Lab 13b: Two-Sample Hypothesis Test

Stat 131A, Fall 2018

Learning Objectives:

- Inference for the difference between two population means.

Problem 1

Identify the research question that involves an inference about a difference between two population means.

a. Do community college students prefer computer-based homework or pencil-and-paper homework in math classes?

b. Do community college students receive less in financial aid than students attending public universities?

c. Do female community college students prefer to work on group projects more than male students?

d. Is the average amount of financial aid received in a year by community college students greater than $2,600?

Problem 2

A publishing company wanted to test whether typing speeds differ when using word processor A or word processor B. The typing speeds (in words per minute) are recorded for a random sample of 25 typists using processor A, and for another (different) random sample of 25 typists using word processor B. Which "case" does this study fall into?

a. Comparing two means—independent samples.

b. Comparing two means—matched pairs.

Problem 3

Identify the situations that involve inference about a difference between two population means.

a. The National Assessment of Educational Progress (NAEP) is the largest national assessment of what students in the U.S. know and can do in various subject areas. Is the mean score for 8th graders in Texas on the NAEP math test higher than the national average of 281?

b. The mean score on the NAEP math test for 8th graders in Texas is compared to the mean score for 8th graders in California.

c. A school district uses questions from the NAEP math test to assess the effectiveness of a new computer-based math instruction. Students take the test before and after the intervention and the district looks for improvement.

d. A school district compares a computerized math program to individualized tutoring for 4th graders who have difficulty in math. They use questions from the NAEP math test in a pre-test and post-test design to assess improvement in math skills for the two groups.

Problem 4

A Canadian study measuring depression level in teens (as reported in the Journal of Adolescence, vol. 25, 2002) randomly sampled 112 male teens and 101 female teens, and scored them on a common depression scale (higher score representing more depression). The researchers suspected that the mean depression score for male teens is higher than for female teens, and wanted to check whether data would support this hypothesis.

If μ_1 and μ_2 represent the mean depression score for male teens and female teens respectively, which of the following is an appropriate pair of hypotheses in this case? Check all that apply.

a. $H_0 : \mu_1 - \mu_2 = 0, \quad \mu_1 - \mu_2 < 0$

b. $H_0 : \mu_1 - \mu_2 > 0, \quad \mu_1 - \mu_2 = 0$

c. $H_0 : \mu_1 = \mu_2, \quad \mu_1 > \mu_2$

d. $H_0 : \mu_1 - \mu_2 = 0, \quad \mu_1 - \mu_2 > 0$

Problem 5

As a result of dog breeding for certain physical traits, many dog breeds have changed in the last 100 years. A dog breeder wants to measure the effects of breeding on the face length of the boxer breed using a hypothesis test. Suppose that the dog breeder has face length measurements (in inches) from a random sample of 40 female boxer dogs taken in 1915 and a random sample of 40 female boxer dogs taken in 2015.

Which type of hypothesis test should the breeder use?

a. Test for one population proportion

b. Test for one population mean

c. Test for a difference in two population proportions

d. Test for a difference in two population means

Problem 6

In 1990 and 2004, a survey organization tested the 17-year-olds on mathematics as well as reading. The average score went up from 305 to 307. You may assume that the survey organization took simple random samples of size 1000 in each of the two years. The SD for the 1990 data was 34, and the SD for the 2004 data was 27. Can the difference between the 305 and 307 be explained as a chance variation?

a) Should you make a one-sample z-test or a two-sample z-test? Why?

b) Formulate the null and alternative hypothesis in terms of a box model? Do you need one box or two?

c) Is the difference real, or can it be explained by chance?

Problem 7

In 2000, 59% of college freshmen thought that capital punishment should be abolished; by 2010, the percentage had dropped to 35%. Is the difference real, or can it be explained by chance? You may assume that the percentages are based on two independent simple random samples, each of size 1000.

Problem 8

A university takes a simple random sample of 132 male students and 279 female students; 41% of the men and 17% of the women report working more than 10 hours during the survey week. To find out whether the difference in percentages is statistically significant, the investigator starts by computing z = (41 - 17) / 0.048. Is anything worng?

Problem 9

A geography test was given to a simple random sample of 250 high-school students in a certain large school district. One question involved an outline map of Europe, with the countries identified only by number. The students were asked to pick out Great Britain and France. As it turned out, 65.6% could find France, compared to 70.4% for Great Britain. Is the difference statistically different? Or can this be determined from the information given?

Problem 10

Some years, the Gallup Poll asks respondents how much confidence they have in various American institutions. You may assume the results are based on a simple random sample of 1000 persons each year; the samples are independent from year to year.

a) In 2005, only 41% of the respondents had "a great deal or quite a lot" of confidence in the Supreme Court, compared to 50% in 2000. Is the difference real?

b) In 2005, only 22% of the respondents had "a great deal or quite a lot" of confidence in Congress, whereas 24% of the respondents had "a great deal or quite a lot" of confidence in labor. Is the difference between 24% and 25% real?

Problem 11

A fast food chain has been accused of trying to save money by reducing the number of fries in your order when you purchase them in big size. As a stat 2 student, you're prepared to test these allegations. The parameter you are interested in is:

$$d = avg_{regular} - avg_{big}$$

where $avg_{regular}$ is the average number of fries in a regular order, and avg_{big} is the average number of fries in a big-size order. Your null hypothesis is that $d = 0$ and your alternative hypothesis is that $d > 0$. In each of the following scenarios, indicate in which of the options you expect to have a smaller p-value, or whether there is no difference. Explain your reasoning.

a. Option 1: $d = 5$. Option 2: $d = 10$

b. Option 1: you take 30 samples of each type of fry. Option 2: you take 300 samples of each type of fry. Assuming that the value of d is the same in both samples.

c. Option 1: $d = -10$. Option 2: $d = 10$. Assume that the sample siz is the same in both samples.

Problem 12

Luke and Leia are on a statistical mission in the forest moon of Endor. They want to determine whether there is a difference in the average height between two types of Ewoks: brown fur versus gray fur. Leia draws a random sample of 50 Ewoks with brown fur having an average height of 90 cms, and SD of 40 cms. Luke draws a random sample of 50 Ewoks with grey fur having an average height of 110 cms, and SD of 60 cms. Is the average height of gray fur Ewoks greater than the average height of brown fur Ewoks? Or is it just chance variation? Conduct a hypothesis test, using a 5% significance level.