

# Project 4 – Comparing a Sample Mean to a Population Mean Using Simulation

## Simulation

In this case we will begin with the population data for a quantitative variable, and that means that we know the population mean,  $\mu$ . We will draw a random sample and compute its sample mean  $\bar{x}$ . Then, using simulation, we will select many random samples of the same size to determine if the sample is an unusual sample.

- If you have a sample of size  $n$ , draw a sample of size  $n$  from the population using sampling without replacement.
- Compute the mean of that sample.
- Repeat the process many times. (We will repeat the process 1000 times.)
- Find the critical values for these 1000 means by finding the 2.5<sup>th</sup> percentile and the 97.5<sup>th</sup> percentile.  
These values separate the lower 2.5% and upper 2.5% of the sample means from the rest, and can be used to determine if your sample is unusual. These can also be thought of as the boundaries that separate the samples for which we would make an error when estimating the population mean.
- We will also determine what percent of all samples are at least as extreme as your sample.

## Investigation – Ages of Mr. Woodbury’s Students

**(a)** Open the Project 3 data set from the StatCrunch group for this class. You will find the ages of each student in Mr. Woodbury’s classes this semester. In this project we will treat this data set as a population.

- What is the population mean age  $\mu$  for Mr. Woodbury’s students? \_\_\_\_\_

**(b)** Select a sample of 8 nearby students in your class, write down their ages, and compute the sample mean age.

- Sample ages: \_\_\_\_\_
- Sample Mean: \_\_\_\_\_

### StatCrunch Step 1 – Select 1000 Samples of Size $n = 8$

Data > Sample

Select the column that contains the data.

Enter the sample size & the number of samples.

Store samples as “Stacked with a sample id.” Click Compute!.

### StatCrunch Step 2 – Compute the 1000 Sample Means

Stat > Summary Stats > Columns

Click the column that contains the sample data. (It will be named “Sample(variable)”.)

Group by “Sample” from the pull-down menu.

Select the Mean under Statistics.

Check the box “Store in data table.”

Click Compute!.

[Note: A popup window may appear that says, “Whoa! Lots of unique numeric values for Sample.” Click Cancel when this popup appears to avoid “binning” the data.]

### StatCrunch Step 3 – Compute the Critical Values

Stat > Summary Stats > Columns

Select the column labeled Mean.

Type 2.5,97.5 in the box labeled Percentiles and click Compute!.

- Lower Critical Value: \_\_\_\_\_ Upper Critical Value: \_\_\_\_\_

StatCrunch Step 4 – Determining which sample means are greater than (or less than) yours?

Greater Than – Use When $\bar{x} > \mu$	Less Than – Use When $\bar{x} < \mu$
<p>Data &gt; Compute &gt; Expression In the Expression box, type</p> <p><i>between(Mean, value of your sample mean, an upper bound that is definitely higher than all possible values)</i></p> <p>For example, if the population mean age is 22 and your sample mean age is 24, type</p> <p style="text-align: center;"><i>between(Mean,24,200)</i></p> <p>(I chose an upper limit of 200 because no mean age should be above 200.)</p> <p>Click Compute!</p>	<p>Data &gt; Compute &gt; Expression In the Expression box, type</p> <p><i>between(Mean, a lower bound that is definitely less than all possible values, your sample mean)</i></p> <p>For example, if the population mean age is 22 and your sample mean age is 19, type</p> <p style="text-align: center;"><i>between(Mean,0,19)</i></p> <p>(I chose a lower limit of 0 because no mean age should be below 0.)</p> <p>Click Compute!</p>

You will end up with a new column containing the words true or false.

True: That sample mean is at least as extreme as yours.

False: That sample mean is less extreme than yours.

StatCrunch Step 5 – Computing the Percent of Sample Means that Are at Least as Extreme as Yours

Stat > Tables > Frequency

Select the “between” column.

Be sure Frequency and Relative Frequency are highlighted.

Click Compute!.

The relative frequency for true will tell you what percent of the sample means were more extreme than yours.

- What percent of the samples produced a mean at least as extreme as yours?
- Is the sample that you took by hand considered an unusual result?