Building Concepts: Sample Proportions

TEACHER NOTES

Lesson Overview

In this TI-Nspire lesson, students investigate the effect of sample size on variability by comparing the distribution of sample proportions with the population proportion.

A statistic computed from a random sample can be used as an estimate of that same characteristic of the population from which the sample was selected.

Learning Goals

1. Identify *sampling variability* as the variation from sample to sample in the values of a sample statistic;
2. understand that the shape, center, and spread of simulated sampling distributions of sample proportions for a given sample size will be fairly predictable;
3. understand that the sampling variability among samples is related to size of the samples; as the sample size increases, the variability decreases;
4. recognize that using a sample statistic from an unknown population to understand some characteristic of the population is based on knowing how statistics from samples drawn from known populations behave and that simulation can be a tool to approximate this behavior.

Prerequisite Knowledge

Sample Proportions is the nineteenth lesson in a series of lessons that explore the concepts of statistics and probability. This lesson builds on the concepts of the previous lessons. Prior to working on this lesson students should have completed *Probability and Simulation*, *Law of Large Numbers* and *Why Random Sampling*? Students should understand:

- that random sampling is likely to produce a sample that is representative of the population;
- how to use simulation to collect data.

Vocabulary

- **random sample:** representative of the population from which it was drawn
- **sampling variability:** the variation from sample to sample in the values of a sample statistic
- **sampling distribution of a statistic:** the collection of sample statistics from all possible samples of a given size from a specific population
- **simulated sampling distributions:** modeling a collection of sample statistics from a specific population

Lesson Pacing

This lesson should take 50–90 minutes to complete with students, though you may choose to extend, as needed.

Lesson Materials

- Compatible TI Technologies:

  TI-Nspire CX Handhelds, TI-Nspire Apps for iPad®, TI-Nspire Software
- Sample Proportions_Student.pdf
- Sample Proportions_Student.doc
- Sample Proportions.tns
- Sample Proportions_Teacher Notes
- To download the TI-Nspire activity (TNS file) and Student Activity sheet, go to http://education.ti.com/go/buildingconcepts.

Class Instruction Key

The following question types are included throughout the lesson to assist you in guiding students in their exploration of the concept:

Class Discussion: Use these questions to help students communicate their understanding of the lesson. Encourage students to refer to the TNS activity as they explain their reasoning. Have students listen to your instructions. Look for student answers to reflect an understanding of the concept. Listen for opportunities to address understanding or misconceptions in student answers.

Student Activity: Have students break into small groups and work together to find answers to the student activity questions. Observe students as they work and guide them in addressing the learning goals of each lesson. Have students record their answers on their student activity sheet. Once students have finished, have groups discuss and/or present their findings. The student activity sheet can also be completed as a larger group activity, depending on the technology available in the classroom.

Deeper Dive: These questions are provided for additional student practice and to facilitate a deeper understanding and exploration of the content. Encourage students to explain what they are doing and to share their reasoning.

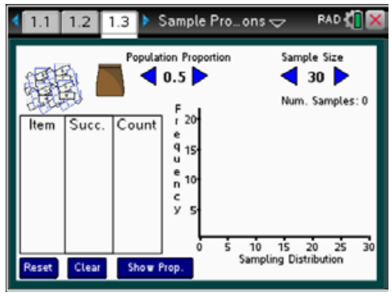
Mathematical Background

A central question in statistics is how to use information from a sample to begin to understand something about the population from which the sample was drawn. Collecting data from every member of the entire population can be time consuming and often impossible. The best way to collect data in such a situation is to use a random sample of the population: A statistic computed from a random sample, such as the sample proportion, can be used as an estimate of that same characteristic of the population from which the sample was selected. In prior lessons, students investigated variability within a single sample. In this lesson, students begin to differentiate between the variability within a single sample and the variability inherent in a statistic computed from each sample when samples of the same size are repeatedly selected from the same population. Understanding variability from this perspective enables students to think about how far a proportion of “successes” in a sample is likely to vary from the proportion of “successes” in the population. The variability in samples can be studied using simulations.

The collection of sample statistics from all possible samples of a given size from a population is called a *sampling distribution*. The complete sampling distribution of all possible values of a sample statistic for samples of a given size is typically difficult to generate but a subset of that distribution based on simulated sample statistics can be used to approximate the theoretical distribution. Students should note that sampling distributions of sample statistics computed from random samples of the same size from a given population tend to have certain predictable attributes. For example, while sample proportions vary from sample to sample, they cluster around the population proportion. In the case of sample means, the sample means cluster around the mean of the population. In both cases, the distributions of the sample statistic for a given sample size have a fairly predictable shape and spread.

Part 1, Page 1.3

Focus: Students develop an understanding of sampling variability by generating random samples from a population where the proportion of successes is known and observing the variability from sample to sample.



On page 1.3, students can select the bag to generate a random sample (size 30) from a population where the proportion of successes is 0.5 and display the result on the dot plot. Selecting the bag again will generate a new random sample. After 10 single samples, selecting the bag generates 10 samples at a time.

Proportion changes the population proportion of "successes."

Sample Size changes the size of the sample.

Show Prop displays in the table and graph the proportion of successes after each count.

Clear Data clears sample data but maintains population proportion.

Reset resets the page.

TI-Nspire Technology Tips

menu accesses page options.

tab cycles through proportion, sample size, show proportion/show count, and draw.

enter selects highlighted segments and displays length.

ctrl del resets the page to the original screen.

Class Discussion

Teacher Tip: In the following questions, students use a simulated distribution of successes from random samples of a given size (30) to make conjectures about the general behavior of such a distribution—the shape, center, and spread (as measured by the spread of successes) are relatively predictable after about 50 samples. Be sure students understand what the rows and columns represent in the table. Students may find it useful to use the scratchpad to do some of the calculations.

Have students…

Look for/Listen for…

Car colors vary from year to year and brand to brand, but the most popular color for a car is white. About 25% of all cars sold in the United States are white. Suppose you randomly sampled 30 cars in a grocery store parking lot and counted the number of white cars.

Class Discussion (continued)

• ***Describe a way to choose a random sample of 30 cars in the parking lot.***	Answers will vary. You could number the cars and randomly choose 30 of the numbers or you could number the rows and the cars in the rows, randomly choose five rows and then randomly choose six cars from each row. (Note that you might want students to use page 2.2 in Activity 18 to generate such a random sample, where the blocks would correspond to rows and the cells to the cars in a row.)
• ***About how many white cars would you expect to see in your sample?***	Answer: About 7 or 8.
• ***Would you be surprised to see 10 white cars in your sample? 20? Why or why not?***	Answers may vary. Students might think 10 white cars is possible, and others might think it is not; some students will begin to question whether the sample was random if you see that 20 of the 30 cars are white.
• ***On page 1.3, set the proportion to match the information about the number of white cars. Select the bag on page 1.3 to draw a random sample. Explain what the dot on the number line represents.***	Answer: Set the population proportion to 0.25. A dot at 11 indicates the random sample of 30 cars had 11 white cars.
If you were able to select a different sample of the same size, which of the following do you think is likely to be true about the number of white cars in the sample? Explain your reasoning. ***a. Eight cars in the sample will be white.*** ***b. The number of white cars will be the same as the number in the first sample.*** ***c. The number of white cars will be greater than 10.*** ***d. The number of white cars will be between 5 and 19.***	Answers will vary. Students should comment that: a. is likely to occur since 8 cars is very close to 25% of 30 cars, b. is unlikely since the number of white cars will vary from sample to sample, c. is somewhat unlikely since 10 is a little more than 30% of the sample of size 30, and that d. is almost certain since it would be very unusual to observe less than 5 or more than 19 white cars in a sample of 30 cars from a population with 25% white cars.

Class Discussion (continued)

Have students…	Look for/Listen for…
Select the bag a second time.	
• ***How did the number of white cars in this sample compare to your answer to the question above?***	Answers will vary. Some students might have answers that satisfy several of the choices in the previous question. (e.g., this sample could be the same as the first answer, also be greater than 10 and also be between 5 and 19.)
• ***Select the bag until you have taken ten samples. Describe the dot plot.***	Answers will vary. The plot may center around 8 or 9 white cars and a spread of 8 (one sample with 5 white cars to three samples with 13 white cars).
• ***Select Menu > Draw > Ten Times. What is the smallest number of white cars in any of the samples? The largest?***	Answers will vary. One example might be the smallest number of students in any sample was 2 and the largest 13.

Selecting the bag again will generate another 10 samples. Note that you are simulating the sampling distribution of the possible number of white cars in a random sample of size 30 when 25% of all cars are white.

• ***Continue to select the bag until you have generated 50 samples. By chance did you get at least one sample with 10 white cars? 20?***	Answers will vary. Some of the simulations should have produced several samples that had 10 white cars; it will be very rare to see a sample with 20 white cars.
• ***Describe the shape of the distribution and give the smallest and the largest number of white cars in any of the samples.***	Answers will vary. The distribution should mound shaped and centered around 7 or 8. In one example, the smallest number of white cars is 2, and the largest is 13.
• ***Generate another 50 samples and describe the dot plot.***	Answers will vary. The distribution should be appearing more mound shaped with center around 7 or 8. In one example, the least number of white cars is still 2, and the most is still 13.
• ***Reset. Set the proportion to represent the percent of white cars in the population. Generate 100 new samples. How does the new distribution of the number of white cars in a sample of 30 cars compare to the distribution from the 10 samples modeled above?***	Answers: will vary. In one example, the shape of the distribution is almost the same, but the spread is 11, since the number of white cars in the samples go from 1 to 12.

Class Discussion (continued)

Have students…	Look for/Listen for…
• ***Select*** **Show Prop *(proportion).***	Note: The count column has been replaced by a column with proportions in decimal form.
• ***Explain how the proportion of white cars was calculated.***	Answer: The total count value of white cars divided by 30 is the proportion of white cars in the sample.
About 20% of cars sold are black.	
• ***If you randomly sampled 30 cars from the same grocery store parking lot, about how many black cars would you expect to see? What proportion of the 30 cars is this?***	Answer: about 6 black cars; 6 black cars out of 30 cars is 0.2.
• ***Change the proportion to represent the number of black cars that are sold. (Be careful not to Reset.) Note that the sampling distribution for the number of white cars is greyed out. Draw one sample from the bag. Describe what the pink dot represents.***	Answer: The pink dot, for example 8, indicates that in one sample of 30 cars, 8 of the cars were black. Notice the overlap between the distributions is in a color different from the non-overlapping parts of the distributions.
• ***How does the pink dot fit into the sampling distribution of the number of white cars in samples of 30 cars?***	Answer: Using the example from the previous answer, 8 is in the lower part of the distribution.
• ***Make a conjecture about how the sampling distribution of the number of black cars in samples of 30 cars will relate to the sampling distribution of the number of white cars in samples of 30 cars. Generate many samples to check your conjecture.***	Answers will vary. The two distributions of the sample proportions will overlap quite a bit.
Using the proportion 0.45 (45%), generate a sampling distribution of 100 random samples drawn from car dealers in a region.	
• ***How does this distribution compare to the distribution in the sampling above where 20% of cars sold are black?***	Answers will vary. The two simulated sampling distributions overlap from about 0.2 to 0.5, which means that samples from both a population with 20% black cars and one with 45% black cars could have between 6 and 15 black cars
• ***By chance did you have a sample in which the proportion of black cars was greater than 0.6? 0.7?***	Answers will vary. Typically students will have a few samples in which the proportion of black cars was greater than or equal to 0.6. A sample where the proportion of black cars was more than 0.7 will not occur often.

Class Discussion (continued)

• ***What is the smallest number of black cars in any of the samples? The greatest?***	Answers will vary. For example, the smallest in one sampling distribution was 6; the largest was 19. The spread of the number of black cars in the samples was 13.
• ***What decimal proportions are associated with your answers to the previous question?***	Answers will vary. One interval using the above example could be from 0.2 to 0.63.
Generate sampling distributions for population proportions of 0.60, 0.50, 0.25, and 0.20. Compare the spreads for the number of cars in samples of size 30 for each of these population proportions..	Answer: For population proportions of 0.6 (60%) and 0.5 (50%) the spread is about 14 cars; for population proportions of 0.25 (25%), the spread is from about 1 to 14 or 13 cars; for population proportion of 0.20 (20%) the spread is from 0 to 12 or 12 cars. Overall the spreads for samples of size 30 from population proportions from about 20% to 60% (and maybe higher) are fairly constant, about 12 to 14 cars. Note: this spread will decrease for population proportions very close to 0% and to 100%.

Suppose you took 20 different random samples of 30 cars each to investigate a claim you read in the newspaper. Describe where you would sample, and what you would count in each sample. What do you think will be a typical spread for the number of cars in each situation below? Explain your reasoning, then use the TNS activity to check.

• ***30% of luxury cars are silver.***	Answers may vary. I would sample car dealers who specialized in luxury cars. The spread will probably be from about 2 to 16 silver cars (proportions 0.067 to 0.53). Reasons for each scenario might reference the answer to the question above.
• ***70% of the cars in a community have four doors.***	Answers may vary. I would sample all of the car dealers in the area, or I would sample the cars in a major parking lot. The span will probably be from about 14 to 28 four-door cars (proportions 0.47 to 0.93).
• ***40% of the cars sold to women are white or silver.***	Answers may vary. I would sample a roster of cars sold to women who bought cars. The span will probably be from about 5 to 19 (proportions 0.17 to 0.63).

Class Discussion (continued)

> **Teacher Tip:** In the following questions, students look at typical results from simulating outcomes for different population proportions of success, make conjectures about what seems to be commonly occurring and what outcomes might rarely occur. One caution: when comparing simulated sampling distributions, be sure the number of samples is the same unless the file has been set to proportions rather than counts.

When you simulate a distribution of sample proportions from a specific population, a range of typical or "plausible" sample proportions will arise. The word* plausible *indicates that a sample could have come from a population if the count/proportion falls within the expected spread of sample counts/proportions drawn from that population.

Suppose you observed 19 cars that were red in a random sample of 30 cars.

- ***Is it plausible that this sample came from a population with 10% red cars? Use the file to simulate the sampling distribution and justify your response.***

  Answer: No because 19 was not included in the simulated sampling distribution when 10% of the cars were red.

- ***Is a population with 50% red cars a plausible population for that sample? Explain your reasoning.***

  Answer: Yes, because the spread of sample proportions of cars that were red in a simulated sampling distribution from a population where 0.5 were red is from 7 to 20, which includes 19.

- ***Is a population with 60% cars that are red a plausible population for that sample? Explain your reasoning?***

  Answer: Yes, because the spread of the proportions of cars that were red in a simulated sampling distribution from a population where 0.6 were red is from 11 to 24, which includes 19.

- ***Find at least one other plausible population proportion of red cars that might have produced a random sample with 19 red cars.***

  Answers will vary. Any population proportion between 0.5 and 0.7 would work, (other proportions are acceptable as well as long as the count shows up in the simulated sampling distribution).

Student Activity Questions—Activity 1

1. **Do you agree with the following statements about the number of white cars in a random sample of 30 cars from a population where 25% of the cars white? Why or why not?**

   a. **It would be surprising if the sample had 5 white cars.**

   b. **It would be surprising if the sample had 18 white cars.**

   c. **The proportion of white cars has to be between 0.2 and 0.3.**

   d. **By chance you may have a random sample where the proportion of white cars is 0.45.**

   Answers will vary. B and d are likely to be true. a and c are not because the simulated sampling distributions seemed to go from about 3 or 4 white cars in a sample to about 11 or 12 white cars in a sample.

2. **In a certain community, a two-door car is not as popular as a four-door car. 60% of all cars sold in the area have four doors. Suppose you take random samples of 30 cars from 100 different car dealers.**

   a. **Estimate the minimum number and maximum number of four-door cars you would typically see in the sampling distribution of the number of four-door cars. Give both a count and the associated sample proportion.**

   Answers may vary. In one distribution of the number of four-door cars, the least number was 10 in one sample of 30 (proportion about 0.33) and the most was 26 (approximately a proportion of 0.87).

   b. **Describe what you would expect the sampling distribution of the number of four-door cars in random samples of 30 cars to look like.**

   Answers may vary. The distribution is symmetric and mound shaped centered at 18 (proportion of 0.6) with a spread of about 14 (a spread in proportions of 0.47).

   c. **Generate the sampling distribution and compare it to your description above.**

   Answer: The distribution goes from 12 to 25 (proportions of 0.4 to 0.833) for a spread of 13 (a spread in proportions of 0.433), and peaks around 18 or 0.6.

   d. **Does the distribution shape of the distribution change if you increase the number of samples to 200?**

   Answer: The distribution has greater frequencies for each outcome, but the shape remains basically the same.

3. **Do not Reset. In another community 50% of all cars sold are four-door. Change the proportion of successes in the population to 0.5.**

   a. **Generate ten samples. How do these relate to the sampling distribution when 0.6 of the cars had four doors?**

   Answer: The number of four-door cars in each sample for this community seems to be a bit to the left of the center of the simulated sampling distribution when 0.60 were four-door cars.

Student Activity Questions—Activity 1 (continued)

b. **Generate one more set of 10 samples of size 30, then repeat until you have a total of 50 samples, each with 30 cars. How does the shape compare to the distribution of samples from a population where 0.6 of the cars in the dealer lots had four doors?**

Answer: The center of the distribution seems to be shifting from around 18 to around 15. (Note this is a shift of 3 or 10% of 30.) In one example, the spread of the number of four-door cars was about the same, 12, from about 8 to 20 four-door cars; however, the interval from smallest to largest has shifted to the left.

c. **Generate samples until you have 100 samples of 30 cars each. Describe the distribution of the number of four-door cars.**

Answer: Using the example above, the center of the distribution, which is mound shaped and symmetric, is around 15, and the spread of the number of four-door cars in the samples is 12, from 8 to 20 cars.

d. **What is common between the distribution of the number of four-door cars when the 60% of the cars had four doors and when 50% of the population had four doors?**

Answer: Using the examples above, the distributions were both mound shaped and symmetric, and they both had a spread of about 12 or 13 four-door cars.

4. **Indicate whether you agree or disagree with each of the following statements about drawing random samples of size 30 from a population with a given proportion of successes. Explain your thinking in each case.**

a. **If you take 100 random samples of size 30 from the same population where the proportion of successes in the population is from about 30% to 70%, the spread for the number of counts in each sample will be about 14.**

Answer: 13 or 14 seems to be what has happened in the sampling distributions we simulated.

b. **A simulated sampling distribution of all possible samples of size 30 drawn from a population with 0.5 chance of a success will center around 15 successes.**

Answer: This seems to be true; in each case the simulated distribution of sample proportions seems to center around the product of the 30 and the proportion of successes.

c. **Just by chance, you might get sample counts that are five away from what you would expect.**

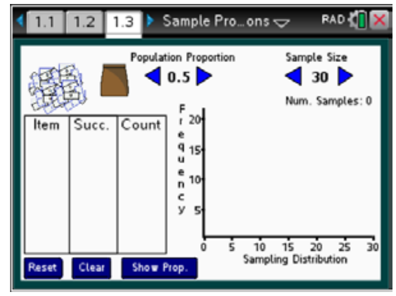
Answer: This is true. Random samples of the same size from the same population vary; a sample count 5 larger than what you might expect seemed to happen quite a bit.

d. **If you calculate a proportion of successes from a sample, you can use that number to predict exactly what the population proportion will be.**

Answer: No, because the spreads of possible numbers for different population proportions overlap and you cannot tell which population might have generated a sample with the observed proportion.

Part 2, Page 1.3

Focus: Students investigate the relationship between sample size and the shape of the simulated distribution of the proportion of successes in random samples.



Class Discussion

The following question confronts students with absolute and relative comparisons to understand why proportions are necessary when comparing quantities with different size bases.

Have students…	Look for/Listen for…
***Select* menu> Sample Size > Sixty*. Use 25% as the proportion of white cars made*.**	
• ***Generate a simulated sampling distribution of the number of white cars in 100 samples of size 60. Describe the distribution.***	Answers will vary. The distribution should be mound-shaped and symmetric with a center around 15 and may go from about 9 to 25, for a spread of 16.
• ***Compare the sampling distribution from above for samples of size 60 with the sampling distribution you found using samples of size 30.***	Answer: The center and spread are different. For the samples of size 30, the center is about 7 or 8 and the spread was about 11 (from 1 to 12). For samples of size 60, the center is around 15 and the spread of the counts was 16.
• ***Select* Show Prop *for the distribution of samples of size 60. Then redo the simulation for samples of size 30, when 25% of the cars made were white. Select* Show Prop *for this distribution and compare the two spreads using proportions. Explain the result.***	Answers may vary. The spread of the proportions of successes for samples of size 60 is from 0.12 to 0.4 or 0.28, while the spread of the proportions of successes for samples of size 30 is from 0.03 to 0.4 or 0.37. The spread for the larger sample size is smaller. The number of success in sampling distributions for samples of size 60 have a possible spread from 1 to 60 compared to a possible spread of 1 to 30 that is associated with sampling distributions for samples of size 30. Thus, to compare spreads, or even centers, you need to account for the difference in possible counts, which is what happens with proportions.

Class Discussion (continued)

The following question has students work in small groups to examine the difference in the sampling distributions of sample proportions for different sample sizes.

Work in a group with three other students. Select menu> Sample Size. One member of the group should choose sample size 10, another sample size 20, a third sample size 50, and a fourth sample size 100. Be sure they generate many samples to have a good representation of the sampling distribution at each sample size.

Students might compare their work to that of other groups.

- **Generate simulated sampling distributions of the proportion of each characteristic in the samples of your given size: silver cars (population proportion is 25%), four-door cars (population proportion is 60%) and two–door cars (population proportion is 40%). In each case, describe the simulated sampling distribution. (You might find a table useful for displaying the information about each distribution.)**

  Answers will vary.

Proportion/ Sample size	**0.25 (silver)**	**0.6 (four-door)**	**0.4 (two-door)**
Sample size 10	A bit skewed, peaks at 2 or 3 (0.25), spread is 7 (0.7) from 0 to 7 (0 to 0.7)	Mound-shaped and symmetric centered around 6 (0.6); spread is 7 (0.7), from 2 to 9 (0.2 to 0.9)	Nearly mound shaped and symmetric around 4 (0.4), spread is 7 (0.7), from 1 to 8 (0.1 to 0.8)
Sample size 20	Mound shaped and symmetric, centered around 5 (0.25); spread is 10 (0.5) from 0 to 10 (0 to 0.5)	Mound shaped and symmetric, centered around 12 (0.6); spread is 9 (0.45), from 8 to 17 (0.4 to 0.85)	Mound shaped and symmetric, centered around 8 (0.4); spread is 11 (0.55) from 2 to 13 (0.1 to 0.65)
Sample size 50	Mound shaped and symmetric, centered around 12 or 13 (0.25); spread is 16 (0.32), from 4 to 20 (0.08 to 0.4)	Mound shaped and symmetric, centered around 30 (0.6); spread is 20 (0.4), from 18 to 38 (0.36 to 0.76)	Mound shaped and symmetric, centered around 20 (0.4); spread is 18 (0.36), from 11 to 29 (0.22 to 0.58)
Sample size 100	Mound shaped and symmetric, centered around 25; spread is 20 (0.2), from 15 to 35 (0.15 to 0.35)	Mound shaped and symmetric, centered around 60; spread is 26 (0.26), from 45 to 71 (0.45 to 0.71)	Mound shaped and symmetric, centered around 40; spread is 20 (0.2), from 30 to 50 (0.3 to 0.5)

Class Discussion (continued)

- ***Compare your results, then use them to make a conjecture about the connection between your simulated sampling distributions and sample size.***

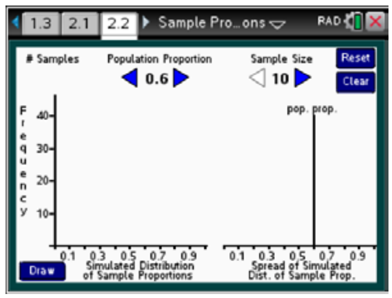
Answer: Students should see that the spreads of the sampling distributions of sample proportions decrease as the sample size increases. Note that the spread of counts of successful outcomes increases as the sample size increases, but the spread of the proportions of successful outcomes decreases as the sample size increases.

Part 3, Page 2.2

Focus: Students continue to investigate the relationship between sample size and the shape of simulated distribution of the proportion of successes in random samples.

On page 2.2, the arrow keys on the screen can be used to select a population proportion and a sample size. The spread of the proportions of success from each of the samples is represented by the horizontal bar on the right.

Draw generates 150 random samples of the given size from the population.



TI-Nspire Technology Tips

menu accesses page options.

tab cycles through population proportion and sample size.

enter draws new samples.

ctrl del resets the page to the original screen.

Student Activity Questions—Activity 2

1. **Go to page 2.2. In a certain community 60% of all pickups sold are black.**

   a. **Select Draw to generate a sampling distribution of 150 random samples of size 10, How does the horizontal bar on the right relate to the simulated sampling distribution?**

   Answer: The horizontal bar represents the spread of the proportions of black pickups in each of the 150 random samples.

   b. **Change the sample size to 20 and select *Draw*. Compare the two horizontal bars.**

   Answer: The spread of the results from samples of size 20 is smaller than the spread from samples of size 10.

Student Activity Questions—Activity 2 (continued)

c. Without resetting, select *Draw* for sample sizes 40, 60, 80, and 100. What do you notice about the spread of the proportions of black pickups as the sample size changes? Reset and generate the simulated sampling distributions again. Did your observation change?

Answer: Both times, overall, as the sample size increases, the spread in the proportions of black pickups in the samples decreases.

d. If it is possible to use any sample size, which one would you choose and why?

Answers will vary. Some may say 100 because the variability is smaller and that means you have a better chance to get close to the actual population value. Others might offer practical suggestions about costing more, taking more time, and so on as reasons for choosing a smaller sample size. Others might notice that the difference between the spread for samples of size 80 and for samples of size 100 was not that much, so you might as well use samples of size 80.

Deeper Dive — Page 1.3

- ***When the population proportion of a success is 20% to 65%, the simulated sampling distributions are typically mound shaped and symmetric. Investigate the shapes of the distributions for very small and large population proportions of success.***

Answers may vary. The distributions for 0.1 and 0.15, 0.85 and 0.9 are a bit skewed (right around 0.1 and left around 0.9) because the number of successes cannot go below 0 or above 1.

Deeper Dive — Page 2.2

Make a conjecture about the spread of a simulated sampling distribution of proportions of black pickup trucks in samples of size 70 for a population where 60% were black pickup trucks. Then use the file to check your conjecture.

Answers may vary. For example, the spread of the number of black trucks might go from 0.43 to 0.83.

What happens to the spread of simulated sampling distributions of proportions of successes in each case below? Use the TNS activity to check your thinking.

- ***double the sample size?***

Answer: the spread is not doubled but it does decrease.

- ***half the sample size?***

Answer: the spread is not cut in half but it does increase.

Deeper Dive

Project: ***Working in a group use page 2.2 in Activity 18 to design a strategy for taking a random sample of the number of cars parked on the street in nine Blocks that are a given color (or in a comparable situation: parking lots for grocery stores, malls, car dealers, etc.). You can choose any color you think is popular, and use a sample size that is 30 or less. Check the proportion of success you find to see if it is contained in any of the possible sampling distributions for your sample size that you can simulate using the file. Identify the population proportions of success that could have generated a sample with the same proportion as the one you found in your random sample. Use this thinking to make a conjecture about the actual proportion of that color car in the population of all the cars in those blocks. Record your work and findings on a poster to share with the class.***

Answers will vary. Students might find a random sample of size 27 by choosing randomly 3 cars from each of the nine blocks. Such a sample might have 8 red cars. This could have come from any population that has a proportion of successes from 0.15 to 0.55 (15% to 55%).

Sample Assessment Items

After completing the lesson, students should be able to answer the following types of questions. If students understand the concepts involved in the lesson, they should be able to answer the following questions without using the TNS activity.

1. If the proportion of those buying coffee at a coffee house is 0.65 female, how many females would you expect to see when there are 40 people at the coffee house?

   a. 13 b. 21 c. 26 d. 32

   Answer: c. 26

2. A random sample of 50 students found that 22 of them wanted to change the lunch menu. Which of the following would be surprising?

   a. Another random sample of 50 students found that 26 of them wanted to change the lunch menu.

   b. When the administration surveyed all 850 students in the school, they found that 721 of them wanted to change the lunch menu.

   c. When the administration surveyed all 850 students in the school, they found that 365 of them wanted to change the lunch menu.

   Answer: b. When the administration surveyed all 850 students in the school, they found that 721 of them wanted to change the lunch menu.

3 If the sample size increases, the variability from sample to sample

   a. increases b. decreases c. stays the same d. not enough information to tell.

   Answer: b. decreases

4. Would you rather go to a small hospital where a sample of 10 operations showed that 8 of them were successful or to a large hospital where on average 75% of the operations are successful? Give a reason for your answer.

   Answer: The large one because a small sample can have a lot of variability and it did not say it was random. When something happens "On average" there will be some variability but the variability will be close to the 75%.

Student Activity Solutions

In these activities you will generate random samples from a population where the proportion of successes is known then observe and describe the variability from sample to sample. After completing the activities, discuss and/or present your findings to the rest of the class.

Activity 1 [Page 1.3]

1. Do you agree with the following statements about the number of white cars in a random sample of 30 cars from a population where 25% of the cars white? Why or why not?

   a. It would be surprising if the sample had 5 white cars.

   b. It would be surprising if the sample had 18 white cars.

   c. The proportion of white cars has to be between 0.2 and 0.3.

   d. By chance you may have a random sample where the proportion of white cars is 0.45.

   Answers will vary. b) and d) are likely to be true. a) and c) are not because the simulated sampling distributions seemed to go from about 3 or 4 white cars in a sample to about 11 or 12 white cars in a sample.

2. In a certain community, a two-door car is not as popular as a four-door car. 60% of all cars sold in the area have four doors. Suppose you take random samples of 30 cars from 100 different car dealers.

   a. Estimate the minimum number and maximum number of four-door cars you would typically see in the sampling distribution of the number of four-door cars. Give both a count and the associated sample proportion.

      Answers may vary. In one distribution of the number of four-door cars, the least number was 10 in one sample of 30 (proportion about 0.33) and the most was 26 (approximately a proportion of 0.87).

   b. Describe what you would expect the sampling distribution of the number of four-door cars in random samples of 30 cars to look like.

      Answers may vary. The distribution is symmetric and mound shaped centered at 18 (proportion of 0.6) with a spread of about 16 (a proportion of 0.54).

   c. Generate the sampling distribution and compare it to your description above.

      Answer: The distribution goes from 12 to 25 (proportions of 0.4 to 0.833) for a spread of 13 (a spread in proportions of 0.433), and mounds around 18 or 0.6.

   d. Does the distribution shape of the distribution change if you increase the number of samples to 200?

      Answer: The distribution has greater frequencies for each outcome, but the shape remains basically the same.

3. Do not select **Reset**. In another community 50% of all cars sold are four-door. Change the proportion of successes in the population to 0.5.

   a. Generate ten samples. How do these relate to the sampling distribution when 0.6 of the cars had four doors?

      Answer: The number of four-door cars in each sample seems to be a bit to the left of the center for the simulated sampling distribution when 0.60 were four-door cars.

   b. Generate one more set of 10 samples of size 30, then repeat until you have a total of 50 samples, each with 30 cars. How does the shape compare to the distribution of samples from a population where 0.6 of the cars in the dealer lots had four doors?

      Answer: The center of the distribution seems to be shifting from around 18 to around 15. (Note this is a shift of 3 or 10% of 30!) In one example, the spread of the number of four-door cars was about the same, 12, from about 8 to 20 four-door cars; however, the interval from smallest to largest has shifted to the right.

   c. Generate samples until you have 100 samples of 30 cars each. Describe the distribution of the number of four-door cars.

      Answer: Using the example above, the center of the distribution, which is mound shaped and symmetric, is around 15, and the spread of the number of four-door cars in the samples is 12, from 8 to 20 cars.

   d. What is common between the distribution of the number of four-door cars when the 60% of the cars had four doors and when 50% of the population had four doors?

      Answer: Using the examples above, the distributions were both mound shaped and symmetric, and they both had a spread of about 12 or 13 four-door cars.

4. Indicate whether you agree or disagree with each of the following statements about drawing random samples of size 30 from a population with a given proportion of successes. Explain your thinking in each case.

   a. If you take 100 random samples of size 30 from the same population where the proportion of successes in the population is from about 30% to 70%, the spread for the number of counts in each sample will be about 14.

      Answer: 13 or 14 seems to be what has happened in the sampling distributions we simulated.

   b. A simulated sampling distribution of all possible samples of size 30 drawn from a population with 0.5 chance of a success will center around 15 successes.

      Answer: This seems to be true; in each case the simulated distribution of sample proportions seems to center around the product of the 30 and the proportion of successes.

c. Just by chance, you might get sample counts that are five away from what you would expect.

Answer: This is true. Random samples of the same size from the same population vary; a sample count 5 larger than what you might expect seemed to happen quite a bit.

d. If you calculate a proportion of successes from a sample, you can use that number to predict exactly what the population proportion will be.

Answer: No, because the spreads of possible numbers for different populations proportions overlap and you cannot tell which population might have generated a sample with the observed proportion.

Activity 2 [Page 2.2]

1. Go to page 2.2. In a certain community 60% of all pickups sold are black.

a. Select **Draw** to generate a sampling distribution of 150 random samples of size 10. How does the horizontal bar on the right relate to the simulated sampling distribution?

Answer: The horizontal bar represents the spread of the proportions of black pickups in each of the 150 random samples.

b. Change the sample size to 20 and select Draw. Compare the two horizontal bars.

Answer: Sample response, the spread for the results from a sample of size 20 is smaller than the spread for a sample of size 10.

c. Without resetting, select **Draw** for sample sizes 40, 60, 80 and 100. What do you notice about the spread of the proportions of black pickups as the sample size changes? Reset and generate the simulated sampling distributions again. Did your observation change?

Answer: Both times, overall, as the sample size increases, the spread in the proportions of black pickups in the samples decreases.

d. If it is possible to use any sample size, which one would you choose and why?

Answers will vary. Some may say 100 because the variability is smaller and that means you have a better chance to get close to the actual population value. Others might offer practical suggestions about costing more, taking more time, and so on as reasons for choosing a smaller sample size. Others might notice that the difference between the spread for samples of size 80 and for samples of size 100 was not that much, so you might as well use samples of size 80.

