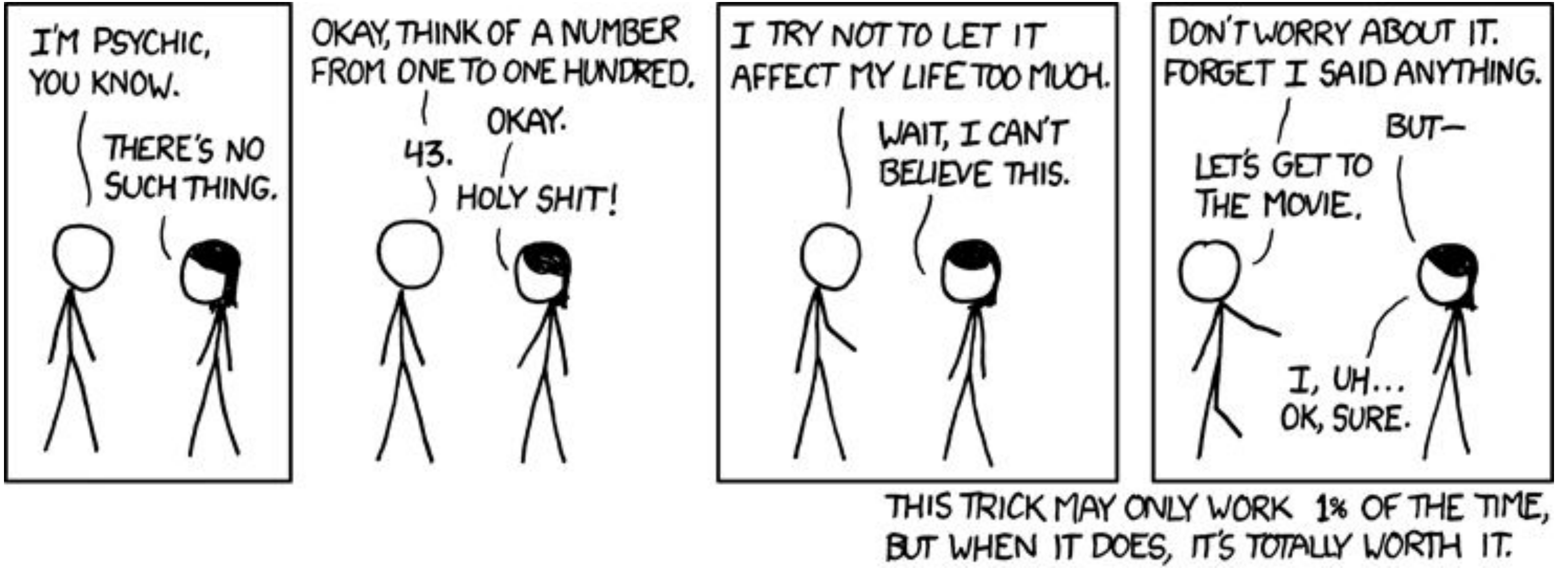


Lecture 4: Introduction to Probability … probably

Announcements and reminders

- Homework 1 due Friday 1 Feb at 5 PM

→ Go through tutorials & nb01

Monday: nb02 & 03

data vis., plotting

Why we need probability

Aspects of the world seem random and unpredictable

- Are we tall or short?
- Do we have Mom's eyes or Dad's?
- Is the eye of the hurricane going to pass over New Orleans?
- **Which team will win a best-of-seven series?**
- How long will it take to drive to the airport?
- Which grocery store line should I get in?

Why we need probability

Aspects of the world seem random and unpredictable

- Are we tall or short?
- Do we have Mom's eyes or Dad's?
- Is the eye of the hurricane going to pass over New Orleans?
- **Which team will win a best-of-seven series?**
- How long will it take to drive to the airport?
- Which grocery store line should I get in?

Probability is a way of thinking about unpredictable phenomena as if they were each generated from some **random process** …

Why we need probability

Aspects of the world seem random and unpredictable

- Are we tall or short?
- Do we have Mom’s eyes or Dad’s?
- Is the eye of the hurricane going to pass over New Orleans?
- **Which team will win a best-of-seven series?**
- How long will it take to drive to the airport?
- Which grocery store line should I get in?

Probability is a way of thinking about unpredictable phenomena as if they were each generated from some **random process** …

… and there are specific classes of random processes we can **describe with math**!

Basic definitions

\omega

\Omega

Think of a random process as a **trial** or **experiment**

Definition: The **sample space** Ω is the set of all possible outcomes of the experiment.

Example: If we flip a fair coin a single time, what is the sample space?

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

Example: If we are doing a poll, and ask each person their birth month, what is the sample space?

$$\Omega = \{\text{Jan}, \text{Feb}, \text{Mar}, \ldots\}$$

Observation: These are **discrete** sample spaces because there are a finite number of outcomes.

Basic definitions

Think of a random process as a **trial** or **experiment**

Definition: The **<u>sample space</u>** Ω is the set of all possible outcomes of the experiment.

Example: If we flip a fair coin a single time, what is the sample space?

$$\Omega = \{\text{Heads}, \text{Tails}\}$$

Example: If we are doing a poll, and ask each person their birth month, what is the sample space?

$$\Omega = \{\text{Jan}, \text{Feb}, \text{Mar}, \ldots, \text{Nov}, \text{Dec}\}$$

Observation: These are **<u>discrete</u>** sample spaces because there are a finite number of outcomes.

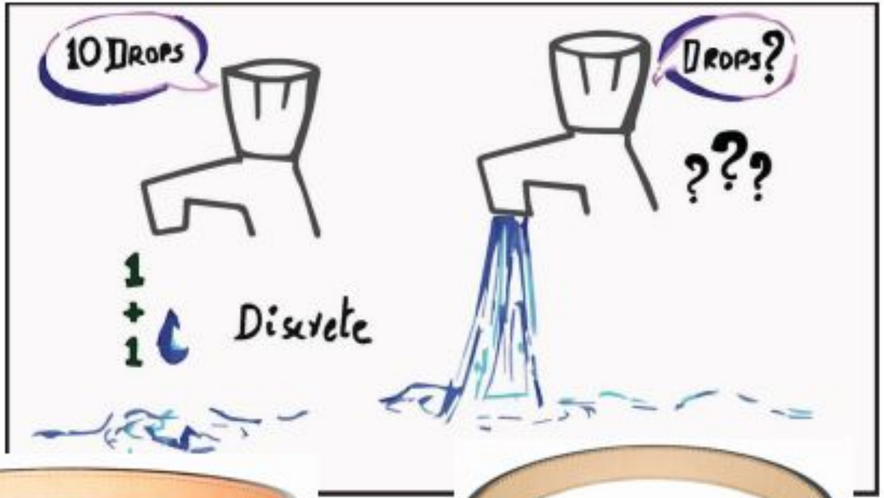
What does "Discrete Structures" mean?

It's the computer sciency way of saying... "**discrete math**"

Okay... then what is "**discrete math**"?

Well, there's *continuous* math
- like derivatives and integrals
- or the flow of water out of a faucet

and then there's *discrete* math
- like counting, sorting, enumeration
- or individual droplet of water



Basic definitions

Definition: For each event in Ω the **probability** is a measure between 0 and 1 of how likely it is for the event to occur.

Observation: The sum of the probability of each distinct outcome in Ω is 1. Why?

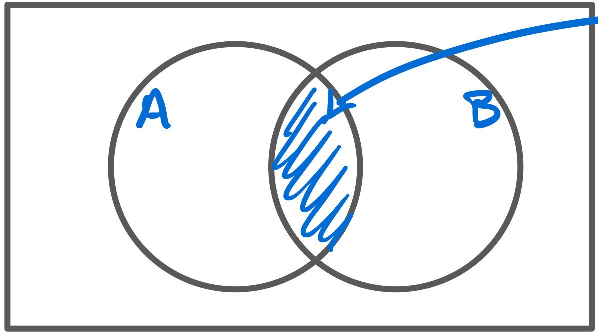
$$\boxed{p(\Omega) = 1}$$

all possible outcomes. & SOME thing must occur

Set operations

Definition: The **intersection** of two events is the subset of outcomes in **both** events

intersection = "and"

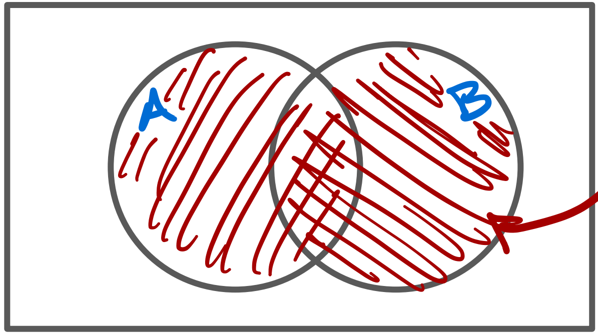


$A \cap B$

A \cap B

Definition: The **union** of two events is the subset of outcomes in **one or both** events

union = "or"

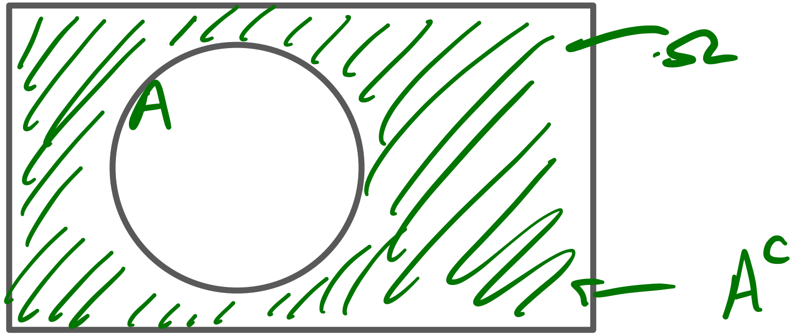


$A \cup B$

A \cup B

Set operations

Definition: The **complement** of an event A is the set of all outcomes in Ω that are **not** in A



A ^ c

Notation:

\overline{A} or \bar{A}

- Complement: A^c $(\bar{A})$
- Intersection: $A \cap B$
- Union: $A \cup B$

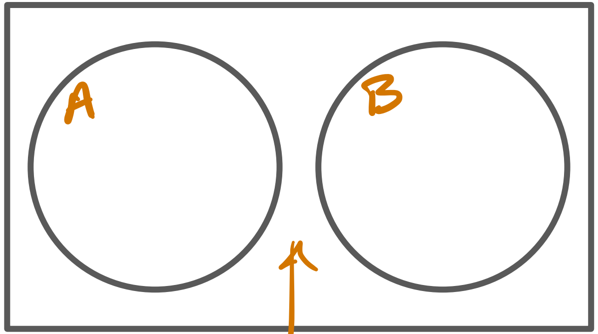
Set operations

Definition: When the intersection of two events is empty, we call those two events **disjoint** or **mutually exclusive**.

$A = \{H\}$

$B = \{T\}$

disjoint



$A = \{1, 2\}$

$B = \{5, 6\}$

$A \cap B =$ NOTHING $= \phi$

Notation:

- Null (empty) set: ϕ \emptyset

Set operations

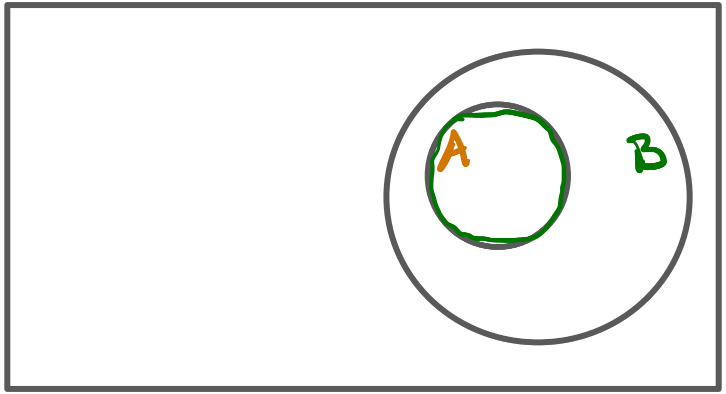
Definition: If all outcomes of event A are also outcomes of event B, we say A is a **subset** of B

Roll a fair 6-sided die:

$A = \{1,2\}$

$B = \{1,2,3,4\}$

$A \subset B$



Notation:

- Subset: $A \subseteq B$

If A is a subset of B AND $A \neq B$, then A is a proper subset of B: $A \subset B$

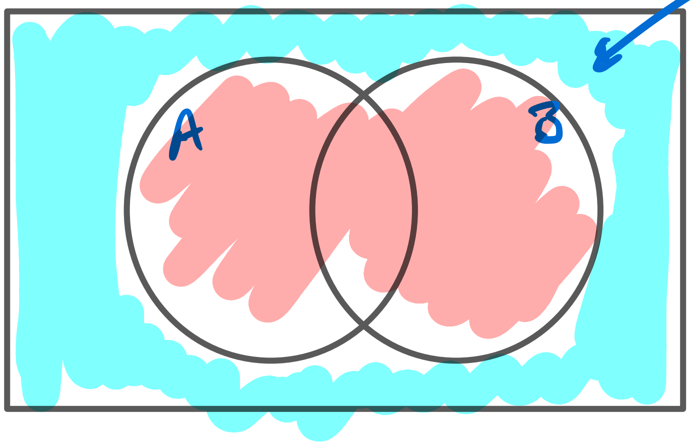
De Morgan's Laws

- Complement of a union: $(A \cup B)^c = A^c \cap B^c$ *
- Complement of an intersection: $(A \cap B)^c = A^c \cup B^c$

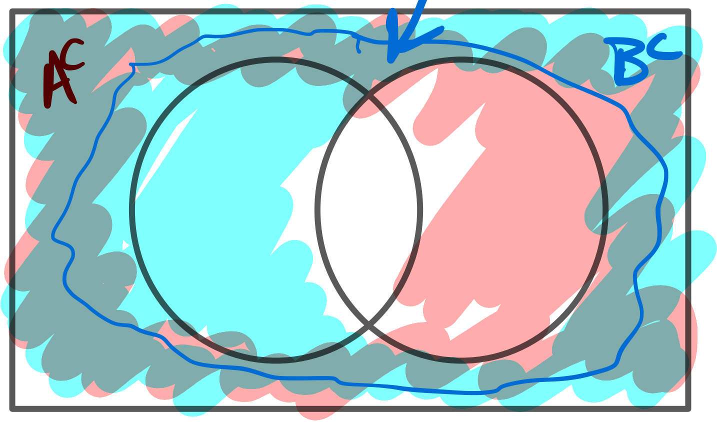
Question: Can we convince ourselves these are true using pictures?

First one:

$(A \cup B)^c$



$A^c \cap B^c$



Probability functions

A **biased coin** is a coin with a modified probability function

$P(H)$ $P(T)$

Instead of P({H, T}) = {½ , ½} , a biased coin's probability function is P({H, T}) = {p, q}

Question: What can we say about q?

$$P(\Omega) = 1$$

$$P(H) + P(T) = 1$$

$$q = P(T) = 1 - P(H) \longrightarrow q = 1 - p$$

Looking ahead: A random process with two outcomes with fixed probabilities assigned to each outcome is called a **Bernoulli Trial**.

Probability functions

A **biased coin** is a coin with a modified probability function

Instead of $P(\{H, T\}) = \{½ , ½\}$, a biased coin's probability function is $P(\{H, T\}) = \{p, q\}$

Question: What can we say about q?

$p(\Omega) = 1$

→ $q = 1-p$ (because the probabilities of all possible outcomes need to sum to 1)

→ $P(\{H, T\}) = \{p, 1-p\}$

Looking ahead: A random process with two outcomes with fixed probabilities assigned to each outcome is called a **Bernoulli Trial**.

Probability functions

Note that a probability function has two key properties:

1) The probability of the entire sample space is 1

2) The probability of the union of disjoint events is the sum of the probability of each event

Formal definition: A **probability function** P assigns to each event A a number P(A) in [0, 1] s.t.

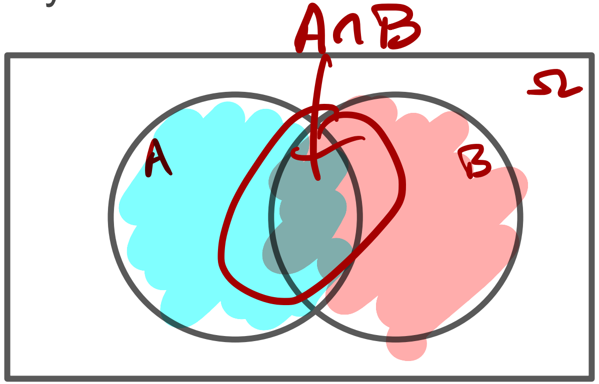
1) $P(\Omega) = 1$
2) $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint events

Probability of non-disjoint events

Question: What is the probability of the union of events A and B if they are *not* disjoint?

$p(A \cup B) = p(A) + p(B) - p(A \cap B)$

disjoint: $A \cap B = \phi$ & $p(A \cap B) = 0$



$p(A) = \frac{|A|}{|\Omega|}$

$p(B) = \frac{|B|}{|\Omega|}$

Example: A = {2, 4, 6}, B = {3, 6}

(S'pose we roll a fair 6-sided die)

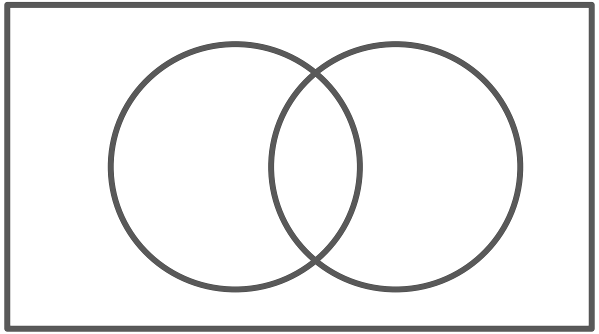
$p(A \cup B) = p(A) + p(B) - p(A \cap B)$

$= \frac{3}{6} + \frac{2}{6} - \frac{1}{6}$

$A \cap B = ? = \{6\}$

Probability of non-disjoint events

Question: What is the probability of the union of events A and B if they are ***not*** disjoint?



Example: $A = \{2, 4, 6\}$, $B = \{3, 6\}$

Disjoint

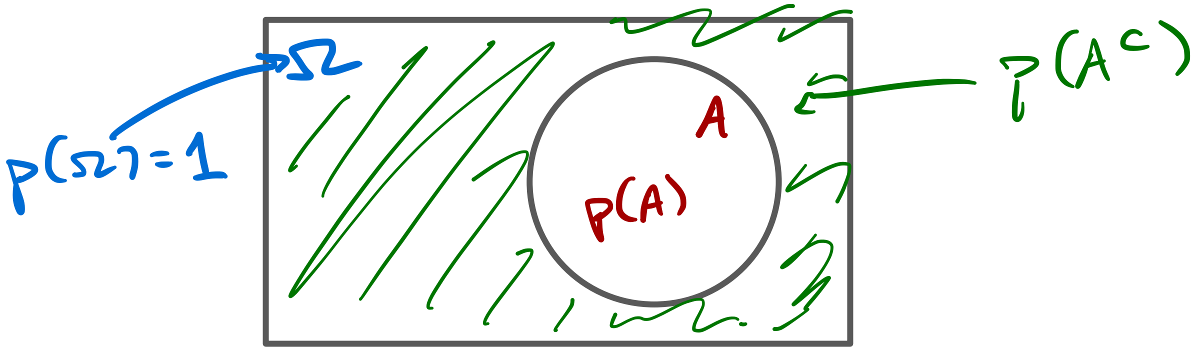
$A \cap B = \phi$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{6} = \tfrac{2}{3}$$

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$ **(always true)**

Probability of the complement

Question: What is the probability of the complement of an event A?



Example: A = {2, 4}, Ω = {1, 2, 3, 4, 5, 6} $\longrightarrow$ $A^c = \{1, 3, 5, 6\}$

$p(A^c) = 4/6$

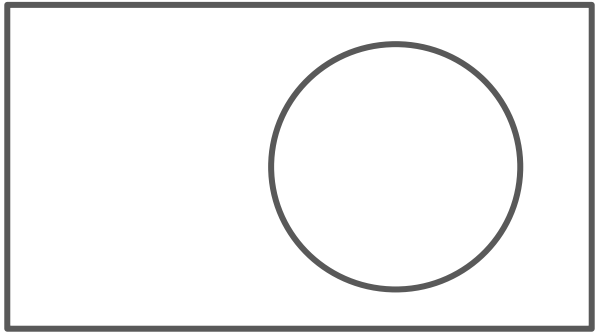
know: $A \cap A^c = \phi \longrightarrow$ disjoint

$\longrightarrow p(A \cup A^c) = p(A) + p(A^c)$

$\longrightarrow p(\Omega) = 1 = p(A) + p(A^c)$

Probability of the complement

Question: What is the probability of the complement of an event A?



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

$\Omega = A \cup A^c$, and A and A^c are **disjoint** $\rightarrow$ $P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$

$\rightarrow$ $1 = P(A) + P(A^c)$

$\rightarrow$ $\boxed{P(A^c) = 1 - P(A)}$

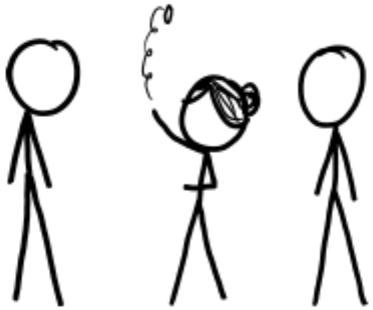
More complicated coins

Question: What is the probability that I flip a biased coin twice and both flips come up heads?

$p(HH) = p * p$

H on first flip

H on second flip



The sample space for a single coin flip is: Ω = {H, T}

The sample space for two coin flips is:
Ω = {H, T} × {H, T} = {(H, H), (H, T), (T, H), (T, T)}

→ This is an example of the <u>**product**</u> of sample spaces

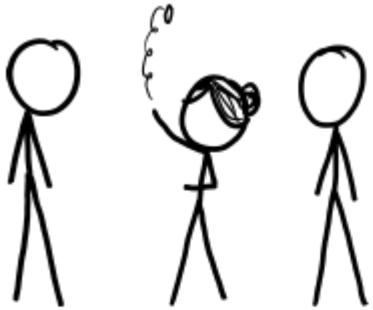
More complicated coins

Question: What is the probability that I flip a biased coin twice and both flips come up heads?

$P(\{H, T\}) = \{p, 1\text{-}p\}$

$P(HH) = P(H) \cdot P(H) = p{\cdot}p = p^2$

The sample space for a single coin flip is: $\Omega = \{H, T\}$



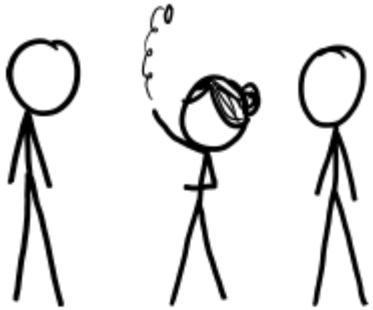
The sample space for two coin flips is:
$\Omega = \{H, T\} \times \{H, T\} = \{(H, H), (H, T), (T, H), (T, T)\}$

→ This is an example of the **product** of sample spaces

More complicated coins

Question: What is the probability that I flip a biased coin twice and both flips come up heads?

Sanity check: Does the result of the first flip affect the result of the second flip?



More complicated coins

Question: What is the probability that I flip a biased coin twice and both flips come up heads?

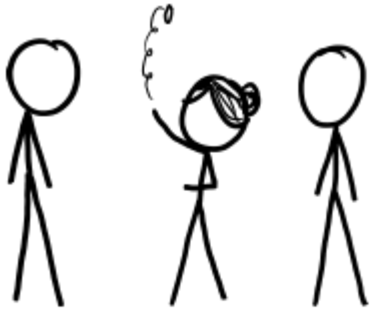
Sanity check: Does the result of the first flip affect the result of the second flip?

More later

Definition: When two (or more) trials do not affect each other, we say they are **<u>independent</u>**

Fun fact: When two events are independent, we can **multiply** their probabilities to find the probability of **both** occurring:

$$P((H, H)) = P(H) \cdot P(H) = p \cdot p = p^2$$



More complicated coins

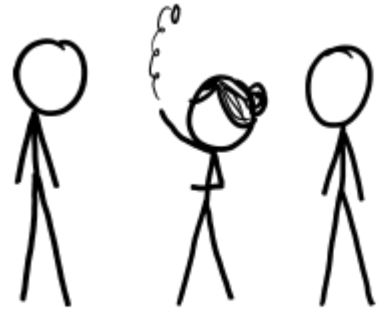
Question: What is the probability that I flip a biased coin twice and get one H and one T?

→ We want to know the probability of events (H, T) **OR** (T, H)

→ If the outcomes are independent, then **OR** means addition: (b/c they're **disjoint**)

P((H, T) **or** (T, H)) =

$p(1-p)$



More complicated coins

Question: What is the probability that I flip a biased coin twice and get one H and one T?

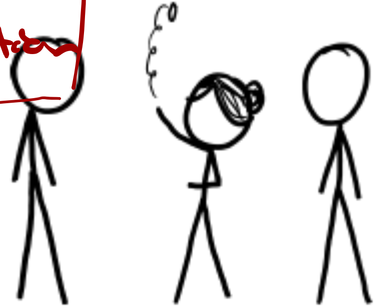
→ We want to know the probability of events (H, T) **OR** (T, H)

→ If the outcomes are independent, then **OR** means addition: (b/c they're **disjoint**)

no way for BOTH to occur

AND means multiplication

$$\begin{aligned} P(\,(H, T)\ \textbf{or}\ (T, H)\,) &= P((H, T)) + P((T, H)) \\ &= P(H)\cdot P(T) + P(T)\cdot P(H) \\ &= p\cdot(1-p) + (1-p)\cdot p \\ &= 2p(1-p) \end{aligned}$$



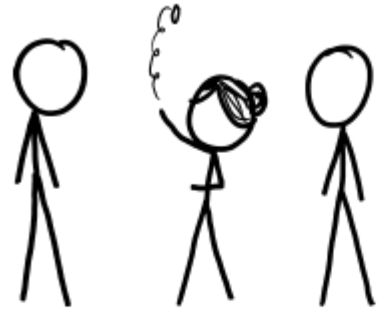
More complicated coins

Question: What is the probability that I flip 5 (possibly biased) coins and get exactly one H?

HTTTT, OR THTTT, OR TTHTT, TTTHT, TTTTH

$$p(1-p)^4 + (1-p)\,p\,(1-p)^3 + \dots$$

$$\dots = 5p(1-p)^4$$

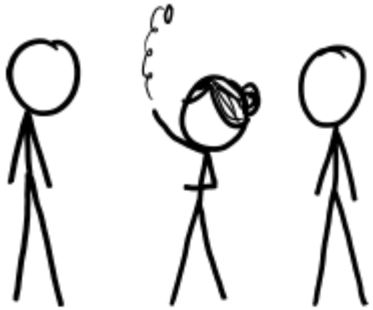


More complicated coins

Question: What is the probability that I flip 5 (possibly biased) coins and get exactly one H?

→ Want P({HTTTT, THTTT, TTHTT, TTTHT, TTTTH})

disjoint, so we can add their individual probabilities



More complicated coins

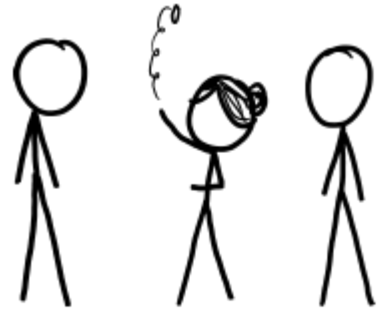
Question: What is the probability that I flip 5 (possibly biased) coins and get exactly one H?

→ Want P({HTTTT, THTTT, TTHTT, TTTHT, TTTTH})

→ All independent, disjoint, so…

$= 5 \cdot p \cdot (1-p)^4$

$= 5/32$ (if $p=½$)



An empirical estimate

Question: S'pose we have a coin but do not know if it is biased. But we want to estimate the probabilities of H and T.

→ What could we do?!

flip it so so many times

hats = estimate

& see $\hat{p} = \dfrac{\#\text{heads}}{\#\text{flips}} \approx 0.5$

ALL SPORTS COMMENTARY

An empirical estimate

Question: S'pose we have a coin but do not know if it is biased. But we want to estimate the probabilities of H and T.

→ What could we do?!

Let's find out!

- Get in groups, get out laptop, and open **nb04**
- **Today:** *How to approximate probabilities of events using random simulation!*

ALL SPORTS COMMENTARY