introduction to inference

- case study: gender discrimination
- introduction to inference via simulation

Duke UNIVERSITY

Dr. Mine Çetinkaya-Rundel
Duke University

gender discrimination

- 48 male bank supervisors given the same personnel file, asked to judge whether the person should be promoted
- files were identical, except for gender of applicant
- random assignment
- 35 / 48 promoted
- are females are unfairly discriminated against?

data

		promotion		
		promoted	not promoted	total
gender	male	21	3	24
	female	14	10	24
	total	35	13	48

% of males promoted = 21/24 ≈ 88%

% of females promoted = 14/24 ≈ 58%

null hypothesis

"There is nothing going on"

promotion and gender are independent, no gender discrimination, observed difference in proportions is simply due to chance

alternative hypothesis

"There is something going on"

promotion and gender are dependent, there is gender discrimination, observed difference in proportions is not due to chance.

two competing claims

null hypothesis

H_0: Defendant is innocent

alternative hypothesis

H_A: Defendant is guilty

burden of proof

present the evidence

collect data

judge the evidence

"Could these data plausibly have happened by chance if the null hypothesis were true?"

yes

Fail to reject H_0

no

Reject H_0

Image source: http://en.wikipedia.org/wiki/File:Trial_by_Jury_Usher.jpg

recap: hypothesis testing framework

- start with a null hypothesis (H_0) that represents the status quo
- set an alternative hypothesis (H_A) that represents the research question, i.e. what we're testing for
- 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, stick with the null hypothesis
  - if they do, then reject the null hypothesis in favor of the alternative

simulation scheme

[use a deck of playing cards to simulate this experiment]

1. face card: not promoted, non-face card: promoted
   - set aside the jokers, consider aces as face cards
   - take out 3 aces → 13 face cards left in the deck (face cards: A, K, Q, J)
   - take out a number card → 35 number (non-face) cards left in the deck (number cards: 2-10)

Step 1:

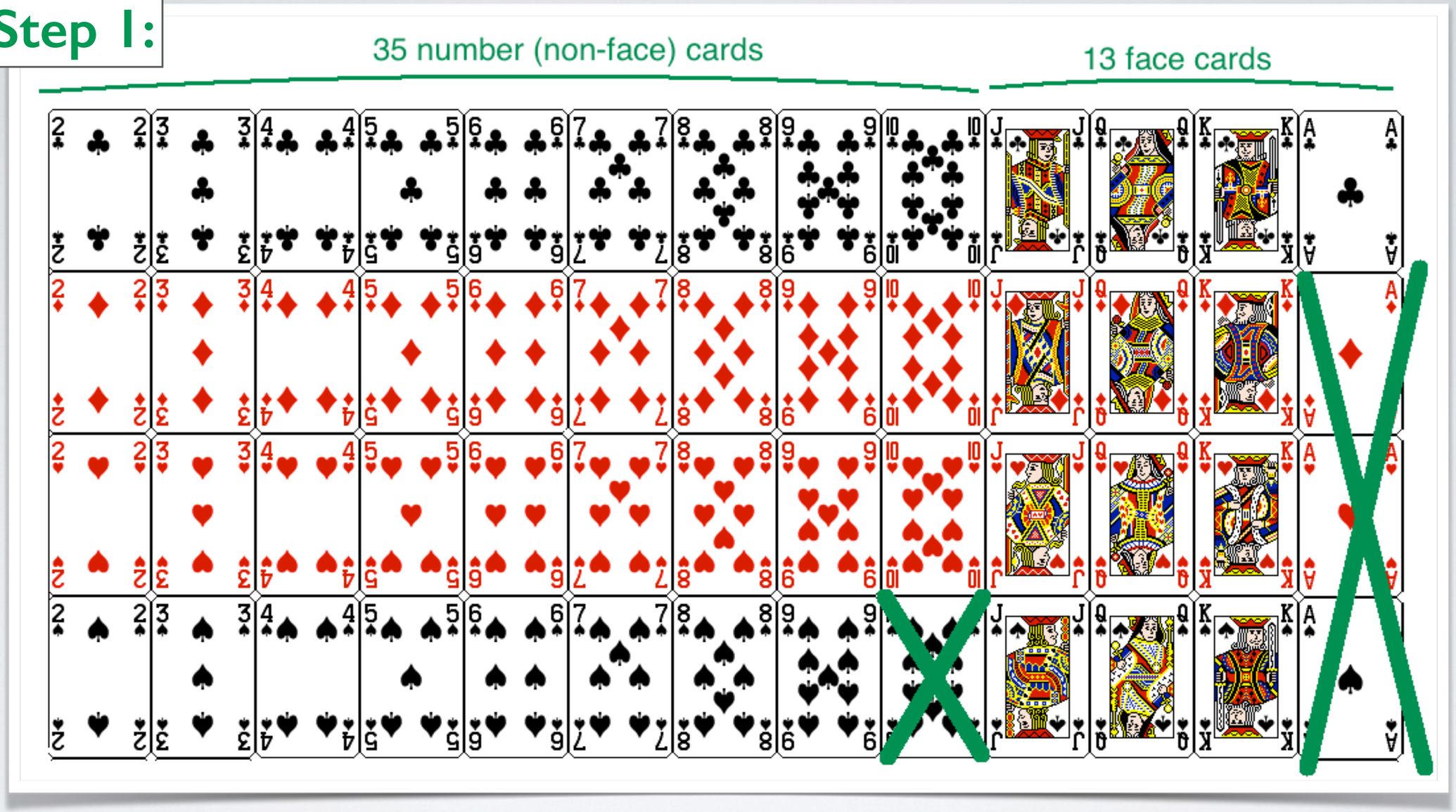


Image source: http://www.jfitz.com/cards/

simulation scheme

[use a deck of playing cards to simulate this experiment]

1. face card: not promoted, non-face card: promoted
   - set aside the jokers, consider aces as face cards
   - take out 3 aces → exactly 13 face cards left in the deck (face cards: A, K, Q, J)
   - take out a number card → 35 number (non-face) cards left in the deck (number cards: 2-10)
2. shuffle the cards, deal into two groups of size 24, representing males and females

Step 2:

Shuffle and split into two groups of 24

(males and females)

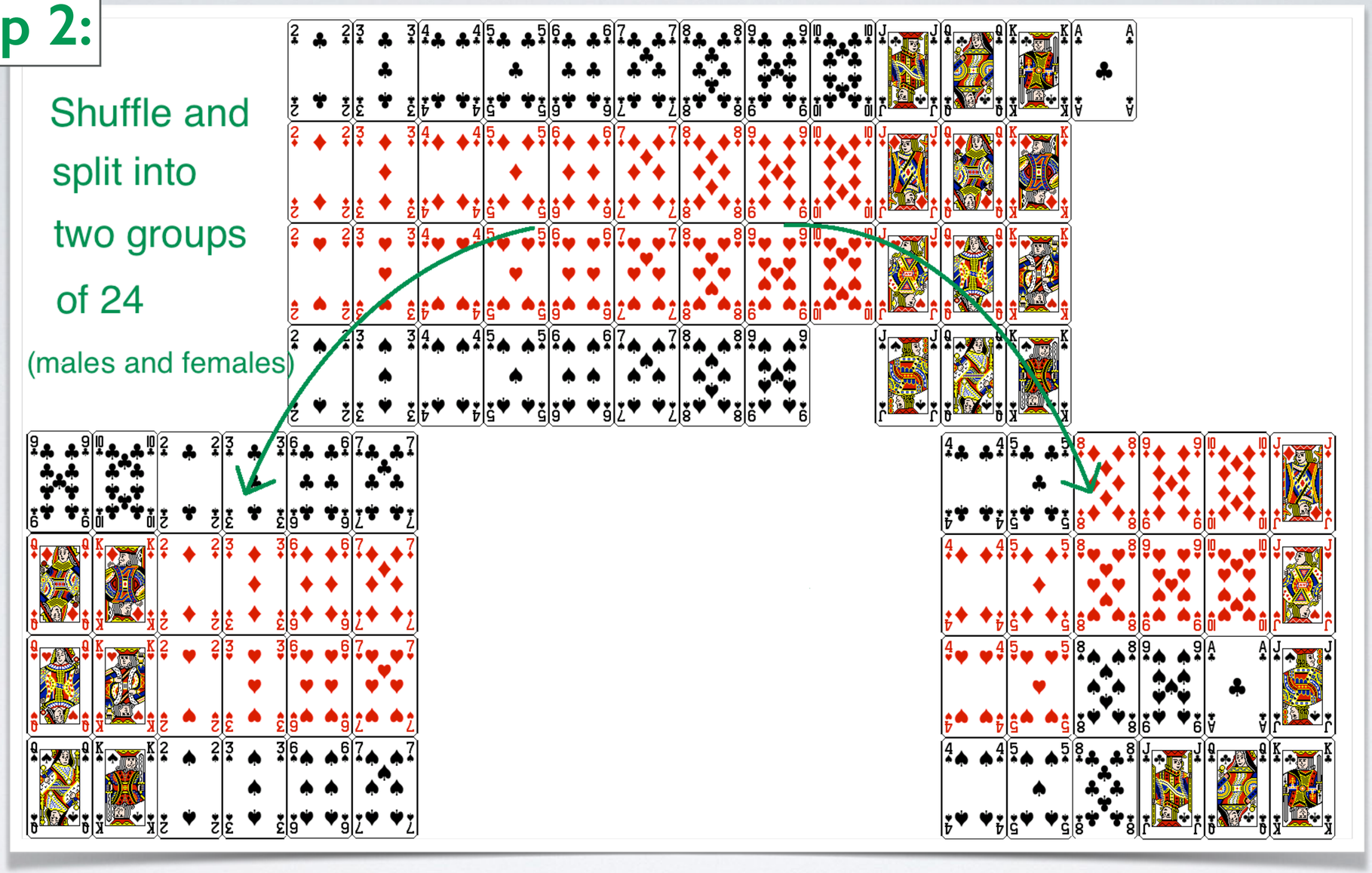


Image source: http://www.jfitz.com/cards/

simulation scheme

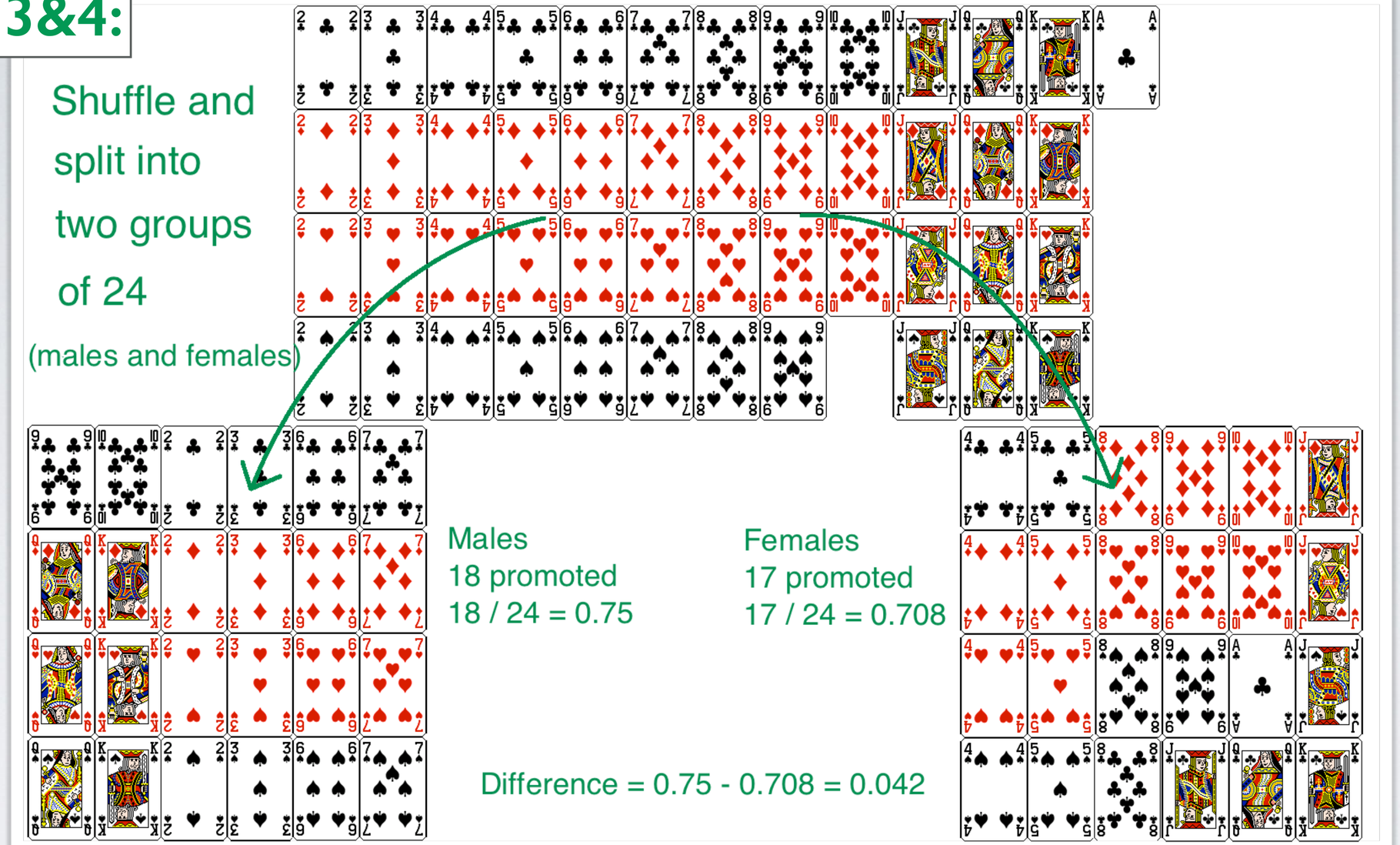
[use a deck of playing cards to simulate this experiment]

1. face card: not promoted, non-face card: promoted
   - set aside the jokers, consider aces as face cards
   - take out 3 aces → exactly 13 face cards left in the deck (face cards: A, K, Q, J)
   - take out a number card → 35 number (non-face) cards left in the deck (number cards: 2-10)
2. shuffle the cards, deal into two groups of size 24, representing males and females
3. count how many number cards are in each group (representing promoted files)
4. calculate the proportion of promoted files in each group, take the difference (male - female), and record this value

Steps 3&4:

Shuffle and split into two groups of 24

(males and females)

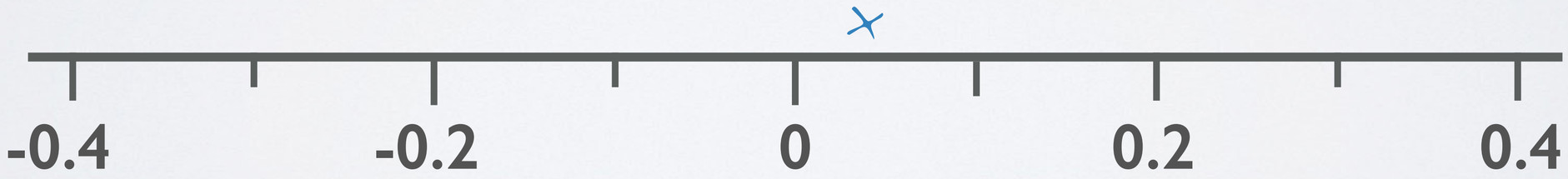


Males
18 promoted
18 / 24 = 0.75

Females
17 promoted
17 / 24 = 0.708

Difference = 0.75 - 0.708 = 0.042

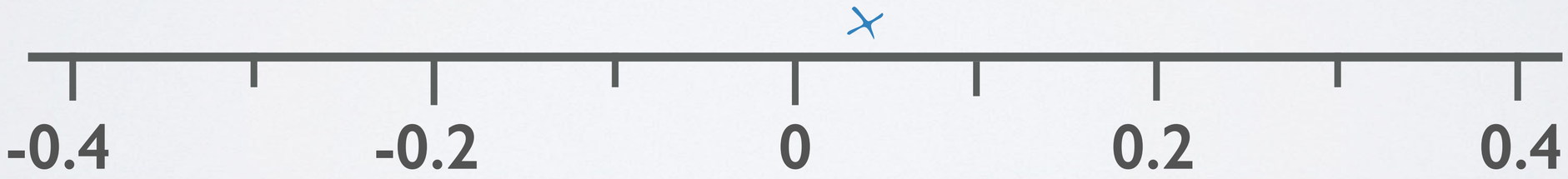
Image source: http://www.jfitz.com/cards/


-0.4
-0.2
0
0.2
0.4

simulation scheme

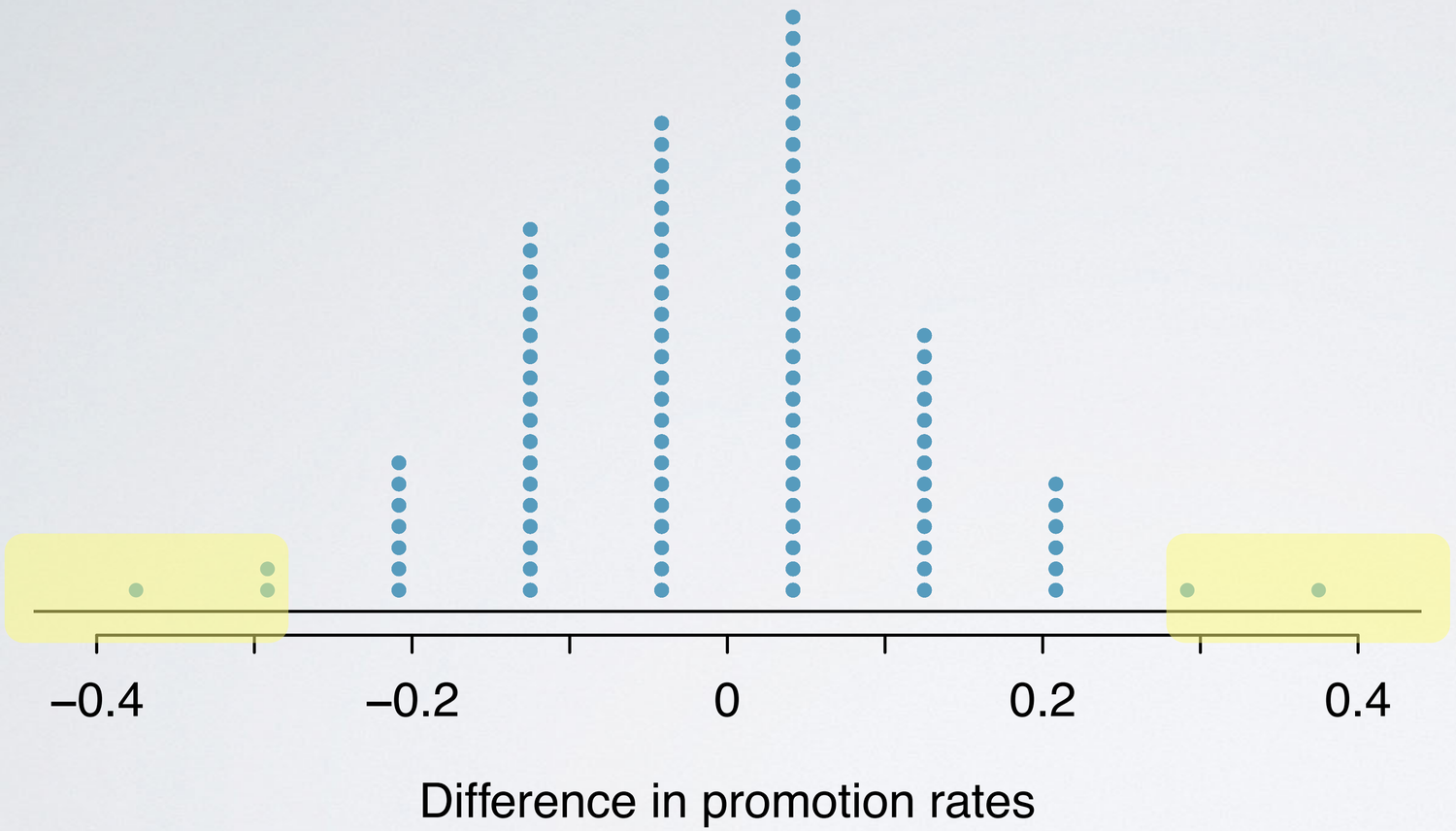
[use a deck of playing cards to simulate this experiment]

1. face card: not promoted, non-face card: promoted
   - set aside the jokers, consider aces as face cards
   - take out 3 aces → exactly 13 face cards left in the deck (face cards: A, K, Q, J)
   - take out a number card → 35 number (non-face) cards left in the deck (number cards: 2-10)
2. shuffle the cards, deal into two groups of size 24, representing males and females
3. count how many number cards are in each group (representing promoted files)
4. calculate the proportion of promoted files in each group, take the difference (male - female), and record this value
5. repeat steps 2 - 4 many times


-0.4
-0.2
0
0.2
0.4

making a decision

- results from the simulations look like the data → the difference between the proportions of promoted files between males and females was due to chance (promotion and gender are independent)
- results from the simulations do not look like the data → 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)


–0.4
–0.2
0
0.2
0.4
Difference in promotion rates

summary

- set a null and an alternative hypothesis
- simulate the experiment assuming that the null hypothesis is true
- evaluated the probability of observing an outcome at least as extreme as the one observed in the original data

p-value

- and if this probability is low, reject the null hypothesis in favor of the alternative