SOCY 392 Exam III Name_____ Date_____

STOP. THINK. GET ORGANIZED.

- You have an hour to complete this exam.
- The first 14 questions are multiple choice (4 points each).
- The last 2 questions are short answer (15 points each).

---

1. What is the difference between a sample distribution and a sampling distribution?
    a. A sample distribution contains the true population parameter; a sampling distr. does not.
    b. A sample distribution is a normal distribution; a sampling distribution is skewed
    c. A sample distribution represents one sample drawn from a population; a sampling distribution represents many samples drawn from a population.
    d. A sample distribution contains more data than a sampling distribution
    e. None of these

2. Which of the following questions would *require* you to use a method of statistical inference to answer?
    a. Calculating the average age of a sample of patients in the ER
    b. Finding the proportion of shoes in your closet that are sneakers
    c. Determining the exact number of Hot Cheetos you ate last month
    d. Asking whether UofSC yearly football's record is significantly better or worse than Clemson's
    e. None of the above

3. An alpha level of .05 corresponds to…
    a. A p-value of .95
    b. A cutoff that corresponds to a 5% chance of incorrectly rejecting the null hypothesis
    c. A cutoff that corresponds to a 5% chance of incorrectly failing to reject the null hypothesis
    d. The point at which you should use a t-test instead of a z-test
    e. None of the above

4. What is the measure of spread **for a sampling distribution** that is calculated when we perform hypothesis tests, and used when we construct 95% confidence intervals around estimates?
    a. The standard error
    b. The population standard deviation
    c. The sample standard deviation
    d. The variance
    e. None of the above

Scenario A (for questions 5 to 8): A researcher wants to learn how far people in Columbia live from the nearest grocery store by foot. They randomly sample 100 households and use Google Maps to find how many minutes it would take to walk from each house to the nearest grocery store. They find that the average home in their sample is a 27-minute walk to the nearest grocery store, with a sample standard deviation of 9 minutes.

5. The researcher wants to tell people how long it takes for the average person in Columbia to get to a grocery store without being overly precise. Based on the information above, they could do so by…
    a. Constructing a 95% confidence interval around their sample mean
    b. Performing a one-sample t-test
    c. Performing a one-sample z-test
    d. Asking people in Charleston the same question and estimating another mean
    e. None of the above

6. The researcher heard on the radio that the true average walking commute time to a grocery store for the population of households in Columbia is 31 minutes. What method would you recommend to the researcher to test whether these numbers are significantly different from each other?
    a. Use a one-sample t-test to compare their sample mean of 27 minutes to the null hypothesis of 31 minutes.
    b. Use a one-sample z-test to compare their sample mean of 27 minutes to the null hypothesis of 31 minutes.
    c. Use a two-sample t-test to compare their sample mean of 27 minutes to the radio-reported sample mean of 31 minutes.
    d. Use a two-sample z-test to compare their sample mean of 27 minutes to the radio-reported sample mean of 31 minutes.
    e. None of the above.

7. Let's say that the researcher performs a two-tailed hypothesis test, looks up their critical t or z value, and computes a p-value of .043. This means…
    a. They fail to reject the null hypothesis at an alpha level of .05
    b. They reject the null hypothesis at an alpha level of .05
    c. They accept the null hypothesis at an alpha level of .05
    d. They can accept the null hypothesis with 95% confidence
    e. None of the above.

8. The researcher instead opts to compute a 95% confidence interval around his sample mean – their best estimate for the true population parameter – and finds that this interval ranges from 25.2 minutes to 28.8 minutes. What might the researcher conclude, given that the radio claimed the true average time was 31 minutes?
    a. That they may have not actually gotten a random sample
    b. That the radio's claim might be wrong
    c. That the confidence interval does not contain the estimate
    d. Both A and B
    e. None of the above.

9. What does it mean to "reject the null hypothesis at an alpha level of .05" when conducting a two-tailed one-sample hypothesis test?
    a. It means you have compelling evidence that the mean for the group/population you sampled from is probably significantly different than the null hypothesis value.
    b. It means you can say with certainty that your evidence that your sample mean is significantly different from the population mean.
    c. It means you can say with certainty that your evidence that your sample mean is significantly different from the null hypothesis.
    d. It means you have compelling evidence that the mean for the group/population from which your sample was drawn is the same as the null hypothesis value.
    e. None of the above.

10. What is another way of interpreting the claim from #9?
    a. The probability of getting the sample mean you observed under the null hypothesis is < .05
    b. The probability of getting a sample mean as low or lower than the one you observed under the null hypothesis is < .05
    c. The probability of getting a sample mean as high or higher than the one you observed under the null hypothesis is < .05
    d. The probability of getting a sample mean as extreme or more extreme than the one you observed under the null hypothesis is < .05
    e. None of the above.

Scenario B (for questions 11-14): You work for the CDC – they deploy you to a small town that's concerned it has a public health problem. Some citizens have been reporting that they began to experience tachycardia (a persistent resting heart rate of > 100 beats per minute, or BPM) after drinking from public water fountains. You can't be sure – maybe this is paranoia. But you want to see if this town's mean resting heart rate is higher than the typical heart rate in the US. If not, maybe this is just small-town paranoia. The average resting heart rate in the US is 80 BPM, with a standard deviation of 10 BPM. You use county records to contact 100 random people and meet with them to obtain their resting heart rate. You find that your sample mean for the town is 82.3 BPM, with a sample standard deviation of 11.1. You want to determine if your estimate for mean heart rate of people in this town is *significantly greater than* the US mean resting heart rate. So you conduct a one-sided t-test comparing your sample estimate for the town to the population mean for the US.

11. Using a t-table, what is the critical t-value associated with an alpha level of .05 for this hypothesis test?
    a. 1.960
    b. 1.962
    c. 1.984
    d. 1.660
    e. None of the above

12. You then perform this hypothesis test. What critical value for t do you get?
    a. 2.3
    b. 2.4
    c. 1.9
    d. 0.5
    e. None of the above

13. Interpret the results of your test. Is there a reason to suspect that the people in the town may indeed have a health concern on their hands?
    a. Yes – the probability of observing a sample mean of 82.3 BPM or greater if it came from a population with a true mean of 80 BPM is < 5%
    b. No – the probability of observing a sample mean of 82.3 BPM or greater if it came from a population with a true mean of 80 BPM is > 5%
    c. Yes – the probability of observing a sample mean of 82.3 BPM or greater if it came from a population with a true mean of 80 BPM is < 95%
    d. No – the probability of observing a sample mean of 82.3 BPM or greater if it came from a population with a true mean of 80 BPM is > 95%
    e. We cannot tell because we have not constructed a confidence interval.

14. You recruit another member of the CDC to replicate your study. They obtain a sample of 1,000 people from the down and get a mean of 81.7. Do they get the same result for their hypothesis test?
    a. Yes
    b. No
    c. The question does not provide enough information for me to determine the answer.
    d. None of the above.

Refer to the following scenario to answer questions 15-18. Pick 2 of the three questions and answer them completely and comprehensively. **SHOW ALL THE MATH/STEPS NECESSARY TO ANSWER EACH QUESTION.**

15. There are 450 players in the NBA. 312 of these players shot 75% or higher from the free-throw line. Patrick randomly sampled 200 players and found that the proportion of players in his sample who made 75% of their shots or more from the free-throw line was .72.

    a. What's the probability that of obtaining a proportion as extreme or more extreme than the sample proportion that Patrick found, given the known population proportion?

    b. Construct a 95% confidence interval around Patrick's sample proportion. Then explain what that confidence interval means with respect to the population parameter.

16. You randomly sample 100 UofSC students to find the mean number of drinks they consumed on Halloween night, and found its value of 4.1 with a standard deviation of 1.5. You do the same thing at Clemson with a sample of 110 students, and find a sample mean of 4.9 with a standard deviation of 1.7.

    a. Do you have evidence to suggest that the average number of drinks consumed at UofSC and Clemson by Halloween-party-people are significantly different? Conduct and interpret a hypothesis test to convince us of your answer.

    b. Now compute a confidence interval around each of the mean estimates. Explain how the result is consistent with, or differs from, your answer to A.

17. What's the difference between a "two-sided/two-tailed test" and a "two-sample test?" Is it possible to do a two-sided one-sample test? If no, explain why. If yes, provide an example, and explain why your example represents a two-sided one-sample test.