Randomization Inference

Alex Stephenson

Introduction

This simulation is an example of randomization inference.

```
set.seed(8675309)
library(dplyr)
library(ggplot2)
library(readr)
```

To understand how the public perceives Donald Trump's tweets, YouGov runs a poll that asks a representative sample of the US population to rate each tweet the day they are published. Trump's writing style when tweeting is often hyperbolic, with certain words in all-caps and extending out others ("sooooo") for effect.

To show how randomization inference works, let's take 20 of Trump's recent tweets, and suppose that Trump randomly inserts hyperbolic phrases into his tweets. The trump.csv dataset has twenty tweets with scores for both Democrats and Republicans.

```
tweets = read_csv("trump.csv")
```

```
## Parsed with column specification:
## cols(
##   tweet = col_integer(),
##   tweet_text = col_character(),
##   score_obs = col_integer(),
##   exclamation = col_integer(),
##   Y_i1 = col_integer(),
##   Y_i0 = col_integer()
## )
```

```
glimpse(tweets)
```

```
## Observations: 20
## Variables: 6
## $ tweet       <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,...
## $ tweet_text  <chr> "Congratulations to the Philadelphia Eagles on a g...
## $ score_obs   <int> 53, -23, -35, -49, -26, -17, -6, -4, 7, 11, 37, 44...
## $ exclamation <int> 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0,...
## $ Y_i1        <int> NA, -23, -35, NA, -26, NA, -6, NA, 7, NA, 37, NA, ...
## $ Y_i0        <int> 53, NA, NA, -49, NA, -17, NA, -4, NA, 11, NA, 44, ...
```

Let's suppose for this simulation that exclamation is a treatment assignment and we want to know the effect that it has on the score of the tweets.

Sharp Null Hypothesis

```
# Fill in potential outcomes to make the Sharp Null
tweets_ri = tweets %>%
    mutate(Y_i1 = ifelse(is.na(Y_i1), Y_i0, Y_i1),
           Y_i0 = ifelse(is.na(Y_i0), Y_i1, Y_i0))%>%
```

```
    select(Y_i1, Y_i0, exclamation, score_obs)
glimpse(tweets_ri)
```

```
## Observations: 20
## Variables: 4
## $ Y_i1        <int> 53, -23, -35, -49, -26, -17, -6, -4, 7, 11, 37, 44...
## $ Y_i0        <int> 53, -23, -35, -49, -26, -17, -6, -4, 7, 11, 37, 44...
## $ exclamation <int> 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0,...
## $ score_obs   <int> 53, -23, -35, -49, -26, -17, -6, -4, 7, 11, 37, 44...
```

First we take a difference in means in our observed values, our average treatment effect (ATE). We are going to compare this value to a distribution created by randomizing treatment assignment under the assumption that the true potential outcomes are identical and so there is no difference in treatment and control. This is the Sharp Null Hypothesis.

```
ATE = mean(tweets_ri$score_obs[tweets_ri$exclamation == 1]) - mean(tweets_ri$score_obs[tweets_ri$exclama
ATE
```

```
## [1] -4.3
```

Simulation

To apply randomization inference, we first create all possible treatment vectors.

```
poss_treatments = matrix(NA, 10000, 20)
for(i in 1:nrow(poss_treatments)){
    poss_treatments[i,] = sample(tweets_ri$exclamation, 20, replace = F)
}

# Keep only unique treamtent assignments
poss_treatments = unique(poss_treatments)
```

Next we calculate the average treatment effect for each possible randomization

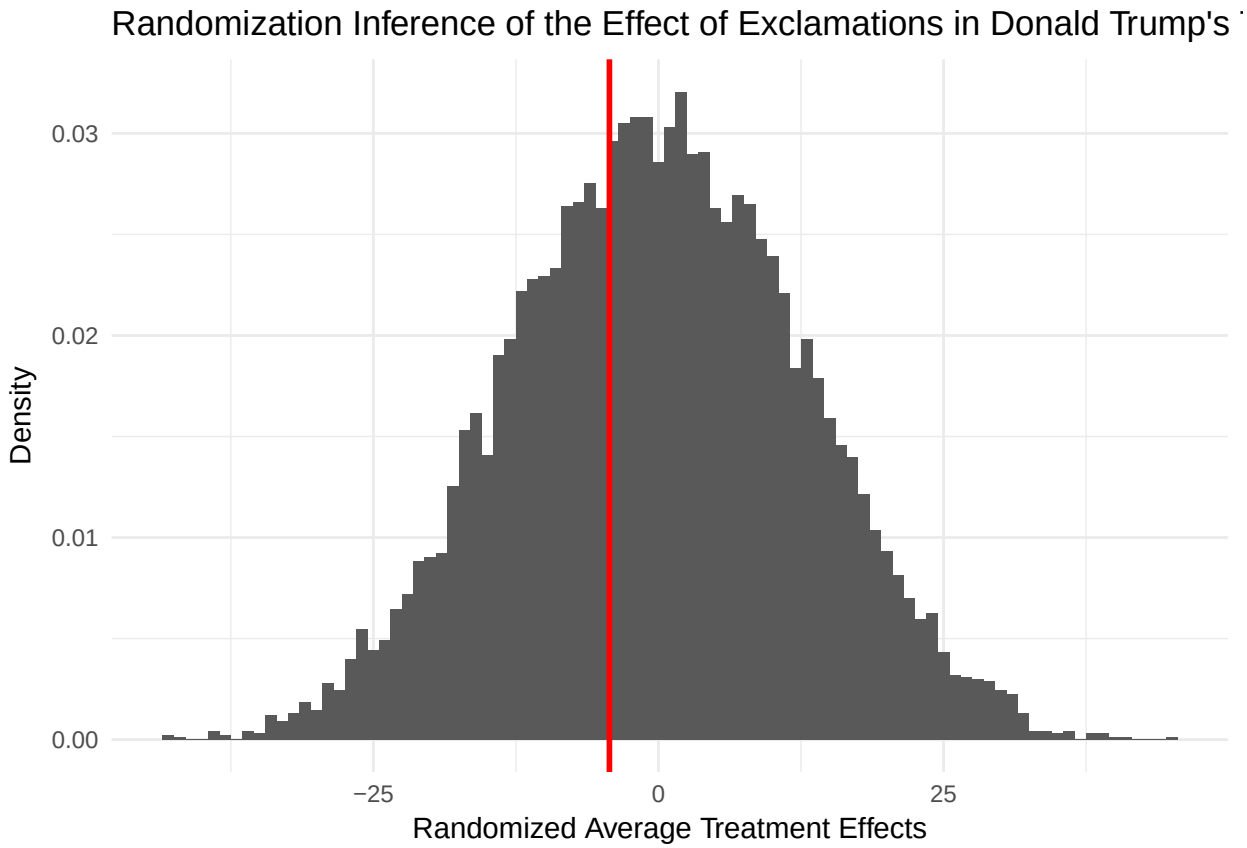
```
poss_ate = NA
for(i in 1:nrow(poss_treatments)){
    mean_w_exclam = mean(tweets_ri$score_obs[poss_treatments[i,]== 1])

    mean_wo_exclam = mean(tweets_ri$score_obs[poss_treatments[i, ]== 0])

    poss_ate[i] = mean_w_exclam - mean_wo_exclam
}
```

To evaluate whether our observed average treatment effect is significant, we can plot the distribution of our randomization ## Results

```
ggplot(as.data.frame(poss_ate), aes(x = poss_ate))+
    geom_histogram(aes(y=..density..), binwidth = 1)+
    geom_vline(xintercept = ATE, color = "red", size = 1)+
    theme_minimal()+
        xlab("Randomized Average Treatment Effects")+
        ylab("Density")+
        ggtitle("Randomization Inference of the Effect of Exclamations in Donald Trump's Tweets")
```



Our graph indicates that exclamations do not seem to have a major effect. We can also calculate a p-value.

```
# One tailed
sum(poss_ate>=ATE)/length(poss_ate)
```

```
## [1] 0.6316925
```

```
# Two tailed
sum(abs(poss_ate)>=ATE)/length(poss_ate)
```

```
## [1] 1
```

From this we can conclude (given our stylized assumptions) that there is not a significant effect of Trump using exclamations in the impact of the observed reaction scores to his tweets.