CSE 312

Foundations of Computing II

Lecture 5: Introduction to probability

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Stefano Tessaro
tessaro@cs.washington.edu

Organization

- HW1 – Gradescope info soon
- Note office hours change
  - My own are on Monday 3-4pm right now.
  - Few more have moved – check homepage!
- https://forms.gle/QA2ASR4LzepGEVKT7
  - Slides online

Today

- Combinatorics wrap-up: Pigeonhole principle
- Introduction to probability

Pigeonhole Principle – Goal

Mostly showing that a set S has **at least** a certain size.

Pigeonhole Principle

If there are n pigeons in $k < n$ holes, then one hole must contain at least two pigeons!

Pigeonhole Principle – Refined version

If there are n pigeons in $k < n$ holes, then one hole must contain at least $\frac{n}{k}$ pigeons!

Pigeonhole Principle – Refined version

If there are n pigeons in $k < n$ holes, then one hole must contain at least $\frac{n}{k}$ pigeons!

Proof. Assume there are $< n/k$ pigeons per hole.

Then, there are $< k \times \frac{n}{k} = n$ pigeons overall.

A contradiction!

Pigeonhole Principle – Even stronger version

If there are n pigeons in $k < n$ holes, then one hole must contain at least $\left\lceil \frac{n}{k} \right\rceil$ pigeons!

Here: $\lceil x \rceil$ = first integer $\geq x$ (verbalized "ceiling of x")
e.g., $\lceil 2.731 \rceil = 3$, $\lceil 3 \rceil = 3$

Reason: There can only be an integer number of pigeons in a hole …

Pigeonhole Principle – Example

In a room with 367 people, there are at least two with the same birthday.

Solution:

- **367** pigeons = people
- **366** holes = possible birthdays
- Person goes into hole corresponding to own birthday

Pigeonhole Principle – Example (Surprising?)

In every set S of 38 numbers, there are two whose difference is a multiple of 37.

Solution:

- **38** pigeons = numbers $n \in S$
- **37** Holes = Numbers $\{0, \ldots, 36\}$
- A number $n \in S$ goes into hole $n \bmod 37$

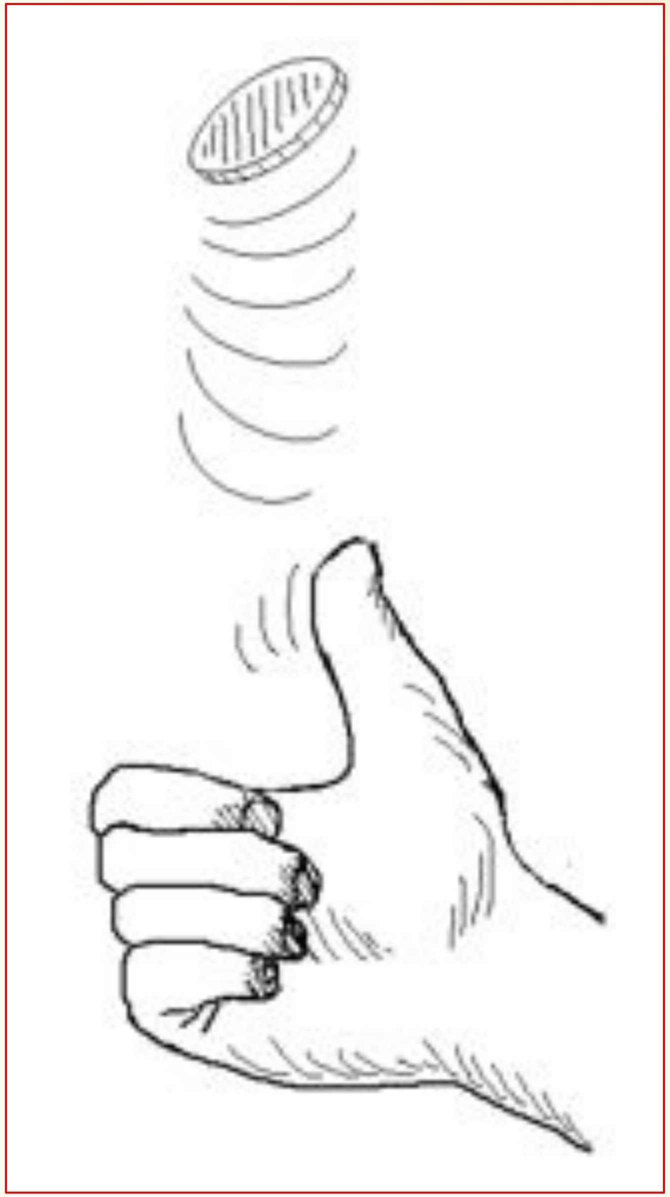
PHP → there are distinct $n_1, n_2 \in S$ s.t. $n_1 \bmod 37 = n_2 \bmod 37$
→ $n_1 - n_2$ multiple of 37

Next – Probability

- We want to model a non-deterministic process.
  - i.e., outcome not determined a-priori
  - E.g. throwing dice, flipping a coin, …
  - We want to numerically measure likelihood of outcomes = probability.
  - We want to make complex statements about these likelihoods.
- We will not argue why a certain physical process realizes the probabilistic model we study
  - Why is the outcome of the coin flip really "random"?
- First part of class: "Discrete" probability theory
  - Experiment with finite / discrete set of outcomes.

Example – Coin Tossing

Imagine we toss coins – each one can be **heads** or **tails**.



Probability space

Either finite or infinite countable (e.g., integers)

Definition. A (discrete) **probability space** is a pair $(\Omega, \mathbb{P})$ where:

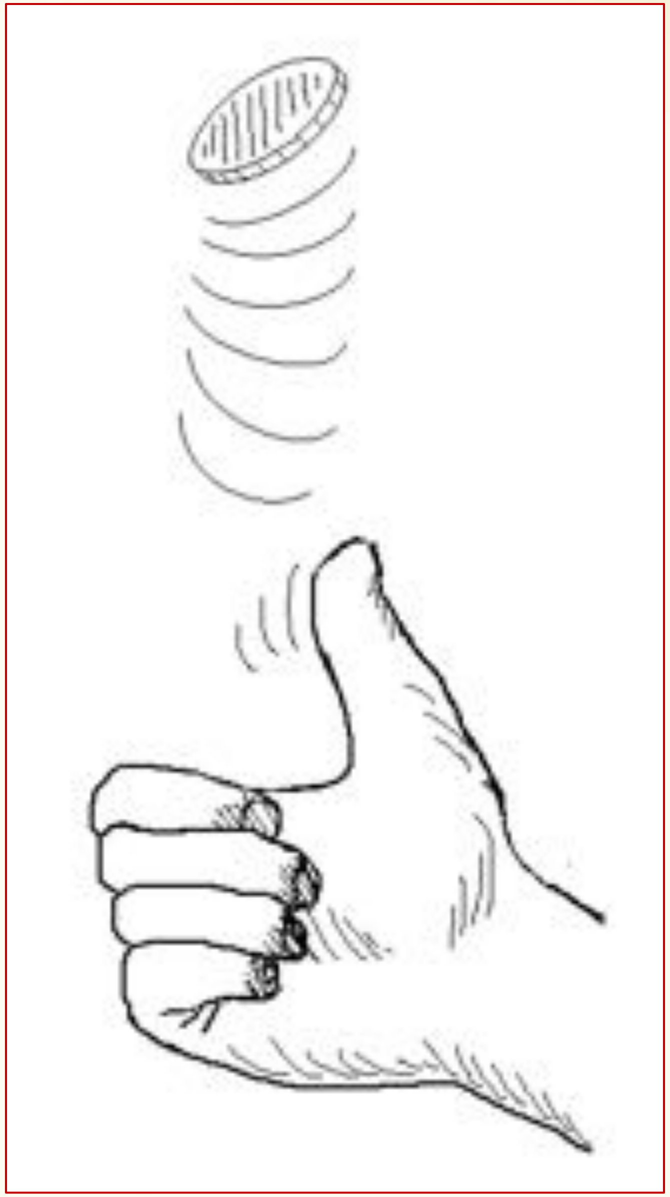
- Ω is a set called the **sample space.**
- $\mathbb{P}$ is the **probability measure,** a function $\mathbb{P}: \Omega \to \mathbb{R}$ such that:
  - $\mathbb{P}(x) \geq 0$ for all $x \in \Omega$
  - $\sum_{x \in \Omega} \mathbb{P}(x) = 1$

Set of possible **elementary outcomes**

Likelihood of each **elementary outcome**

Example – Coin Tossing

Imagine we toss one coin – outcome can be **heads** or **tails.**



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

$\mathbb{P}$? Depends! What do we want to model?!

Fair coin toss

$$\mathbb{P}(\mathrm{H}) = \mathbb{P}(\mathrm{T}) = \frac{1}{2} = 0.5$$

Example – Unfair Coin Toss

Imagine we toss an unfair coin – outcome can be **heads** or **tails**.

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

$$\mathbb{P}(\mathrm{H}) = p$$

$$\mathbb{P}(\mathrm{T}) = 1 - p$$

p

$1 - p$

p can be determined by tossing coin several times and observing frequency of heads ("frequentist interpretation")

Example – Two Coin Tosses

Imagine we toss **two** fair coins

25% 25%

25% 25%

$$\Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$$

$$\mathbb{P}(\mathrm{HH}) = \mathbb{P}(\mathrm{HH}) = \mathbb{P}(\mathrm{HH}) = \mathbb{P}(\mathrm{HH}) = \frac{1}{4} = 0.25$$

Example – Glued Coin Tosses

50%

50%

Imagine we toss **two** coins, glued to always show opposite faces.

$$\Omega = \{\mathrm{HT}, \mathrm{TH}\}$$

$$\mathbb{P}(\mathrm{HT}) = \mathbb{P}(\mathrm{TH}) = 0.5$$

Example – Fair Dice

We throw two fair dice.

$$\Omega = \{(i,j) \mid i,j \in [6]\}$$

$$\mathbb{P}\big((i,j)\big) = \frac{1}{36}.$$

Uniform Probability Space

Definition. A **uniform probability space** is a pair $(\Omega, \mathbb{P})$ such that

$$\mathbb{P}(x) = \frac{1}{|\Omega|}$$

for all $x \in \Omega$.

All of the above are uniform, except the unfair coin!

Summary – Probability spaces

- The probability space describes only a **single experiment,** sampling a **single outcome.**
- Two-toss experiment ≠ 2 x one-toss experiment
  - We need to explicitly explain how the two coin tosses are related.

Next – Events

Typical questions we would like to answer in a random experiment.

Assume that we flip three fair coins, then:

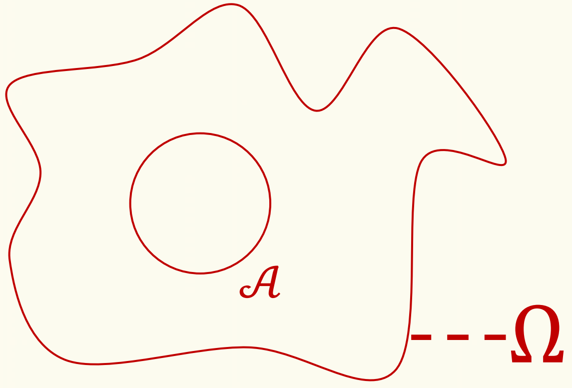
- *What is the probability that all tosses give us tails?*
- *What is the probability that we get heads at least once?*
- *What is the probability that we get an even number of heads?*

These are not basic outcomes!

Events

Definition. An **event** in a probability space $(\Omega, \mathbb{P})$ is a subset $\mathcal{A} \subseteq \Omega$. Its probability is

$$\mathbb{P}(\mathcal{A}) = \sum_{\omega \in \mathcal{A}} \mathbb{P}(\omega)$$



Convenient abuse of notation: $\mathbb{P}$ is extended to be defined over **sets**

$$\mathbb{P}(\omega) = \mathbb{P}(\{\omega\})$$

Events - Examples

Definition. An **event** in a probability space $(\Omega, \mathbb{P})$ is a subset $\mathcal{A} \subseteq \Omega$. Its probability is

$$\mathbb{P}(\mathcal{A}) = \sum_{\omega \in \mathcal{A}} \mathbb{P}(\omega)$$

$$\Omega = \{\mathrm{HHH}, \mathrm{HHT}, \dots, \mathrm{TTT}\}$$

$$\mathbb{P}(\mathrm{HHH}) = \mathbb{P}(\mathrm{HHT}) = \cdots = \mathbb{P}(\mathrm{TTT}) = 1/8$$

"all tosses give us tails"

$$\mathcal{A} = \{\mathrm{TTT}\} \qquad \mathbb{P}(\mathcal{A}) = \mathbb{P}(\mathrm{TTT}) = \frac{1}{8}$$

"we get heads at least once"

$$\mathcal{B} = \{\mathrm{TTH}, \mathrm{THT}, \mathrm{THH}, \mathrm{HTT}, \mathrm{HTH}, \mathrm{HHT}, \mathrm{HHH}\}$$

$$\mathbb{P}(\mathcal{B}) = 7 \times \frac{1}{8} = \frac{7}{8}$$

"we get an even number of heads"

$$\mathcal{C} = \{\mathrm{TTT}, \mathrm{THH}, \mathrm{HHT}, \mathrm{HTH}\}$$

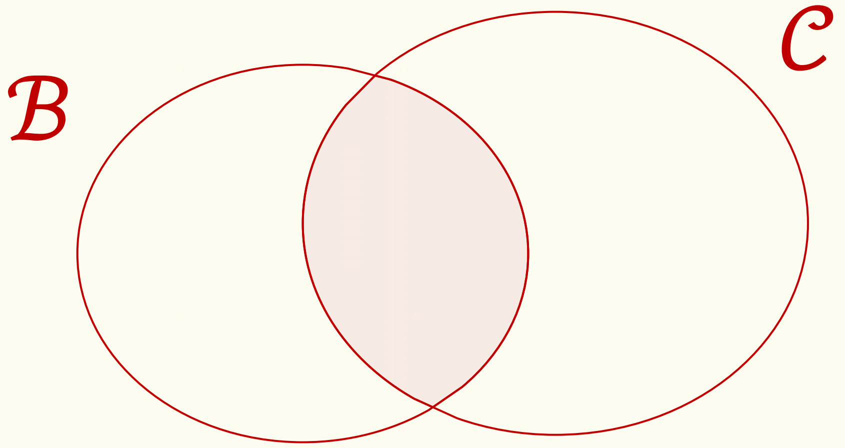
$$\mathbb{P}(\mathcal{C}) = 4 \times \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

Events – AND

"we get heads at least once" $\mathcal{B} = \{\mathrm{TTH}, \mathrm{THT}, \mathrm{THH}, \mathrm{HTT}, \mathrm{HTH}, \mathrm{HHT}, \mathrm{HHH}\}$

"we get an even number of heads" $\mathcal{C} = \{\mathrm{TTT}, \mathrm{THH}, \mathrm{HHT}, \mathrm{HTH}\}$

*"we get heads at least once **and** we get an even number of heads"*



$\mathcal{B} \cap \mathcal{C} = \{\mathrm{THH}, \mathrm{HHT}, \mathrm{HTH}\}$

$$\mathbb{P}(\mathcal{B} \cap \mathcal{C}) = \frac{3}{8}$$

Events – OR

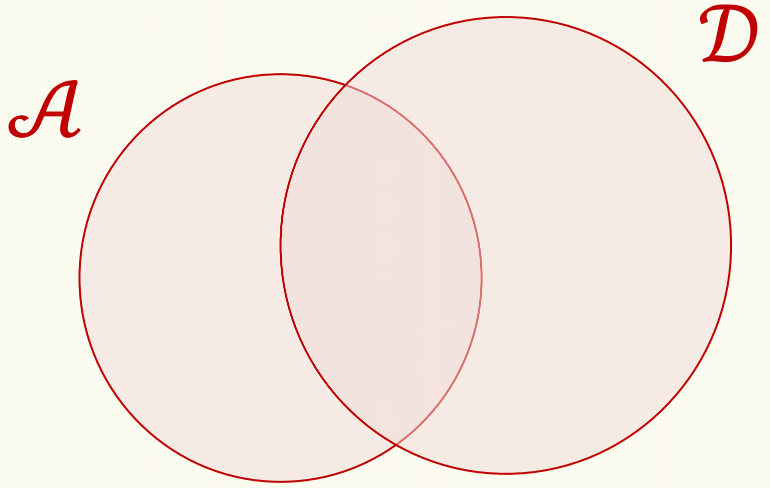
"all tosses give us tails" $\mathcal{A} = \{\mathrm{TTT}\}$

"all tosses give us heads" $\mathcal{D} = \{\mathrm{HHH}\}$

*"all tosses give us tails **or*** all tosses give us heads"

$$\mathcal{A} \cup \mathcal{D} = \{\mathrm{TTT}, \mathrm{HHH}\}$$

$$\mathbb{P}(\mathcal{A} \cup \mathcal{D}) = \frac{2}{8} = \frac{1}{4}$$



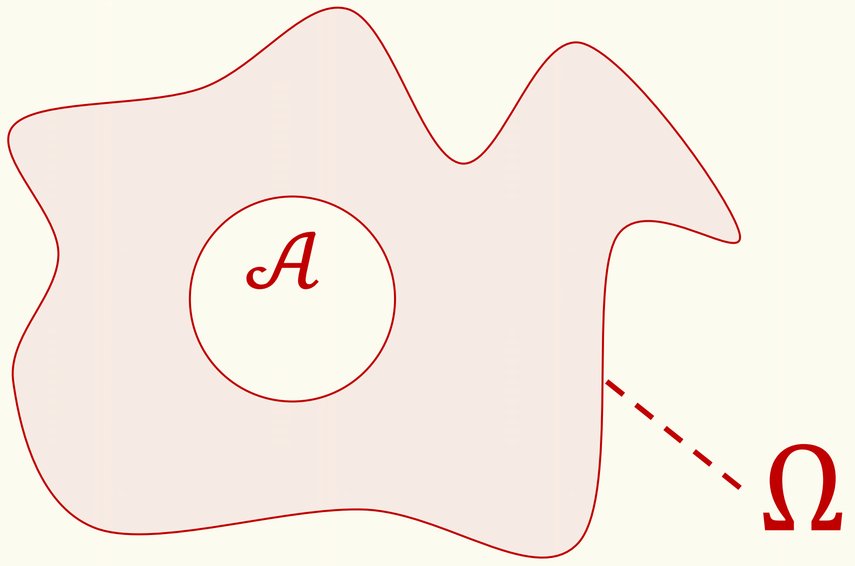
Events – NOT

"all tosses give us tails" $\mathcal{A} = \{\mathrm{TTT}\}$

"not all tosses are tails"

$$\mathcal{A}^c = \{\mathrm{TTH}, \mathrm{THT}, \mathrm{THH}, \mathrm{HTT}, \mathrm{HTH}, \mathrm{HHT}, \mathrm{HHH}\}$$

$$\mathbb{P}(\mathcal{A}^c) = \frac{7}{8}$$



Properties of Probability

$$\mathbb{P}(\mathcal{A}) = \sum_{\omega \in \mathcal{A}} \mathbb{P}(\omega)$$

For all events $\mathcal{A}, \mathcal{B}$

$$0 \leq \mathbb{P}(\mathcal{A}) \leq 1$$

$$\mathbb{P}(\emptyset) = 0$$

$$\mathbb{P}(\Omega) = 1$$

$$\mathbb{P}(\mathcal{A} \cup \mathcal{B}) = \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

$$\mathbb{P}(\mathcal{A}^c) = 1 - \mathbb{P}(\mathcal{A})$$

If $\mathcal{A} \subseteq \mathcal{B}$ then $\mathbb{P}(\mathcal{B} \setminus \mathcal{A}) = \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A})$

Example – Fair Dice

We throw two fair dice.

$$\Omega = \{(i,j) \mid i,j \in [6]\}$$

$$\mathbb{P}\big((i,j)\big) = \frac{1}{36}.$$

What is the probability the two dice add up to 7?

What is the probability the two dice do not add up to 7?

Example – Fair Dice

$$\Omega = \{(i,j) \mid i,j \in [6]\} \qquad \mathbb{P}\big((i,j)\big) = \frac{1}{36}.$$

What is the probability the two dice add up to 7?

$$\mathcal{A} = \{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}$$

$$\mathbb{P}(\mathcal{A}) = \frac{6}{36} = \frac{1}{6}$$

What is the probability the two dice do not add up to 7?

$$\mathbb{P}(\mathcal{A}^c) = 1 - \frac{1}{6} = \frac{5}{6}$$