

Central Limit Theorem – ID: 9998

By Mark Howell

Time required
45 minutes

Activity Overview

This activity introduces students to the Central Limit Theorem by letting them explore sample distributions. By changing sample size they observe and identify the effect on distributions of sample means. The activity can serve as a first introduction to sampling distributions. First, students investigate the idea of a random sample, and look at means of random samples of different sizes. Second, they look at histograms made from sample means of various sizes. Students are led to make qualitative statements about the behaviors of such distributions. Finally, students use a slider to control the sample size, and watch the shape of the sampling distribution change as the size of the sample changes.

Concepts

- Mean of a sample
- Sample size
- Sampling distribution
- Central Limit Theorem

Teacher Preparation

This activity is designed for use in an introductory Statistics classroom. It can be used to introduce the idea of random samples, sample mean, sampling distributions, and the Central Limit Theorem.

- *Prior to beginning this activity, students should understand the idea of mean, and have experience interpreting the shape of a histogram. The student should have a basic understanding of a normal distribution.*
- *The screenshots on pages 2–4 (top) demonstrate expected student results. Refer to the screenshots on pages 4 (bottom) and 5 for a preview of the student TI-Nspire document (.tns file).*
- **To download the student .tns file, go to education.ti.com/exchange and enter “9998” in the quick search box.**

Classroom Management

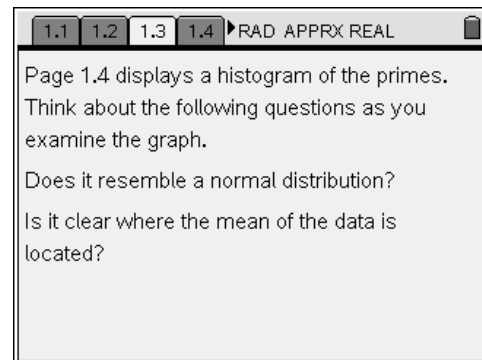
- *This activity is designed to be **student-centered** with the teacher acting as a facilitator while students work cooperatively.*
- *For the most part, students will perform simple calculations and input data. Therefore, a basic working knowledge of the TI-Nspire handheld is sufficient.*
- *The ideas contained in the following pages are intended to provide a framework as to how the activity will progress. Suggestions are also provided to help ensure that the activity is completed successfully.*

TI-Nspire™ Applications

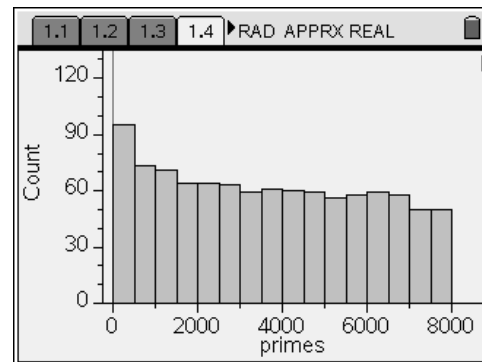
Calculator, Graphs & Geometry, Lists & Spreadsheet, Notes, Data & Statistics

Problem 1 – Sampling Distributions

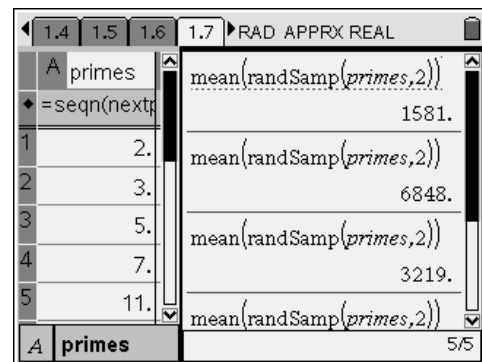
Step 1: Students will be introduced to the data set of primes on page 1.2 that they will use to analyze sampling distributions. Help students answer the questions on page 1.3 (shown to the right). Students may need to be reminded of the normal distribution and how to identify skew and mean in a histogram.



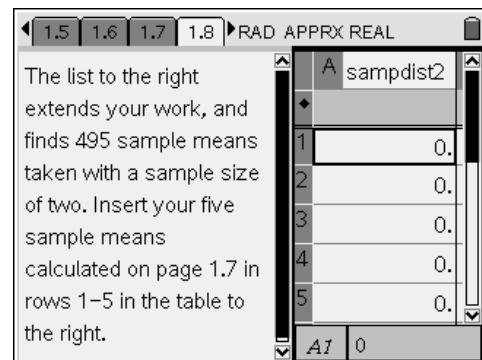
Step 2: Students should observe that data is not normally distributed and is skewed to the left slightly. The Central Limit Theorem is introduced informally. Students will explore the Central Limit Theorem and define it more formally later in the activity. Help students to understand that the distribution on page 1.4 can also be thought of as a sampling distribution with sample size of one.



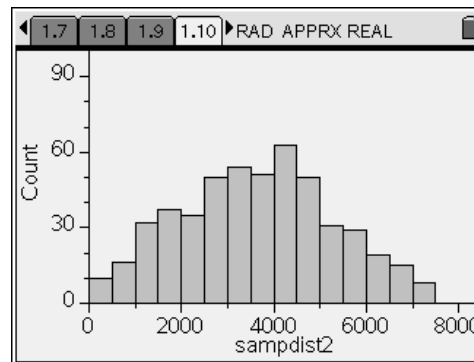
Step 3: On page 1.7, students will take five random samples, each consisting of two prime numbers, and take the sample mean. They will do this using the **Mean** command (**MENU > Statistics > ListMath > Mean**). Students may need clarification on the difference between sample and population mean.



Step 4: Students will insert the data they calculated on page 1.7 into the table on page 1.8. Make sure students are overwriting the zeros in rows 1–5. Students will now be able to see their data actually included in the histogram on page 1.10.



Step 5: Students observe the histogram of the sampling distribution with a sample size of two. Students should be able to articulate some of the characteristics of the graph. To further understanding have the students answer the questions from page 1.3 for this new histogram.

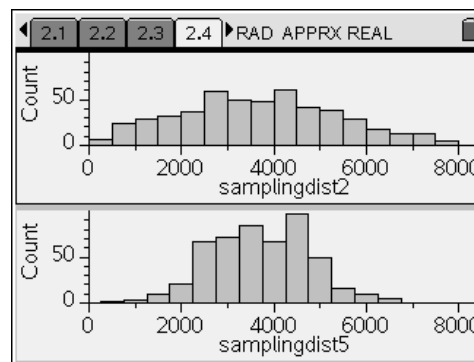


Step 6: Students should spend some time predicting how the distribution will change when the sample size increases to five.

Problem 2 – Increasing Sample Size

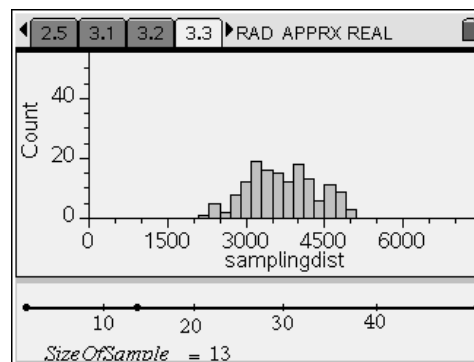
Step 1: The student will now look at a sampling distribution with a sample size of five. The command **samplemeans(primes,5,500)** is used to create a column of data that consists of 500 sample means, with each sample having a size of $n = 5$. The function **samplemeans** is a hidden function and understanding how it works is not necessary for understanding the concepts in this activity. Students should think about the shape of the histogram before they advance to page 2.2.

Step 2: Now that they have seen the sampling distribution for a sample size of five, students will compare it with the sampling distribution they saw on page 1.10. Have students articulate at least three comparisons about the two graphs.



Problem 3 – Sampling Distributions and Population Mean

Step 1: Students will use a slider to control the sample size, and generate many sampling distributions. They see the highest frequency in the histograms generally increases with increased sample size, and the variability generally decreases with increased sample size. If the program is running slowly on the TI-Nspire handheld device, allow students to use the computer software. Have students change the size of the sample using the slider and observe how the distribution changes, comparing their graphs with other students, to show that even though each student's data is unique the graphs will resemble each other.



Step 2: Students calculate the population mean to use as an indicator for evaluating their histograms.

Step 3: Students will plot the population mean on their histograms to show that the sampling distribution reaches its highest frequency at the population mean, and that this point is a point of symmetry for the distribution. Reinforce that the sample means in the distribution will approach the population mean when the sample size is increased. Students will now see a more formal definition of the Central Limit Theorem.

Central Limit Theorem – ID: 9998

(Student)TI-Nspire File: *StatAct10_CentralLimitThm_EN.tns*

1.1 1.2 1.3 1.4 ▸ RAD APPRX REAL

CENTRAL LIMIT THEOREM

Statistics

Using the Central Limit Theorem

1.1 1.2 1.3 1.4 ▸ RAD APPRX REAL

You will use the first 1000 prime numbers, listed to the right in Column A, to analyze sampling distributions.

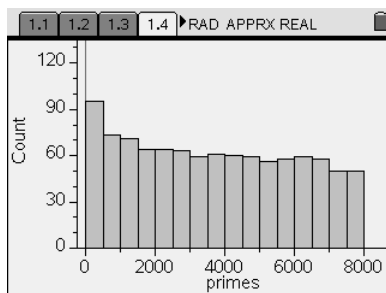
A primes	
1	2.
2	3.
3	5.
4	7.
5	11.

A primes

1.1 1.2 1.3 1.4 ▸ RAD APPRX REAL

Page 1.4 displays a histogram of the primes. Think about the following questions as you examine the graph.

Does it resemble a normal distribution?
Is it clear where the mean of the data is located?



1.2 1.3 1.4 1.5 ▸ RAD APPRX REAL

The graph on page 1.4 shows clearly that the data is not normally distributed.

Even though the original population is not normally distributed, the Central Limit Theorem states that as the sample size increases, the distribution of the sample mean will approach a normal distribution.

1.3 1.4 1.5 1.6 ▸ RAD APPRX REAL

Start by looking at a distribution with a sample size of two. The next page contains a calculator screen ready to calculate the mean, or average, of samples of two prime numbers taken from our list of primes. Press enter five times to execute the `mean(randSamp(primes,2))` command and get five means.

1.4 1.5 1.6 1.7 ▸ RAD APPRX REAL

A primes	
1	2.
2	3.
3	5.
4	7.
5	11.

A primes

mean(randSamp(primes,2))

1.5 1.6 1.7 1.8 ▸ RAD APPRX REAL

The list to the right extends your work, and finds 495 sample means taken with a sample size of two. Insert your five sample means calculated on page 1.7 in rows 1–5 in the table to the right.

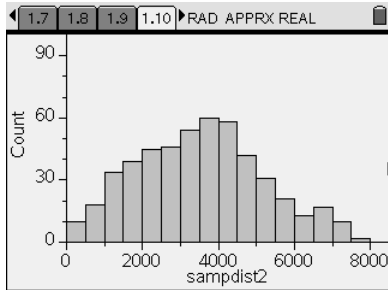
A sampdist2	
1	0.
2	0.
3	0.
4	0.
5	0.

A7 0

1.6 1.7 1.8 1.9 ▸ RAD APPRX REAL

The data collected on page 1.8 is graphed on page 1.10. Think about the following questions as you examine the graph.

Does it resemble a normal distribution?
How does it compare with the parent distribution on page 1.4?

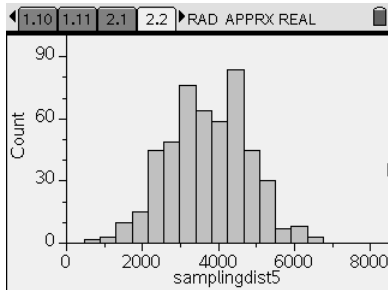


The data is much closer to a normal distribution. The data is less varied and more uniform.

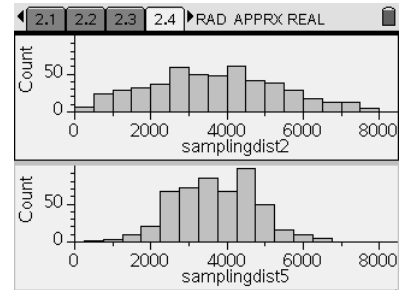
What do you think will happen if the sample size is increased to five?

Column B contains 500 means of random samples of five prime numbers. Page 2.2 displays a histogram of the sample means. What would you predict as the population mean of this data?

	B
1	2. 4849.
2	3. 3329.
3	5. 5781.
4	7. 3491...
5	11. 3908...



On the next page the distribution for a sample size of two is compared with a sample size of five. What do you notice about the two graphs?

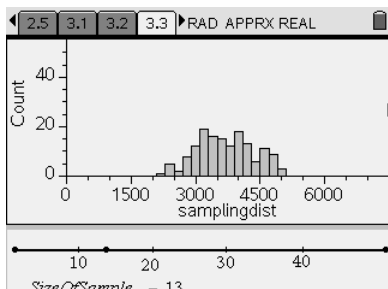


You should have seen that the means became less "spread out" with an increase in sample size. The larger the sample size, the more clumped the sample means became.

Now you will extend these observations, by looking at more sample sizes and observing the effect on the shape of the distribution of sample means.

The next page will adjust the data in the table to the left and create a histogram for various sample sizes. Drag the point under the histogram to change the sample size. Observe how the shape of the sampling distribution changes.

	prime
1	30.85
2	39.15
3	19.15
4	78.69
5	02.54



Execute the command `mean(primes)` in the calculator below to calculate the population mean of the first 1000 prime numbers.

Locate the population mean on the histogram on page 3.3 by choosing plot value from the Tools menu, and inputting the mean you calculated on page 3.4. Adjust the histogram to see how the sample means approach the population mean as the sample size increases.

Evidenced by the distribution of primes, the Central Limit Theorem implies that regardless of the original distribution of data, the distribution of sample means will approach a normal distribution, and the sample means will approach the population mean, as the sample size increases.