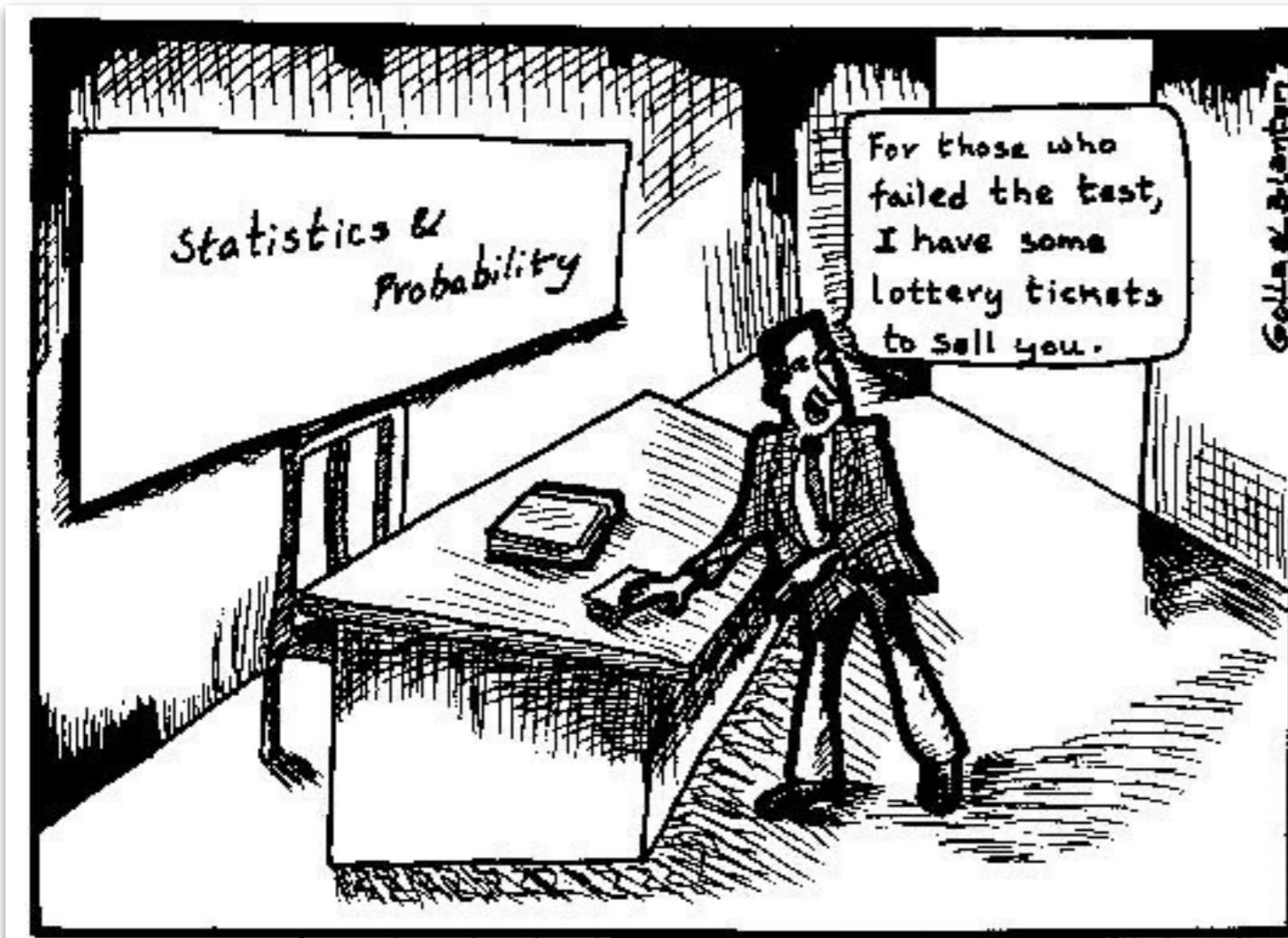


Probability & Causality



Chat

What's the present you most wanted as a kid?

To: Everyone ▾ More ▾

Type message here...

A screenshot of a messaging interface. The top right corner shows a "Chat" button. Below it, a blue message bubble contains the text "What's the present you most wanted as a kid?". Underneath the message, there are "To:" and "More" dropdown menus, and a text input field with the placeholder "Type message here...".

O COLLABORATIVE PLAYLIST
psych252

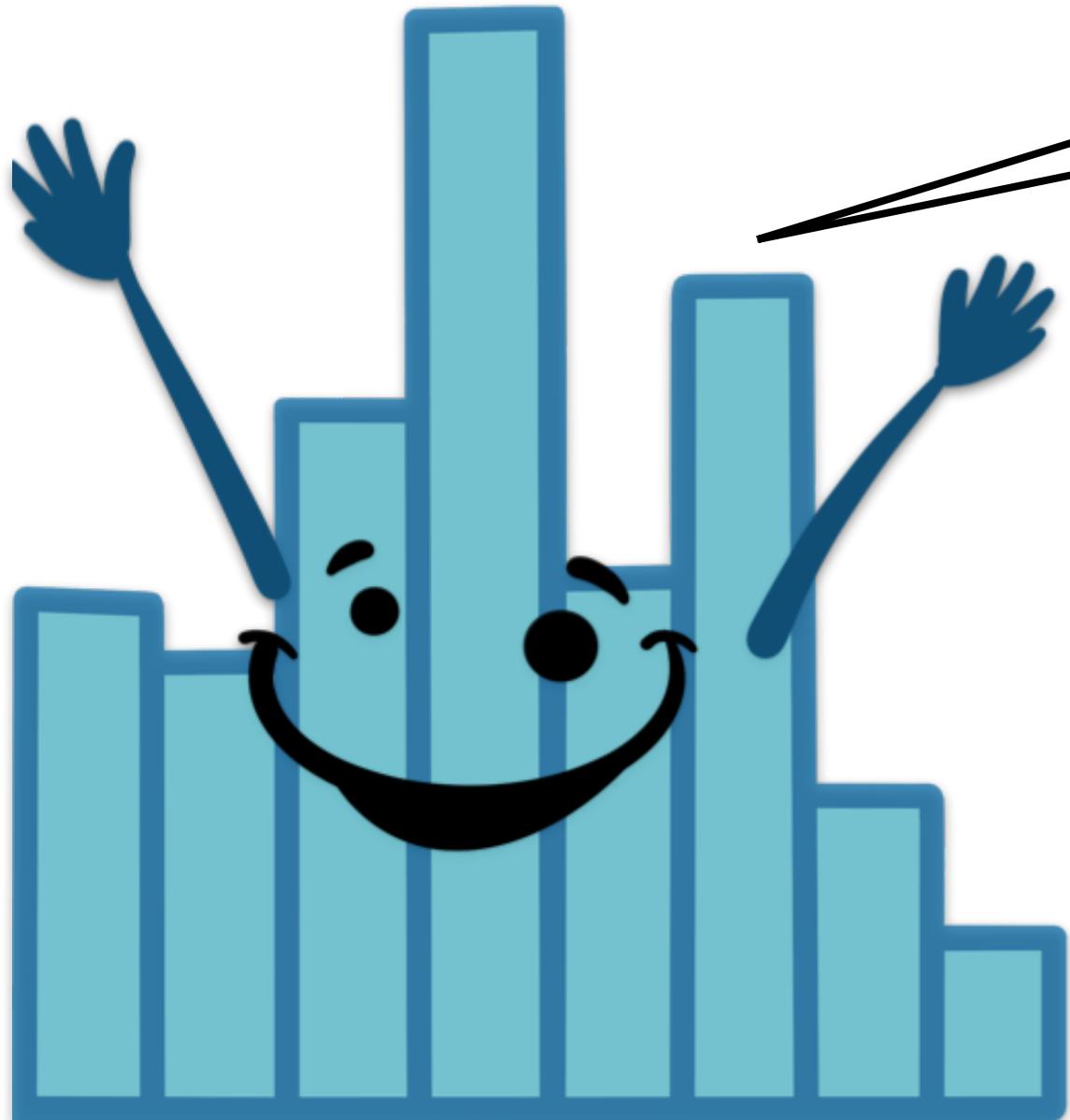
<https://tinyurl.com/psych252spotify21>

PLAY ...

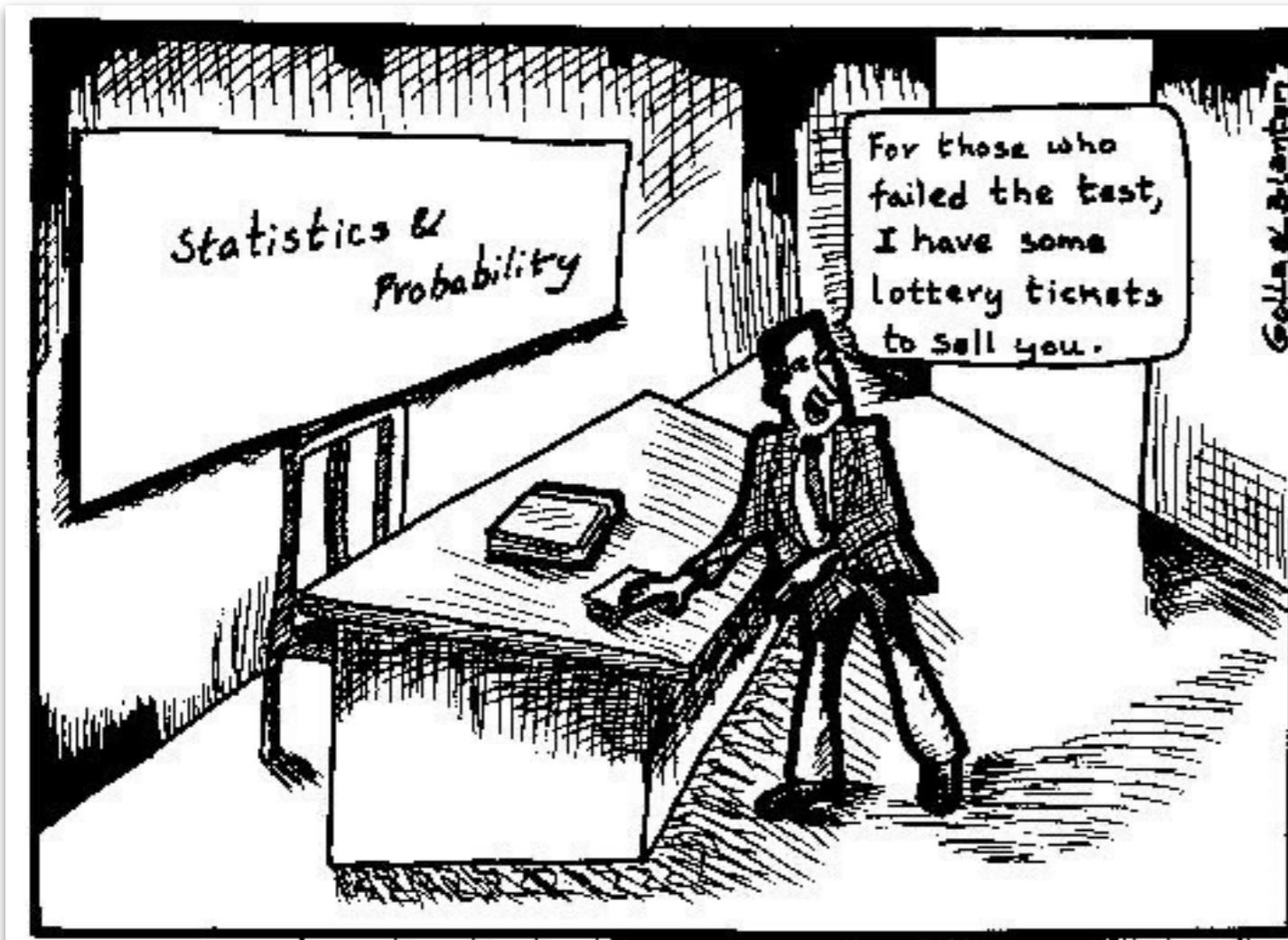
A screenshot of a Spotify collaborative playlist page. The title is "psych252". Below the title is a link: "https://tinyurl.com/psych252spotify21". At the bottom are two buttons: "PLAY" and "...". To the left of the title is a graphic of a bar chart with a smiling face, where the bars are blue and the chart has a friendly appearance.

01/25/2021

Remember to
record the
lecture!



Probability & Causality



Chat

What's the present you most wanted as a kid?

To: Everyone ▾ More ▾

Type message here...

O COLLABORATIVE PLAYLIST
psych252

<https://tinyurl.com/psych252spotify21>

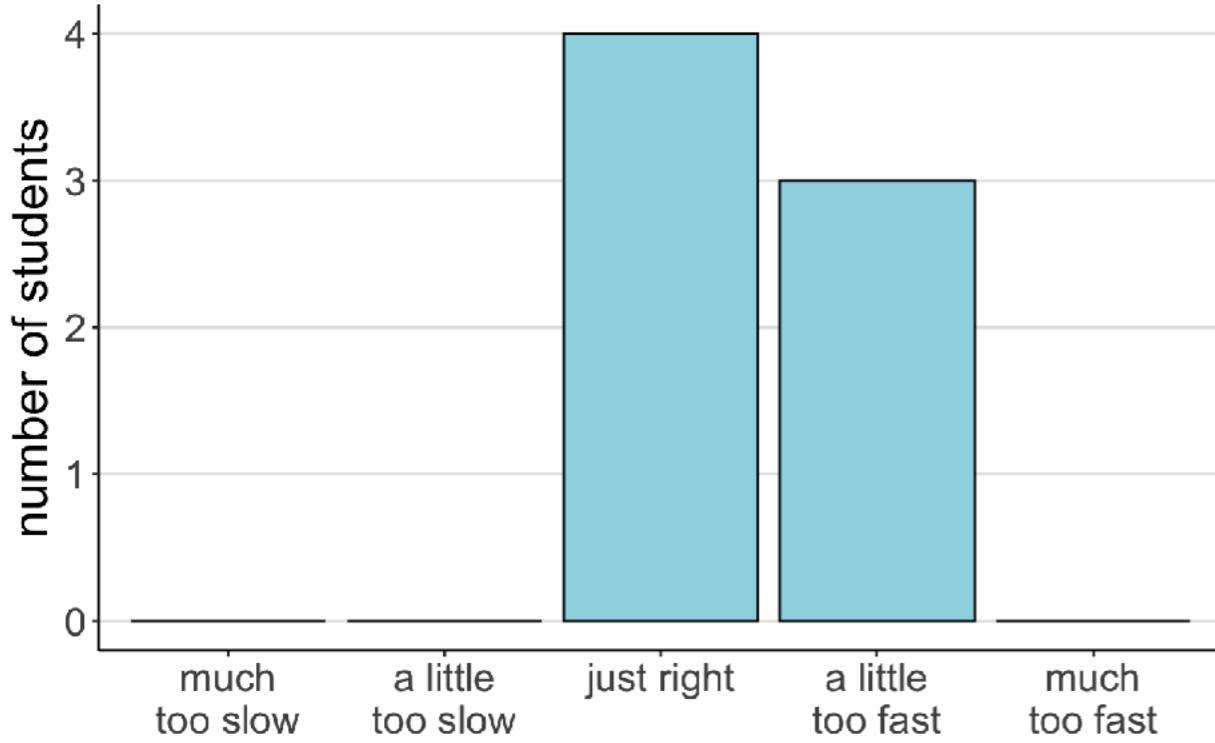
PLAY ...

01/25/2021

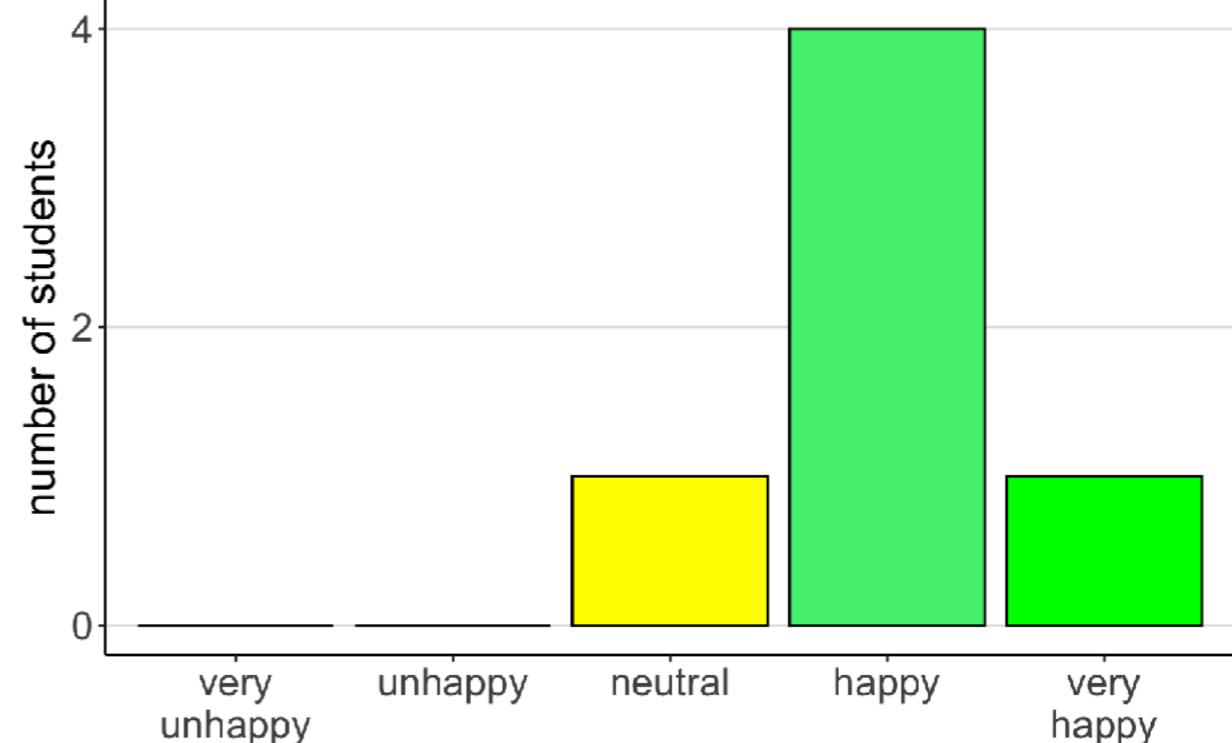
Your feedback

Your feedback

How was the pace of today's class?



How happy were you with today's class overall?



I really liked the pace! The breakout rooms were a little unbalanced so it might be better to ensure equal numbers in all of them for a better learning experience.

Outline

- Introduction to probability / Recap
 - Motivation
 - Counting possibilities
 - **Clue** guide to probability
- Causal Bayesian Networks
 - seeing (prediction) vs. doing (explanation)
 - correlation is not causation

Motivation

What does statistics have to do with probability?



Theory

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



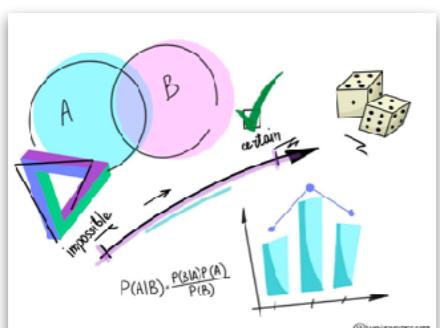
Prediction

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.

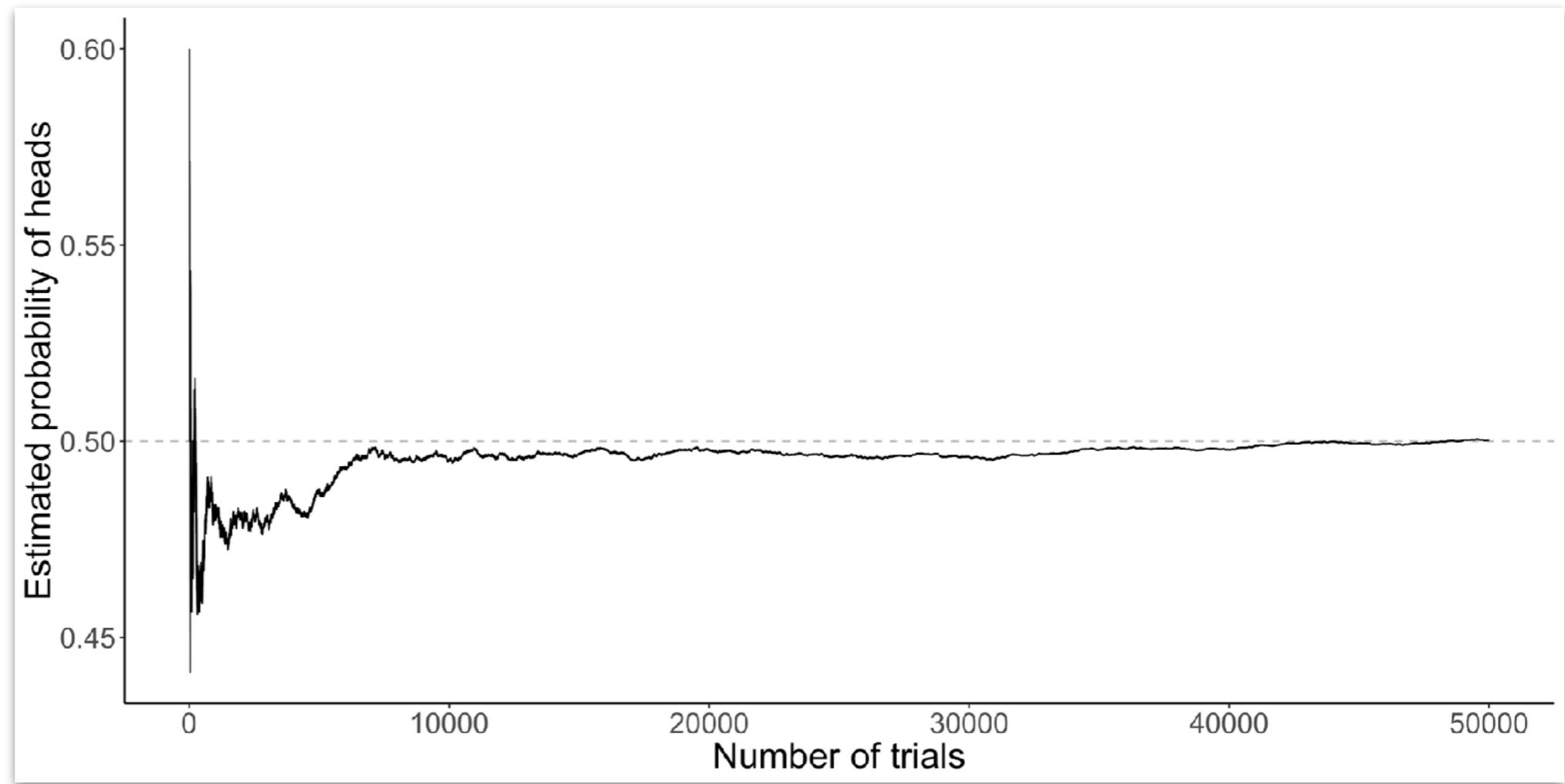


Probability

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

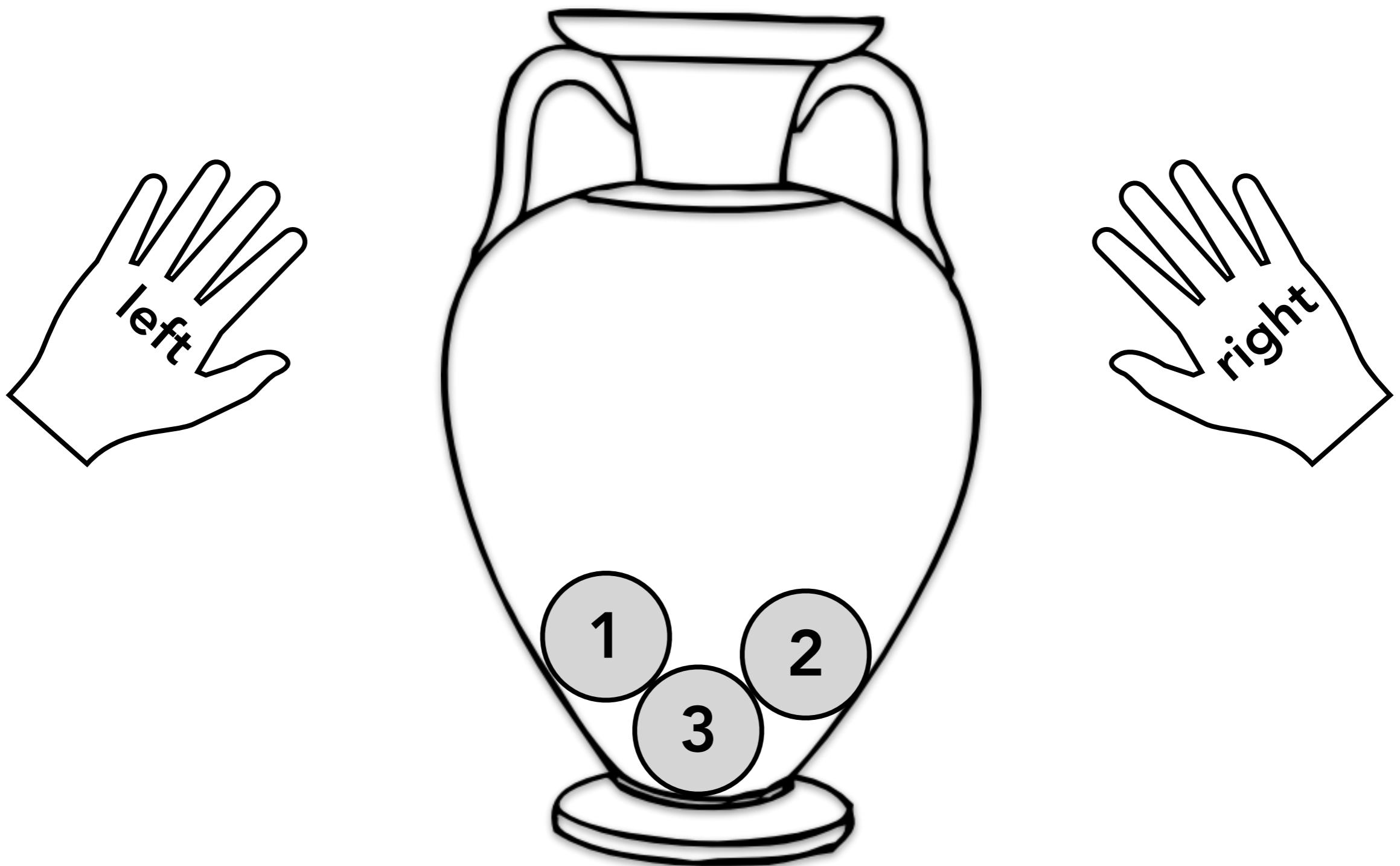
Subjective/Bayesian interpretation

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

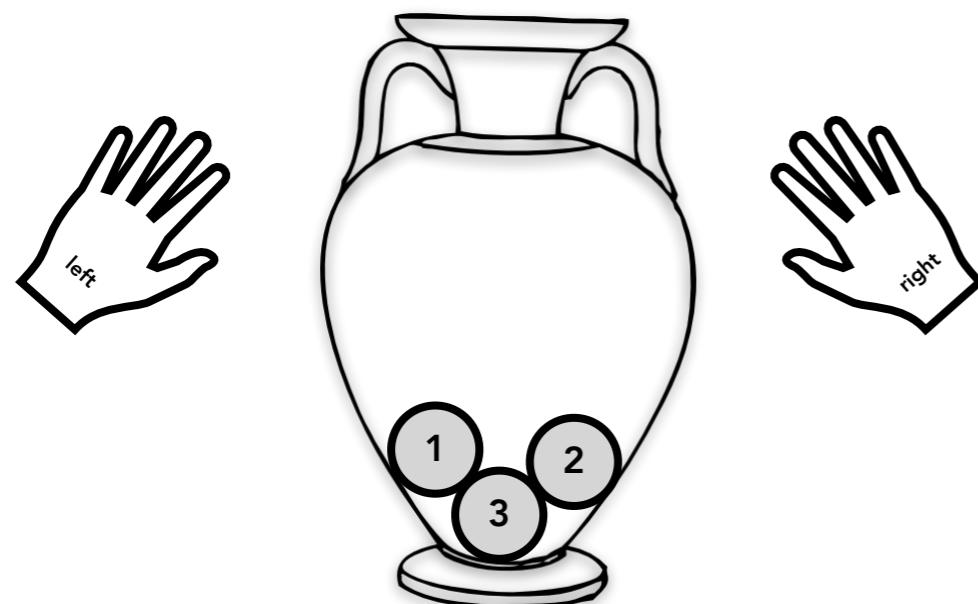
Counting possibilities

no stats class without urns!



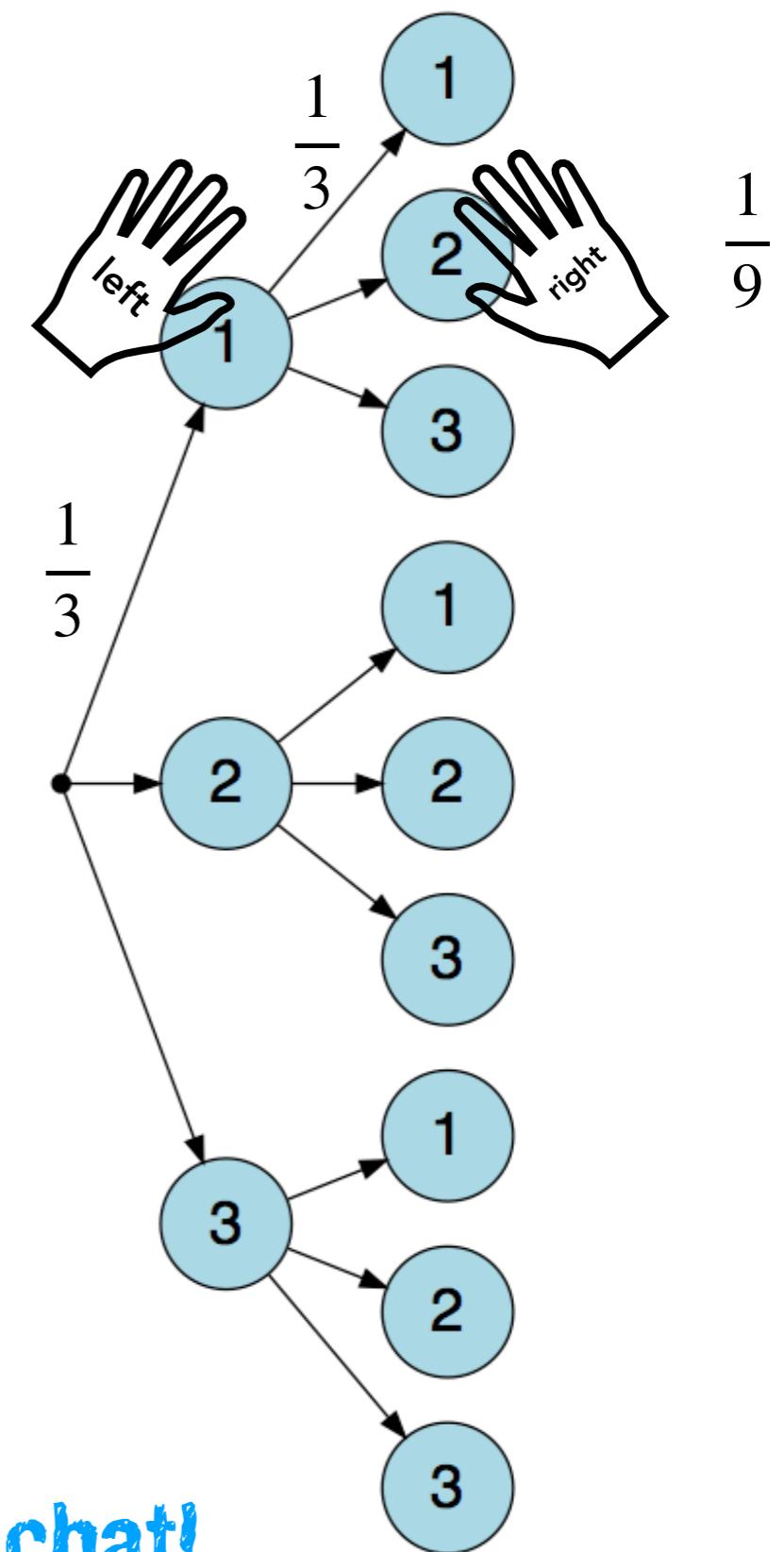
Sampling with replacement

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



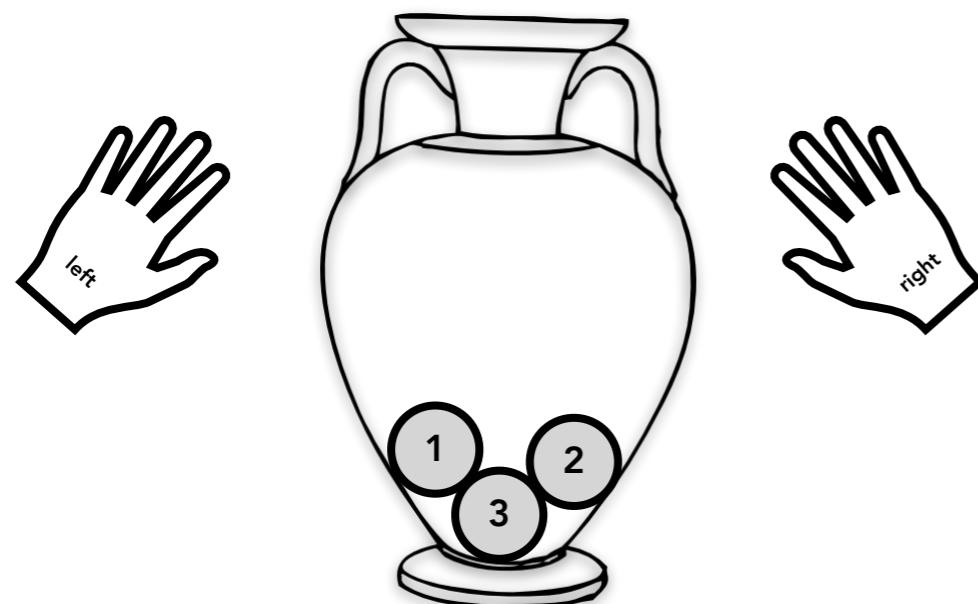
What is the probability that I first draw the 1 with my left hand, and then, after putting the 1 back into the urn again, draw the 2 with my right hand?

→ write in the chat!



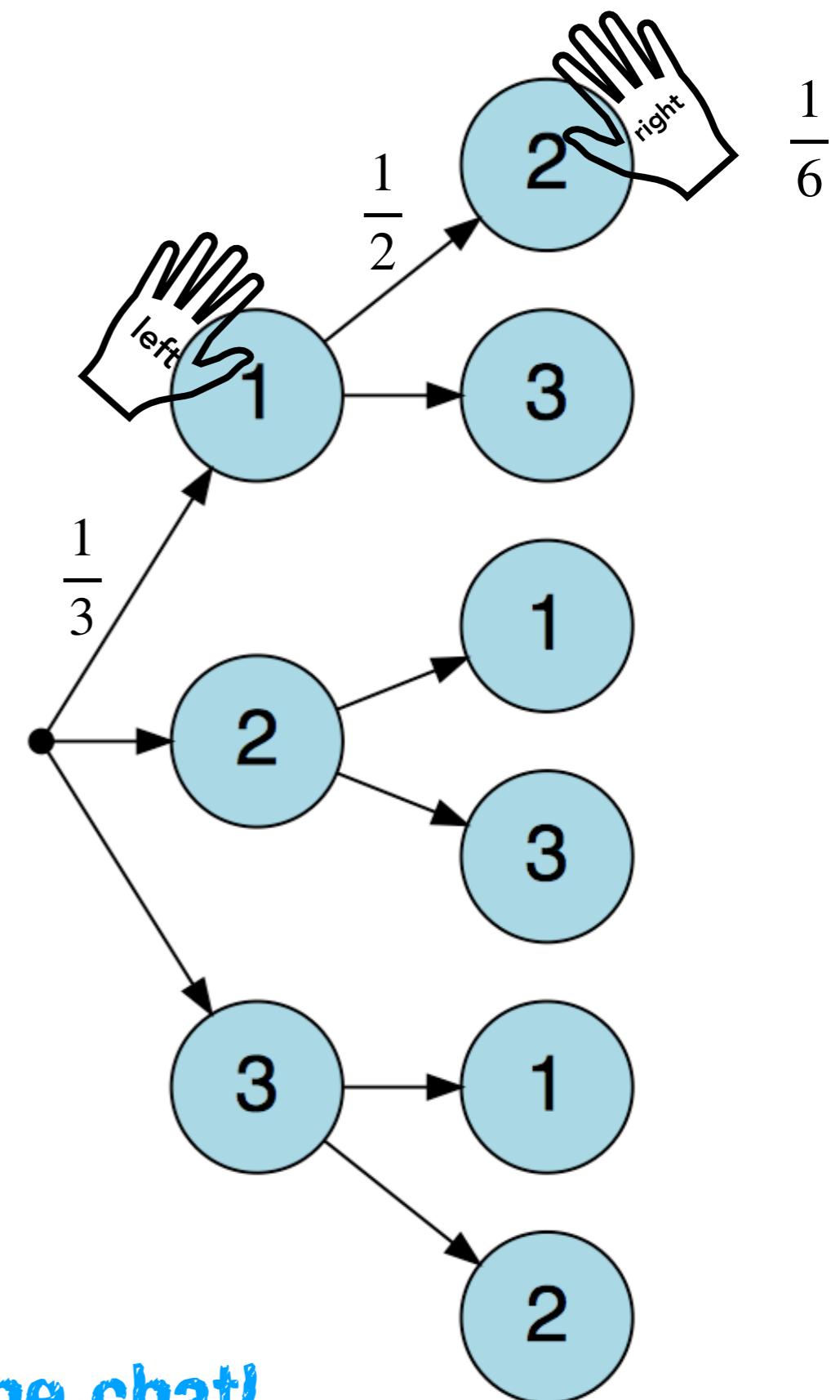
Sampling without replacement

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



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?

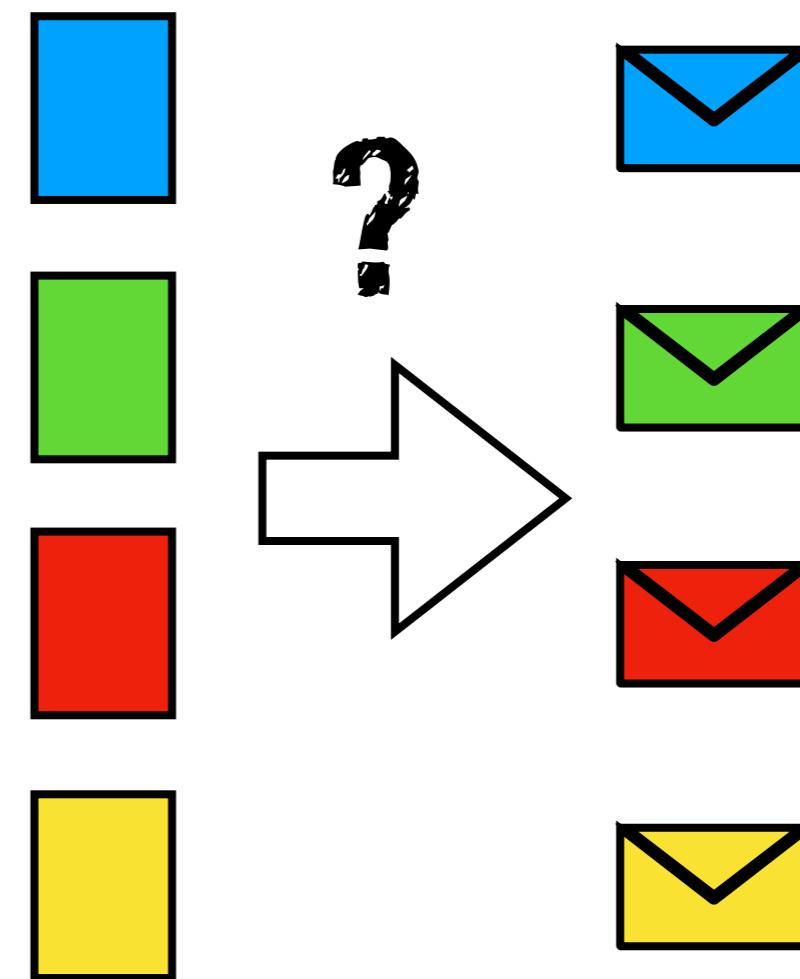
write in the chat!



Random secretary



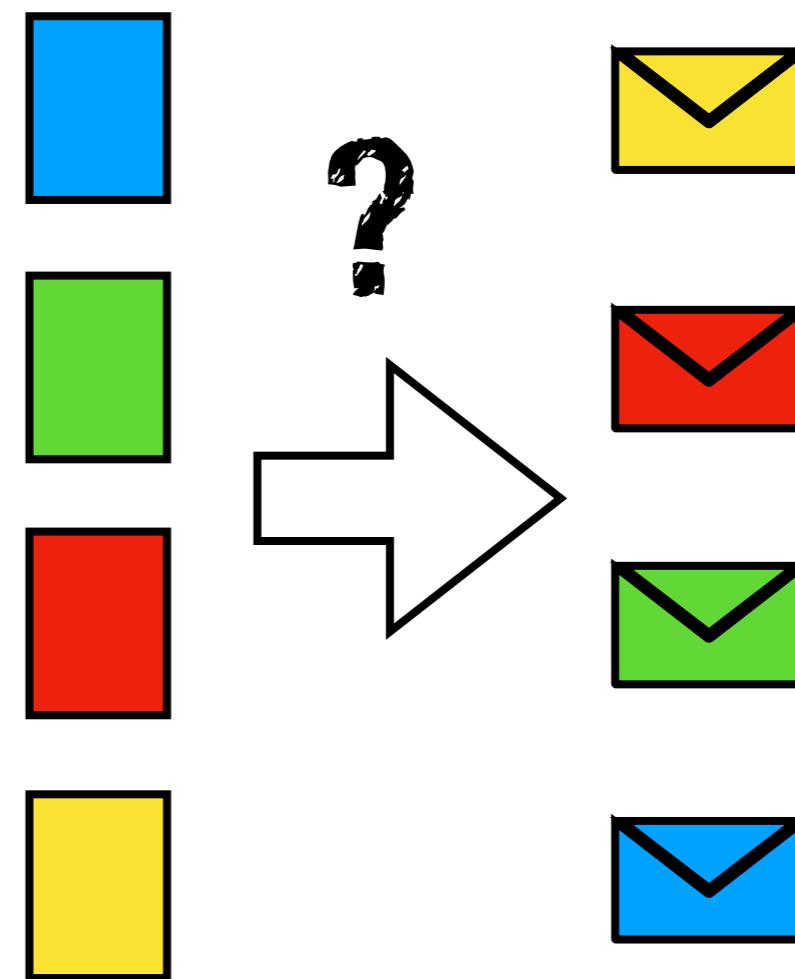
A secretary types four letters to four people and addresses the four envelopes. If he inserts the letters at random, each in a different envelope, what is the probability that exactly three letters will go into the right envelope?



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?

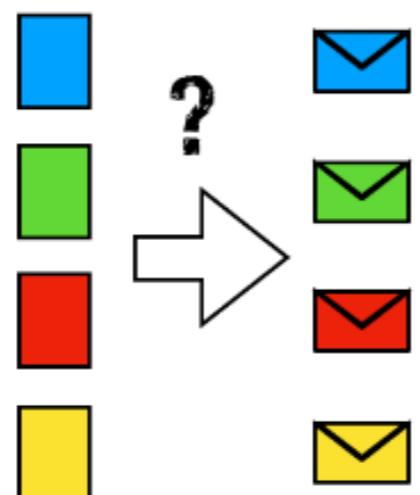


What is the probability that exactly three letters will go into the right envelope?

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?



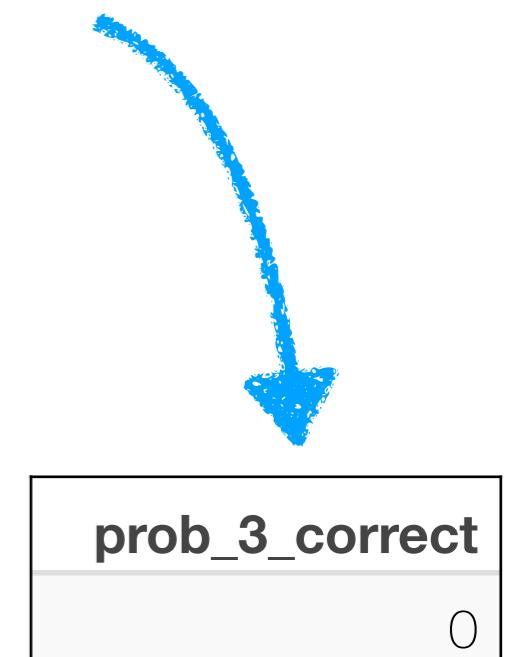
0% 25% 50% 75% 100%

Random secretary

```
1 df.letters = permutations(x = 1:4, k = 4) %>%
2   as_tibble(.name_repair = ~ str_c("person_", 1:4))
```

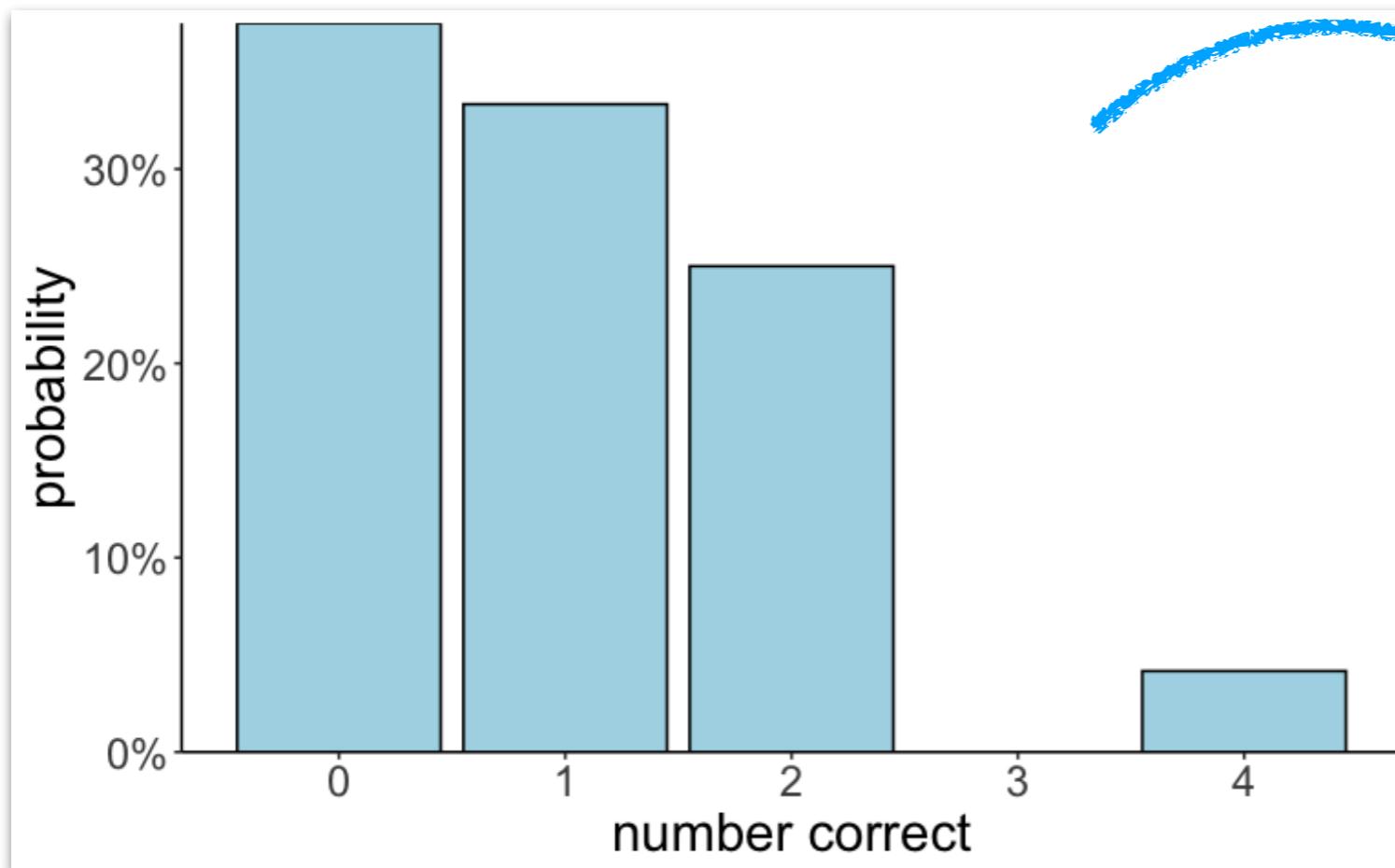
person_1	person_2	person_3	person_4
1	2	3	4
1	2	4	3
1	3	2	4
1	3	4	2
1	4	2	3
1	4	3	2
2	1	3	4
2	1	4	3
2	3	1	4

⋮ 24 rows total



Random secretary

```
1 ggplot(data = df.letters,
2         mapping = aes(x = n_correct)) +
3   geom_bar(aes(y = stat(count)/sum(count)),
4           color = "black",
5           fill = "lightblue") +
6   scale_y_continuous(labels = scales::percent,
7                      expand = c(0, 0)) +
8   labs(x = "number correct",
9        y = "probability")
```



probability of getting
0, 1, 2, 3, or 4
envelopes to the
correct person

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!

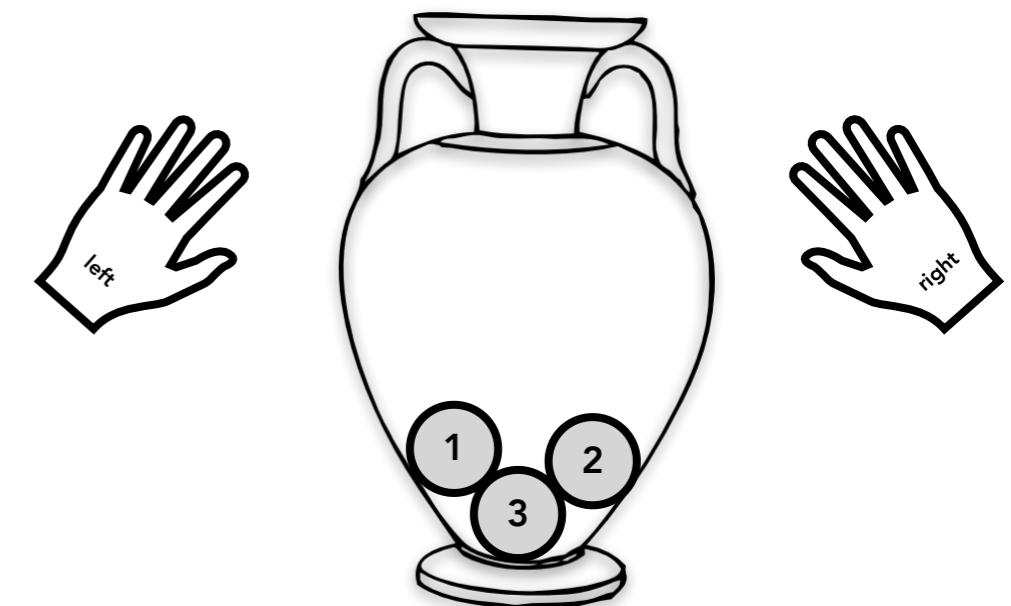
Definitions

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 outcomes for an experiment.

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



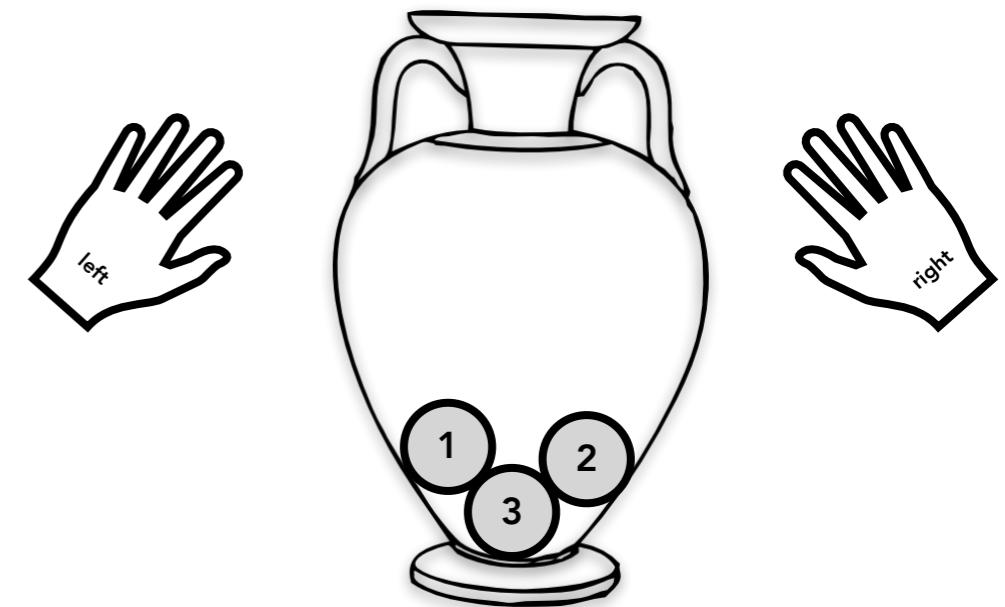
Event: Subset of the sample space. $(1, 1)$

Definitions

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

1. Probability cannot be negative.

$$P(X_i) \geq 0$$



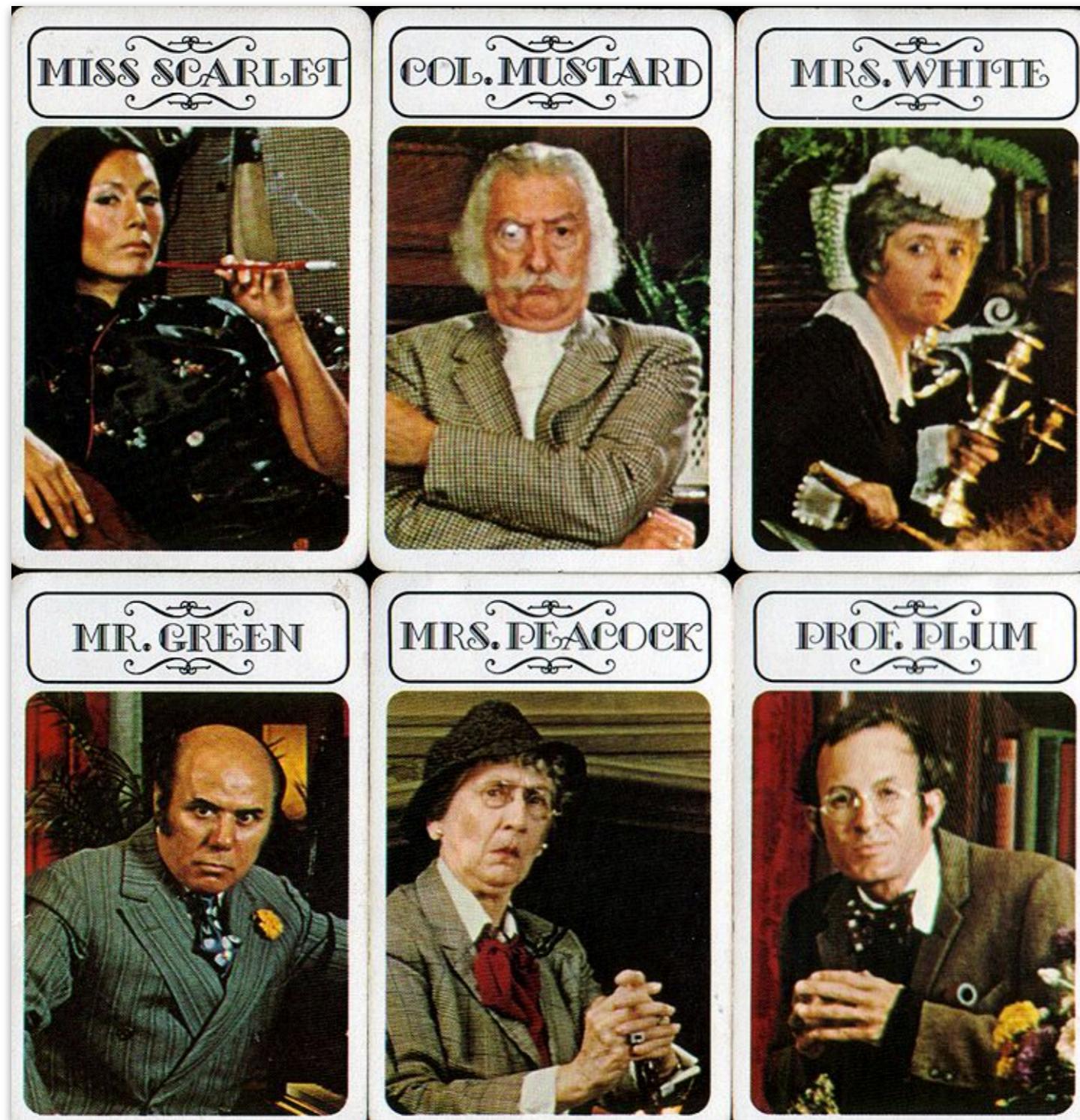
2. Total probability of all outcomes in the sample space is 1.

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

clue guide to probability

Clue guide to probability

Who killed Mr Boddy?



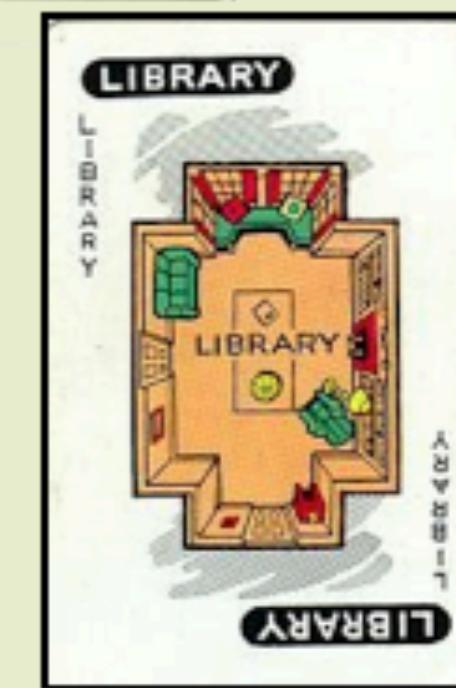
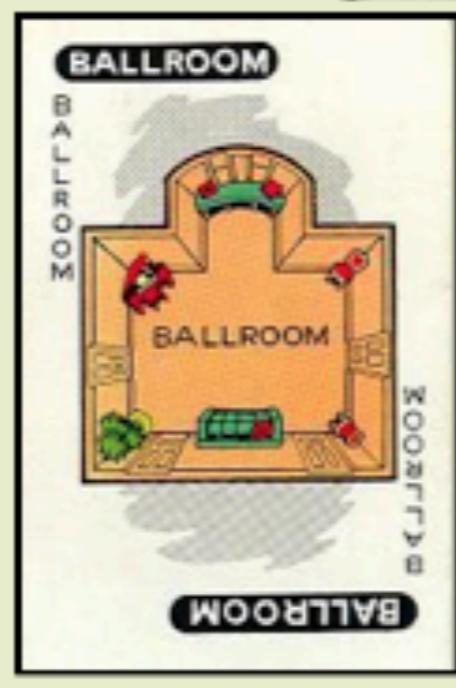
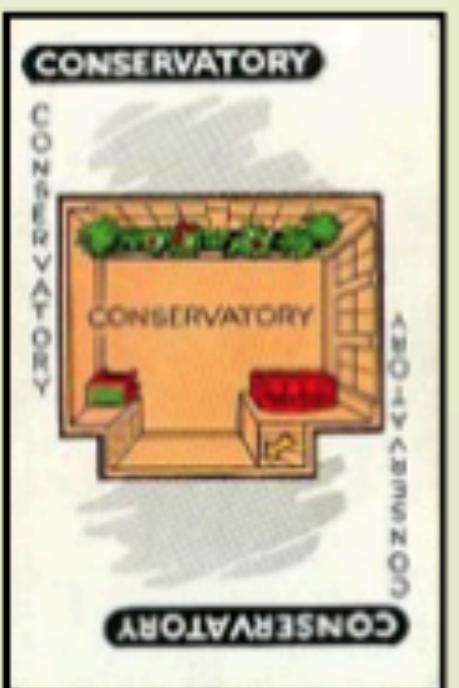
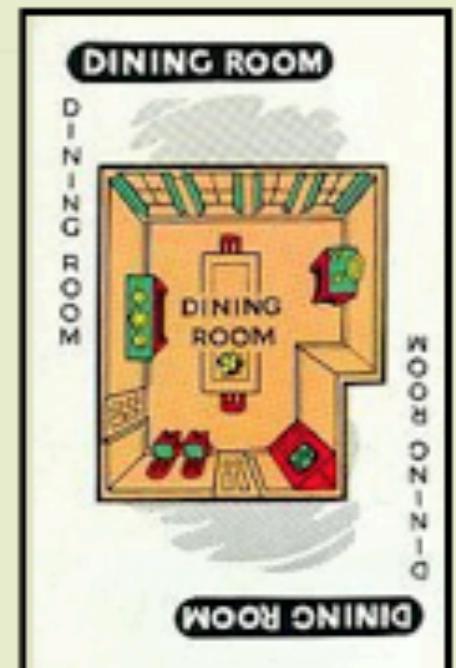
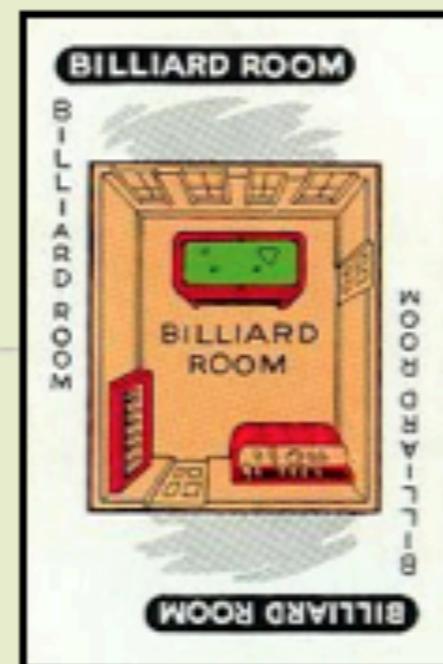
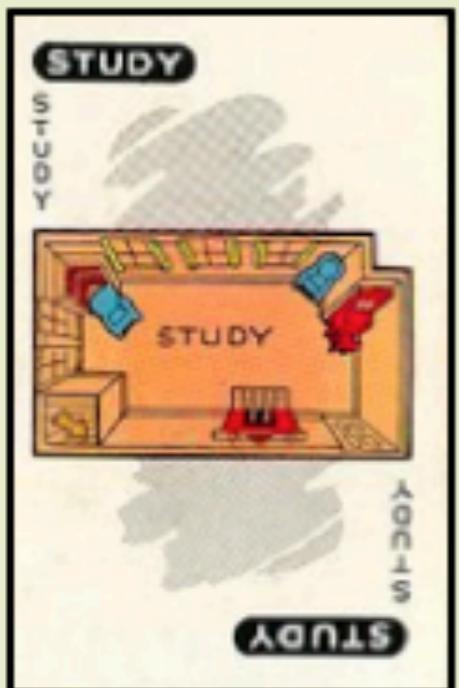
Clue guide to probability

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

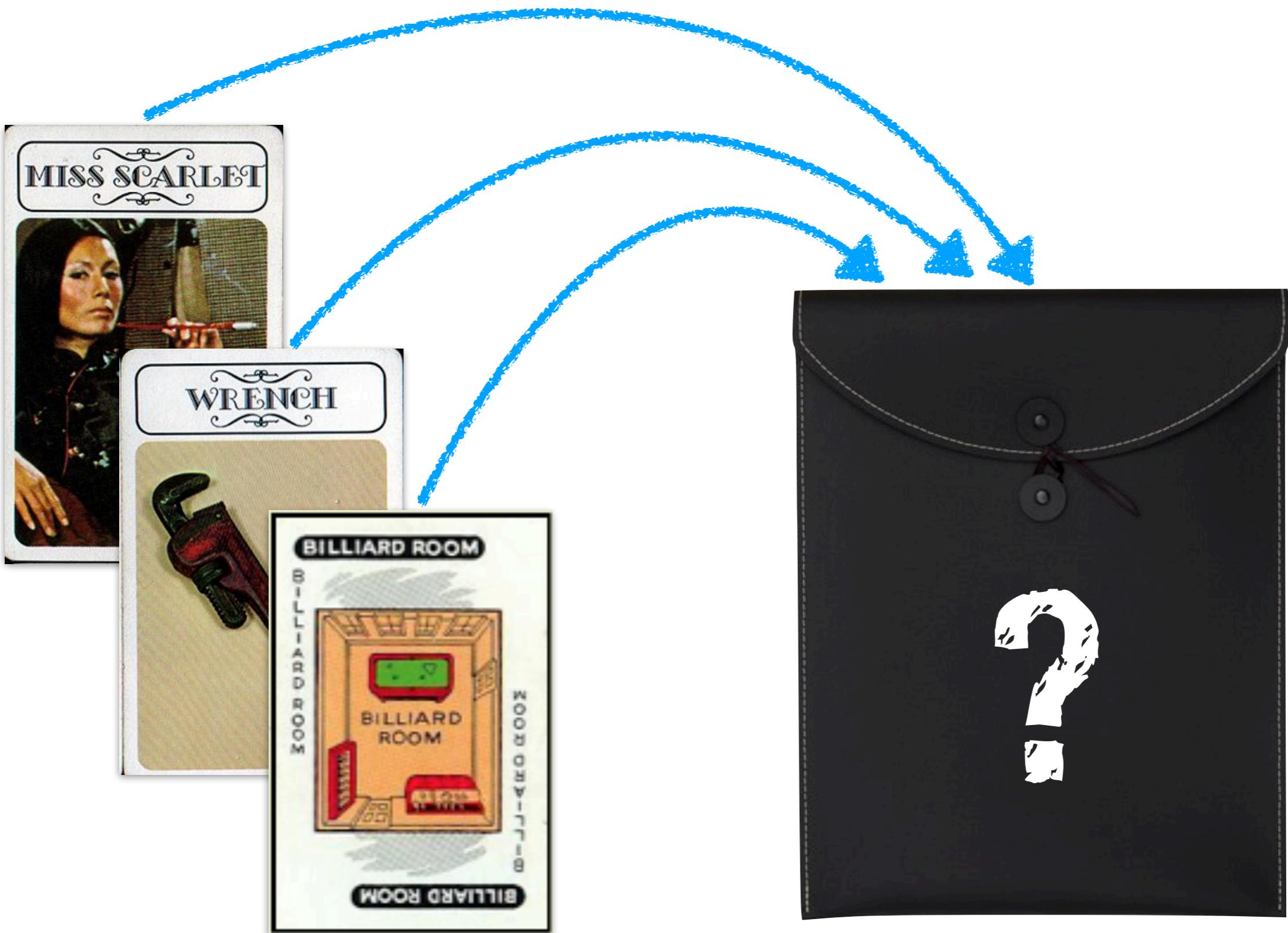


Clue guide to probability

Who killed Mr Boddy, with what, and where?



Clue guide to probability



Clue guide to probability

```
1 who = c("ms_scarlet", "col_mustard", "mrs_white",
2       "mr_green", "mrs_peacock", "prof_plum")
3 what = c("candlestick", "knife", "lead_pipe",
4         "revolver", "rope", "wrench")
5 where = c("study", "kitchen", "conservatory",
6           "lounge", "billiard_room", "hall",
7           "dining_room", "ballroom", "library")
8
9 df.clue = expand_grid(who = who,
10                      what = what,
11                      where = where)
```

all combinations

Ω

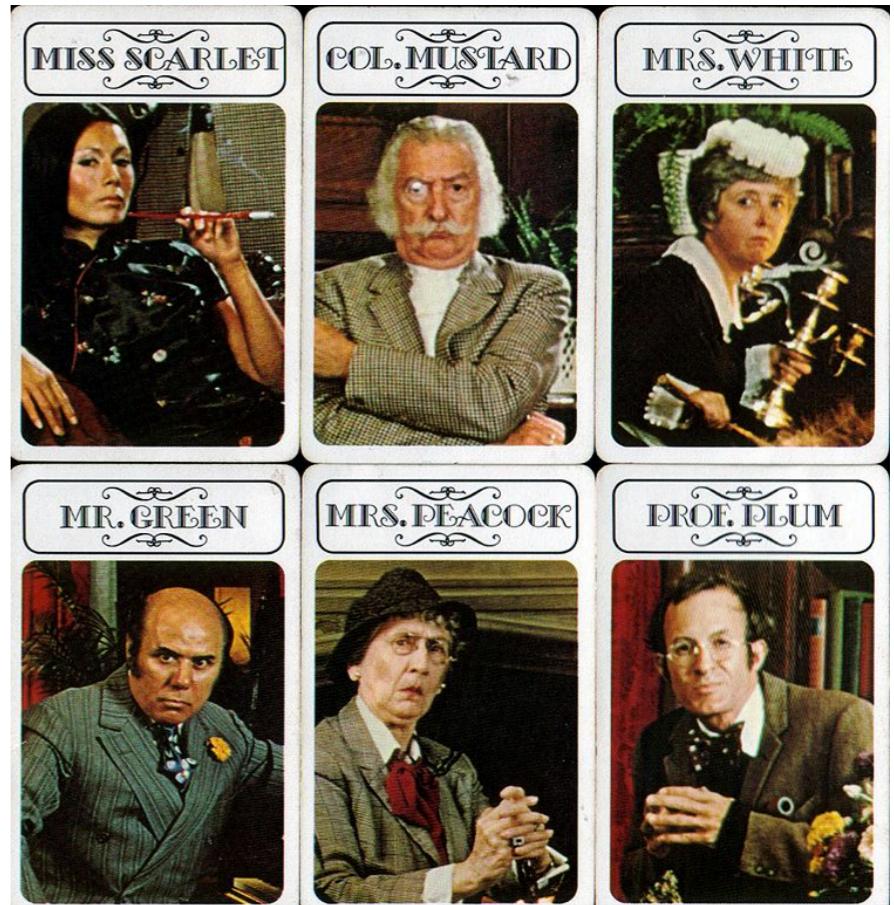
who	what	where
ms_scarlet	candlestick	study
ms_scarlet	candlestick	kitchen
ms_scarlet	candlestick	conservatory
ms_scarlet	candlestick	lounge
	⋮	

nrow(df.clue) = 324

Clue guide to probability

Who?

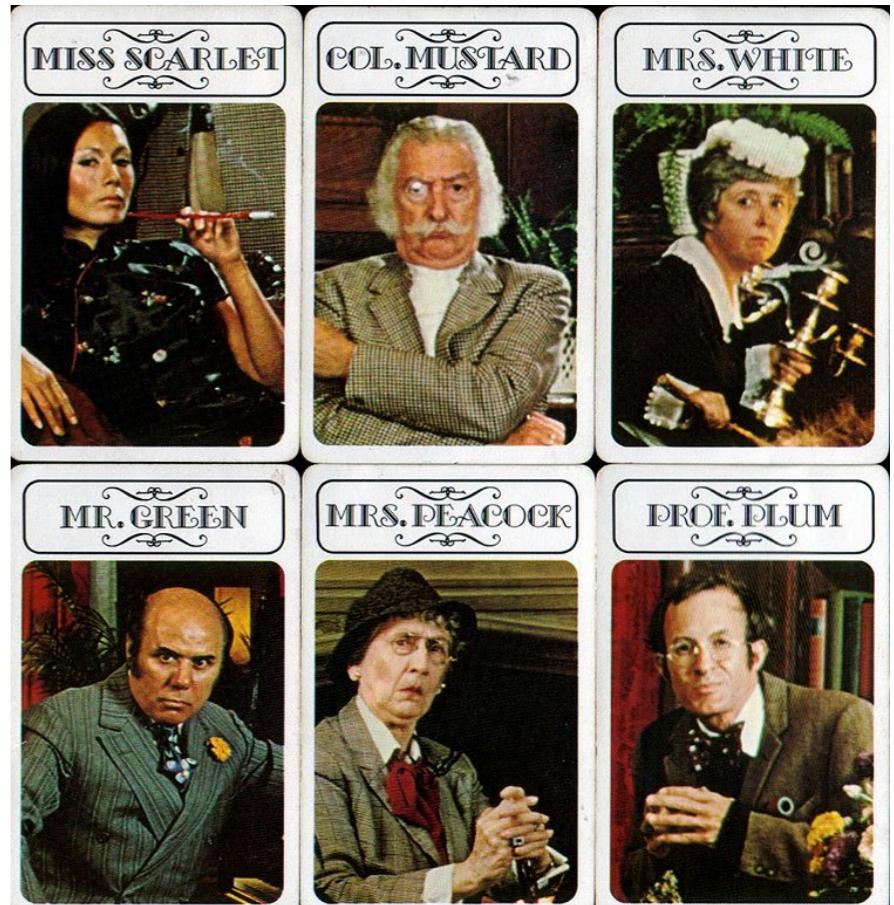
- 6 suspects
- mutually exclusive and exhaustive
- $p(\text{who} = \text{one of the six}) = 1$
- each equally likely a priori
- $p(\text{who} = \text{Prof. Plum}) = \frac{1}{6}$



Clue guide to probability

Who?

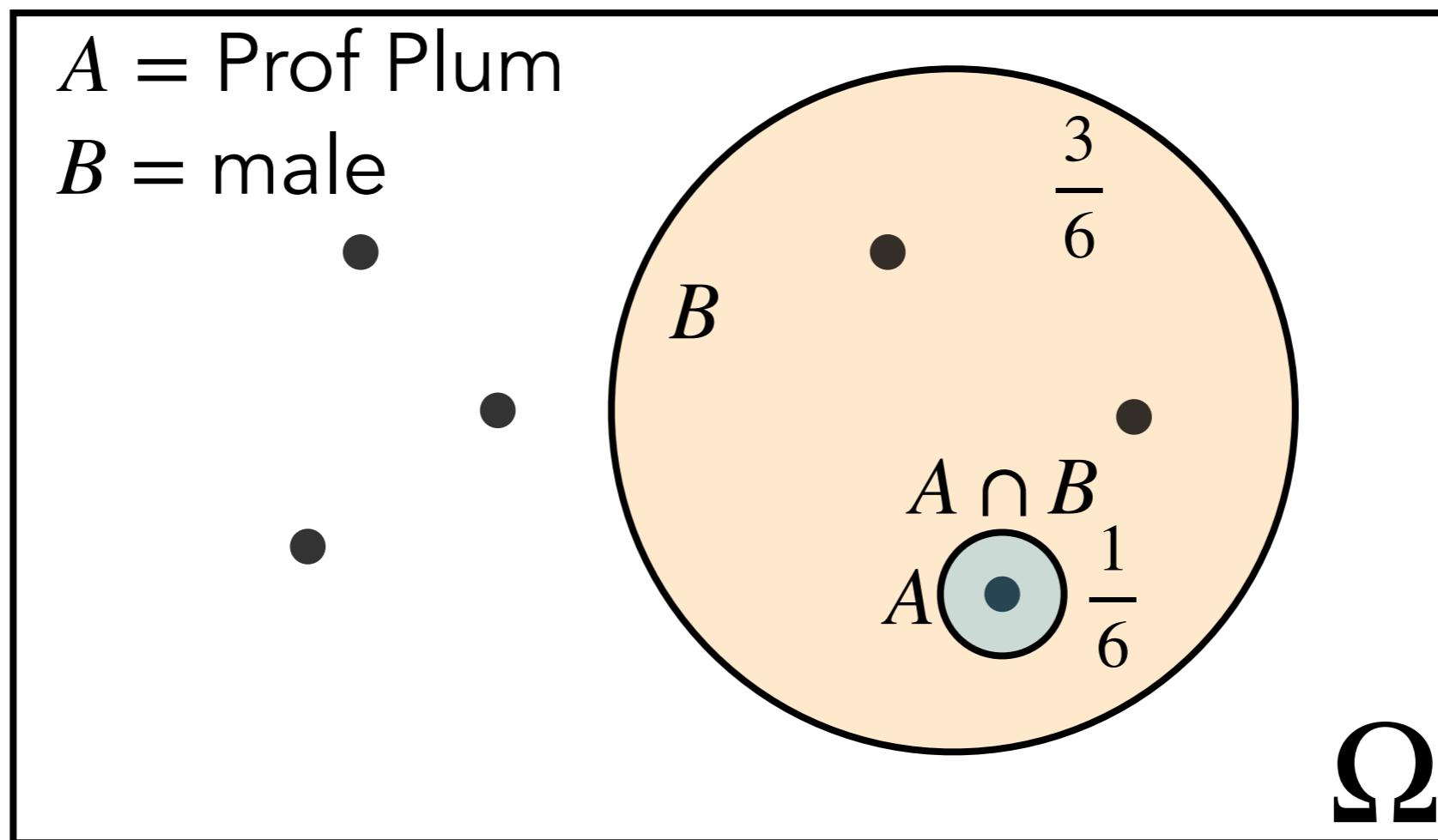
- *conditional probability*:
- $p(A | B)$ (probability of A given B)
- **Definition:** $p(A | B) = \frac{p(A, B)}{p(B)}$



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

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

Clue guide to probability



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

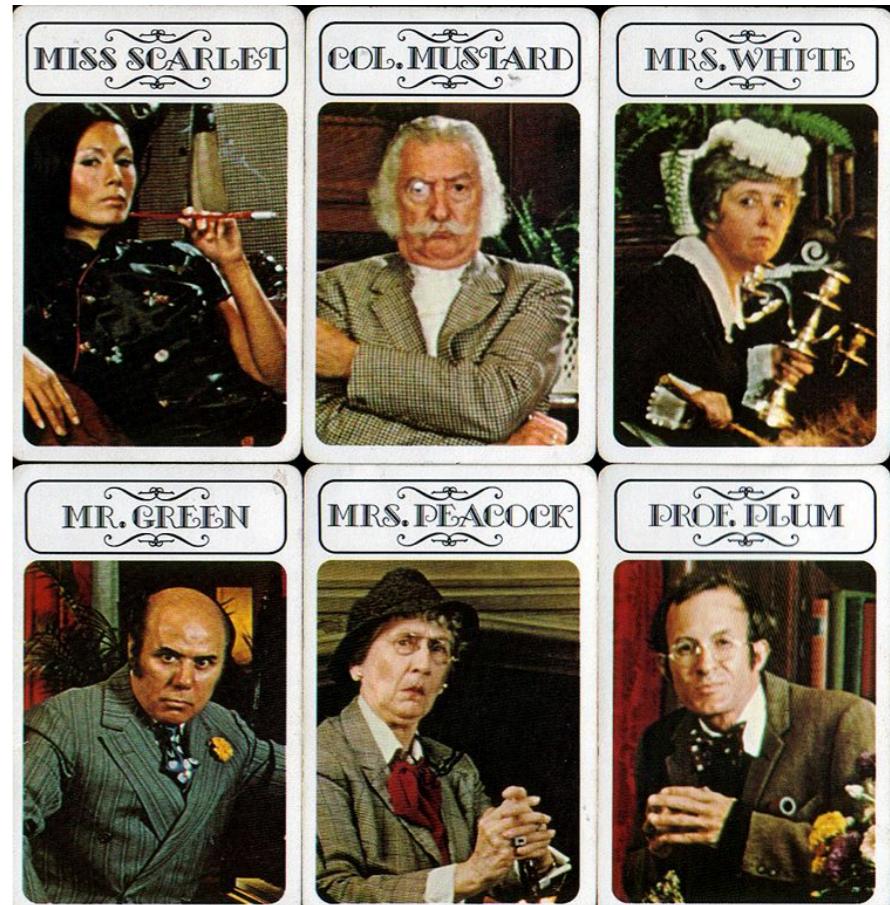
Definition: $p(A | B) = \frac{p(A, B)}{p(B)} = \frac{1}{3}$

$$p(A) = \frac{1}{6} \quad p(A, B) = \frac{1}{6} \quad p(B) = \frac{3}{6}$$

Clue guide to probability

Who?

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



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

```

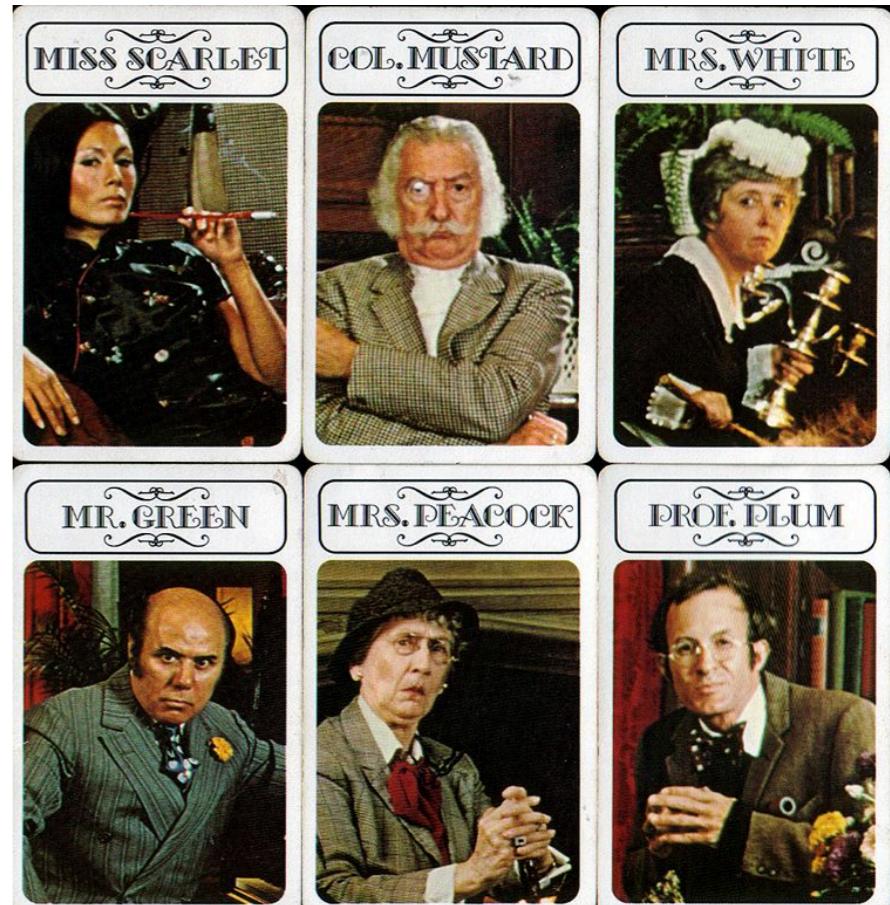
1 df.suspects = df.clue %>%
2   distinct(who) %>%
3   mutate(gender = ifelse(
4     test = who %in% c("ms_scarlet",
5                           "mrs_white",
5                           "mrs_peacock"),
6     yes = "female",
6     no = "male"))

```

Clue guide to probability

Who?

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



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

```

1 df.suspects %>%
2   summarize(p_prof_plum_given_male =
3     sum(gender == "male" &
4       who == "prof_plum") /
5     sum(gender == "male"))

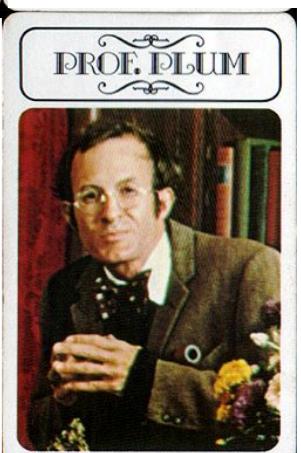
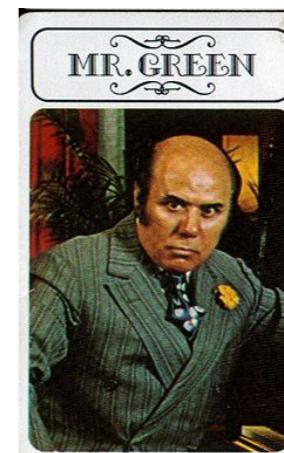
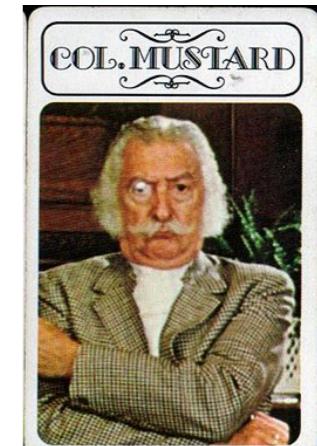
```

use naive definition of probability

Clue guide to probability

Who?

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

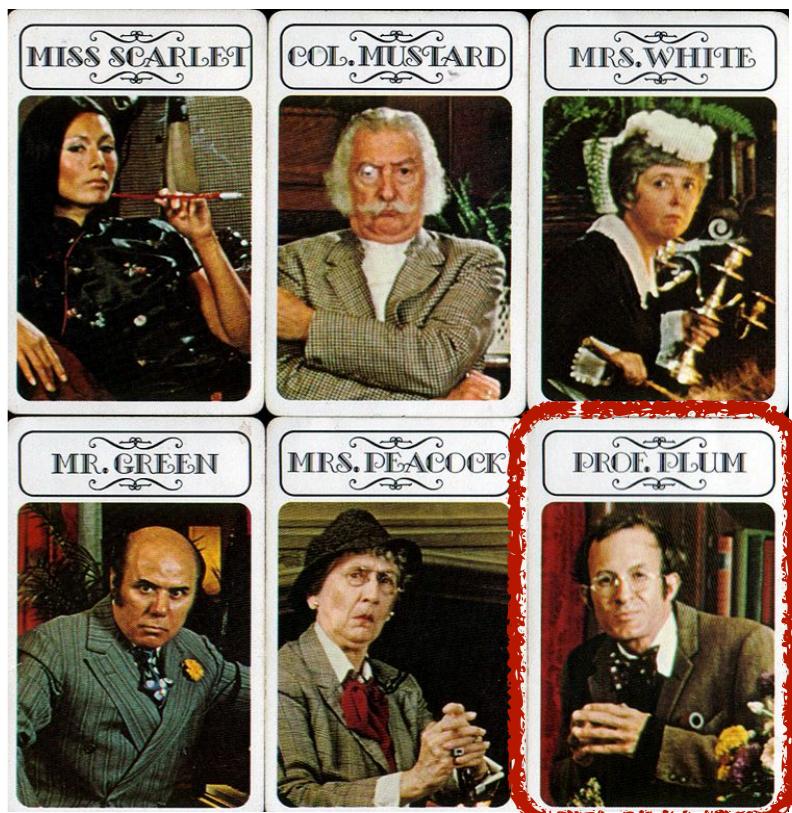


who	gender
col_mustard	male
mr_green	male
prof_plum	male

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

Clue guide to probability

Who?



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

What?



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

Clue guide to probability

Who?

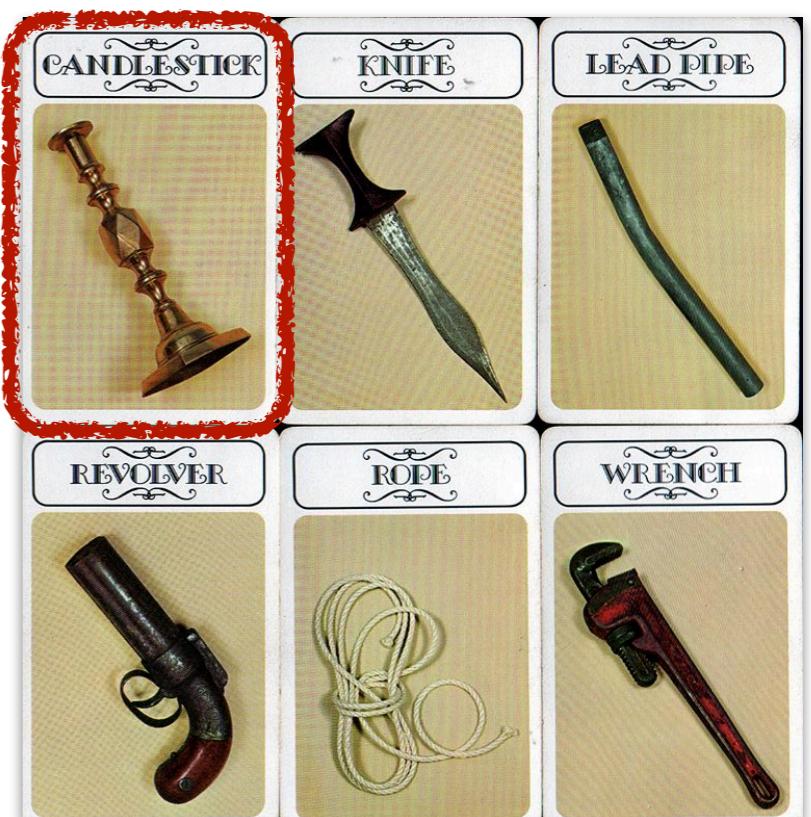


- joint probability:
- if A and B are independent then
- **Definition:** $p(A, B) = p(A) \cdot p(B)$

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

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

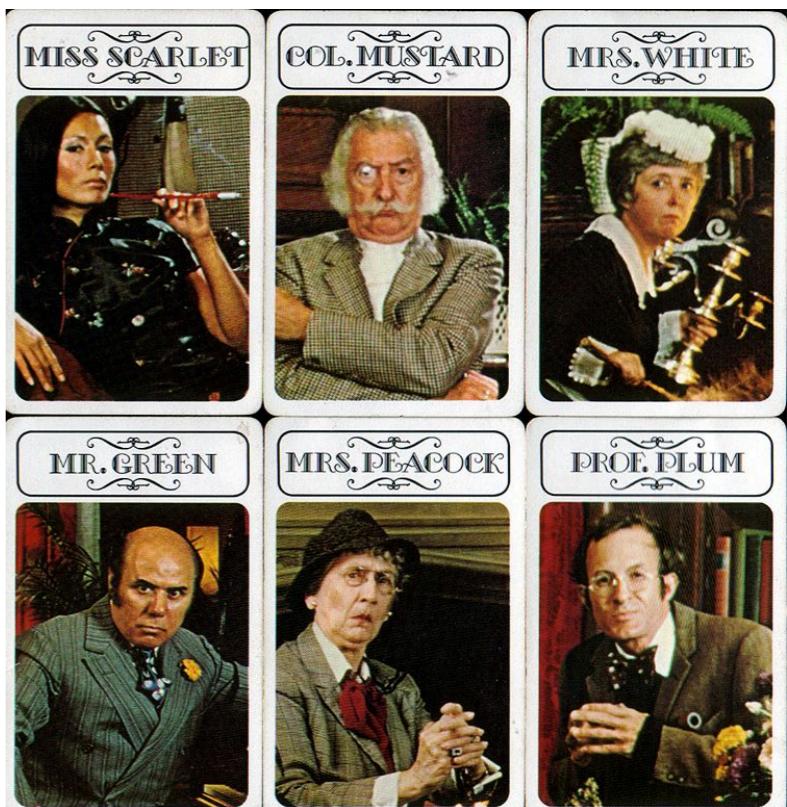
What?



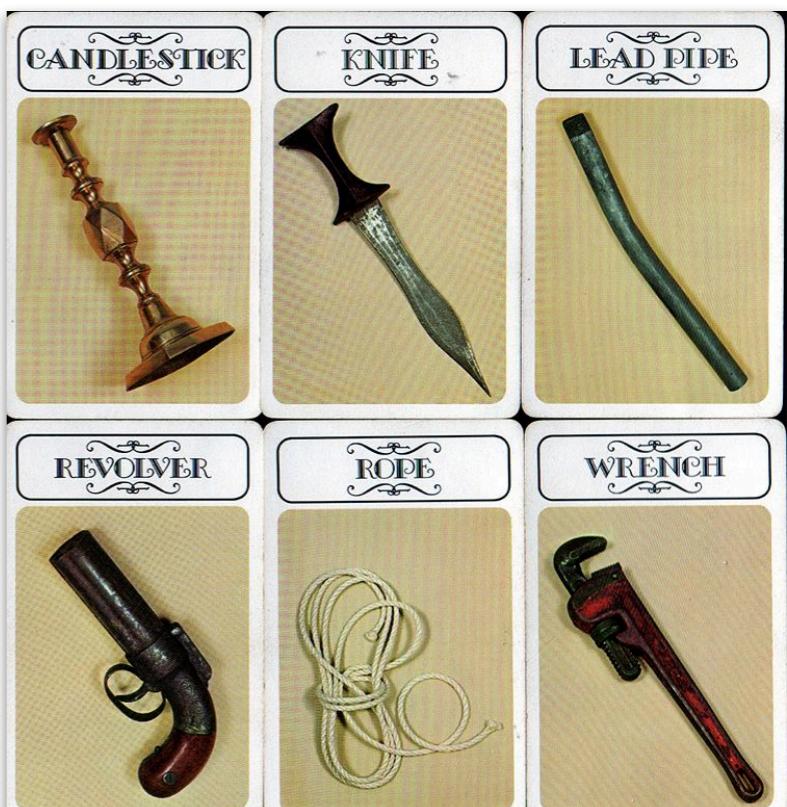
Clue guide to probability

- dependence:
 - **Definition:** $p(A | B) \neq p(A)$
 - **Definition:** $p(A, B) = p(A) \cdot p(B | A)$
-
- if women were more likely than men to use the revolver then
 - $p(\text{Mrs. White} | \text{Revolver}) > p(\text{Mrs. White})$

Who?



What?



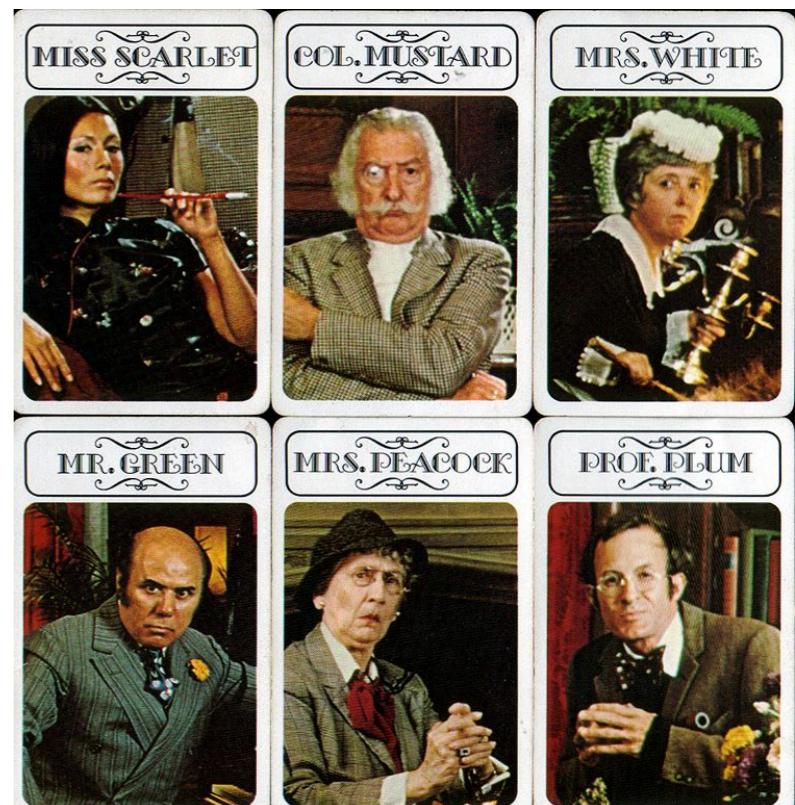
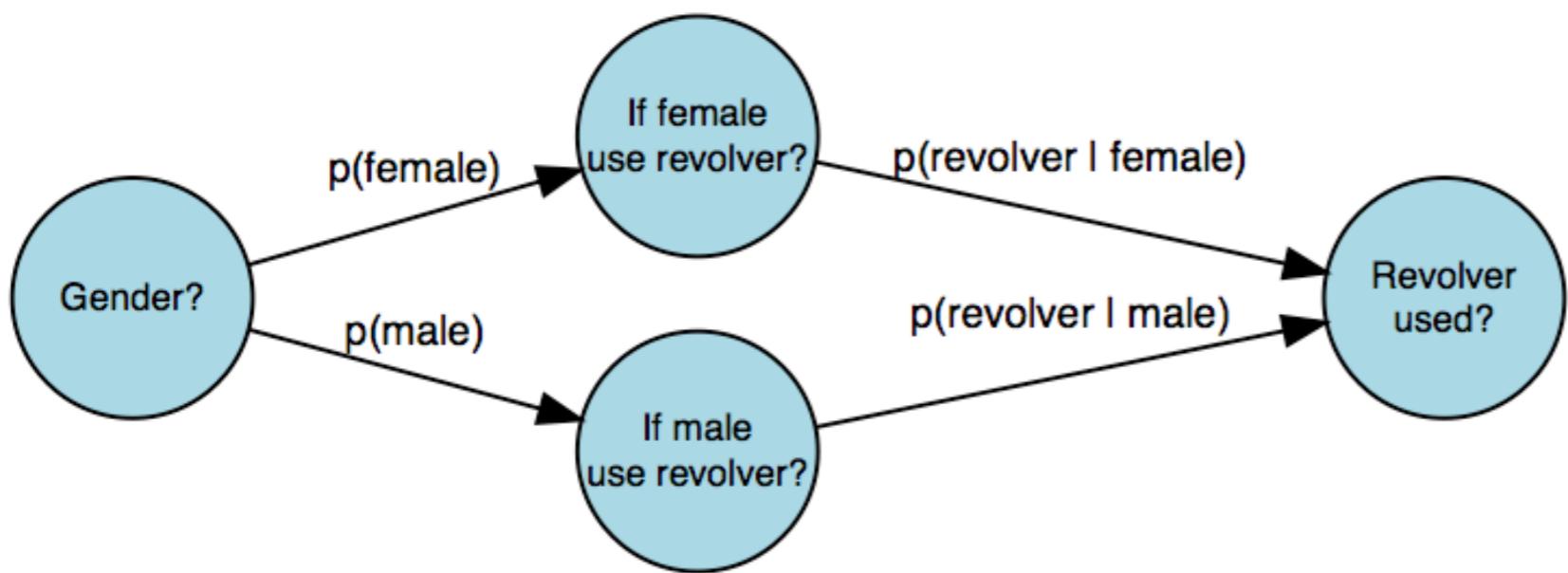
Clue guide to probability

- law of total probability
- Definition:

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

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

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



Who?



Clue guide to probability

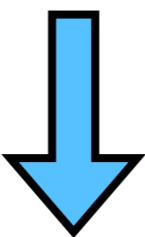
Bayes Theorem in a few steps



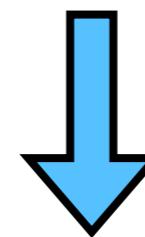
Clue guide to probability

- Bayes' theorem (derivation)

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



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



$$p(A, B) = p(B | A) \cdot p(A) = p(A | B) \cdot p(B)$$



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

Clue guide to probability

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

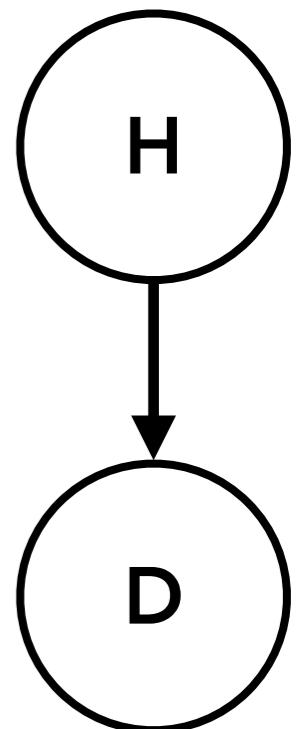
posterior **likelihood** **prior**

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

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

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

H = Hypothesis
 D = Data

probability of the data?!

law of total probability:

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

**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 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(¬T|¬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.

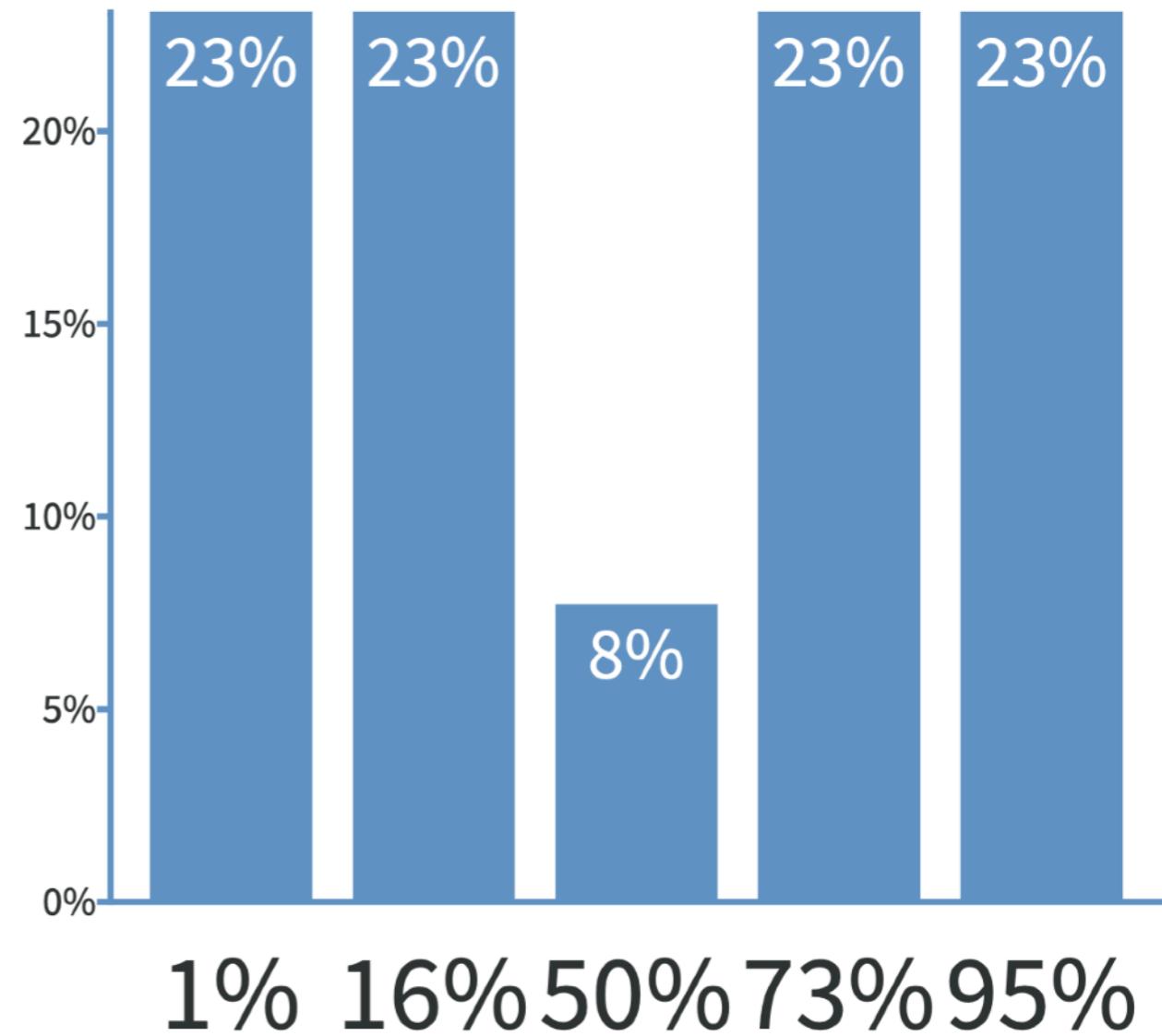
1% 16% 50% 73% 95%

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.



Breakout rooms

Tasks: 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.

Size: ~3 people

Time: 5 minutes

Report: We will vote again.

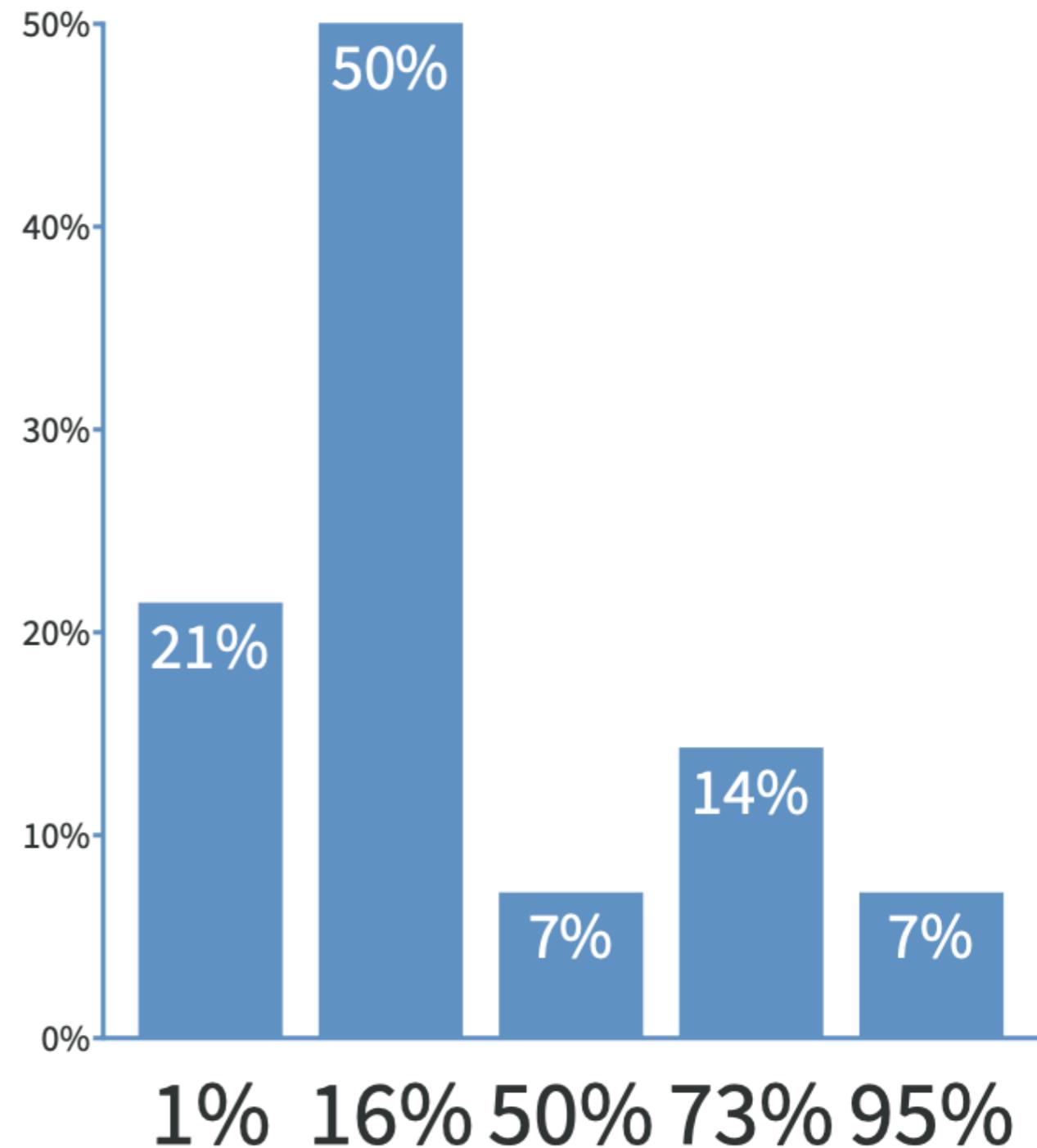


What's the probability that Fred has conditionitis?

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

$$P(D) = 0.01$$

$$P(T|D) = 0.95$$

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

what we want to know

$$P(D|T) = ?$$

$$p(D|T) = \frac{p(T|D) \cdot p(D)}{p(T)} \text{ Bayes' rule}$$

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



law of total
probability

$$p(D|T) = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.05 \cdot 0.99} \approx 0.16$$

Clue guide to probability

what we know

$$P(D) = 0.01$$

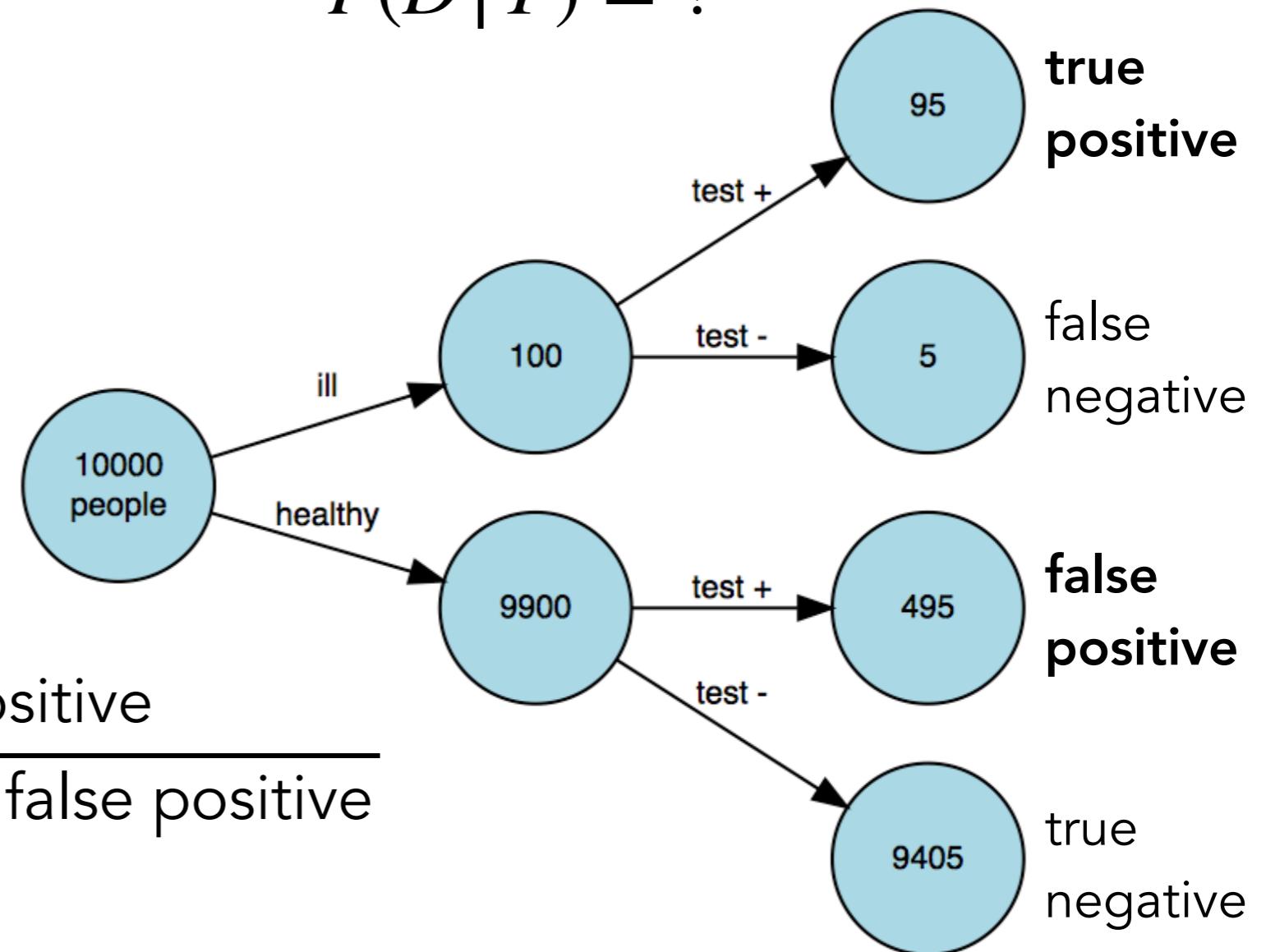
$$P(T|D) = 0.95$$

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

$$\begin{aligned} P(D|T) &= \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \\ &= \frac{95}{95 + 495} \\ &\approx 0.16 \end{aligned}$$

what we want to know

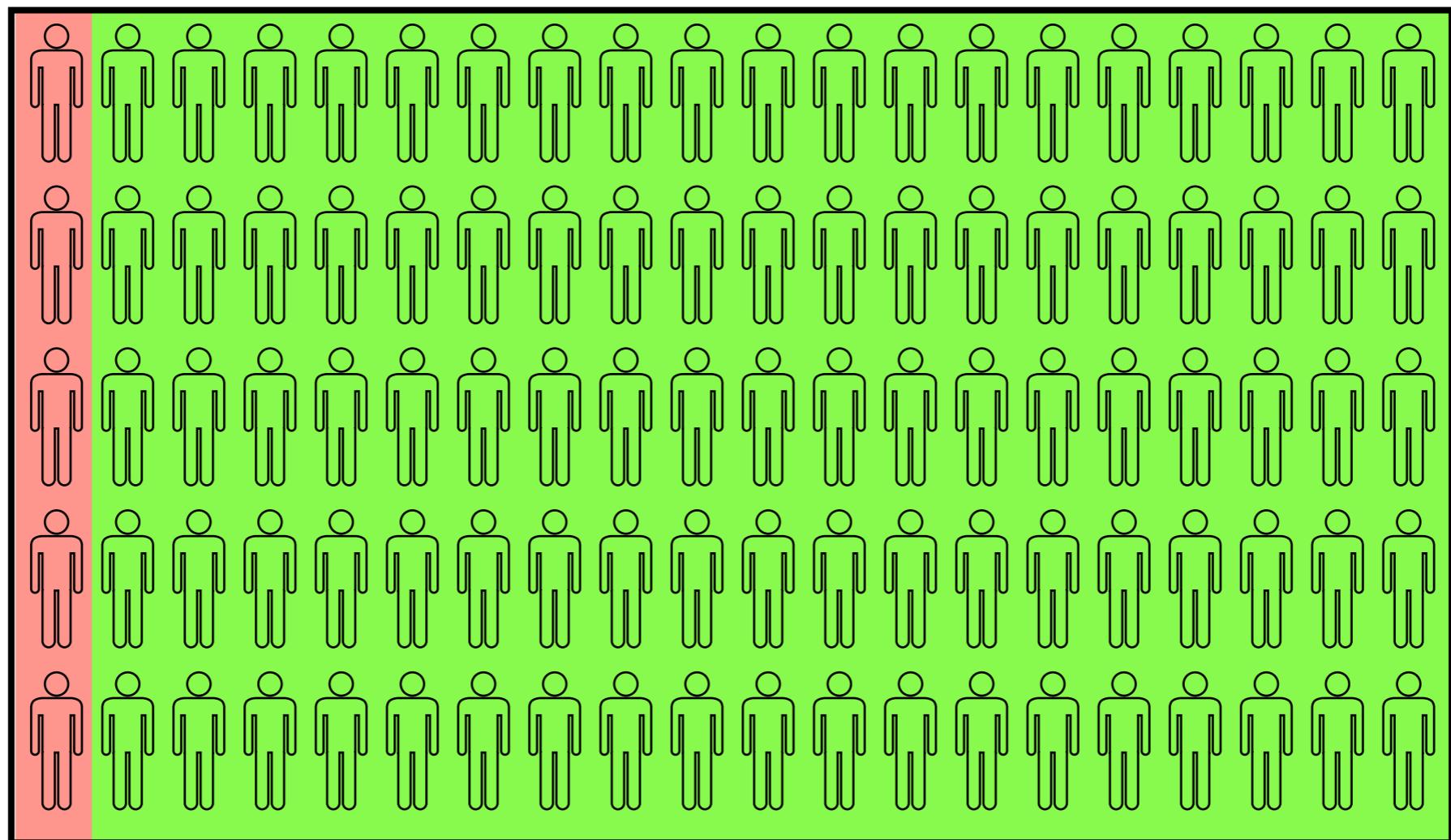
$$P(D|T) = ?$$



Clue guide to probability

this example uses different probabilities than the conditionitis one

$n = 100$



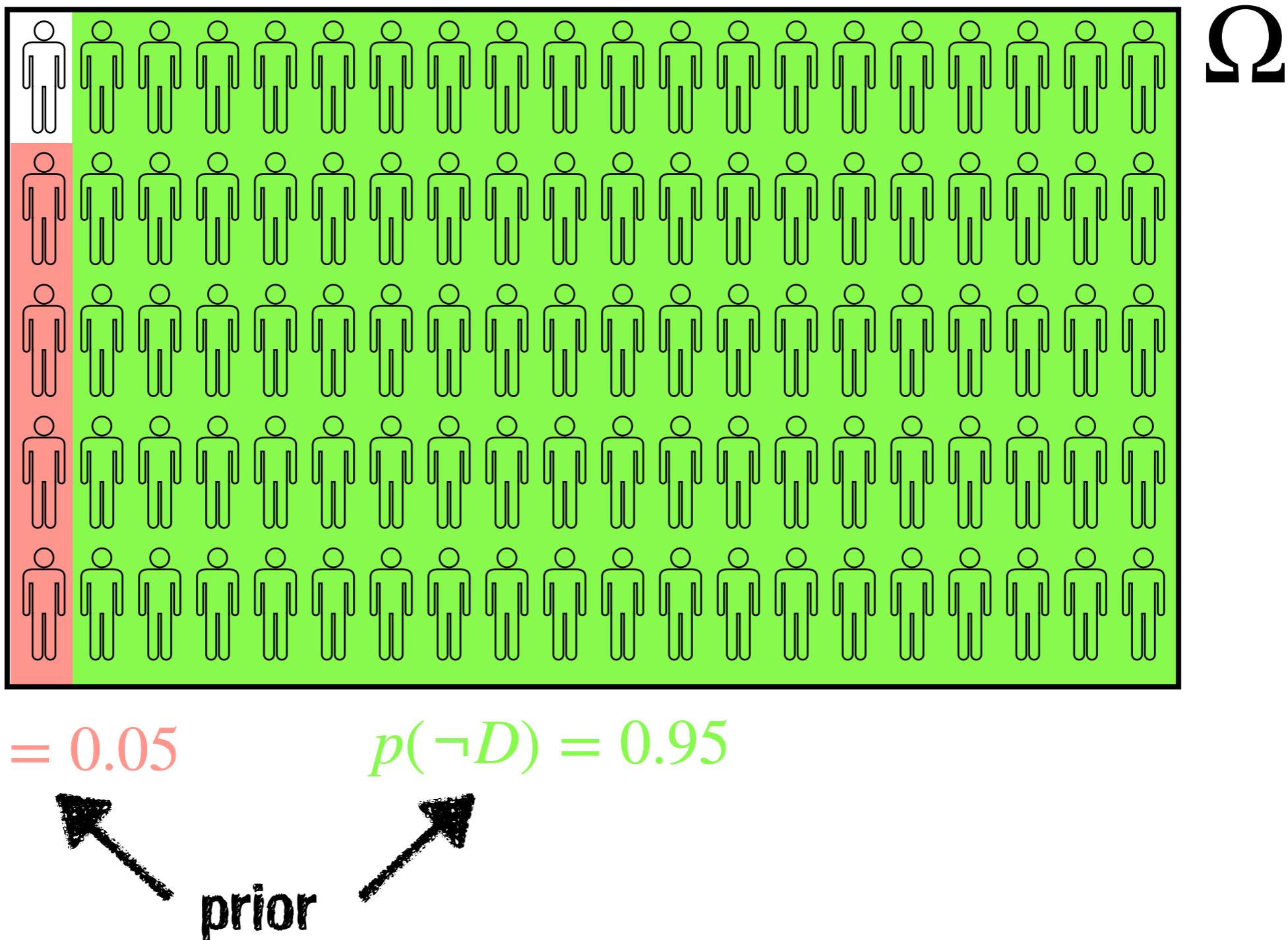
$$p(D) = 0.05$$

$$p(\neg D) = 0.95$$

prior

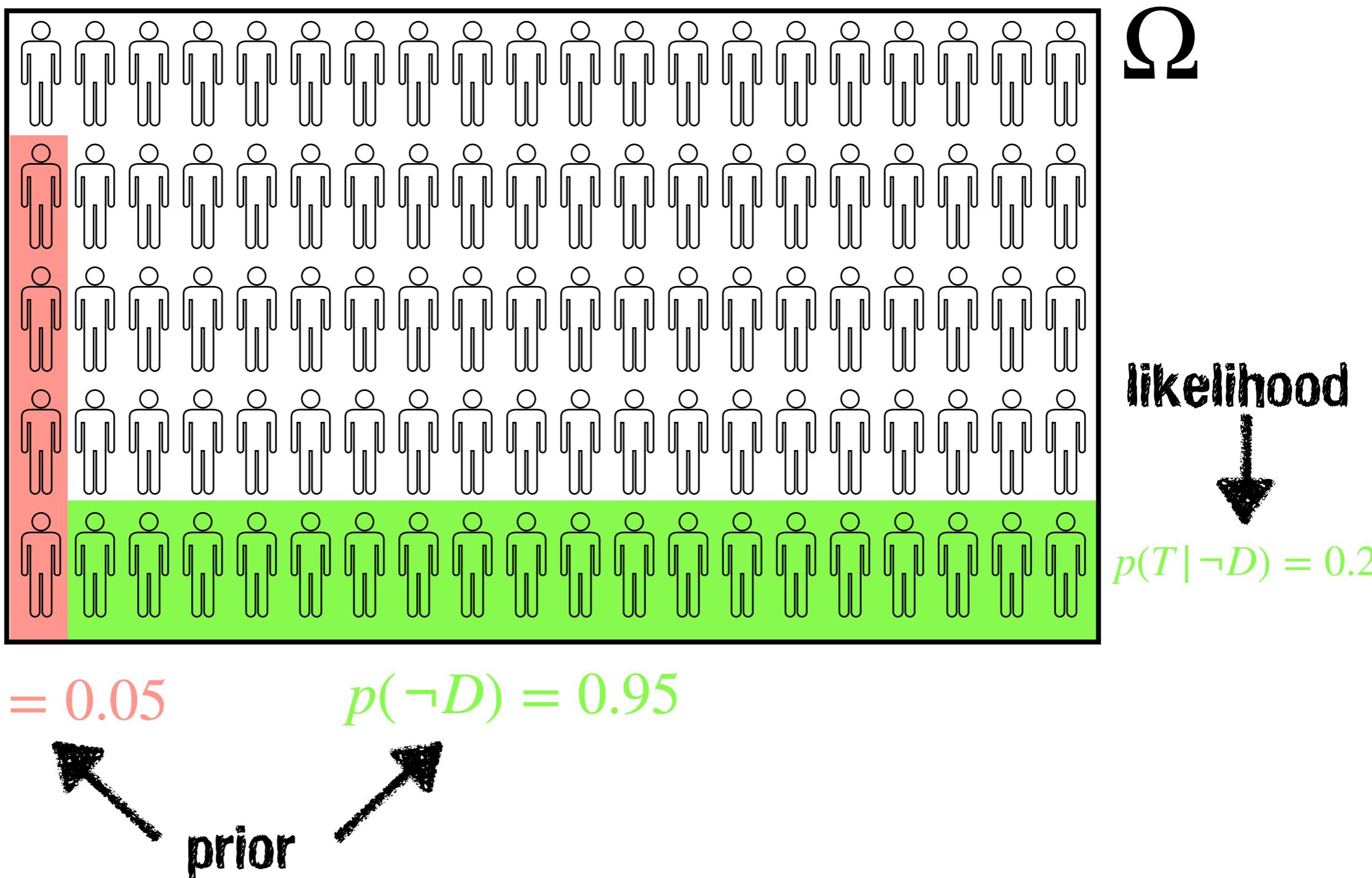
Clue guide to probability

$n = 100$



Clue guide to probability

$n = 100$

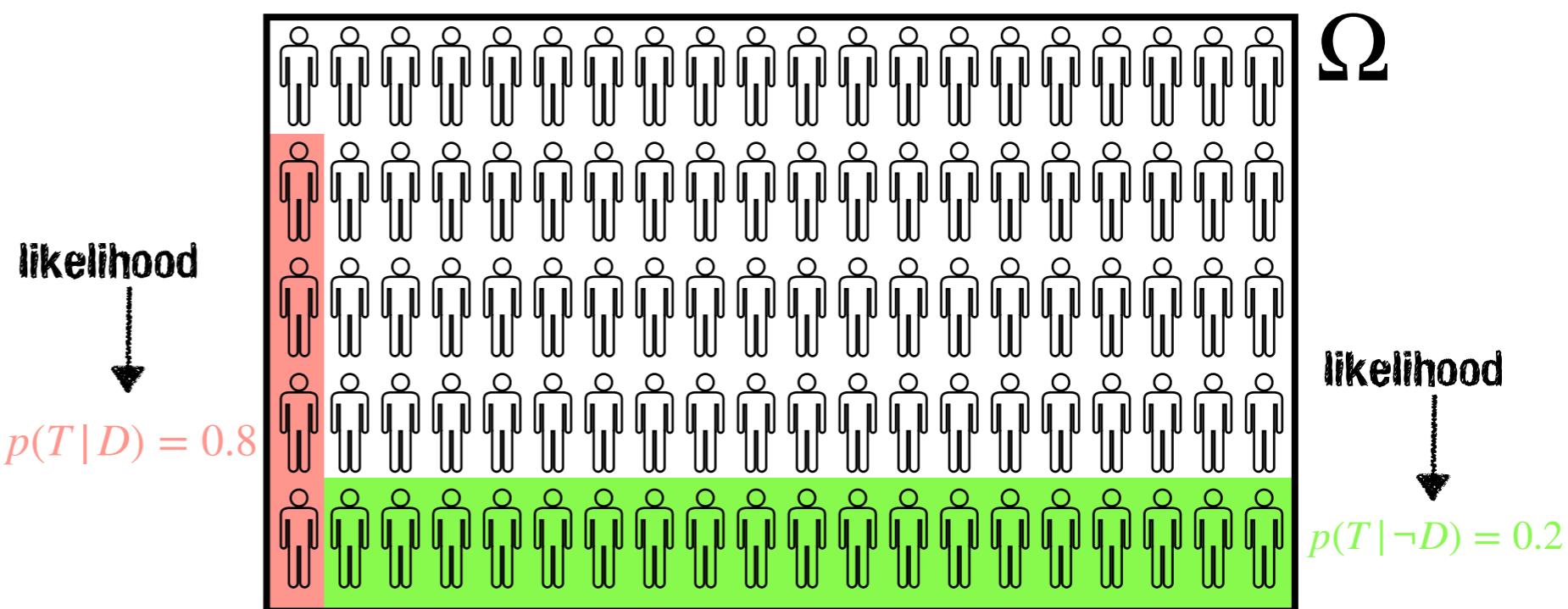


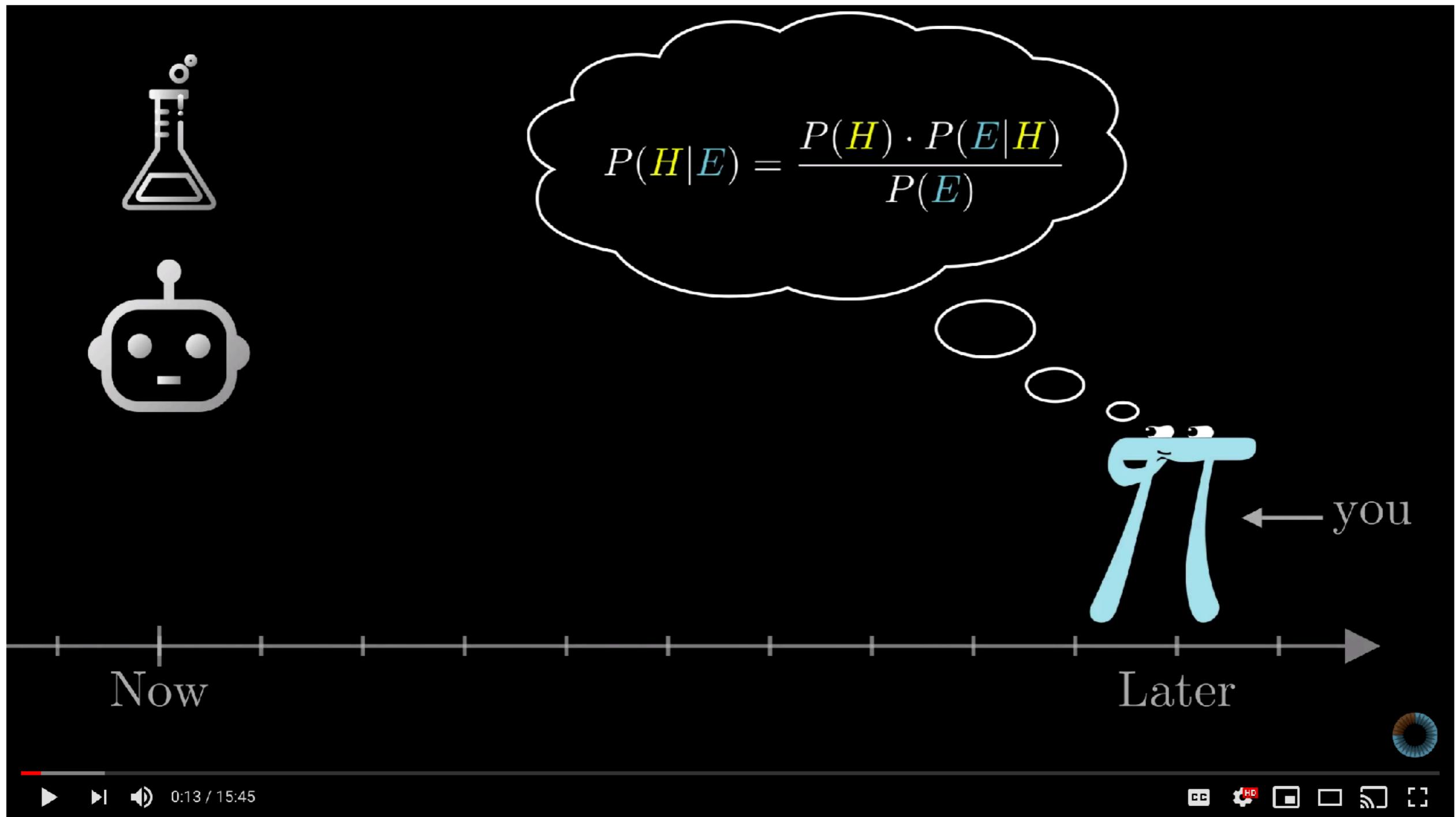
Clue guide to probability

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

$$p(D|T) = \frac{4}{4 + 19} = 0.174$$

$n = 100$





Bayes theorem, and making probability intuitive

461,105 views • Dec 22, 2019

26K

228

SHARE

SAVE

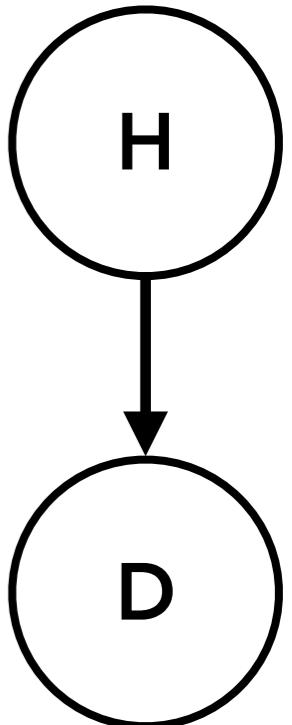
...

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

Bayes' theorem in three panels

posterior $p(H|D) = \frac{\text{likelihood} \quad \text{prior}}{p(D)} \quad \begin{matrix} \text{subjective probability} \\ \text{interpretation} \end{matrix}$

$H = \text{Hypothesis}$
 $D = \text{Data}$



Getting Bayes right matters!

NAS PNAS

Officer characteristics and racial disparities in fatal officer-involved shootings

David J. Johnson^{a,b,1}, Trevor Tress^b, Nicole Burkell^b, Carley Taylor^b, and Joseph Cesario^b

^aDepartment of Psychology, University of Maryland at College Park, College Park, MD 20742; and ^bDepartment 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

Check for updates

Psychological and Cognitive Sciences

Original claim:

Requires Bayes' rule

$$\Pr(\text{shot}|\text{minority civilian, white officer}, X) - \Pr(\text{shot}|\text{minority civilian, minority officer}, X) = \frac{\Pr(\text{min. civ.}|\text{shot, white off.}, X) \times \Pr(\text{shot}|\text{white off.}, X)}{\Pr(\text{minority civilian}|\text{white officer}, X)} - \frac{\Pr(\text{min. civ.}|\text{shot, min. off.}, X) \times \Pr(\text{shot}|\text{min. off.}, X)}{\Pr(\text{minority civilian}|\text{minority officer}, X)}.$$

[2]

Claim:

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

$$\Pr(\text{shot}|\text{minority civilian, white officer}, X)$$

$$- \Pr(\text{shot}|\text{minority civilian, minority officer}, X) \leq 0,$$

[1]

What the statistic says:

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

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

paper was retracted

Johnson, D. J., Tress, T., Burkell, 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!

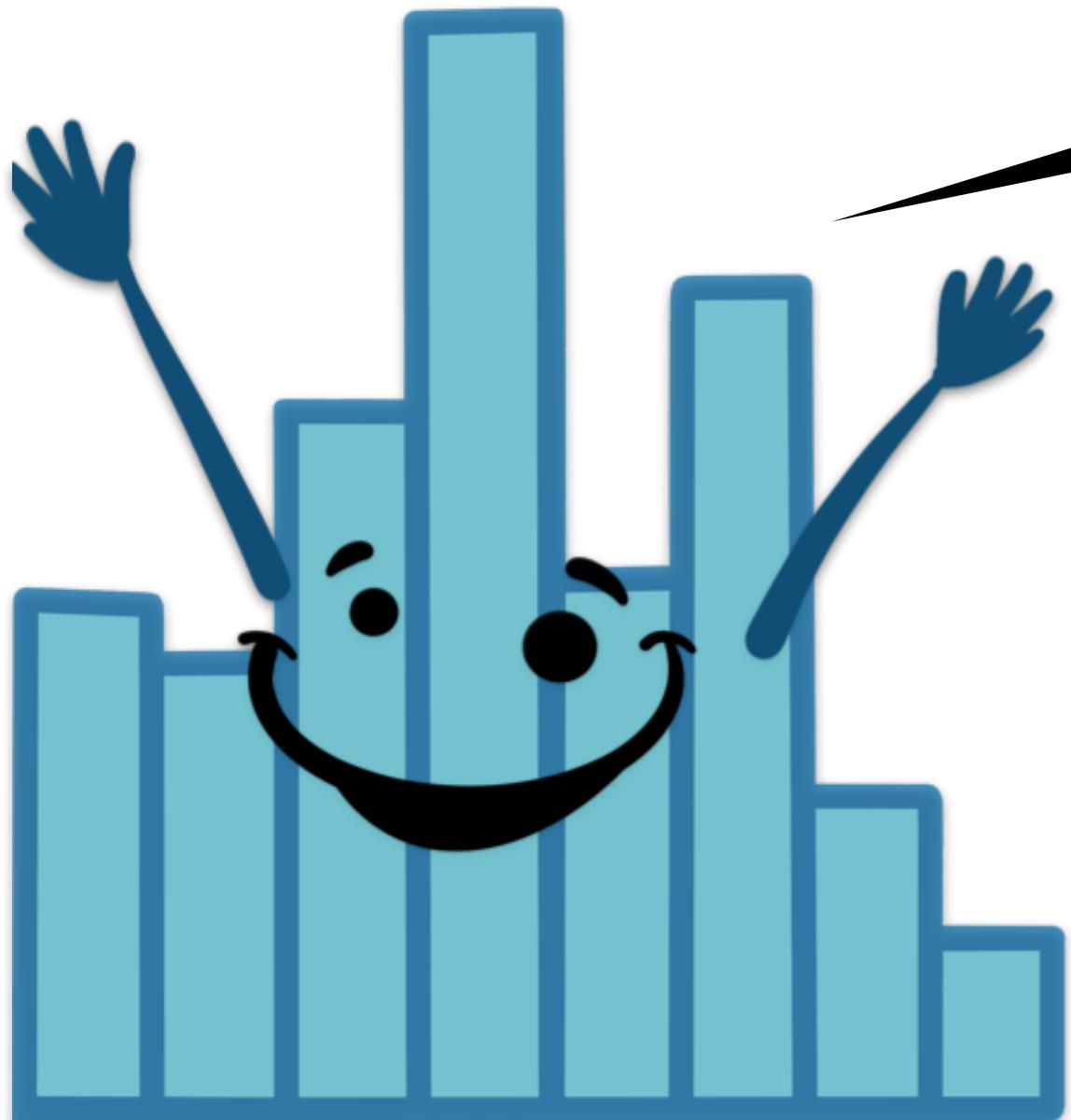
The screenshot shows the Zotero application interface. On the left is a library list with various items, including a highlighted entry: "Officer characteristics and racial disparities in f..." by Johnson et al., which has been retracted. On the right, a detailed view of this item is shown. A red banner at the top states "This work has been retracted." Below it, the retraction date is listed as "Retracted on 7/10/2020". Under "Concerns/Issues About Results", it says "Any question, controversy or dispute over the validity of the results". Under "Upgrade/Update of Prior Notice", it says "Either a change to or affirmation of a prior notice". A link to "Retraction Notice" is provided. At the bottom, there is a "More details:" section with a link to "https://retractionwatch.com/2020/07/08/retraction-of-paper-on-police-killings-and-race-not-due-to-mob-pressure-or-distaste-for-the-political-views-of-people-citing-the-work-approvingly-say-authors/". A note at the bottom right indicates "Data from Retraction Watch".

Title	Creator	Date Added
The R Inferno	Burns	... 9/29/2020, 1:4...
Officer characteristics and racial disparities in f...	Johnson et al.	... 6/15/2020, 4:5...
Reply to Knox and Mummolo and Schimmack and ...	Johnson and C...	... 6/15/2020, 4:5...
Young unarmed nonsuicidal male victims of fatal ...	Schimmack an...	... 6/15/2020, 4:5...
Making inferences about racial disparities in polic...	Knox and Mum...	... 6/15/2020, 4:5...
Quasi-experimental causality in neuroscience and ...	Marinescu et al.	... 5/29/2020, 10:...
Equivalence Testing for Psychological Research: A ...	Isager et al.	... 5/29/2020, 10:...
TreeBUGS: An R package for hierarchical multinom...	Heck et al.	... 5/27/2020, 1:4...
Analysis of variance with unbalanced data: an upd...	Hector et al.	... 4/15/2020, 1:4...
Bayesian analysis of factorial designs.	Rouder et al.	... 4/2/2020, 12:5...
Model comparison in ANOVA	Rouder et al.	... 4/2/2020, 12:4...
Bayesian Data Analysis	Gelman et al.	... 4/2/2020, 12:2...
d-separation		... 3/22/2020, 4:0...
To Bayes or Not To Bayes? That's no longer the qu...	Fokoue	... 3/22/2020, 4:0...
Science is not a signal detection problem	Wilson et al.	... 3/22/2020, 4:0...
Credible Confidence: A Pragmatic View on the Fre...	Albers et al.	... 3/22/2020, 4:0...
Effective Analysis of Reaction Time Data	Whelan	... 3/16/2020, 11:...
Are Reaction Time Transformations Really Benefici...	Schramm and 3/16/2020, 11:...
To transform or not to transform: using generaliz...	Lo and Andrews	... 3/16/2020, 11:...

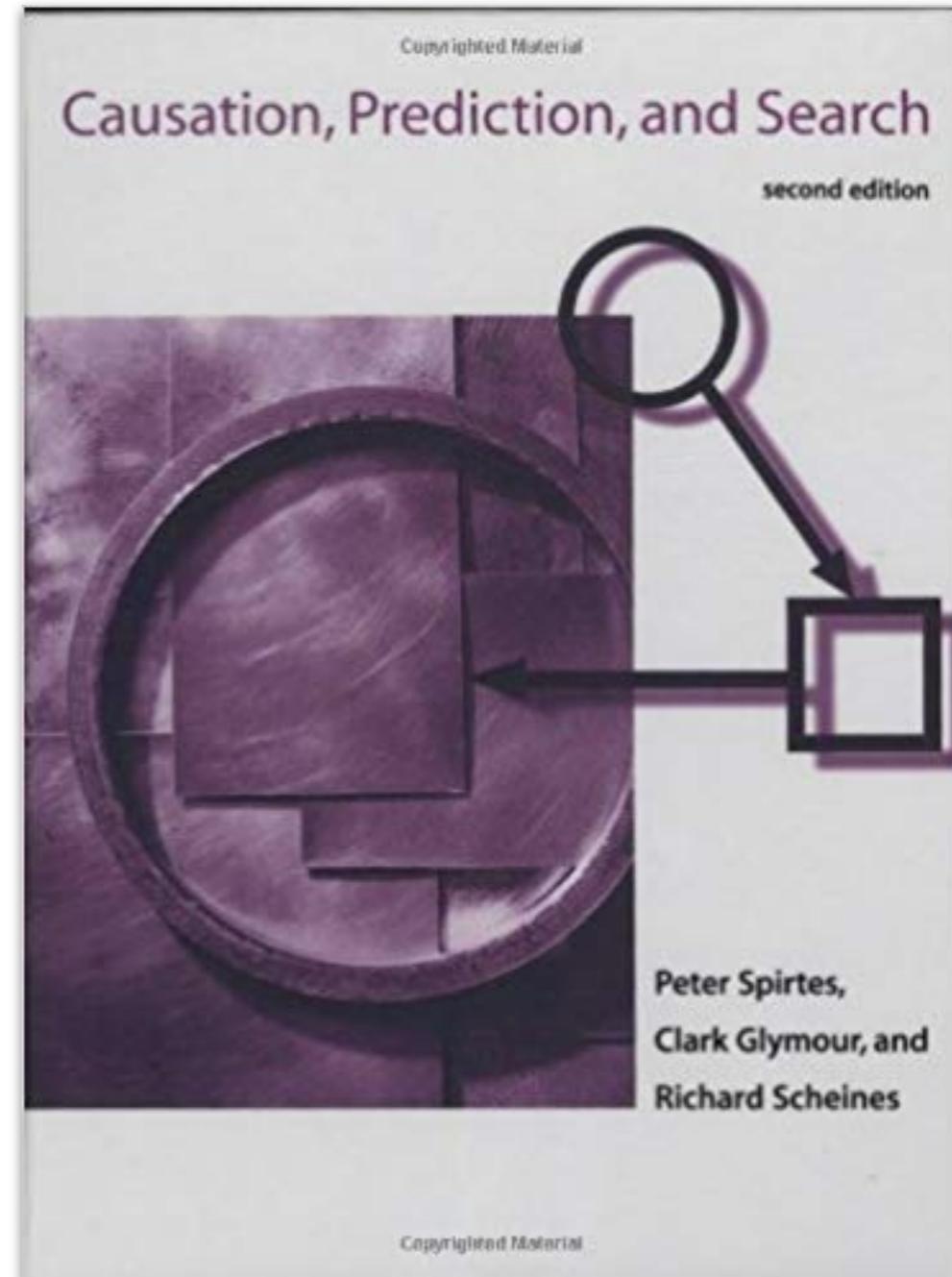
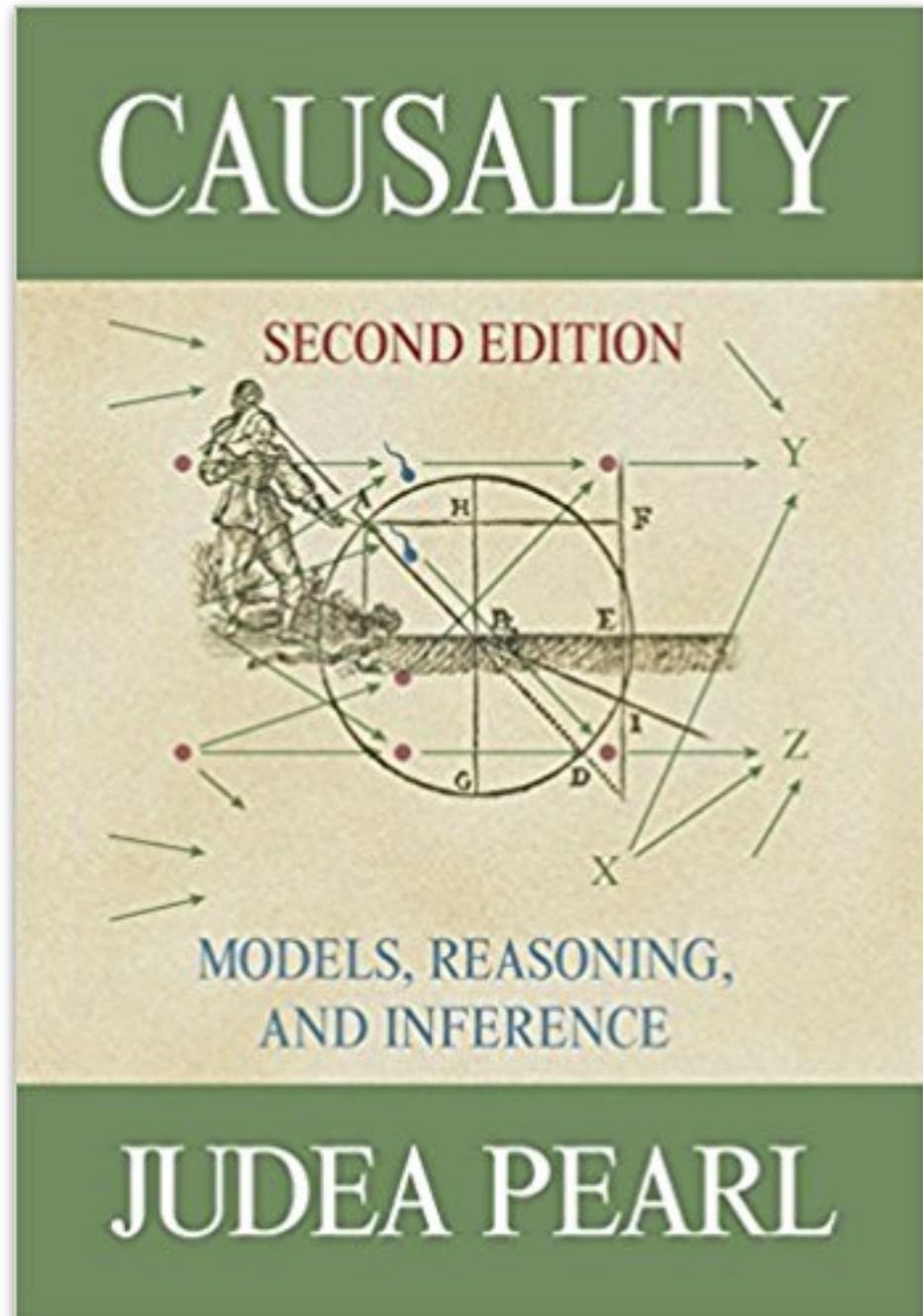
Tip: Use Zotero as a reference manager!

01:00

stretch break!



Causal Bayesian Networks

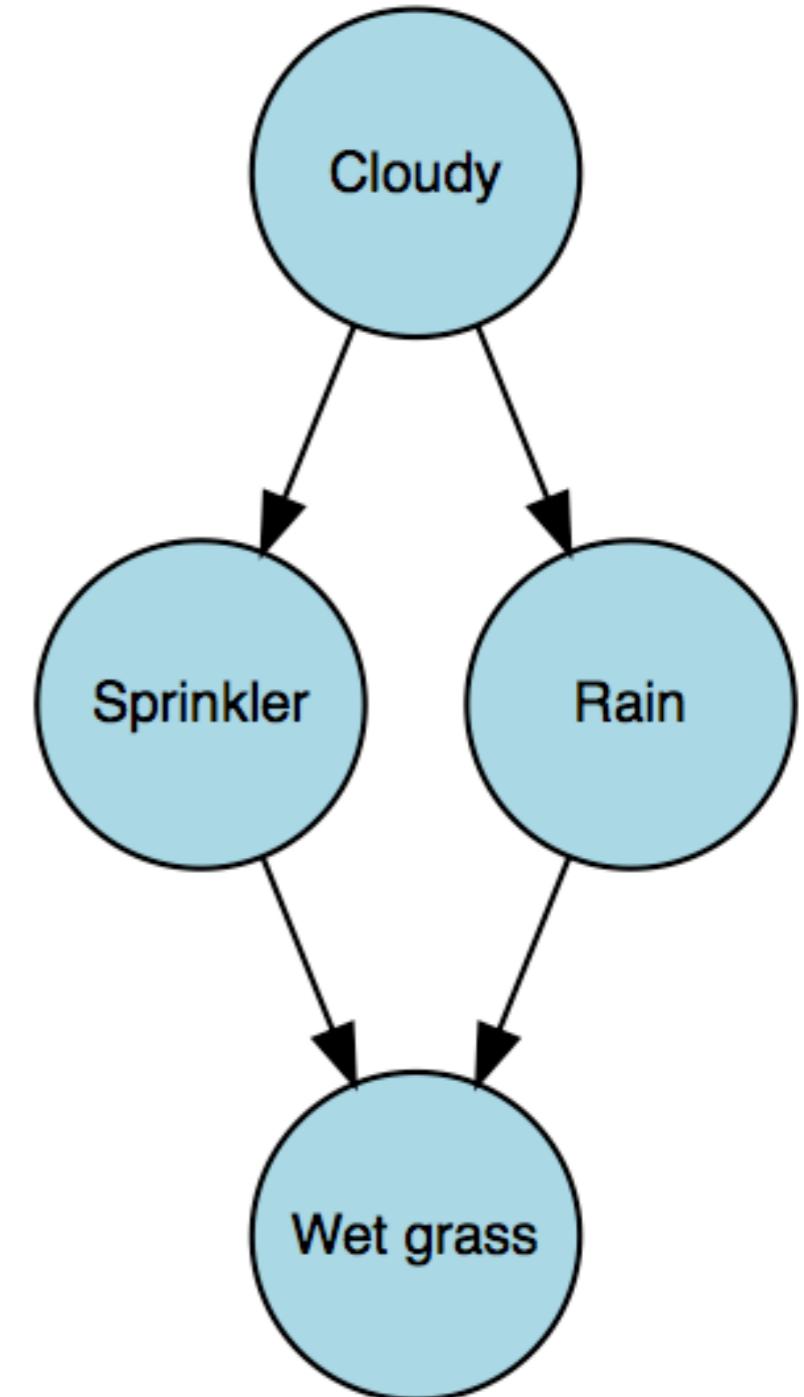


Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, England: Cambridge University Press.

Spirtes, P., Glymour, C. N., & Scheines, R. (2000). *Causation, prediction, and search*. The MIT Press. 60

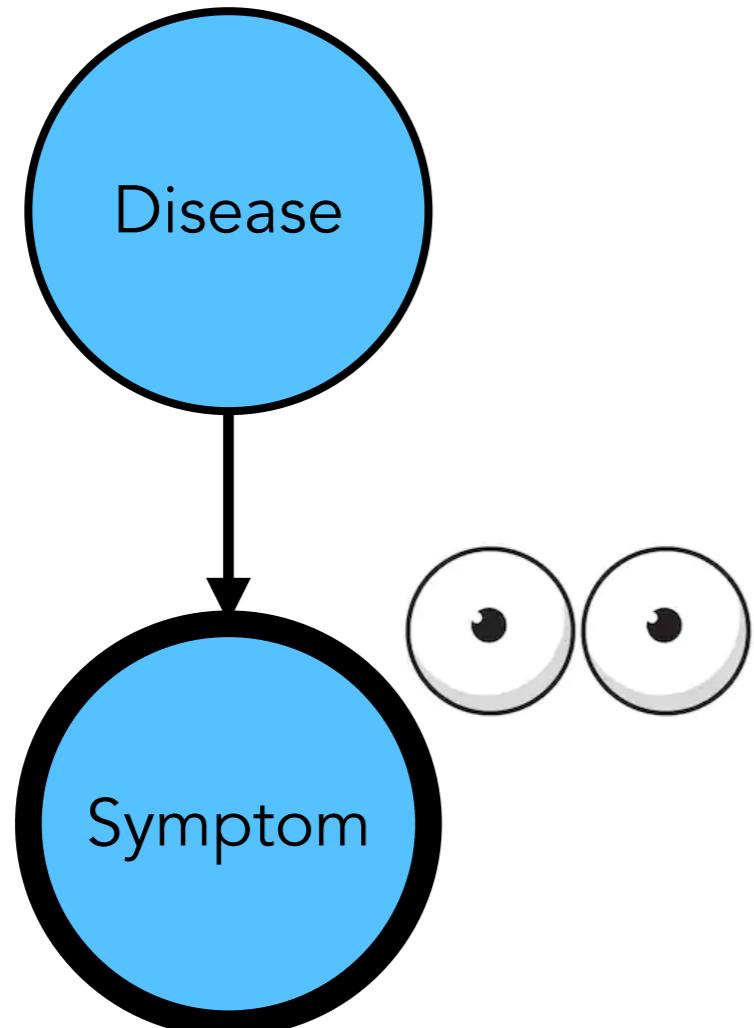
Representation

- **nodes** represent variables of interest
- **links** represent causal relationships between variables
- **conditional probability tables** parameterize the model



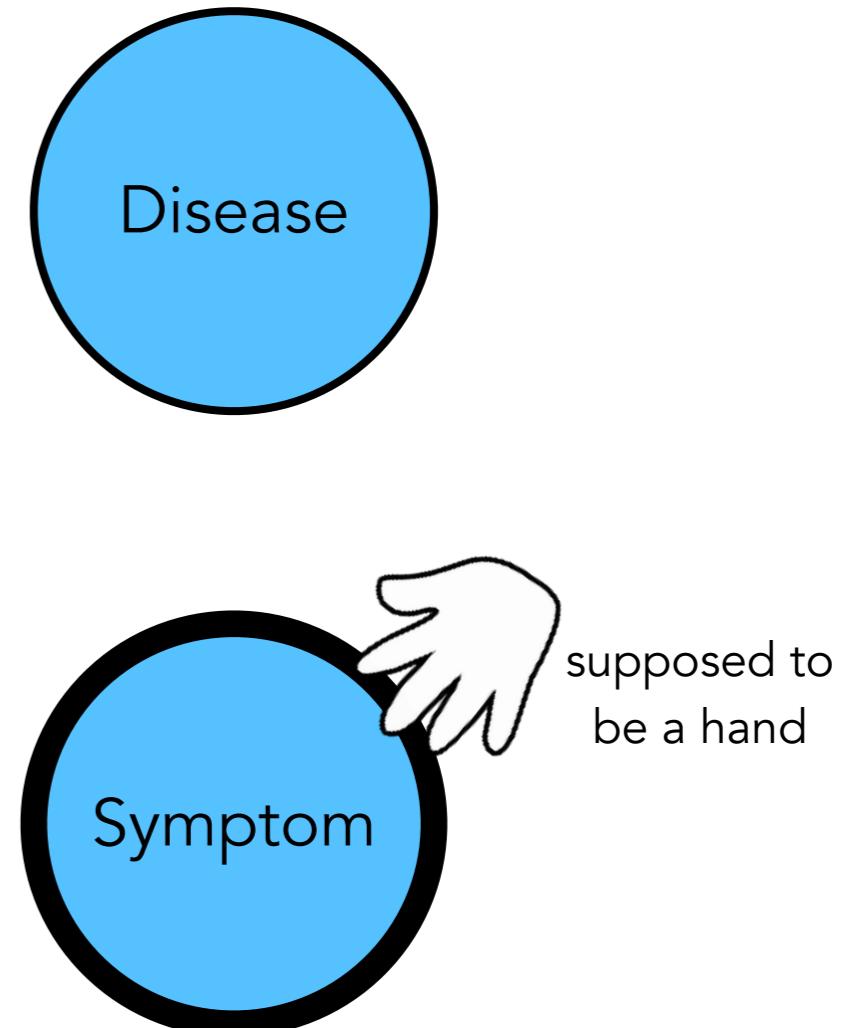
Observation vs. Intervention

seeing



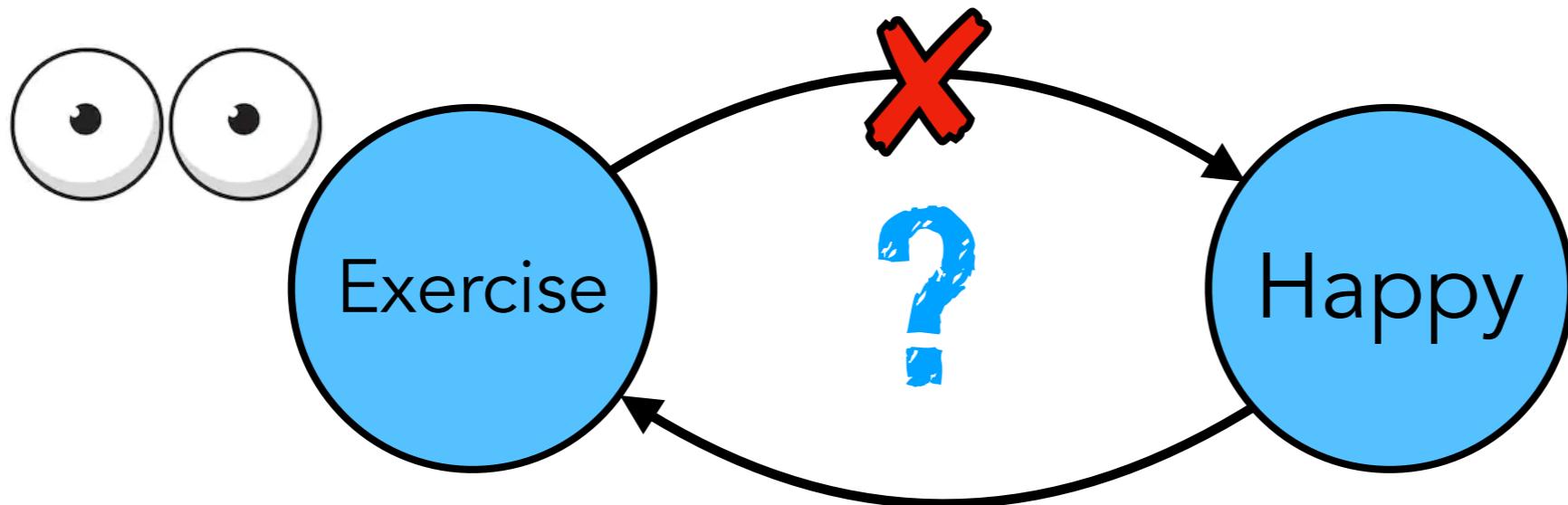
$$p(D | S) > p(D)$$

doing

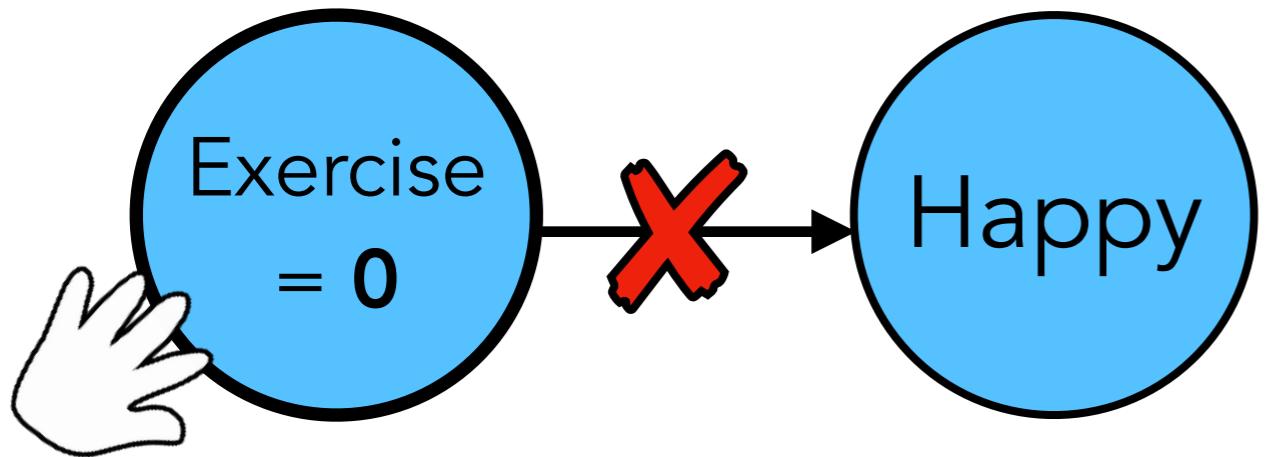


$$p(D | \text{do}(S)) = p(D)$$

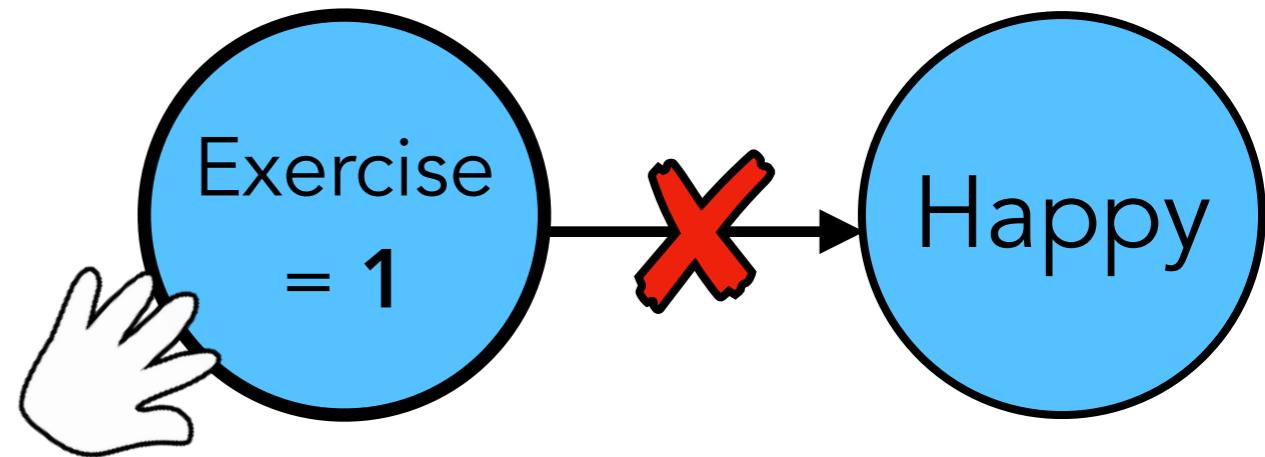
Inferring causal structure through intervention



Experiment 1

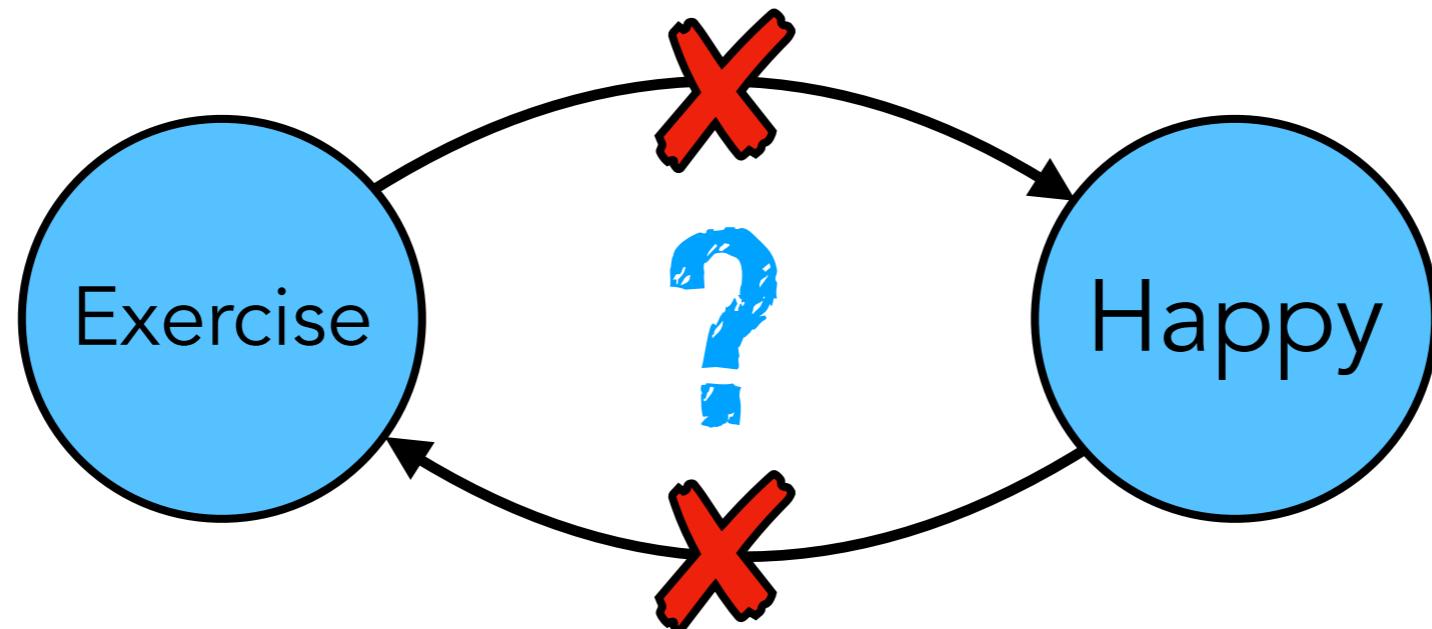


$$p(\text{Happy} \mid \text{do}(\text{Exercise} = 0)) = 0.3$$

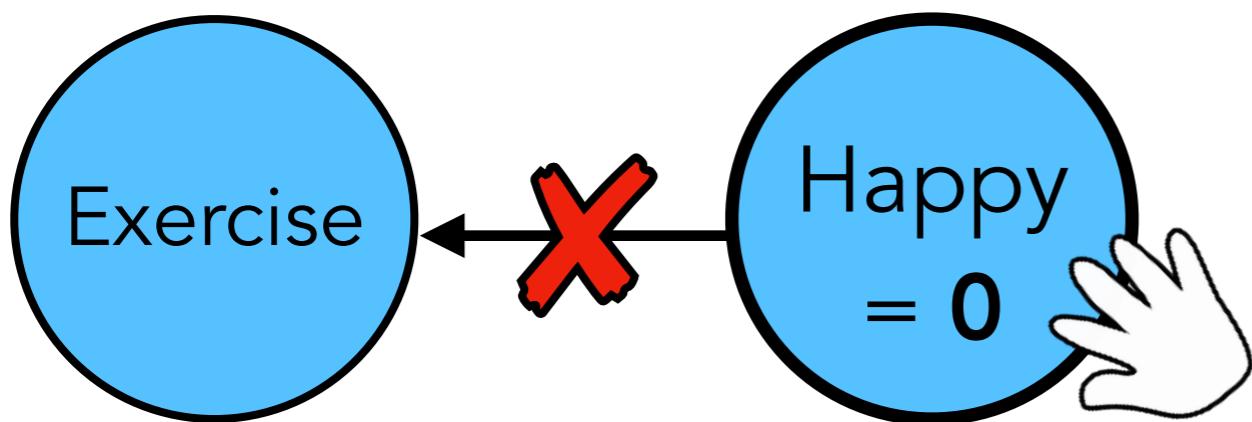


$$p(\text{Happy} \mid \text{do}(\text{Exercise} = 1)) = 0.3$$

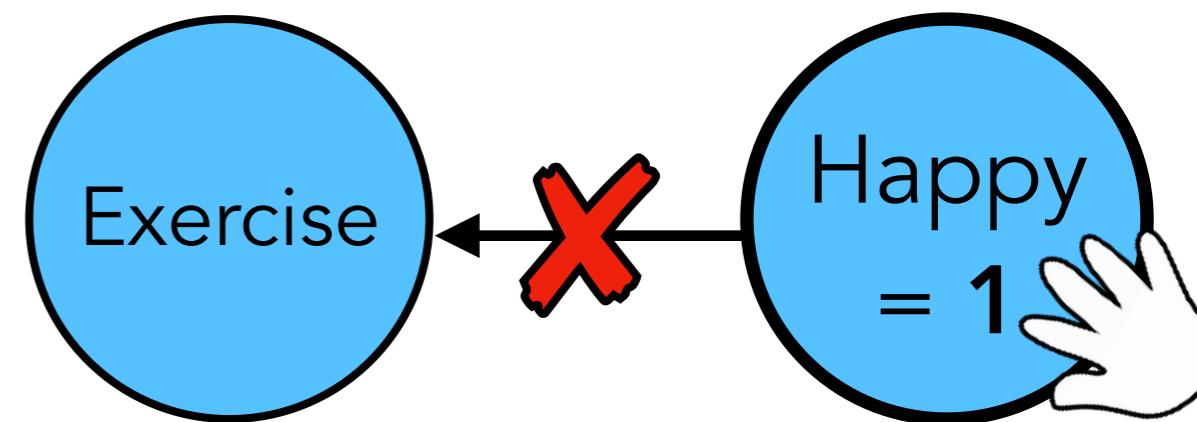
Inferring causal structure through intervention



Experiment 2

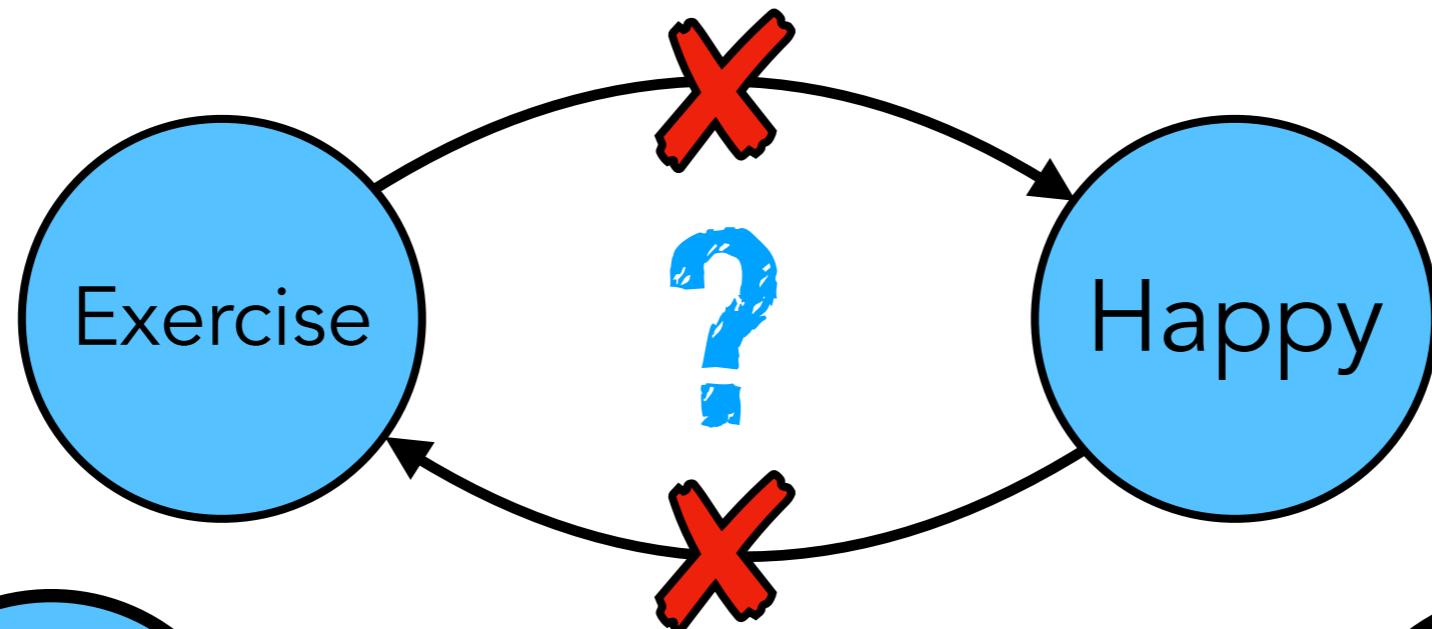


$$p(\text{Exercise} | \text{do}(\text{Happy} = 0)) = 0.1$$

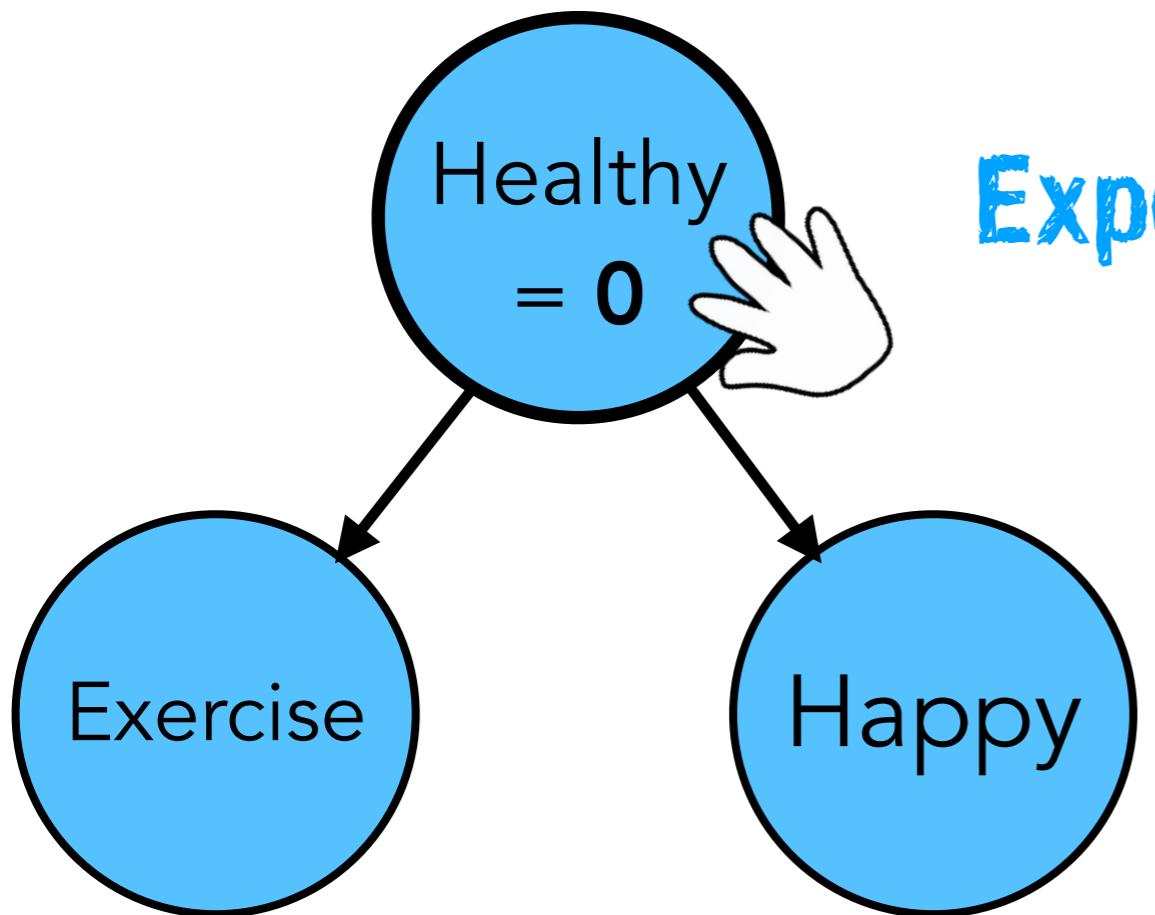


$$p(\text{Exercise} | \text{do}(\text{Happy} = 1)) = 0.1$$

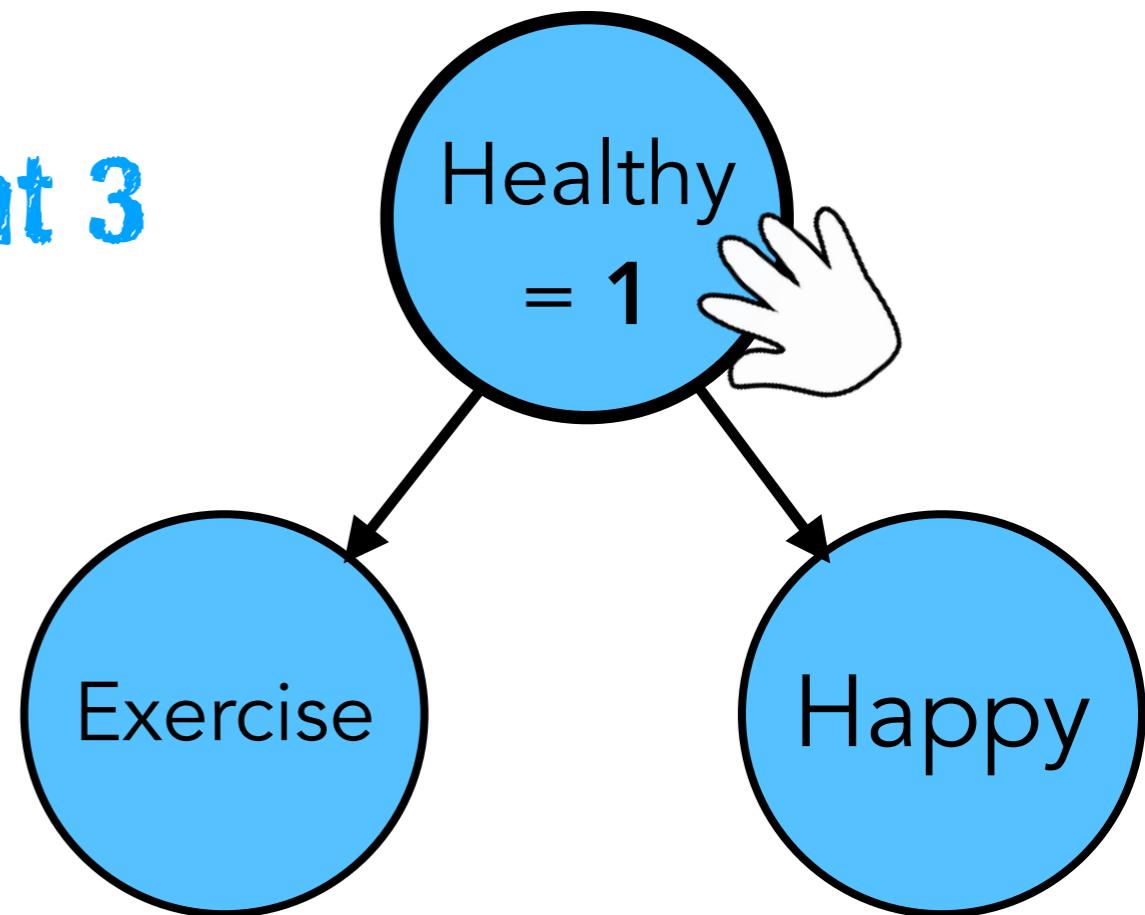
Inferring causal structure through intervention



Experiment 3



$$p(\text{Happy} | \text{do}(\text{Healthy} = 0)) = 0.05$$
$$p(\text{Exercise} | \text{do}(\text{Healthy} = 0)) = 0.1$$

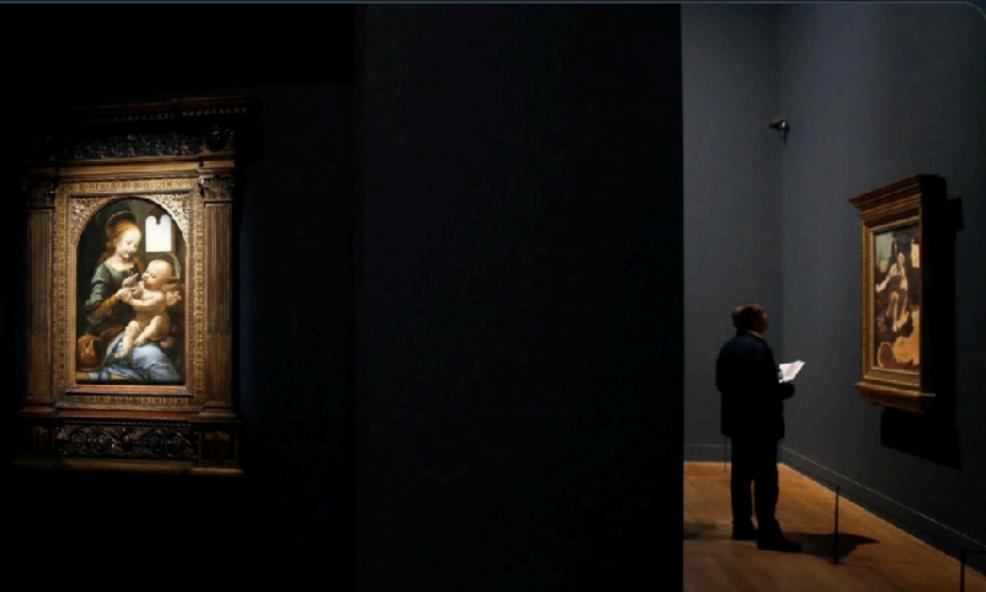


$$p(\text{Happy} | \text{do}(\text{Healthy} = 0)) = 0.5$$
$$p(\text{Exercise} | \text{do}(\text{Healthy} = 0)) = 0.75$$



NYT Health
@NYTHealth

Want to live longer? Try going to the opera. Researchers in Britain have found that people who reported going to a museum or concert even once a year lived longer than those who didn't.



Another Benefit to Going to Museums? You May Live Longer

Researchers in Britain found that people who go to museums, the theater and the opera were less likely to die in the study period than those who didn't.

[nytimes.com](#)

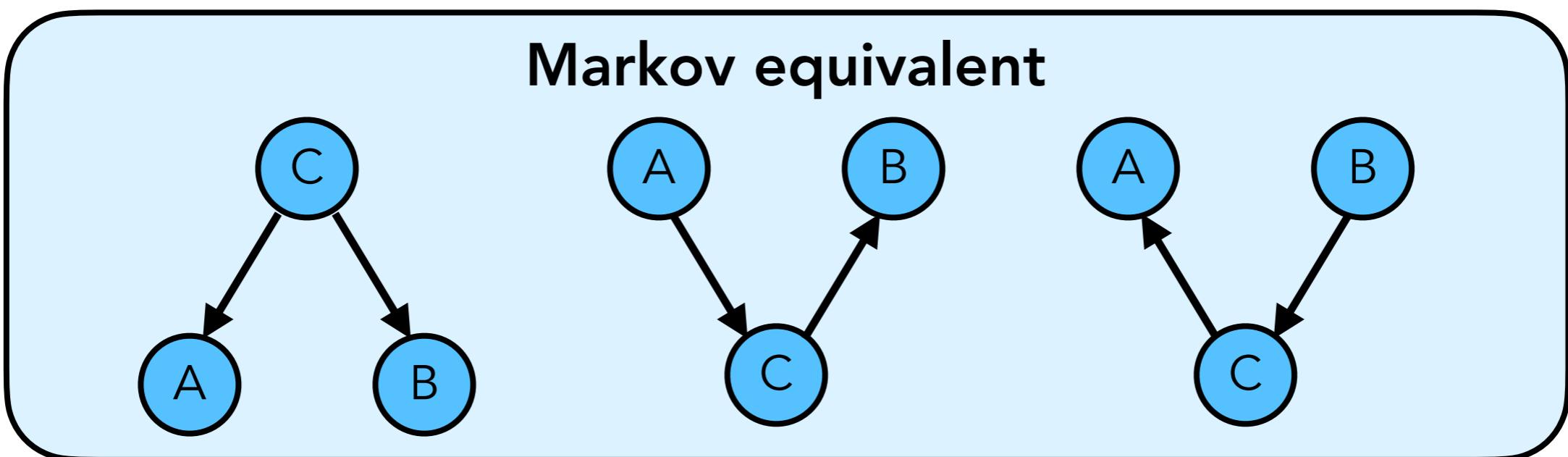
9:19 AM · Dec 22, 2019 · SocialFlow

336 Retweets 1.3K Likes



Important take home message

- correlation is not causation
- correlation (= probabilistic dependence)
suggests that there is some causal relationship
- but we don't know which one it is



- **causal interventions** / experiments can reveal the underlying causal structure

Outline

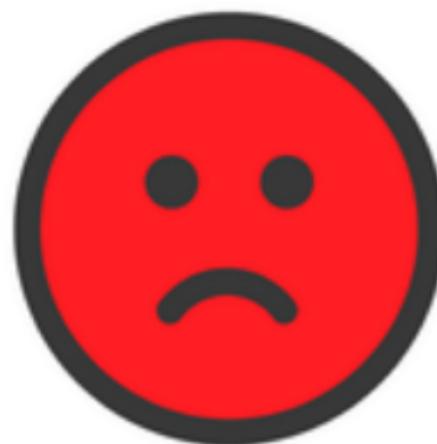
- Introduction to probability / Recap
 - Motivation
 - Counting possibilities
 - **Clue** guide to probability
- Causal Bayesian Networks
 - seeing (prediction) vs. doing (explanation)
 - correlation is not causation

Feedback

How was the pace of today's class?

much a little just a little much
too too right too too
slow slow fast fast

How happy were you with today's class overall?



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

Thank you!