_\_\_\_

Exercise 8.2.38

How much area is in each tail?  $\frac{\alpha}{2} =$  \_\_\_\_\_

**Answer**

0.025

Exercise 8.2.39

Identify the following specifications:

- a. lower limit
- b. upper limit
- c. error bound

Exercise 8.2.40

The 95% confidence interval is: \_\_\_\_\_.

**Answer**

(24.52,36.28)

### Exercise 8.2.41

Fill in the blanks on the graph with the areas, upper and lower limits of the confidence interval, and the sample mean.

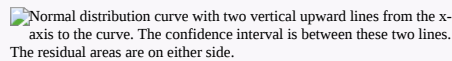
 Normal distribution curve with two vertical upward lines from the x-axis to the curve. The confidence interval is between these two lines. The residual areas are on either side.

Figure 8.2.5.

### Exercise 8.2.42

In one complete sentence, explain what the interval means.

#### Answer

We are 95% confident that the true mean age for Winger Foothill College students is between 24.52 and 36.28.

### Exercise 8.2.43

Using the same mean, standard deviation, and level of confidence, suppose that  $n$  were 69 instead of 25. Would the error bound become larger or smaller? How do you know?

### Exercise 8.2.44

Using the same mean, standard deviation, and sample size, how would the error bound change if the confidence level were reduced to 90%? Why?

#### Answer

The error bound for the mean would decrease because as the CL decreases, you need less area under the normal curve (which translates into a smaller interval) to capture the true population mean.

## Chapter Review

In many cases, the researcher does not know the population standard deviation,  $\sigma$ , of the measure being studied. In these cases, it is common to use the sample standard deviation,  $s$ , as an estimate of  $\sigma$ . The normal distribution creates accurate confidence intervals when  $\sigma$  is known, but it is not as accurate when  $s$  is used as an estimate. In this case, the Student's  $t$ -distribution is much better. Define a  $t$ -score using the following formula:

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \quad (8.S.8)$$

The  $t$ -score follows the Student's  $t$ -distribution with  $n - 1$  degrees of freedom. The confidence interval under this distribution is calculated with  $E = \left( t_{\frac{\alpha}{2}} \right) \frac{s}{\sqrt{n}}$  where  $t_{\frac{\alpha}{2}}$  is the  $t$ -score with area to the right equal to  $\frac{\alpha}{2}$ ,  $s$  is the sample standard deviation, and  $n$  is the sample size. Use a table, calculator, or computer to find  $t_{\frac{\alpha}{2}}$  for a given  $\alpha$ .

## Formula Review

$s$  = the standard deviation of sample values.

$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$  is the formula for the  $t$ -score which measures how far away a measure is from the population mean in the Student's  $t$ -distribution

$df = n - 1$ ; the degrees of freedom for a Student's  $t$ -distribution where  $n$  represents the size of the sample

$T \sim t_{df}$  the random variable,  $T$ , has a Student's  $t$ -distribution with  $df$  degrees of freedom

$E = t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$  = the error bound for the population mean when the population standard deviation is unknown

$t_{\frac{\alpha}{2}}$  is the  $t$ -score in the Student's  $t$ -distribution with area to the right equal to  $\frac{\alpha}{2}$

The general form for a confidence interval for a single mean, population standard deviation unknown, Student's  $t$  is given by (lower bound, upper bound)

$$= (\text{point estimate} - E, \text{point estimate} + E) \quad (8.S.9)$$

$$= \left( \bar{x} - \frac{ts}{\sqrt{n}}, \bar{x} + \frac{ts}{\sqrt{n}} \right) \quad (8.S.10)$$

Use the following information to answer the next five exercises. A hospital is trying to cut down on emergency room wait times. It is interested in the amount of time patients must wait before being called back to be examined. An investigation committee randomly surveyed 70 patients. The sample mean was 1.5 hours with a sample standard deviation of 0.5 hours.

### Exercise 8.3.3

Identify the following:

- $\bar{x} =$  \_\_\_\_\_
- $s_x =$  \_\_\_\_\_
- $n =$  \_\_\_\_\_
- $n - 1 =$  \_\_\_\_\_

### Exercise 8.3.4

Define the random variables  $X$  and  $\bar{X}$  in words.

**Answer**

$X$  is the number of hours a patient waits in the emergency room before being called back to be examined.  $\bar{X}$  is the mean wait time of 70 patients in the emergency room.

### Exercise 8.3.5

Which distribution should you use for this problem?

### Exercise 8.3.6

Construct a 95% confidence interval for the population mean time spent waiting. State the confidence interval, sketch the graph, and calculate the error bound.

**Answer**

$CI : (1.3808, 1.6192)$


 This is a normal distribution curve. The peak of the curve coincides with the point 1.5 on the horizontal axis. A central region is shaded between points 1.38 and 1.62.

Figure 8.3.1.

$E = 0.12$

### Exercise 8.3.7

Explain in complete sentences what the confidence interval means.

Use the following information to answer the next six exercises: One hundred eight Americans were surveyed to determine the number of hours they spend watching television each month. It was revealed that they watched an average of 151 hours each month with a standard deviation of 32 hours. Assume that the underlying population distribution is normal.

### Exercise 8.3.8

- $\bar{x} =$  \_\_\_\_\_
- $s_x =$  \_\_\_\_\_
- $n =$  \_\_\_\_\_
- $n - 1 =$  \_\_\_\_\_

**Answer**

- $\bar{x} = 151$
- $s_x = 32$
- $n = 108$
- $n - 1 = 107$

### Exercise 8.3.9

Define the random variable  $X$  in words.

### Exercise 8.3.10

Define the random variable  $\bar{X}$  in words.

**Answer**

$\bar{X}$  is the mean number of hours spent watching television per month from a sample of 108 Americans.

### Exercise 8.3.11

Which distribution should you use for this problem?

### Exercise 8.3.12

Construct a 99% confidence interval for the population mean hours spent watching television per month. (a) State the confidence interval, (b) sketch the graph, and (c) calculate the error bound.

**Answer**

$CI : (142.92, 159.08)$

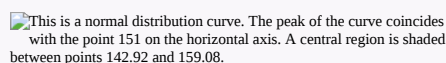
 This is a normal distribution curve. The peak of the curve coincides with the point 151 on the horizontal axis. A central region is shaded between points 142.92 and 159.08.

Figure 8.3.2.

$E = 8.08$

### Exercise 8.3.13

Why would the error bound change if the confidence level were lowered to 95%?

Use the following information to answer the next 13 exercises: The data in Table are the result of a random survey of 39 national flags (with replacement between picks) from various countries. We are interested in finding a confidence interval for the true mean number of colors on a national flag. Let  $X$  = the number of colors on a national flag.

$X$	Freq.
1	1
2	7
3	18
4	7
5	6

### Exercise 8.3.14

- $\bar{x} =$  \_\_\_\_\_
- $s_x =$  \_\_\_\_\_
- $n =$  \_\_\_\_\_

**Answer**

- 3.26
- 1.02
- 39

### Exercise 8.3.15

Define the random variable  $\bar{X}$  in words.

### Exercise 8.3.16

What is  $\bar{x}$  estimating?



**Answer**

$\mu$

### Exercise 8.3.17

Is  $\sigma_x$  known?

### Exercise 8.3.18

As a result of your answer to Exercise, state the exact distribution to use when calculating the confidence interval.

**Answer**

$t_{38}$

Construct a 95% confidence interval for the true mean number of colors on national flags.

### Exercise 8.3.19

How much area is in both tails (combined)?

### Exercise 8.3.20

How much area is in each tail?

**Answer**

0.025

### Exercise 8.3.21

Calculate the following:

- lower limit
- upper limit
- error bound

### Exercise 8.3.22

The 95% confidence interval is \_\_\_\_\_.

**Answer**

(2.93, 3.59)

### Exercise 8.3.23

Fill in the blanks on the graph with the areas, the upper and lower limits of the Confidence Interval and the sample mean.


 This is a template of a normal distribution curve with the central region shaded to represent a confidence interval. The residual areas are on either side of the shaded region. Blanks indicate that students should label the confidence level, residual areas, and points that define the confidence interval.

Figure 8.3.3.

### Exercise 8.3.24

In one complete sentence, explain what the interval means.

**Answer**

We are 95% confident that the true mean number of colors for national flags is between 2.93 colors and 3.59 colors.

### Exercise 8.3.25

Using the same  $\bar{x}$ ,  $s_x$ , and level of confidence, suppose that  $n$  were 69 instead of 39. Would the error bound become larger or smaller? How do you know?

**Answer**

The error bound would become  $E = 0.245$ . This error bound decreases because as sample sizes increase, variability decreases and we need less interval length to capture the true mean.

### Exercise 8.3.26

Using the same  $\bar{x}$ ,  $s_x$ , and  $n = 39$ , how would the error bound change if the confidence level were reduced to 90%? Why?

## References

1. Jensen, Tom. "Democrats, Republicans Divided on Opinion of Music Icons." Public Policy Polling. Available online at <http://www.publicpolicypolling.com/Day2MusicPoll.pdf> (accessed July 2, 2013).
2. Madden, Mary, Amanda Lenhart, Sandra Coresi, Urs Gasser, Maeve Duggan, Aaron Smith, and Meredith Beaton. "Teens, Social Media, and Privacy." PewInternet, 2013. Available online at <http://www.pewinternet.org/Reports/2...d-Privacy.aspx> (accessed July 2, 2013).
3. Prince Survey Research Associates International. "2013 Teen and Privacy Management Survey." Pew Research Center: Internet and American Life Project. Available online at <http://www.pewinternet.org/~media/...al%20Media.pdf> (accessed July 2, 2013).
4. Saad, Lydia. "Three in Four U.S. Workers Plan to Work Pas Retirement Age: Slightly more say they will do this by choice rather than necessity." Gallup® Economy, 2013. Available online at <http://www.gallup.com/poll/162758/th...ement-age.aspx> (accessed July 2, 2013).
5. The Field Poll. Available online at <http://field.com/fieldpollonline/subscribers/> (accessed July 2, 2013).
6. Zogby. "New SUNYIT/Zogby Analytics Poll: Few Americans Worry about Emergency Situations Occurring in Their Community; Only one in three have an Emergency Plan; 70% Support Infrastructure 'Investment' for National Security." Zogby Analytics, 2013. Available online at <http://www.zogbyanalytics.com/news/2...analytics-poll> (accessed July 2, 2013).
7. "52% Say Big-Time College Athletics Corrupt Education Process." Rasmussen Reports, 2013. Available online at [http://www.rasmussenreports.com/publ...cation\\_process](http://www.rasmussenreports.com/publ...cation_process) (accessed July 2, 2013).

## Chapter Review

Some statistical measures, like many survey questions, measure qualitative rather than quantitative data. In this case, the population parameter being estimated is a proportion. It is possible to create a confidence interval for the true population proportion following procedures similar to those used in creating confidence intervals for population means. The formulas are slightly different, but they follow the same reasoning.

Let  $p'$  represent the sample proportion,  $\frac{x}{n}$ , where  $x$  represents the number of successes and  $n$  represents the sample size. Let  $q' = 1 - p'$ . Then the confidence interval for a population proportion is given by the following formula:

$$(\text{lower bound, upper bound}) = (p' - E, p' + E) = \left( p' - z\sqrt{\frac{p'q'}{n}}, p' + z\sqrt{\frac{p'q'}{n}} \right)$$

The "plus four" method for calculating confidence intervals is an attempt to balance the error introduced by using estimates of the population proportion when calculating the standard deviation of the sampling distribution. Simply imagine four additional trials in the study; two are successes and two are failures. Calculate  $p' = \frac{x+2}{n+4}$ , and proceed to find the confidence interval. When sample sizes are small, this method has been demonstrated to provide more accurate confidence intervals than the standard formula used for larger samples.

## Formula Review

$p' = \frac{x}{n}$  where  $x$  represents the number of successes and  $n$  represents the sample size. The variable  $p'$  is the sample proportion and serves as the point estimate for the true population proportion.

$$q' = 1 - p' \quad (8.S.11)$$

$$p' - N \left( p, \sqrt{\frac{pq}{n}} \right) \quad (8.S.12)$$

The variable  $p'$  has a binomial distribution that can be approximated with the normal distribution shown here.

$E$  = the error bound for a proportion =  $z_{\frac{\alpha}{2}} \sqrt{\frac{p'q'}{n}}$

Confidence interval for a proportion:

(lower bound, upper bound) =  $(p' - E, p' + E) = \left( p' - z_{\frac{\alpha}{2}} \sqrt{\frac{p'q'}{n}}, p' + z_{\frac{\alpha}{2}} \sqrt{\frac{p'q'}{n}} \right)$

$n = \frac{z_{\frac{\alpha}{2}}^2 p'q'}{E^2}$  provides the number of participants needed to estimate the population proportion with confidence  $1 - \alpha$  and margin of error  $E$ .

Use the normal distribution for a single population proportion  $p' = \frac{x}{n}$

$E = \left( z_{\frac{\alpha}{2}} \right) \sqrt{\frac{p'q'}{n}}$

The confidence interval has the format  $(p' - E, p' + E)$ .

- $\bar{x}$  is a point estimate for  $\mu$
- $p'$  is a point estimate for  $\rho$
- $s$  is a point estimate for  $\sigma$

Use the following information to answer the next two exercises: Marketing companies are interested in knowing the population percent of women who make the majority of household purchasing decisions.

#### Exercise 8.4.6

When designing a study to determine this population proportion, what is the minimum number you would need to survey to be 90% confident that the population proportion is estimated to within 0.05?

#### Exercise 8.4.7

If it were later determined that it was important to be more than 90% confident and a new survey were commissioned, how would it affect the minimum number you need to survey? Why?

**Answer**

It would decrease, because the  $z$ -score would decrease, which reducing the numerator and lowering the number.

Use the following information to answer the next five exercises: Suppose the marketing company did do a survey. They randomly surveyed 200 households and found that in 120 of them, the woman made the majority of the purchasing decisions. We are interested in the population proportion of households where women make the majority of the purchasing decisions.

#### Exercise 8.4.8

Identify the following:

- $x$  = \_\_\_\_\_
- $n$  = \_\_\_\_\_
- $p'$  = \_\_\_\_\_

#### Exercise 8.4.9

Define the random variables  $X$  and  $P'$  in words.

**Answer**

$X$  is the number of “successes” where the woman makes the majority of the purchasing decisions for the household.  $P'$  is the percentage of households sampled where the woman makes the majority of the purchasing decisions for the household.

#### Exercise 8.4.10

Which distribution should you use for this problem?

#### Exercise 8.4.11

Construct a 95% confidence interval for the population proportion of households where the women make the majority of the purchasing decisions. State the confidence interval, sketch the graph, and calculate the error bound.

**Answer**

$CI : (0.5321, 0.6679)$

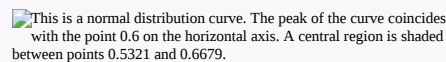
 This is a normal distribution curve. The peak of the curve coincides with the point 0.6 on the horizontal axis. A central region is shaded between points 0.5321 and 0.6679.

Figure 8.4.1.

$E : 0.0679$

**Exercise 8.4.12**

List two difficulties the company might have in obtaining random results, if this survey were done by email.

Use the following information to answer the next five exercises: Of 1,050 randomly selected adults, 360 identified themselves as manual laborers, 280 identified themselves as non-manual wage earners, 250 identified themselves as mid-level managers, and 160 identified themselves as executives. In the survey, 82% of manual laborers preferred trucks, 62% of non-manual wage earners preferred trucks, 54% of mid-level managers preferred trucks, and 26% of executives preferred trucks.

**Exercise 8.4.13**

We are interested in finding the 95% confidence interval for the percent of executives who prefer trucks. Define random variables  $X$  and  $P'$  in words.

**Answer**

$X$  is the number of “successes” where an executive prefers a truck.  $P'$  is the percentage of executives sampled who prefer a truck.

**Exercise 8.4.14**

Which distribution should you use for this problem?

**Exercise 8.4.15**

Construct a 95% confidence interval. State the confidence interval, sketch the graph, and calculate the error bound.

**Answer**

$CI : (0.19432, 0.33068)$

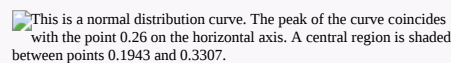
 This is a normal distribution curve. The peak of the curve coincides with the point 0.26 on the horizontal axis. A central region is shaded between points 0.1943 and 0.3307.

Figure 8.4.2.

$E : 0.0707$

**Exercise 8.4.16**

Suppose we want to lower the sampling error. What is one way to accomplish that?

**Exercise 8.4.17**

The sampling error given in the survey is  $\pm 2$ . Explain what the  $\pm 2$  means.

**Answer**

The sampling error means that the true mean can be 2% above or below the sample mean.

Use the following information to answer the next five exercises: A poll of 1,200 voters asked what the most significant issue was in the upcoming election. Sixty-five percent answered the economy. We are interested in the population proportion of voters who feel the economy is the most important.

**Exercise 8.4.18**

Define the random variable  $X$  in words.

**Exercise 8.4.19**

Define the random variable  $P'$  in words.

**Answer**

$P'$  is the proportion of voters sampled who said the economy is the most important issue in the upcoming election.

#### Exercise 8.4.20

Which distribution should you use for this problem?

#### Exercise 8.4.21

Construct a 90% confidence interval, and state the confidence interval and the error bound.

**Answer**

$CI : (0.62735, 0.67265)$

$E : 0.02265$

#### Exercise 8.4.22

What would happen to the confidence interval if the level of confidence were 95%?

Use the following information to answer the next 16 exercises: The Ice Chalet offers dozens of different beginning ice-skating classes. All of the class names are put into a bucket. The 5 P.M., Monday night, ages 8 to 12, beginning ice-skating class was picked. In that class were 64 girls and 16 boys. Suppose that we are interested in the true proportion of girls, ages 8 to 12, in all beginning ice-skating classes at the Ice Chalet. Assume that the children in the selected class are a random sample of the population.

#### Exercise 8.4.23

What is being counted?

**Answer**

The number of girls, ages 8 to 12, in the 5 P.M. Monday night beginning ice-skating class.

#### Exercise 8.4.24

In words, define the random variable  $X$ .

#### Exercise 8.4.25

Calculate the following:

- a.  $x =$  \_\_\_\_\_
- b.  $n =$  \_\_\_\_\_
- c.  $p' =$  \_\_\_\_\_

**Answer**

- a.  $x = 64$
- b.  $n = 80$
- c.  $p' = 0.8$

#### Exercise 8.4.26

State the estimated distribution of  $X$ .  $X \sim$  \_\_\_\_\_

#### Exercise 8.4.27

Define a new random variable  $P'$ . What is  $p'$  estimating?

**Answer**

$p$

#### Exercise 8.4.28

In words, define the random variable  $P'$ .

#### Exercise 8.4.29

State the estimated distribution of  $P'$ . Construct a 92% Confidence Interval for the true proportion of girls in the ages 8 to 12 beginning ice-skating classes at the Ice Chalet.

**Answer**

$$P \sim N\left(0.8, \sqrt{\frac{(0.8)(0.2)}{80}}\right). (0.72171, 0.87829)$$

#### Exercise 8.4.30

How much area is in both tails (combined)?

#### Exercise 8.4.31

How much area is in each tail?

**Answer**

0.04

#### Exercise 8.4.32

Calculate the following:

- lower limit
- upper limit
- error bound

#### Exercise 8.4.33

The 92% confidence interval is \_\_\_\_\_.

**Answer**

(0.72; 0.88)

#### Exercise 8.4.34

Fill in the blanks on the graph with the areas, upper and lower limits of the confidence interval, and the sample proportion.


 Normal distribution curve with two vertical upward lines from the x-axis to the curve. The confidence interval is between these two lines. The residual areas are on either side.

Figure 8.4.3.

#### Exercise 8.4.35

In one complete sentence, explain what the interval means.

**Answer**

With 92% confidence, we estimate the proportion of girls, ages 8 to 12, in a beginning ice-skating class at the Ice Chalet to be between 72% and 88%.

#### Exercise 8.4.36

Using the same  $p'$  and level of confidence, suppose that  $n$  were increased to 100. Would the error bound become larger or smaller? How do you know?

#### Exercise 8.4.37

Using the same  $p'$  and  $n = 80$ , how would the error bound change if the confidence level were increased to 98%? Why?

**Answer**

The error bound would increase. Assuming all other variables are kept constant, as the confidence level increases, the area under the curve corresponding to the confidence level becomes larger, which creates a wider interval and thus a larger error.

### Exercise 8.4.38

If you decreased the allowable error bound, why would the minimum sample size increase (keeping the same level of confidence)?

### Contributors

- Barbara Illowsky and Susan Dean (De Anza College) with many other contributing authors. Content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at <http://cnx.org/contents/30189442-699...b91b9de@18.114>.

This page titled [8.S: Confidence Intervals \(Summary\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.