

Data Science – Fall 2016
Homework #5

1. Persi Diaconis is a famous magician turned statistician who studies random number generation. One thing for which he is famous is the ability to ask you whether you want heads or tails, then flip a coin and get it right almost every time! You are interested in testing whether this is really true (in actuality, you are just going to test whether he gets it right more than half the time, which is what a random flip would produce.) Please include screenshots of your output, similar to the hypothesis testing solutions document on the website.
 - a. State the null model for this experiment.
 - b. If Persi flips the coin 25 times and gets it right 21 times, what is the p-value using...
 - i. ...the randomization/simulation method
 - ii. ...the “chi-squared test” (Pearson method), which is obtained from the “test probabilities” in the distribution of a categorical variable output, and setting the true probability is 50% success and 50% failure.
 - c. What decision do you make for each method above about whether his flips are random or not?

2. On January 28, 1986, the space shuttle Challenger exploded. This was due to O-ring failure. O-rings are supposed to prevent the leak from combustible gases - their damage can lead to explosions. The data below are from 23 previous shuttle flights and give temperatures at time of launch by whether there was O-ring damage. The temperature (in degrees F) is that of the joints between sections of the solid rocket boosters prior to ignition. Could this tragedy have been prevented? We are interested in testing to see whether the temperature was related to whether there was damage to the O-ring.

Damage: 53, 57, 58, 63, 70, 70, 75

Not Damage: 66, 67, 67, 67, 68, 69, 70, 70, 72, 73, 75, 76, 76, 78, 79, 81

- a. State the null model for this analysis.
- b. Calculate the p-value using...
 - i. ...the randomization/simulation method
 - ii. ...the t-test, which is obtained from the Fit Y by X model and selecting “t-test”.
- c. What decision do you make for each method above about whether his flips are random or not?
- d. (Optional, and only for those who have taken statistics before) The simulation test is a better test to use in this context compared to the t-test. Explain why (hint: what are the assumptions for the t-test?).