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. Name the attributes that are typically discussed when describing a sampling distribution. The mean (or location) of the sampling distribution, the standard error (or spread) of the sampling distribution, and the shape of the sampling distribution are usually discussed.

B. In your own words, describe the information conveyed by the standard error of an estimator. The standard error gives us a measure of how the estimator varies from sample to sample.

C. What does it mean to say an estimator is unbiased? Unbiased estimators cluster around the target parameter on the number line so that the average value of the estimator is equal to the target parameter.

D. In your own words, explain what the central limit theorem tells us about the sample mean. The central limit theorem tells us that for random variables that do not follow a normal distribution, the distribution of the sample means is approximately normal when the sample size is sufficiently large.

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

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)
Define Point Estimators (1)
Define Sampling Distributions (1)
Describe the Standard Error of an Estimator (2)
Discuss the Desired Traits of a Point Estimator (2)
Know the Mean of the Distribution of Sample Means (1)
Know the Standard Deviation (Error) of the Distribution of Sample Means (2)
Compare the Variation of the Sample Means to the Variation of the Random Variable (2)
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)

Exercises:

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

2. The average height for Dutch males is 71.9 inches (That's almost 6 feet tall!). The standard deviation for their heights is 2.6 inches. Assuming that these heights have a bell-shaped distribution, what is the probability that a randomly selected Dutch man is between 68 and 70 inches tall?
   $z = \frac{(68-71.9)}{2.6} = -1.50$, $z = \frac{(70-71.9)}{2.6} \approx -0.73$, 0.4332 - 0.2673 = 0.1659

3. True or false: If the standard deviation for IQ scores is 15$(\sigma = 15)$, it is a certainty that the standard error$(\sigma_{\bar{x}})$ for the sample mean IQ scores derived from samples of five people will be more than 15. In other words, $\sigma_{\bar{x}} > 15$. For any n > 1, the standard error of the mean is always less than the standard deviation for the random variable.

4. The average shower taken in the USA uses 17.2 gallons of water. The standard deviation for these measurements is 6.6 gallons. It is safe to assume these water usage amounts are normally distributed. What is the probability that a randomly monitored shower in the USA uses between 10 and 17.2 gallons of water?
   $z = \frac{(10-17.2)}{6.6} \approx -1.09$, 0.3621

5. 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 33% 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 - 0.3300 = 0.1700, The z score associated with that area is z = -0.44. Then, we plug that z score into the formula $x = z\sigma + \mu = -0.44(210) + 1250 = 1157.6$. They consume 1,157.6 or less calories of popcorn.

6. True or false: A point estimate is an interval on the number that contains the point (value) we are trying to estimate. A point estimate is a single value calculated from sample data that is used to estimate an unknown parameter.

7. A large study tracked the amount of coffee consumed in one day by coffee drinkers. The data seems to follow a normal distribution. The average number of ounces consumed in one day by the coffee drinkers was 27.9 ounces, and the standard deviation for these values was 7.5 ounces. Assuming the results of this study are representative of the coffee drinking population, what is the probability that one randomly selected coffee drinker consumes more than 36 ounces of coffee in one day?
$z = \frac{(36 - 27.9)}{7.5} = 1.08$, 0.5000 - 0.3599 = **0.1401**

8. The average height for Dutch males is 71.9 inches (That's almost 6 feet tall!). The standard deviation for their heights is 2.6 inches. Assuming that these heights have a bell-shaped distribution, determine the height separating the shortest 15% of Dutch men from the rest. We need to look up the area between the unknown z score and the mean, and that area is equal to 0.5000 - 0.1500 = 0.3500, The z score associated with that area is z = -1.04. Then, we plug that z score into the formula $x = z\sigma + \mu = -1.04(2.6) + 71.9 = 69.196$.

9. True or false: A sampling distribution is the probability distribution for a given sample statistic.

10. 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 40 monitored showers in the USA have an average water usage that is greater than 18 gallons of water? First find the standard error of the sample means: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6.6}{\sqrt{40}} \approx 1.04355$, Then convert 18 into a z score: $z = \frac{(18-17.2)}{1.04355} \approx 0.77$, 0.5000 - 0.2794 = 0.2206

11. The average shower taken in the USA uses 17.2 gallons of water. The standard deviation for these water usage measurements is 6.6 gallons. Assuming these water usage amounts are normally distributed, determine the amount of water used by the top 10% of US showers in terms of water usage. We need to look up the area between the unknown z score and the mean, and that area is equal to 0.5000 - 0.1000 = 0.4000, The z score associated with that area is z = 1.28. Then, we plug that z score into the formula $x = z\sigma + \mu = 1.28(6.6) + 17.2 = 25.648$. The top ten percent use at least 25.6 gallons of water per shower.

12. Which of the following point estimators would be the best choice to estimate θ?
A. Estimator A has a variance of 9.1 and $E(A) = \theta + 1.2$
B. Estimator B has a variance of 8.2 and $E(B) = \theta$
C. Estimator C has a variance of 7.6 and $E(C) = \theta$ (smallest variance and it is unbiased)
D. Estimator D has a variance of 7.6 and $E(D) = \theta - 2.8$

13. 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 20 women randomly from the population for a study. The sample mean weight will be calculated for the 20 women. If all of the possible random samples of 20 women were taken from the population, what would the mean be for all of those sample means? $\mu_{\bar{x}} = 164.7$

14. A study in 2014 indicated that the average height for women in Guatemala was 58.7 inches (a little under 4ft 11in). The standard deviation for these heights is 1.9 inches. These heights are normally distributed. Find the height separating the tallest 1% of Guatemalan women from the rest. We need to look up the area between the unknown z score and the mean, and that area is equal to 0.5000 - 0.0100 = 0.4900, The z score associated with that area is z = 2.33. Then, we plug that z score into the formula $x = z\sigma + \mu = 2.33(1.9) + 58.7 = 63.127$.

15. 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 20 women randomly from the population for a study. The sample mean weight will be calculated for the 20 women. If all of the possible random samples of 20 women were taken from the population, what would the standard error be for the set of sample means? $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{37.5}{\sqrt{20}} \approx 8.385$

16. True or false: The random variable X has a skewed distribution. The distribution of the sample means for samples of size 5 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.

17. True or false: The random variable X has a left-skewed distribution. The distribution of the sample means for samples of size 49 can be reasonably assumed to be approximately normal. The sample size is large enough to assume normality using the central limit theorem.

18. Find $P(Z > 2.33)$ = 0.5000 - 0.4901 = 0.0099

19. 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 probability that a sample of four Dutch men have an average height that is under 70 inches. First find the standard error of the sample means: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.6}{\sqrt{4}} = 1.3$, Then convert 70 into a z score: $z = \frac{(70 - 71.9)}{1.3} \approx -1.46$, 0.5000 - 0.4279 **= 0.0721**

20. Statisticians would like to estimate the value of a parameter represented by the symbol Ω. They have found an unbiased estimator for Ω. What is the expected value of this estimator? Ω, because unbiased estimators have an expected value that is equal to the parameter they are estimating.