Online CDS 101 *Introduction to Computational and Data Sciences*

Inference and simulation

Two-sided hypothesis tests using `infer`

These slides are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Download and load the dataset

You can follow along by downloading and loading the dataset by placing the following *setup* code block at the top of a R Markdown file.

````
```{r setup, include = FALSE}
# Load required packages
library(tidyverse)
library(infer)
# Load datasets
college_apps <- read_rds(
  url("http://data.cds101.com/college_applications.rds")
)
applicants_data <- read_rds(
  url("http://data.cds101.com/gender_discrimation.rds")
)
# Observed result
experiment_result <- (21/24) - (14/24)
```
````

Simulations recap

In the previous lecture, we ran the following simulations:

```
# College applications null distribution
college_apps_null <- college_apps %>%
  specify(formula = number_colleges ~ NULL) %>%
  hypothesize(null = "point", mu = 8) %>%
  generate(reps = 10000, type = "bootstrap") %>%
  calculate(stat = "mean")

# Gender discrimation null distribution
simulation_results <- applicants_data %>%
  specify(outcome ~ sex, success = "Promoted") %>%
  hypothesize(null = "independence") %>%
  generate(reps = 10000, type = "permute") %>%
  calculate(stat = "diff in props", order = combine("Male", "Female"))
```

Two-sided hypothesis testing with p-values

- If the research question was "Do the data provide convincing evidence that the average amount of schools that GMU students apply to is **different** than the national average?", the alternative hypothesis would be different.

$$H_0 : \mu = 8$$
$$H_A : \mu \neq 8$$

- Hence the p-value could change as well:

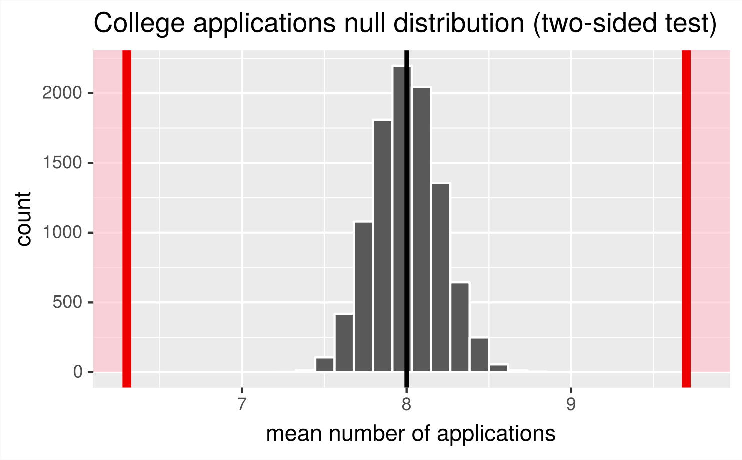
```
college_apps_p_value_right <- college_apps_null %>%
  get_pvalue(obs_stat = 9.7, direction = "right")

college_apps_p_value_left <- college_apps_null %>%
  get_pvalue(obs_stat = 6.3, direction = "left")

college_apps_p_value_two_sided <- college_apps_p_value_right +
  college_apps_p_value_left
```

Two-sided hypothesis testing with p-values

- If the research question was "Do the data provide convincing evidence that the average amount of schools that GMU students apply to is **different** than the national average?", the alternative hypothesis would be different.

$$H_0 : \mu = 8$$
$$H_A : \mu \neq 8$$

- Hence the p-value could change as well:



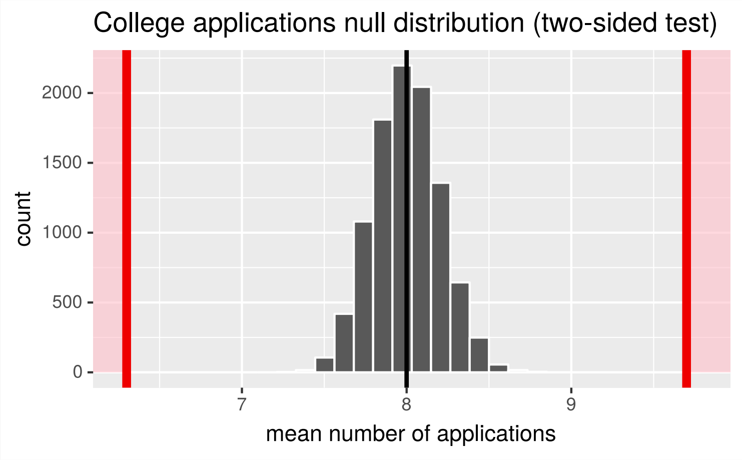
p-value = 0

Two-sided hypothesis testing with p-values

- If the research question was "Do the data provide convincing evidence that the average amount of schools that GMU students apply to is **different** than the national average?", the alternative hypothesis would be different.

$$H_0 : \mu = 8$$
$$H_A : \mu \neq 8$$

- Hence the p-value could change as well:



p-value = 0

Although in this example, it does not change.

Gender discrimination dataset: two-sided hypothesis test

- We can use the same null distribution that we generated earlier
- In the two-sided hypothesis test, we need to count when the difference in the men and women hiring fractions is **larger** than 0.292 and also when it is in the opposite extreme, which would be when the bias is towards hiring more women than men
- The opposite extreme corresponds to a difference in hiring fractions that is **less than** -0.292
- As before, we can filter the data just to keep these extreme outcomes, then divide the remaining rows and divide by 10,000

```
sim_p_value_right <- simulation_results %>%
  get_pvalue(obs_stat = experiment_result, direction = "right")

sim_p_value_left <- simulation_results %>%
  get_pvalue(obs_stat = -experiment_result, direction = "left")

sim_p_value_left + sim_p_value_right
```

Gender discrimination dataset: two-sided hypothesis test

- We can use the same null distribution that we generated earlier
- In the two-sided hypothesis test, we need to count when the difference in the men and women hiring fractions is **larger** than 0.292 and also when it is in the opposite extreme, which would be when the bias is towards hiring more women than men
- The opposite extreme corresponds to a difference in hiring fractions that is **less than** -0.292
- As before, we can filter the data just to keep these extreme outcomes, then divide the remaining rows and divide by 10,000

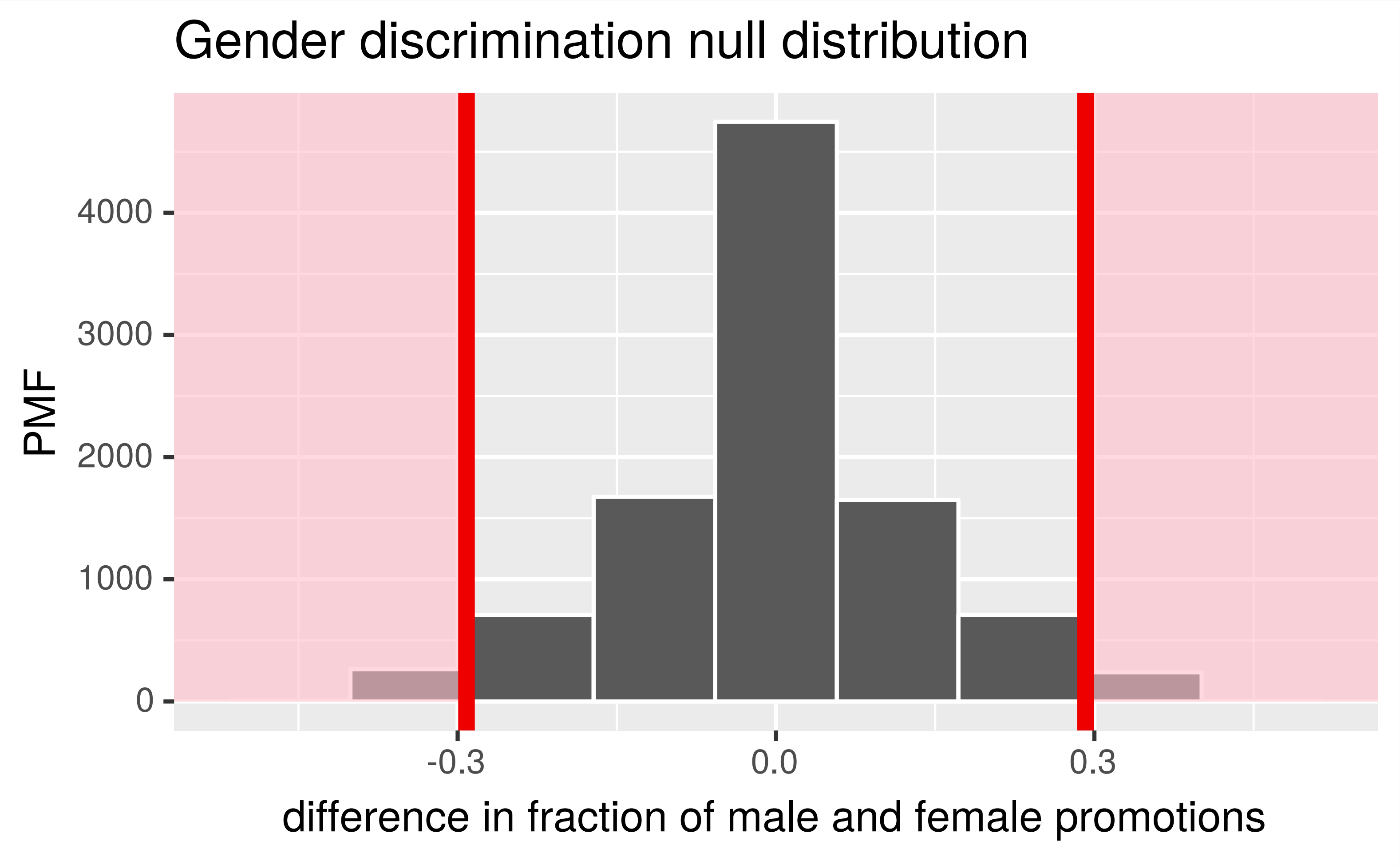
```
sim_p_value_right <- simulation_results %>%
  get_pvalue(obs_stat = experiment_result, direction = "right")

sim_p_value_left <- simulation_results %>%
  get_pvalue(obs_stat = -experiment_result, direction = "left")

sim_p_value_left + sim_p_value_right
```

p_value
0.0512

Visualization of null distribution (two-sided)

```
simulation_results %>%
  visualize(bins = 9) +
  shade_p_value(obs_stat = experiment_result, direction = "right") +
  shade_p_value(obs_stat = -experiment_result, direction = "left") +
  labs(
    x = "difference in fraction of male and female promotions",
    y = "PMF",
    title = "Gender discrimination null distribution"
  )
```

Visualization of null distribution (two-sided)



Credits

License

Creative Commons Attribution-NonCommerical-ShareAlike 4.0 International

Acknowledgments

Content adapted from the Chapter 3 OpenIntro Statistics slides developed by Mine Çetinkaya-Rundel and made available under the CC BY-SA 3.0 license.