Power Analysis

Allison Fox, Hannah George, and Carlie McCleary

Introduction

The team conducted a power analysis of the experiment by simulating the data generating process in three separate outcome scenarios. The purpose of the analysis was to ensure an adequate amount of samples are collected to generate enough power to properly reject the null hypothesis of no effect. The three scenarios that were explored were related to the expected treatment effect of our age, gender, and news platform (Youtube vs Facebook) covariates. In each subsequent simulation the magnitude of the effects increased, which was the political news bias score, while the standard deviations remained constant. Measurements of bias used for the experiment range from -35 to 35 with left leaning values being <=0 and right leaning >=0.

Assumed treatment effects used in the simulations were influenced by the findings of a Pew Research article (https://www.pewresearch.org/fact-tank/2020/04/01/americans-main-sources-for-political-news-vary-by-party-and-age/). The article suggests there are demographic differences apparent among those who rely on left or right leaning news sources. More specifically, Fox News is one of the most well known Republican leaning sources and it's viewership is predominantly male and older individuals. The rest of the news sources outlined in the article were Democratic leaning and almost all had majority female viewers. This evidence aided in setting the baseline treatment effect directions and magnitudes.

Implementation and Analysis

The first step of the power analysis was to simulate the data generating process. We created a function to produce expected experimental data given an estimate of the co-variate treatment effects (tau) and a baseline sample size of 126.

Next, we created a function that uses the newly generated experimental data to produce multiple outcomes for power, varying by sample size. This was accomplished by collecting 1,000 p-values for each percentage of sample size taken from the experimental data. The percentages ranged from 10% to 200% of the original sample size of 126. The actual power calculation was performed at each sample size value by creating 1,000 linear regression models with the covariates and saving the associated p-values for each one. Power for each covariate was determined by the finding the fraction of models that resulted in p-values at or below our statistical significance level of 0.05.

Below are the three different scenarios the team explored. We aimed to examine how our experiment's power would change if it resulted in different treatment effect means. The tau distributions were created using the adjusted mean value, a constant standard deviation of 3 because it best resembled the data, and the estimated sample size. The functions to generate experimental data and then calculate power were called, resulting in power lists for our 'Low Treatment Effect', 'Regular Treatment Effect' and 'Large Treatment Effect' scenarios.

Low Treatment Effect

```
y0 <- rnorm(estimated_sample_size, mean = 0, sd = 3)
tau_male <- rnorm(estimated_sample_size, mean = 3, sd = 3)
```

```
tau_female <- rnorm(estimated_sample_size, mean = -1, sd = 3)
tau_old <- rnorm(estimated_sample_size, mean = 2, sd = 3)
tau_facebook <- rnorm(estimated_sample_size, mean = -1, sd = 3)
```

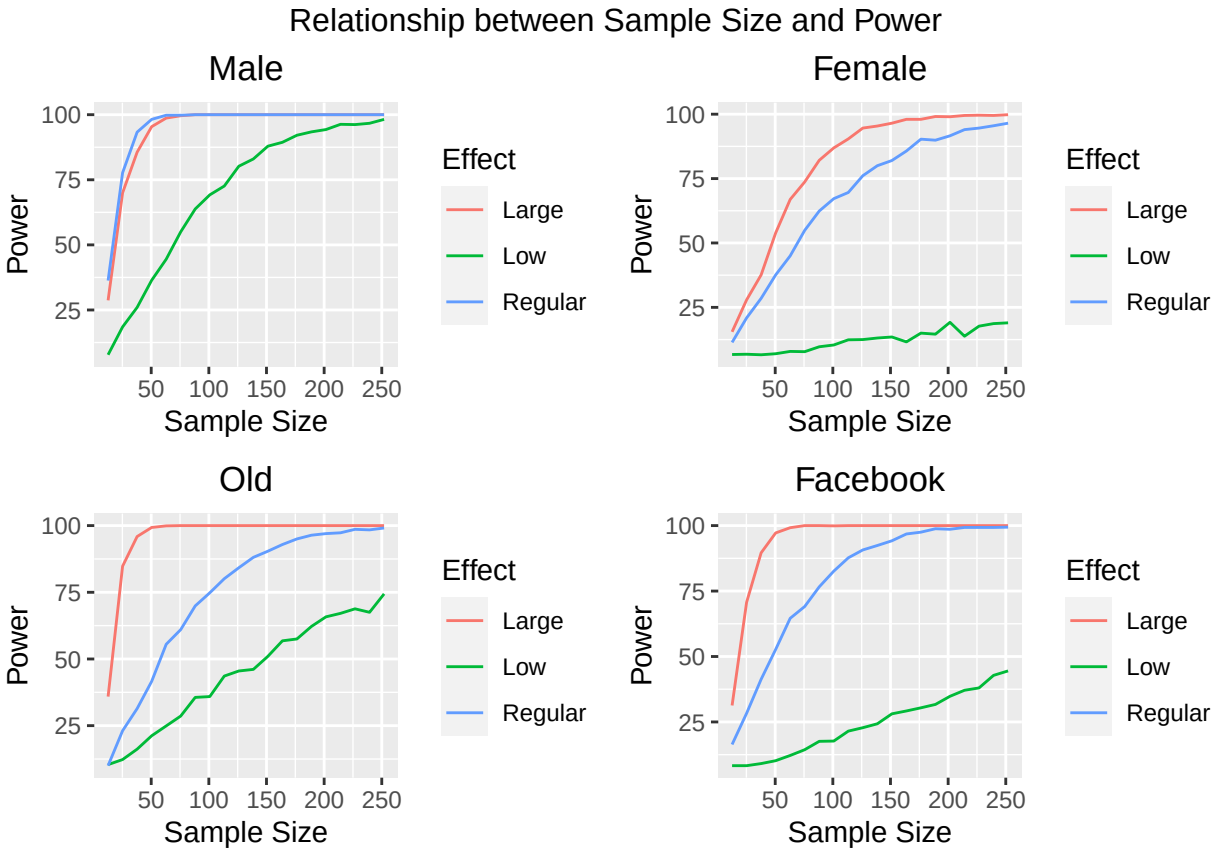
Regular Treatment Effect

```
y0 <- rnorm(estimated_sample_size, mean = 0, sd = 3)
tau_male <- rnorm(estimated_sample_size, mean = 5, sd = 3)
tau_female <- rnorm(estimated_sample_size, mean = -3, sd = 3)
tau_old <- rnorm(estimated_sample_size, mean = 4, sd = 3)
tau_facebook <- rnorm(estimated_sample_size, mean = -3, sd = 3)
```

Large Treatment Effect

```
y0 <- rnorm(estimated_sample_size, mean = 0, sd = 3)
tau_male <- rnorm(estimated_sample_size, mean = 7, sd = 3)
tau_female <- rnorm(estimated_sample_size, mean = -5, sd = 3)
tau_old <- rnorm(estimated_sample_size, mean = 6, sd = 3)
tau_facebook <- rnorm(estimated_sample_size, mean = -5, sd = 3)
```

Sample Size and Power



The relationship between sample size and power for each covariate is highlighted in the graphs above. Our

existing experimental targets are to generate a sample size of 126 and to achieve a power of 80% or greater. For most of these covariates, it is evident that a Low treatment effect will not allow us to hit our power goal. If our treatment effects at the conclusion of this experiment turn out to be more aligned with the Regular and Large effects that we simulated, we would be able to generate a power of greater than 80% with a sample of 126. We can conclude that our current experimental framework should allow us to successfully reject the null hypothesis of no effect, if indicated by the result. Our intention is to use this same statistical technique to calculate the power for our real experimental data and findings.