Probability Spaces and Events

For this course, we will only be concerned with discrete probabilities. This section formalizes some notions you should already be familiar with: probability spaces, events and probability distributions.

Definition: Probability space

A (discrete) probability space $(\Omega, \mathcal{F}, P)$ consists of a discrete, non-empty *sample space* Ω, an *event space* $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ (where $\mathcal{P}(\Omega)$ is the powerset of Ω) and a *probability measure* P which is a function $P : \Omega \to \mathbb{R}_{\geq 0}$ that satisfies

$$\sum_{\omega \in \Omega} P(\omega) = 1.$$

The event space $\mathcal{F}$ is required to be non-empty and closed under intersection, union and complements. For convenience, we will most often assume that $\mathcal{F}$ equals the powerset $\mathcal{P}(\Omega)$ of Ω, i.e., it contains all possible subsets of events, and therefore fulfils the required properties.

Definition: Event

An event $\mathcal{A}$ is an element of the event space $\mathcal{F} \subseteq \mathcal{P}(\Omega)$, i.e., a subset $\mathcal{A}$ of the sample space Ω. Its probability is defined as

$$P[\mathcal{A}] := \sum_{\omega \in \mathcal{A}} P(\omega),$$

where by convention $P[\emptyset] = 0$.

As a notational convention, we write $P[\mathcal{A}, \mathcal{B}]$ for $P[\mathcal{A} \cap \mathcal{B}]$, and $P[\overline{\mathcal{A}}]$ for $P[\Omega \setminus \mathcal{A}]$. The following identities hold for arbitrary events $\mathcal{A}, \mathcal{B} \subseteq \Omega$ (try to prove them for yourself):

- $P[\overline{\mathcal{A}}] = 1 - P[\mathcal{A}]$
- $P[\mathcal{A} \cup \mathcal{B}] = P[\mathcal{A}] + P[\mathcal{B}] - P[\mathcal{A}, \mathcal{B}]$
- $P[\mathcal{A}] = P[\mathcal{A}, \mathcal{B}] + P[\mathcal{A}, \overline{\mathcal{B}}]$.

It is often useful to consider the probability of an event *given* that some other event happened:

Definition: Conditional probability

For events $\mathcal{A}$ and $\mathcal{B}$ with $P[\mathcal{A}] > 0$, the conditional probability of $\mathcal{B}$ given $\mathcal{A}$ is defined as

$$P[\mathcal{B}|\mathcal{A}] := \frac{P[\mathcal{A},\mathcal{B}]}{P[\mathcal{A}]}\ .$$

Example: Fair die

We throw a six-sided fair die once, and consider the number that comes up. The sample space for this experiment is $\Omega = 1, 2, 3, 4, 5, 6$, with event space $\mathcal{F} = \mathcal{P}(\Omega)$ and probability measure $P[i] = \frac{1}{|\Omega|} = \frac{1}{6}$ for all $i \in \Omega$ (this is a **uniform** probability measure). Consider the events $\mathcal{A} = 2, 4, 6$ and $\mathcal{B} = 3, 6$. Using the formulas in the definitions of events and conditional probabilities, we can compute the following probabilities:

$$\begin{aligned}
P[\mathcal{A}] &= \frac{1}{2} (\text{the outcome is even}) \\
P[\mathcal{B}] &= \frac{1}{3} (\text{the outcome is a multiple of } 3) \\
P[\mathcal{A},\mathcal{B}] = P[6] &= \frac{1}{6} (\text{the roll is even and a multiple of } 3) \\
P[\mathcal{A}|\mathcal{B}] = \frac{1/6}{1/3} &= \frac{1}{2} (\text{the roll is even, given that it is a multiple of } 3) \\
P[\mathcal{B}|\mathcal{A}] = \frac{1/6}{1/2} &= \frac{1}{3} (\text{the roll is a multiple of } 3\text{, given that it is even})
\end{aligned}$$

This example shows that in general, $P[\mathcal{A}|\mathcal{B}]$ is *not necessarily equal* to $P[\mathcal{B}|\mathcal{A}]$. In fact, they are related through Bayes' rule.
decide whether this button is necessary, and whether the title of this block should be question instead of example