Lab 12a: Hypothesis Tests for Proportions

Stat 131A, Fall 2018

Learning Objectives:

- Hypothesis tests for proportions

Problem 1

Which of the following research questions involves testing a claim about a single population proportion?

  a. According to The College Board, 62% of students graduating from a community college with an associate degree in 2007-2008 had no student loan debt. Has this figure increased since then?

  b. What proportion of students at Naugatuck Valley Community College commute more than 5 miles each way to campus?

  c. According to the National Postsecondary Student Aid Study, conducted by the National Center for Education Statistics, the average student loan debt for community college students who received an associate degree was $5,879. Has the average student loan debt for these students increased since then?

  d. Is there a difference in the proportion of males who have student loan debt and the proportion of females who have student loan debt when they graduate from community college?

Problem 2

According to the American Association of Community Colleges, 23% of community college students receive federal grants. The California Community College Chancellor's Office anticipates that the percentage is smaller for California community college students. They collect a sample of 1,000 community college students in California and find that 210 received federal grants.

  a. For which group are we making a hypothesis?

    i) the population of community college students

    ii) the population of community college students in California

    iii) a sample of community college students in California

  b. Which are the correct hypotheses for this scenario?

i) $H_0 : p = 0.23$, $H_a : p < 0.210$

ii) $H_0 : p = 0.23$, $H_a : p < 0.23$

iii) $H_0 : p = 0.210$, $H_a : p > 0.210$

iv) $H_0 : p = 0.23$, $H_a : p \neq 0.23$

c. What does p represent in the hypotheses?

   i) the proportion of community college students who received federal grants

   ii) the proportion of California community college students who received federal grants

   iii) the proportion of 1,000 community college students who received federal grants

   iv) the number of federal grant recipients attending community colleges in California

Problem 3

A psychic claims to be able to predict the outcome of coin flips before they happen. Someone who guesses randomly will predict about half of coin flips correctly. In 100 flips, the psychic correctly predicts 57 flips. Do the results of this test indicate that the psychic does better than random guessing? The hypotheses are

- $H_0 : p = 0.50$
- $H_a : p > 0.50$

where p is the proportion of correct coin flip predictions by the psychic.

Give the test statistic for this problem.

   a. Z = -1.40

   b. Z = -9.43

   c. Z = 1.40

Problem 4

In 2007, a Gallup poll estimated that 45% of U.S. adults rated their financial situation as "good." We want to know if the proportion is smaller this year. We gather a random sample of 100 U.S. adults this year and find that 39 rate their financial situation as "good."

After carrying out a hypothesis test for $p = 0.45$ compared to $p < 0.45$, we obtain a p-value of 0.114. Determine whether each of the following statements about the p-value is valid or invalid.

   a. There is an 11.4% chance that the percentage of adults who rate their financial situation as "good" is still 45%.

b. There is an 11.4% chance that the percentage of adults who rate their financial situation as "good" is now less than 45%.

c. If the percentage of adults who rate their financial situation as "good" is still 45%, there is an 11.4% chance that a random sample of size 100 would have a sample percentage of 39% or lower.

Problem 5

A die is rolled 100 times. The total number of spots is 368 instead of the expected 350. Can this be explained as a chance variation, or is the die laoded?

Problem 6

A die is rolled 1000 times. The total number of spots is 3680 instead of the expected 3500. Can this be explained as a chance variation, or is the die laoded?

Problem 7

According to one investigator's model, the data are like 400 draws made at random from a large box. The null hypothesis says that the average of the box equals 50; the alternative says that the average of the box is more than 50. In fact, the data averaged out to 52.7, and the SD was 25. Compute the test statistic Z and p-value. What do you conclude?

Problem 8

In 2009 news organizations reported that 47% of U.S. adults favored providing a legal way for illegal immigrants already in the U.S. to become citizens. In 2010, the percentage with this opinion rose to 50%. Suppose that a current poll of 100 randomly selected U.S. adults shows this percentage is higher than 50% this year.

A researcher conducts a hypothesis test with the data from the current poll. She finds that the increase this year is not statistically significant. What conclusion can we draw?

a. It must be true that 50% of the public still has this opinion this year.

b. The researcher must have misinterpreted the results. There should be a significant increase.

c. The sample size may be too small to detect a statistically significant increase.

Problem 9

Suppose a common drug used to treat high blood pressure causes nausea in 30% of patients who take it. A clinical trial with a new drug for high blood pressure shows a "statistically significant" decrease in the percentage of patients who experience nausea. What can we conclude?

    a. A smaller percentage of patients have side effects with the new drug, but we do not know how much smaller.

    b. The new drug is much less dangerous than the old drug.

    c. A much smaller percentage (much lower than 30%) of patients have side effects with the new drug.

Problem 10

Data from the Center for Disease Control estimates that about 30.4% of American teenagers were overweight in 2008. The definition of overweight is a body mass index (BMI) of over 25. The percentage was very similar for boys and girls.

A professor in public health at a major university wants to determine whether the proportion has changed since 2008. He samples 800 randomly selected incoming freshman at universities around the country. Using the BMI measurements, he finds that 210, or about 26%, of them are overweight.

The professor tests the hypotheses $H_0 : p = 0.304$ versus $H_a : p \neq 0.304$. The P-value is about 0.011. If the professor uses a significance level of 0.05, what conclusion can he draw?

    a. The evidence from the sample shows that the proportion of all American teenagers who are overweight is now less than than 30.4%.

    b. The evidence from the sample shows that the proportion of all American teenagers who are overweight is now different than than 30.4%.

    c. Because our sample is not representative of the population of all American teenagers, the professor cannot draw any reasonable conclusion.

    d. The evidence from the sample shows that the proportion of all incoming freshmen at American universities who are overweight is different than than 30.4%.