Central Limit Theorem Exercise

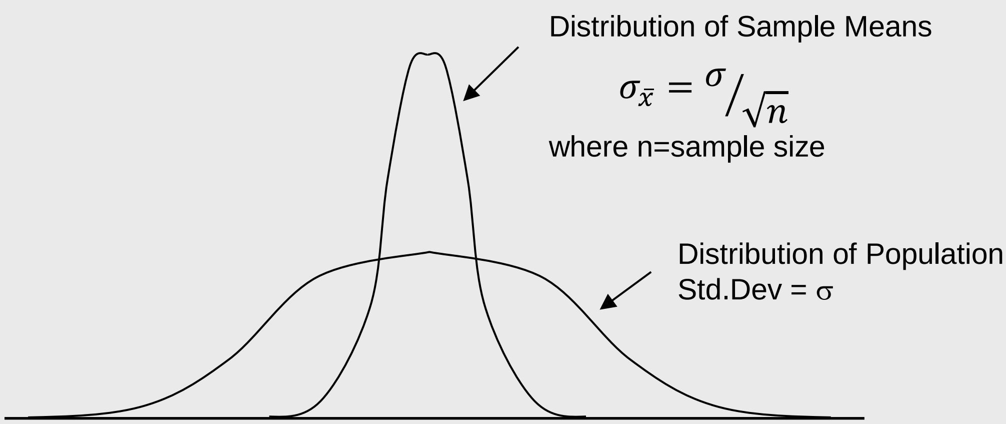
In this exercise, you will be using a simulation to explore and help better understand the concepts of the Central Limit Theorem.

Central Limit Theorem

If a random sample of size n is drawn from a population with mean μ and variance, $\boldsymbol{\sigma}^2$,

then the sample mean, $\bar{x}$, has approximately a normal distribution with mean μ and variance $\boldsymbol{\sigma}^2/n$.

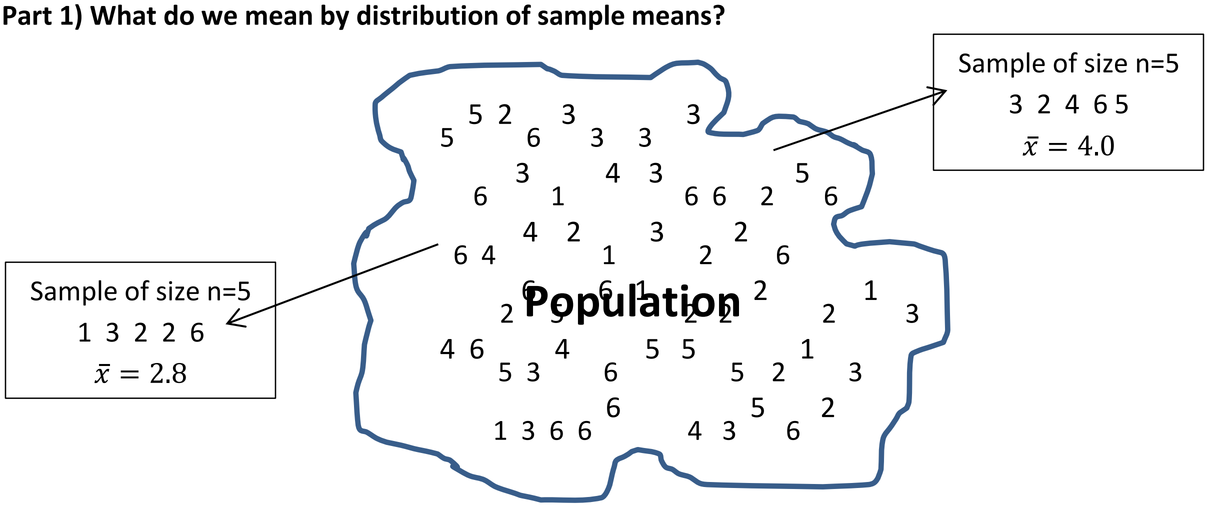
That is, the distribution function of $\frac{\bar{x}-u}{\sigma/\sqrt{n}}$ is approximately a standard normal. The approximation improves as the sample size increases.



Open **the Rice Virtual Lab in Statistics** by clicking on the link in Blackboard. The click on **Simulation/Demonstrations**, then **Sampling Distributions Simulation**.

Click on the **Begin** button – it may take a few minutes to load.

Part 1) What do we mean by distribution of sample means?



In the simulation, the histogram at the top is of all possible outcomes of a population. The default when the simulation is opened is a normal distribution with a mean of 16 and standard deviation of 5.

Click on the **Animated** button. You should see 5 “items” being taken from the parent population falling into the next graph. We just took a sample size n=5.

What fell into the third graph was the average of the 5 values, $\bar{x}$ for the sample.

Clicking Animated again, will show another sample of size 5 falling, and then the $\bar{x}$ falling.

To the left of the third chart you will see the Reps, which is the number of samples of size n that have dropped. This is then the number of $\bar{x}$’s that are in the histogram.

If you click the 5 button, it will drop 5 samples at a time. Click the 5 button several times to see the histogram build, until you have around 200 samples of size 5 (200 reps).

You should see the **distribution of sample means** begin to appear to look like a normal distribution.

If you click the 10,000 button, it will include the sample means for 10,000 more samples of size n. Now there should be no mistaking that the distribution of sample means follows a normal distribution.

Part 2) Standard error of the means?

We know that, theoretically, the relationship between the standard deviation for a parent population and the standard deviation for the distribution of sample means is given by the equation for the standard error of the means:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

The parent population is normally distributed with a mean of 16 and a standard deviation, σ, of 5 as shown in the upper left hand corner in the simulation.

For the sample size of 5, what would you expect the standard deviation, $\sigma_{\bar{x}}$, to be for the distribution of means?

How does your calculation compare to the standard deviation (sd) in the simulation for the sample means (next to the third graph area?

Part 3) Impact of sample size

Click on the **Clear lower 3** button at the top right.

In the bottom right hand corner, use the dropdowns to set bottom graph to be used for Mean, with N=25.

Click on the 10,000 button to drop samples. In the third graph you will still have a graph of the distribution of means for n=5. The bottom graph will show the distribution of the sample means for n=25.

How do the graphs compare to the population and to each other?

Calculate the standard error of the mean for n=25. How does it compare with the standard deviation for the simulated data?

From the simulation you should be able to clearly see that as the sample size increases, the distribution of the means becomes more narrow (lower standard deviation), and the mean values will fall closer to the mean of the population.

Part 4) Use with different types of distributions

Clear the lower three again.

Use the dropdown box (where normal now appears) to select the Uniform distribution.

What is the mean and standard deviation of this parent population?

Mean = *Standard deviation =*

Based on the formula for the standard error of the means, what would you expect the standard deviation to be for the distribution of the means for samples

of size n=5?

of size n=25?

Drop 10,000 samples - sample means will fall into the lower boxes for samples of size n=5 and 25.

Compare the two distributions of sample means to the population and each other. Comment.

Compare the theoretical standard deviations with the standard deviations in the simulation. Comment.

Clear the lower three again.

Use the dropdown box to select a **Skewed** distribution.
Drop 10,000 samples - sample means will fall into the lower boxes for samples of size n=5 and 25.

Comment on the shape and spread of the distribution of the means in the lower boxes.

Optional: Use the dropdown box to select a **Custom** distribution. Use the mouse over the parent population to draw in a distribution. Drop the samples to see the resulting distributions.

Part 5. Key Learning Points

From the simulation I learned that (check all that apply).

____ As the sample size increases, the standard deviation of the distribution of the means will decrease.

____ As the sample size increases, the means will fall closer to the true mean of the population.

____The standard error of the mean, calculated with the sample size and the original population's standard deviation, will provide the theoretical standard deviation for the distribution of sample means.

____The Central Limit Theorem is applicable to any distribution.

____ This simulation helped me understand the Central Limit theorem.