

Practice problems on last topics:

- 1) We want to analyze the “pain data” described in “Pain Study” on the webpage.
 - a) Does there appear to be a difference in pain thresholds between students and dancers?
 - b) Does there appear to be a difference in pain thresholds between men and women?
 - c) Does there appear to be an interaction between sex and occupation? Plot the 4 group means to help answer this question (informal) and use the SAS output to answer the question (formally). Which subgroup seems to experience the least pain?
- 2) A researcher observes the number of births at a hospital in four consecutive quarterly periods observing 110, 57, 53, and 80 births. It is conjectured that twice as many babies are born during the first quarter (Jan.-March) than are born in any of the other three quarters. Test at $\alpha = .05$ if these data strongly contradict the stated conjecture.
- 3) 9.65

Some other practice problems:

- 1) A British weatherman keeps track of the temperature in the month of September in Lubbock, TX. He finds that the average (sample mean) high temperature in the 30 days was 35 degrees Celsius with a sample standard deviation of 3 degrees Celsius. His boss, being a good American, realizes no one will understand these meaningless numbers and asks him to convert them into degrees Fahrenheit. With the knowledge that $F = 32 + C \times (9/5)$ what is the mean and standard deviation in degrees Fahrenheit?
 - a) mean=35 std. dev.=32+ 3×(9/5)
 - b) mean=95 std. dev.=32+ 3×(9/5)
 - c) mean=95 std. dev.=3×(9/5)
 - d) mean=95 std. dev.=3
- 2) A researcher wants to know if there is a connection between PCB’s in the environment and brain function of infants. In a certain group of children, 10% have a “high” level of PCB’s, 80% have “good” brain function, and 6% of the children have both “high” levels of PCB’s and “good” brain function.
 - A) Are the two attributes (“high” levels of PCB’s) and (“good” brain function) independent of each other? (circle your answer and numerically explain your answer)
 - a) Yes b) No c) we need more information to be sure.
 - B) If you know that someone has “high” levels of PCB’s, how likely is it that they have “good” brain function? (explain your reasoning)
 - a) .08 b) .50 c) .60 d) .80
- 3) A company wants to do a poll to see what proportion of its employees favor their new flexible hour program. The vice president decided to choose 10 employees, and he observes the proportion that favor the new program. His secretary thinks it is a better idea to let 100 employees answer the question, and observe the proportion out of 100 that favor the new program.

Let \hat{p}_{10} denote the sample proportion of its employees that favor the program in the vice presidents’ poll and let \hat{p}_{100} denote the average number from the secretary’s poll.

 - A)

- a) \hat{p}_{10} and \hat{p}_{100} should be identical.
- b) \hat{p}_{10} and \hat{p}_{100} should be approximately the same.
- c) \hat{p}_{10} should be smaller than \hat{p}_{100} .
- d) \hat{p}_{10} should be larger than \hat{p}_{100} .
- B)
- a) \hat{p}_{10} and \hat{p}_{100} have identical variability.
- b) \hat{p}_{10} and \hat{p}_{100} have approximately the same variability.
- c) \hat{p}_{10} is less variable than \hat{p}_{100} .
- d) \hat{p}_{10} is more variable than \hat{p}_{100} .
- 4) We want to estimate, μ , the mean height of adult magnolia trees in the Brazos Valley. To do so we sample 8 trees and compute \bar{x} , s^2 and compute a 90% confidence interval.
- A) If we now want to compute a 95% confidence interval the resulting interval be:
- a) wider than the first interval
- b) narrower than the first interval
- c) approximately the same length
- d) there is absolutely no way to tell if it will be wider or narrower.
- B) A colleague tells us that the current 90% interval must become one quarter of it's current width. How many trees should we sample the next time (it will still be a 90% interval)?
- a) 2
- b) 8
- c) 32
- d) 64
- 5) I compute a 90% interval (CI1) for the % of women who will develop breast cancer in the next 10 years based on data from a random sample of 1000 women. In a related study, the proportion who develop cancer is to be estimated from 4000 women and used to construct CI2. Which is the most likely thing to happen?
- a) CI2 will be about 4 times as wide as CI1.
- b) CI2 will be about 2 times as wide as CI1.
- c) CI2 will be about 50% as wide as CI1.
- d) CI2 will be about 25% as wide as CI1.
- 6) In my place of employment we always compute 80% confidence intervals for a population mean, μ , and in December we will compute 50 such intervals. Which of the following is true?
- a) The *best guess* is that 40 of the intervals *will not* contain μ .
- b) It is *guaranteed* that 40 of the intervals *will not* contain μ .
- c) The *best guess* is that 40 of the intervals *will* contain μ .
- d) It is *guaranteed* that 40 of the intervals *will* contain μ .
- What's the approximate chance that less than 75% of the intervals contain μ .
- 7) A nutritionist is interesting in assessing the effect of eating x grams of cashews per day on heart rate. After observing hundreds of people they want to know what the average heart rate is of people who who eat 20 grams daily. What is the appropriate interval to gauge plausible values for this unknown rate?
- a) A confidence interval.
- b) A Prediction interval.
- c) Either a confidence or prediction interval.
- d) Neither a confidence or prediction interval.

8) An entomologist wants to prove that native ants are more likely to attack fire ants if they are starved (thus aggression depends on diet). After collecting lots of data from starved and fed ants, she does a hypothesis test at $\alpha = .05$. In this case there is:

- a) A 95% chance that she concludes that aggression depends on diet even though it doesn't.
- b) A 95% chance that she concludes that aggression doesn't depend on diet, though it does.
- c) A 5% chance that she concludes that aggression depends on diet even though it doesn't.
- d) A 5% chance that she concludes that aggression doesn't depend on diet, though it does.

9) In an experiment to see if a new diet helps people lose weight, the researchers weigh each of 40 people at the beginning of the study and then weigh them again at the end of the study. Which of the following is the best way to see if the diet works:

- a) Use a *2 sample test* to compare the before and after weights, test H_0 : the two *population standard deviations* (before and after) are equal.
- b) Use a *1 sample test* to analyze the change in weights between before and after, test H_0 : the *population standard deviation* of the changes is equal to 0.
- c) Use a *2 sample test* to compare the before and after weights, test H_0 : the two *population means* (before and after) are equal.
- d) Use a *1 sample test* to analyze the change in weights between before and after, test H_0 : the *population mean* of the changes is equal to zero.

10) A patient is said to be hypokalemic if the measured level of potassium is 3.5 meq/l or less. An individual's potassium level is not constant, however, but varies from day to day. Suppose that the variation follows a normal distribution. Assume your mean level is 3.8 with a standard deviation of .2.

a) If you are measured on many days, on what proportion of days will the measurements indicate hypokalemia?

b) It is known that the average in Brazos county in 1995 was 3.7 meq/l with a standard deviation of A . A county official is worried about hypokalemia in the county. They sample 36 people and if the average of the 36 measurements is less than 3.5 meq/l they will do more testing. How likely is it they will do more testing if the true average is still 3.7 meq/.

c) Actually, the average in the U.S. is about 3.7 meq/l with a standard deviation of A . It appears that now 1000 counties will do the same testing as Brazos, but now U.S. officials want to find a cutoff value, u , such that only 5% of the counties need more testing (using the same criterion as in b) to do more testing). They can only fund more extensive testing in 5% of the counties. What should u be?

11) A sample of households is randomly selected from all the households in the city to estimate the percentage in which the head of the household is unemployed. It is assumed that 15% of the people are unemployed.

- a) If they sample 3 people, how likely is it that exactly 1 is unemployed?
- b) What is the approximate probability that out of 100 people less than 12 are unemployed?

12) A researcher is interested in the number of eggs that a California Condor lays in a single litter. It is known that this number is always 1 or 2 eggs. Assume that 40% of the time the number of eggs is 1, the remaining times it is 2. Let X = the number of eggs laid by a randomly chosen condor.

a) What is the *distribution* of \bar{X} , the average number of eggs laid by two condor mothers, i.e., give the possible outcomes with associated probabilities in a table.

b) 100 Moms in a certain flock of condors will lay eggs, what is the expected number of total eggs they will have?

c) If out of these 100 Moms, a total of less than 150 eggs are laid, there will be serious concern for the survival of the condor. How likely is it this will happen?