Math Objectives

- Students will recognize that samples have smaller variability than the population.
- Students will recognize that the variability in samples is a function of sample size, *n*, and that the standard deviation of the distribution of sample means is the sample standard deviation divided by the square root of *n* when the sampling is done with replacement.
- Students will look for and make use of structure (CCSS Mathematical Practices).
- Students will use appropriate tools strategically (CCSS Mathematical Practices).

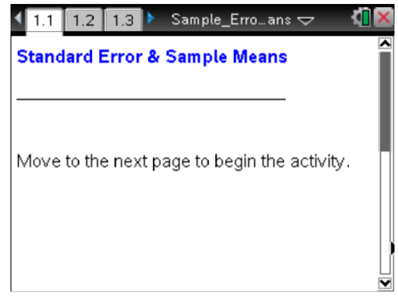
Vocabulary

- mean
- sample
- sample distribution of sample means
- standard deviation
- standard error
- variance

About the Lesson

- This lesson involves investigating the relationship between the standard deviation of a population, the area of a set of rectangles, and the standard deviation of the sampling distribution of sample mean areas of the rectangles.
- As a result, students will:
  - Simulate a sampling distribution of sample mean areas for samples of size 50 and note the mean and spread.
  - Generate a simulated sampling distribution of mean areas for samples of size 40 and compare it to the distribution for samples of size 50.
  - Use arrows to produce a series of simulated sampling distributions of sample mean areas for different sample sizes and make a conjecture about the possible relationship between the sample size and the standard deviation.
  - Observe the graph of the two variables (sample size and standard deviation of simulated sampling distribution of sample means) and model the relationship with a function, reflecting on why the function makes sense.

TI-Nspire™ Navigator™ System



TI-Nspire™ Technology Skills:

- Download a TI-Nspire document
- Open a document
- Move between pages
- Hover over a point in a scatter plot
- Click a minimized slider

Tech Tips:

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

Lesson Files:

Student Activity
Standard Error & Sample Means_Student.pdf
Standard Error & Sample Means_Student.doc

TI-Nspire document
Standard Error & Sample Means.tns

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

- Transfer a File.
- Use Screen Capture to compare different random samples generated by students.
- Use Teacher Edition computer software to review student documents.
- Use Quick Poll to assess students' understanding.

Prerequisite Knowledge

- Students should be familiar with sampling distributions.

Related Activities

- Statistics Nspired activity *Sampling Distributions*

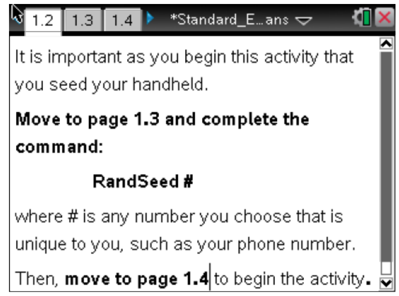
Discussion Points and Possible Answers

Tech Tip: Page 1.2 gives instructions on how to seed 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.)

Once students have seeded their random number generators, they do not have to do it again unless they have cleared all of the memory. But it is important that this be done if the memory has been cleared or with a new device, as otherwise the "random" numbers will all be the same as those on other similarly cleared devices.

Teacher Tip: The sampling in this activity is done with replacement because the population consists of the area of 100 rectangles, Sampling without replacement with this small population would violate the 10% rule. You might have students explore the Statistics Nspire activity **10% Rule**.

Move to pages 1.2 and 1.3.



This activity involves generating a number of random samples from a population. In order to avoid having your results be identical to those for another student in the room, it is necessary to "seed" the random number generator. Read the instructions on Page 1.2 for seeding your random number generator, then carry them out on Page 1.3.

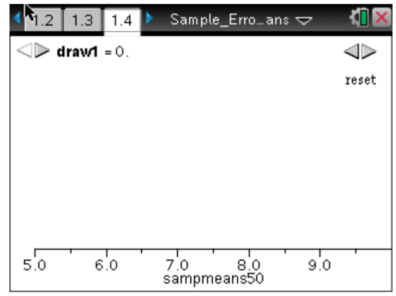
1. From your earlier work, what do you know about the mean of the sampling distribution of sample means?

   Sample Answer: As the number of samples increases, the mean of the sampling distribution of the sample means will tend to the mean of the population.

Teacher Tip: You might want students to examine the actual areas of the 100 rectangles, choose a random sample of size 10, and calculate the mean area for their sample as a precursor to the simulations that follow. The areas of the rectangles are on Page 1.8 of the .tns file. Students can use the right side of the screen on Page 1.8 to generate random rectangles and their areas.

- With the cursor at the end of RandSamp(**population**,10)|, press enter.
- After students find the mean area of their sets of randomly chosen rectangles, collect their means either having them report out and recording on the board in a stem-and-leaf plot or using the TI-Nspire Navigator data accumulation feature.
- Discuss the distribution with the class. You can use this experience to help them make sense of the distributions they obtain in Question 2.

Move to page 1.4.



2. The top arrow on Page 1.4 will generate a sample of size 50, chosen with replacement, from a population of rectangles and will display the mean area of the rectangles in the sample.
   a. Generate 10 samples. Describe the simulated distribution of the mean areas for the samples, and make a rough sketch of what you see.

   <u>Sample Answer</u>: The simulated sampling distribution goes from a mean area of 6.4 sq. units to 8.2 sq. units with most around 7.2 sq. units. Sketches will vary.

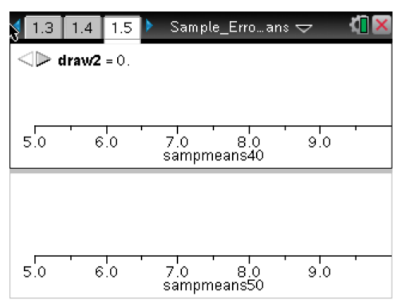
   b. The vertical line will display the mean of the simulated sampling distribution of the mean areas. Click on the line, and record the result on your sketch above.

   <u>Sample Answer</u>: The mean area of the ten sample mean areas is 7.3037.

3. Click the top arrow on Page 1.4 to generate five new samples and plot the sample mean for each of the new samples. Generate 45 more samples. How did the simulated sampling distribution change as the number of samples changed?

   <u>Sample Answer</u>: The simulated distribution of sample means became more symmetric and mound shaped, with the center around 7.4.

Move to page 1.5.



The bottom work area on Page 1.5 is the same as Page 1.4. The arrow in the top work area will generate samples of size 40 from the population and display the simulated distribution of the sample means.

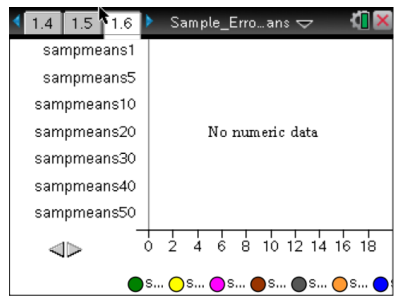
4. Generate 50 samples of size 40. How do the simulated sampling distributions for the two sample sizes compare?

   <u>Sample Answer</u>: The means of the two simulated sampling distributions seem to be about the same, 7.42 and 7.39, but the simulated distribution for sample size n = 40 seems to have a slightly larger spread than the one for n = 50.

Teacher Tip: It is important to note that, by chance, the simulated distributions might not follow the pattern where increasing the sample size corresponds to a decrease in the spread. Sharing student work at this point will help make the point that, in general, as *n* increases the spread decreases. Students might use TI-Nspire Navigator as in the note below, share sketches on the board, or line up their handhelds and walk around observing all of the simulated sampling distributions.

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

Move to page 1.6.



The mean of the sampling distribution of all sample means will be the population mean. A question of interest is how the standard deviation of the sampling distribution of sample means behaves.

5. On Page 1.6, click the arrow to display simulated sampling distributions for 50 samples of different sample sizes.

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

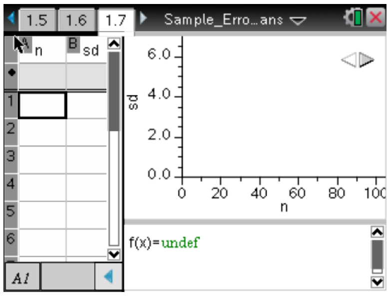
a. What conjecture can you make about the standard deviation of distributions of sample means for samples of size *n* drawn from the same population? Justify your answer by referring to the graphs. (Note that using the left arrow will generate 50 new samples for each sample size.)

<u>Sample Answer</u>: It looks like the standard deviation of the distribution of sample means decreases as the sample size increases. When the sample size is 1, the spread goes from areas of 1 sq. unit to 18 sq. units; for 5 it goes from 4 to 16 sq. units; and continues to decrease until for sample size 50, it goes from a little under 6 to over 8 sq. units. The standard deviation is getting smaller and smaller as *n* increases.

b. What value would you expect for the standard deviation of the actual sampling distribution of all sample means for sample size 1? Explain your reasoning.

<u>Sample Answer</u>: The standard deviation should be the same as the population standard deviation because if you draw samples of size 1, you are drawing individual members of the population and eventually will have the distribution of the entire population.

Move to page 1.7.



6. The spreadsheet shows the standard deviations for simulated sampling distributions of 100 sample means for different sample sizes.
   a. Hover over a point in the graph. Explain what the coordinates represent.

Sample Answer: The coordinates (explanatory variable, response variable) are (sample size, standard deviation for the simulated sampling distribution of sample means for that sample size). For example, the coordinates (10, 1.67) indicate that for a sample size 10, the standard deviation of the simulated sampling distribution of sample mean areas is 1.67 square units.

   b. What function do you think might model the relationship between the variables? Explain your reasoning.

Sample Answer: It looks like the standard deviation could be decreasing exponentially as the sample size increases, but the standard deviation values are not decreasing at a constant multiplicative rate. It might be a power relationship.

   c. Click the arrow once to see a power regression as a model of the relationship between the sample size and standard deviation. Does the equation support your conjecture? Why or why not?

Sample Answer: The equation of the graph is $y = 5.07311x^{-0.49}$ or $y = \frac{5.07311}{x^{0.49}}$. It seems like the relationship could be a power regression.

   d. The mean of the areas, the population in the activity, is 7.4 sq. units, and the standard deviation is 5.19 units. Click the arrow again to see the theoretical model for the relationship between the sample size and the standard deviation in the simulated sampling distribution of sample means. How does the theoretical model compare to the regression equation for your data? (Note the model uses the variable *x* to represent *n*, the sample size.)

Sample Answer: The theoretical model, $y = 5.19x^{-0.5}$, was very close to the regression equation.

7. The standard deviation of the sampling distribution of sample means is called the **standard error** of the distribution.

   a. Suppose you had a sampling distribution of sample means for samples of size 64. If the population standard deviation is 9.6, what is the standard error of the sampling distribution?

   Sample Answer: The standard error would be $9.6/\sqrt{64}$ or 1.2 units.

   b. In general, how can you find the standard error for a sampling distribution of sample means for samples of any size?

   Sample Answer: The standard error of the sampling distribution of the sample means for samples of size *n* is the population standard deviation divided by the square root of the sample size when the sampling is done with replacement.

> **Teacher Tip:** Depending on your class, you might want to discuss the mathematics that underlies the formula. The variance of a sum of independent random variables is the sum of the variances of the individual random variables. The statistic called "sample mean" is obtained by adding "*n*" independent, but identically-distributed, random variables (namely, *x*), and then dividing by *n*. Therefore, the variance of the sampling distribution of sample means will be *n* times the variance of x (that is, the population variance), divided by n^2. After reducing, this becomes Var(x-bar) = Var(x)/*n* which gives, after square roots, the formula.

TI-Nspire Navigator Opportunity: *Quick Poll*
See Note 3 at the end of this lesson.

Wrap Up

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

- The standard deviation of the sampling distribution of sample means decreases as the sample size increases.
- The formula for finding the standard error (deviation) for the sample distribution of sample means is a function of the sample size and the population standard deviation. $s = \frac{\sigma}{\sqrt{n}}$.

Assessment

Choose the most appropriate response. Be prepared to support your answer from the work you did in the activity. (Note an * marks the correct answers.)

1. The standard deviation of the sampling distribution of sample means is
   a. the same as the standard deviation of the population.
   b. less than the standard deviation of the population.*
   c. more than the standard deviation of the population.
   d. cannot tell how it is related to the population standard deviation.

2. As the sample size increases, the standard deviation of the sampling distribution of sample means
   a. increases.
   b. decreases.*
   c. stays the same.
   d. cannot tell whether it increases or decreases.

3. Suppose the population standard deviation is 36. If you wanted to make the standard deviation of the sampling distribution of sample means less than 9, your sample size should be
   a. 4
   b. 9
   c. 16
   d. 25*

TI-Nspire Navigator

Note 1

Question 4, Name of Feature: Screen Capture

A Screen Capture can be taken at this point to compare the simulated sampling distributions of sample means for samples of size 40 and size 50.

Note 2

Question 5, Name of Feature: Screen Capture

A Screen Capture can be used to compare the similarity and differences among the sampling distributions of the sample means for the different sample sizes.

Note 3

Name of Feature: Quick Poll

A Quick Poll can be given at the conclusion of the lesson. You can save the results and show a Class Analysis at the start of the next class to discuss possible misunderstandings students might have.