P Chapter 7 Projects

Project A: Central Limit Theorem Experiment

You will need a standard six-sided die and at least six sets of data to complete this project.

Consider the distribution of the possible outcomes from rolling a single die; that is, 1, 2, 3, 4, 5, and 6. Let's use this distribution as our theoretical population distribution. We want to use this population distribution to explore the properties of the Central Limit Theorem. Let's begin by determining the shape, center, and dispersion of the population distribution.

1. What would you expect the distribution of the outcomes from repeated rolls of a single die to look like; in other words, what is its shape? (**Hint:** What is the probability of getting each value?)

   Shape: ____________

2. Calculate the mean of the population. (**Hint:** What is the mean outcome for rolling a single die?)

   $\mu =$ ________________

3. Calculate the standard deviation of the population. (**Hint:** What is the standard deviation of all possible outcomes from rolling a single die?)

   $\sigma =$ ________________

Let's continue by exploring the distribution of the original population empirically. To do so, follow these steps.

Step 1: Roll your die 60 times and record each outcome.

Step 2: Combine your results with at least two other students and tally the frequency of each roll of the die from the combined results. Record your results in a table similar to the following.

Outcome	Frequency
1	
2	
3	
4	
5	
6	

Step 3: Draw a bar graph of these frequencies.

Step 4: Does the distribution appear to be a normal distribution? Is this what you expected from question 1?

The Central Limit Theorem is not about individual rolls like we just looked at, but is about the averages of sample rolls. Thus we need to create samples in order to explore the properties of the Central Limit Theorem.

Step 5: Return to your original data from Step 1. To create samples from your data you can group the rolls into sets of 10. For each sequence of 10 rolls, calculate the mean of that sample. Round your answers to one decimal place. (You should have six sample means.)

Step 6: Combine your sample means with those of as many of your classmates as you can. Record the sample means of each of your classmates' six samples.

Step 7: Tally the frequencies of the sample means from your combined results in a table like the one that follows.

Sample Mean	Frequency
1.0–1.2	
1.3–1.5	
1.6–1.8	
1.9–2.1	
2.2–2.4	
2.5–2.7	
2.8–3.0	
3.1–3.3	
3.4–3.6	
3.7–3.9	
4.0–4.2	
4.3–4.5	
4.6–4.8	
4.9–5.1	
5.2–5.4	
5.5–5.7	
5.8–6.0	

Step 8: Draw a histogram of the sample means.

Step 9: What is the shape of this distribution?

Step 10: What is the mean of your sample means? (**Hint:** Use the sample means you collected in Step 6.)

$\mu_{\bar{x}}$ = __________

How does $\mu_{\bar{x}}$ compare to μ from question 2?

Step 11: What is the standard deviation of the sample means? (Again, go back to the sample means you collected in Step 6 and use a calculator or statistical software.)

$\sigma_{\bar{x}}$ = __________

How does $\sigma_{\bar{x}}$ compare to σ from question 3?

Since our samples were groups of 10 rolls, $n = 10$. Using σ from question 3, calculate $\frac{\sigma}{\sqrt{n}}$.

$\frac{\sigma}{\sqrt{n}}$ = __________

How does $\sigma_{\bar{x}}$ compare to $\frac{\sigma}{\sqrt{n}}$?

The Central Limit Theorem says that the distribution of the sample means should be closer to a normal distribution when the sample size becomes larger. To see this effect, group your original data from Step 1 into two samples of 30 rolls instead of six sets of 10.

Repeat Steps 5–11 using the new sample size of $n = 30$.

Step 12: Do your results seem to verify the Central Limit Theorem?

Project B: Sampling Distribution Simulation

In the Hawkes Learning courseware, *Beginning Statistics*, open Lesson 7.1, Introduction to the Central Limit Theorem. This lesson is a simulation designed to help you better understand sampling distributions as well as the Central Limit Theorem. Begin the simulation by choosing a parent distribution. Select the Settings icon at the top right of the screen, choose Distribution, then pick a distribution from the menu and press OK. If you do not have a preference, the computer will automatically begin with a uniform parent distribution. For each iteration, the computer randomly chooses 30 numbers from the parent distribution and displays them in the parent histogram. At the same time, the computer chooses samples of size 5, 15, and 30 from those same 30 numbers. The mean of each sample is calculated and then displayed in its respective histogram. Answer the following questions.

1. How many numbers are displayed in the parent histogram?
2. How many numbers are displayed in each of the sampling distribution histograms?

Click **Next** to obtain a second iteration.

3. After the second iteration, how many numbers are displayed in the parent histogram?
4. After the second iteration, how many numbers are displayed in each of the sampling distribution histograms?

Now let's see what happens after many iterations. Click **Auto** and let the simulation run until about 100 iterations have passed. Click the same button, which now says **Stop**, to stop the process.

5. How many numbers are displayed in the parent histogram? (This number will vary, depending on how many iterations have passed.)
6. How many numbers are displayed in each of the sampling distributions?
7. Which, if any, of the sampling distributions appear to have a normal shape?

Finally, click **Auto** again and allow the program to process at least 1500 iterations before clicking **Stop**.

8. Which of the sampling distributions appear to have a normal shape?
9. Compare the means of sample means listed in the table. Do these numbers behave as you would expect?
10. Compare the standard deviations of sample means listed in the table. Do these numbers behave as you would expect?

Take some time to explore the other parent distributions available. Although the shape of the parent histogram will differ with each parent used, the Central Limit Theorem will always remain constant.