Math Objectives

- Students will recognize that the means of different samples from a normal population will vary symmetrically around the mean of the population, with values near the mean occurring more frequently than those further from the mean, but in a narrower interval than that of individual elements of the population.
- Students will recognize that as sample size increases, variability (spread) in the sampling distribution of sample means decreases.
- Students will look for and make use of structure (CCSS Mathematical Practices).
- Students will model with mathematics (CCSS Mathematical Practices).

Vocabulary

- dot plot
- histogram
- mean
- normal distribution
- population
- sample
- sample mean
- sampling distribution
- standard deviation

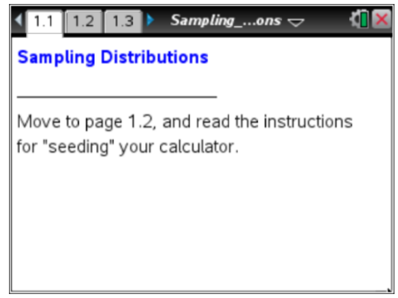
About the Lesson

Note that this lesson is not about establishing the central limit theorem but rather is focused on helping students understand what a sampling distribution is. This would be a good activity prior to introducing the central limit theorem.

- This lesson involves examining samples from a normal population and observing the distribution of the means of those samples.
- As a result, students will:
  - Understand that the sample mean varies from sample to sample, having its own distribution.
  - Estimate descriptive measures for the sampling distribution, and use those measures to approximate the simulated sampling distribution by selecting the mean and standard deviation for overlaying a normal curve.

TI-Nspire™ Navigator™ System

- Send the TI-Nspire document to students.
- Use Class Capture to display multiple distributions.
- Use Quick Poll to compare student sample means.



TI-Nspire™ Technology Skills:

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

Tech Tips:

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

Lesson Files:
Student Activity
Sampling_Distributions_Student.pdf
Sampling_Distributions_Student.doc
TI-Nspire document
Sampling_Distributions.tns

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

Prerequisite knowledge

- Normal distributions and the empirical rule
- Random samples
- Descriptive measures (mean and standard deviation)

Prerequisite Activities

- Normal Curve Family
- Z-Scores

Discussion Points and Possible Answers

Tech Tip: Page 1.2 gives instructions on how to seed the random number generator on the TI-Nspire. 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.) To complete the command, press enter.

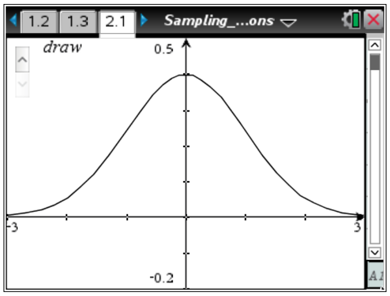
Teacher Tip: The main goal of Questions 1–5 is to have students realize that the sample mean is itself a variable that has a distribution whose shape, center, and spread can be described in its own right.

By thinking about the averaging process, students should conclude that the set of means should have a center near the population mean but should be less variable than the values that make up the population. It is easier to get a single value from the population that is, say, above $x = 2$ than it is to get 10 separate values whose mean is above $x = 2$.

The remainder of the lesson provides simulation experiences aimed at quantifying these properties, at least approximately.

Tech Tip: Page 2.1 initially shows only a normal density curve and minimized slider labeled "draw." Instruct students to use only the "up" arrow on the slider as they proceed through this lesson.

Move to page 2.1.



1. The graph on this page shows a normal distribution with mean 0 and standard deviation 1, representing a population of values. Recall the empirical rule and your knowledge of normal distributions. Predict the shape, center and spread of a sample randomly selected from this population.

   Sample Answers: Sample values will bounce around, with some above 0 and some below 0. There will be no real pattern, but values very near 3 or –3 will be rare.

Teacher Tip: The elements that make up the sample come from the population. Thus, the population's characteristics should be reflected in the distribution of those sample elements. The values of individual elements in samples should fall approximately symmetrically around 0, with a mean near 0. It will be rare to find values as large as 3 or as small as –3, but values as extreme as –2 or 2 should occur occasionally. Discuss these ideas after students have completed Question 2.

2. Each time you click the arrow (▲) labeled *draw*, you generate a random sample of size 10 from the given population. The elements of the sample will be displayed as points on the *x*-axis. Click to select the first sample. Look particularly at the center and spread of the values selected in the sample. Even though you have selected only 10 values from the population, do they seem to support your predictions in Question 1? Explain.

   Sample Answers: For most students, the answer will be yes, at least somewhat. For example: The values are not quite symmetrical around 0, but the prediction about not often getting close to 3 or –3 was right. The high value appears to be around 1.7, and the low is about –1.9, with four values below the mean and six above it. That all seems reasonable and matches the prediction pretty well.

3. In addition to the 10 values that make up the sample, a vertical line is displayed on the plot. What do you think it represents? Explain your reasoning.

   Sample Answers: It is labeled $\bar{x}$ and appears to be around the middle of the sample, so it represents the mean of the sample. If you estimated the mean of the distribution or the "balance point," it would be about where $\bar{x}$ falls.

Teacher Tip: Be sure students recognize that the plotted line is the mean and not the median. It might be worthwhile to have students verify that the line is not exactly midway between the fifth and sixth dots in the sample.

Teacher Tip: To help students begin to see the variability among values of $\bar{x}$, you might want to have the class share their graphs and/or values of $\bar{x}$ with each other after completing Question 3.

TI-Nspire Navigator Opportunity: *Class Capture*
See Note 1 at the end of the lesson.

4. Think about the variability among the values that make up your sample and the variability among the vertical $\bar{x}$ lines for several different samples.
   a. Use *draw* to select another sample (still of size 10). Record the value of $\bar{x}$, and write a short description of how the individual values of your sample are distributed (look back at Question 2).

      Sample Answers: The values are somewhat similar to those in Question 2 but not exactly the same. They still bounce around 0 and stay away from 3 or –3. The high and low values changed. $\bar{x}$ is 0.262.

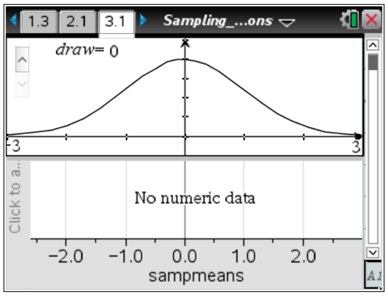
   b. Repeat part a four more times. Describe how the values of $\bar{x}$ vary compared to how the individual values in the samples vary.

      Sample Answers: The individual values remain similar to those in Question 2 but not exactly the same. They still bounce around 0 and, for the most part, still stay away from 3 or –3. The high and low values change for each sample. $\bar{x}$ seems to vary a lot less than the individual values in the sample. Some $\bar{x}$ values fall above 0, and some fall below 0, but none are anywhere near the high and low values for the individual values.

   c. Predict the center, spread, and shape of the distribution that would be formed by the $\bar{x}$ values from a large number of samples. Explain.

      Sample Answers: The $\bar{x}$ values should fall roughly symmetrically around 0. The mean of a sample of size 10 will fall somewhere in the "middle" of the 10 values that make up the sample itself, so $\bar{x}$ should not be very far from 0. Getting a really large mean would require that all 10 individual values were really large, and, based on the answers to parts *a* and *b*, that does not seem likely.

Move to page 3.1.



5. The top work area on page 3.1 shows the same population from Questions 1–4. An axis for a new dot plot has been added in the lower work area.
   a. Click *draw* five times, and describe what seems to be happening in the dot plot.

      Sample Answers: Each click produces a new value plotted in the dot plot directly below the value of $\bar{x}$ in the sample in the upper screen.

b. What variable do you think is being plotted in the dot plot?

Sample Answers: The variable is $\bar{x}$, the sample mean.

6. Click *draw* five more times to generate more values in the dot plot. Does the dot plot seem to confirm the predictions you made in Question 4 about the center, spread, and shape of the distributions of the $\bar{x}$ values from a large number of samples? Explain.

   Sample Answers: Yes. The values are accumulating on either side of 0, roughly symmetrically. It looks as though the center of this set of values might be 0, and the spread is noticeably smaller than that of any of the individual samples themselves.

7. You know several measures of spread. Think of one of those measures.

   a. Without doing any actual calculations, estimate the value of that measure of spread, first for the set of individual values in the samples themselves, and then for the set of values in the dot plot in the lower screen on Page 3.1.

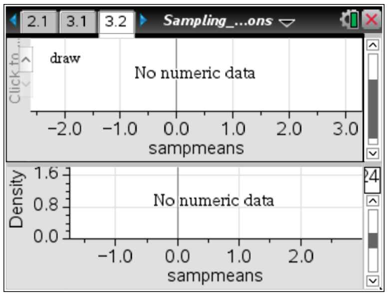
   Sample Answers: IQR (interquartile range) describes the spread of the middle half of the values being examined. An estimate of the IQR for the individual values in the samples is about 2. For the dot plot of $\bar{x}$, the IQR seems to be less than 1, maybe 0.5.

   b. How do the measures of spread for the individual elements and the sample means compare? Explain any differences you see.

   Sample Answers: The IQR for sample means is definitely less than the IQR for the individual elements in the samples. Since the sample means are always "in the middle of the sample," they do not bounce around as much as the individual values in the sample.

Teacher Tip: Be sure students understand that the mean of the sample, $\bar{x}$, represents a measure of the center, and high and low individual values have been "averaged out" toward that center. This decrease in variability is very visible in comparing the distribution of sample means with the population distribution. The extremes in the tails of the population visible in the top screen on Page 3.1 that would occur in a sample have been averaged with the other values in the sample, and consequently, the spread is smaller for the distribution of sample means displayed in the lower screen on Page 3.1.

Move to page 3.2.



8. The top work area of Page 3.2 is an exact copy of the lower work area on Page 3.1 with which you have been working. The lower work area displays the same data ($\overline{x}$ values from your samples) in a histogram.

Teacher Tip: When a histogram is graphed, the y-axis scale can be set to Frequency, Percent, or Density. When the scale on the vertical axis is density, the histogram has been scaled so that the total area of the histogram is 1 square unit.

Tech Tip: Depending upon the sample size and number of draws, the vertical axis on page 4.2 may need to be adjusted to display all of the histogram. Grab and drag a tic mark on the vertical axis to adjust, as needed.

a. Comment on which display seems better for seeing the overall shape of the distribution. Explain your reasoning.

Sample Answers: Answers will vary. The dot plot may appear better if the histogram's bars are separated and/or short.

Move back to page 3.1.

b. Each click of the *draw* arrow will generate ten more samples, and the means of those samples will be added to the plot in the lower work area. Click *draw* until you have about 100 samples, and then look at the graphs on page 3.2**.** Which graph type seems most appropriate for this larger simulation? Explain.

Sample Answers: Answers may vary, but the histogram is likely the better display of the overall shape of the distribution. The histogram appears less cluttered than the 100-point dot plot.

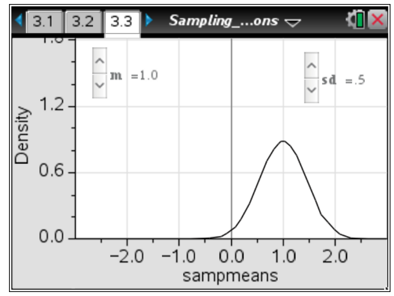
c. Describe the shape of the distribution of sample means. Estimate the mean and standard deviation of this distribution.

Sample Answers: It appears symmetric and mound-shaped, approximately normally distributed. The mean is approximately 0. The standard deviation might be 0.5.

Teacher Tip: While it is not a major point of the lesson, students should become aware that histograms tend to be better representations of data when there are more data present, since histograms display values in a “grouped” manner. Dot plots involve no grouping, so they do better where there are few data.

TI-Nspire Navigator Opportunity: ***Class Capture and Quick Poll***
See Notes 2 and 3 at the end of the lesson.

Move to page 3.3.

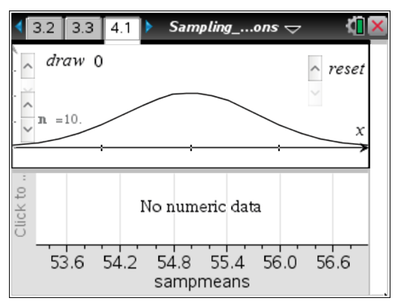


The plot on Page 3.3 is an exact copy of the histogram you examined on Page 3.2 but has an “adjustable” normal curve in the window. You control the appearance of that curve by clicking on the arrows (▲ or ▼) to select a mean and standard deviation.

9. Use the arrows (▲ or ▼) to set the mean and standard deviation to match the estimates you made in Question 8c. Then, if necessary, re-adjust the values so that they fit your histogram as well as possible. Record your final values for mean and standard deviation. Comment on the accuracy of your predictions.

   <u>Sample Answers:</u> The mean was close, but the standard deviation is even smaller than my guess. The actual standard deviation value is more like 0.3.

Move to page 4.1.



Teacher Tip: The sample size should not be changed once the student starts to draw samples in a particular exploration.

All of the students in grades 9-12 at a school measured the circumferences of their heads. The *draw* arrow will select different samples of size ten from the school population, and the mean circumference of the students' heads will be displayed in the lower plot.

10. a Use pages 4.1 and 4.2 to repeat the “sample-size-10” explorations you carried out in Questions 8 and 9 with this new population. Describe the sampling distribution of mean head circumferences for samples of size 10 taken from all grades 9-12 students in the school.

Sample Answers: The means of the head circumferences for 10 students appear to be around 55 and are more concentrated than the distribution of individual head sizes of the high school students (smaller standard deviation, around 0.2 or 0.3). The shape appears roughly symmetric and mound-shaped.

Note: Click the reset arrow on page 4.1 to erase an exploration using one sample size in order to begin a new exploration with another sample size.

b. The lower and upper arrows will generate samples of sizes other than 10 by changing the value of *n*. Repeat these explorations once more, this time using a different sample size. Comment on how changing the sample size affects the center, spread, and shape of the distribution of mean head circumference. Be as specific as possible, indicating what happens when the sample size increases and what happens when the sample size decreases.

Sample Answers: The overall results are the same, with the exception of spread. Larger samples lead to smaller spreads in the sampling distribution of mean head circumference. Smaller samples let the mean head size vary more.

11. Write a brief description to explain what you learned about a sampling distribution of sample means for someone who did not do this activity.

Sample Answers: Student responses should mention that when a random sample is drawn from a population, its sample mean can be calculated. Repeated samplings give a set of sample means that can be plotted to make a simulated sampling distribution. This distribution will have about the same mean as the original population and will be symmetric around that mean, but its spread will be much smaller than that of the original population with smaller spreads occurring for larger sample sizes.

Wrap Up

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

- Elements sampled from a normal population vary according to that normal distribution.
- Means of different samples of a fixed size from a given population vary but differently from the individual elements from the population.
- The simulated sampling distribution of sample means looks approximately normally distributed.
- As sample size increases, variability in the sampling distribution of sample means decreases.

TI-Nspire Navigator

Note 1

Question 3, Class Capture and Quick Poll

A Class Capture would allow students to see other results and note the variability. Another good alternative would be to send a Quick Poll to collect each student's displayed $\bar{x}$ value.

Note 2

Question 8, Class Capture

Using Class Capture can reinforce the idea that sample means vary from sample to sample, but that they do so in a somewhat predictable manner. Students should notice that while their individual graphs will be different in detail, most will be centered somewhere near 0, and the spreads of their plots will be about the same.

Note 3

Before Question 9, Quick Poll

Send students a screen shot of page 3.3 and ask them to predict what the mean and standard deviation should be. Show the results, without comment. Have them take 5 minutes to convince a partner that their prediction is correct. Then resend the Quick Poll. Show the results and have students discuss what they were thinking when they made the prediction. Then go to page 3.3 and have them test their predictions.