

# STA 100 Homework 4

Due 11:59 pm Friday, July 28 onto Gradescope

1. According to an official statement released by the state of California, the average height of a Redwood tree in California is greater than 350 ft. To assess this claim, a scientist collects a sample of 51 randomly chosen Redwood trees in California. She finds that the average heights in her sample is 355 ft, with a standard deviation 42 ft. Conduct a hypothesis test to assess whether the true average height is greater than 350 ft. Use  $\alpha = 0.1$ .
  - (a) State the null and alternative hypotheses.
  - (b) Calculate the test statistic and determine the critical value for this problem.
  - (c) Do you reject the null, or fail to reject?
  - (d) State the conclusion in terms of the problem.

2. A researcher wants to investigate the effectiveness of a new teaching method in improving students' test scores. The researcher selects a random sample of 20 students and administers a pre-test and post-test to each student. The scores (out of 100) are recorded in the table below:

	Score
Pre-test	70, 75, 82, 68, 72, 80, 77, 73, 78, 85, 67, 75, 88, 71, 74, 79, 76, 81, 84, 69
Post-test	78, 80, 86, 74, 79, 83, 82, 77, 81, 89, 76, 81, 91, 78, 81, 83, 79, 84, 87, 75

Using a significance level of 0.05, test whether there is a significant difference in the mean test scores before and after the teaching method. Assume the test score is normally distributed.

- (a) State the null and alternative hypotheses.
  - (b) Calculate the test statistic and determine the critical value for this problem.
  - (c) Do you reject the null, or fail to reject?
  - (d) State the conclusion in terms of the problem.
  - (e) Construct a 95% confidence interval for the mean difference in test scores. Is the confidence interval consistent with the conclusion in (d)?
3. To evaluate the policy of routine vaccination of infants for whooping cough, adverse reactions were monitored in 340 infants who received their first injection of the vaccine. Reactions were noted in 68 of the infants.
    - (a) Find the 95% confidence interval for the true probability of an adverse reaction to the vaccine.
    - (b) Interpret the confidence interval from (a) in terms of the problem.
    - (c) Does your interval suggest that under 25% of infants had an adverse reaction?
    - (d) If we made many, many 95% confidence intervals, what percentage would we expect to cover the true proportion?
  4. In a study of non-human primates, a sample of 71 orangutans were tested, and 14 were found to be blood type B.
    - (a) Find the 95% confidence interval for the true proportion of type B orangutans.
    - (b) Interpret your confidence interval from (a) in terms of the problem.

- (c) Does your interval support the claim that approximately 25% of orangutans have type B blood?
- (d) If we made many, many 95% confidence intervals, what percentage would we expect does not cover the true proportion?
5. An experiment was conducted in which two types of acorn squash were crossed. According to the genetic model, 1/2 of the resulting plants should have dark stems and dark fruit (A), 1/4 should have light stems and dark fruit (B), and 1/4 should have light stems and light fruit (C). The actual recorded counts are:

	A	B	C
Count	220	129	105

- (a) State the null and alternative hypotheses for testing the genetic model.
- (b) Calculate the test-statistic.
- (c) Find the range of  $p$ -value.
- (d) Do we reject, or fail to reject the null if  $\alpha = 0.05$ ?
- (e) State your conclusion in terms of the problem.
6. At a Midwestern hospital there were a total of 932 births in 20 consecutive weeks. Of those births, 216 occurred on weekends (Saturday and Sunday) and 716 occurred on weekdays (Monday, Tuesday, Wednesday, Thursday, Friday). The goal is to determine if approximately the same about of births occur on the weekends that would happen by chance (by chance is 2/7).
- (a) State the null and alternative hypotheses.
- (b) Calculate the test-statistic.
- (c) Find the range of  $p$ -value.
- (d) Interpret your  $p$ -value in terms of the problem.
- (e) Do we reject, or fail to reject the null if  $\alpha = 0.01$ ?
- (f) State your conclusion in terms of the problem.
7. Mongolian gerbils are thought to be equally likely to be brown, white, or black in color. A random sample showed the following frequencies:

	Black	Brown	White
Count	40	59	42

- (a) State the null and alternative hypotheses.
- (b) Calculate the test statistic.
- (c) Find the range of  $p$ -value.
- (d) Using  $\alpha = 0.1$ , do we reject or fail to reject the null?
- (e) State your conclusion in terms of the problem.
- R is necessary for the remaining questions. Attach source codes and any plots you produce to your homework submission. You may write down your numerical results.
8. A researcher wants to investigate whether a certain treatment has an effect on the growth rate of a particular species of plants. The researcher measures the growth rates of 10 plants before and after applying the treatment.

	Growth rates (mm/day)
Pre-treatment	2.1, 1.8, 2.4, 1.6, 2.0, 2.3, 1.7, 2.2, 2.5, 2.0
Post-treatment	2.6, 2.3, 2.8, 2.1, 2.4, 2.7, 2.0, 2.5, 2.9, 2.3

Use this data and R, perform a paired-sample  $t$  test to test whether there is a significant difference in the mean difference in growth rates before and after treatment. Use the significance level of 0.05.

- (a) Find the test statistic.
  - (b) Find the  $p$ -value.
  - (c) Do you reject or fail to reject the null?
  - (d) Construct a 95% confidence interval for the mean difference in growth rates.
9. Biologists wanted to know if a particular parasite (weevils) had a preference for a particular bean plant. The results are stored in `beans.csv`, where the result was which plant the weevil preferred (Pinto, Cowpea, Navy, Northern). Use this data and R, and assume the claim is that weevils are equally likely to select all bean plants.
- (a) Find the test statistic.
  - (b) Find the  $p$ -value.
  - (c) Do you reject or fail to reject the null if  $\alpha = 0.05$ ?
  - (d) Which bean plant(s) did the weevils prefer more than expected if the null was true?
  - (e) What category contributed most to the value of the test statistic?
10. On Canvas (Files  $\rightarrow$  Data) you will find the dataset `DRP.csv`. Results of an experiment to test whether directed reading activities in the classroom help elementary school students improve aspects of their reading ability. A treatment class of 21 third-grade students participated in these activities for eight weeks, and a control class of 23 third-graders followed the same curriculum without the activities. After the eight-week period, students in both classes took a Degree of Reading Power (DRP) test which measures the aspects of reading ability that the treatment is designed to improve. The first column is the group (treatment or control), and the second is the DRP score (the higher the better).
- (a) Find the 95% confidence interval for the difference in average DRP score.
  - (b) If the experimenter is interested in if the treatment group has a higher average than the control group, find the test statistic.
  - (c) If the experimenter is interested in if the treatment group has a higher average than the control group, find the  $p$ -value.
  - (d) Do we fail to reject or reject the null hypothesis if  $\alpha = 0.05$ ?