Lecture 1

The sample space, events, and outcomes

- Need a math model for describing "random" events that result from performing an "experiment."
- Ω denotes a "sample space," that is more appropriately also called an "outcome space." We think of the elements of Ω as "outcomes" of the experiment.
- $\mathcal{F}$ is a collection of subsets of Ω; elements of $\mathcal{F}$ are called "events." We wish to assign a "probability" $\mathrm{P}(A)$ to every $A \in \mathcal{F}$. When Ω is finite, $\mathcal{F}$ is always taken to be the collection of all subsets of Ω.

Example 1. Roll a six-sided die; what is the probability of rolling a six? First, write a sample space. Here is a natural one:

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

In this case, Ω is finite and we want $\mathcal{F}$ to be the collection of all subsets of Ω. That is,

$$\mathcal{F} = \Big\{ \varnothing\,, \Omega\,, \{1\}\,, \ldots, \{6\}\,, \{1, 2\}\,, \ldots, \{1, 6\}\,, \ldots, \{1, 2, \ldots, 6\} \Big\}.$$

Example 2. Toss two coins; what is the probability that we get two heads? A natural sample space is

$$\Omega = \Big\{ (H_1\,, H_2)\,, (H_1\,, T_2)\,, (T_1\,, H_2)\,, (T_1\,, T_2) \Big\}.$$

Once we have readied a sample space Ω and an event-space $\mathcal{F}$, we need to assign a probability to every event. This assignment cannot be made at whim; it has to satisfy some properties.

Rules of probability

Rule 1 (Nonnegative)**.** $0 \leq P(A) \leq 1$ for every event A.

Rule 2 (Total one)**.** $P(\Omega) = 1$. "Something will happen with probability one."

Rule 3 (Addition)**.** If A and B are disjoint events [i.e., $A \cap B = \varnothing$], then the probability that at least one of the two occurs is the sum of the individual probabilities. More precisely put,

$$P(A \cup B) = P(A) + P(B).$$

Lemma 1. *Choose and fix an integer $n \geq 1$. If $A_1, A_2, \ldots, A_n$ are disjoint events, then*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1) + \cdots + P(A_n).$$

Proof. The proof uses *mathematical induction.*

Claim. If the assertion is true for $n - 1$, then it is true for n.

The assertion is clearly true for $n = 1$, and it is true for $n = 2$ by Rule 3. Because it is true for $n = 2$, the Claim shows that the assertion holds for $n = 3$. Because it holds for $n = 3$, the Claim implies that it holds for $n = 4$, etc.

Proof of Claim. We can write $A_1 \cup \cdots \cup A_n$ as $A_1 \cup B$, where $B = A_2 \cup \cdots \cup A_n$. Evidently, A_1 and B are disjoint. Therefore, Rule 3 implies that $P(A) = P(A_1 \cup B) = P(A_1) + P(B)$. But B itself is a disjoint union of $n - 1$ events. Therefore $P(B) = P(A_2) + \cdots + P(A_n)$, thanks to the assumption of the Claim ["the induction hypothesis"]. This ends the proof. □

Rules 1–3 suffice if we want to study only finite sample spaces. But infinite samples spaces are also interesting. This happens, for example, if we want to write a model that answers, "what is the probability that we toss a coin 12 times before we toss heads?" This leads us to the next, and final, rule of probability.

Rule 4 (Extended addition rule)**.** If $A_1, A_2, \ldots$ are [countably-many] disjoint events, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

This rule will be extremely important to us soon. It looks as if we might be able to derive this as a consequence of Lemma 1, but that is not the case … it needs to be assumed as part of our model of probability theory.

Rules 1–4 have other consequences as well.