CSE 312

Foundations of Computing II

Lecture 6: Probability of Events

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Stefano Tessaro

tessaro@cs.washington.edu

Last time - recap

- Probability spaces
- Events

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 $x \in \Omega$
  - $\sum_{x \in \Omega} \mathbb{P}(x) = 1$

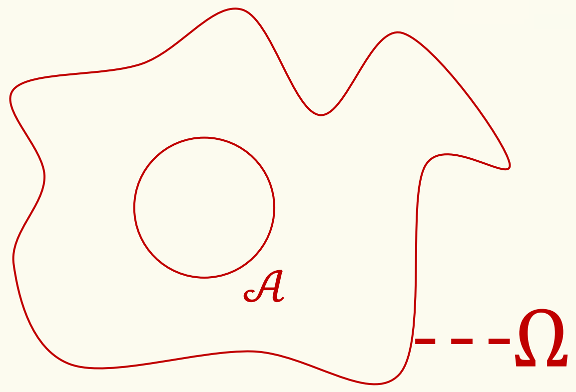
Set of possible **elementary outcomes**

Likelihood of each **elementary outcome**

Events

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

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



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

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

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

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?

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

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

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?

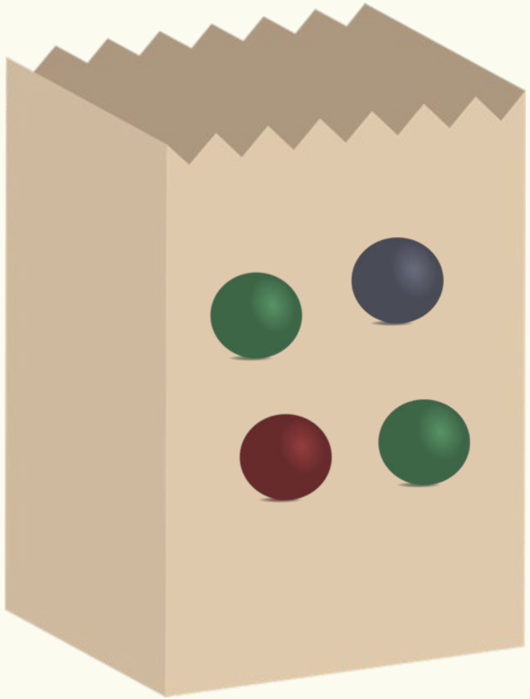
$$\mathbb{P}(\mathcal{A}) = \frac{6}{36} = \frac{1}{6}$$

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

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

Picking Balls

Pick two random balls, w/o replacement



$$\Omega = \{\mathrm{rg}, \mathrm{rb}, \mathrm{gr}, \mathrm{gg}, \mathrm{gb}, \mathrm{bg}, \mathrm{br}\}$$

$$\mathbb{P}(\mathrm{rg}) = \mathbb{P}(\mathrm{bg}) = \frac{1}{4} \times \frac{2}{3} = \frac{1}{6}$$

$$\mathbb{P}(\mathrm{rb}) = \mathbb{P}(\mathrm{br}) = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$$

$$\mathbb{P}(\mathrm{gr}) = \mathbb{P}(\mathrm{gg}) = \mathbb{P}(\mathrm{gb}) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$

"we do not get black" $\mathcal{A} = \{\mathrm{rg}, \mathrm{gr}, \mathrm{gg}\}$ $\mathbb{P}(\mathcal{A}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

"first ball is red" $\mathcal{B} = \{\mathrm{rg}, \mathrm{rb}\}$ $\mathbb{P}(\mathcal{B}) = \frac{1}{6} + \frac{1}{12} = \frac{1}{4}$

Example – Birthdays

Context: 5 people are assigned random birthdays (out of 365 possible ones). What is the probability that they all have different birthdays?

$$\Omega = [365]^5$$

$$\mathbb{P}(\omega) = \frac{1}{365^5}$$

$$\mathcal{E} = \{(x_1, \ldots, x_5) \in [365]^5 \mid x_1, \ldots, x_5 \text{ are distinct}\}$$

Example – Birthdays

$$\mathbb{P}(\omega) = \frac{1}{365^5}$$

$$|\mathcal{E}| = 365\times364\times363\times362\times361$$

$$\mathbb{P}(\mathcal{E}) = \frac{365\times364\times363\times362\times361}{365^5}$$

$$= \frac{365}{365}\times\frac{364}{365}\times\frac{363}{365}\times\frac{362}{365}\times\frac{361}{365}$$

$$= 0.97286\ldots$$

Example – Birthdays - Generalization

Context: $n \leq 365$ people are assigned random birthdays (out of 365 possible ones). What is the probability that they all have different birthdays?

$$\Omega = [365]^n \qquad \mathbb{P}(\omega) = \frac{1}{365^n}$$

$$\mathcal{E} = \{(x_1, \ldots, x_n) \in [365]^n \mid x_1, \ldots, x_n \text{ are distinct}\}$$

Example – Birthdays

$$\mathbb{P}(\omega) = \frac{1}{365^n}$$

$$|\mathcal{E}| = 365 \times 364 \times \cdots \times (365 - n + 1)$$

$$\mathbb{P}(\mathcal{E}) = \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

$$= \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{365 - n + 1}{365}$$

$$= ?$$

Example – Birthdays

$$\mathbb{P}(\mathcal{E}) = \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{365 - n + 1}{365}$$

$$= \left(1 - \frac{0}{365}\right) \times \left(1 - \frac{1}{365}\right) \times \cdots \times \left(1 - \frac{n-1}{365}\right)$$

$$\leq e^{-\frac{0}{365}} \times e^{-\frac{1}{365}} \times e^{-\frac{2}{365}} \times \cdots \times e^{-\frac{n-1}{365}}$$

Fact. $1 - x \leq e^{-x}$ where $e = 2.71828 \ldots$

Example – Birthdays

$$\mathbb{P}(\mathcal{E}) \leq e^{-\frac{0}{365}} \times e^{-\frac{1}{365}} \times e^{-\frac{2}{365}} \times \cdots \times e^{-\frac{n-1}{365}}$$

$$= e^{-\left(\frac{0}{365}+\frac{1}{365}+\cdots+\frac{n-1}{365}\right)} = e^{-\frac{n(n-1)}{2\times 365}}$$

Fact. $\sum_{i=0}^{n-1} i = \frac{n(n-1)}{2}$

(cf. section 0)

e.g., $n = 23$

$$\mathbb{P}(\mathcal{E}) \leq e^{-0.69315\ldots} \approx 0.5$$

[Near matching lower bound can be proved]

Birthday Paradox

Interpretation. With 23 people we are already guaranteed that two people have the same Birthday with probability ~ 0.5.

(In class: we are almost certain!)

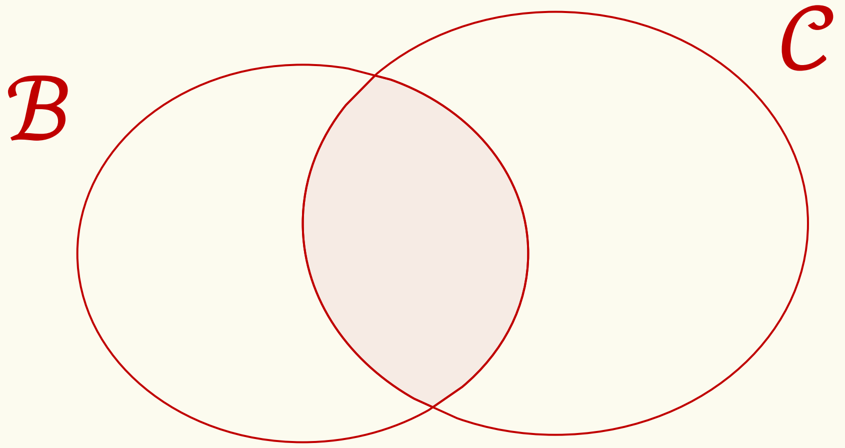
Next: Composing events

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

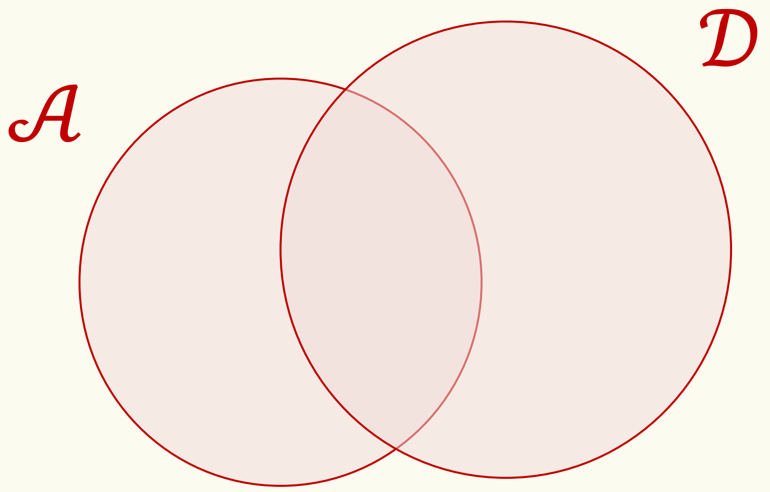
"all tosses give us tails" $\mathcal{A} = \{\mathrm{TTT}\}$

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

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

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

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



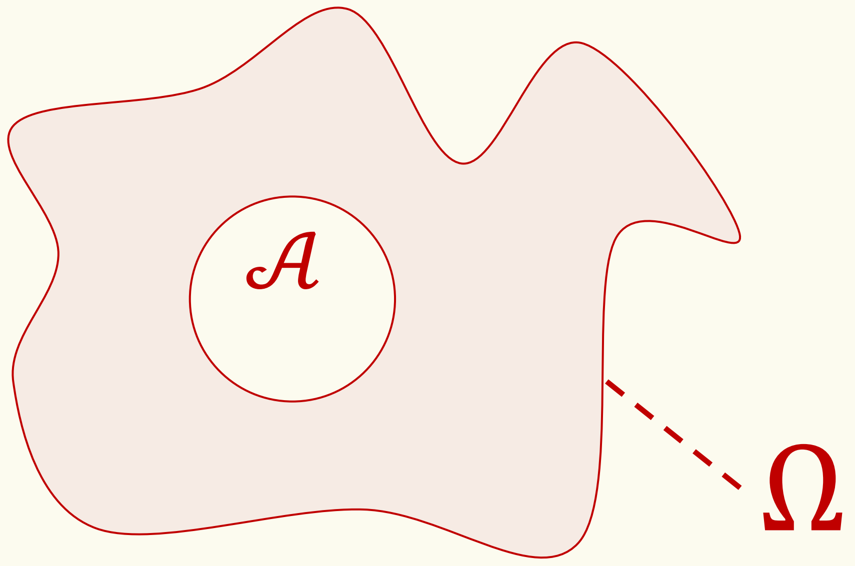
Events – NOT

"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

Theorem. If the events $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ are pairwise disjoint, i.e., $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$ for all $i \neq j$, then

$$\mathbb{P}(\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots \cup \mathcal{A}_n) = \mathbb{P}(\mathcal{A}_1) + \mathbb{P}(\mathcal{A}_2) + \cdots + \mathbb{P}(\mathcal{A}_n)$$

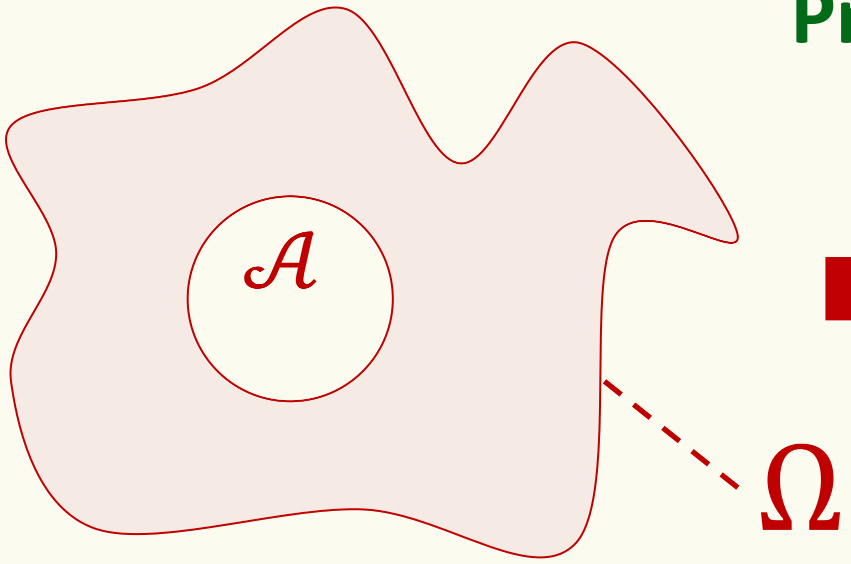
"Additivity of probability"

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

Consequences of Additivity – Complement

Fact. For any $\mathcal{A}$

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



Proof: $\mathcal{A}^c$ and are $\mathcal{A}$ disjoint

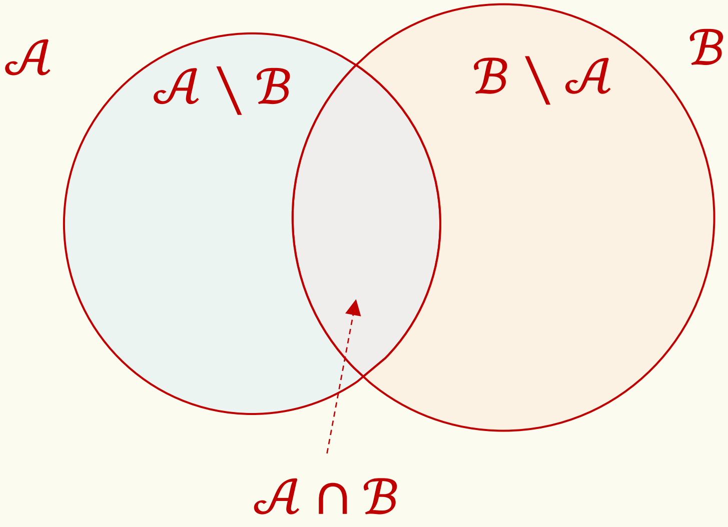
$$\mathcal{A}^c \cup \mathcal{A} = \Omega$$

$$\Rightarrow \quad 1 = \mathbb{P}(\Omega) = \mathbb{P}(\mathcal{A}^c) + \mathbb{P}(\mathcal{A})$$

Consequences of Additivity – Inclusion-Exclusion

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

Proof. $\mathbb{P}(\mathcal{A} \cup \mathcal{B}) = \mathbb{P}(\mathcal{A} \setminus \mathcal{B}) + \mathbb{P}(\mathcal{B} \setminus \mathcal{A}) + \mathbb{P}(\mathcal{A} \cap \mathcal{B})$



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

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

$$\Rightarrow \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) = \mathbb{P}(\mathcal{B} \setminus \mathcal{A}) + \mathbb{P}(\mathcal{B} \setminus \mathcal{A}) + 2\mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

Consequences of Additivity – Law of Total Probability

Theorem. If the events $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ are a partition of Ω, then for any event $\mathcal{B}$,

$$\mathbb{P}(\mathcal{B}) = \mathbb{P}(\mathcal{A}_1 \cap \mathcal{B}) + \cdots + \mathbb{P}(\mathcal{A}_n \cap \mathcal{B})$$