

# Mid Project

2025-10-29

```
library(infer)

data(gss)

# Calculating the observed statistic
F_hat <- gss |>
  specify(age ~ partyid) |>
  calculate(stat = "F")
F_hat

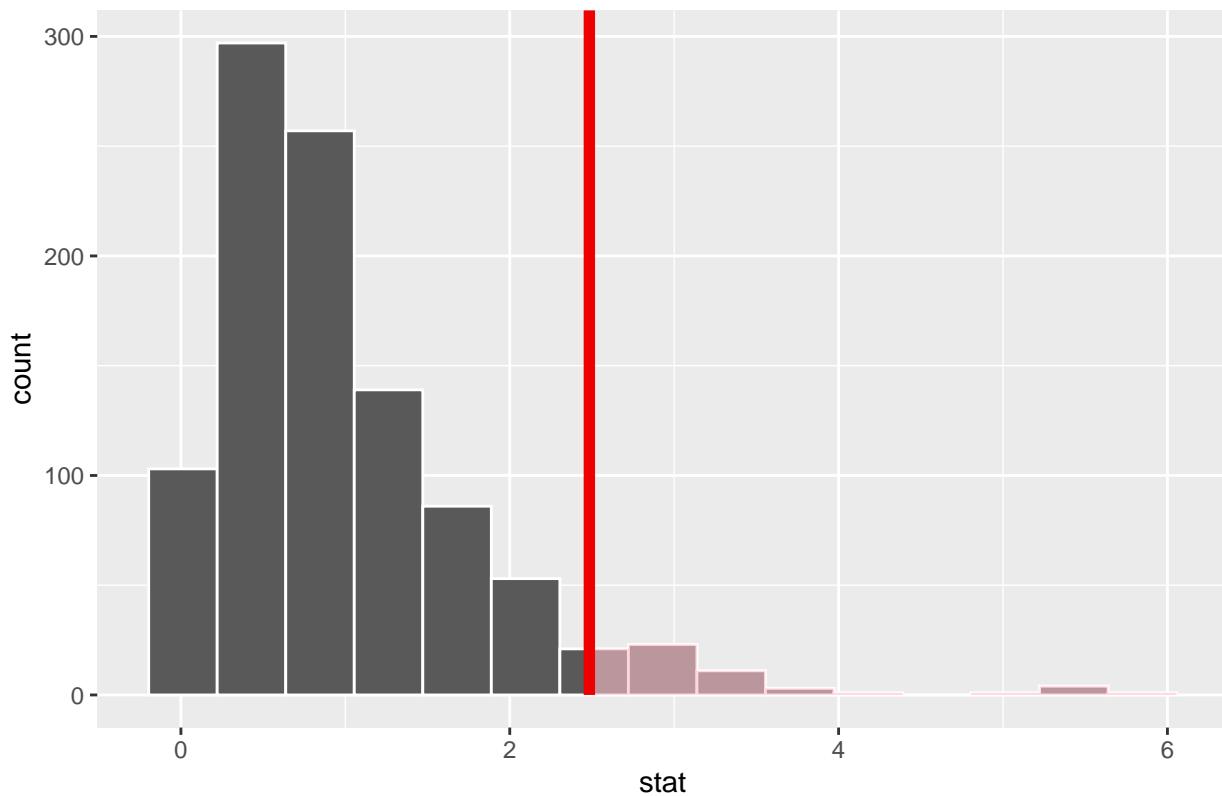
## Response: age (numeric)
## Explanatory: partyid (factor)
## # A tibble: 1 x 1
##       stat
##   <dbl>
## 1  2.48

# generate a null distribution
null_dist <- gss |>
  specify(age ~ partyid) |>
  hypothesize(null = "independence") |>
  generate(reps = 1000, type = "permute") |>
  calculate(stat = "F")
null_dist

## Response: age (numeric)
## Explanatory: partyid (factor)
## Null Hypothesis: indepe...
## # A tibble: 1,000 x 2
##       replicate     stat
##       <int>    <dbl>
## 1          1  0.573
## 2          2  0.871
## 3          3  0.0124
## 4          4  0.896
## 5          5  0.737
## 6          6  0.162
## 7          7  0.568
## 8          8  0.847
## 9          9  0.206
## 10         10  0.387
## # i 990 more rows

visualize(null_dist) +
  shade_p_value(obs_stat = F_hat, direction = "greater")
```

## Simulation-Based Null Distribution



```
null_dist |>  
  get_p_value(obs_stat = F_hat, direction = "greater")  
  
## # A tibble: 1 x 1  
##   p_value  
##     <dbl>  
## 1    0.055
```