

MATH/STAT 4850: Foundations of Statistics, Spring 2018

Homework 1

Due Friday, January 26 at 3:30 PM.

This homework counts 30 points toward your total homework grade.

1. Suppose in the rat experiment, we didn't know whether to expect that the drug would increase or decrease the time for the rat to solve the maze. There were 5 rats in the experiment. The three treated rats had times of 8, 10 and 15 seconds, while the two untreated rats had times of 5 and 9 seconds.
 - (a) Compute the difference in mean times of the treated group and the control (untreated) group.
 - (b) For all possible permutations, compute the difference in mean times of treated group and untreated group.
 - (c) Find the percentage of time that the difference in mean times of treated and untreated is as large or larger (in absolute value) than what was observed.
 - (d) Interpret your answer above in terms of a p -value.
2. Consider the dataset `flights` that you can obtain by installing the package `nycflights13`. This gives information about flights out of New York City in the calendar year 2013.
 - (a) Compute the mean delay time for flights that departed from Lagaardia (airport code LGA).
 - (b) Compute the mean delay time for flights that departed from JFK.
 - (c) Use a permutation technique to estimate the null distribution under $H_0 : \mu_{JFK} = \mu_{LGA}$.
 - (d) What is the p -value associated with the test of H_0 versus H_a there is a difference in means?
 - (e) Would you reject the null hypothesis at the $\alpha = .05$ level?
3. Repeat the above hypothesis test using the following test statistics. Use the same sampling from permutations for computing all 3 p -values.
 - (a) The sum of the delay times from LGA
 - (b) The mean of the delay times from LGA
 - (c) The difference of means in delay times from LGA and from JFK.

4. Load the data contained in the `data/homework_1_4` file in the github page for the class by typing `load("<filepath>/homework_1_4")`, where `filepath` is the path to the directory where `homework_1_4` is stored. This data is from an ultimate league that I organize. People who signed up for the league are asked to rate their skill level as “high”, “mid” or “low”.

Imagine that this is a random sample of people who might sign up for the league that I organize. Use a permutation test of independence to determine whether there is a difference in how males and females who are likely to sign up for my league self-rate their skill levels.