STAT 477/STAT 577
HW 2 - Module 1: Sections 3 and 4

1. 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 large bag of 100 of the candies.

   (a) Determine the sampling distribution of the sample proportion $\hat{p}$.

   (b) Find the probability a large bag of 100 of the candies would have more than 20% green M&Ms.

   (c) Find the probability a large bag of 100 of the candies would have less than 10% green M&Ms.

2. Many dog owners teach their dogs to "shake hands". Your friend has a dog and you notice over time that the dog seems to favor his right paw in doing this trick. You decide to test the accuracy of your perception. For 15 consecutive times the trick is done in the same way with the same person, you find that the dog extended his right paw 10 times.

   (a) Use R to give the summary table and bar graph of the sample data. You can use the data file **Dogs.csv**.

   (b) Use R to conduct a binomial exact test for determining whether the dog favors his right paw when "shaking hands". Make sure to include the null and alternative hypotheses, test statistic, p-value, and conclusion.

   (c) Use R to find the rejection region for this test. Use $\alpha = 0.05$.

   (d) Based on the rejection region you found in part (c), what is the observed Type I error rate for this test?

   (e) Based on the rejection region you found in part (c), what is the power of your hypothesis test if the dog favors his right paw with probability 0.6, 0.75, or 0.9?

3. A company's old antacid formula provided relief from heartburn for 75% of the people who used it. The company develops a new formula in hopes of improving on the proportion of users who obtain relief. In a random sample of 400 people, 312 had relief of their heartburn.

   (a) Use R to give the summary table and bar graph of the sample data. You can use the data file **Antacid.csv**.

   (b) Explain why you can use the score test for this hypothesis test.

   (c) Use R to conduct a score test for determining whether the new formula is better than the old formula. Make sure to include the null and alternative hypotheses, test statistic, p-value, and conclusion.

(d) Use R to calculate the power of this score test if the true proportion of users who obtain relief from their heartburn with the new formula is either $p = 0.8$, 0.85, or 0.9 and $\alpha = 0.01$, 0.05, and 0.1.

(e) Discuss the effect of the value of the population proportion p and the value of α on the power of this hypothesis test.

(f) After the above analysis, the company decided to switch production to the new antacid formula. After several years in production, they found the new formula provided relief to 80% of the people who used it. Suppose the company would like to test another formula in the future. What sample size will they need to use to have a power of 0.9 to detect an improvement in the proportion of users who obtain relief of 0.05 if $\alpha = 0.05$.