2 Basic Probability

2.1 Chance Events

A **probability** is a number between 0 and 1 that describes how likely an event is to occur. When setting up an experiment, we first specify a set of outcomes, which we call the **sample space**, typically denoted with the symbol Ω. Collections of items in Ω are called **events**, and to these events, we associate probabilities. We can think of the probability of an event as the "chance" of it occurring.

Let's see how to set up sample spaces in some familiar settings.

2.1.1 Sample Spaces for Coin and Dice Experiments

Example 2.1. Imagine we are flipping a coin. The set of possible outcomes is

$$\Omega = \{\texttt{H}, \texttt{T}\},$$

i.e. we can flip either a "Heads" or a "Tails".

Now suppose we flip two coins. The set of possible outcomes becomes

$$\Omega = \{\texttt{HH}, \texttt{HT}, \texttt{TH}, \texttt{TT}\}.$$

Example 2.2. Now suppose we are rolling a die. The sample space of a die roll is similar to that of a coin flip, except we have more outcomes. Since a standard casino die has 6 sides, we can write the sample space as

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

An example of a **subset** A of Ω (we write $A \subset \Omega$) could be the set of even rolls, $\{2, 4, 6\}$.

2.1.2 Assigning Probabilities to Dice Rolls and Coin Flips

To each element of our sample space Ω, we associate a probability. We require the probabilities of all the outcomes in Ω to sum to 1. To find the probability of a subset of our sample space, we would add up the probabilities of the elements contained in that subset.

Example 2.3. Assume that the die we rolled above was a fair die. Then each of the outcomes is equally likely to occur, i.e.

$$P(\text{roll } 1) = P(\text{roll } 2) = P(\text{roll } 3) = P(\text{roll } 4) = P(\text{roll } 5) = P(\text{roll } 6) = \frac{1}{6}$$

To find the probability of rolling an even number, we add up all the probabilities of the even numbers in Ω. This gives us

$$\begin{aligned} P(\text{roll an even number}) &= P(\text{roll } 2) + P(\text{roll } 4) + P(\text{roll } 6) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= \frac{1}{2}. \end{aligned}$$

Example 2.4. Suppose that we flip a single fair coin. Then as stated above, the possible outcomes are H or T. Since flipping an H is equally as likely as flipping a T, we have

$$P(\mathtt{H}) = P(\mathtt{T}) = \frac{1}{2}.$$

Similarly, if we flip the coin twice, each of the possible outcomes should be equally likely, so

$$P(\mathtt{HH}) = P(\mathtt{HT}) = P(\mathtt{TH}) = P(\mathtt{TT}) = \frac{1}{4}.$$

Now assume that our coin has a bias, that is, the probability of flipping H is some number between 0 and 1, denoted p. Then since the probabilities of all the outcomes must sum to 1, we have

$$\begin{aligned} P(\mathtt{H}) &= p \\ P(\mathtt{T}) &= 1 - p. \end{aligned}$$

Similarly, the probabilities for elements in the second sample space become

$$\begin{aligned} P(\mathtt{HH}) &= p^2 \\ P(\mathtt{HT}) &= p(1-p) \\ P(\mathtt{TH}) &= (1-p)p \\ P(\mathtt{TT}) &= (1-p)^2, \end{aligned}$$

since if we flip a H with probability p, the probability of then flipping another H in sequence would be p^2. The probability of flipping a H and then a T is the product of their probabilities, $p \cdot (1-p)$. The last two probabilities above are obtained similarly.

We check that these probabilities sum to 1 below.

$$\begin{aligned} P(\mathtt{HH}) + P(\mathtt{HT}) + P(\mathtt{TH}) + P(\mathtt{TT}) &= p^2 + p \cdot (1-p) + (1-p) \cdot p + (1-p)^2 \\ &= p^2 + (p - p^2) + (p - p^2) + (1 - 2p + p^2) \\ &= 2p - p^2 + (1 - 2p + p^2) \\ &= 1. \end{aligned}$$

Exercise 2.5. What is the probability that we get at least one $\mathtt{H}$?

Solution. One way to solve this problem is to add up the probabilities of all outcomes that have at least one $\mathtt{H}$. We would get

$$\begin{aligned} P(\text{flip at least one } \mathtt{H}) &= P(\mathtt{HH}) + P(\mathtt{HT}) + P(\mathtt{TH}) \\ &= p^2 + p \cdot (1-p) + (1-p) \cdot p \\ &= p^2 + 2 \cdot (p - p^2) \\ &= 2p - p^2 \\ &= p \cdot (2-p). \end{aligned}$$

Another way to do this is to find the probability that we **don't** flip at least one $\mathtt{H}$, and subtract that probability from 1. This would give us the probability that we **do** flip at least one $\mathtt{H}$.

The only outcome in which we don't flip at least one $\mathtt{H}$ is if we flip $\mathtt{T}$ both times. We would then compute

$$P(\text{don't flip at least one } \mathtt{H}) = P(\mathtt{TT}) = (1-p)^2.$$

Then to get the **complement** of this event, i.e. the event where we **do** flip at least one $\mathtt{H}$, we subtract the above probability from 1. This gives us

$$\begin{aligned} P(\text{flip at least one } \mathtt{H}) &= 1 - P(\text{don't flip at least one } \mathtt{H}) \\ &= 1 - (1-p)^2 \\ &= 1 - (1 - 2p + p^2) \\ &= 2p - p^2 \\ &= p \cdot (2-p). \end{aligned}$$

Incredible! Both methods for solving this problem gave the same answer. Notice that in the second approach, we didn't have to sum up any probabilities to get the answer. It can often be the case that computing the probability of the complement of an event and subtracting that from 1 to find the probability of the original event requires less work than computing the probability directly. □

2.1.3 Independence

If two events A and B don't influence or give any information about the other, we say A and B are independent. Remember that this is not the same as saying A and B are disjoint. If A and B were disjoint, then given information that A happened, we would know with certainty that B did *not* happen. Hence if A and B are disjoint they could never be independent. The mathematical statement of independent events is given below.

Definition 2.6. *Let A and B both be subsets of our sample space Ω. Then we say A and B are independent if*

$$P(A \cap B) = P(A)P(B)$$

In other words, if the probability of the intersection factors into the product of the probabilities of the individual events, they are independent.

We haven't defined set intersection in this section, but it is defined in the set theory chapter. The $\cap$ symbol represents A *and* B happening, i.e. the intersection of the events.

Example 2.7. Returning to our double coin flip example, our sample space was

$$\Omega = \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}.$$

Define the events

$$\begin{aligned} A &\doteq \{\text{first flip is heads}\} = \{\mathtt{HH}, \mathtt{HT}\} \\ B &\doteq \{\text{second flip is tails}\} = \{\mathtt{HT}, \mathtt{TT}\} \end{aligned}$$

Notation: We write the sign $\doteq$ to represent that we are defining something. In the above expression, we are defining the arbitrary symbols A and B to represent events.

Intuitively, we suspect that A and B are independent events, since the first flip has no effect on the outcome of the second flip. This intuition aligns with the definition given above, as

$$P(A \cap B) = P(\{\mathtt{HT}\}) = \frac{1}{4}$$

and

$$P(A) = P(B) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

We can verify that

$$P(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B).$$

Hence A and B are independent. This may have seemed like a silly exercise, but we will often encounter pairs of sets where it is not intuitively clear whether or not they are independent. In these cases, we can simply verify this mathematical definition to conclude independence.