# Lecture 2

## 1. Properties of probability

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

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

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 1.3, 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.

**Example 2.1.** Recall that $A^c$, the complement of $A$, is the collection of all points in $\Omega$ that are not in $A$. Thus, $A$ and $A^c$ are disjoint. Because $\Omega = A \cup A^c$ is a disjoint union, Rules 2 and 3 together imply then that

$$\begin{aligned} 1 &= \mathrm{P}(\Omega) \\ &= \mathrm{P}(A \cup A^c) \\ &= \mathrm{P}(A) + \mathrm{P}(A^c). \end{aligned}$$

Thus, we obtain the physically–appealing statement that

$$\mathrm{P}(A) = 1 - \mathrm{P}(A^c).$$

For instance, this yields $\mathrm{P}(\varnothing) = 1 - \mathrm{P}(\Omega) = 0$. "Chances are zero that nothing happens."

**Example 2.2.** If $A \subseteq B$, then we can write $B$ as a disjoint union: $B = A \cup (B \cap A^c)$. Therefore, $\mathrm{P}(B) = \mathrm{P}(A) + \mathrm{P}(B \cap A^c)$. The latter probability is $\geq 0$ by Rule 1. Therefore, we reach another physically-appealing property:

$$\text{If } A \cup B, \text{ then } \mathrm{P}(A) \leq \mathrm{P}(B).$$

**Example 2.3.** Suppose $\Omega = \{\omega_1, \ldots, \omega_N\}$ has $N$ distinct elements ("$N$ distinct outcomes of the experiment"). One way of assigning probabilities to every subset of $\Omega$ is to just let

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{N},$$

where $|E|$ denotes the number of elements of $E$. Let us check that this probability assignment satisfies Rules 1–4. Rules 1 and 2 are easy to verify, and Rule 4 holds vacuously because $\Omega$ does not have infinitely-many disjoint subsets. It remains to verify Rule 3. If $A$ and $B$ are disjoint subsets of $\Omega$, then $|A \cup B| = |A| + |B|$. Rule 3 follows from this. In this example, each outcome $\omega_i$ has probability $1/N$. Thus, these are "equally likely outcomes."

**Example 2.4.** Let

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

There are four possible outcomes. Suppose that they are equally likely. Then, by Rule 3,

$$\begin{aligned}
\mathrm{P}(\{H_1\}) &= \mathrm{P}\Big(\{H_1\,, H_2\} \cup \{H_1\,, T_2\}\Big) \\
&= \mathrm{P}(\{H_1\,, H_2\}) + \mathrm{P}(\{H_1\,, T_2\}) \\
&= \frac{1}{4} + \frac{1}{4} \\
&= \frac{1}{2}.
\end{aligned}$$

In fact, in this model for equally-likely outcomes, $\mathrm{P}(\{H_1\}) = \mathrm{P}(\{H_2\}) = \mathrm{P}(\{T_1\}) = \mathrm{P}(\{T_2\}) = 1/2$. Thus, we are modeling two fair tosses of two fair coins.

**Example 2.5.** Let us continue with the sample space of the previous example, but assign probabilities differently. Here, we define $\mathrm{P}(\{H_1\,, H_2\}) = \mathrm{P}(\{T_1\,, T_2\}) = 1/2$ and $\mathrm{P}(\{H_1\,, T_2\}) = \mathrm{P}(\{T_1\,, H_2\}) = 1/2$. We compute, as we did before, to find that $\mathrm{P}(\{H_1\}) = \mathrm{P}(\{H_2\}) = \mathrm{P}(\{H_3\}) = \mathrm{P}(\{H_4\}) = 1/2$. But now the coins are not tossed fairly. In fact, the results of the two coin tosses are the same in this model.

The following generalizes Rule 3, because $\mathrm{P}(A \cap B) = 0$ when $A$ and $B$ are disjoint.

**Lemma 2.6** (Another addition rule)**.** *If $A$ and $B$ are events (not necessarily disjoint), then*

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).$$

**Proof.** We can write $A \cup B$ as a disjoint union of three events:

$$A \cup B = (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B).$$

By Rule 3,

$$\mathrm{P}(A \cup B) = \mathrm{P}(A \cap B^c) + \mathrm{P}(A^c \cap B) + \mathrm{P}(A \cap B). \tag{1}$$

Similarly, write $A = (A \cap B^c) \cup (A \cap B)$, as a disjoint union, to find that

$$\mathrm{P}(A) = \mathrm{P}(A \cap B^c) + \mathrm{P}(A \cap B). \tag{2}$$

There is a third identity that is proved the same way. Namely,

$$\mathrm{P}(B) = \mathrm{P}(A^c \cap B) + \mathrm{P}(A \cap B). \tag{3}$$

Add (2) and (3) and solve to find that

$$\mathrm{P}(A \cap B^c) + \mathrm{P}(A^c \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - 2\mathrm{P}(A \cap B).$$

Plug this in to the right-hand side of (1) to finish the proof. □

## 2. An example

Roll two fair dice fairly; all possible outcomes are equally likely.

**2.1.** A good sample space is

$$\Omega = \begin{Bmatrix} (1,1) & (1,2) & \cdots & (1.6) \\ \vdots & \vdots & \ddots & \vdots \\ (6,1) & (6,2) & \cdots & (6.6) \end{Bmatrix}$$

We have seen already that $\mathrm{P}(A) = |A|/|\Omega|$ for any event $A$. Therefore, the first question we address is, "how many items are in $\Omega$?" We can think of $\Omega$ as a 6-by-6 table; so $|\Omega| = 6 \times 6 = 36$, by second-grade arithmetic.

Before we proceed with our example, let us document this observation more abstractly.

**Proposition 2.7** (The first principle of counting)**.** *If we have $m$ distinct forks and $n$ distinct knives, then $mn$ distinct knife–fork combinations are possible.*

. . . not to be mistaken with . . .

**Proposition 2.8** (The second principle of counting)**.** *If we have $m$ distinct forks and $n$ distinct knives, then there are $m+n$ utensils.*

. . . back to our problem . . .

**2.2.** What is the probability that we roll doubles? Let

$$A = \{(1,1)\,,(2,2)\,,\ldots,(6,6)\}.$$

We are asking to find $\mathrm{P}(A) = |A|/36$. But there are 6 items in $A$; hence, $\mathrm{P}(A) = 6/36 = 1/6$.

**2.3.** What are the chances that we roll a total of five dots? Let

$$A = \{(1,4)\,,(2,3)\,,(3,2)\,,(4,1)\}.$$

We need to find $\mathrm{P}(A) = |A|/36 = 4/36 = 1/9$.

**2.4.** What is the probability that we roll somewhere between two and five dots (inclusive)? Let

$$A = \left\{ \overbrace{(1,1)}^{\text{sum }=\text{ 2}}\,,\underbrace{(1,2)\,,(2,1)}_{\text{sum }=3},\overbrace{(1,3)\,,(2,2)\,,(3,1)}^{\text{sum }=4},\underbrace{(1,4)\,,(4,1)\,,(2,3)\,,(3,2)}_{\text{sum}=5} \right\}.$$

We are asking to find $\mathrm{P}(A) = 10/36$.

**2.5.** What are the odds that the product of the number of dots thus rolls is an odd number? The event in question is

$$A := \begin{Bmatrix} (1,1), & (1,3), & (1,5) \\ (3,1), & (3,3), & (3,5) \\ (5,1), & (5,3), & (5,5) \end{Bmatrix}.$$

And $\mathrm{P}(A) = 9/36 = 1/4$.