

# Inference and simulation

---

A manual simulation of the gender discrimination experiment



# Recap: hypothesis testing framework

- We start with a null hypothesis ( $H_0$ ) that represents the status quo
- We also have an alternative hypothesis ( $H_A$ ) that represents our research question, i.e. what we're testing for
- We conduct a hypothesis test under the assumption that the null hypothesis is true, either via simulation or theoretical methods
  - If the test results suggest that the data do not provide convincing evidence for the alternative hypothesis, we stick with the null hypothesis
  - If they do, then we reject the null hypothesis in favor of the alternative

# Simulating the experiment

Let's simulate the scenario from the case study under the assumption of independence, i.e. leave things up to chance.

- If results from the simulations based on the **chance model** look like the data, then we can determine that the difference between the proportions of promoted files between males and females was simply **due to chance** (promotion and gender are independent).
- If the results from the simulations based on the chance model do not look like the data, then we can determine that the difference between the proportions of promoted files between males and females was not due to chance, but **due to an actual effect of gender** (promotion and gender are dependent).

# Simulating the experiment with a deck of cards

*Pretend for a moment that we didn't have a computer available, how could we simulate this experiment using playing cards?*

# Simulating the experiment with a deck of cards

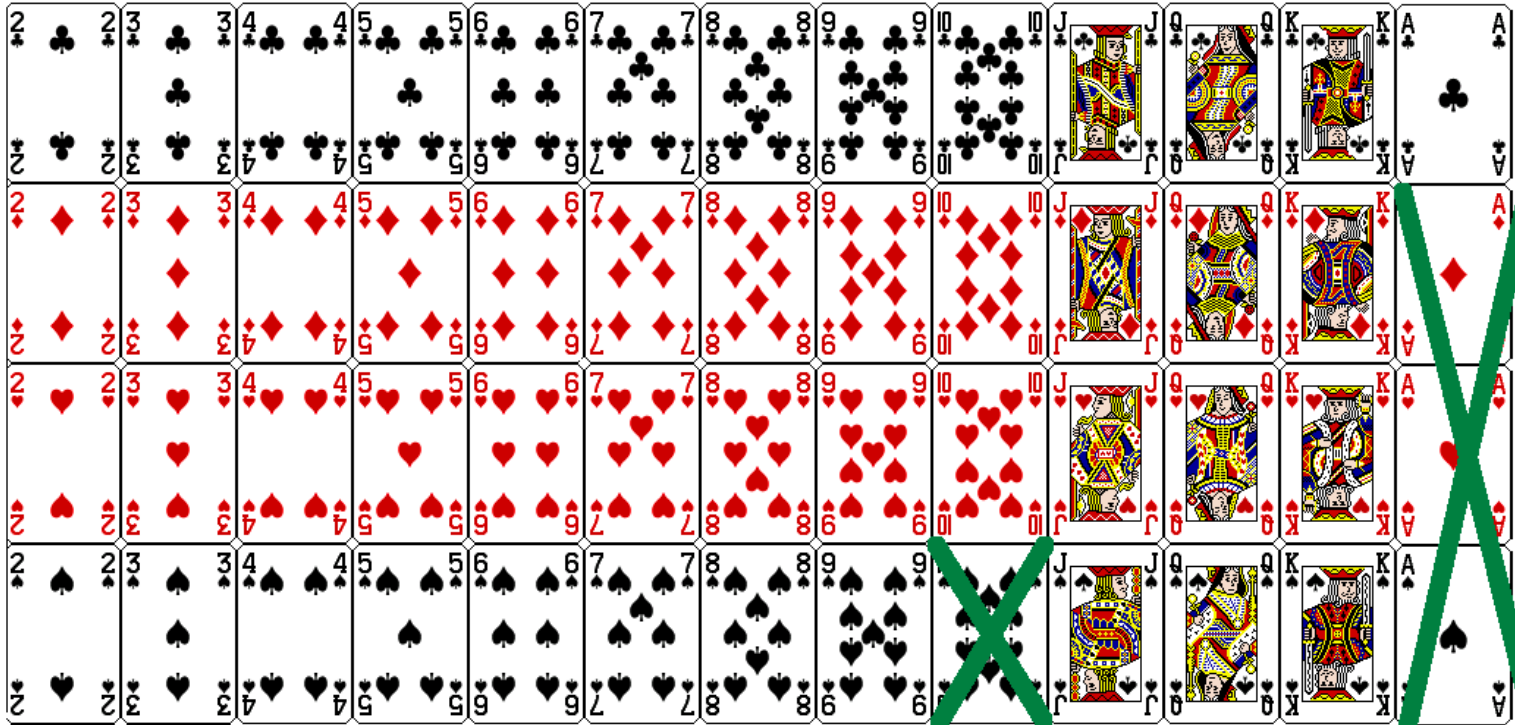
*Pretend for a moment that we didn't have a computer available, how could we simulate this experiment using playing cards?*

1. Let a face card represent *not promoted* and a non-face card represent *promoted*
  - Consider aces as face cards
  - Set aside the jokers
  - Take out 3 aces → there are exactly 13 face cards left in the deck (face cards: A, K, Q, J)
  - Take out a number card → there are exactly 35 number (non-face) cards left in the deck (number cards: 2-10)
2. Shuffle the cards and deal them into two groups of size 24, representing males and females
3. Count and record how many files in each group are promoted (number cards)
4. Calculate the proportion of promoted files in each group and take the difference (male - female), and record this value
5. Repeat steps 2 – 4 many times

# Step 1

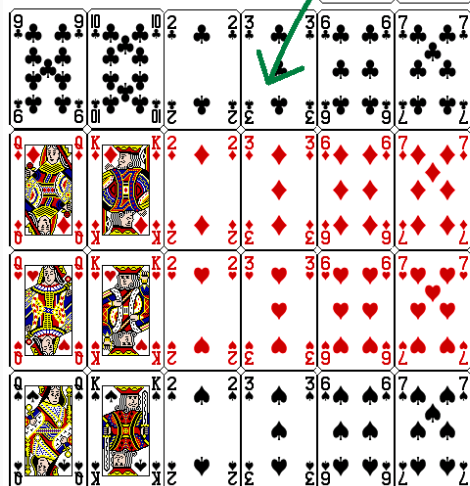
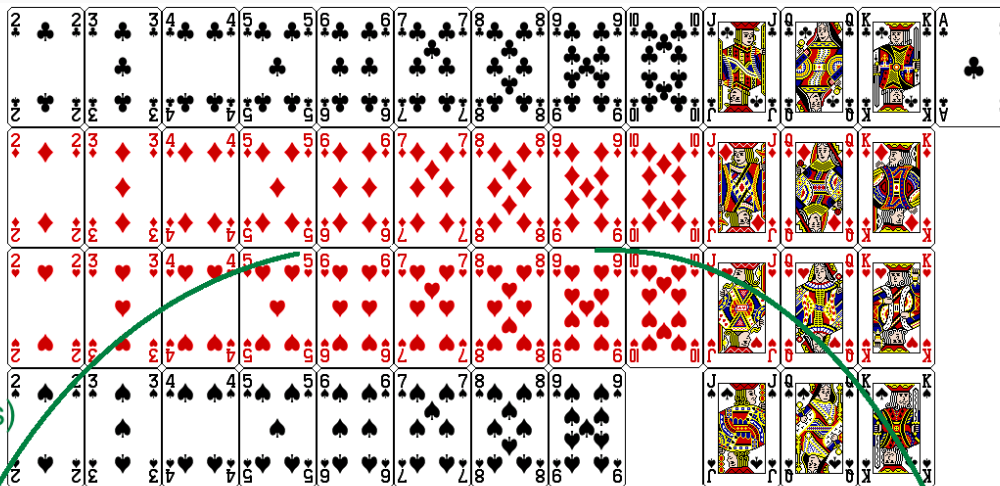
35 number (non-face) cards

13 face cards



# Step 2

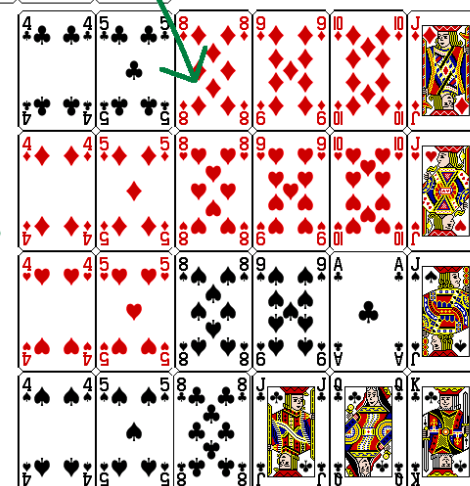
Shuffle and  
split into  
two groups  
of 24  
(males and females)



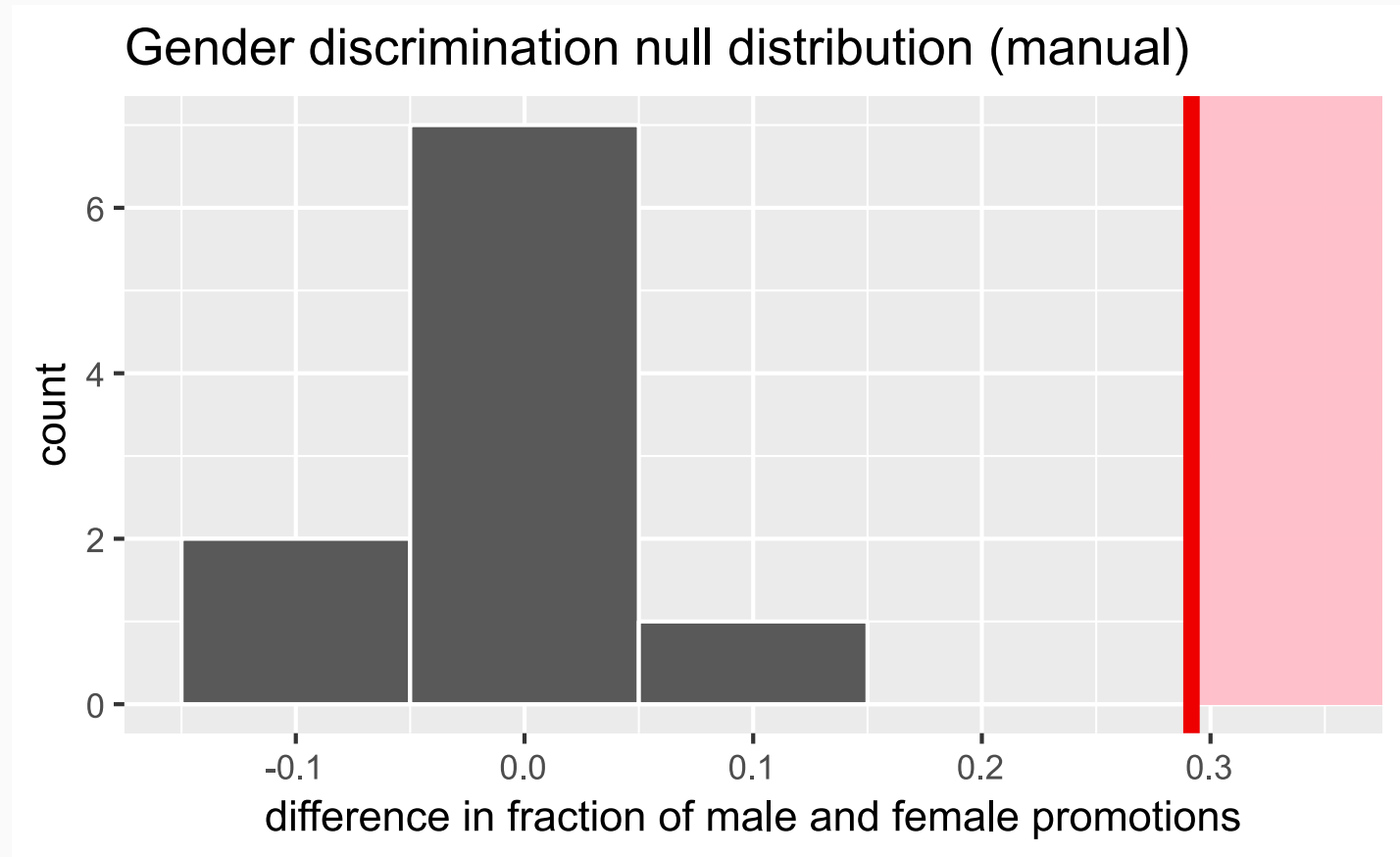
Males  
18 promoted  
 $18 / 24 = 0.75$

Females  
17 promoted  
 $17 / 24 = 0.708$

Difference =  $0.75 - 0.708 = 0.042$



# Null distribution using playing cards





# Null distribution results table

<b>Trial</b>	<b>Promoted men</b>	<b>Promoted women</b>	<b>Difference in proportions</b>
1	18	17	0.0416667
2	18	17	0.0416667
3	18	17	0.0416667
4	17	18	-0.0416667
5	16	19	-0.1250000
6	18	17	0.0416667
7	17	18	-0.0416667
8	19	16	0.1250000
9	16	19	-0.1250000
10	18	17	0.0416667

# Credits

License

[Creative Commons Attribution-NonCommercial-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](#).