



Math Objectives

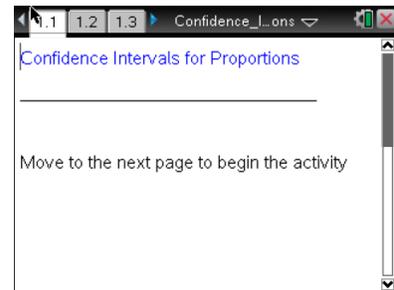
- Students will interpret a confidence interval for a population proportion as a set of populations that seem reasonable as candidates for populations that could produce a sample with the observed proportion.
- Students will recognize that a confidence interval is calculated from a given set of sample data.
- Students will recognize that a confidence interval from a sample statistic from an unknown population is based on knowing how statistics from samples drawn from known populations behave.
- Students will reason abstractly and concretely (CCSS Mathematical Practices).
- Students will use appropriate tools strategically (CCSS Mathematical Practices).

Vocabulary

- confidence interval
- population
- proportion
- sample
- sampling distribution

About the Lesson

- This lesson involves the concept of confidence intervals as a tool to make statements about a population proportion based on a given sample.
- As a result, students will:
 - Generate a random sample of men and women and count the proportion of the sample that is male, the observed proportion.
 - Examine a sampling distribution of proportions for samples drawn from several known populations, in particular those in which the proportion of men is 0.5 and is 0.75.
 - Set a population proportion, generate a sampling distribution for that given population proportion and decide whether their observed proportion of men seems likely to have come from this population. If so, a point is marked on a number line.
 - Repeat this for a sequence of population proportions until they have an interval of 'yeses'—a confidence interval—marked on the number line.



TI-Nspire™ Technology Skills:

- Download a TI-Nspire document
- Open a document
- Move between pages
- Use a minimized slider

Tech Tips:

- Make sure the font size on your TI-Nspire handhelds is set to Medium.

Lesson Files:

Student Activity

Confidence_Intervals_for_Proportions_Student.pdf

Confidence_Intervals_for_Proportions_Student.doc

TI-Nspire document

Confidence_Intervals_for_Proportions.tns

Visit www.mathnspired.com for lesson updates and tech tip videos.



TI-Nspire™ Navigator™ System

- Transfer a File.
- Use Screen Capture to compare different graphical displays.
- Use Quick Poll to check student understanding.

Prerequisite Knowledge

- familiarity with sampling distributions

Related Activities

- Sampling Distributions
- Confidence Levels for Proportions
- Confidence Intervals for Means

Discussion Points and Possible Answers

Tech Tip: Page 1.2 gives instructions for seeding the random number generator on the handheld. Page 1.3 is a *Calculator* page for the seeding process. Ensuring that students carry out this step will prevent students from generating identical data. (Syntax: RandSeed #, where # is a number unique to each student.)

Teacher Tip: The approach to confidence intervals in this activity is intended to provide the conceptual understanding of a confidence interval for proportions as a foundation for the formula usually used to calculate the interval.

This activity does not investigate the *level* of confidence or the probability that intervals produced by the method will include the true value of the parameter, which is done in the Statistics Nspired activity, *Confidence Levels for Proportions*. The investigation as presented is deliberately informal where the exact definition of "likely" is based on inspection. Teachers can choose to quantify this definition as indicated in the teacher notes in the activity. It is recommended, however, that the concept of the confidence interval be developed using this activity prior to mixing the concept of confidence intervals with the concept of confidence levels—the probability that a particular method will produce an interval that contains the true parameter.



Teacher Tip: You can choose to present this activity either as a teacher-led whole-class activity or with students working individually or in groups. If you choose the whole-class activity, students should all use the same number on Page 1.3 to seed the random number generator. This will give everyone the same observed proportion. If each student uses different seed numbers on Page 1.3, they will have different observed proportions and sharing results will not make sense for Question 6.

If 40% of the people polled from a random sample support a stance, does this indicate that exactly 40% of the population from which the sample was drawn will support the stance? Statisticians use a **confidence interval** to describe how to make an inference about a population from a sample.

Move to pages 1.2 and 1.3.

Follow your teacher's directions to seed the random number generator.

Move to page 1.4.

A coffee house was interested in whether their early morning customers were mostly men or women. A survey indicated that, on average, the total number of customers was 62 per day for that time period.

1. Why might the company care about whether their customers were male or female?

Sample Answers: They might want to target their advertising to attract more males or females depending on their typical customers; they might want to know who the typical customer was so they could ensure that the stores catered to the needs of those customers (i.e., certain music, food items).

2. Click the arrow on Page 1.4 to display the results of the survey. The men are marked with a blue dot and women with a pink x. What proportion of the customers in your sample was male?

Sample Answer: I counted 21 males out of 62 people, so the proportion of men was about 0.34. Students will have the same proportion if they used the same seed for their random number generator or different proportions if they used different seeds.





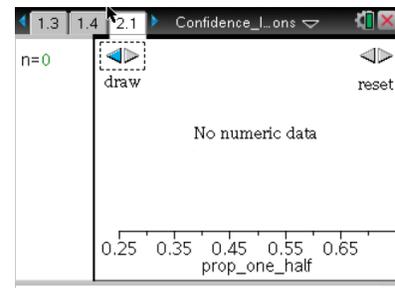
Tip: Count as accurately as possible, but do not worry about missing one or two – the key idea is to determine about what proportion of men is in the sample. Be sure to record this value; it will be important throughout the activity.

Teacher Tip: Do not worry about an exact counting – the key idea is that students are looking for some sense of what proportion of men is in the sample. Be sure students record this value; it will be important throughout the activity.

The value you found is called the *observed sample proportion* and is often represented by the symbol \hat{p} , called ‘*p-hat*.’ The question the owners want answered is what does this observed outcome—the proportion of males in the sample—indicate about the proportion of males in the population as a whole. The sample proportion will be important in the final steps of the activity, but for now, just note it as the answer to Question 2.

Move to page 2.1.

Samples vary, so before any statement can be made about the population, it is necessary to look carefully at what might happen when you know what the constraints are. The following questions explore how sampling variability can affect situations in which you know the population proportions (the “truth”).



3. Suppose the proportions for males and females who frequented the coffee house were the same, 0.5.
 - a. If you sample 62 customers, how many males would you expect in your sample? Do you think your expected number will be exactly correct? Explain your reasoning.

Sample Answer: I would expect about 31 men, but not exactly because things vary too much. Maybe 29 or 32 or so.

- b. Click the arrow on the left to display the proportion of men in a sample of 62 customers. What proportion of the customers were males?

Tip: Moving the cursor over a data point will display the exact proportion represented by that point.

Tech Tip: Moving the cursor over a data point will display the exact proportion represented by that point.

Sample Answers: Student results will vary, for example 0.66 or 0.45, if they used different seeds for the random number generator.



- c. Click the arrow five more times. Write a statement describing the results and what they mean in terms of the proportion of men likely to be found in samples drawn from a population in which exactly 50% are male.

Sample Answer: The distribution of proportions goes from 0.32 men to 0.66 men and is pretty evenly spread out with three more than 0.5 and three less than 0.5. So, when we know that the proportion of men is really one half, variability from sample to sample produces samples in which about 1/3 are men or 21 of the 62 to where 2/3 are men or 42 out of the 62.

Teacher Tip: This point—that the number of men varies among random samples where the expected number is exactly 31—is critical to understanding the "story" line for confidence intervals. Sample results vary even when parameters for a population are known. The focus is on capturing some sense of this variability.

4. Continue to generate samples until you have 100 samples of size 62 from this 50-50 population. (Note that after the tenth sample, the results are displayed in sets of ten samples.)
- a. Describe the distribution.

Sample Answer: The distribution is mound-shaped with the center around 0.5. The proportions range from 0.24 to 0.66, but most of them are between 0.42 and 0.58.

- b. Based on your experience with these 100 samples, if you thought a population consisted of 50% men and 50% women, would you be surprised if the proportion of men in a sample of 62 people was 0.55? 0.20? Explain your reasoning in each case.

Sample Answer: I wouldn't be surprised if the proportion of men was 0.55 because that is in the middle of the distribution of my 100 sample proportions, and 0.55 is close to the population proportion of 0.5. I would be surprised if the proportion was 0.2 because that never happened in my 100 samples from a population that is actually 0.5 male. So, although it could happen, it would be pretty surprising.



- c. Recall your observed proportion of men from Question 2. Would it seem likely that this sample proportion came from a population where the proportion of men was exactly 0.5? Why or why not?

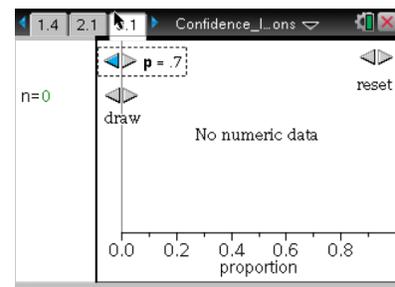
Sample Answers: Depending on the observed values, students might answer, “Yes, because my observed value was 0.44 and that is in the middle of the distribution of proportions that occurred when I sampled from a population known to have 0.5 males.” Others might say, “No because my observed value was 0.65, and that happened only once out of 100 samples by chance in the distribution of outcomes when the population proportion was really 0.5. For example, an observed proportion of 0.34 could have come from this population, but only twice did a number that small or smaller show up in the distribution.

Teacher Tip: Try not to focus on the meaning of "likely" at this point. Have students give general impressions about whether their observed outcome might be part of the sampling distribution given the variability in the distribution. Some students might say if a value occurs in their simulated sampling distribution, it is likely; others might leave off some of the extreme values in either tail. Either interpretation is fine at this stage. You can return to this at the end of the activity and quantify the meaning of "likely" or typical if it seems appropriate for your class.

Move to page 3.1.

The top arrow on the left of Page 3.1 allows you to set the proportion of men, p , for a given population. The bottom arrow on the left will generate samples of size 62 and display the proportion of men in each sample. After the first ten samples, the results are displayed in sets of ten until you have 100 samples.

The arrow on the right will erase the distribution and allow you to start the simulation over.



5. Use the top arrow on the left to set the population proportion, p , to 0.75. Use the bottom arrow on the left to generate a simulated sampling distribution for that proportion.
- a. Write a statement describing the results and what they mean in terms of the number of men in samples drawn from a population known to contain exactly 75% males.

Sample Answer: The simulated sampling distribution is mound-shaped and centered around 0.75 with proportions as small as 0.61 and as large as 0.89 – from 38 to 55 men even though the expected number of men was 46.5 or about 47.



- b. Recall your observed proportion of men from Question 2. Would it seem likely that this sample proportion came from a population where the proportion of men was exactly 0.75? Why or why not?

Sample Answer: My observed proportion of 0.34 or 21 men does not seem like it came from a sample from a population where 75% of the population was male. That result did not show up at all in the simulated sampling distribution.

- c. Use the reset arrow on the right to erase the distribution. Using the same population proportion, p , generate another simulated sampling distribution of the sample proportions of males. How will your answers to Questions 4a and b change for this new distribution?

Sample Answer: The new simulated sampling distribution of proportions is still mound-shaped and centered around 0.75, going from 0.63 to 0.9 or from about 39 to 56 men, which is really close to the first simulated sampling distribution. My observed outcome still does not seem likely to have come from this population.

Follow directions from your teacher to answer Question 6.

Teacher Tip: If you had all of the students use the same observed proportion when they seeded their random number generator for Question 2, you might want to assign each student a different population for which they generate a simulated sampling distribution for the proportion of men. Assign the populations in increments that range from 1% to 10% depending on class size, but be sure to include those from 10% to 90% by 10s. After students have answered Question 6, ask whether the observed sample proportion would have been a likely outcome for the distribution of sample proportions they generated for their given population. Collect the responses either orally or using TI-Nspire Navigator and note on the board which of the populations produced simulated sampling distributions that seemed likely to contain the observed proportion.

TI-Nspire Navigator Opportunity: Quick Poll

See Note 1 at the end of this lesson.

6. Click the top arrow to choose a population proportion (this might be assigned by your teacher). Generate 100 samples.



Teacher Tip: Be sure students recognize they are simulating sampling distributions of proportions of men in samples of size 62 taken from a population having a known proportion of males (defined by their choice with the arrow). They are selecting samples from a population for which they know the "truth"—the proportion of that population that is really male—and observing the simulated sampling distribution for that true proportion, in particular observing the variation around the "truth" that is possible from sample to sample.

- a. Describe the distribution of the proportion of men in your 100 samples. In particular, what proportions seem typically to occur?

Sample Answers: Descriptions will vary depending on the population proportion but should be mound-shaped with the center of the distribution close to the population proportion used. Typical values will vary: For a p of 0.65, most of the sample proportions were from 0.55 to 0.75.

- b. Recall the observed proportion of males you found in Question 2. Does it seem likely that observed proportion could have come from the population you used in this simulation? Explain your reasoning.

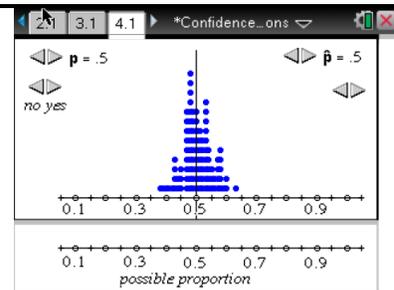
Sample Answers: Results will vary depending on the population proportions students used to generate their sampling distributions of the proportion of men. Some might note that the original sample proportion is in the center of their simulated distribution of sample proportions; others might find it never occurred in the distribution for their 100 samples.

TI-Nspire Navigator Opportunity: Screen Capture
See Note 2 at the end of this lesson.

Move to page 4.1.

Click the top arrow on the right to choose the observed sample proportion of men you generated on Page 1.4. The display will show a vertical line marking that observed proportion.

The top arrow on the left allows you to choose a given population proportion of men, and the display shows a simulated sampling distribution of the proportion of men based on 100 samples of size 62 drawn from that population, just as you saw in Question 6.





7. Choose a population proportion and use the yes/no arrow to indicate whether it seems likely the observed sample proportion of men, represented by the vertical line, could have come from that chosen population. Note that when you choose yes, the population proportion is plotted on the horizontal axis, **possible populations**, in the bottom panel.
- a. Write down the population proportions for which you chose “yes.” Continue to choose different population proportions until you have marked and recorded an interval that seems to contain all reasonable “Yeses.” Describe your interval.

Sample Answer: A set of populations that seems likely to have produced a sample with 34% men, the observed outcome, ranged from a population where the percent of men was 27% to a population with the percent of men was 40%.

- c. Use the reset arrow under \hat{p} to simulate a new sampling distribution of proportions from samples of size 62 with the same population proportion. Repeat the directions for Question 7a to find a set of populations that might have produced the sample with the observed proportion of men. How does the new interval compare to the one you found in Question 7a?

Sample Answer: The interval did not change much; this one went from 0.26 to 0.42.

The interval you found is called a *confidence interval* for a proportion. A confidence interval is an estimated range of values that is likely to contain an unknown population parameter, in this case proportion.

8. Suppose six months later the coffee house surveyed their customers on a randomly chosen day of the week. Again, the number of customers was 62. Return to Page 1.4, and generate a new sample.
- a. Record the proportion of men in the sample, return to Page 4.1, and find a confidence interval for the population proportion based on your new sample proportion.

Sample Answer: My new sample had 29 men out of 62, so the proportion of men was about 47%. This produced a confidence interval of 0.4 to 0.54.



- b. Write a few sentences to explain to someone who was not in class how to interpret your confidence interval.

Sample Answer: Any population where the true proportion of men is from 40% to 54% is a possible candidate for the unknown population from which I drew the sample. The simulations showed that rarely if ever did a sample from a population with more than 54% men have any samples having 47% men, which was what I found in my sample. Likewise, no simulated sampling distributions that I looked at in which the population proportion of men was less than 40% ever had a sample with 47% men. So it would seem reasonable that my sample of 47% men came from a population that had somewhere between 40% to 54% men.

Teacher Tip: If you would like to quantify "likely," you might have students return to Page 3.1 and examine the 100 elements in a sampling distribution for some given proportion. To have a 90% probability that the interval produced by this method includes the true proportion for the population, for each of the sampling distributions you examine, you would describe any of the five largest and 5 smallest outcomes as "unlikely"; reserving likely outcomes to be those that occurred in 90% of the samples. (When values coincide, use at least 90% of the samples, excluding 5 or fewer in each tail.)

Wrap Up

Upon completion of the lesson, the teacher should ensure that students are able to understand:

- A confidence interval for a population proportion is a set of populations that seem reasonable as candidates for populations that could produce a sample with the observed proportion.
- A confidence interval is calculated from a given set of sample data.
- A confidence interval for a sample statistic from an unknown population is based on knowing how statistics from samples drawn from known populations behave.



Assessment

Identify the following as true or false and explain why.

1. If you simulate the sampling distribution for an observed population proportion of 0.6, the confidence interval is the interval spanned by the distribution.

Answer: False: a confidence interval is not based on a simulation of observed outcomes but on examining simulations of samples from populations whose proportion of interest (the proportion of men) is known.

2. The observed proportions from two different samples from the same population will determine the same confidence interval.

Answer: False: Each observation might determine a different confidence interval; they will only be identical if the two observations are identical.

3. A confidence interval is a statement about populations that are likely to contain the true proportion based on information from a sample from an unknown population.

Answer: True: A confidence interval is a set of populations that seem reasonable as candidates for populations that could produce a sample with the observed proportion of "yeses."

4. If a confidence interval for the proportion of men in the population is from 0.4 to 0.7, in most cases this would mean the population proportion of men is halfway between, or 0.55.

Answer: False: Each proportion in the interval is just as likely to be the true population proportion as any other value.

TI-Nspire Navigator

Note 1

Question 6, Quick Poll

Collect the responses either orally or using TI-Nspire Navigator and note on the board which of the populations produced simulated sampling distributions that seemed likely to contain the observed proportion.

Quick Poll can be used to check for student understanding throughout the activity. The suggested assessment questions could be used to check for understanding at the end of the activity.

Note 2

Question 7, Screen Capture

A Screen Capture can be used in Question 6 to share sampling distributions of proportions of men from different populations and discuss which of them seem reasonable candidates as a population that produced the sample with the observed proportion of men.