Math Objectives

- Students will recognize that elements of a sample 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.
- 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.

Vocabulary

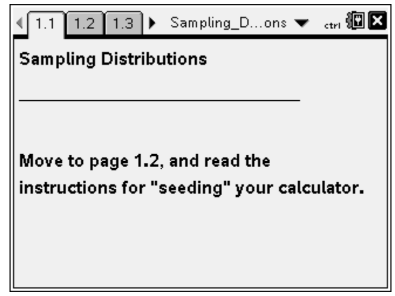
- dot plot
- histogram
- mean
- normal distribution
- population
- sample
- sample mean
- sampling distribution
- standard deviation

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.
- Students will 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.

The overall message is that the sample mean is itself a varying quantity, changing from sample to sample. Thus, it has a distribution of its own. That sampling distribution of sample means is



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 Materials:
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.

approximately normal, is centered at the population mean, and has a smaller standard deviation than that of the population, with that decrease becoming more pronounced as the sample size increases.

TI-Nspire Navigator™

- Send the .tns file to students.
- Use Screen Capture to display multiple distributions.
- Use Quick Poll to compare student sample means.

Prerequisite knowledge

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

Prerequisite Stat Nspired 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.)

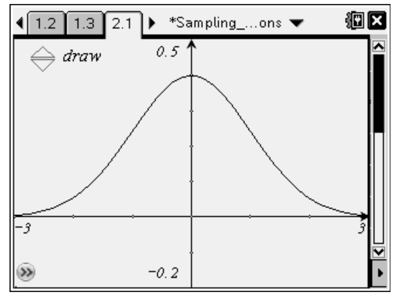
Teacher Tip: The main goal of Questions 1–5 is to have students realize that the sample mean is itself a variable that has a distribution whose shape, center, and spread can be described in its own right.

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. That is, 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 average 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 what you expect as the shape, center, and spread of the distribution of values randomly selected from this population.

   Sample Answer: 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 up arrow (▲) on the slider labeled “draw,” you will generate a random sample of size 10 from the given population. The members of your sample will be displayed as dots on the *x*-axis. Click to select your first sample. Even though you have selected only 10 values from the population, do they seem to support your predictions in Question 1? Look particularly at the center and spread of the values selected in your sample. Explain.

   Sample Answer: 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 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 dots that make up your sample, a vertical line is displayed on your plot. What do you think it represents? Explain your reasoning.

   Sample Answer: It is labeled $\bar{x}$ and appears to be around the middle of the sample, so maybe 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: The focus of this activity is on having students understand that a sample statistic (the sample mean in this activity) varies and has a distribution of its own. To help students begin to see the variability among values of $\bar{x}$, you may 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: *Screen Capture*
A TI-Navigator class Screen Capture would be great for this. Another good alternative would be to collect a Quick Poll of each student's displayed $\bar{x}$ value. If these are not available, students could place their calculators, with page 2.1 displayed, on their desks and then take a stroll around the room to see all the graphs.

4. Think about the variability among the dots 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 dots of your sample are distributed (look back at Question 2).

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

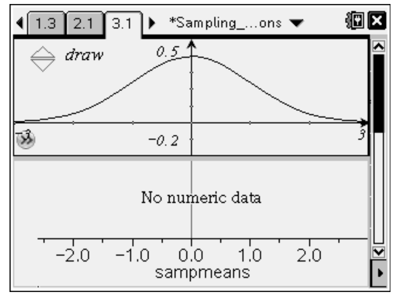
   b. Repeat part a four more times. Then describe how the $\bar{x}$ lines vary compared to how the individual dots of the samples vary.

      Sample Answer: The dots remain similar to those in Question 2, but not exactly the same. They still bounce around 0 and still stay away from 3 or –3. The high and low values change for each sample. $\bar{x}$ seems to vary a lot less. 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 dots.

   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 Answer: The $\bar{x}$ values should fall roughly symmetrically around 0. The mean of a sample of size 10 will always fall somewhere in the "middle" of the 10 dots that make up the sample itself, so $\bar{x}$ should not be very far from 0. Getting a really large average would require having all 10 dots be really big, and that never happened in the five tries of parts a and b.

Move to page 3.1.



5. The top screen on page 3.1 shows the same population that you dealt with in Questions 1–4. A new dot plot has been added in the lower screen.
   a. Click *draw* a few times and describe what seems to be happening in the lower dot plot.

   Sample Answer: Each click produces a new dot directly below the value of $\bar{x}$ in the sample in the upper frame.

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

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

   > **Teacher Tip:** You may wish to discuss with students the difference between the variable, $\bar{x}$, and its values. Only one variable is displayed in the plot, but that variable takes on additional values as more samples are selected.

6. Click *draw* 10 or 15 times to generate more dots 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 Answer: Yes. The dots are accumulating on either side of 0, roughly symmetrically. It looks as though the center of this set of dots 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 values of that measure of spread, first for the set of individual values in your samples themselves, and then for the set of dots in the dot plot of the lower screen on page 3.1.

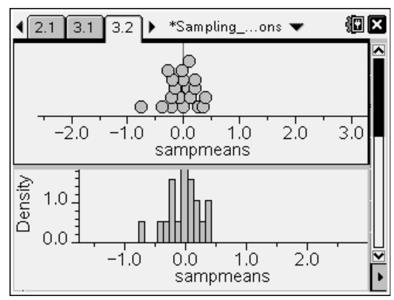
   Sample Answer: IQR (interquartile range) describes the spread of the middle half of the values being examined. An estimate of the IQR for the individual values in the samples is about 2. For the dot plot of $\bar{x}$, the IQR seems to be less than 1, maybe 0.5.

b. How do these two numbers (measures of spread) compare? Explain any differences you see.

Sample Answer: 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 dots in the sample can.

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 bottom screen on page 3.1.

Move to page 3.2.



8. The top screen of page 3.2 is an exact copy of the lower screen on page 3.1, with which you have been working. The lower screen displays the same data ($\bar{x}$ values from your samples) in a histogram. Comment on which display seems best for seeing the overall shape of the distribution. Explain your reasoning.

Sample Answer: The dot plot appears better. The histogram's bars are too separate and are all short, so the overall shape is not apparent.

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: ***Screen Capture***
Using a TI-Navigator class Screen 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 very different in detail, most will be centered somewhere near 0, and the apparent spreads of their plots will be about the same.

Return to page 3.1.

9. Click *draw* <u>many</u> more times (for about a total of 100).
   a. Does the dot plot seem to confirm your earlier predictions about the distribution of sample means? Explain why your prediction seems reasonable.

      <u>Sample Answer:</u> Yes. The dots still accumulate on either side of 0, roughly symmetrically. The center of this distribution still seems to be about 0, and the spread seems about the same as before.

   b. In Question 7 you estimated a measure of spread for the sampling distribution of sample means and compared the variability among sample means to the variability within the population itself. Revise that estimate and comparison based on this larger simulation.

      <u>Sample Answer:</u> For the dot plot of $\bar{x}$, the IQR still seems to be around 0.5.

Return to page 3.2.

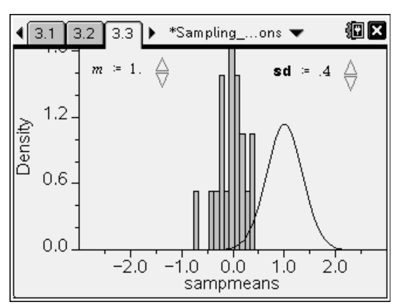
10. a. Which graph type seems most appropriate for this larger simulation? Explain.

      <u>Sample Answer:</u> Now the histogram shows the shape very well and appears less cluttered than the 100-point dot plot.

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

      <u>Sample Answer:</u> It appears to be normally distributed. The mean looks like 0. The standard deviation might be 0.5.

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 up arrows (▲) to select your choices for mean and standard deviation.

11. Use the up arrows (▲) to set the mean and standard deviation to match the estimates you made in Question 10. Then re-adjust the values if necessary so that they seem to fit your histogram as well as possible. Record your final values for mean and standard deviation. Comment on the accuracy of your predictions.

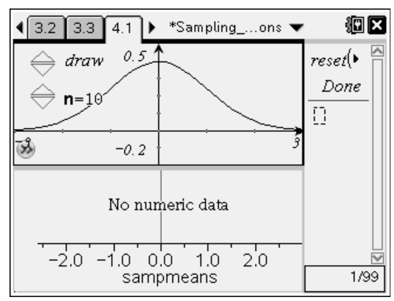
    Sample Answer: The mean was close, but the standard deviation is even smaller than the guess. The actual standard deviation value is more like 0.3.

TI-Nspire Navigator Opportunity: *Screen Capture*
A TI-Navigator class Screen Capture of students' page 3.2 should make clear the fact that although students all have very different samples and sample means, the overall distribution of those sample means follows a very clear pattern that can be summarized by its shape, center, and spread.

Move to page 4.1.

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

12. On page 4.1, you can take samples of sizes other than 10 by changing the value of *n*. Repeat the "sample-size-10" explorations you carried out in Questions 9–11, this time using a different sample size. Comment on how changing the sample size affects the center, spread, and shape of the distribution of $\bar{x}$. Be as specific as possible, indicating what happens when the sample size increases and what happens when the sample size decreases.



    Note: Type **reset()** in the right-hand panel of this page to erase an exploration using one sample size in order to begin another exploration with another sample size.

    Sample Answer: The overall results are the same, with the exception of spread. Larger samples lead to smaller spreads in the distribution of $\bar{x}$. Smaller samples let $\bar{x}$ vary more.

Teacher Tip: After completing an exploration, the right pane in page 4.1 may be used to reset the sampling process for another exploration. Students should type **reset()** as shown in the window when it first opens.

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

    Sample Answer: 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 the original population.

Wrap Up

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

- 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 than 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.