
Homework for Lecture 12

The following homework assignment should be completed as an RMarkdown Document. Only show your working code and place comments/explanations in the “white space” to help me evaluate your work. Please submit both the knitted RMarkdown document (.doc) and the .Rmd file. You may discuss the assignment with your classmates, but your write-up and work should be done independently of your classmates. The use of AI is considered a violation of academic policy for this assignment. When in doubt, ask me what is and is not allowed.

For all of these questions, use the **cherry** dataset in the **openintro** library. Make sure you load the **openintro** library in one of your opening chunks of code. We will use the **volume** column for these problems and all sampling should be done **with replacement**.

1. Explore the Original Data

- (a) Visualize the **volume** variable with a histogram. Describe the shape of the original data. Does the distribution appear normal? Explain your answer in complete sentences.
- (b) Calculate the mean and standard deviation of the **volume** variable.

2. Sampling Distribution with Sample Size 2

Generate 1,000 sample means by repeatedly taking samples of size 2 from the **volume** variable with replacement and finding the mean of each sample.

- (a) What is the expected mean of this sampling distribution according to the Central Limit Theorem?
- (b) What is the expected standard deviation of this sampling distribution according to the Central Limit Theorem?
- (c) What is the actual mean of your 1,000 sample means?
- (d) What is the actual standard deviation of your 1,000 sample means?
- (e) How do the actual values compare to the expected values?
- (f) Visualize the sampling distribution with a histogram. Does it appear normal?

3. Sampling Distribution with Sample Size 5

Repeat the process from Question 2, but now generate 1,000 sample means using samples of size 5.

- (a) What is the expected mean and expected standard deviation of this sampling distribution?
- (b) What is the actual mean and actual standard deviation of your sampling distribution?
- (c) Visualize the sampling distribution with a histogram.
- (d) How does this sampling distribution compare to the sampling distribution when the sample size was 2?

4. **Sampling Distribution with Sample Size 10** Repeat the process again, but now generate 1,000 sample means using samples of size 10.

- (a) What is the expected mean and expected standard deviation of this sampling distribution?
- (b) What is the actual mean and actual standard deviation of your sampling distribution?
- (c) Visualize the sampling distribution with a histogram.
- (d) How does this sampling distribution compare to the previous two sampling distributions?

5. **Sampling Distribution with Sample Size 30**

Repeat the process one final time, but now generate 1,000 sample means using samples of size 30.

- (a) What is the expected mean and expected standard deviation of this sampling distribution?
- (b) What is the actual mean and actual standard deviation of your sampling distribution?
- (c) Visualize the sampling distribution with a histogram.
- (d) Does this sampling distribution appear approximately normal? Explain your answer.

6. **Final CLT Reflection**

Write a short paragraph answering the following questions:

- What happened to the shape of the sampling distribution as the sample size increased?
- What happened to the standard deviation of the sample means as the sample size increased?
- Did the mean of the sampling distribution stay close to the mean of the original volume variable?
- How does this homework demonstrate the Central Limit Theorem?