

# How Long



## Activity

7 8 9 10 11 12



TI-84Plus CE



Investigation



Student



50 min

## Problem Statement

Imagine two points randomly placed in a  $1 \times 1$  square on the Cartesian plane. What would be the average length of the line connecting these two points?

## Investigation

### Question: 1.

Determine the range of lengths that would be possible.

Range:  $[0, \sqrt{2}]$

### Question: 2.

Estimate the mean length of the line and justify your reasoning.

Answers will vary, however student reasoning should include reference to the mean length of a line segment located on a straight line of length one unit. (Mean = 0.5) It is reasonable to assume that the mean length of the line segment in the  $1 \times 1$  square would not exceed half the previously mentioned range (Q.1). Student estimates for the mean length should therefore lie between 0.50 and 0.71. Limitations imposed by the square may also prompt students to estimate closer to 0.50 than 0.71.

## Sampling

Load the HOWLONG program onto your calculator. The program provides three options:

- **Observe**

Watch as a series of line segments are generated and their length computed.

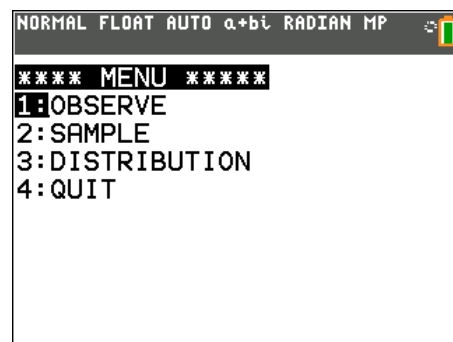
- **Sample**

Generate multiple line segments and record their lengths and compute the mean length.

- **Distribution**

Multiple samples are generated and the mean of each sample is recorded. The results are progressively graphed to show the distribution of the sample means.

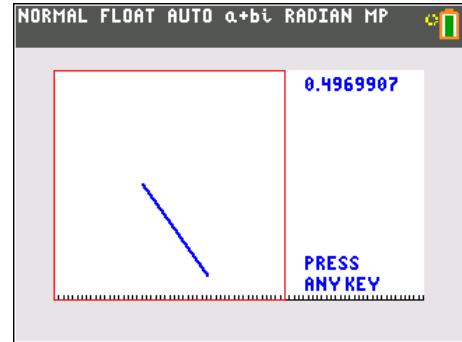
Run the HOWLONG program and select "observe" from the menu.



Generate and record the lengths of 10 line segments using the Observe menu option.

**Note:**

The results are not recorded on the calculator. Results must be recorded manually for this section of the program. Press any key to generate the next sample, after all the samples have been generated the program will return to the main menu.



**Question: 3.**

Record the lengths of the 10 randomly generated line segments and calculate the mean.

Answers will vary, typically the mean length will vary between 0.30 and 0.74

**Question: 4.**

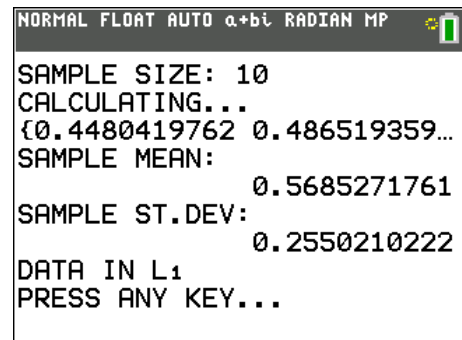
If more line segments were generated, would this produce a more reliable estimate of the mean?

“Yes”.

**Teacher Notes:** The purpose of this question is to drive the intuitive understanding that a larger sample will provide a more accurate estimate of the mean. It is not the intention for students to provide any mathematical evidence to support their response.

From the main menu select: Sample

The sample will generate and record the results for the requested number of line segments, compute the sample mean and standard deviation accordingly. Use this menu option to complete the following questions.



**Question: 5.**

Generate a sample of 10 line segments. Repeat this process 6 times and record the mean segment length for each sample.

Sample	1	2	3	4	5	6
Mean	0.519	0.595	0.418	0.509	0.597	0.398

Results will vary; the data shown above is representative of 6 samples consisting of 10 line segments each.

**Question: 6.**

Generate a sample of 100 line segments. Repeat this process 6 times and record the mean segment length for each sample.

Sample	1	2	3	4	5	6
Mean	0.516	0.523	0.540	0.541	0.521	0.474

Results will vary; the data shown above is representative of 6 samples consisting of 10 line segments each.

**Question: 7.**

Compare the distribution of the sample means in Questions 5 and 6.

Answers will vary depending on the data sets. It is reasonable to assume that students should identify:

- ✓ The means of each sample are approximately the same.
- ✓ The spread (standard deviation) of data is smaller in the second sample, more consistent.

The purpose of the question is for students to develop an intuitive understanding that the standard deviation in the second 'distribution' is smaller as the sample size was larger.

**Question: 8.**

Daisy generates 50 samples and records the mean of the 10 line segments generated in each sample. Which sample distribution will most likely resemble Daisy's, the sample distribution from Question 5 or 6?

Daisy's sample distribution is more likely to resemble the one in Question 5 since each sample is the mean of 10 segments. Students should support their decision by referring to the distribution (variation) of the sample means.

**Question: 9.**

Craig generates 50 samples each consisting of the mean of 10 lines segments. Chris generates 50 samples each consisting of the mean 50 line segments. Unfortunately both Chris and Craig were recording their results on scrap pieces of paper. A sample mean of 0.35 was recorded, who is most likely to have generated this sample mean?

Craig – His samples involve a smaller quantity of line segments, therefore there will be a much greater variation in the sample means. Since 0.35 is much lower than the expected value it is more likely to have come from Craig.

This question assumes students have already obtained a reasonable estimate of the mean. (Question 2)

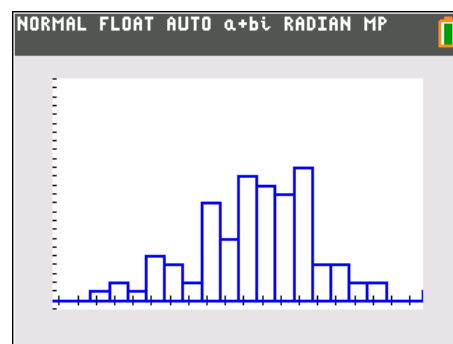
The HOWLONG calculator program can generate multiple samples and record the mean of each.

From the main menu select: Distribution

Multiple samples will be generated; the mean of each sample will be recorded in List 1 ( $L_1$ ) and plotted progressively. (Sample opposite)

**Note:**

This is a relatively slow procedure, depending on the quantity of line segments in each sample. For example generating 50 samples each consisting of 100 line segments will take approximately four minutes to generate. Remember that is 5000 line segments and 5000 length calculations for the 10,000 randomly generated points!

**Question: 10.**

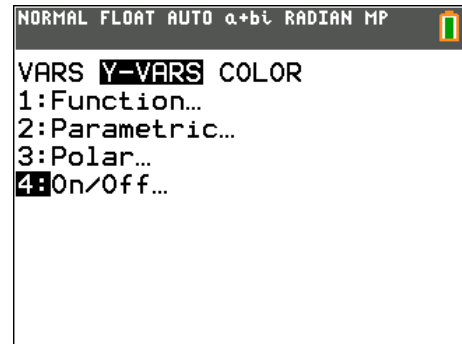
Using the distribution option in the program, generate 100 samples each consisting of 10 line segments. Describe the approximate shape of the sampling distribution for the sample means.

The distribution is approximately normal, students may not realise this based on 100 samples, and may wish to do 200 samples, this will take approximately 3 minutes to generate the samples and graph.

Once the large sampling distribution from Question 10 is completed, select the option to exit the program.

Use the [Vars] key on the calculator and locate Y-Vars.

Select the option to switch the Y-Variables on. This will produce a graph to approximate the data.



### Teacher Notes:

The purpose of this activity is to informally introduce the Central Limit Theorem with a focus on:

- Normality of the sampling distribution for the sample means
- Difference between sample size and the number of samples
- Understanding that increasing the sample size corresponds to a reduction in the standard deviation for the sampling distribution.

In the last question students may also choose to generate more samples. Generating 200 samples each of size 15 takes several minutes; however, the resulting distribution is more readily identifiable as approximately normal.

When exiting the program a normal distribution is generated and stored in  $Y_0$  and can be graphed without checking the equation by switching “Functions On” using the instructions provided. A sample is provided below.

