CS 237: PROBABILITY IN COMPUTING

INSTRUCTORS: TIAGO JANUARIO, SOFYA RASKHODNIKOVA

BOSTON UNIVERSITY

"I wish we hadn't learned probability 'cause I don't think our odds are good."

LECTURE 2

Last time

- Course information
- Introduction to probability theory
- Sample Spaces and Events
- Examples

Today

- Probability function
- Symmetry
- Probability axioms

GET READY FOR VOTING!

- Students from A1 section:
    1. Enter this URL: **https://app.tophat.com/e/**037447
    2. Login as needed using your Top Hat account info.
    3. If asked about enrolling, click *Enroll*.
- Students from A2 section:
    1. Enter this URL: **https://app.tophat.com/e/**606363
    2. Login as needed using your Top Hat account info.
    3. If asked about enrolling, click *Enroll*.

GAMES OF CHANCE

- Games of chance popular throughout recorded history

Dogs and Jackals

Egyptian board game (c. 3500 BC)

Astragalus

Egyptian 4-sided die

GAMES OF CHANCE

Gambling house in France (during renaissance)

GAMES OF CHANCE

- France (1650s): gambling very popular (and unregulated)
- Games grew more and more complex
- Need for mathematical methods to analyze chance of winning

Blaise Pascal

[1623-1662]

Pierre de Fermat

[1607-1665]

THE SCHOOL OF PROBABILITY

- Introduce three self-evident and indisputable properties of probability (the axioms)
- Develop the mathematical theory of probability from these axioms

Andrey Kolmogorov

[1903 - 1987]

> "The theory of probability as a mathematical discipline can and should be developed from axioms in exactly the same way as geometry and algebra."

PROBABILITY: CAST OF CHARACTERS

- Experiment: a repeatable procedure
  - Toss a coin
  - Toss a coin 3 times
  - Roll two dice
  - Shuffle a deck of cards
  - Draw a sock from a drawer containing red and black socks; if the sock is red, draw another sock

PROBABILITY: CAST OF CHARACTERS

lower-case Ω

- Outcome: result of the experiment ω
- Sample space Ω: set of all possible outcomes
  - Toss a coin

tails

$\Omega = \{H, T\}$

heads

PROBABILITY: CAST OF CHARACTERS

- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
  - Roll a 4-sided die twice

$$\Omega = \{(i, j) \mid 1 \le i,j \le 4 \text{ and } i,j \in \mathbb{N}\}$$

i → 1st roll

j → 2nd roll

2nd roll \ 1st roll	1	2	3	4
1	(1,1)	(2,1)	(3,1)	(4,1)
2	(1,2)	(2,2)	(3,2)	(4,2)
3	(1,3)	(2,3)	(3,3)	(4,3)
4	(1,4)	(2,4)	(3,4)	(4,4)

PROBABILITY: CAST OF CHARACTERS

- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
  - Toss a coin until we get a heads

$\Omega = \{H, TH, TTH, \ldots\}$

The experiment stops when you get heads

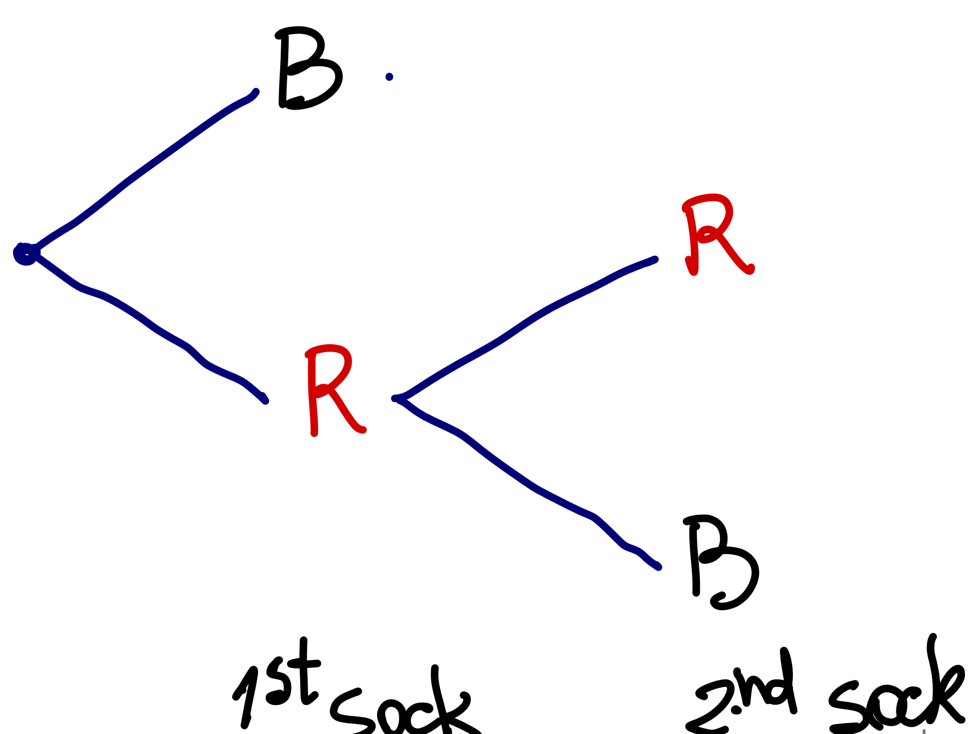
PROBABILITY: CAST OF CHARACTERS

- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
  - Shuffle a deck of cards and display all cards

$\Omega = 52!$

PROBABILITY: CAST OF CHARACTERS

- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
  - Draw a sock from a drawer containing red and black socks; if the sock is red, draw another sock



outcome

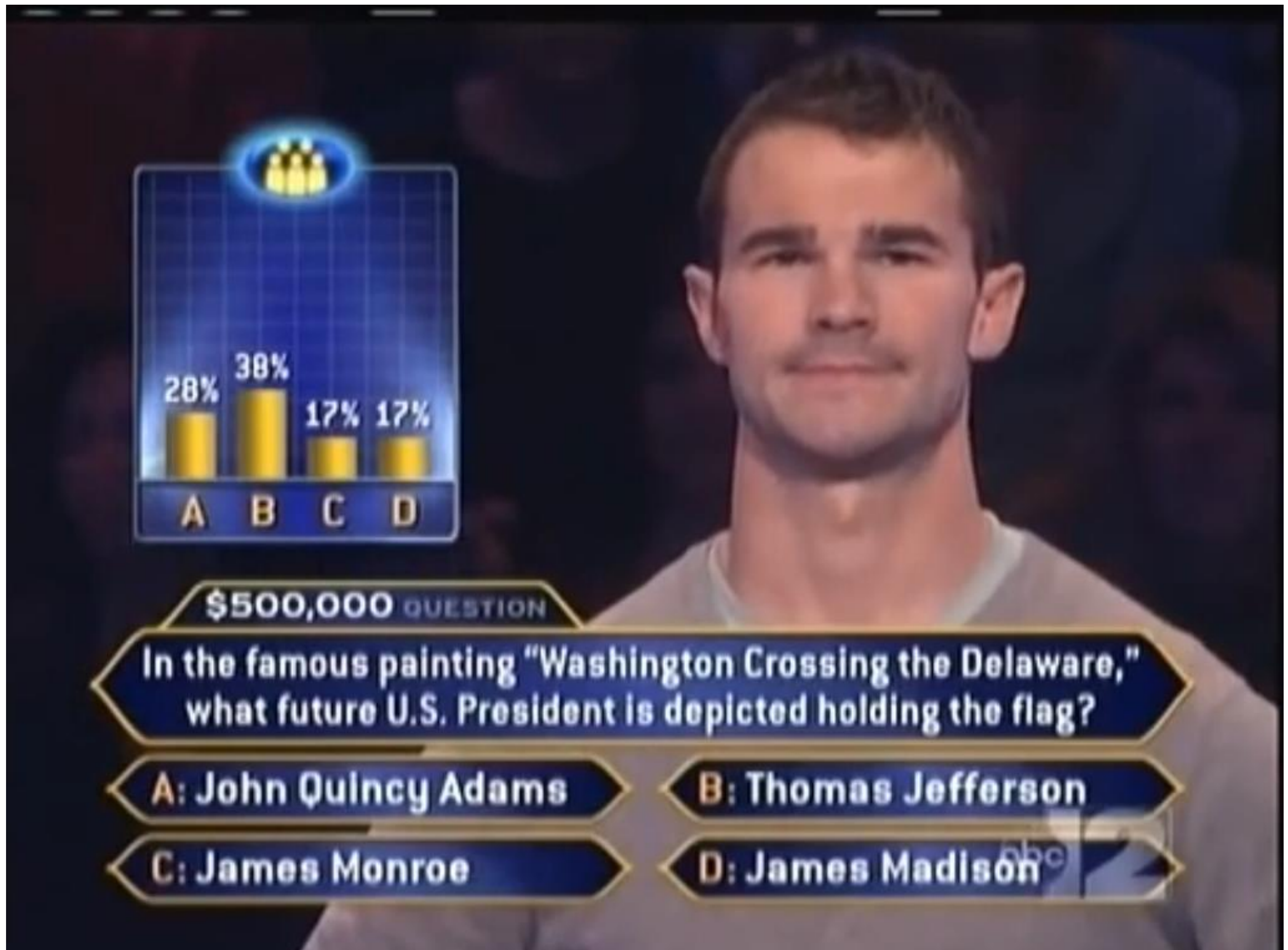
B

RR

RB

$\Omega = \{B, RR, RB\}$

TOP HAT QUESTION



TOP HAT QUESTION

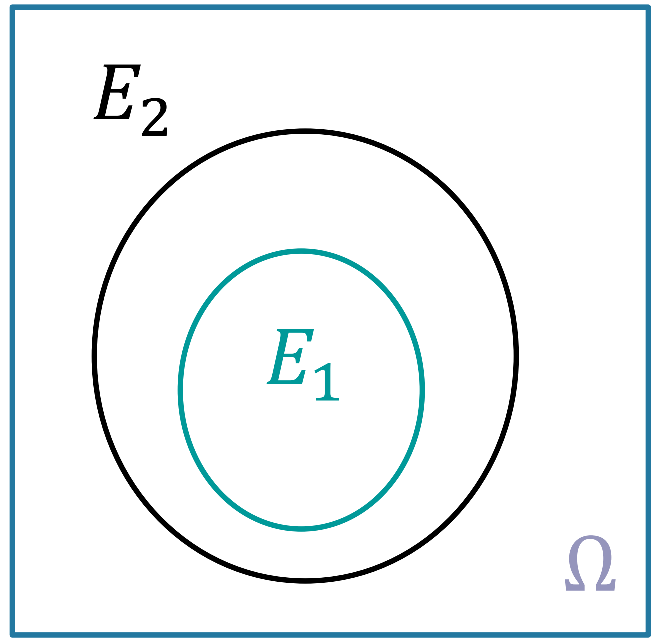
- Experiment: toss a coin 3 times

Does the first event imply the second?

$E_1 \Rightarrow E_2$

$E_1 =$ "at least 2 heads"

$E_2 =$ "exactly 2 heads"



A. YES

B. NO

PROBABILITY FUNCTION

- Each outcome in the sample space Ω is assigned a probability, which is a number greater or equal to 0.
- All probabilities of outcomes in Ω must add up to 1.

$\frac{1}{4}$

$\frac{1}{4}$

$\frac{1}{4}$

$\frac{1}{4}$

- Probability of event E, denoted Pr(E), is the sum of probabilities of all outcomes in E.

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H, T\}$
$Pr(E)$				

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H,T\}$
$Pr(E)$				

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

Symmetry:

Each outcome is equally likely

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H,T\}$
$Pr(E)$		$\frac{1}{2}$	$\frac{1}{2}$	

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H, T\}$
$Pr(E)$		$\frac{1}{2}$	$\frac{1}{2}$	

IN THIS COURSE, COINS LAND EITHER ON HEADS OR TAILS

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H, T\}$
$Pr(E)$	0	$\frac{1}{2}$	$\frac{1}{2}$	

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H, T\}$
$Pr(E)$	0	$\frac{1}{2}$	$\frac{1}{2}$	

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H, T\}$
$Pr(E)$	0	$\frac{1}{2}$	$\frac{1}{2}$	1

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin

$$\Omega = \{H, T\}$$

E	$\emptyset$	$\{H\}$	$\{T\}$	$\{H, T\}$
$Pr(E)$	0	$\frac{1}{2}$	$\frac{1}{2}$	1

Compact description: $Pr(E) = \frac{|E|}{|\Omega|}$ for all events $E \subseteq \Omega$

RECAP: PROBABILITY CAST OF CHARACTERS

- Experiment: a repeatable procedure
- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
- Event: a subset of the sample space
- Probability function $\Pr$: assigns a probability Pr(E) to each event E

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: roll a die

$$\Omega = \{1,2,3,4,5,6\}$$

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: roll a die

$$\Omega = \{1,2,3,4,5,6\}$$

Symmetry:

Each outcome is equally likely

$$Pr(\{1\}) = Pr(\{2\}) = \cdots = Pr(\{6\}) = \frac{1}{6}$$

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: roll a die

Symmetry:

Each outcome is equally likely

$$\Omega = \{1,2,3,4,5,6\}$$

$$Pr(\{1\}) = Pr(\{2\}) = \cdots = Pr(\{6\}) = \frac{1}{6}$$

For larger events $E,\ Pr(E)$ is the sum of the probabilities of the outcomes in E

$$Pr(\text{"even"}) = Pr(\{2,4,6\}) = Pr(\{2\}) + Pr(\{4\}) + Pr(\{6\}) = \frac{3}{6}$$

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: roll a die

Symmetry:

Each outcome is equally likely

$$\Omega = \{1,2,3,4,5,6\}$$

$$Pr(\{1\}) = Pr(\{2\}) = \cdots = Pr(\{6\}) = \frac{1}{6}$$

For larger events E, $Pr(E)$ is the sum of the probabilities of the outcomes in E

$$Pr(E) = \frac{|E|}{|\Omega|} \text{ for all events } E \subseteq \Omega$$

PROBABILITY FUNCTION

Symmetry:
Each outcome is equally likely

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin 3 times

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$$Pr(\{HHH\}) = Pr(\{HHT\}) = \cdots = Pr(\{TTT\}) = \frac{1}{8}$$

PROBABILITY FUNCTION

Symmetry:
Each outcome is equally likely

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin 3 times

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$$Pr(\{HHH\}) = Pr(\{HHT\}) = \cdots = Pr(\{TTT\}) = \frac{1}{8}$$

For larger events E, $Pr(E)$ is the sum of the probabilities of the outcomes in E

$$Pr(E) = \frac{|E|}{|\Omega|} \text{ for all events } E \subseteq \Omega$$

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin 3 times

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

- The coin tosses are independent (the outcome of a toss does not change the probability that a different toss is H or T)

$$Pr(\{HHT\}) = Pr(\text{"}1^{st}\ toss\ is\ H\text{"} \cap \text{"}2^{nd}\ toss\ is\ H\text{"} \cap \text{"}3^{rd}\ toss\ is\ T\text{"})$$

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin 3 times

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

- The coin tosses are independent (the outcome of a toss does not change the probability that a different toss is H or T)

$$Pr(\{HHT\}) = Pr(\text{"}1^{st}\ toss\ is\ H\text{"} \cap \text{"}2^{nd}\ toss\ is\ H\text{"} \cap \text{"}3^{rd}\ toss\ is\ T\text{"})$$

$$= Pr(\text{"}1^{st}\text{ toss is H"}) \cdot \Pr(\text{"}2^{nd}\text{ toss is H"}) \cdot \Pr(\text{"}3^{rd}\text{ toss is T"})$$

product rule for independent events (will justify it later)

PROBABILITY FUNCTION

- Assigns a probability Pr(E) to each event E
- Experiment: toss a coin 3 times

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

- The coin tosses are independent (the outcome of a toss does not change the probability that a different toss is H or T)

$$Pr(\{HHT\}) = Pr(\text{"}1^{\text{st}}\text{ toss is H"} \cap \text{"}2^{\text{nd}}\text{ toss is H"} \cap \text{"}3^{\text{rd}}\text{ toss is T"})$$

$$= \Pr(\text{"}1^{\text{st}}\text{ toss is H"}) \cdot \Pr(\text{"}2^{\text{nd}}\text{ toss is H"}) \cdot \Pr(\text{"}3^{\text{rd}}\text{ toss is T"})$$

$$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

PROBABILITY FUNCTION

Not symmetric:
Outcomes are not equally likely

- Assigns a probability Pr(S) to each event E
- Experiment: toss a coin until heads for the first time

$$\Omega = \{\underbrace{TT \dots T}_{i \text{ times}} H : i \in \mathbb{N}\}$$

The coin tosses are independent (the outcome of a toss does not change the probability that a different toss is H or T)

$$\Pr(\{\underbrace{TTT \dots T}_{i \text{ tails}} H\}) = \left(\frac{1}{2}\right)^i \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{i+1}$$

PROBABILITY FUNCTION

Not symmetric:
Outcomes are not equally likely

- Assigns a probability Pr(S) to each event E
- Experiment: toss a coin until heads for the first time

$$\Pr(\{\underbrace{TTT \dots T}_{i \text{ tails}} H\}) = \left(\frac{1}{2}\right)^{i} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{i+1}$$

- For larger events E, Pr(E) is the sum of the probabilities of the outcomes in E

$$\Pr(E) = \sum_{\omega \in S} \Pr(\{\omega\}) \text{ for all } E \subseteq \Omega$$

PROBABILITY FUNCTION

What principles did we use to come up with those probability functions?

- Symmetry: each outcome of the coin toss (or die roll) is equally likely
- The probability of each outcome is a number between 0 and 1
- Additivity: for events with more than one outcome, the probability of the event is the sum of the probabilities of its outcomes

Note: some experiments are not symmetric (toss a coin until H)

PROBABILITY FUNCTION: AXIOMS

- Probability function assigns a probability to each event
- For a function to be a probability function, it must satisfy the following properties, called axioms of probability
- Non-negativity: $Pr(E) \geq 0$ for all events $E \subseteq \Omega$
- Additivity: if A and B are disjoint events then

$$Pr(A \cup B) = Pr(A) + Pr(B)$$

- Normalization: $Pr(\Omega) = 1$

PROBABILITY FUNCTION: AXIOMS

- Additivity (generalized): If $P(A \cup B) = P(A) + P(B)$ whenever A, B are disjoints sets, then for $A_1 \ldots A_n$ disjoints sets

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1 \cup B)$$

- where $B = A_2 \cup \cdots \cup A_n$ is disjoint from A.
- From the previous slide

$$P(A_1 \cup B) = P(A_1) + P(B)$$

- If $A_1, A_2, \ldots$ are disjoint events then

$$Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = Pr(A_1) + Pr(A_2) + \cdots + \Pr(A_n)$$

TOP HAT QUESTION

- 50 students showed up to class today.
- 20 are Red Sox fans (R), 25 are Patriots fans (P).
- Prof. Tiago chooses a student at random.
- What is the range of possible values for $p = Pr(R \cup P)$?

A. $p \geq 0.5$

B. $p \leq 0.4$

C. $0.4 \leq p \leq 0.9$

D. $0.5 \leq p \leq 0.9$

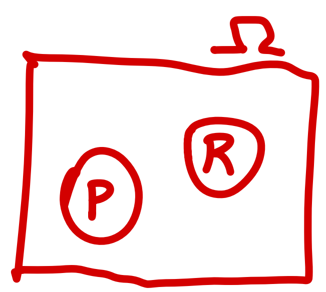
E. $0.4 \leq p \leq 0.5$

maximize $Pr(R \cup P)$

$Pr(R \cup P) = Pr(R) \cup Pr(P)$

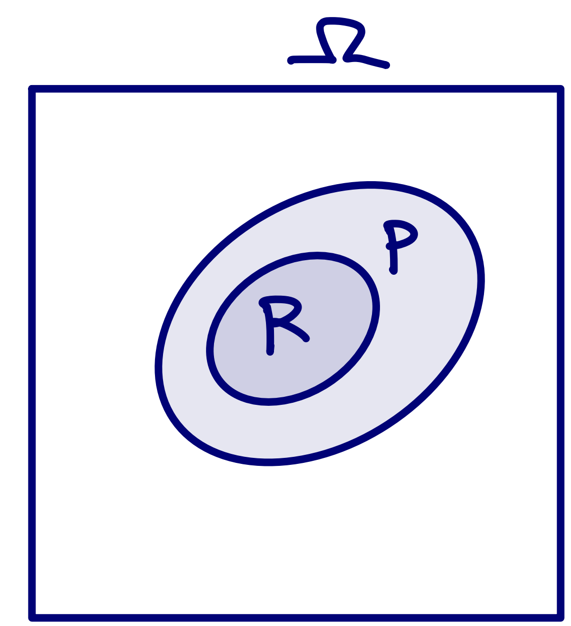
$= \frac{20}{25} + \frac{25}{25}$

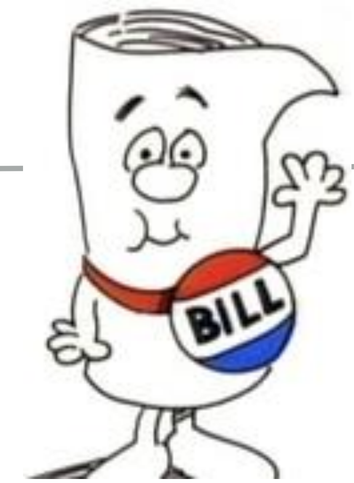
$= 0.9$

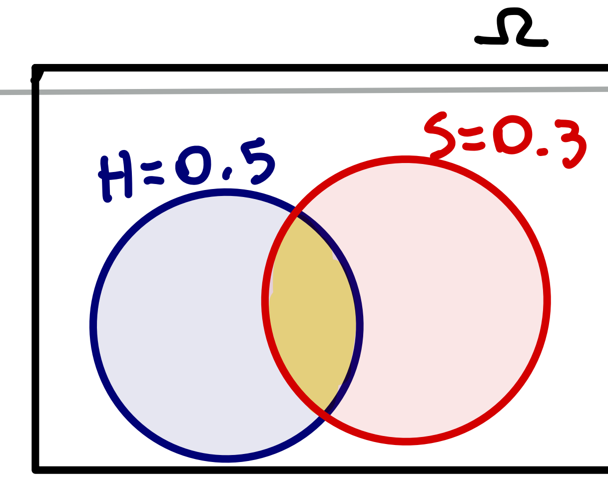


minimize $Pr(R \cup P)$

$Pr(R \cup P) = Pr(P) = \frac{25}{50} = 0.5$







I'M JUST A BILL

- For a bill to come before the US president, it must be passed by both the House and the Senate
- Suppose that 40% of bills pass the House, 30% the Senate, and 50% pass at least one of the two
- What is the Pr the next bill will come before the president?

$\Pr(H \cup S) = \Pr(H) + \Pr(S) - \Pr(H \cap S)$

$\Pr(H \cap S) = \Pr(H) + \Pr(S) - \Pr(H - S) = 0.3 + 0.4 - 0.5 = 0.2$

(A) 0.2 (B) 0.4 (C) 0.5 (D) 0.7

(E) None of the above

TOP HAT QUESTION

> ``Linda is 31, single, outspoken, and very bright. She majored in philosophy in college. As a student, she was deeply concerned with racial discrimination and other social issues, and participated in anti-nuclear demonstrations."

Rank the likelihood of the following alternatives:

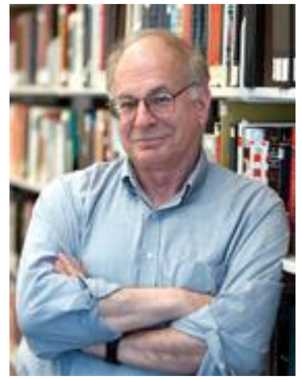
(a) Linda is active in the feminist movement

(b) Linda is a bank teller

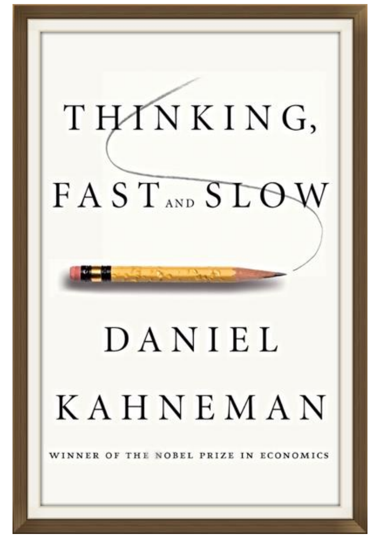
(c) Linda is a bank teller and active in the feminist movement

CONJUNCTION FALLACY

Amos Tversky
[1937-1996]



Daniel Kahneman
[1934-2024]



Between 80 and 95 percent of the subjects said (a) > (c) > (b)

Rank the likelihood of the following alternatives:

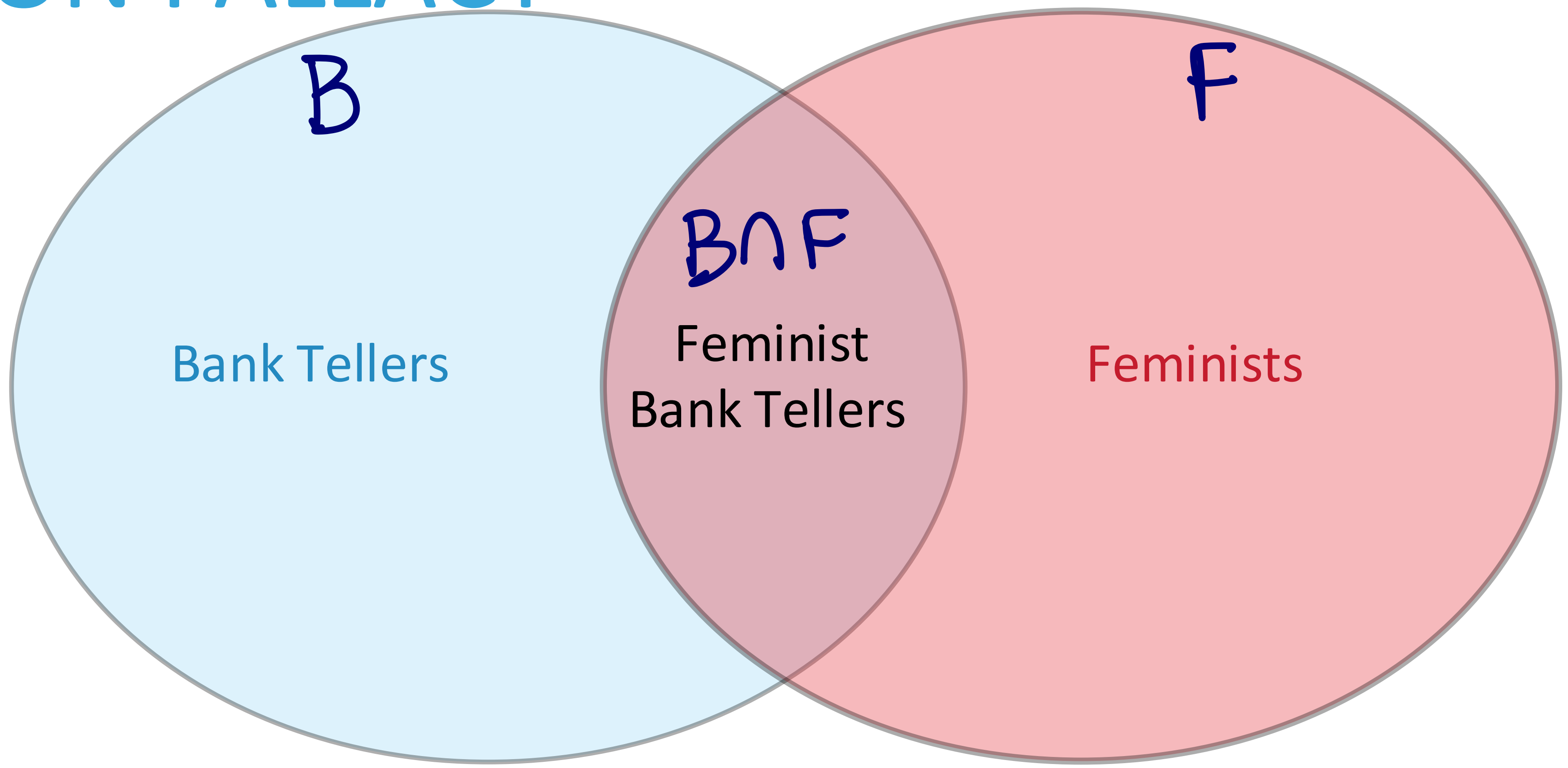
(a) Linda is active in the feminist movement

(b) Linda is a bank teller

(c) Linda is a bank teller and active in the feminist movement

$Pr(B \cap F) > Pr(B)$?

CONJUNCTION FALLACY



Rank the likelihood of the following alternatives:

(a) Linda is active in the feminist movement

(b) Linda is a bank teller

(c) Linda is a bank teller and active in the feminist movement

BETTING ON DICE

Chevalier de Mere
[1607-1684]

- We have a standard die and we roll it 4 times
- Should we bet on "at least one 6 turns up"?

A = "at least one 6 turns up in 4 rolls of a die"

A. YES

B. NO

BETTING ON DICE

Chevalier de Mere
[1607-1684]

- We have a standard die and we roll it 4 times
- Should we bet on "at least one 6 turns up"?

A = "at least one 6 turns up in 4 rolls of a die"

$$Pr(A) = 1 - Pr(\overline{A}) = 1 - \left(\frac{5}{6}\right)^4 \approx 0.5177$$

BETTING ON DICE


Chevalier de Mere

- We have 2 dice and we roll them 24 times
- Should we bet on "at least one pair of 6s turns up"?

B = "at least one pair of 6s turns up in 24 rolls of 2 dice"

$$Pr(B) = 1 - Pr(\overline{B}) = 1 - \left(\frac{35}{36}\right)^{24} \approx 0.4914$$

BETTING ON DICE

Chevalier de Mere

- We have 2 dice and we roll them 24 times
- Should we bet on "at least one pair of 6s turns up"?

C = "at least one pair of 6s turns up in 25 rolls of 2 dice"

$$Pr(B) = 1 - Pr(\overline{B}) = 1 - \left(\frac{35}{36}\right)^{25} \approx 0.5055$$

UN GRAND SCANDALE


Chevalier de Mere

- We have 2 dice and we roll them n times
- Should we bet on "at least one pair of 6s turns up"?
- A gambling rule of thumb at the time predicted:
  - 4 repetitions are favorable for an event with Pr = 1/6
  - This event is 6 times less likely (Pr = 1/36)
  - So one should use 6 times more repetitions: $6 \cdot 4 = 24$

UN GRAND SCANDALE

Chevalier de Mere

- We have 2 dice and we roll them n times
- Should we bet on "at least one pair of 6s turns up"?
- A gambling rule of thumb at the time predicted:
  - n = 24 rolls lead to favorable changes of winning bet
- de Mere's gambling experience
  - n = 25 rolls are needed for favorable chances

SEE YOU NEXT LECTURE!