Hypothesis Testing

Prabidhik KC

2022-11-25

```
## Loading the necessary packages

library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.8     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.1
## v readr   2.1.2     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(gov50data)
library(infer)
```

```
tea
```

```
## # A tibble: 8 x 2
##   truth      guess
##   <chr>      <chr>
## 1 tea-first  tea-first
## 2 milk-first milk-first
## 3 milk-first milk-first
## 4 tea-first  tea-first
## 5 tea-first  tea-first
## 6 milk-first milk-first
## 7 tea-first  tea-first
## 8 milk-first milk-first
```

```
## Eandomly guessing the order
tea %>%
  mutate(random_guess = sample(guess))
```

```
## # A tibble: 8 x 3
##   truth      guess      random_guess
##   <chr>      <chr>      <chr>
```

```
## 1 tea-first   tea-first   tea-first
## 2 milk-first milk-first tea-first
## 3 milk-first milk-first milk-first
## 4 tea-first   tea-first   tea-first
## 5 tea-first   tea-first   milk-first
## 6 milk-first milk-first milk-first
## 7 tea-first   tea-first   tea-first
## 8 milk-first milk-first milk-first
```

```r
null_dist <- tibble(
  trump_share = rbinom(n = 1000, size = 1363, prob = 0.475)/1363
)

null_dist
```

```
## # A tibble: 1,000 x 1
##     trump_share
##           <dbl>
##  1        0.481
##  2        0.478
##  3        0.465
##  4        0.475
##  5        0.466
##  6        0.465
##  7        0.471
##  8        0.475
##  9        0.463
## 10        0.456
## # ... with 990 more rows
```

```r
null_dist %>%
  ggplot(mapping = aes(x = trump_share)) +
  geom_histogram(binwidth = 0.01) +
  geom_vline(xintercept = 0.475, size = 1.5) +
  geom_vline(xintercept = 0.44, size = 1.25, color = "red") +
  geom_vline(xintercept = 0.51, size = 1.25, color = "red")
```
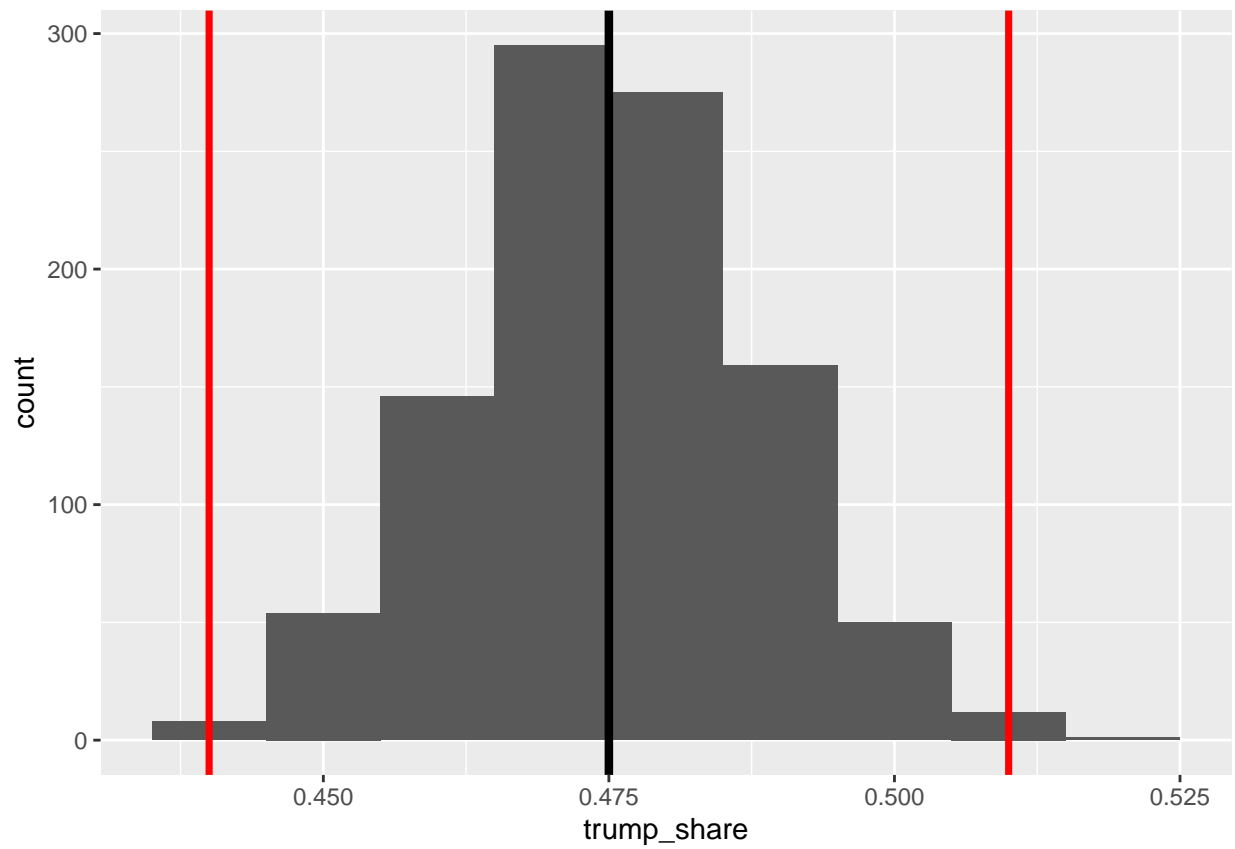
```
mean(null_dist$trump_share < 0.44) + mean(null_dist$trump_share > 0.51)
```

```
## [1] 0.006
```