Project 1 **Math 121**

Please type your solutions to the questions below in Microsoft Word or Google Docs (or other document editor) and turn in your solutions in class on ***Monday, November 10****. Your grade will be based on three factors: completeness, correctness, and style. To get full style credit you should write all answers in complete sentences. It is okay to discuss the problems with other students, but all of your solutions must be explained in your own words.*

A study published in *Nature* in 2014 reported on a study where 38 rhesus monkeys were placed on a restricted calorie diet while another 38 rhesus monkeys were given a normal diet. Below is a table that shows how many of the monkeys in each group had died by age 33.

	Restricted calorie diet	Regular diet
Died by age 33	26	32
Lived past 33 years	12	6
Total	38	38

1. Write a short introductory paragraph about this study. What were the explanatory and response variables that the researchers were interested in? What were the survival rates (proportion that survived) for each of the two groups of monkeys, and how big was the difference? How did the researchers investigate the relationship between the two variables? Did they use an experiment or an observational study?

   This was an experiment, although it doesn't say if the monkeys were randomly assigned to the treatment groups. The explanatory variable was the diet and the response variable was survival to age 33. Survival rates were 12 out of 38 (31.6%) in the treatment group versus 6 out of 38 (15.8%) in the control group.

2. From the data, it looks restricted calorie diets help monkeys live longer, but is this result statistically significant? Carry out an appropriate hypothesis test and explain what the results mean. Be sure to include a statement of the hypotheses, the test statistic (z-value), and the p-value as part of your answer.

   - $H_0 : p_{\text{treatment}} = p_{\text{control}}$
   - $H_A : p_{\text{treatment}} \neq p_{\text{control}}$

   The z-value for this test is

   $$z = \frac{\frac{12}{38} - \frac{6}{38}}{\sqrt{\frac{18}{76}\left(1 - \frac{18}{76}\right)\left(\frac{1}{38} + \frac{1}{38}\right)}} = 1.62.$$

   This corresponds to a 2-sided p-value of 10.5% which is weak evidence and we should not reject the null hypothesis. These results are inconclusive about whether calorie restriction works extend lifespan.

3. Find a 95% confidence interval for the difference in survival rates for the two groups of monkeys. Explain in words you are 95% confident is in the interval.

   The confidence interval is -3.0% to 34.6%. We can be 95% sure that monkeys on a restricted diet are between 3% less likely and 34.6% more likely to live to 33 years old.

4. Are the samples sizes big enough to trust the results of the two sample inference techniques used here? Explain how you can tell. For the hypothesis test, we do have at least 5 successes and 5 failures in each group, so that is trustworthy. For the confidence interval we only have 6 monkeys in the control group that lived to 33 years old, which is below the 10 successes and 10 failures threshold. So that means we should trust our confidence interval a little less.

5. Write a concluding paragraph to summarize your results. Should we be worried about lurking variables, and if not, why not? Should we conclude that low calorie diets cause monkeys to live longer? Explain the reasons for your conclusion.

   If the monkeys were randomly assigned to the treatment groups, then we probably don't need to worry about lurking variables, but that isn't specified. The p-value was not significant, so we cannot conclude that restricted calorie diets increase life span.