What do I know already? (Do not look up the answers to these questions. The purpose is to assess your current level of knowledge on these topics.)

A. What is a 95% confidence interval for the population mean designed to accomplish? What does it mean to say the confidence level is 95%? A 95% confidence interval is designed to capture (contain) the population mean. A 95% confidence level indicates that, when repeated samples of the same size are taken a large number of times from the population, 95% of the intervals calculated from those samples will contain the population mean.

B. What attributes do the z and t distributions share? Both distributions are bell-shaped, and both distributions have a mean of zero.

C. In what ways are the z and t distributions different? The z distribution is a single case of the normal distribution. It has a standard deviation of 1. The t distribution is a family of curves. The standard deviation for the different t curves is always greater than 1. This means there is more area in the tails of the t curves, and less area clustered around the mean. However, as n approaches infinity, the t distribution approaches the z distribution.

D. A scholarly journal randomly samples 16 papers that have been submitted and approved for publication. The journal records the amount of time it took them to reach the final conclusion to publish each of the successful papers. The data is used to create the following 98% confidence interval which estimates the true mean time to reach a positive decision. The interval is: 10 days to 30 days. The journal is not happy with the resulting interval because it is too wide. In other words, the margin of error is too large. If the journal decides to collect new data and try again, what steps can they take to ensure the interval has a smaller margin of error this time? They could increase the sample size, and/or they could reduce the confidence level.

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

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)

Exercises:

1. Researchers want to independently confirm the numbers reported by Amazon for its Prime members. A random sample of 36 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.
   $\left(2584-1.96\frac{1490}{\sqrt{36}}, 2584+1.96\frac{1490}{\sqrt{36}}\right)=(\$2,097.3, \$3,070.7)$ We are 95% confident the true mean is between $\$2,097.3$ and $\$3,070.7$.

2. True or false: If the random variable X has a normal distribution, the distribution of the sample means for samples of size 7 is normal. Since X has a normal distribution, at any sample size the sample means have a normal distribution.

3. A 95% confidence interval for the true mean commute time in Los Angeles (LA) was constructed. Interpret the interval: 29 minutes ± 2 minutes. Based on this interval, does it appear that the true mean commute time in LA is longer than 35 minutes? Since the interval ranges from 27 min to 31 min, we believe the mean commute time is less than 35 minutes because it should be between 27 and 31 minutes.

4. In an effort to determine if the Statistics lab is effective at improving final exam scores, the department selected a random sample of 49 final exam scores for students who attended the lab. These scores had an average of 72.1 points and a standard deviation of 14.9 points. Use the sample of exam scores to form a 90% confidence interval for the true mean final exam score for students attending the lab.
$\left(72.1-1.645\frac{14.9}{\sqrt{49}}, 72.1+1.645\frac{14.9}{\sqrt{49}}\right)=(68.60, 75.60)$ We are 90% confident the true mean is between 68.60 and 75.60 points.

5. Use the following data to form a margin of error for a 99% confidence interval to estimate the population mean:
$n=60$
$\bar{x}=5.5$ $E=2.576*\frac{0.9}{\sqrt{60}}\approx 0.299$
$s=0.9$

6. The average height for Dutch males is 71.9 inches. The standard deviation for their heights is 2.6 inches. Assuming that these heights have a bell-shaped distribution, determine the height separating the tallest 2.5% of Dutch men from the rest. $x=1.96*2.6+71.9=76.996$

7. Researchers have randomly selected 120 commuters in Los Angeles (LA) in an effort to determine the average length of the commute in this famous American city. Find the appropriate critical value that would be used to construct a 98% confidence interval for the true mean length of commute in LA. $z_{\alpha/2}=2.326$

8. 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 less than 19 gallons of water? $Z=\frac{(19-17.2)}{\frac{6.6}{\sqrt{36}}}\approx 1.64$, 0.4495+0.5000 = 0.9495

9. Use the following data to form a margin of error for a 98% confidence interval to estimate the population mean:
$n=20$
$\bar{x}=28.5$ $E=2.539*\frac{1.2}{\sqrt{20}}\approx 0.681$
$s=1.2$

10. The average weight of women in the United States is 164.7 pounds. The standard deviation for these weights is 37.5. Researchers plan to select 25 women randomly from the population for a study, and the sample mean weight will be calculated for the group of 25 women. If all of the possible random samples of 25 women were taken from the population, what would the mean be for all of those sample means? 164.7 pounds

11. The average height of women in the United States is 63.6 inches. The standard deviation for these heights is 2.5 inches. Researchers plan to select 100 women randomly from the population for a study. The sample mean height will be calculated for the 100 women. If all of the possible random samples of 100 women were taken from the population, what would the standard error be for the resulting set of sample means? $\frac{2.5}{\sqrt{100}} = 0.25$

12. True or false: If the random variable X has a right-skewed distribution, the distribution of the sample means for samples of size 10 can be reasonably assumed to be approximately normal. The sample size is too small for us to be reasonably sure the sample means are normal.

13. An FIU student wants to estimate the true time it takes to have his order completed at Tropical Smoothie on campus. Assuming that the standard deviation for these times is 2.5 minutes and that he wants to estimate the true average time to complete his order to within 1.0 minutes using a 99% confidence interval, how many order-completion times should he get for his random sample?
$n = \left[\frac{2.576 * 2.5}{1.0}\right]^2 = 41.4736$ **round up to 42**

14. True or false: If the random variable X has a right-skewed distribution, the distribution of the sample means for samples of size 36 can be reasonably assumed to be approximately normal. The sample size is large enough to assume normality using the central limit theorem.

15. Scientists would like to estimate the true mean waist circumference (μ) for adult males living in the US. The sample mean waist circumference $(\overline{x})$ for a random sample of 50 adult males from the US is an unbiased estimator of the true mean waist circumference for adult males living in the US. What is the expected (or average) value for the set of sample means described here? μ

16. Finding a parking spot on campus at FIU can be a challenge. The standard deviation for the time it takes to find a spot on campus is 7.1 minutes. If you want to create a 95% confidence interval estimate of the true mean time it takes to find a parking spot at FIU to within 2 minutes, on how many occasions should you randomly time your search for a parking spot?
$n = \left[\frac{1.960 * 7.1}{2}\right]^2 \approx 48.4138$ **round up to 49**

17. Environmentalists have randomly selected 12 rainfall volume measurements from various spots in Broward County Florida during a tropical storm. Find the appropriate critical value that would be used to construct a 95% confidence interval for the true mean amount of rainfall in Broward during this storm.
$t_{\alpha/2} = 2.201$

18. True or false: If the standard deviation for STA 2122 final exam scores is 20 $(\sigma = 20)$, the standard error $(\sigma_{\bar{x}})$ for the sample mean final exam score derived from samples of 25 people will also be equal to 20. $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{25}} = 4$

19. A 95% confidence interval for the true mean commute time in South Florida (Palm Beach, Broward, and Miami) was constructed. Interpret the interval: 28 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 25 min to 31 min, we believe the mean commute time is less than 35 minutes because it should be between 25 and 31 minutes.

20. 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 22 of these times from the log and finds that the sample mean amount of time is 179.1 minutes. The standard deviation for these times is 18.2 minutes. Use the data to construct a 95% confidence interval for the true mean amount of time the CEO spends looking at her email each day.
$\left(179.1 - 2.08\frac{18.2}{\sqrt{22}}, 179.1 + 2.08\frac{18.2}{\sqrt{22}}\right) = (171.03, 187.17)$ We are 95% confident the true mean is between171.03 minutes and 187.17 minutes.