Probability

Definition. *The* ***sample space*** Ω *for a random scenario is the set of all possible outcomes. An* ***event*** *is a subset of* Ω. *The* ***probability measure*** *is a function*

$$\mathbb{P} : \Big\{ \textit{all events} \Big\} \to [0,1]$$

obeying three axioms:

(i) $\mathbb{P}(A) \geq 0 \ \ \forall A \subset \Omega$.
(ii) $\mathbb{P}(\Omega) = 1$.
(iii) *If* $A_1 \cap A_2 = \emptyset$ *(in which case we say* A_1 *and* A_2 *are* ***disjoint*** *or* ***mutually exclusive****), then we have that*

$$\mathbb{P}\Big(A_1 \cup A_2\Big) = \mathbb{P}(A_1) + \mathbb{P}(A_2).$$

The first axiom is essentially already satisfied since we have defined $\mathbb{P}$ to always take values in $[0,1]$. Nonetheless, you should understand that this implies probabilities can never be negative. The second axiom says that the sample space contains all possible outcomes, in other words, the probability that one of the outcomes in the sample space is realized is 1. The third axiom says that to find the probability that disjoint events occur we can simply add the probabilities of the individual events.

In a random scenario in which all outcomes are equally likely we have that:

$$\text{probability of an event} = \frac{\text{number of outcomes in an event}}{\text{total number of outcomes in the sample space}}$$

This only holds when the outcomes are equally likely.

Proposition.

(i) $\mathbb{P}(\emptyset) = 0$.
(ii) $\mathbb{P}(A') = 1 - \mathbb{P}(A)$.
(iii) *If* $A \subset B$, *then* $\mathbb{P}(A) \leq \mathbb{P}(B)$.
(iv) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

In the above proposition A' is the **complement** of A. It is the unique subset of Ω containing precisely those elements not in A.

Two events A, B are said to be **independent** if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Suppose $B \subset \Omega$ is such that $\mathbb{P}(B) > 0$, then we can talk about the probability of an event $A \subset \Omega$ occurring given that we know B has occurred. This is denoted $\mathbb{P}(A|B)$, and is called the **conditional probability** of A given that B has occurred. We define

$$\mathbb{P}\big(A|B\big) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Theorem (Law of Total Probability). *Suppose* $A_1, A_2 \subset \Omega$ *partition our sample space. That is,* $A_1 \cap A_2 = \emptyset$, *and* $A_1 \cup A_2 = \Omega$. *Then,*

$$\mathbb{P}(B) = \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\mathbb{P}(A_2).$$

Theorem (Bayes' Rule). *Suppose* $A, B \subset \Omega$ *where* $\mathbb{P}(A) > 0$ *and* $\mathbb{P}(B) > 0$. *Then,*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}.$$

Suppose that we conduct n identical independent trials in each of which only two possible outcomes are possible: success or failure. If the probability of success on any individual trial is fixed at p (so the probability of failure is $1-p$), then the probability of having precisely k successes (in n trials) is given by

$$\binom{n}{k} p^k (1-p)^{n-k}.$$

Such a random scenario is said to have a **binomial distribution.** The Binomial Theorem guarantees that these probabilities sum to 1.