TÉCNICO LISBOA

Planning, Learning and Decision Making

Lecture 1

What is probability?

- Has roots in games of chance
- Used to **measure** the likelihood of occurrence of events

Natural tool to model uncertainty

What is probability?

- Classical definition of probability of event A
  - N possible events
  - N_A ways by which A can occur

$$\mathbb{P}[A] = \frac{N_A}{N}$$

  - Examples:
    - Throw a die 10 times: 1, 2, 3, 2, 3, 5, 4, 6, 2, 1
    - What is the probability of drawing an even number in a die?

TÉCNICO LISBOA

What is probability?

- Frequentist definition of probability of event A
  - Relative frequency of event A
  - N observed events
  - N_A times that event A occurred

$$\mathbb{P}[A] = \lim_{N \to \infty} \frac{N_A}{N}$$

  - Examples:
    - What is the probability of drawing a $\spadesuit$ in a card deck?

What is probability?

- Subjective definition of probability of event *A*
  - *Degree of belief* that event *A* may occur
- Example:
  - Probability that Robert de Niro has won three Oscars

Formally, ...

- **Probability space is a triplet (Ω, $\mathscr{F}$, P) where:**
  - Ω is the *sample space* ← Things that can happen
  - $\mathscr{F}$ is the set of events ← Things we want to measure
  - **P** is a probability measure ← Way to measure them

Formally, ...

- **Sample space Ω**
  - Space of possible *outcomes*
  - Each and every thing that may happen
  - Example:
    In a die throw, the possible outcomes are {1, 2, 3, 4, 5, 6}
  - Example:
    When drawing a card, the possible outcomes are {A♤,
    2♤, ..., J♢, Q♢, K♢}

Formally, ...

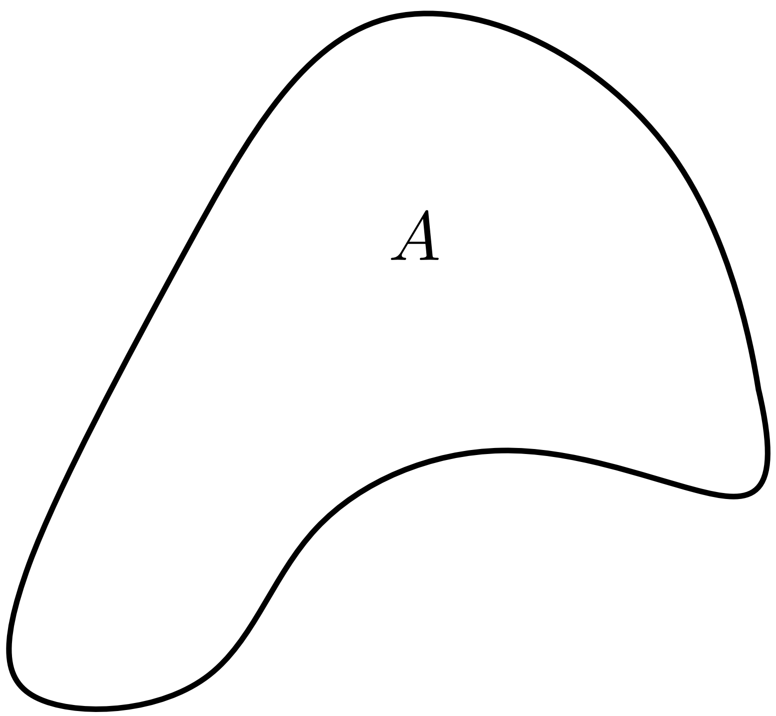
- **Set of events, $\mathscr{F}$**
  - Subsets of Ω that we can "measure", i.e., assign a probability
  - Includes the empty set $\varnothing$ and the full set Ω
  - Examples of events (die throw):
    - Drawing an even number: $\{2, 4, 6\}$
    - Drawing a number larger than 3: $\{4, 5, 6\}$
    - Drawing no number: $\varnothing$

Formally, ...

- **Set of events, $\mathscr{F}$**
  - Subsets of Ω that we can "measure", i.e., assign a probability
  - Includes the empty set $\varnothing$ and the full set Ω
  - Examples of events (card draw):
    - Drawing a spade: {A♤, 2♤, 3♤, 4♤, 5♤, 6♤, 7♤, 8♤, 9♤, 10♤, J♤, Q♤, K♤}
    - Drawing a figure: {J♤, Q♤, K♤, J♧, Q♧, K♧, J♡, Q♡, K♡, J♢, Q♢, K♢}

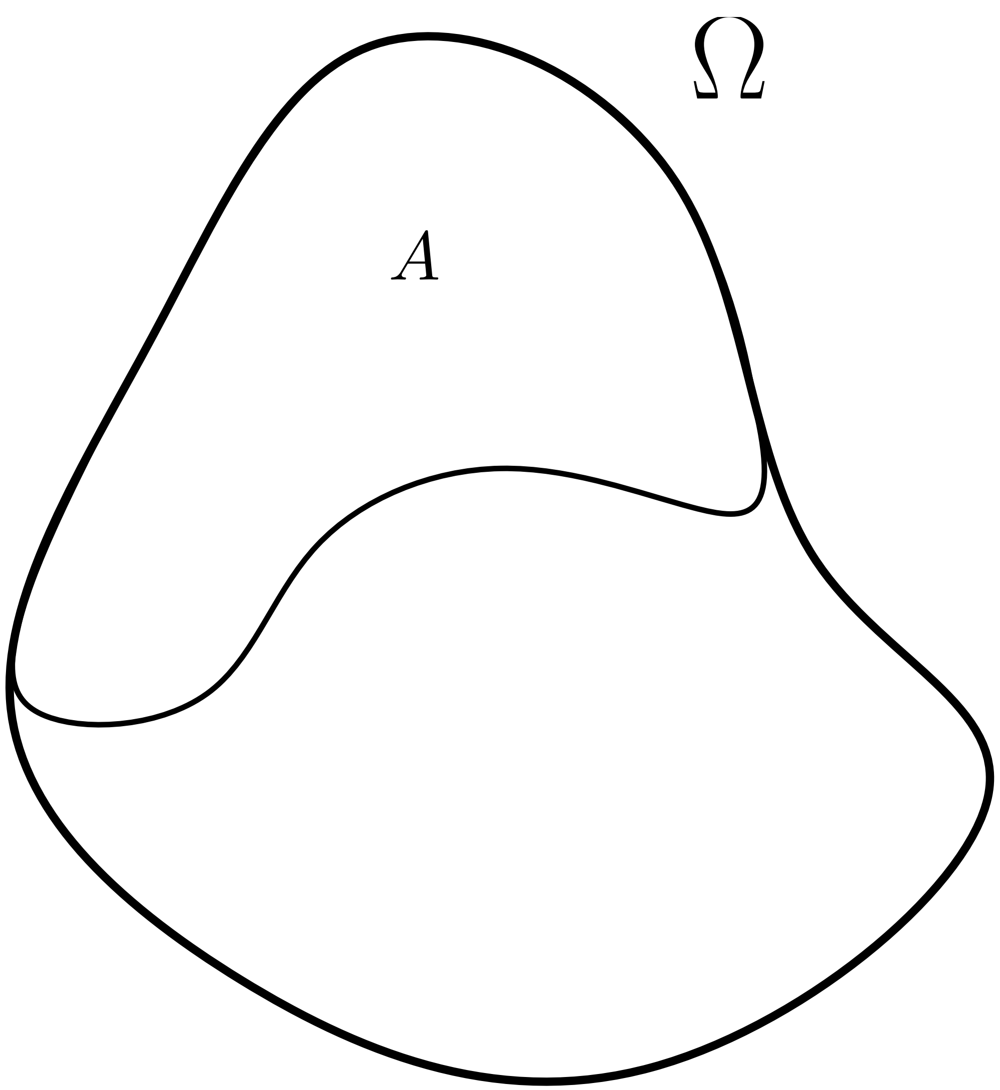
Formally, ...

- **Probability, P**
  - "Measures" each event in $\mathscr{F}$
  - Axioms of probability:
    - $\mathbf{P}(A) \geq 0$



Formally, ...

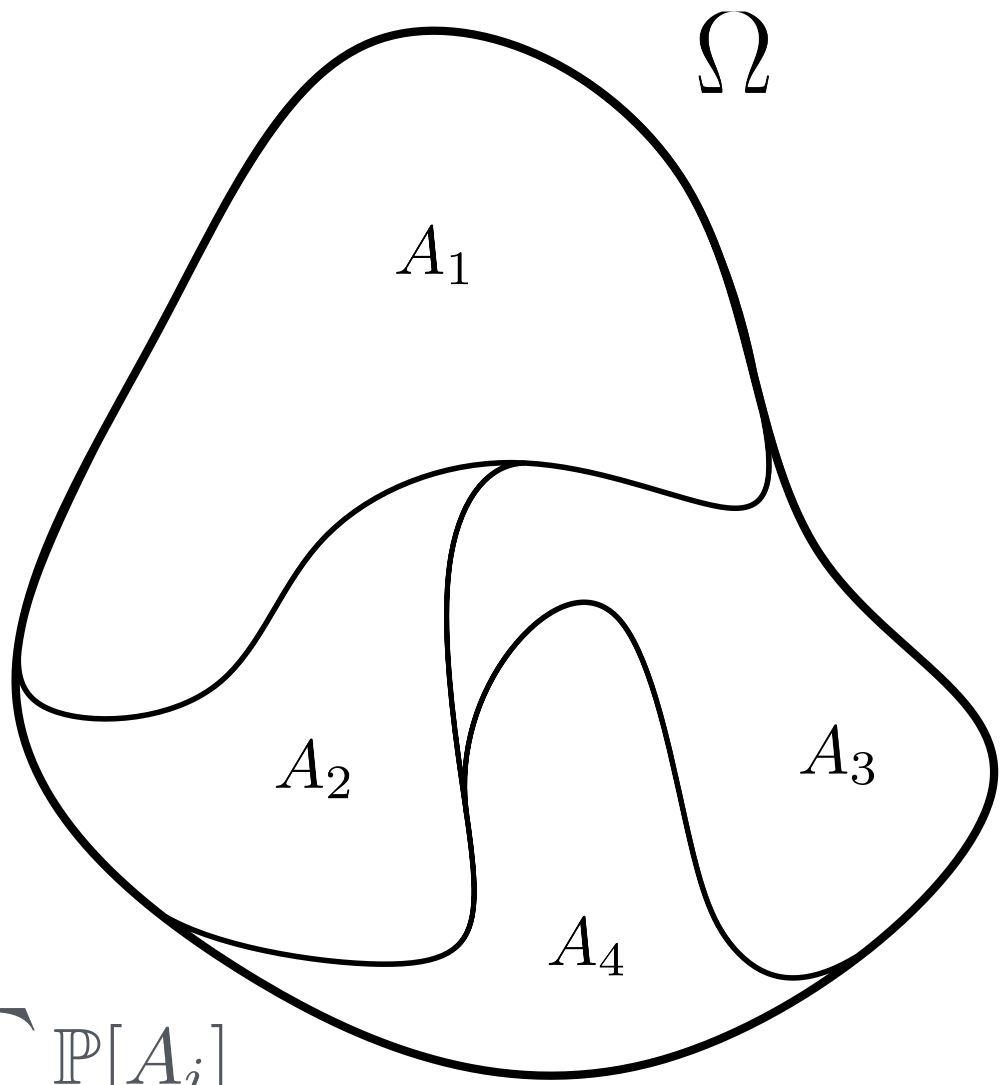
- **Probability, P**
  - "Measures" each event in $\mathscr{F}$
  - Axioms of probability:
    - $\mathbf{P}(A) \geq 0$
    - $\mathbf{P}(\Omega) = 1$



Formally, ...

- **Probability, P**
  - "Measures" each event in $\mathscr{F}$
  - Axioms of probability:
    - **P**(A) ≥ 0
    - **P**(Ω) = 1
    - Given disjoint events $A_1, \ldots, A_n$

$$\mathbb{P}[A_1 \cup \ldots \cup A_n] = \sum_{i=1}^{n} \mathbb{P}[A_i]$$



- **P**(A) ≥ 0
- **P**(Ω) = 1
- $\mathbb{P}[A_1 \cup \ldots \cup A_n] = \sum_{i=1}^{n} \mathbb{P}[A_i]$

Can you show:

- **P**(∅) = 0
- P(Ω) = 1 = P(Ω U ∅) = P(Ω) + P(∅).
- Then P(∅) = 1 - P(Ω) = 0.

Conditional probability

- Conditional probability of event *A* given *B*:

- Probability of *A* occurring, given that *B* occurred

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

… assuming **P**(B) > 0.

- Verifies all axioms of probability

Conditional probability

- **Important properties:**

- Independent events: $\mathbb{P}[A \mid B] = \mathbb{P}[A]$

- Law of total probability: given disjoint events $A_1, \ldots, A_n$, with

  $A_1 \cup \ldots \cup A_n = \Omega$,

$$\mathbb{P}[B] = \sum_{i=1}^{n} \mathbb{P}[B \mid A_i] \cdot \mathbb{P}[A_i]$$

TÉCNICO LISBOA

Conditional probability

- **Important properties:**

- Bayes theorem:

$$\mathbb{P}[A \mid B] = \frac{\overbrace{\boxed{\mathbb{P}[B \mid A]}}^{\text{likelihood}}\overbrace{\boxed{\mathbb{P}[A]}}^{\text{prior}}}{\underbrace{\boxed{\mathbb{P}[B]}}_{\text{evidence}}}$$

hypothesis observations

Random variable

- Given a probability space $(\Omega, \mathscr{F}, \mathbf{P})$,
  - Sometimes Ω is a cumbersome set to work with
  - It would be more convenient to instead work in some other space, more mathematically convenient (for example, $\mathbf{R}$)
  - A **random variable** X is a map $X : \Omega \longrightarrow E$, where E is some convenient space (usually $\mathbf{R}$)
  - We usually work with random variables, ignoring the underlying probability space

Random variable

- We write **P**(X = x) to represent

$$\mathbb{P}[X = x] = \mathbb{P}[\{\omega \in \Omega : X(\omega) = x\}]$$

- If the r.v. takes values in a discrete set, we call it a *discrete random variable*
- If the r.v. takes values in a continuous set, we call it a *continuous random variable*

Expectation

- Given a r.v. X and a function $f : X \longrightarrow \mathbf{R}$, the *expectation* of f is the quantity

$$\mathbb{E}[f(X)] = \sum_x f(x)\mathbb{P}[X = x]$$

- We can have conditional expectations:

$$\mathbb{E}[f(X) \mid Y = y] = \sum_x f(x)\mathbb{P}[X = x \mid Y = y]$$

Properties

- Law of total probability with expectations:

$$\mathbb{E}[f(X)] = \sum_y \mathbb{E}[f(X) \mid Y = y]\mathbb{P}[Y = y]$$

- Conditioning "eliminates" expectation:

$$\mathbb{E}[f(X) \mid X = x] = f(x)$$

Some useful results

- **Law of large numbers:** Given i.i.d. real-valued r.v.s $X_1, \ldots, X_n$, such that

$$\mathbb{E}[X_1] = \mathbb{E}[X_2] = \ldots = \mathbb{E}[X_N] = \mu$$

then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} X_i \stackrel{a.s.}{=} \mu$$

Some useful results

- **Jensen's inequality:** Given a real-valued r.v. X and a convex function f, then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$$