

STAT 477/577 - Homework Assignment 3

Due Tuesday, February 8, by 11:59 PM

General homework guidelines: All homework assignments should be submitted using Canvas. Please submit your answers separately, for each of the problems, as set-up in the Canvas submission portal. You are allowed to either type in your answers, as well as submit graphs directly, within each submission item. You are also allowed to submit a scanned copy of your answer, as long as the answers are submitted separately for each question, as instructed. You can either scan your answers (ask if you don't have access to a scanner or do not know how to use your phone to do so), or just submit a picture of your answer. Please note that if we can't read your answer, we won't be able to award any (partial) credit.

You have one attempt to submit your answers. If technical issues appear and your submission portal has closed for some reason, please email Prof. Caragea explaining the situation and requesting permission to resubmit. Please note that such requests must be made before the deadline.

For full credit, please make sure you submit your HW by the deadline. A late submission is possible, with a 20% penalty, as long as it is turned in before the end of the day on Thursday following the due day (for this HW, it is February 10-th). No submissions will be accepted past this date.

Homework problems.

1. [6 pts.] For each of the distributions described below, simulate the sampling distribution for \hat{p} and include an image of each of your simulated distributions with your answers. Use your simulated distributions to describe the mean, variance, and shape of the sampling distribution for the sample proportion in the following scenarios.
 - (a) According to the company's website, the proportion of Green milk chocolate M&Ms produced is 0.16. Let the sample proportion \hat{p} be the proportion of Green milk chocolate M&Ms in a small bag of 50 candies.
 - (b) Let the sample proportion \hat{p} be the proportion of heads when flipping a fair coin 50 times.
 - (c) Let the sample proportion \hat{p} be the proportion of times a 0, 1, or 2 appears when rolling a 10-sided die 65 times.
2. [9 pts.] Suppose the current treatment for a disease cures 62% of all cases. A new treatment method has been proposed and studied. In a sample of 80 subjects with the disease that were treated with the new method, 63 were cured. Do the results of this study support the claim that the new method has a higher cure rate than the existing method?
 - (a) Use R to give the summary table and bar graph of the sample data. You can use the data file **treatment.csv**.
 - (b) Explain why the binomial exact test should be used to conduct this hypothesis test.
 - (c) Use R to conduct a binomial exact test for determining whether or not the study supports the claim that the new method has a higher cure rate than the existing method. Make sure to include all steps in the hypothesis test. Use $\alpha = 0.05$.
 - (d) Use R to find the rejection region for this test. Use $\alpha = 0.05$.
 - (e) Based on the rejection region you found above, what is the observed Type I error rate for this test?
 - (f) Suppose the cure rate for the new method is 75%. What is the power of your hypothesis test?
 - (g) Use R to obtain a power curve for this hypothesis test when $\alpha = 0.05$. Describe this curve.

3. [7pts.] A start-up company is about to market a new computer printer. It decides to gamble by purchasing commercials during the Super Bowl. The company is hoping the name recognition will be worth the high cost of the commercials. The company's goal is to have over 40% of the public recognize its brand name and associate it with computer printers. The day after the game, a pollster contacts 420 randomly selected adults and finds that 181 of them know this company makes printers. Is this evidence that the company met their goal?
- (a) Use R to give the summary table and bar graph of the sample data. You may use the data file **printer.csv**.
 - (b) Explain why the score test should be used to conduct this hypothesis test.
 - (c) Use R to conduct a hypothesis test for determining whether or not the company met its goal. Make sure to include the null and alternative hypotheses, test statistic, p -value, and conclusion for the hypothesis test.
 - (d) Use R to calculate the power of this hypothesis test if the true proportion of the public who recognizes the company's brand name and product is either 0.45, 0.5, or 0.55 and when $\alpha = 0.01$, 0.05, or 0.1 (to make it easier to keep track of these calculations, fill in the table below with the corresponding values). Discuss the effect of the value of the true proportion and the value of α on the power of this hypothesis test.

p_a	α		
	0.01	0.05	0.1
0.45			
0.50			
0.55			

4. [3 pts.] Suppose instead that the company in problem above had contacted you before they conducted their survey. They state they want to be able to detect when the proportion of the public who recognizes the company's brand name and product is 0.44 with a probability of 0.8 when $\alpha = 0.05$. What sample size will they need to use to achieve this?