Name:_____________________________________

Purpose

After completing this lab, you should be able to distinguish between the distribution of the population, distribution of the sample, and the sampling distribution of a statistic. You should also be able to explain the circumstances under which the Central Limit Theorem (p. 281) ensures that sampling distributions of statistics are approximately normal.

In this activity, you will run the Sampling Sim program to investigate the sampling distribution for $\bar{X}$ and thus see the Central Limit Theorem in action. The first part of this activity takes you through how to use the Sampling Sim program and reviews some basic concepts along the way (parts (a)-(i)). Once you know how to use the simulation and understand what information it is providing, continue to complete the remainder of the activity.

Scenario: Professor Lectures Overtime

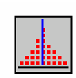
Let $X =$ amount of time a professor lectures after class should have ended. Suppose these times follow a Normal distribution with mean $\mu = 5$ min and standard dev $\sigma = 1.804$ min.

(a) Draw a rough sketch (and label) this distribution.

(b) Is μ a parameter or a statistic?

(c) Suppose you record these times for 5 days $X_1, X_2, \ldots X_5$ and calculate the sample mean $\bar{X}$. Is $\bar{X}$ a parameter or a statistic?

To investigate the sampling distribution of $\bar{X}$, we will take many samples from this population and calculate $\bar{X}$ value for each sample. Open the program *Sampling SIM* by double clicking on its icon.


Sampling Distribution...

- Click the DISTRIBUTION button and select "Normal" from the list. You should see a sketch similar to what you drew in (a).
- From the Window menu, select "Samples."
- Click Draw Samples and one observation from the population is selected at random (Note: the program may be very slow the first couple of times you click this button). This is one realization of the random variable X.

(d) How long did the professor run over this time?

(e) Click Draw Samples again, did you observe the same time?

(f) Change the value in the Sample Size box from 1 to 5 and click Draw Samples. How does this distribution compare (roughly) to the population distribution?

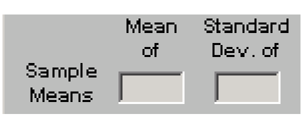
(g) Click Draw Samples again. Did the distribution of your 5 sample values change?

(h) Change the sample size from 5 to 25 and click Draw Samples. Describe how this distribution differs from the ones in (f) and (g). How does the shape, center, and spread of this distribution compare to that of the population (roughly)? (The mean of the sample values is represented by $\bar{X}$, the standard deviation of the sample values is represented by s. Compare these values to μ and σ.)

(i) Click Draw Samples again. Did you get the same distribution? The same $\bar{X}$ and s values?

The main point here is that results vary from sample to sample. In particular, statistics such as $\bar{X}$ and s change from sample to sample. This is *sampling variability* in action. You will now look at the distribution of these statistics (rather than the distribution of individual sample values).

From the Windows menu, select Sampling Distribution. Move this window to the right so you can see all three windows at once. You should see one very small green bar in this window (it will be small and on the *x*-axis). This is the $\bar{X}$ value from the sample you generated in (i). In the Sampling Distribution window, click on "New Series" so it reads "Add More." Click the Draw Samples button. A new sample appears in the Sample Window and a second green bar appears in the Sampling Distribution window for this new sample mean. Click the Draw Samples button until you have 10 sample means displayed in the Sampling Distribution window. *Note*: You can click the F button in the Samples Window to speed up the animation. Record below the values displayed in the "Mean of Sample Means" box and in the "Standard Dev. of Sample Means" box. These values are empirical. Compare these to the theoretical values predicted by the theorem at the top of page 278 of your text ("Mean and Standard Deviation of a Sample Mean").



Mean of Sample Means ___________

Standard Dev. of Sample Means __________

Record below the corresponding *theoretical* values from the theorem (p.278):

Mean of Sample Means ___________

Standard Dev. of Sample Means __________

Be very sure you understand what these numbers represent. If you are not sure, ask me!

(j) In the Population window, click on NORMAL to change the population to one of the following: Bimodal, Skew-, Skew+, Trimodal, U-Shaped, Uniform. Note that this may change the population mean μ and standard deviation σ as well. Please indicate below which population distribution you are using and also the population mean and standard deviation.

(k) Now change Sample Size to 1 and number of samples to 500 (you definitely want to make sure you have the F button pressed in your Samples window to speed up the animation!). Click the Draw Samples button. Record the following information:

- Describe the shape, center (the mean of the sample means), and spread (the standard deviation of the sample means) of the Sampling Distribution of the $\bar{x}$ values. In particular, how do the shape, center, and spread compare to the population distribution? You can click the purple population outline (upper left corner of Sampling Distribution window) for easier visual comparison.

- Now click the blue normal outline in the Sampling Distribution window. Which outline (population or normal) appears to be a better description of the sampling distribution of the sample mean $\bar{x}$ values?

(l) Change the sample size to 5 (keep number of samples at 500) and click the Draw Samples button. Give the information asked for in part (k).

(m) Change the sample size to 25, click the Draw Samples button, and answer the same questions from part (k).

(n) Change the sample size to 50, click the Draw Samples button, and answer the same questions from part (k).

(o) Complete the table below. Are the theoretical values predicted by the Central Limit Theorem (CLT, p. 281) close to the empirical values you got when you ran the simulations above?

Sample Size (n)	Population Mean	Empirical Mean of Sample Means	Theoretical Mean of Sample Means (via the CLT)	Population Standard Deviation	Empirical Standard Deviation of Sample Means	Theoretical Standard Dev. Of Sample Means (via the CLT)
1						
5						
25						
50						

(p) Repeat parts (j)-(o) for another non-normal population. Clearly indicate which population you use below and summarize your results in the table atop the next page.

Sample Size (n)	Population Mean	Empirical Mean of Sample Means	Theoretical Mean of Sample Means (via the CLT)	Population Standard Deviation	Empirical Standard Deviation of Sample Means	Theoretical Standard Dev. Of Sample Means (via the CLT)
1						
5						
25						
50						

(q) Briefly summarize your results in terms of the Central Limit Theorem.