Stat 13 midterm, Prof. Rick Schoenberg, 8/31/23, 10am-11:20am.

1. You may use a calculator, a pencil, and any books and notes you want during the exam, but no internet searching or communicating or use of the computer other than to submit your answers.

2. Final numerical answers have been rounded to 3 significant digits.

3. There are 20 multiple choice questions worth 5 points each.

4. No partial credit is given for multiple choice questions. Choose ONE answer only.

5. Use the exam that matches the FIRST LETTER of your EMAIL ADDRESS!!!!!!!!!!
If THE FIRST LETTER OF YOUR EMAIL ADDRESS is A through J, then use the exam called midA.pdf .
If THE FIRST LETTER OF YOUR EMAIL ADDRESS is K through T, then use the exam called midK.pdf .
If THE FIRST LETTER OF YOUR EMAIL ADDRESS is U through Z, or a number or any other symbol, then use the exam called midU.pdf .

6. You have from 10am to 11:20am. By 11:20am you must EMAIL me your answers, to frederic@stat.UCLA.edu . Your email should just contain your answers, like
ADDBC CDAAB BBCCD DAABC. You do not need to show work. Make sure your answers are exactly in the correct order!

7. You must zoom in to the usual zoom while taking the exam. If you have a question during the exam, ask it using chat.

1. Does the study on infant nightlights and nearsightedness provide strong evidence to conclude that nightlights cause nearsightedness?
a. Yes, because it was a randomized controlled experiment.
b. No, because parental nearsightedness is a plausible confounding factor.
c. Yes, because the study showed that nightlights are too bright for infants and cause damage to the retina, resulting in higher rates of nearsightedness later.
d. No, because the nightlight group had a higher percentage of males than the group that slept in darkness.
e. Yes, because the study had a large sample size.

For the next four problems, consider the following. About 92.2% of American children are immunized for measles, according to the CDC. A scientist wants to see if the frequency of immunization among Americans with college educated parents is different from the overall percentage of 92.2%. She takes a simple random sample of 800 subjects with college educated parents, and finds that 760 of them have been immunized for measles.

2. Which of the following is true?
a. A **parameter** of interest is the percentage of Americans with college educated parents who are immunized for measles, and the **random variable** used to estimate it is the percentage in the sample who are immunized for measles.
b. A **parameter** of interest is the percentage in the sample who are immunized for measles, and it is used to estimate the percentage of Americans who are immunized for measles, which is a **random variable.**
c. A **parameter** of interest is the percentage of immunized Americans who have college educated parents, and the **random variable** used to estimate it is the sample size of 800.
d. A **parameter** of interest is the central limit theorem, and the **random variable** used to estimate it is the conditional proportion of the t-test confidence interval.
e. None of the above.

3. In order to test whether percent immunized depends on whether one's parents are college educated, what would the **null hypothesis** be?
a. The percentage of Americans with college educated parents who are immunized for measles is not 92.2%.
b. 92.2% of Americans have college educated parents.
c. 92.2% of Americans with college educated parents are immunized for measles.
d. 92.2% of Americans immunized for measles have college educated parents.
e. None of the above.

4. Under this null hypothesis, what is the **standardized** Z statistic, summarizing the difference between the sample percentage who are immunized for measles and the population percentage who are immunized for measles? Hint: use the value of π under the null hypothesis to compute the SE.
a. 2.01.
b. 2.80.
c. 2.95.
d. 3.50.
e. None of the above.

5. Which of the following can be concluded?
a. The difference between the immunization percentage in the sample and the immunization percentage for all Americans is statistically significant.
b. The immunization percentage for Americans with college educated parents is 92.2%.
c. The immunization percentage for Americans with college educated parents is 95.0%.
d. Immunization rates do not appear to vary significantly with parental education.
e. None of the above.

6. A survey of a simple random sample of 1000 American adults ages 50-75 finds that those who have no access to the internet have a 30% higher risk of heart attack than those who do not. The study's authors conclude that their data suggests that surfing the internet may be beneficial for the heart. What is the main problem with this conclusion?
a. Those with no access to the internet are probably poorer and thus eat a less healthy diet on average than those who have access to the internet, and diet is related to heart attack risk.
b. The explanatory confidence interval is greater than the significance level.
c. These 1000 Americans are probably much taller than the overall population, and therefore do not adequately represent the overall population of Americans.
d. A confounding factor is genetics, because genes linked to heart disease might make one more likely to want to surf the internet.
e. None of the above.

7. The Physician's Health Study I studied aspirin's effect on reducing the risk of heart attacks. Which of the following was **not** a reason for randomly assigning people to treatment or control in this experiment?
a. To ensure that the sample is more representative of the overall population.
b. To ensure that the treatment and control groups are similar with respect to known potential confounders such as diet and exercise.
c. To ensure that the treatment and control groups are similar with respect to unknown confounding factors.

For the next three problems, use the following information. Scientists wanted to see if there is more caffeine in an average cup of coffee or cup of tea. They took a simple random sample of 100 cups of coffee and found they had on average 95 mg of caffeine with an SD of 10mg. The scientists also took a simple random sample of 84 cups of tea and found they had an average of 26mg of caffeine with an SD of 12mg.

8. Find a 95% CI for how much more caffeine in mg is in the average cup of coffee than in the average cup of tea.
a. 69 +/- 3.23.
b. 69 +/- 3.57.
c. 69 +/- 4.03.
d. 69 +/- 4.54.
e. None of the above.

9. If the scientists took another simple random sample of 100 cups of coffee and 84 cups of tea, computed the difference between the sample means, and used this to estimate the difference between the two population means, how much would the estimate typically be off by?
a. About 1.07 mg.
b. About 1.22 mg.
c. About 1.43 mg.
d. About 1.65 mg.
e. None of the above.

10. How much does a typical cup of coffee's caffeine content differ from 95mg by?
a. About 1 mg.
b. About 1.96 mg.
c. About 3.33 mg.
d. About 10 mg.
e. None of the above.

11. In a study on a certain antihistamine, a sample of 1000 patients with cold symptoms was obtained, and each patient was randomly assigned either to receive the antihistamine or a placebo of sugar and water. The patients in both groups were told to take a pill every morning. After 10 days, the patients returned, and the researchers recorded the body temperature for each patient, and also took a sample of the patient's blood. From the blood samples, the researchers found that of the 502 patients receiving the antihistamine, 407 had actually taken the antihistamine on day 10, and these 407 had statistically significantly lower temperatures than the 95 patients in the antihistamine group who had not taken it on day 10. Can we conclude from this that the antihistamine causes a decrease in body temperature?
a. Yes, because the antihistamine reduces the patient's fever and thus causes a reduction in body temperature.
b. Yes, because the antihistamine reduces cold symptoms in general, resulting in a reduction in body temperature.
c. No, because adherers generally tend to be healthier than non-adherers.
d. No, because the antihistamine might alter the genetic makeup of the patients, and this genetic change might be the source of the apparent reduction in body temperature.
e. No, because the antihistamine group is likely wealthier than the placebo group.

For the next 3 problems, suppose a simple random sample of 100 UCLA students is asked if they know how to ride a bicycle. A 95% confidence interval for the proportion of all UCLA students who know how to ride a bicycle is found to be (0.621 , 0.883).

12. What is the margin of error for this 95%-CI?
a. 0.102.
b. 0.119.
c. 0.131.
d. 0.177.
e. None of the above.

13. If a 99% confidence interval is calculated instead, how will it differ from the 95% confidence interval?
a. The 99% confidence interval would be narrower than the 95% confidence interval.
b. The 99% confidence interval would be the same as the 95% confidence interval.
c. The 99% confidence interval would be wider than the 95% confidence interval.
d. More information is needed in order to determine if the 99% CI would be wider or narrower than the 95% CI.
e. None of the above.

14. Suppose a researcher is considering whether for the whole population of UCLA students, the percentage who know how to ride a bicycle, π, might be 60%, 85%, or 90%. Which of the following is true, based on data from this study? Assume a significance level of 0.05.
a. The null hypothesis that $\pi = 60\%$ would be **accepted**, the null hypothesis that $\pi = 85\%$ would **not be accepted**, and the null hypothesis that $\pi = 90\%$ would be **accepted**.
b. The null hypothesis that $\pi = 60\%$ would be **rejected**, the null hypothesis that $\pi = 85\%$ would **not be rejected**, and the null hypothesis that $\pi = 90\%$ would be **rejected**.
c. The null hypothesis that $\pi = 60\%$ would be **rejected**, the null hypothesis that $\pi = 85\%$ would **be rejected**, and the null hypothesis that $\pi = 90\%$ would be **rejected**.
d. The null hypothesis that $\pi = 60\%$ would be **rejected**, the null hypothesis that $\pi = 85\%$ would **be accepted**, and the null hypothesis that $\pi = 90\%$ would be **accepted**.
e. The null hypothesis that $\pi = 60\%$ would be **accepted**, the null hypothesis that $\pi = 85\%$ would be **regretted**, and the null hypothesis that $\pi = 90\%$ would be **respected**.

15. In the study on Echinacea by O'Neil et al. (2008), why did the researchers find no statistically significant benefit from Echinacea?
a. The sample size was too large for the difference to be considered statistically significant.
b. The sample size was too small for the difference detected to be considered statistically significant.
c. The difference between the two sample means was too large to be considered statistically significant.
d. The publication bias from the standard error of the parameter could have resulted in a skew in the null distribution from the simulations.
e. There was no difference between the sample mean for the treatment group and the sample mean for the control group.

16. If a difference between two groups is studied and the p-value is greater than 5%, what can you conclude?
a. The probability that the null hypothesis is true is greater than 95%.
b. If there is no difference between the two populations, then the probability of observing a difference at least as large as that observed in the data is greater than 5%.
c. The probability that the null hypothesis is true is greater than 5%.
d. The probability, under the null hypothesis, that the confidence interval will have a standardized statistic with a skew greater than the p-value of the normal distribution is greater than 5%.
e. The probability that the alternative hypothesis is false is at least 95%.

17. Suppose Marine the cancer-sniffing dog really had no ability to detect colorectal cancer, and was merely getting lucky. Since the researchers found that Marine was correctly detecting cancer statistically significantly more often than 20% of the time, which of the following would be true?
a. The standard error of the p-value from the normal distribution is significantly larger than the p-value from the skewed t distribution.
b. The researchers would be making a Type I error.
c. The researchers would be making a Type II error.
d. The evidence against the alternative hypothesis is stronger than the evidence in favor of the null hypothesis.
e. None of the above.

18. Suppose you take a simple random sample of 15 UCLA students and a simple random sample of 9 USC students. You find that 73.3% of the UCLA students know how to ride a bicycle, and 55.6% of the USC students know how to ride a bicycle, and you are considering doing a test to determine if the difference between the two percentages is statistically significant. Which of the following is true?
a. You should do a t-test, because the data come from simple random sampling and the sample sizes are small.
b. You should do a Z-test, because the data come from simple random sampling and the sample sizes are large.
c. You should do neither a Z-test nor a t-test, because the assumptions of these tests are not satisfied since the population is not normal and the sample size is small.
d. You should do a Z-test, because the data come from simple random sampling and the population is normally distributed.
e. Either a t-test or a Z-test would be correct, because the data come from simple random sampling, the sample sizes are large, and in such cases the t-test and Z-test will give very similar results.

19. Does the dolphin study provide strong evidence to conclude that swimming with dolphins causes an improvement in depression symptoms?
a. No, because the population might not be normally distributed.
b. Yes, because this was a randomized controlled experiment.
c. No, because the population standard deviation is not known.
d. Yes, because the two-sided confidence interval came from a simple random sample rather than from iid observations drawn from the t distribution with known standard error.
e. No, because genetics might be a confounding factor.

20. Suppose researchers take a simple random sample of 1000 subjects, randomly choose 500 of them to receive a treatment designed to reduce anxiety and the other 500 to receive a placebo. They find that in the treatment group, 320/500 (64.0%) report a reduction in anxiety, and in the placebo group, 290/500 (58.0%) report a reduction in anxiety. Find a 95%-CI for the percentage with reduced anxiety for those given the treatment minus the percentage with reduced anxiety given the placebo.
a. 6% +/- 3.01%.
b. 6% +/- 3.55%.
c. 6% +/- 4.09%.
d. 6% +/- 6.03%.
e. None of the above.