This lab is the first half of a pair of labs designed to help you review for the final exam. The final exam is heavily weighted toward inference, but the exam will contain some earlier concepts too.

Learning Objectives for the Final: (Click the learning objectives below for a short clip on the topic.)

Calculate the Minimum Percentage of Data Inside a Symmetric Interval (3)
Calculate the Expected Value of a Discrete Random Variable (3)
Calculate the Value Corresponding to an Upper Percentile of the Normal Distribution (3)
Calculate the Value Corresponding to a Lower Percentile of the Normal Distribution (3)
Discuss the Central Limit Theorem (2)
Determine if a Sample Size is Large Enough to Employ the Central Limit Theorem (2)
Apply the Central Limit Theorem to Calculating Probabilities for the Sample Mean (3)
Define Interval Estimators (1)
Explain the Meaning of "Confidence Level" (2)
Know How to Find Critical Z Values (3)
Form the Margin of Error when the Sample Size is Large (3)
Construct a Confidence Interval for the Mean when the Sample Size is Large (3)
Interpret a Confidence Interval (2)
Discuss Factors Affecting the Margin of Error or Width of Confidence Intervals (2)
Explain the Special Rounding Rule for Sample Size Calculations (2)
Determine the Sample Size Needed to Estimate the Population Mean (3)
List the Similarities of the t Distribution and the Z Distribution (1)
List the Differences between the t Distribution and the Z Distribution (1)
Find Critical t Values (3)
Form the Margin of Error when the Sample Size is Small (3)
Construct a Confidence Interval when the Sample Size is Small (3)
Know the Meaning of "Hypothesis" in the Context of Statistics (1)
List the Characteristics of the Null Hypothesis (1)
List the Characteristics of the Alternative Hypothesis (1)
Determine the Null and Alternative Hypothesis (2)
Explain the Logic of the Test Statistic (2)
Calculate the Test Statistics from Sample Data (3)
Discuss the Four Possible Outcomes for a Hypothesis Test (2)
Describe a Type I Error (2)
Describe a Type II Error (2)
Classify an Error as Either Type I or Type II (2)
Discuss the Probability of Committing a Type I Error for a One-Tailed Test (2)
Discuss the Probability of Committing a Type I Error for a Two-Tailed Test (2)
Define the Critical Value (1)
Find the Critical Value for a One-Tailed Test when n is Large (3)
Find the Critical Values for a Two-Tailed Test when n is Large (3)
Determine a Decision Rule Based on Critical Values (2)
Use the Classical Approach of Hypothesis Testing to Determine the Initial Conclusion (3)

Exercises:

1. The amount of calories people consume when eating a large bag of popcorn while watching a movie follows a normal distribution. The average number of calories consumed is 1250, and the standard deviation for the number of calories consumed is 210. Find the amount of popcorn calories consumed by the bottom 13.5% of consumers belonging to this population. We need to look up the area between the unknown z score and the mean, and that area is equal to 0.5000 - 0.1350 = 0.3650, The z score associated with that area is z = -1.10. Then, we plug that z score into the formula $x = z\sigma + \mu = -1.10(210) + 1250 = 1019$. They consume 1,019 or less calories of popcorn.

2. Use the following data to form a margin of error for a 98% confidence interval to estimate the population mean:
   $n = 50$
   $\bar{x} = 8.5$ $E = 2.326 * \frac{0.9}{\sqrt{50}} \approx 0.296$
   $s = 0.9$

3. Researchers want to independently confirm the numbers reported by Amazon for its Prime members. Amazon claims that the mean amount of money spent each year by Prime members is more than $3,000. A random sample of 49 Amazon Prime members is selected. The average amount of money spent by these Amazon Prime members last year is $2,584. The sample standard deviation for these amounts is $1490. Use the sample of data to construct a 95% confidence interval to estimate the true mean amount of money that Amazon Prime members spent last year. Does the interval contradict Amazon's claim?
   $\left(2584-1.96\frac{1490}{\sqrt{49}},2584+1.96\frac{1490}{\sqrt{49}}\right)=(\$2,166.8,\$3,001.2)$ We are 95% confident the true mean is between $\$2,166.80 \text{ and } \$3,001.20$. Since the upper limit of the interval is $1.20 higher than $3,000.00, we cannot rule out the possibility that the average annual spending for Prime members is higher than $3,000, so the interval does not contradict Amazon's claim.

4. A hypothesis test of the claim that "18 ounce" boxes of Raisin Bran contain less than 18 oz of cereal on average was conducted using a sample of only 25 boxes of cereal. The results for the test turned out to be:

   H0: $\mu \geq 18.0$. H1: $\mu < 18.0$. Test statistic: t = -2.34. Critical value: t = -2.492. Do not reject H0. There is not sufficient evidence to support the claim that the mean is less than 18.0.

   If an attorney argues the results are flawed because the sample size was too small, is his argument in anyway valid here?

   Yes, there is some validity to the attorney's claim because the small sample size makes it harder to reject false null hypotheses. Since the null was not rejected here, it is possible that the small sample size contributed to the outcome of the test.

5. The average shower taken in the USA uses 17.2 gallons of water. The standard deviation for these measurements is 6.6 gallons. What is the probability that a random sample of 36 monitored showers in the USA have an average water usage that is greater than 19 gallons of water? First find the standard error of the sample means: $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{6.6}{\sqrt{36}}\approx 1.1$, Then convert 19 into a z score:
   $z=\frac{(19-17.2)}{1.1}\approx 1.64$, 0.5000 - 0.4495 = 0.0505

6. Use the following data to form a margin of error for a 95% confidence interval to estimate the population mean:

   $n = 25$

   $\bar{x} = 32.0$ $\quad E = 2.064 * \frac{1.2}{\sqrt{25}} \approx 0.495$

   $s = 1.2$

7. An organic farmer claims that the average weight of his apples is greater than 85 grams. It is safe to assume these apple weights follow a normal distribution. A random sample of 20 of the farmer's apples has a mean weight of 87.2 grams and a standard deviation of 5.1 grams. Test the farmer's claim at the 5% significance level.

   Claim: $\mu > 85$

   $H_0 : \mu \le 85$

   $H_A : \mu > 85$

   Test stat:

   $$t = \frac{(87.2 - 85)}{\frac{5.1}{\sqrt{20}}} \approx 1.929$$

   Critical Value: 1.729 Rejection region: reject if the test stat is more than 1.729.
   Reject the null hypothesis, support the alternative hypothesis.
   The sample data support the claim that the mean is greater than 85 grams.

8. The owner of a coffee chain claims that a cup of their coffee has an average of 150 mg of caffeine or less. A study of 32 randomly selected cups of their coffee was conducted. The mean caffeine amount for the sample was 169.0 mg, and the standard deviation was 26.2 mg. Use a 2% significance level to test the claim.

   Claim: $\mu \le 150$

   $H_0 : \mu \le 150$

   $H_A : \mu > 150$

   Test stat:

   $$z = \frac{(169.0 - 150)}{\frac{26.2}{\sqrt{32}}} \approx 4.10$$

   Critical Values: 2.05 Rejection Region: reject if the test stat is more than 2.05.
   Reject the null, support the alternative
   The sample data allow us to reject the claim that the mean is less than or equal to 150.

9. According to a report by Common Sense Media, teens are spending an average of 9 hours per day in front of a screen for entertainment. The standard deviation for the time teens spend in front of a screen for entertainment is 2.2 hours. What is the minimum percentage of teens that spend between 4 and 14 hours in front of a screen for entertainment each day? Using Chebyshev's theorem: at least 80.64%

10. The head of a bank that provides auto loans claims that US households have an average amount of auto loan debt equal to $29,000. A random sample of 49 US households had an average amount of auto loan debt of $27,802 and a standard deviation of $8,250. Calculate the test statistic and the p-value for a test of this hypothesis.

$$z = \frac{27{,}802 - 29{,}000}{\frac{8{,}250}{\sqrt{49}}} \approx -1.02$$ p-value: (find twice the tail area beyond -1.02)

0.5000-0.3461 = 0.1539 then multiply by two: **0.3078**

11. A 95% confidence interval for the true mean commute time in South Florida (Palm Beach, Broward, and Miami) was constructed. Interpret the interval: 33 minutes ± 3 minutes. Based on this interval, does it appear that the true mean commute time in South Florida is shorter than 35 minutes?
Since the interval ranges from 30 min to 36 min, we cannot use this data to support the claim that the mean commute time is less than 35 minutes. The mean should be between 30 and 36 minutes according to the interval.

12. Consumer Reports claims the average age of cars on the road today is 11.2 years. State the null and alternative hypotheses for a test of Consumer Reports' claim at the 10% significance level. What is the type I error for this test, and what is the probability that the type I error occurs?
$H_0 : \mu = 11.2$
$H_A : \mu \neq 11.2$
Type I error: the error of rejecting the claim that the mean is 11.2 years when it is in fact true that the mean is 11.2 years.
There is a 10% chance of committing the type I error in this problem.

13. True or false: The standard error gives us a measure of how the estimator varies from sample to sample.

14. An organic farmer claims that the average weight of his apples is greater than 85 grams. If a random sample of 100 apples has a mean weight of 86.1 grams and a standard deviation of 5 grams, what is the critical value and rejection region for a test of the farmer's claim at the 1% significance level?
C.V. = 2.326
Rejection region: reject the null if z > 2.326

15. McDonald's claims that the average time customers wait in the drive through is under 2.5 minutes. State the null and alternative hypotheses for a test of McDonald's claim at the 5% significance level. What is the type I error for this test, and what is the probability that the type I error occurs?
$H_0 : \mu \geq 2.5$
$H_A : \mu < 2.5$
Type I error: the error of rejecting the hypothesis that the mean is greater than or equal to 2.5 minutes when it is in fact true that the mean is greater than or equal to 2.5 minutes.
There is an **at most** 5% chance of committing the type I error in this problem.

16. A CEO wants to estimate the true average amount of time she spends reading emails each day. She installs tracking software that records the amount of time she spends looking at messages in her inbox. Her secretary randomly selects 28 of these times from the log and finds that the sample mean amount of time is 180.1 minutes. The standard deviation for these times is 18.1 minutes. Use the data to construct a 90% confidence interval for the true mean amount of time the CEO spends looking at her email each day.
$\left(180.1-1.703\frac{18.1}{\sqrt{28}},180.1+1.703\frac{18.1}{\sqrt{28}}\right)=\left(174.27,185.93\right)$ We are 90% confident the true mean is between174.27 minutes and 185.93 minutes.

17. A pediatrician claims the average weight of newborns is more than 7 pounds. State the null and alternative hypothesis that would be used to test this pediatrician's claim.
$H_0 : \mu \leq 7$
$H_A : \mu > 7$

18. Which population exhibits a greater amount of variation (dispersion) among its values?
a. Population 1 has a mean of 69.1 feet and a standard deviation of 3.2 feet
b. Population 2 has a mean of 69.1 feet and a standard deviation of 5.1 feet