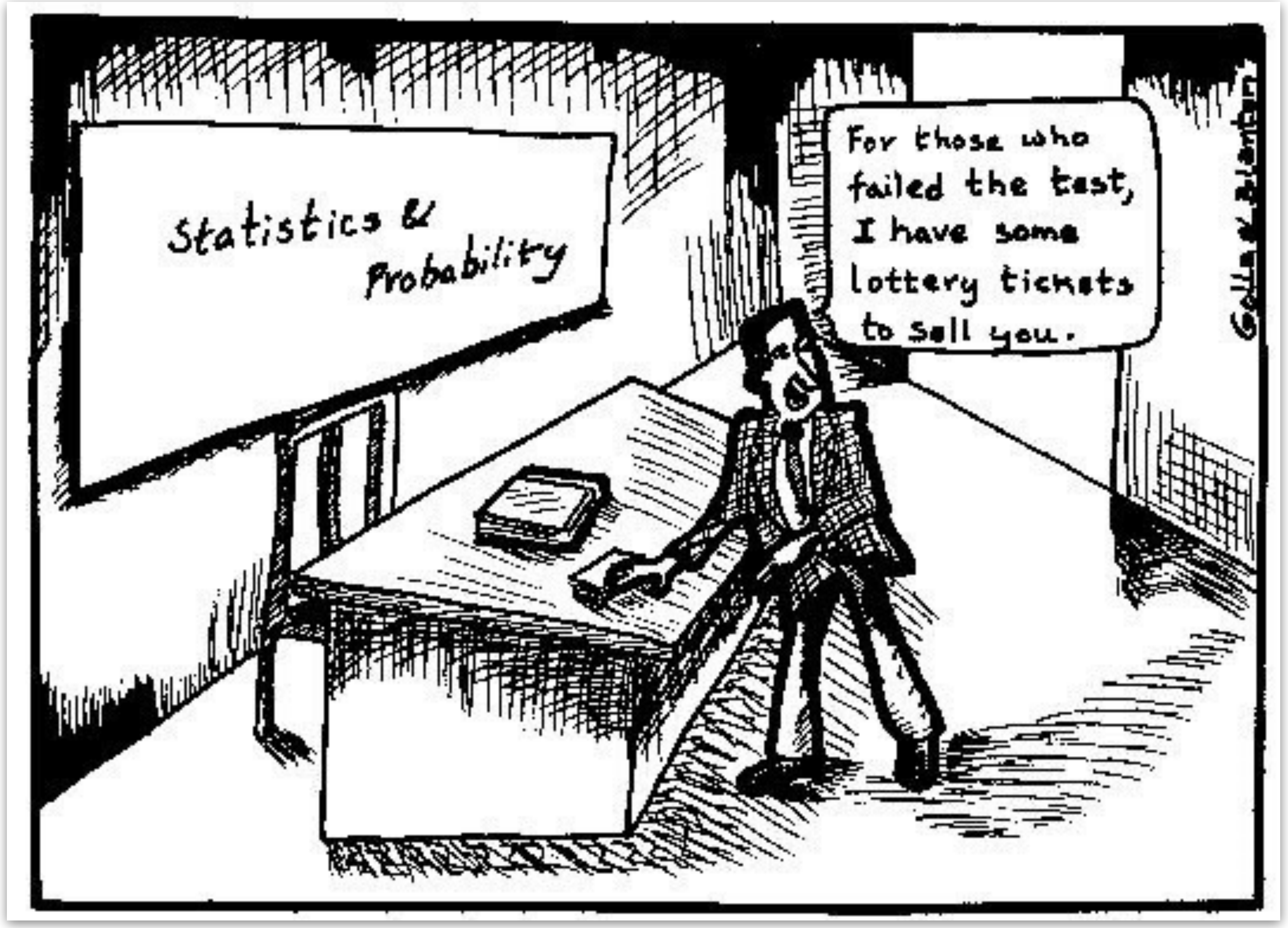
PSYCH 252: Statistical Methods for Behavioral and Social Sciences

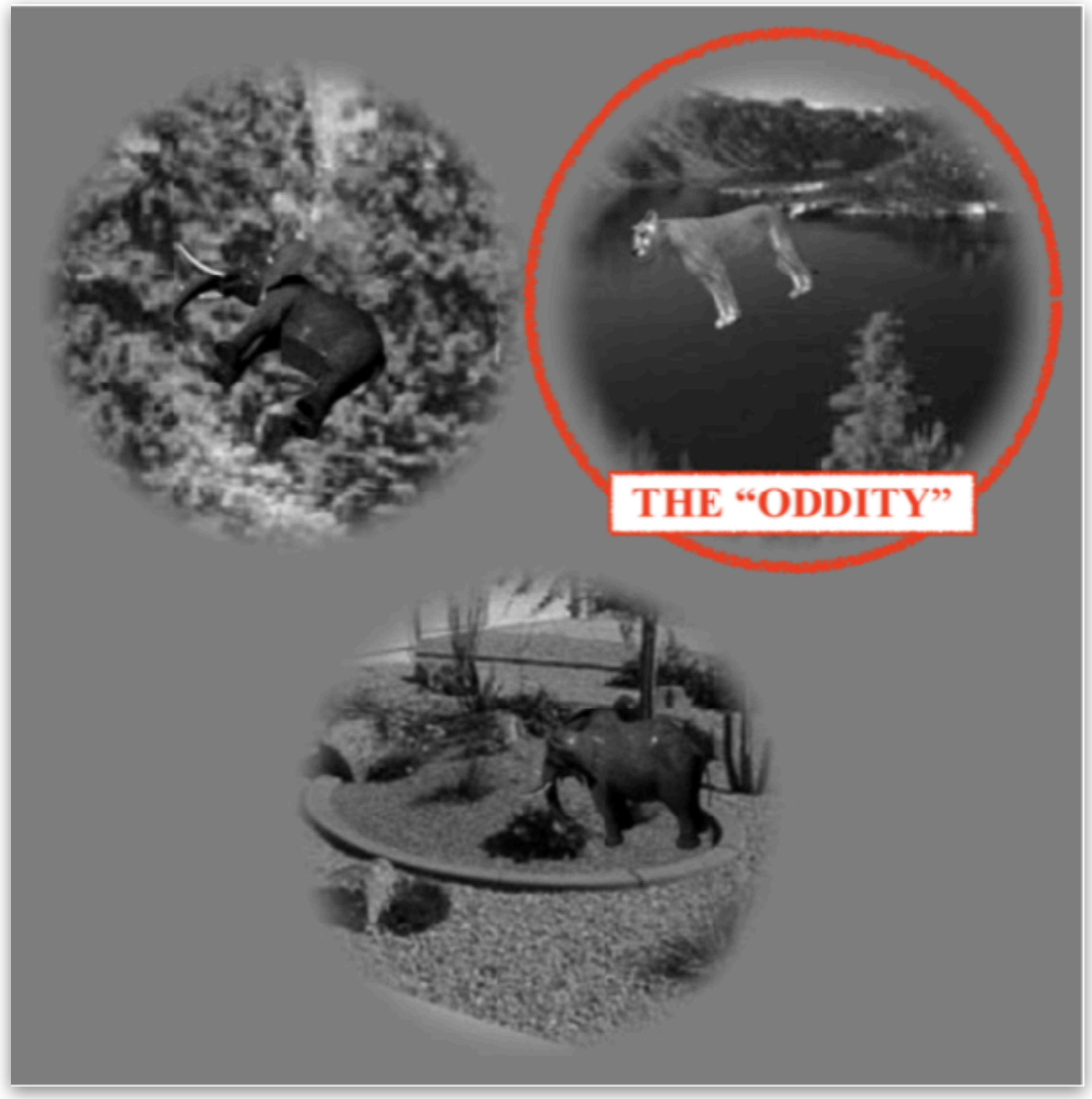
Probability



01/16/2026

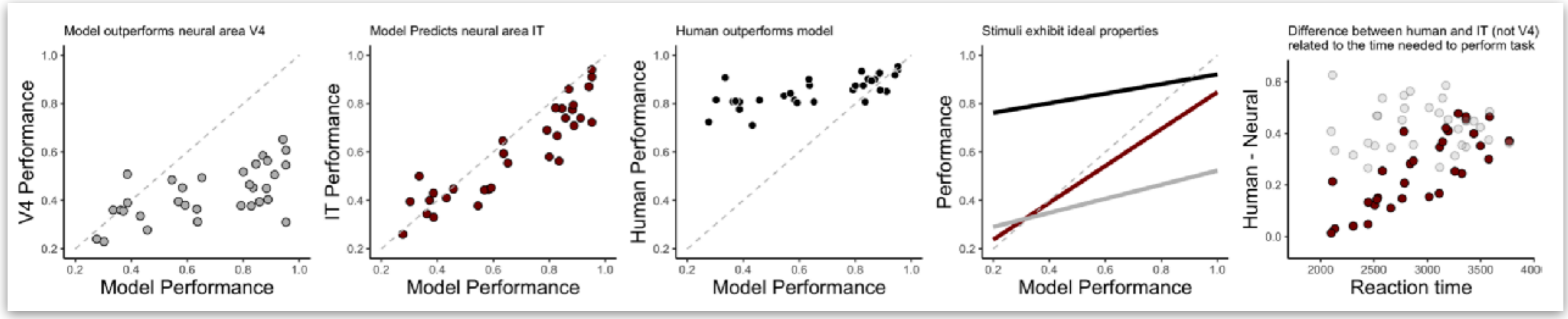
Logistics

Homework 2



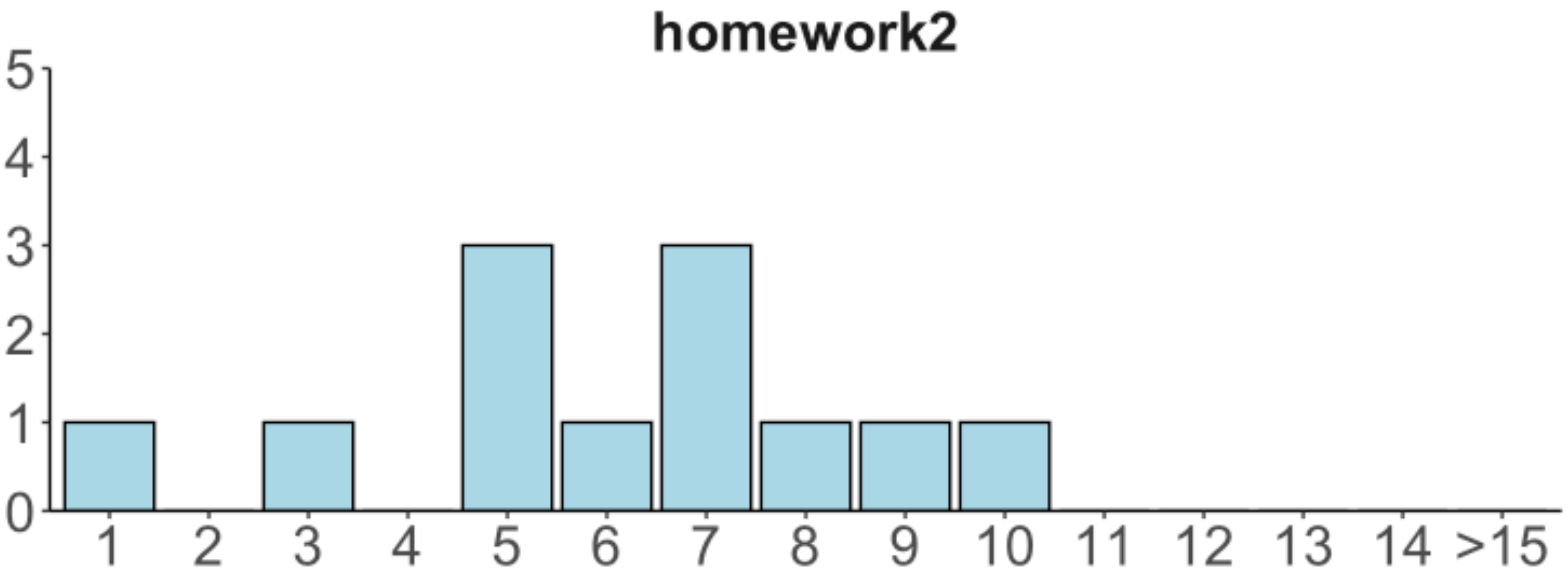
cool experiment

messy dataset



neat plot!

Homework 2



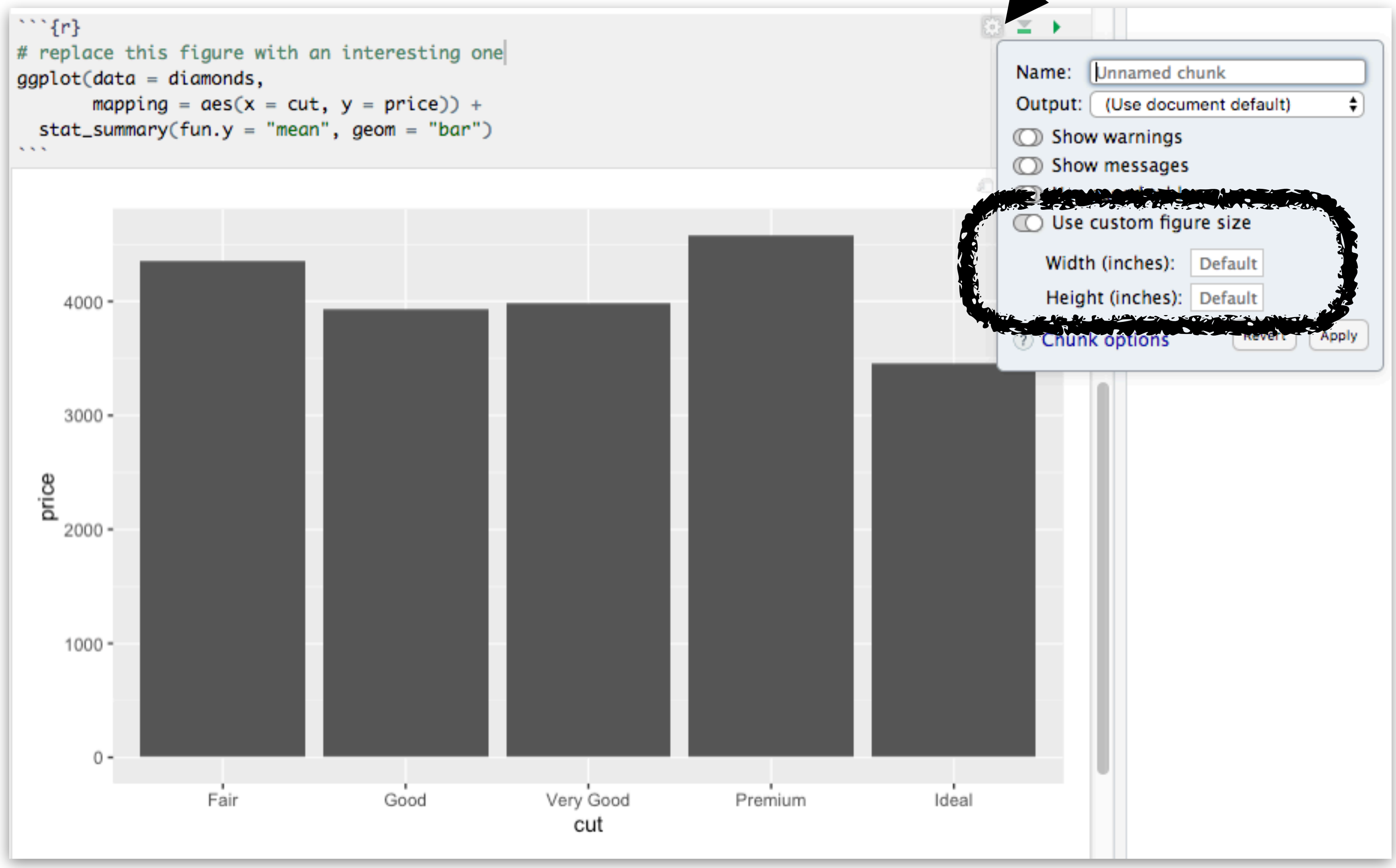
how long it took people in the past

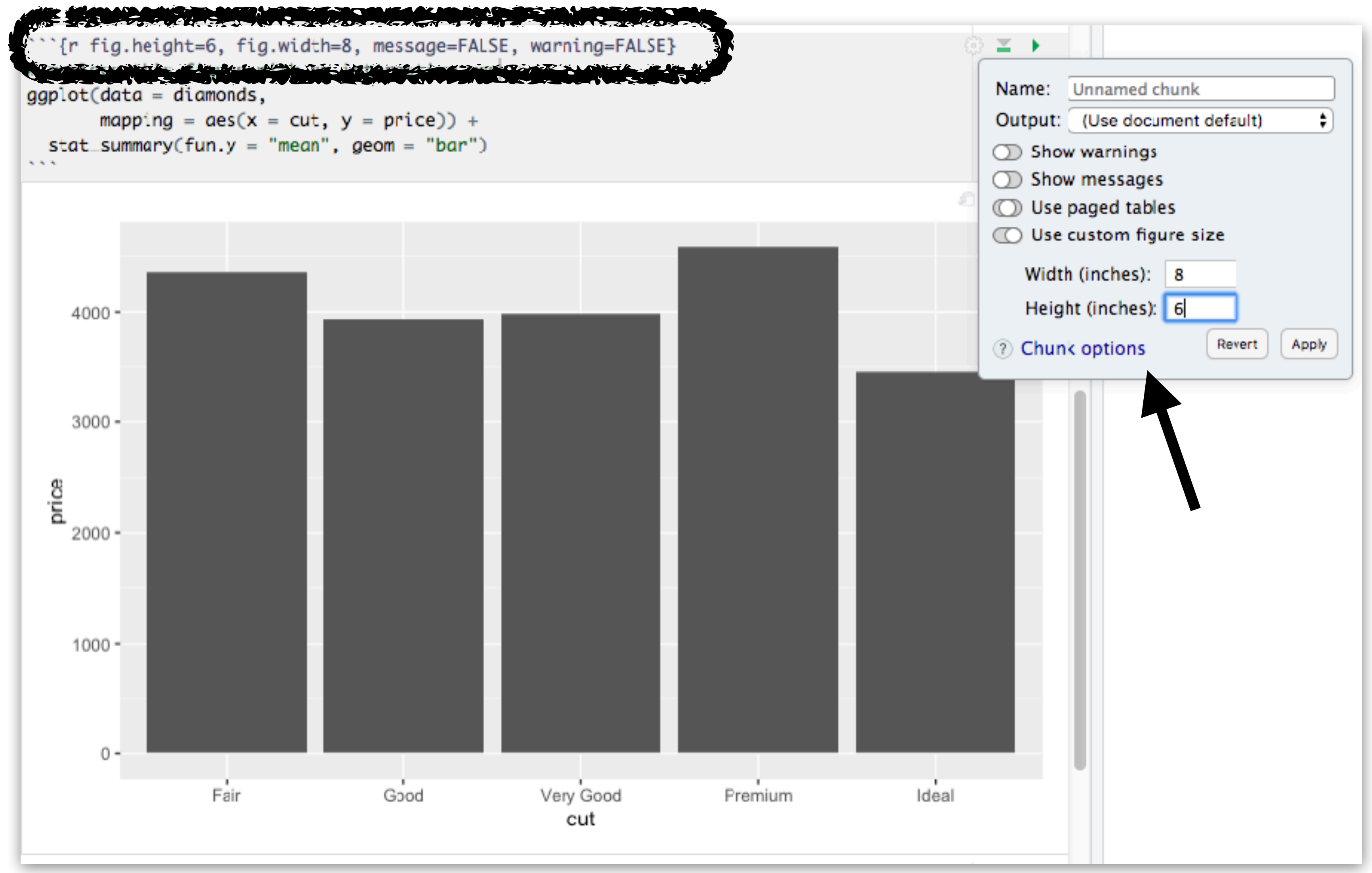
Homework 2

- Due **Thursday 22nd, at 8pm**
- Don't wait until the very last moment to knit your RMarkdown file into a pdf. It may not compile and debugging takes time ...
- You can upload earlier versions of your homework on Canvas and still update until the deadline.
- Get and give help via **Ed Discussion**!
- We encourage you to work in groups!

Homework 2

You can adjust the figure size in your output by using the code chunk options





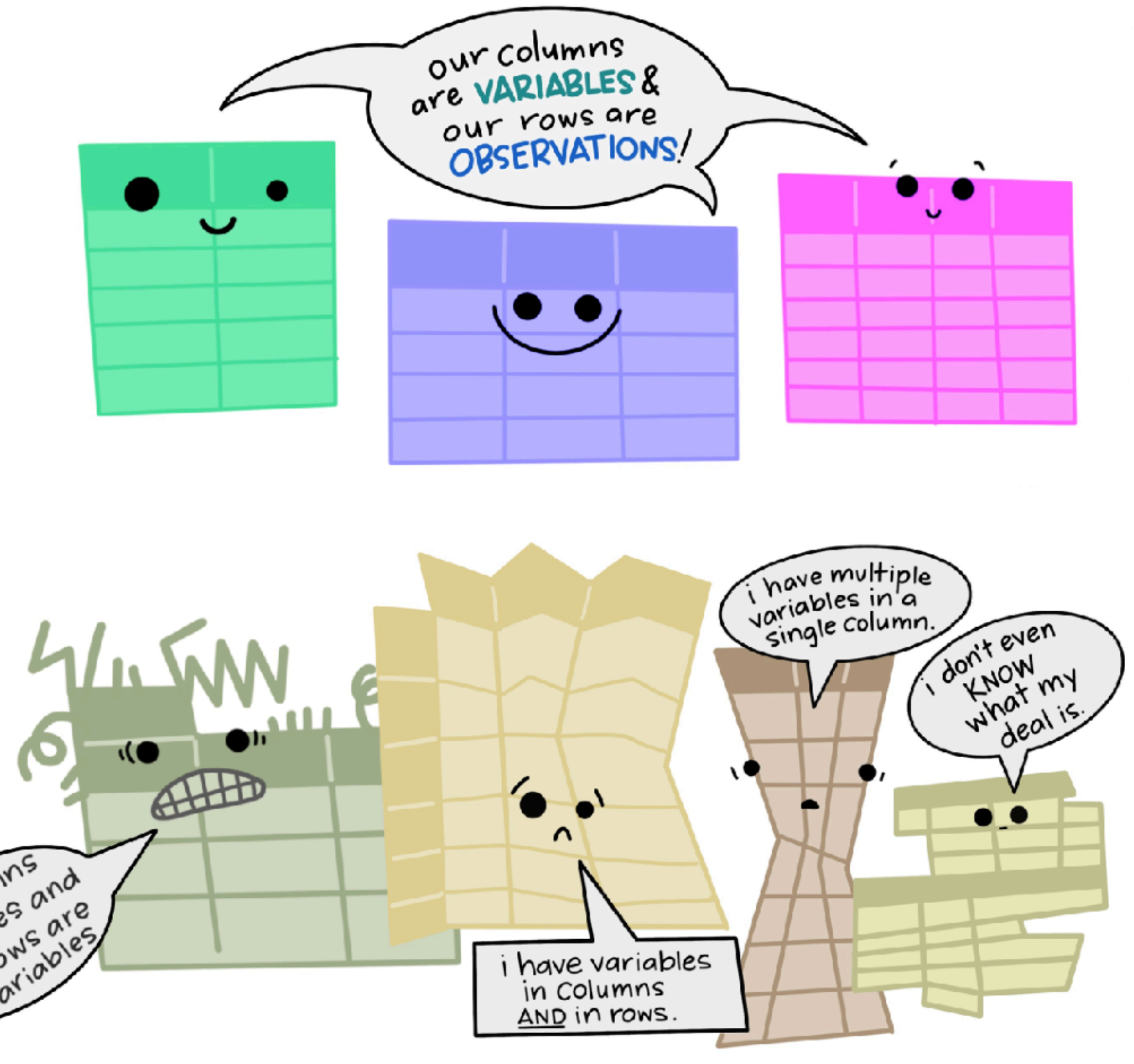
https://yihui.org/knitr/options/

Quick recap

The standard structure of tidy data means that "tidy datasets are all alike..."

"...but every messy dataset is messy in its own way."

—HADLEY WICKHAM



Using functions with tidyverse verbs

```
df.starwars %>%
  select(where(fn = is.numeric))
```

my recommendation

- `fn = is.numeric`
- `fn = "is.numeric"`
- `fn = function(x){is.numeric(x)}`
- `fn = ~ is.numeric(.)`

flexible, short, works well with other verbs we'll learn about later

- `fn = ~ !is.numeric(.)`

select all variables that are not numeric

```
df.starwars %>%
  mutate(across(.cols = c(height, mass, birth_year),
                .fns = list(z = scale,
                            centered = ~ scale(., scale = FALSE)))) %>%
  select(name, contains("height"), contains("mass"), contains("birth_year"))
```

5.4.1 Summarizing data

Let's first load the `starwars` data set again:

```
df.starwars = starwars
```

A particularly powerful way of interacting with data is by grouping and summarizing it. `summarize()` returns a single value for each summary that we ask for:

```
df.starwars %>%
  summarize(height_mean = mean(height, na.rm = T),
            height_max = max(height, na.rm = T),
            n = n())
```

```
# A tibble: 1 × 3
  height_mean height_max     n
        <dbl>      <int> <int>
1        175.        264    87
```

Here, I computed the mean height, the maximum height, and the total number of observations (using the function `n()`). Let's say we wanted to get a quick sense for how tall starwars characters from different species are. To do that, we combine grouping with summarizing:

```
df.starwars %>%
  group_by(species) %>%
  summarize(height_mean = mean(height, na.rm = T))
```

```
df.responses %>%
  left_join(df.stimuli %>%
              select(index, color),
            by = "index") %>%
  group_by(color) %>%
  summarize(response_mean = mean(response))
```

5.4.2.1 `pivot_longer()` and `pivot_wider()`

Let's first generate a data set that is *not* tidy.

```
# construct data frame
df.reshape = tibble(participant = c(1, 2),
                    observation_1 = c(10, 25),
                    observation_2 = c(100, 63),
                    observation_3 = c(24, 45)) %>%
  print()
```

```
# A tibble: 2 × 4
  participant observation_1 observation_2 observation_3
        <dbl>         <dbl>         <dbl>         <dbl>
1           1            10           100            24
2           2            25            63            45
```

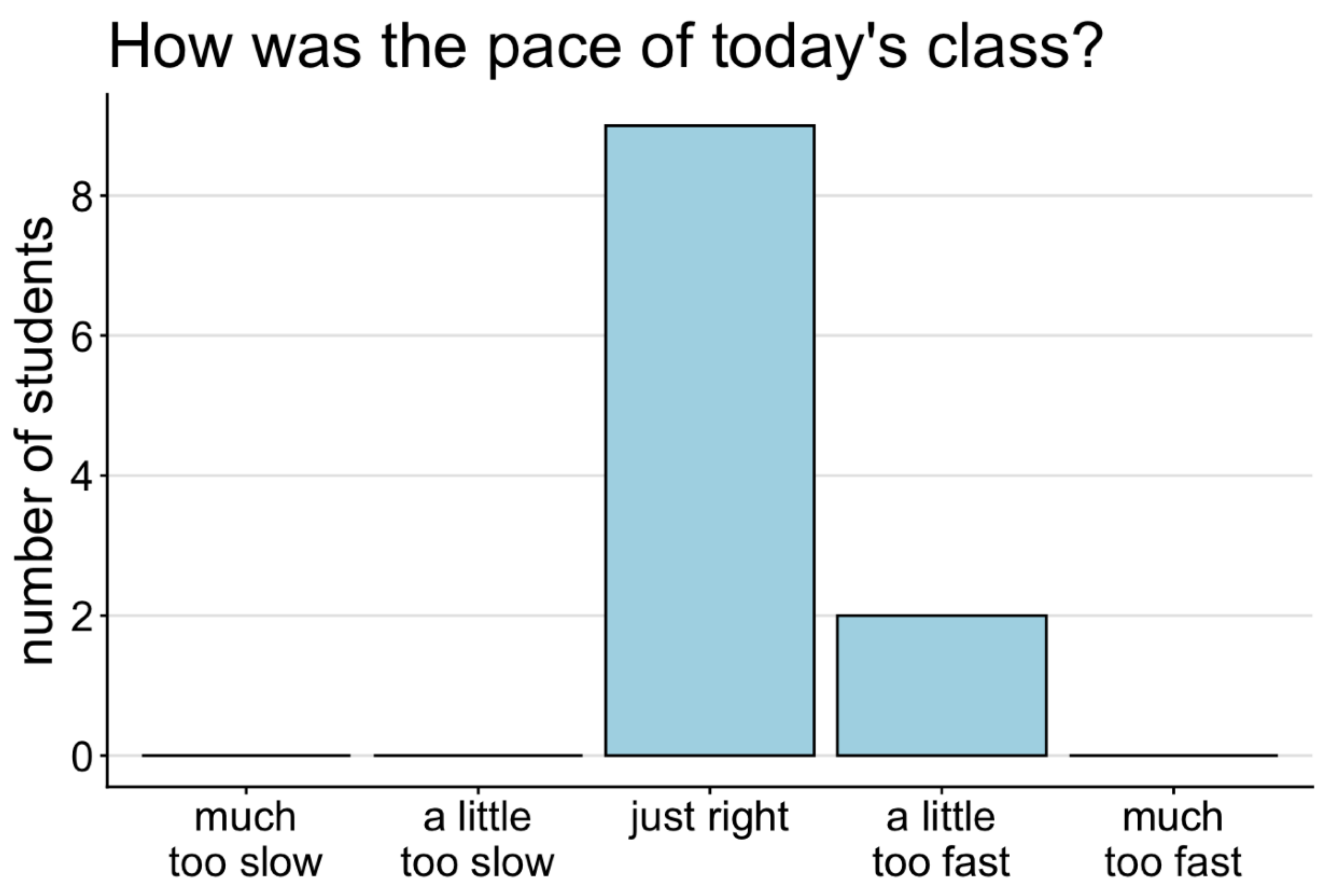
Here, I've generated data from two participants with three observations. This data frame is not tidy since each row contains more than a single observation. Data frames that have one row per participant but many observations are called *wide* data frames.

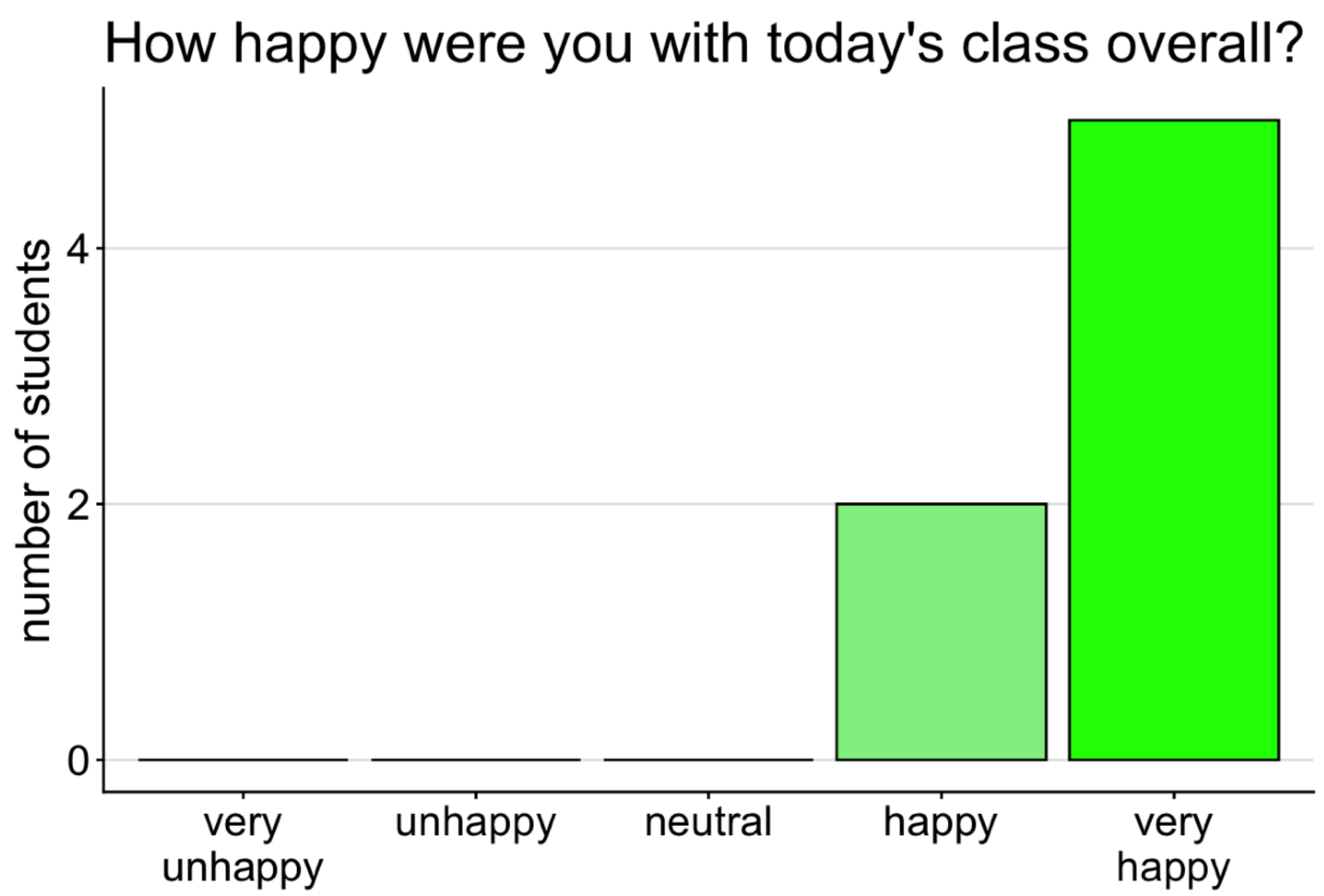
We can make it tidy using the `pivot_longer()` function.

```
df.reshape.long = df.reshape %>%
  pivot_longer(cols = contains("observation"),
               names_to = "index",
               values_to = "rating") %>%
  arrange(participant) %>%
  print()
```

Your feedback

Your feedback





Music to my ears when Prof. Tobi said he would the solutions from the practice problems to canvas. That will be helpful for stepping through on my own time. I still think the pace was a little too fast for seeing the solutions after we stepped through them. What I did appreciate is the stepping through the solution out loud.

Quiz time!

Outline

- Introduction to probability / Recap
  - Motivation
  - Counting possibilities
  - Clue guide to probability
  - Understanding Bayes' Rule
  - Getting Bayes' right matters!
  - Building a Bayesis

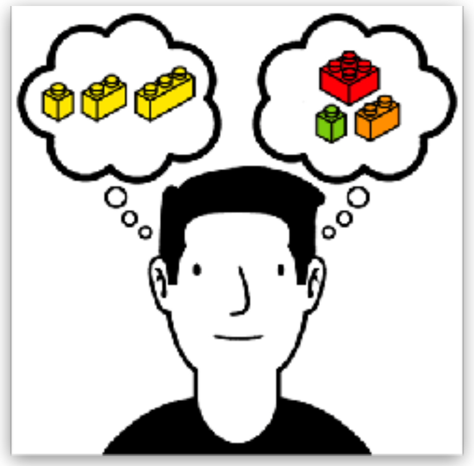
It's a mystery

XX

when I want to hide stuff from you

Motivation

What does statistics have to do with probability?



Theory

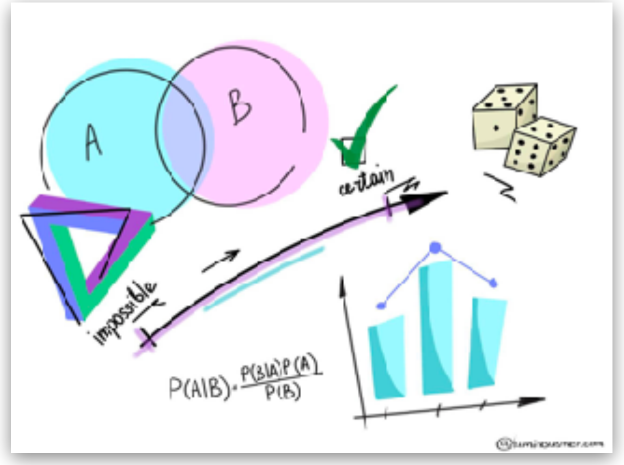
Our goal is to develop theories. In psychology, theories of how the mind works.

Prediction

Our theories need to make testable/falsifiable predictions.

Uncertainty

Because the domains that we are interested in are fundamentally uncertain (e.g. we want to say something about people generally but can only test a sample), we formulate and test these predictions using statistical models.

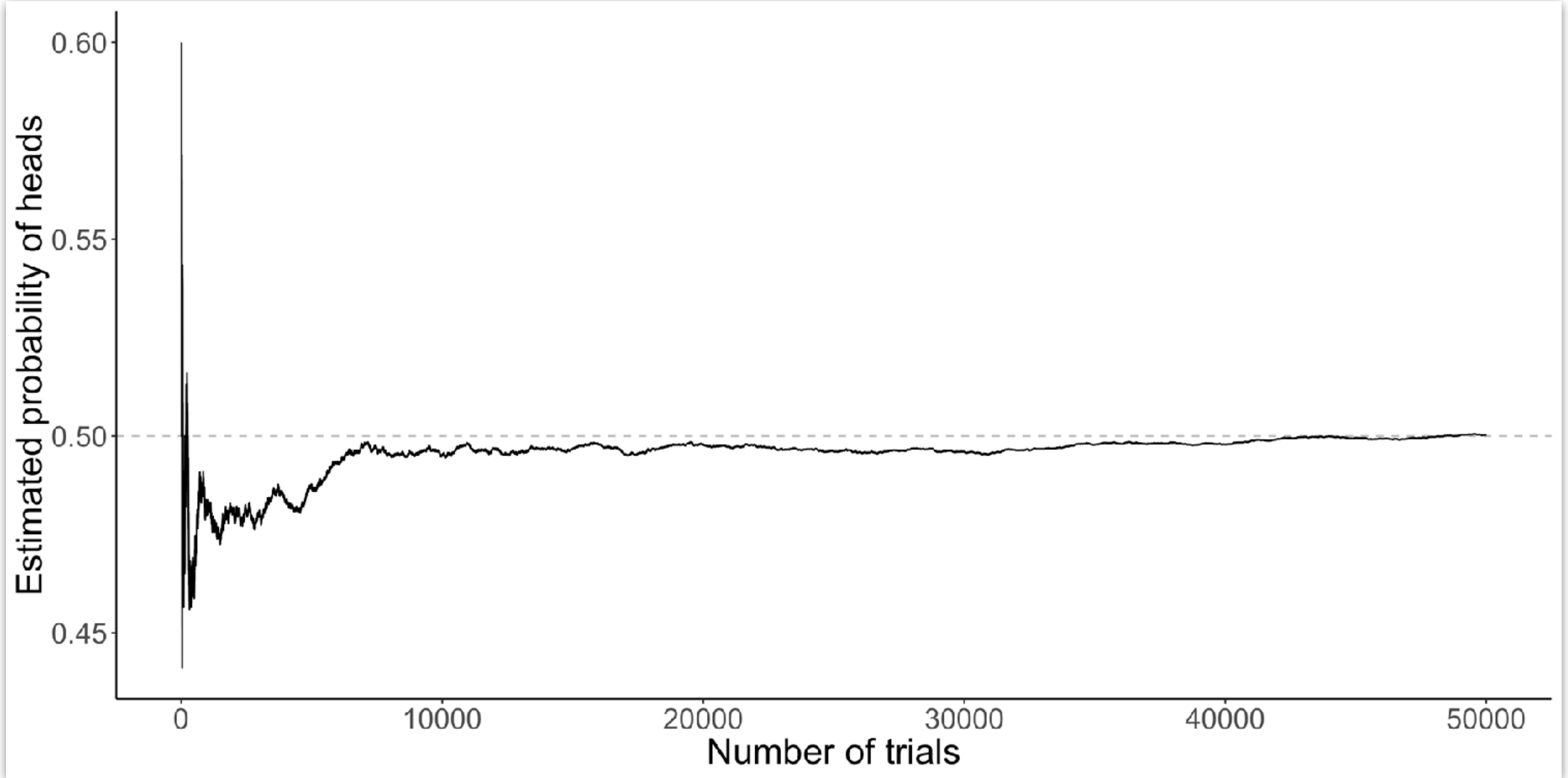


Probability

Probability theory is the formal language for dealing with uncertainty.

Frequentist interpretation

Probabilities = **long-range frequencies**



law of large numbers = empirical probability will approximate the true probability as the sample size increases

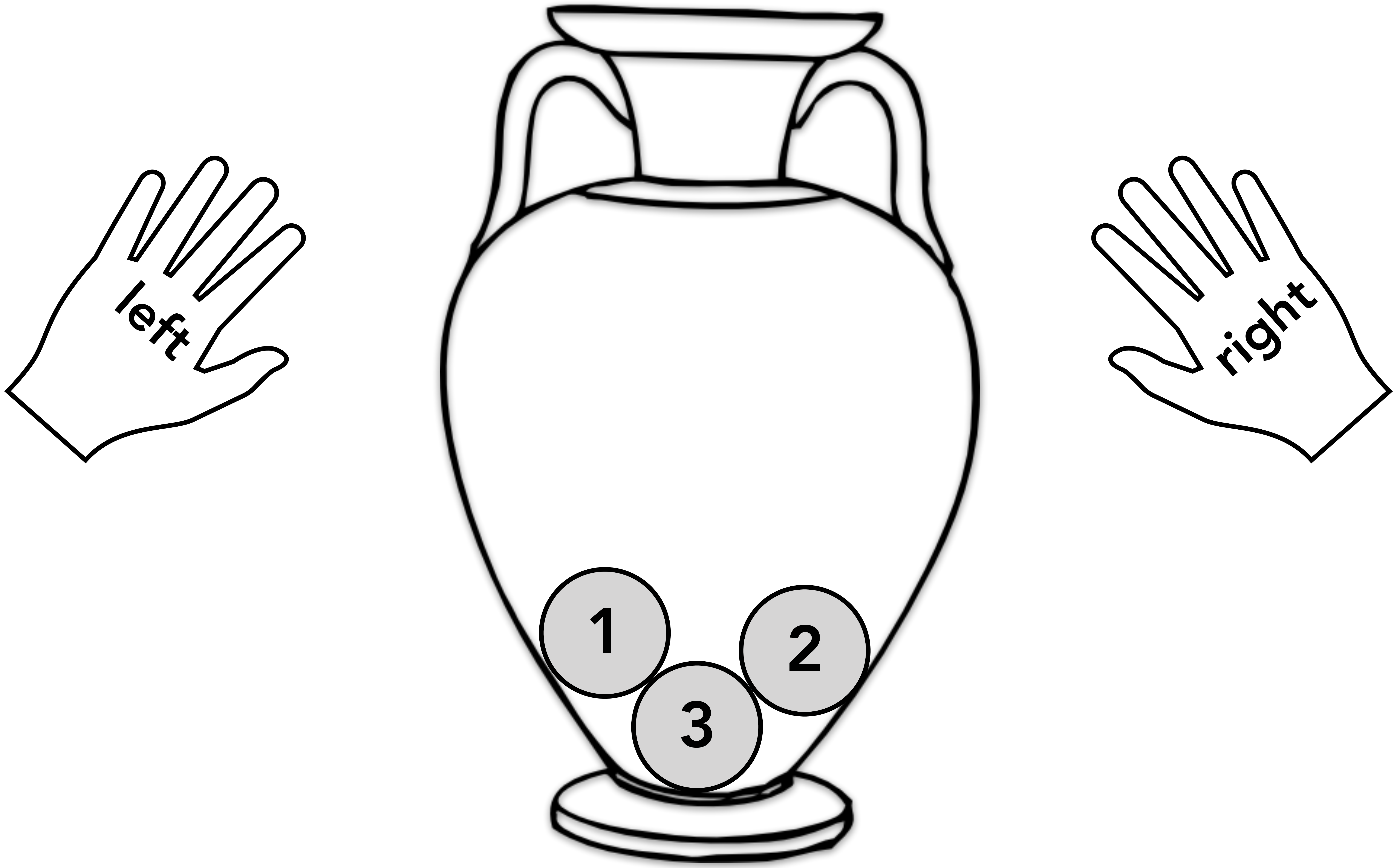
Subjective/Bayesian interpretation

Probabilities = **subjective degrees of belief**

- applies to events which may only happen once
- **"What's the probability that humans will land on Mars someday?"**
- probabilities are not a property of the world, but of a person's beliefs about the world
- at the heart of Bayesian data analysis

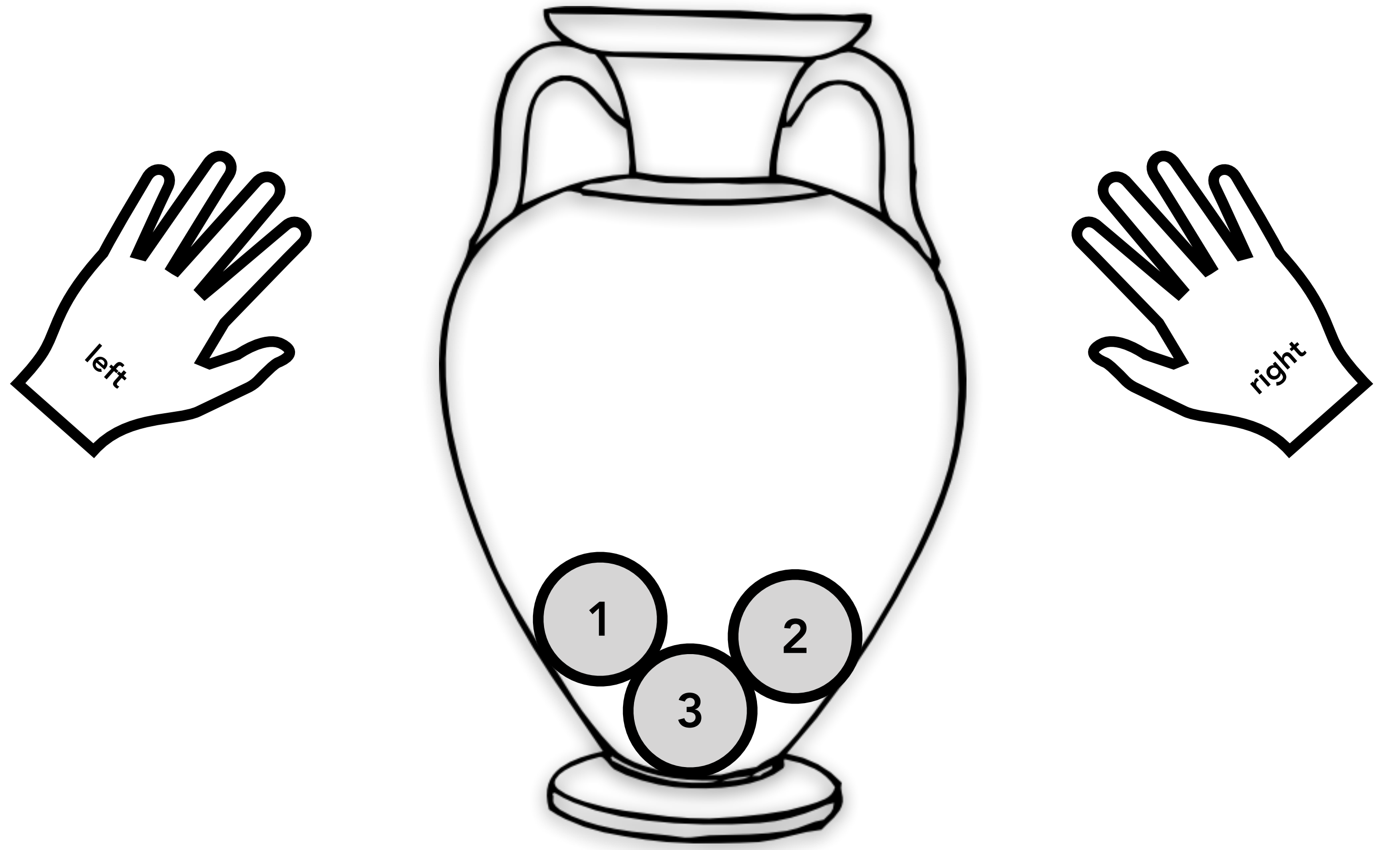
Counting possibilities

no stats class without urns!



Sampling **with** replacement

$$p(\text{left} = 1, \text{right} = 2) = ?$$

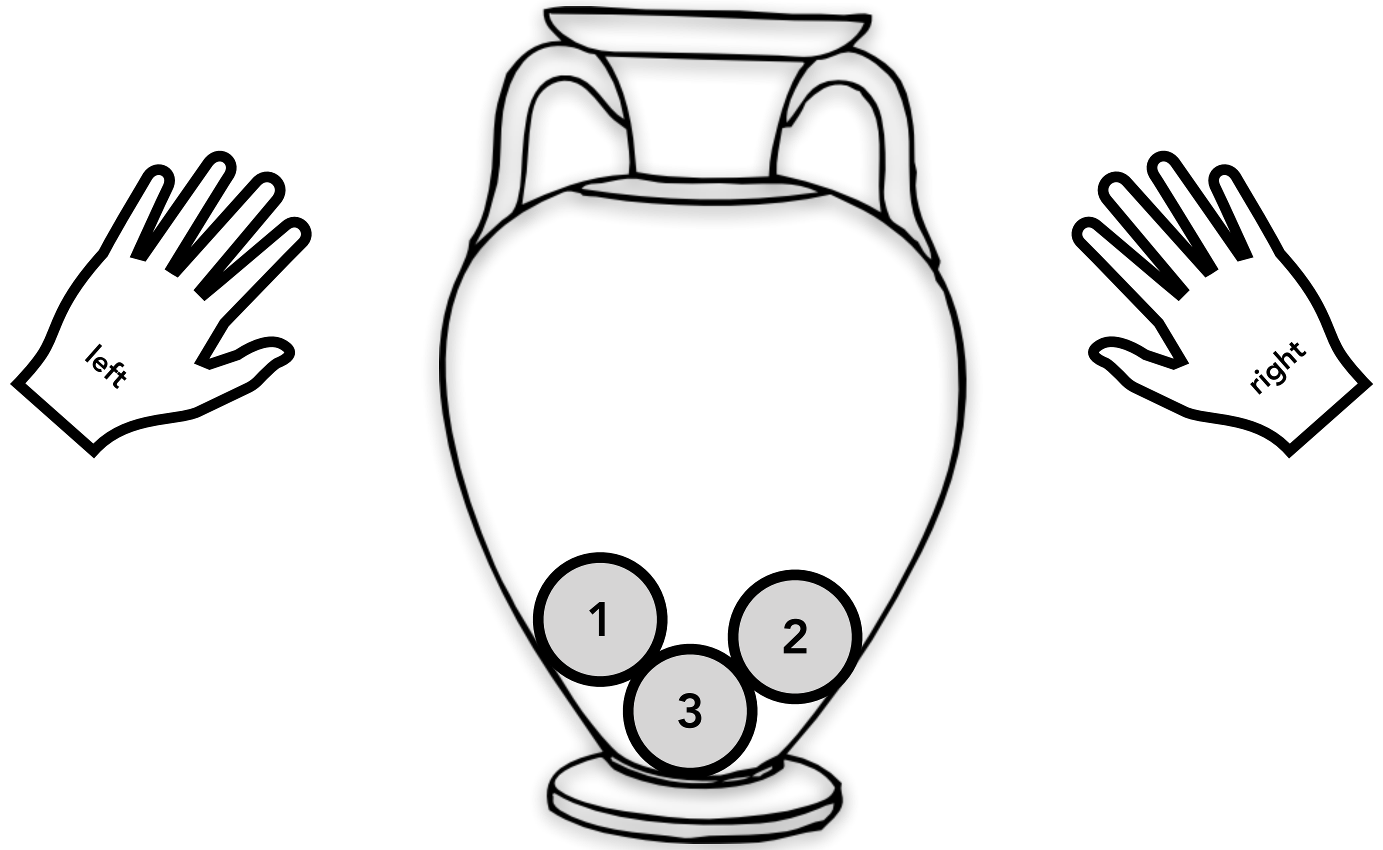


What is the probability that I first draw the 1 with my left hand, and then, after putting the 1 back into the urn again, draw the 2 with my right hand?

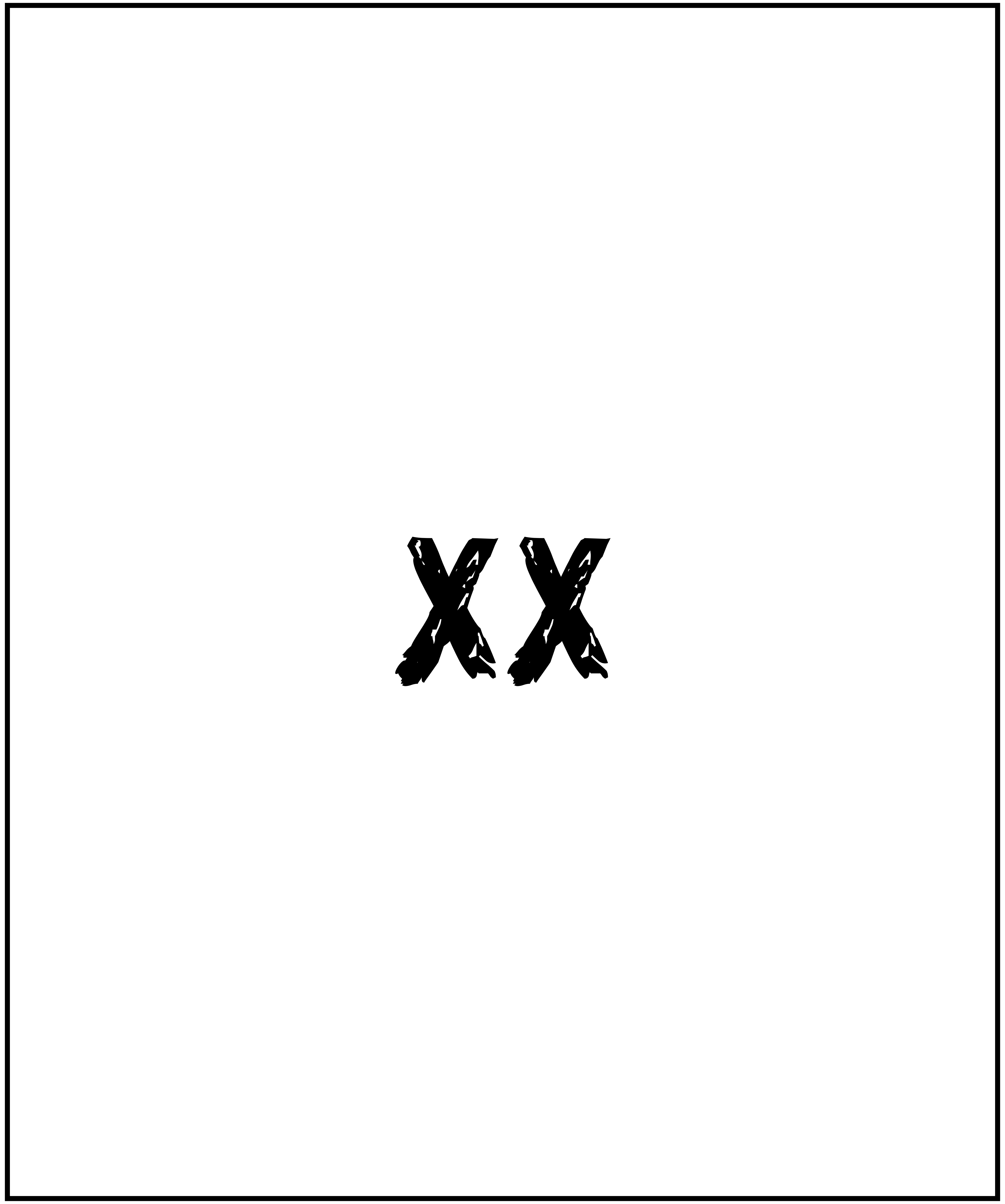
XX

Sampling **without** replacement

$$p(\text{left} = 1, \text{right} = 2) = ?$$

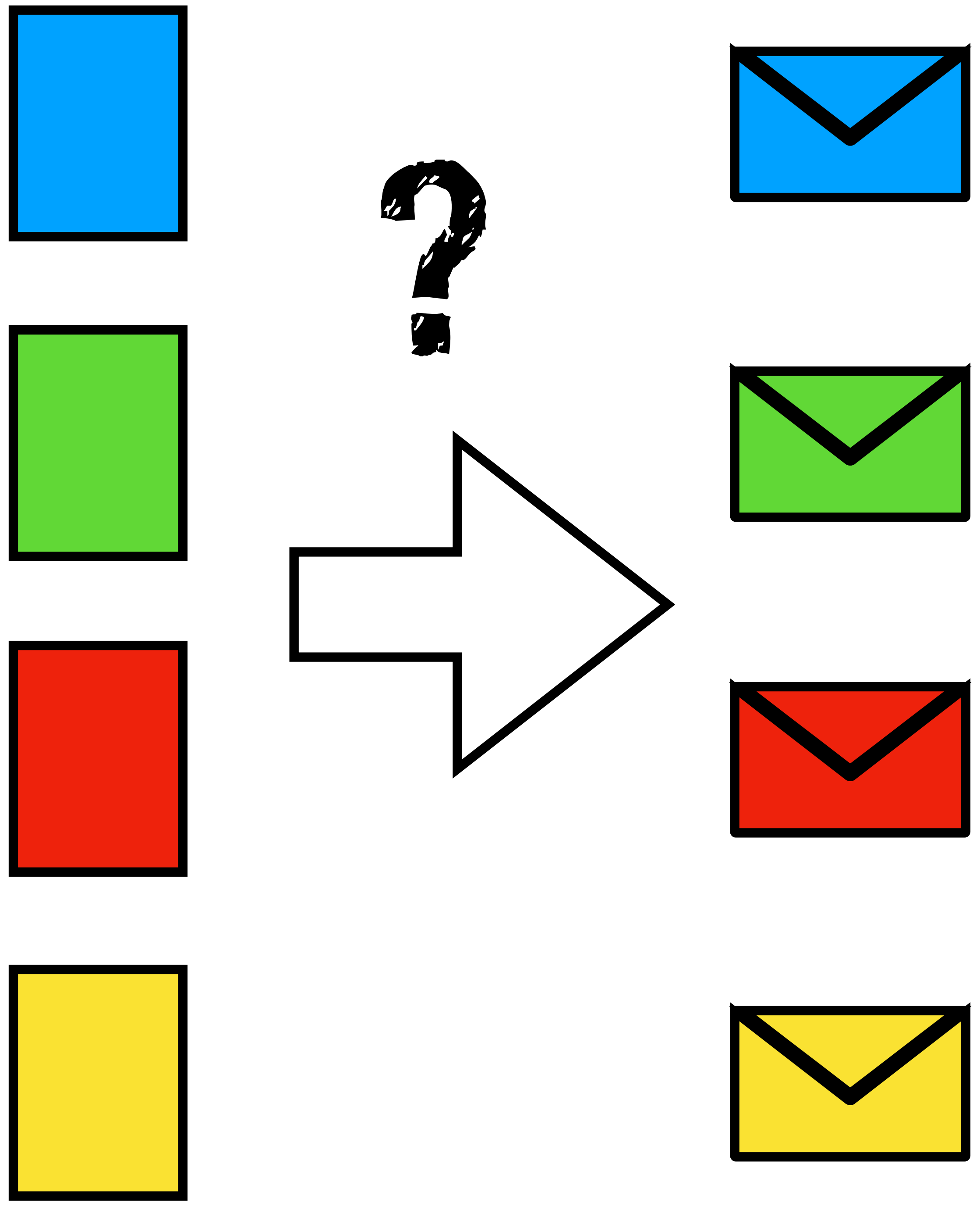


What is the probability that I first draw the 1 with my left hand, and then, without putting the 1 back into the urn, draw the 2 with my right hand?



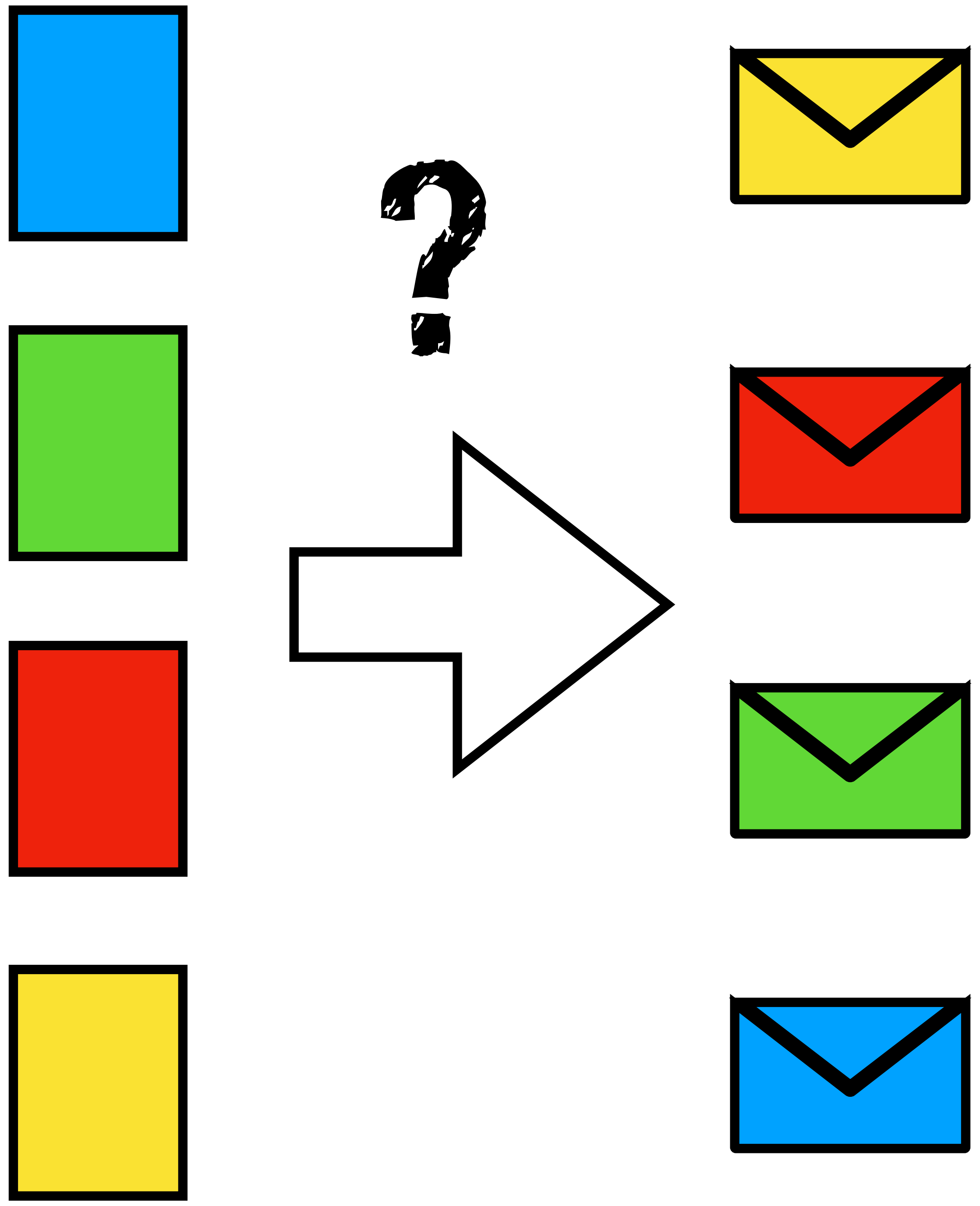
Random secretary

A secretary types four letters to four people and addresses the four envelopes. If he inserts the letters at random, each in a different envelope, what is the probability that exactly three letters will go into the right envelope?



Random secretary

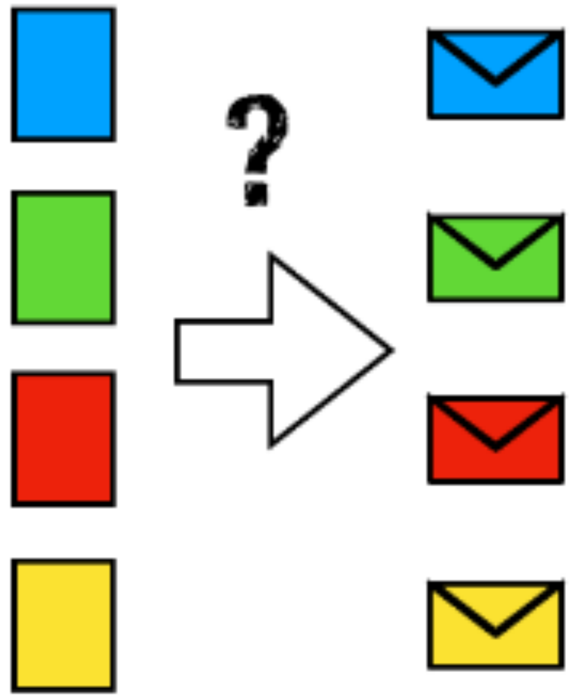
A secretary types four letters to four people and addresses the four envelopes. If he inserts the letters at random, each in a different envelope, what is the probability that exactly three letters will go into the right envelope?

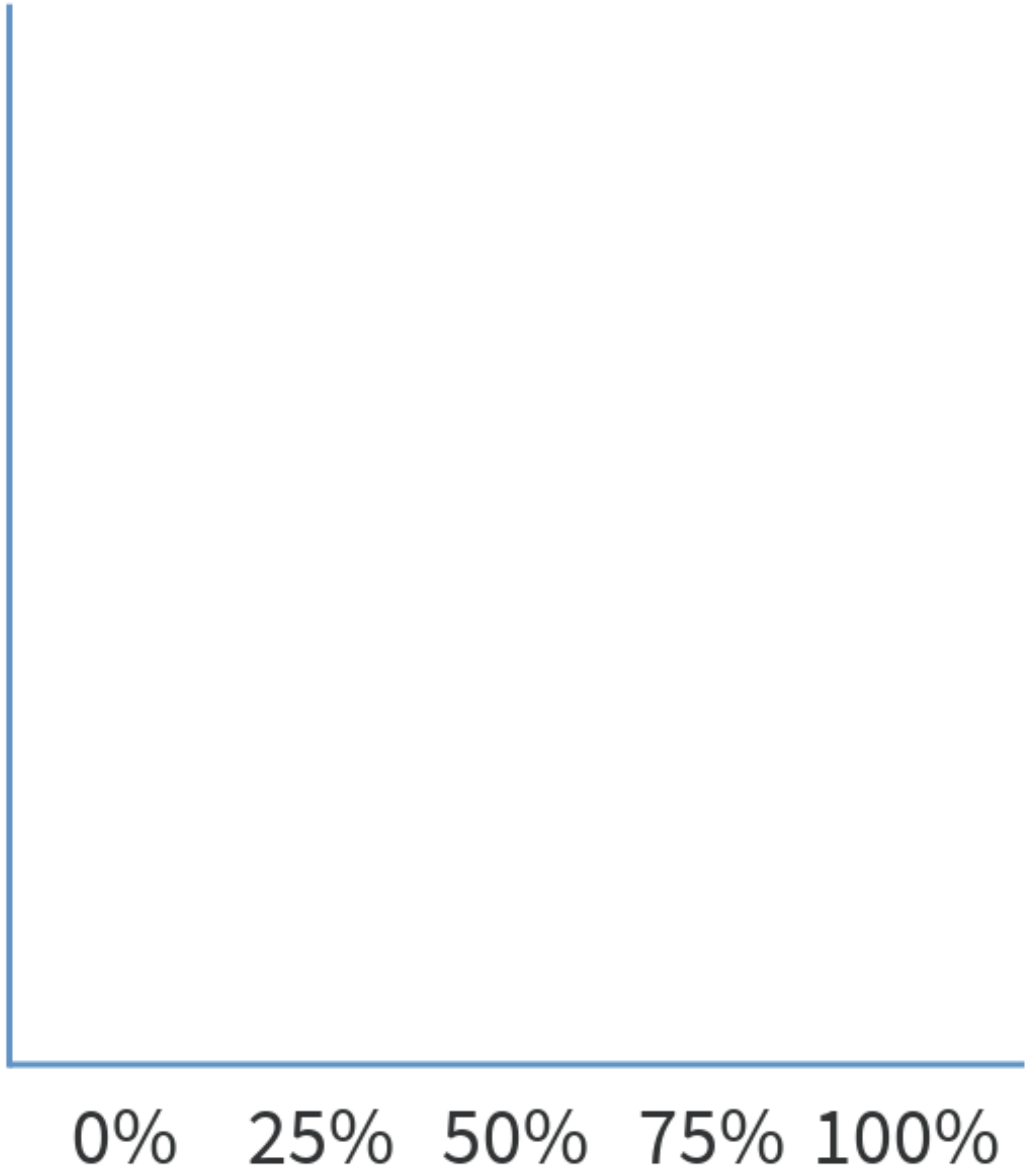


What is the probability that exactly three letters will go into the right envelope?

Random secretary

A secretary types four letters to four people and addresses the four envelopes. If he inserts the letters at random, each in a different envelope, what is the probability that exactly three letters will go into the right envelope?





Random secretary

Random secretary

XX

Naive definition of probability

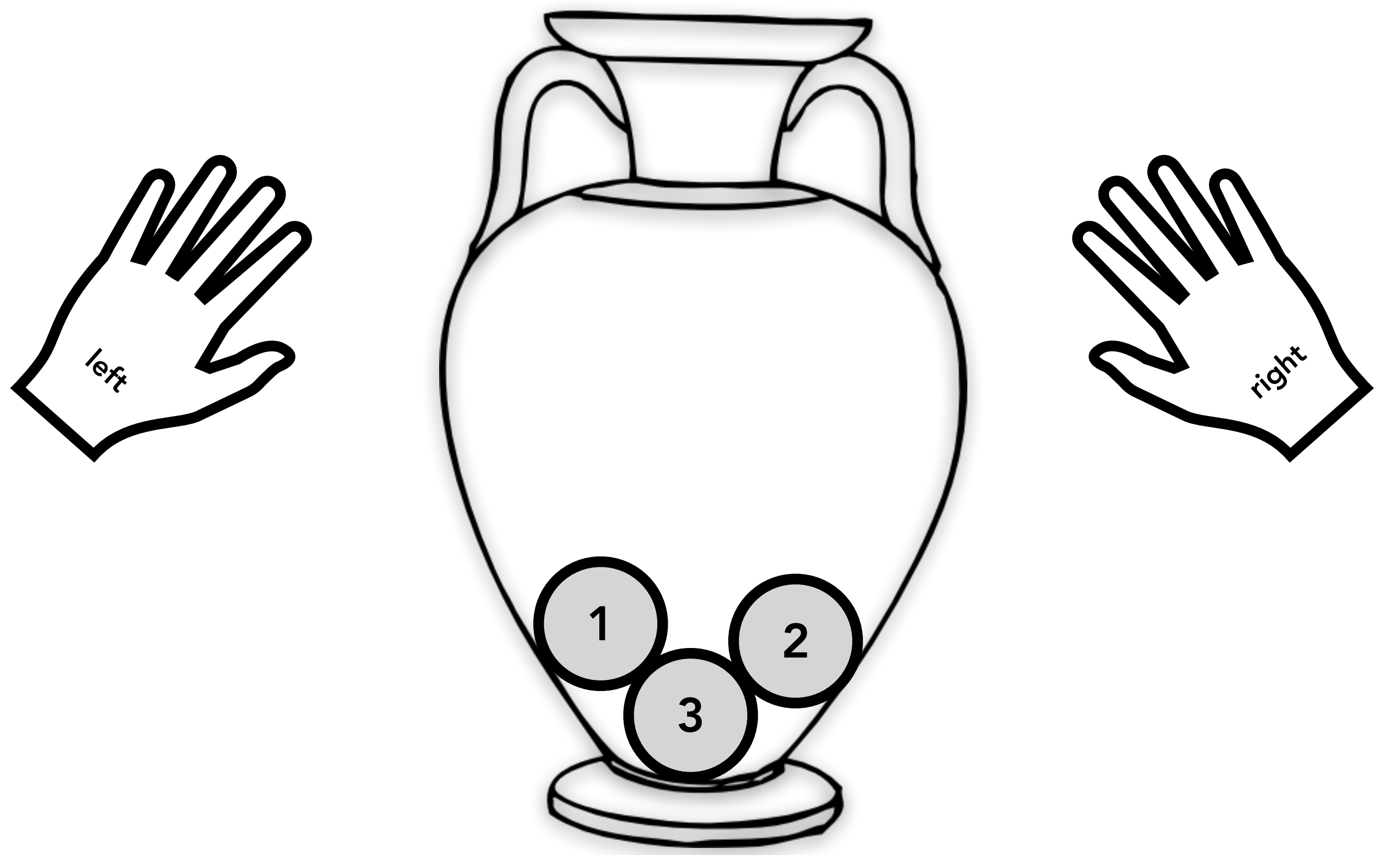
$$P_{\text{naive}}(A) = \frac{\text{number of outcomes favorable to A}}{\text{number of outcomes}}$$

if all outcomes are equally likely!

Definitions

Experiment: Activity that produces or observes an outcome.

Drawing 2 marbles from the urn with replacement, and noting the order.



Sample Space: Set of possible events for an experiment.

$\Omega =$ {(1, 1), (1, 2), ..., (2,1), ..., (3, 3)}

Event: Subset of the sample space.

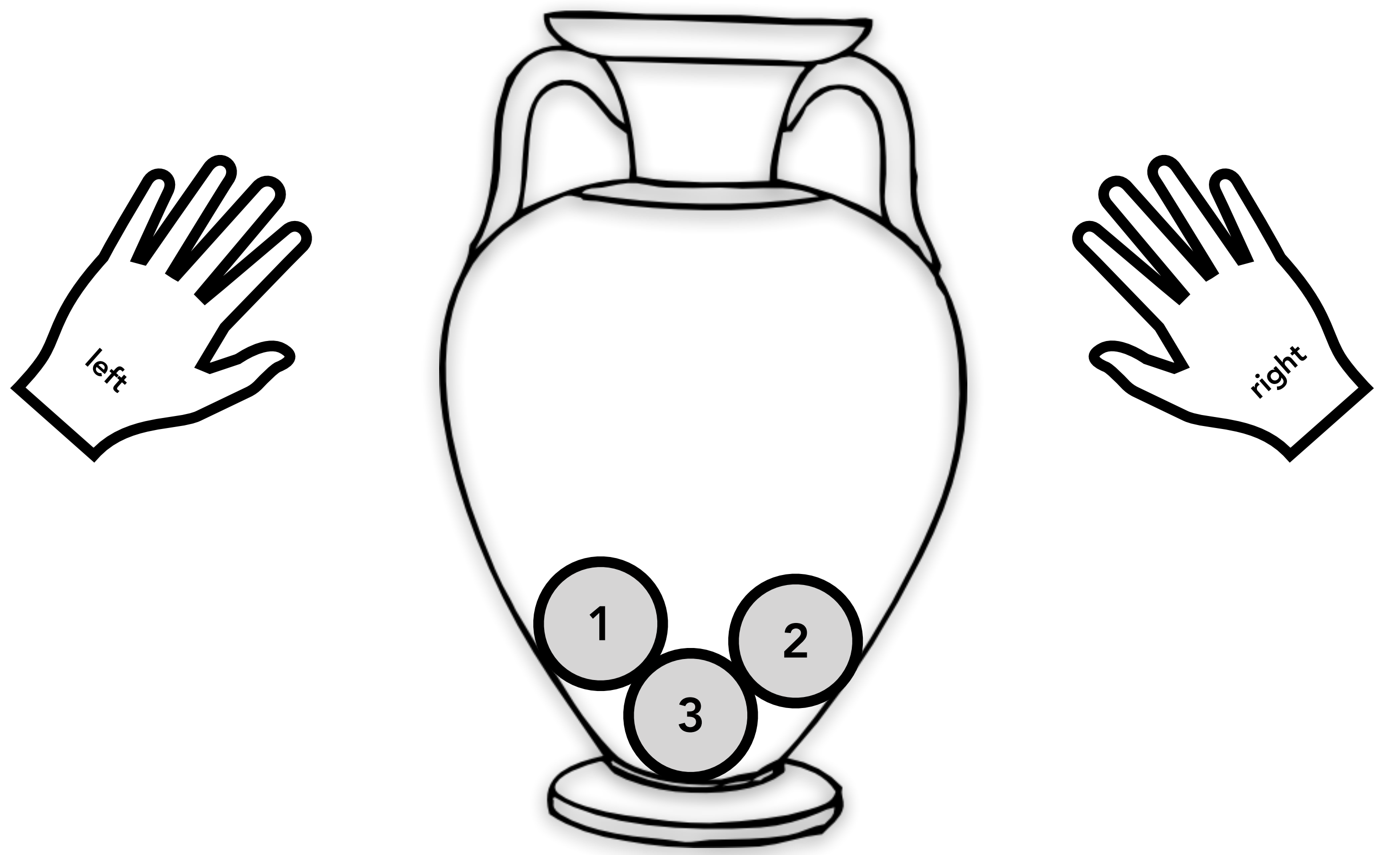
(1, 1)

Definitions

If $P(X_i)$ is the probability of event X_i

1. Probability cannot be negative.

$$P(X_i) \geq 0$$



2. Total probability of all events in the sample space is 1.

$$\sum_{i=1}^{N} P(X_i) = P(X_1) + P(X_2) + \ldots + P(X_N) = 1$$

Clue guide to probability

Clue guide to probability

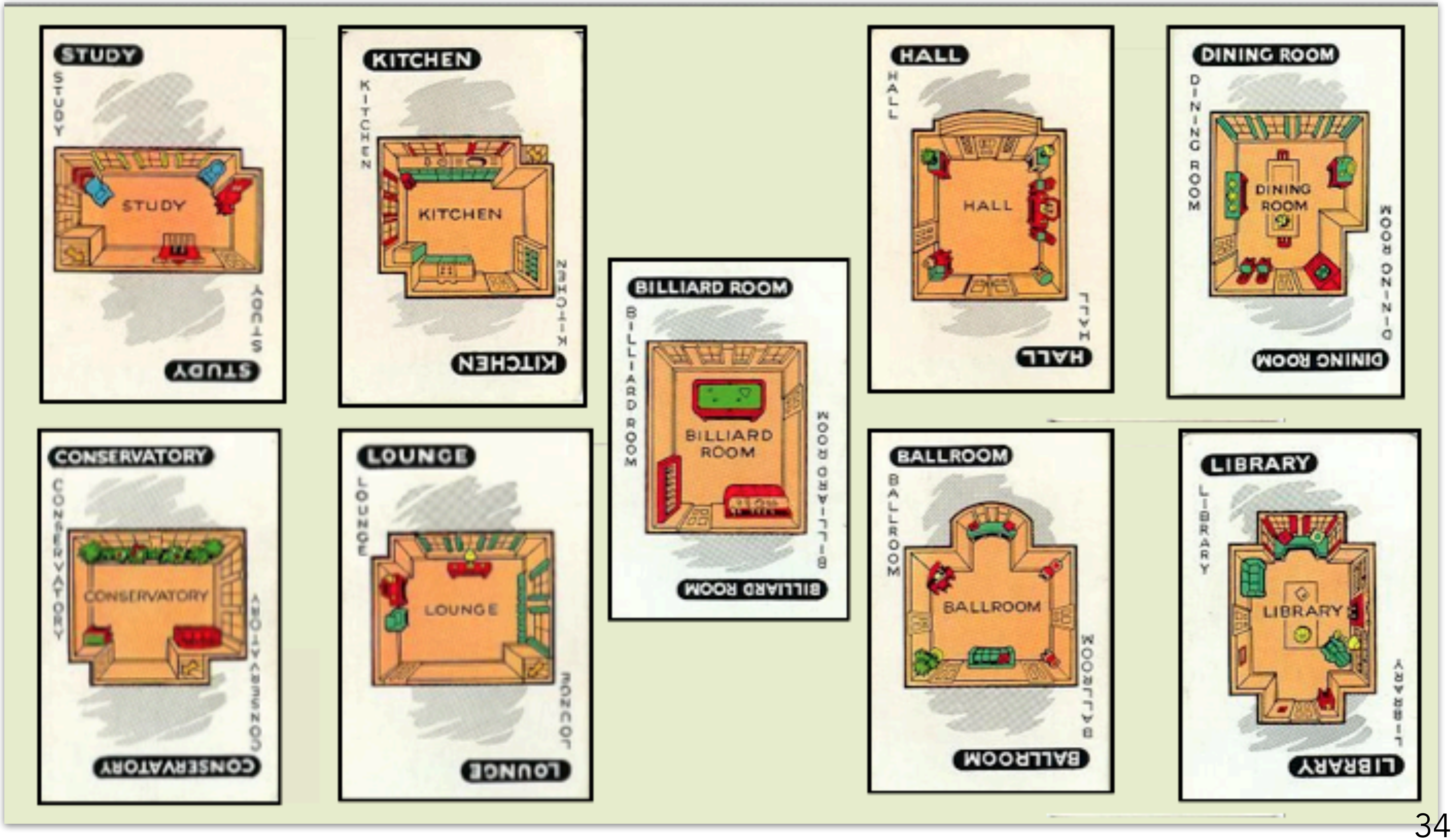
Who killed Mr Boddy?

Clue guide to probability

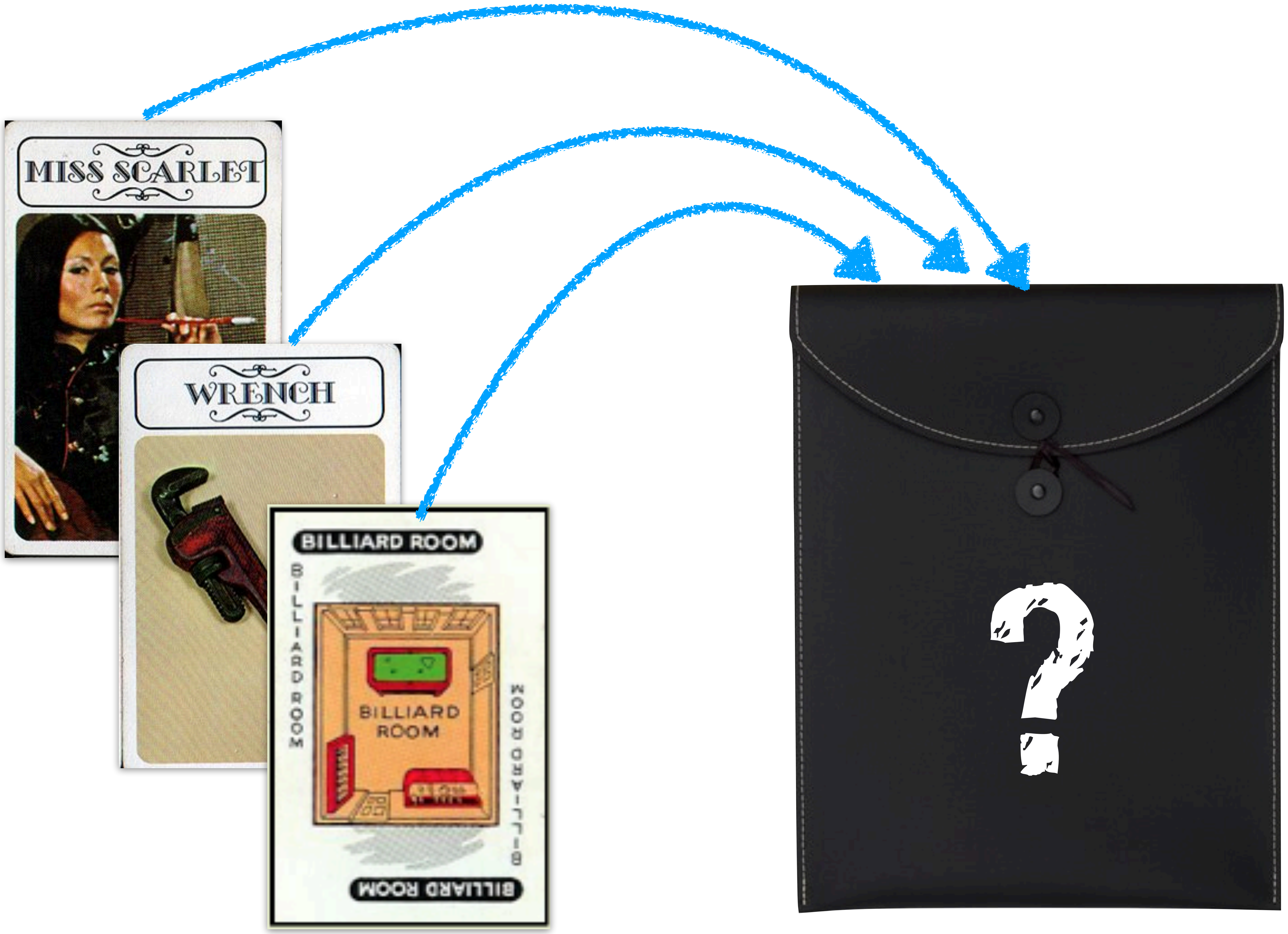
Who killed Mr Boddy, **with what?**

Clue guide to probability

Who killed Mr Boddy, with what, **and where?**



Clue guide to probability



Clue guide to probability

```r
who = c("ms_scarlet", "col_mustard", "mrs_white",
        "mr_green", "mrs_peacock", "prof_plum")
what = c("candlestick", "knife", "lead_pipe",
         "revolver", "rope", "wrench")
where = c("study", "kitchen", "conservatory",
          "lounge", "billiard_room", "hall",
          "dining_room", "ballroom", "library")

df.clue = expand_grid(who = who,
                      what = what,
                      where = where)
```

all combinations

Ω

who	what	where
ms_scarlet	candlestick	study
ms_scarlet	candlestick	kitchen
ms_scarlet	candlestick	conservatory
ms_scarlet	candlestick	lounge

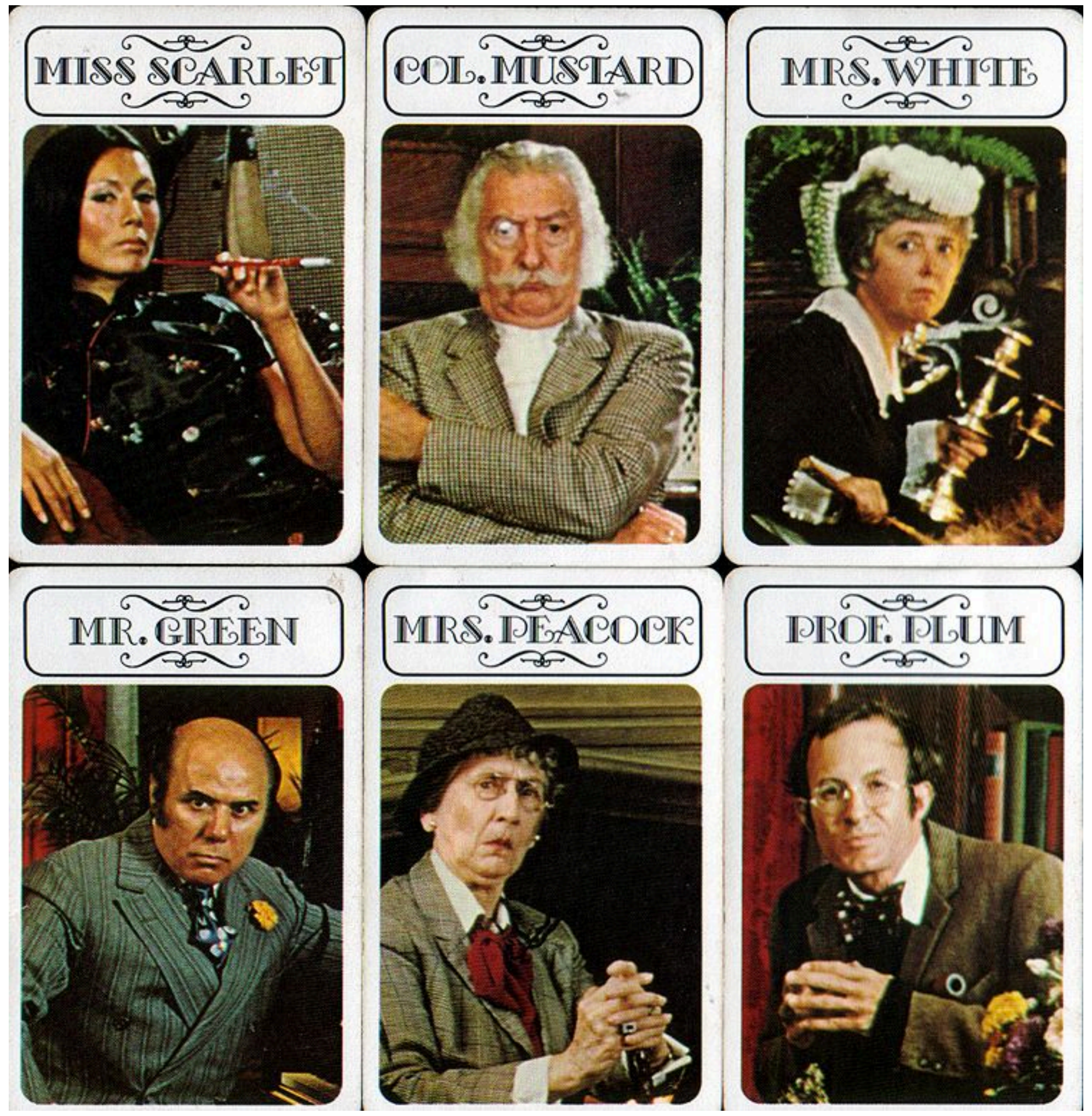
⋮

`nrow(df.clue)` = 324

Clue guide to probability

- 6 suspects
- mutually exclusive and exhaustive
- each equally likely a priori
- $p(\text{who} = \text{Prof. Plum}) = \frac{1}{6}$

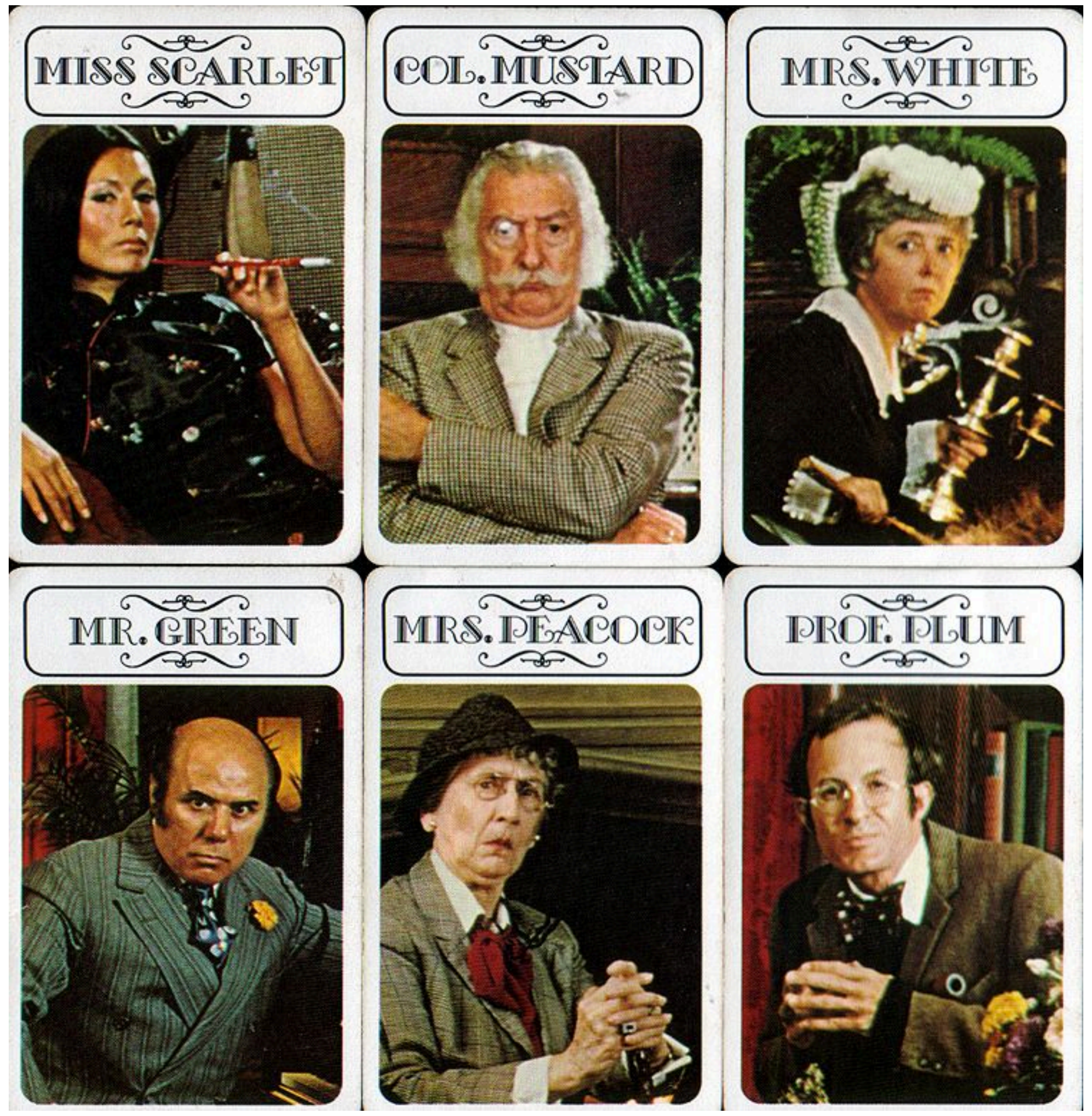
Who?



Clue guide to probability

- ***conditional probability***:
- $p(A \mid B)$ (probability of A given B)
- **Definition**: $p(A \mid B) = \dfrac{p(A, B)}{p(B)}$

Who?

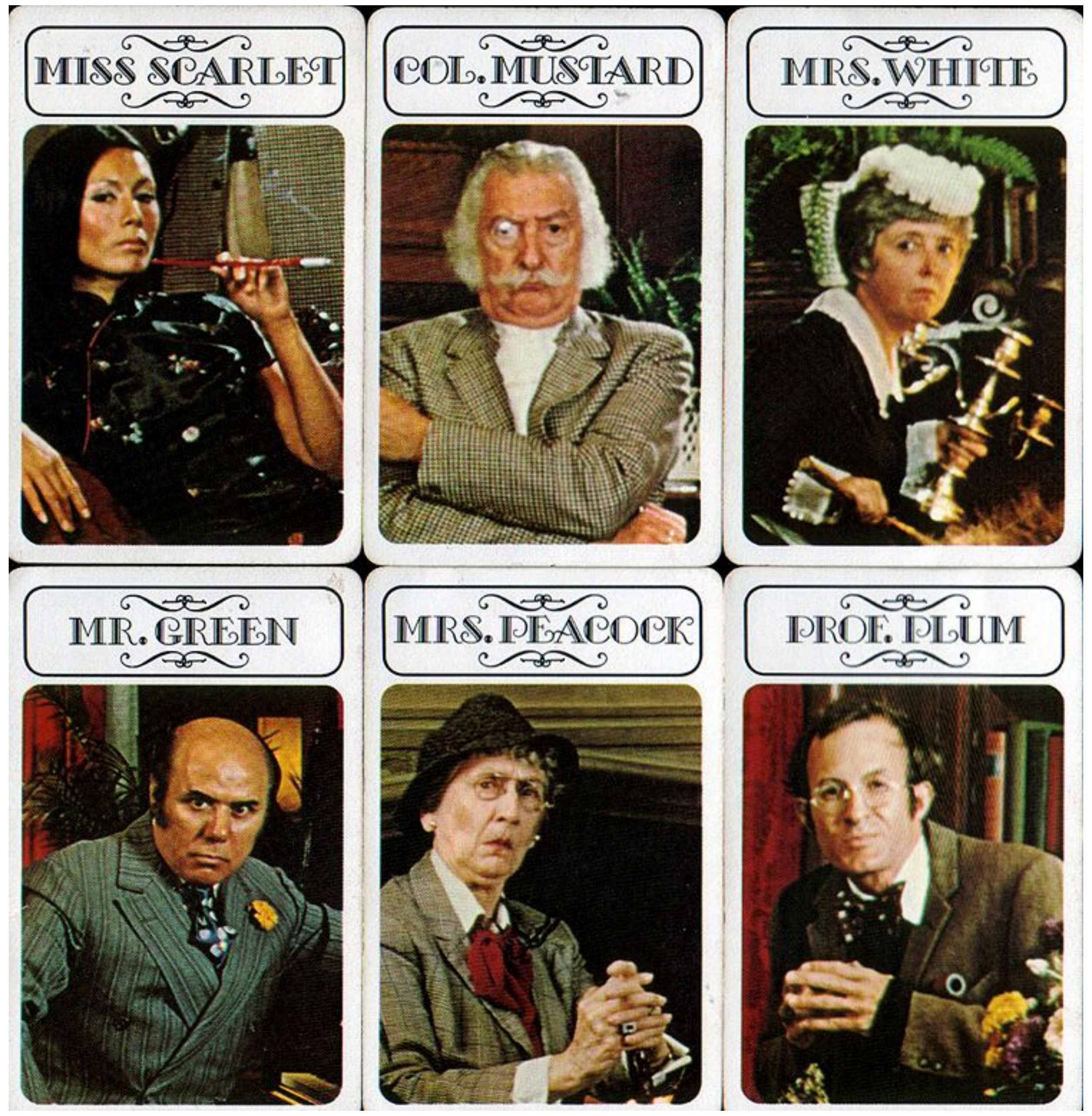


Probability that it was Prof Plum, given that the murderer was male?

$$p(\text{Prof. Plum} \mid \text{male}) = \,?$$

Clue guide to probability

Who?



- ***conditional probability***:
- $p(A \mid B)$ (probability of A given B)
- **Definition**: $p(A \mid B) = \dfrac{p(A, B)}{p(B)}$

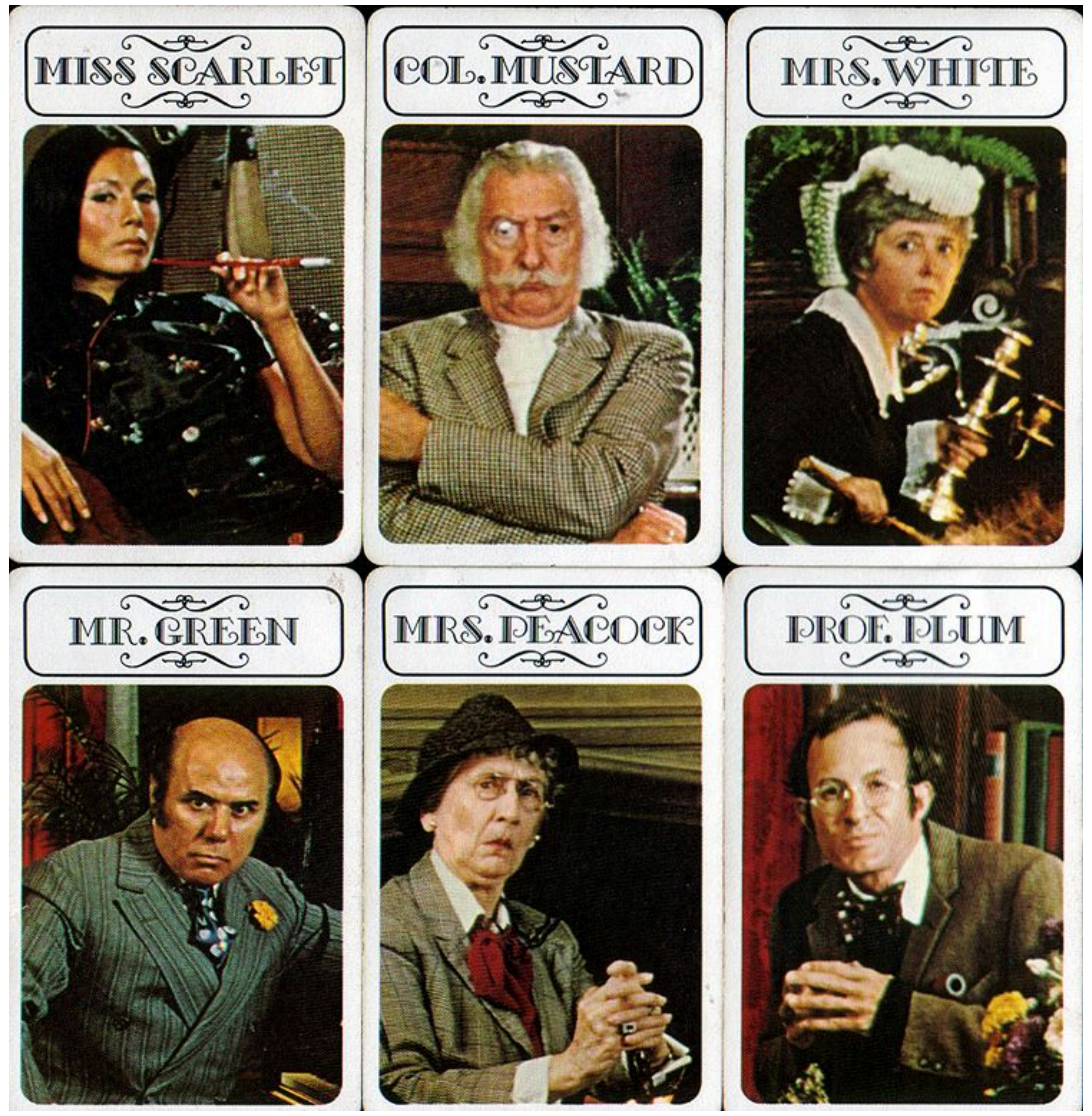
Probability that it was Prof Plum, given that the murderer was male?

$$p(\text{Prof. Plum} \mid \text{male}) =$$

Clue guide to probability

- ***conditional probability***:
- $p(A \mid B)$ (probability of A given B)
- **Definition**: $p(A \mid B) = \dfrac{p(A, B)}{p(B)}$
- $p(\text{Prof. Plum} \mid \text{male}) = \dfrac{1/6}{1/2} = 1/3$

Who?

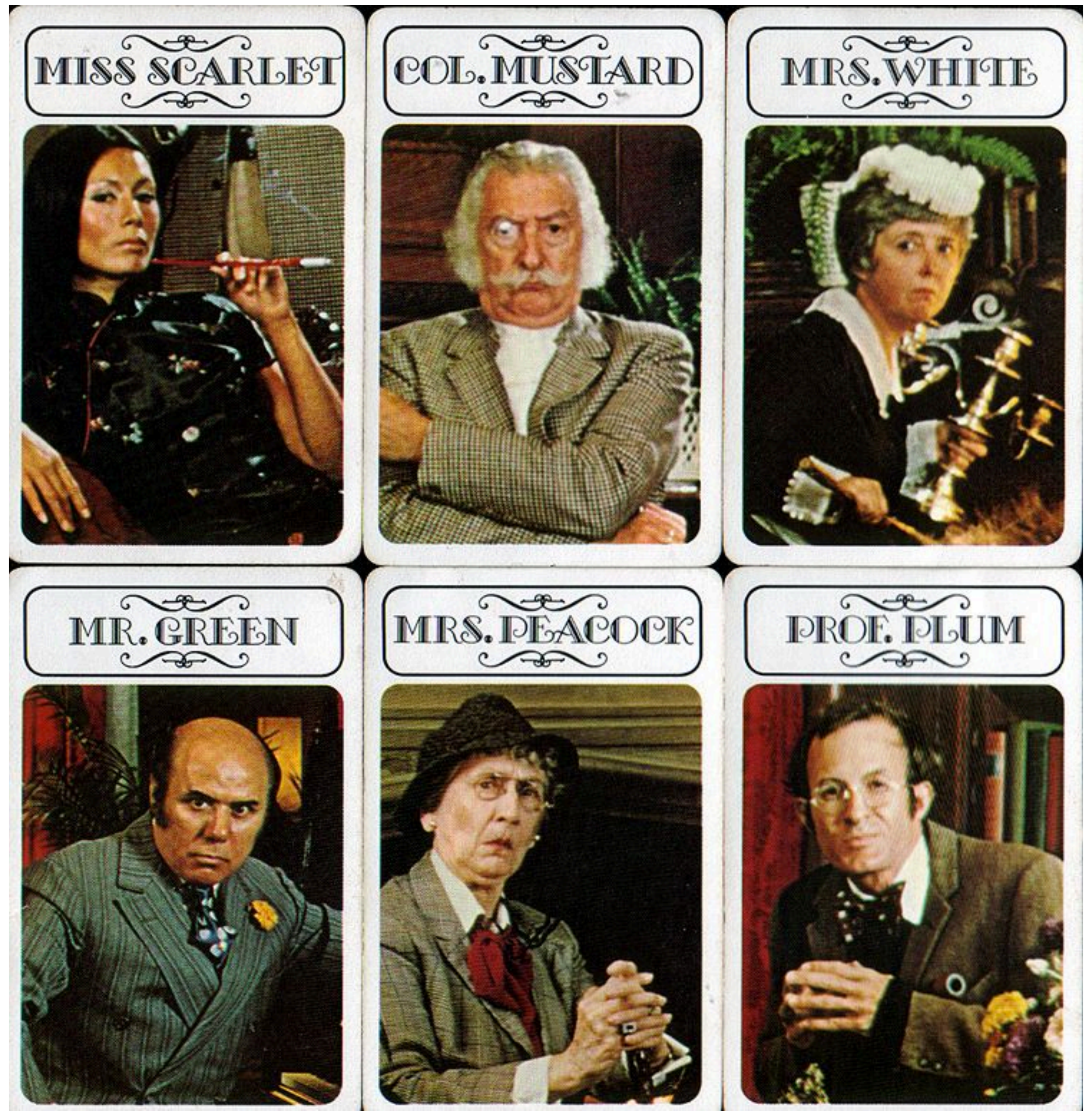


who	gender
col_mustard	male
mr_green	male
prof_plum	male
ms_scarlet	female
mrs_white	female
mrs_peacock	female

```
df.suspects = df.clue %>%
  distinct(who) %>%
  mutate(gender = ifelse(
    test = who %in% c("ms_scarlet",
                      "mrs_white",
                      "mrs_peacock"),
    yes = "female",
    no = "male"))
```

Clue guide to probability

Who?



- ***conditional probability***:
- $p(A \mid B)$ (probability of A given B)
- **Definition**: $p(A \mid B) = \dfrac{p(A, B)}{p(B)}$
- $p(\text{Prof. Plum} \mid \text{male}) = \dfrac{1/6}{1/2} = 1/3$

who	gender
col_mustard	male
mr_green	male
prof_plum	male
ms_scarlet	female
mrs_white	female
mrs_peacock	female

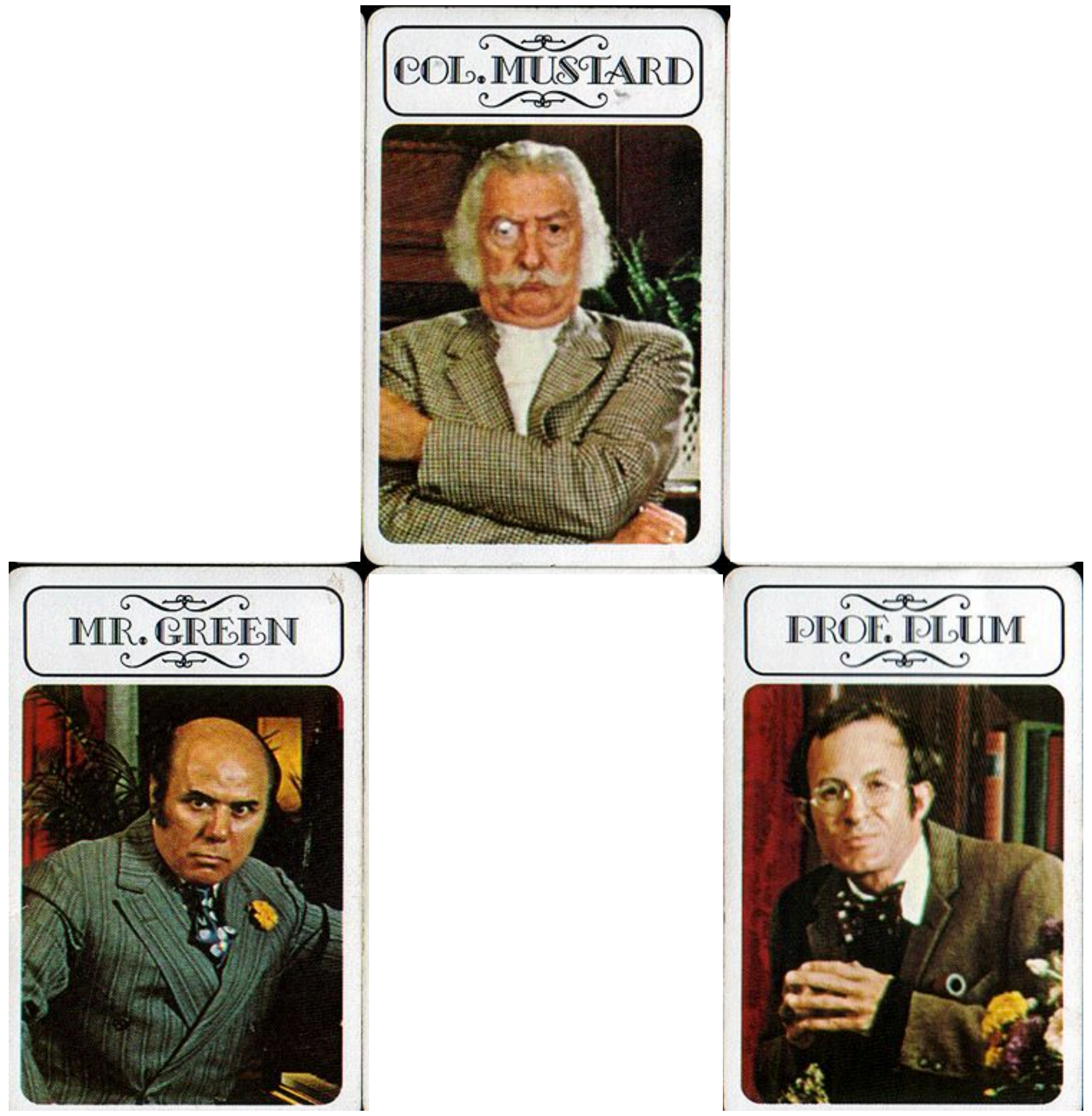
```
df.suspects %>%
  summarize(p_prof_plum_given_male =
              sum(gender == "male" &          who
                  == "prof_plum") /
              sum(gender == "male"))
```

use naive definition of probability

Clue guide to probability

- ***conditional probability***:
- $p(A \mid B)$ (probability of A given B)
- **Definition**: $p(A \mid B) = \dfrac{p(A, B)}{p(B)}$
- $p(\text{Prof. Plum} \mid \text{male}) = \dfrac{1/6}{1/2} = 1/3$

Who?



who	gender
col_mustard	male
mr_green	male
prof_plum	male

```
df.suspects %>%
  filter(gender == "male") %>%
  summarize(p_prof_plum_given_male =
                  sum(who == "prof_plum") /
                  n())
```

Clue guide to probability

- ***independence***:
- A and B are independent if
- **Definition**: $p(A \mid B) = p(A)$
- (probability of A does not change if you know B)

- $p(\text{Prof Plum} \mid \text{candle stick}) = p(\text{Prof Plum})$
- each card (who and what) is drawn from a separate pack of cards

Who?

What?

Clue guide to probability

- ***joint probability***:
- if A and B are independent then
- **Definition**: $p(A, B) = p(A) \cdot p(B)$

- $p(\text{Prof Plum, candle stick}) =$

$p(\text{Prof Plum}) \cdot p(\text{candle stick}) =$

$$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Who?

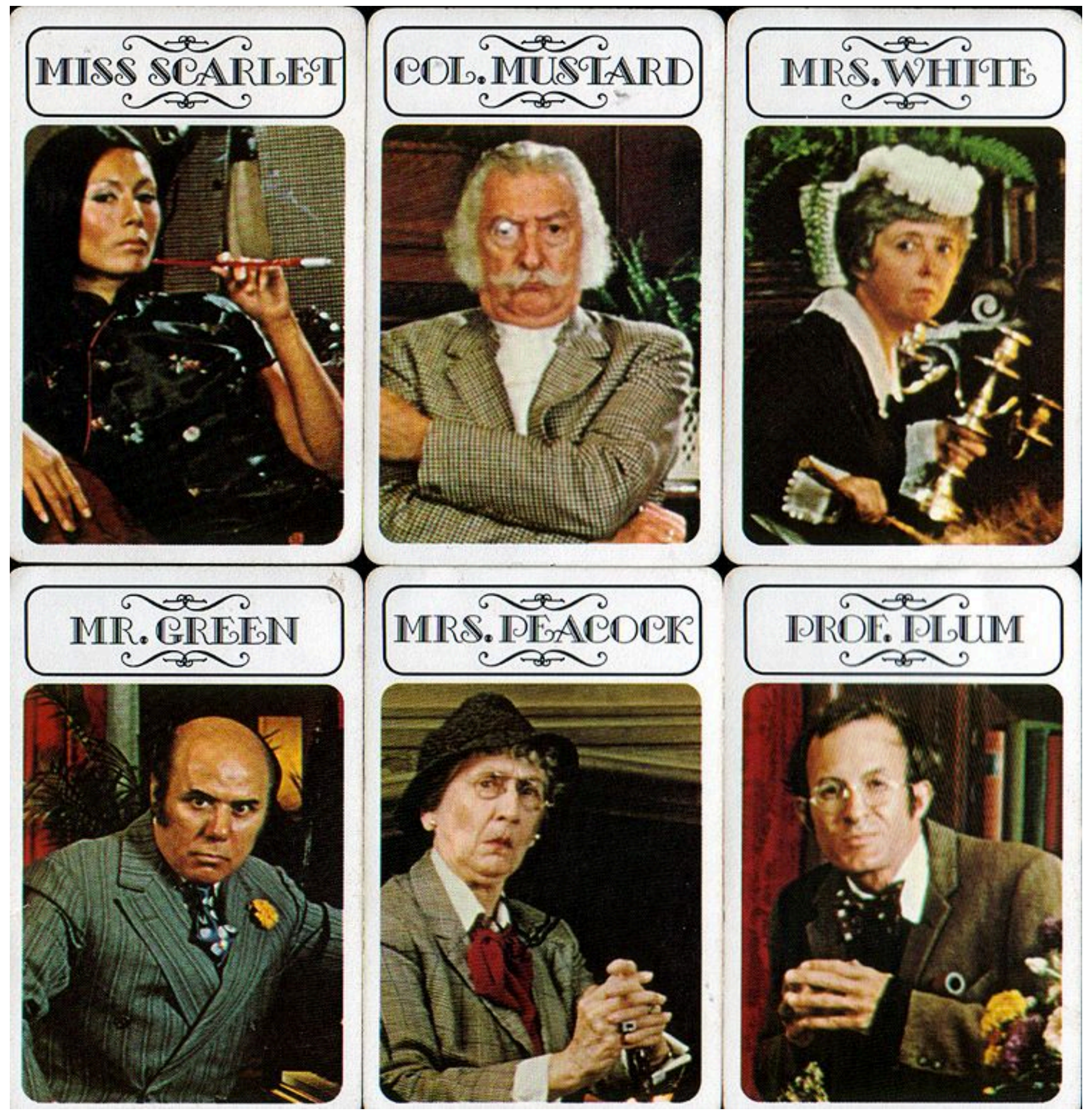
What?

Clue guide to probability

- ***dependence***:
- **Definition**: $p(A \mid B) \neq p(A)$
- **Definition**: $p(A, B) = p(A) \cdot p(B \mid A)$

- if women were more likely than men to use the revolver then
- $p(\text{Mrs. White} \mid \text{Revolver}) > p(\text{Mrs. White})$

Who?



What?

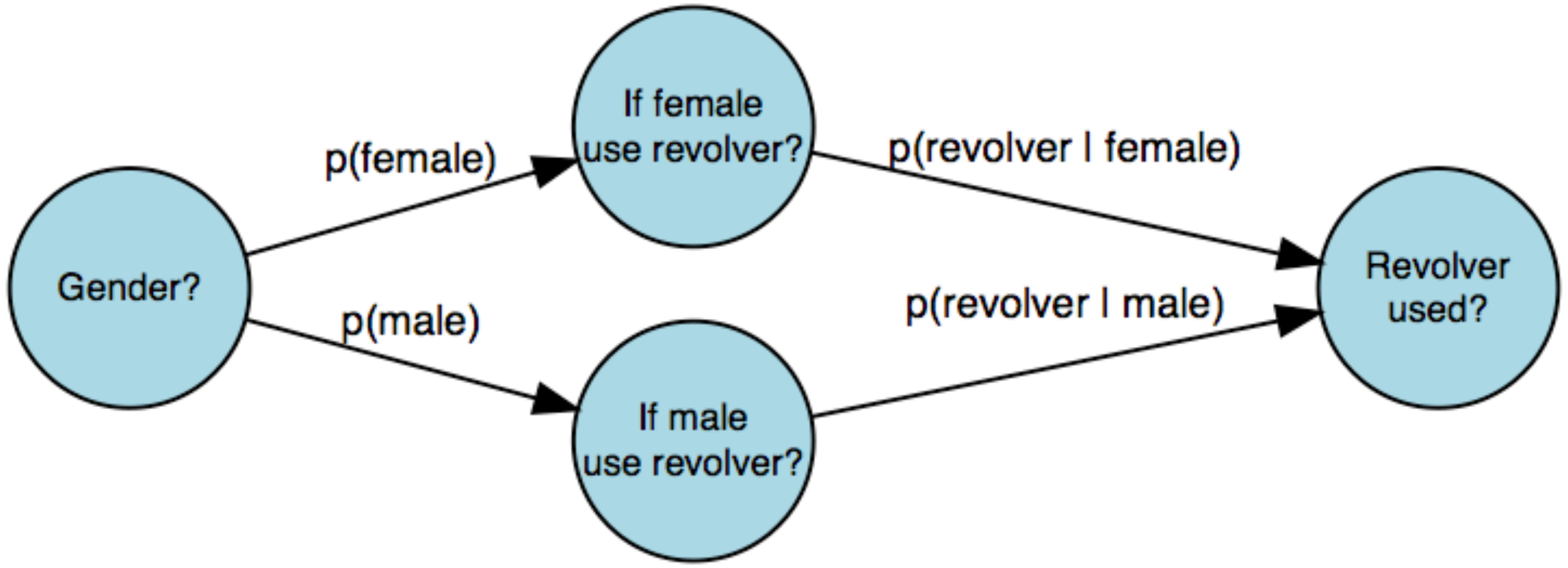
Clue guide to probability

- ***law of total probability***
- **Definition**:

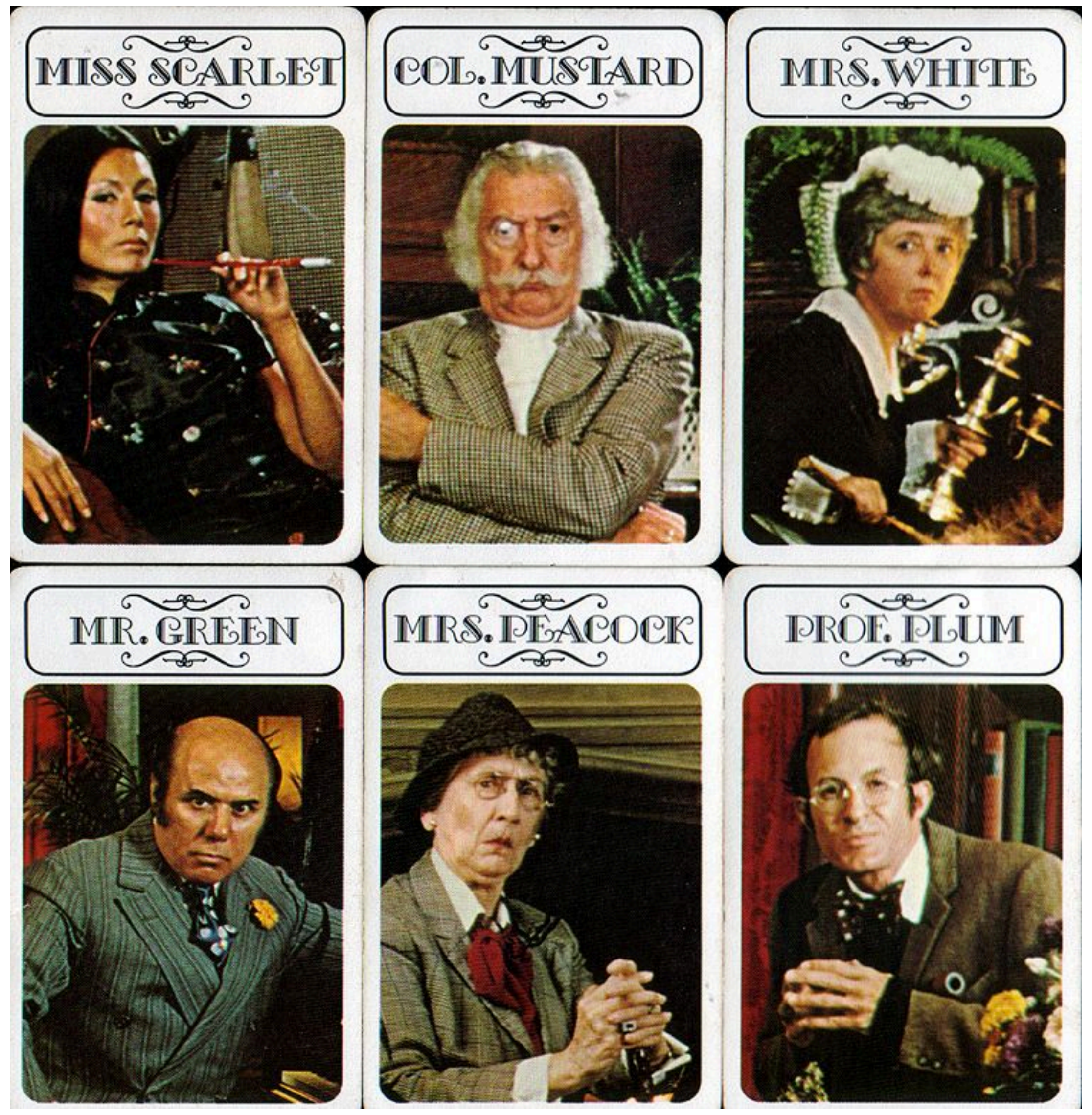
$$p(A) = p(A \mid B) \cdot p(B) + p(A \mid \neg B) \cdot p(\neg B)$$

$$p(A) = \sum_{i=1}^{n} p(A \mid B_i) \cdot p(B_i)$$

$$p(\text{what} = \text{Revolver}) = ?$$



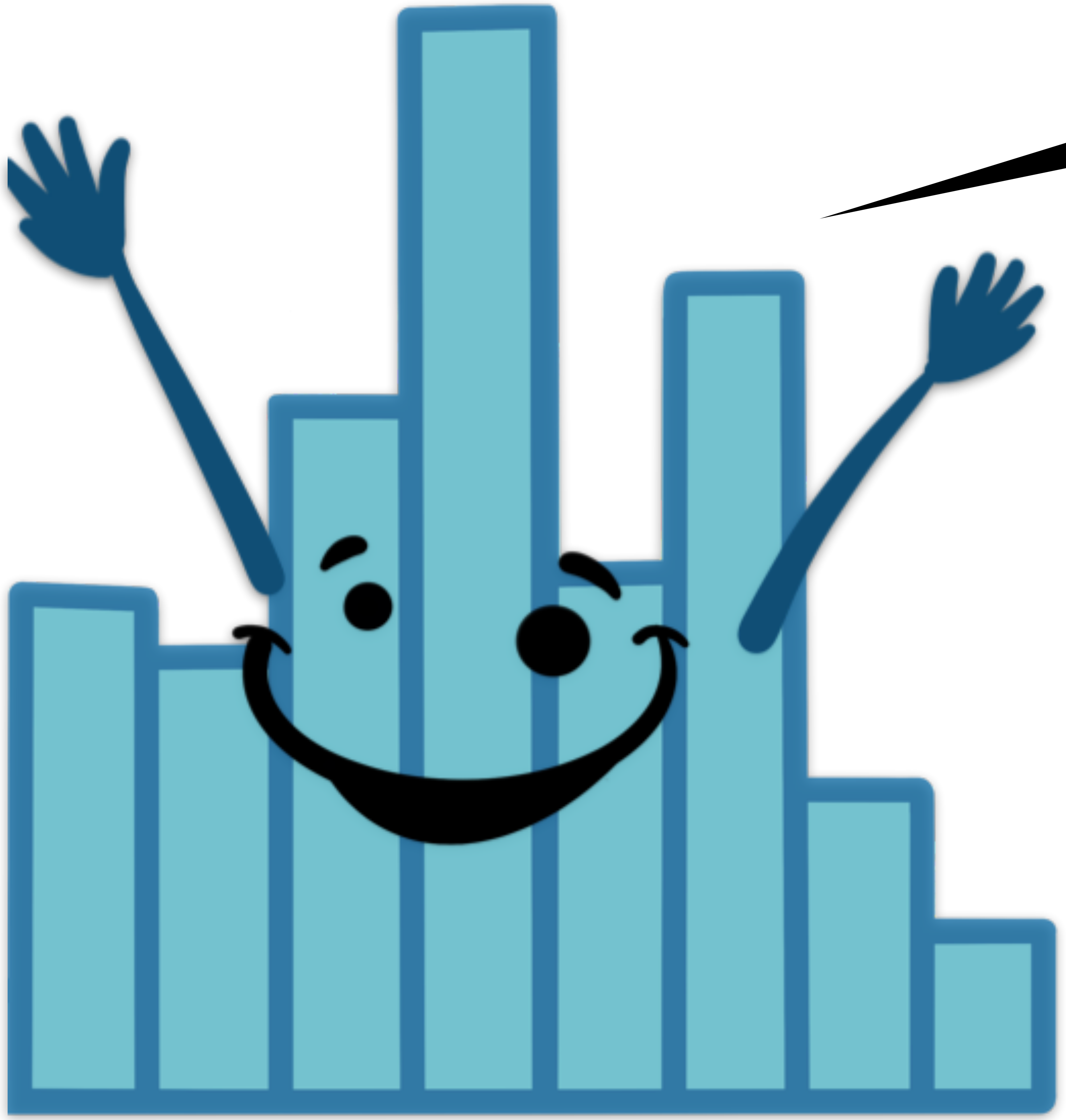
Who?



What?

02:00

stretch break!



Understanding Bayes' rule

Clue guide to probability

Bayes Theorem in a few steps

Clue guide to probability

- *Bayes' theorem (derivation)*

$$p(B \mid A) = \frac{p(A, B)}{p(A)} \qquad p(A \mid B) = \frac{p(A, B)}{p(B)}$$

$$p(A, B) = \boxed{p(B \mid A) \cdot p(A) = p(A \mid B) \cdot p(B)}$$

$$p(B \mid A) = \frac{p(A \mid B) \cdot p(B)}{p(A)}$$

Clue guide to probability

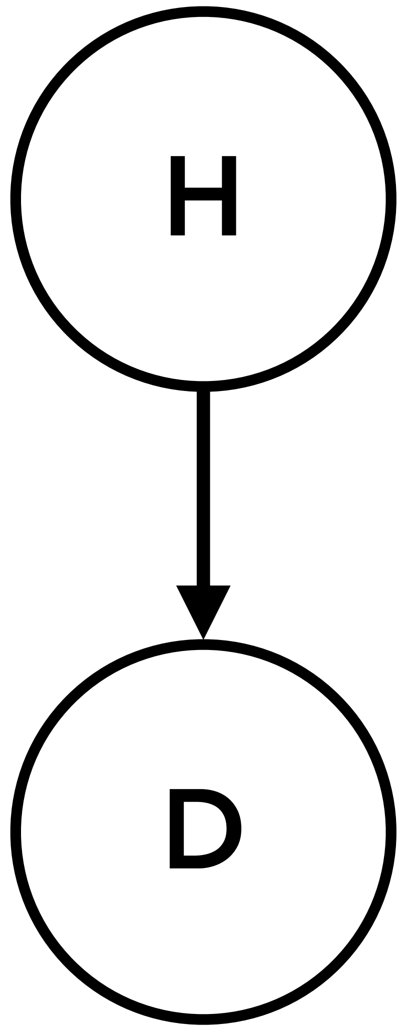
$$p(B\,|\,A) = \frac{p(A\,|\,B)\cdot p(B)}{p(A)}$$

$$\overset{\text{posterior}}{p(H\,|\,D)} = \frac{\overset{\text{likelihood}}{p(D\,|\,H)}\cdot \overset{\text{prior}}{p(H)}}{p(D)}$$

subjective probability interpretation

H = Hypothesis

D = Data



formal framework for learning from data

updating our prior belief $p(H)$, to a posterior belief $p(H\,|\,D)$ given some data

Clue guide to probability

posterior

likelihood prior

$$p(H \mid D) = \frac{p(D \mid H) \cdot p(H)}{p(D)}$$

probability of the data?!

H = Hypothesis

D = Data

law of total probability:

$$p(D) = \sum_{i=1}^{n} p(D \mid H_i) \cdot p(H_i)$$

take into account all the different ways in which the data could have come about

Clue guide to probability

A patient named Fred is tested for a disease called *conditionitis*, a medical condition that afflicts **1% of the population**. The test result is positive, i.e., the test claims that Fred has the disease. Let **D** be the event that Fred has the disease and **T** be the event that Fred tests positive.

Suppose that the test is "95% accurate"; there are different measures of the accuracy of a test, but in this problem it is assumed to mean that $\mathbf{P(T|D) = 0.95}$ and $\mathbf{P(\neg T|\neg D) = 0.95}$.
The quantity $P(T|D)$ is known as the *sensitivity* (= true positive rate) of the test, and $P(\neg T|\neg D)$ is known as the *specificity* (= true negative rate).

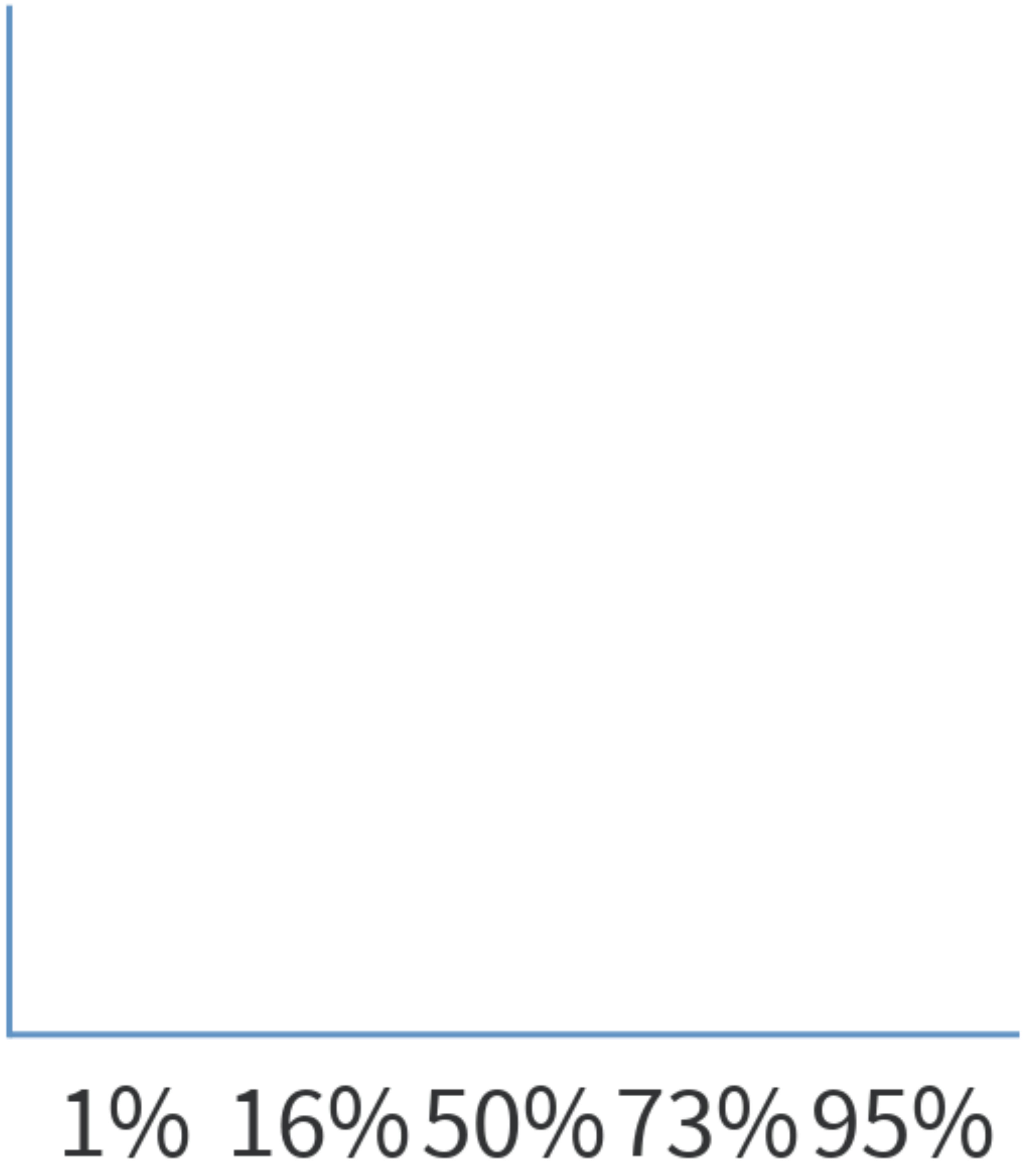
Find the conditional probability that Fred has *conditionitis*, given the evidence provided by the positive test result.

What's the probability that Fred has conditionitis?

A patient named Fred is tested for a disease called *conditionitis*, a medical condition that afflicts **1% of the population**. The test result is positive, i.e., the test claims that Fred has the disease. Let **D** be the event that Fred has the disease and **T** be the event that he tests positive.

Suppose that the test is "95% accurate"; there are different measures of the accuracy of a test, but in this problem it is assumed to mean that **$P(T|D) = 0.95$** and **$P(\neg T|\neg D) = 0.95$**. The quantity $P(T|D)$ is known as the *sensitivity* (= true positive rate) of the test, and $P(\neg T|\neg D)$ is known as the *specificity* (= true negative rate).

Find the conditional probability that Fred has *conditionitis*, given the evidence provided by the positive test result.



Clue guide to probability

what we know

$$P(D) = 0.01$$

$$P(T \mid D) = 0.95$$

$$P(T \mid \neg D) = 0.05$$

what we want to know

$$P(D \mid T) = ?$$

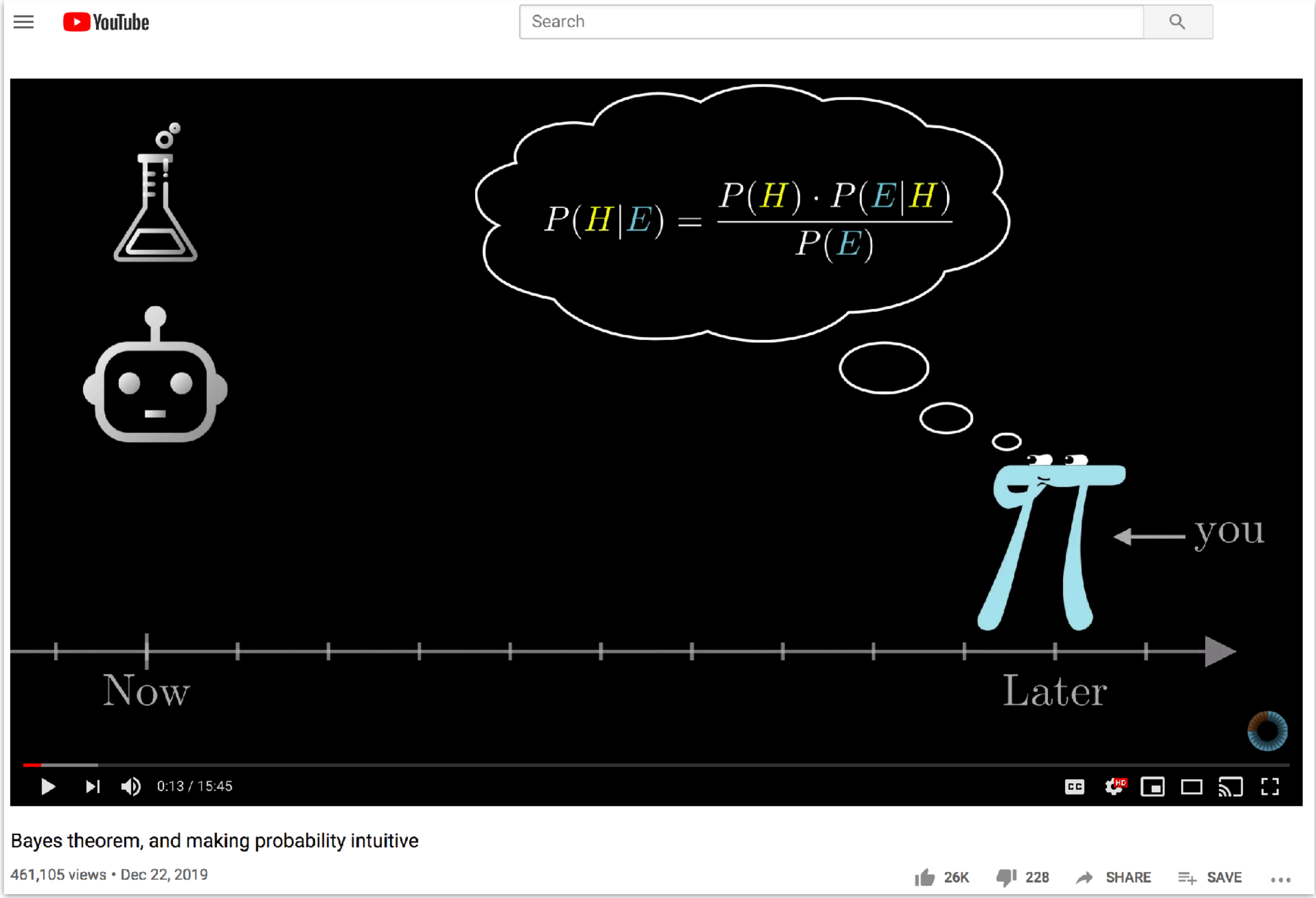
$$p(D \mid T) = \frac{p(T \mid D) \cdot p(D)}{p(T)}$$

Bayes' rule

XX

Clue guide to probability

XX



https://www.youtube.com/watch?v=HZGCoVF3YvM&feature=youtu.be

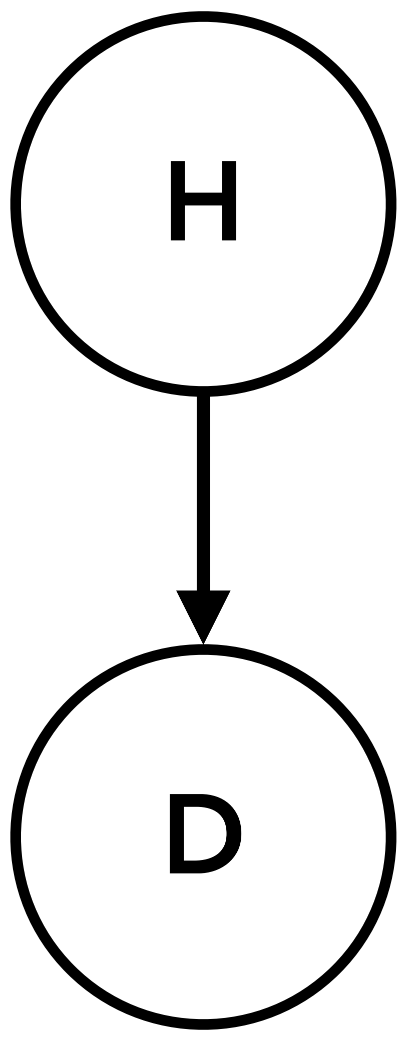
Bayes' theorem in three panels

$$\overset{\text{posterior}}{p(H \mid D)} = \frac{\overset{\text{likelihood}}{p(D \mid H)} \cdot \overset{\text{prior}}{p(H)}}{p(D)}$$

subjective probability interpretation

H = Hypothesis

D = Data



https://www.tjmahr.com/bayes-theorem-in-three-panels/

Getting Bayes' right matters

Getting Bayes right matters!

PNAS

Officer characteristics and racial disparities in fatal officer-involved shootings

David J. Johnson[a,b,1], Trevor Tress[b], Nicole Burkel[b], Carley Taylor[b], and Joseph Cesario[b]

[a]Department of Psychology, University of Maryland at College Park, College Park, MD 20742; and [b]Department of Psychology, Michigan State University, East Lansing, MI 48824

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved June 24, 2019 (received for review March 5, 2019)

Despite extensive attention to racial disparities in police shootings, two problems have hindered progress on this issue. First, databases of fatal officer-involved shootings (FOIS) lack details about officers, making it difficult to test whether racial disparities vary by officer characteristics. Second, there are conflicting views on which benchmark should be used to determine racial disparities when the outcome is the rate at which members from racial groups are fatally shot. We address these issues by creating a database of FOIS that includes detailed officer information. We test racial disparities using an approach that sidesteps the benchmark debate by directly predicting the race of civilians fatally shot rather than comparing the rate at which racial groups are shot to some benchmark. We report three main findings: 1) As the proportion of Black or Hispanic officers in a FOIS increases, a person shot is more likely to be Black or Hispanic than White, a disparity explained by county demographics; 2) race-specific county-level violent crime strongly predicts the race of the civilian shot; and 3) although we find no overall evidence of anti-Black or anti-Hispanic disparities in fatal shootings, when focusing on different subtypes of shootings (e.g., unarmed shootings or "suicide by cop"), data are too uncertain to draw firm conclusions. We highlight the need to enforce federal policies that record both officer and civilian information in FOIS.

officer-involved shootings | racial disparity | racial bias | police use of force | benchmarks

what police departments were involved, and demographic information about civilians. However, even these databases fail to provide information about involved officers, which prevents asking whether certain types of officers are more likely to show racial disparities."

When officers fire their weapons at civilians, there are three possible outcomes: 1) They miss the civilian, 2) they result in a nonfatal hit, or 3) they result in a fatal hit. Not only do officers miss civilians the majority of times they fire [estimates of hit rates range from 20 to 40% (5, 6)], many shootings do not result in fatalities. Thus, it is important to be clear at the outset that our analyses speak to racial disparities in the subset of shootings that result in fatalities, and not officers' decisions to use lethal force more generally.

Why should we expect officer characteristics to relate to the race of a person fatally shot? Decades of research from criminal justice have investigated whether officer characteristics relate to the degree of force used by police. Whereas officer race does not typically predict how much force an officer uses (7–11), male and inexperienced officers use more force (7, 8, 10), perhaps due to their use of more aggressive tactics (e.g., initiating more stops; ref. 11). One issue with this research is that it focuses on whether officer characteristics increase the degree of force used, not whether force is used disproportionately by civilian race. Some researchers have proposed that racial disparities in FOIS might be driven by discrimination by White officers (12), but research on this issue is uncommon due to a lack of officer

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Claim:

"White officers are not more likely to shoot minority civilians than non-White officers"

$$\Pr(\text{shot}|\text{minority civilian}, \text{white officer}, X) - \Pr(\text{shot}|\text{minority civilian}, \text{minority officer}, X) \leq 0, \quad \textbf{[1]}$$

What the statistic says:

"whether a person fatally shot was more likely to be Black (or Hispanic) than White"

Original claim:

Requires Bayes' rule

$$\begin{aligned}&\Pr(\text{shot}|\text{minority civilian}, \text{white officer}, X) \\ &\quad - \Pr(\text{shot}|\text{minority civilian}, \text{minority officer}, X) \\ &= \frac{\Pr(\text{min. civ.}|\text{shot}, \text{white off.}, X) \times \Pr(\text{shot}|\text{white off.}, X)}{\Pr(\text{minority civilian}|\text{white officer}, X)} \\ &\quad - \frac{\Pr(\text{min. civ.}|\text{shot}, \text{min. off.}, X) \times \Pr(\text{shot}|\text{min. off.}, X)}{\Pr(\text{minority civilian}|\text{minority officer}, X)}. \quad \textbf{[2]}\end{aligned}$$

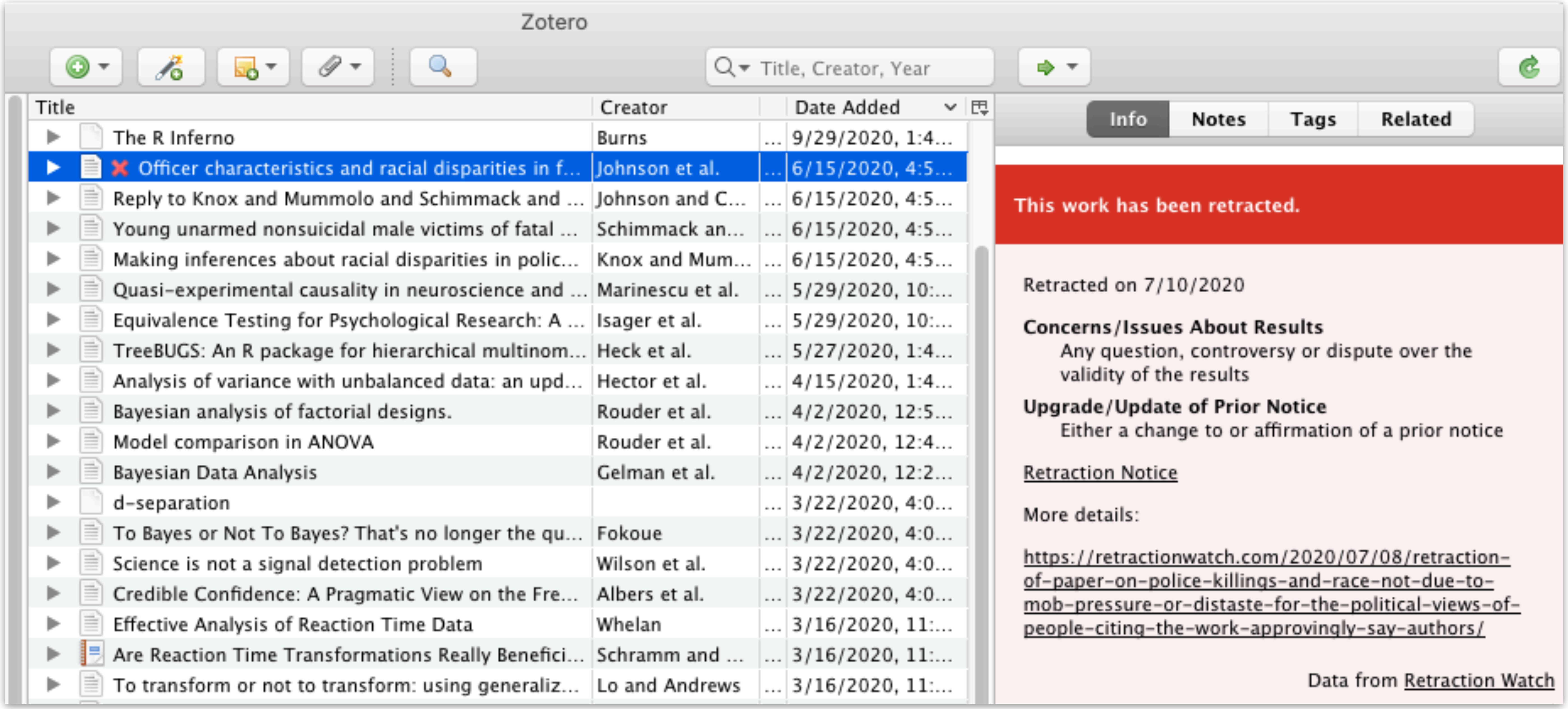
authors didn't have the relevant data to support their claim!

paper was retracted

Johnson, D. J., Tress, T., Burkel, N., Taylor, C., & Cesario, J. (2019). Officer characteristics and racial disparities in fatal officer-involved shootings. *Proceedings of the National Academy of Sciences*, 116(32), 15877–15882.

Knox, D., & Mummolo, J. (2020). Making inferences about racial disparities in police violence. *Proceedings of the National Academy of Sciences*, 117(3), 1261–1262.

Getting Bayes right matters!



Tip: Use Zotero as a reference manager!

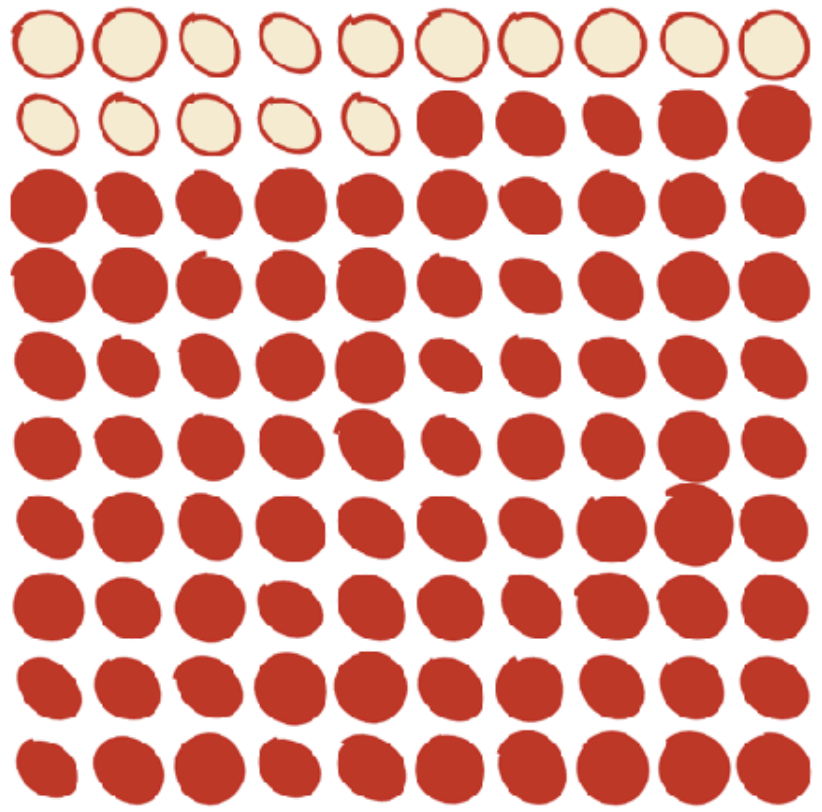
Getting Bayes right matters!

The New York Times

TheUpshot

When They Warn of Rare Disorders, These Prenatal Tests Are Usually Wrong

Some of the tests look for missing snippets of chromosomes. For every 15 times they correctly find a problem …



… they are wrong 85 times

For These Five Tests, Positive Results Are Often Wrong

As prenatal tests have expanded to more rare conditions, a larger share of their positive results are incorrect. Some of the worst-performing tests look for microdeletions, which are small missing snippets of chromosomes.

Chance positive results are wrong

DiGeorge syndrome
Affects 1 in 4,000 births
Can cause heart defects and delayed language acquisition. (May appear on lab reports as "22q.")

81% wrong

1p36 deletion
1 in 5,000 births
Can cause seizures, low muscle tone and intellectual disability.

84% wrong

Cri-du-chat syndrome
1 in 15,000 births
Can cause difficulty walking and delayed speech development.

80% wrong

Wolf-Hirschhorn syndrome
1 in 20,000 births
Can cause seizures, growth delays and intellectual disability.

86% wrong

Prader-Willi and Angelman syndromes
1 in 20,000 births
Can cause seizures and an inability to control food consumption.

93% wrong

https://www.nytimes.com/2022/01/01/upshot/pregnancy-birth-genetic-testing.html

Getting Bayes right matters!

sensitivity: $p(T \mid D) = 0.999$

specificity: $p(\neg T \mid \neg D) = 0.999$

prior: $p(D) = 0.0001$

$T =$ positive test result

$\neg T =$ negative test result

$D =$ disease

$\neg D =$ no disease

data: T (positive test result)

81% wrong

posterior:

$$p(D \mid T) = \frac{p(T \mid D) \cdot p(D)}{p(T \mid D) \cdot p(D) + p(T \mid \neg D) \cdot p(\neg D)}$$

$$= \frac{0.999 \cdot 0.0001}{0.999 \cdot 0.0001 + 0.001 \cdot 0.9999} \approx 0.09$$

https://www.nytimes.com/2022/01/01/upshot/pregnancy-birth-genetic-testing.html

Getting Bayes right matters!

Most people who are in the hospital being treated for Covid are vaccinated.

Getting Bayes right matters!

likelihood: $p(H \mid V) = 0.2$

$p(H \mid \neg V) = 0.5$

prior: $p(V) = 0.8$

$H =$ hospitalized

$\neg H =$ not hospitalized

$V =$ vaccinated

$\neg V =$ no vaccinated

data: H (the person is in the hospital)

posterior: $$p(V \mid H) = \frac{p(H \mid V) \cdot p(V)}{p(H \mid V) \cdot p(V) + p(H \mid \neg V) \cdot p(\neg V)}$$

$$= \frac{0.2 \cdot 0.8}{0.2 \cdot 0.8 + 0.5 \cdot 0.2} \approx 0.62$$

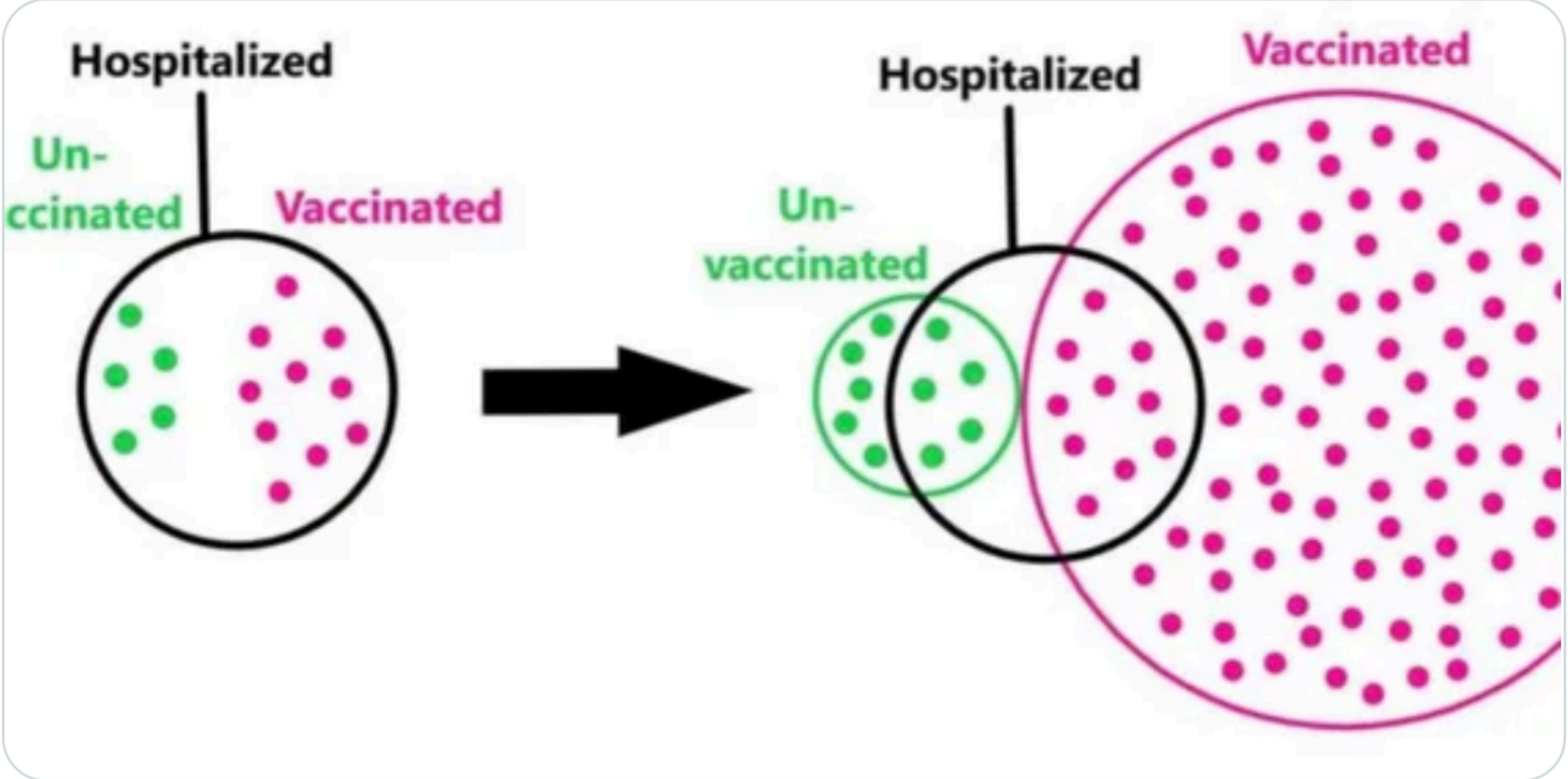
62% of the hospitalized people are vaccinated

https://www.nytimes.com/2022/01/01/upshot/pregnancy-birth-genetic-testing.html

Bayes' rule matters

Nick Brown
@sTeamTraen

Stolen from Reddit. May be of some use.



1:22 PM · Nov 20, 2021 · Twitter Web App

565 Retweets **54** Quote Tweets **2,752** Likes

Building a Bayesis

Rolling the dice

Four sided

Six sided

both dice are equally likely to be picked

p() = p() = 0.5

both dice are equal sided (uniform probability over the different numbers)

Which die do you think was rolled?

4   p( | data) = ?

4, 2, 1   p( | data) = 0.77

4, 2, 1, 3, 1   p( | data) = 0.88

4, 2, 1, 3, 1, 5   p( | data) = 0

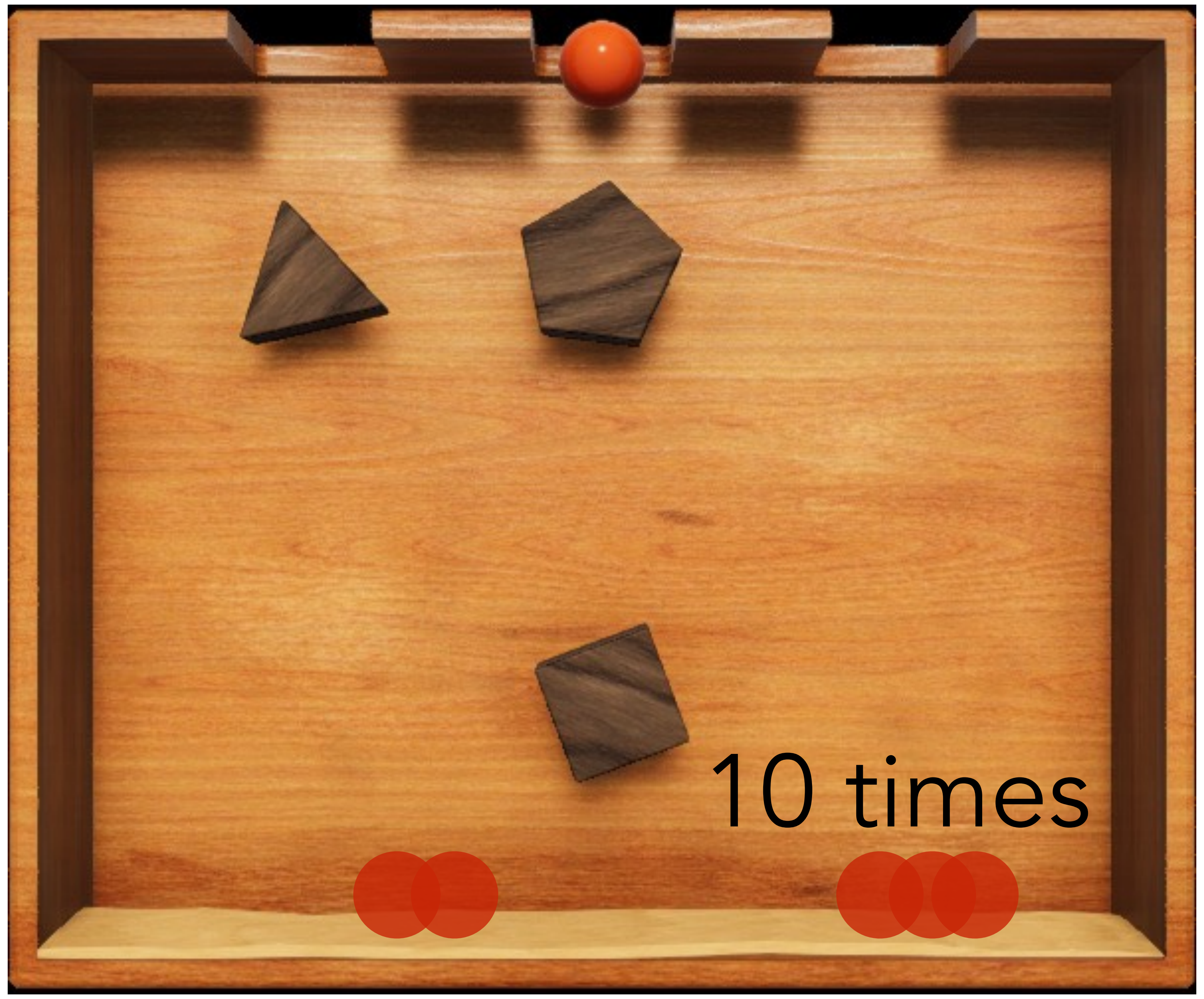
Physical reasoning

Gerstenberg, T., Siegel, M. H., & Tenenbaum, J. B. (2021). What happened? Reconstructing the past from vision and sound. PsyArXiv. https://psyarxiv.com/tfjdk

Physical reasoning

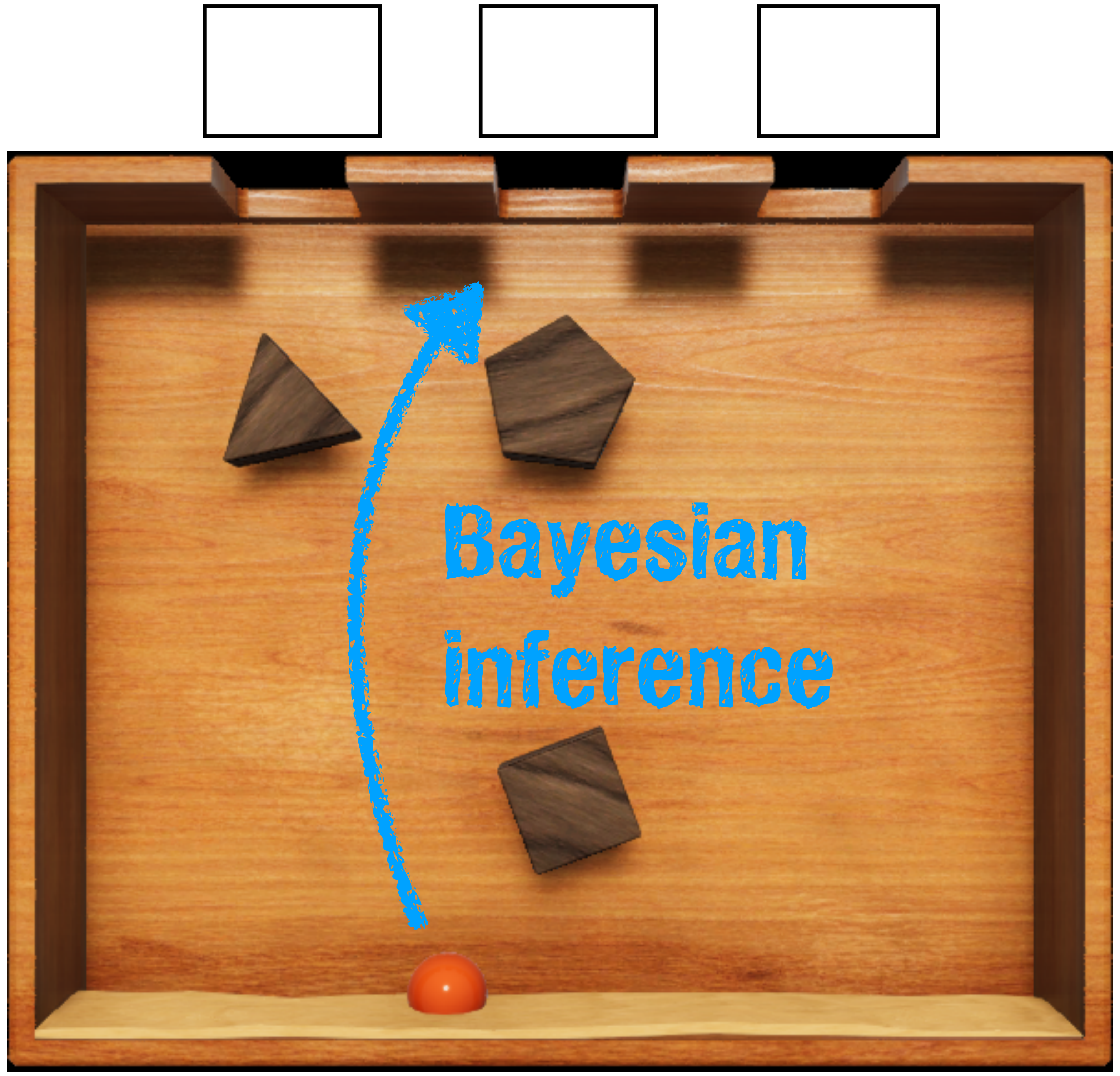
Gerstenberg, T., Siegel, M. H., & Tenenbaum, J. B. (2021). What happened? Reconstructing the past from vision and sound. PsyArXiv. https://psyarxiv.com/tfjdk

Prediction



Where will the ball land?

Inference



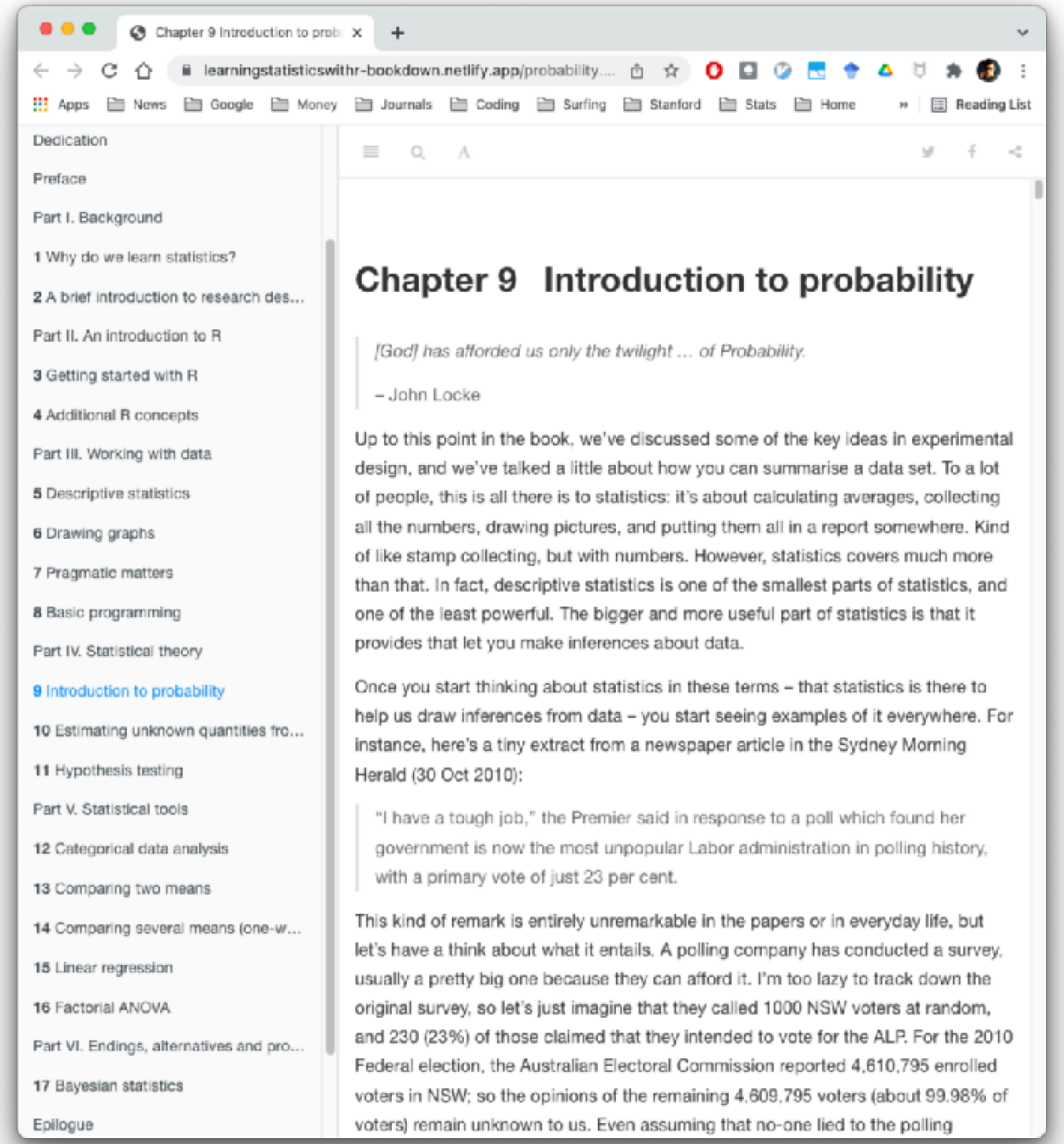
In which hole was the ball dropped?

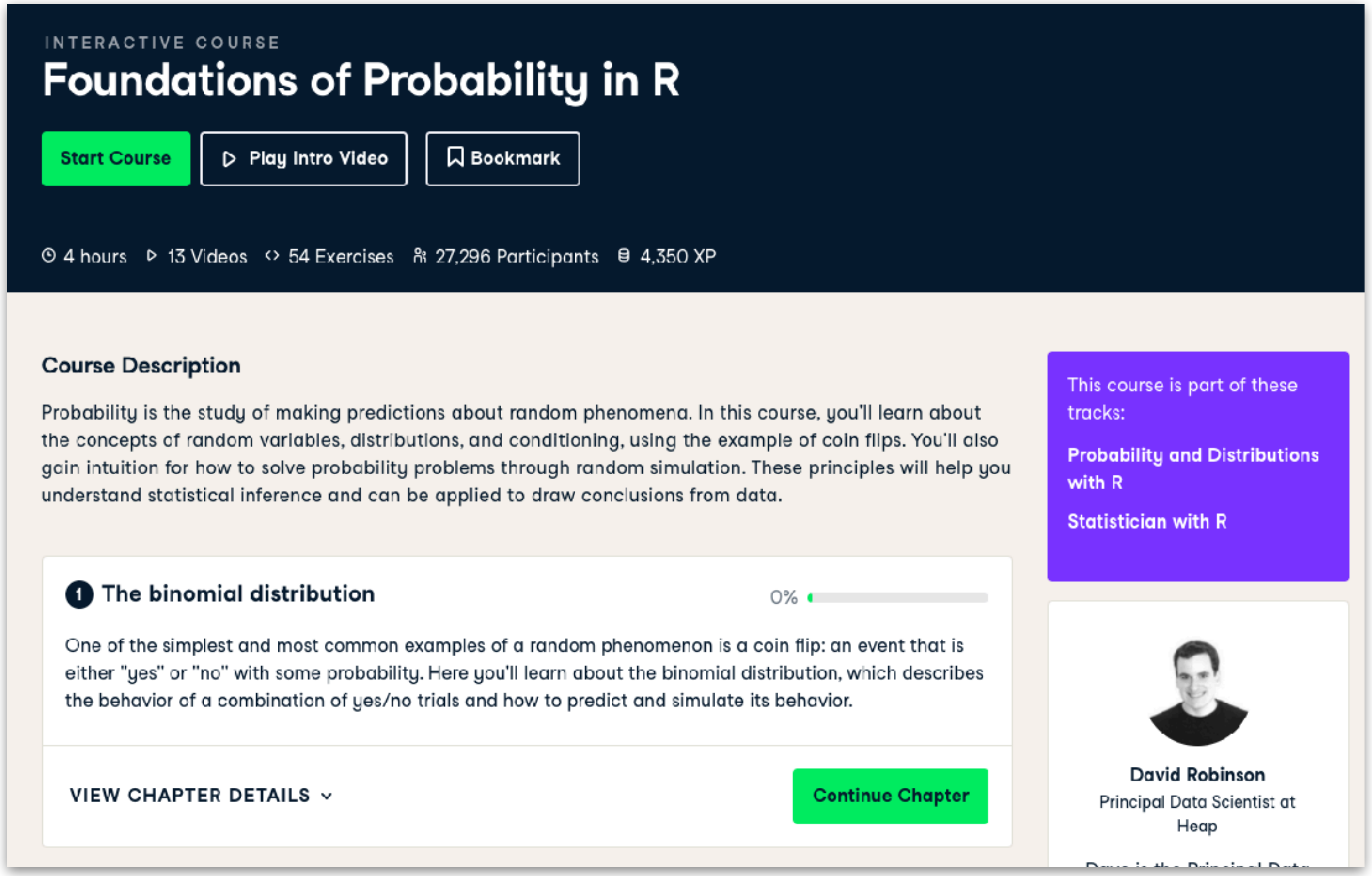
Gerstenberg, T., Siegel, M. H., & Tenenbaum, J. B. (2021). What happened? Reconstructing the past from vision and sound. PsyArXiv. https://psyarxiv.com/tfjdk

Outline

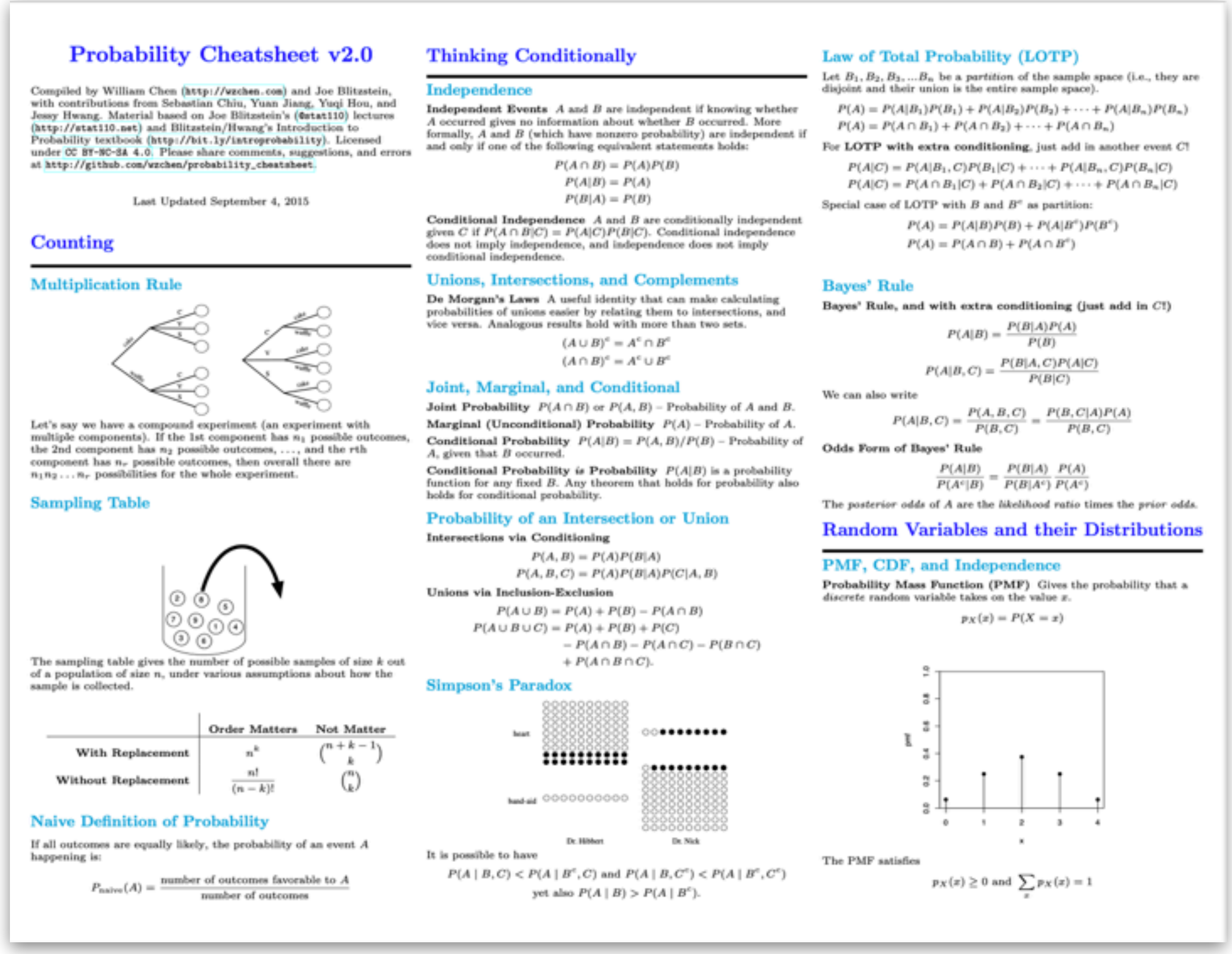
- Introduction to probability / Recap
    - Motivation
    - Counting possibilities
    - Clue guide to probability
    - Understanding Bayes' Rule
    - Getting Bayes' right matters!
    - Building a Bayesis

I want more!



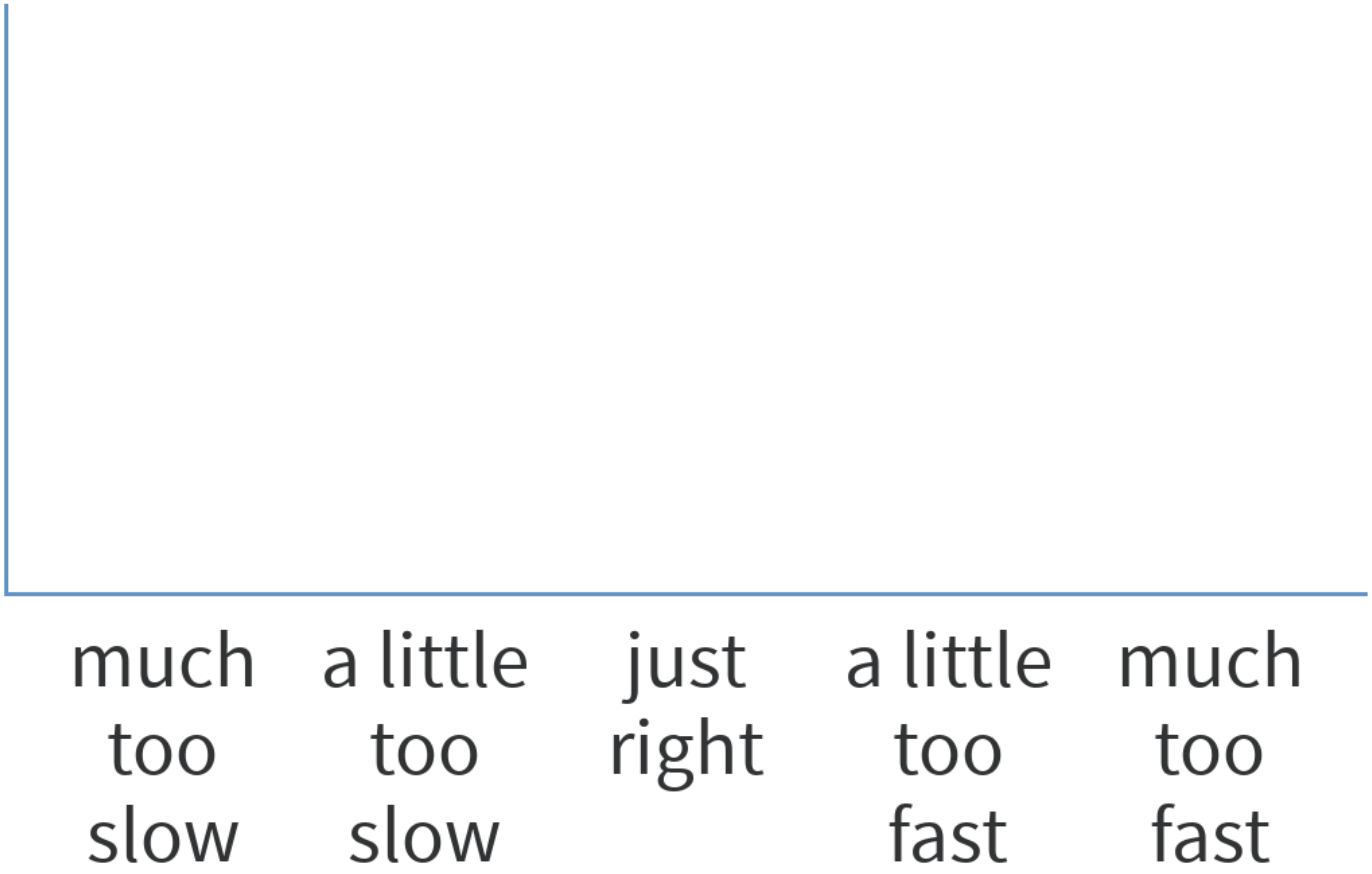


in the `figures/` folder for these materials on canvas



Feedback

How was the pace of today's class?



How happy were you with today's class overall?

What did you like about today's class? What could be improved next time?

Answers to this poll are anonymous

Have a nice Martin Luther King, Jr. day!