

# Lab 12b: Hypothesis Tests for Means

Stat 131A, Fall 2018

## Learning Objectives:

- Hypothesis tests for means

### Problem 1

Which of the following research questions asks us to test a claim about a population mean?

- A bond measure takes a 2/3 majority to pass. Do more than 67% of the voters support the measure?
- What proportion of community college students transfer to a four-year college or university?
- Is the average course load for a community college student greater than 12 units?

### Problem 2

Which of the following situations calls for a hypothesis test about a population mean?

- According to the Center for Disease Control (CDC), roughly 21.5% of all U.S. high-school seniors in 2002 have used marijuana. A sociologist suspects that the rate among African-American high school seniors is lower.
- A recent study estimated that 20% of all college students in the United States smoke. The head of Health Services at Goodheart University suspects that the proportion of smokers may be lower there.
- A certain prescription allergy medicine is suppose to contain an average of 245 parts per million (ppm) of active ingredient. The manufacturer wants to check whether the mean concentration in a large shipment of pills is 245 ppm or not.
- A report on the College Board website stated that in 2003 males scored generally higher than females on the SAT exam. An educational researcher wants to check whether this is true in her school district.

### Problem 3

For the following scenarios, give the null and alternative hypotheses and state in words what  $\mu$  (population average) represents in your hypotheses. A good definition of  $\mu$  describes both the variable and the population.

- a. The National Assessment of Educational Progress (NAEP) is administered annually to 4th, 8th, and 12th graders in the United States. On the math assessment, a score above 275 is considered an indication that a student has the skills to balance a checkbook. In a random sample of 500 young men between the ages of 18 and 20, the mean NAEP math score is 272. Do we have evidence to support the claim that young men nationwide have a mean score below 275?
- b. The National Center for Health Statistics reports that the systolic blood pressure for males 35 to 44 years of age has a mean of 128. In a study of business executives, a random sample of 100 executives has a mean systolic blood pressure of 134. Do the data suggest that the mean systolic blood pressure for business executives is higher than 128?
- c. An analytical chemistry lab is conducting quality control tests on a drug. A single dosage of the drug should contain 8 mg of active ingredient. Of course, there will be a small amount of variability due to imperfections in the production process, but the mean of all dosages produced should be 8 mg. In 20 random dosages, the mean amount of active ingredient is 7.7 mg. Do the data suggest that the mean amount of active ingredient in all dosages produced is different from 8 mg?

#### Problem 4

The National Assessment of Educational Progress (NAEP) gives a math assessment every year to 12th graders in the U.S. On the math test, a score above 275 indicates that a student has the skills to balance a checkbook. For a random sample of 500 young men, the mean NAEP math score is 272 with a standard deviation of 78. Do we have evidence to support that claim that young men nationwide have a mean score below 275?

The null and alternative hypotheses are  $H_0 : \mu = 275$ ,  $H_a : \mu < 275$ . The level of significance is 5%.

The t-test statistic is -0.86 with a P-value of 0.20. What is the correct conclusion?

- a. We do not have enough evidence to conclude that the mean score is less than 275 for young men nationwide.
- b. It is likely that these 500 young men have a mean score less than 275.
- c. The evidence suggests that young men nationwide have a mean score less than 275.
- d. The evidence suggests that young men nationwide have a mean score equal to 275.

#### Problem 5

Suppose we want to assess the effect of a one-day SAT prep class at a 5% level of significance. Scores on the SAT writing exam can range from 200 to 800. A random sample of 50 students takes the SAT writing test before and after a prep class. We test the hypotheses:

- $H_0 : \mu = 0$
- $H_a : \mu > 0$

where  $\mu$  is the mean of the difference in SAT writing scores (after minus before) for all students who take the SAT prep class.

The sample mean is 5 with a standard deviation of 18. Since the sample size is large, we are able to conduct the T-Test. The T-test statistic is approximately 1.96 with a P-value of approximately 0.028. Indicate whether the conclusion is correct or incorrect, and explain.

- Students taking a one-day SAT prep class performed significantly better on the SAT writing exam than students who did not take the class.
- The one-day SAT prep class is associated with statistically significant improvements in SAT writing performance.
- The one-day SAT prep class produces statistically significant improvements in SAT writing performance.
- Students taking a one-day SAT prep class do not show statistically significant improvements in their SAT writing performance. Scores only increased by 5 points, which is not significant on an exam where scores can range from 200 to 800.

## Problem 6

Suppose we want to assess the effect of music on concentration. We measure “concentration” by recording the time it takes a person to complete a simple word puzzle. Each person does a puzzle with and without music. A coin flip determines the order of the treatments. Our subjects are a sample of 18 volunteers from a statistics class.

We test the hypotheses (at the 5% level):

- $H_0 : \mu = 0$
- $H_a : \mu > 0$

where  $\mu$  is the mean of the difference in times (“no music” minus “music”).

The sample mean is 4 seconds with a standard deviation of 9 seconds. The sample is not random, but treatments are randomly ordered. In addition, the distribution of times is not strongly skewed, so we decide to use the t-test despite the small sample size. The t-test statistic is approximately 1.89 with a P-value of approximately 0.038. Indicate whether the conclusion is correct or incorrect, and explain.

- Music is not associated with statistically significant improvements in concentration as measured by time to complete a simple word puzzle.
- Completion times only increased 4 seconds on average. College students listening to music performed significantly faster on word puzzles than students who were not listening to music.

- c. For college students, music produces statistically significant improvements in concentration as measured by time to complete a simple word puzzle.
- d. For college students, music is associated with statistically significant improvements in concentration as measured by time to complete a simple word puzzle.

### Problem 7

Engineers on the Bay Bridge are measuring tower rods to find out if any rods have been corroded from salt water. There are rods on the east and west sides of the bridge span. One engineer plans to measure the length of an eastern rod 25 times and then calculate the average of the 25 measurements to estimate the true length of the eastern rod. A different engineer plans to measure the length of a western rod 20 times and then calculate the average of the 20 measurements to estimate the true length of the western rod.

Suppose the engineer who collected 25 measurements for the eastern rod has a mean length of 23.9 feet with a standard deviation of 0.5 feet. The critical T-value for a 90% (one-sample) confidence interval with  $df = 24$  is  $T=1.71$ .

Which of the following is the resulting 90% confidence interval?

- a. (23.73, 24.07)
- b. (23.05, 24.76)
- c. (23.80, 24.00)

### Problem 8

Commute times in the U.S. are heavily skewed to the right. We select a random sample of 500 people from the 2000 U.S. Census who reported a non-zero commute time. In this sample the mean commute time is 27.6 minutes with a standard deviation of 19.6 minutes. Can we conclude from this data that the mean commute time in the U.S. is less than half an hour?

Conduct a hypothesis test at the 5% level of significance. What can we conclude?

- a. Nothing. The distribution of the variable in the population is heavily skewed, so the conditions for use of a t-model are not met. We cannot trust that the p-value is accurate for this reason.
- b. With a mean of 27.6 minutes, the data supports the claim that the average commute time is less than 30 minutes, but the difference is not statistically significant. We fail to reject the null hypothesis that the mean commute time in the U.S. in the year 2000 was 30 minutes.
- c. With a mean of 27.6 minutes, the data supports the claim that the average commute time is significantly less than 30 minutes. We reject the null hypothesis that the mean commute time in the U.S. in the year 2000 was 30 minutes.

### Problem 9

GPA's of all UC Berkeley students were collected, and were found to have an average of 2.6 with an SD of 1.3. A simple random sample was collected of 100 students who go to tutoring sessions and they were found to have an average GPA of 2.9 with an SD of 1.5. We wish to test whether the tutoring helps raise the GPA of students. Say whether each of the following are true or false:

- a) We cannot use this data, because it is not a simple random sample of all students.
- b) Our null hypothesis should be that students who go to tutoring have the same GPA's as those who do not, and the measured difference was only due to random chance.
- c) The average of the box should be 2.9.
- d) The SD of the box should be 1.3
- e) We should reject the null hypothesis at the 5% level.
- f) The p-value represents the chance that the null hypothesis is true, based on our observed results