

# Chapter 1

## Introduction to Probability

*Probability is the branch of mathematics that studies the possible outcomes of given events together with the outcomes' relative likelihoods and distributions. [...] The analysis of events governed by probability is called statistics. - Wolfram*

Before the 16th century, predicting the outcome of a future event with any degree of accuracy was thought to be impossible. Measuring the likelihood of something occurring was the first approach to probability. In 1620 Galileo published *Sopra le Scoperte dei Dadi* (On the Outcomes of Dice), calculating the chances of certain totals when throwing dice.

Deducing the probability of outcomes is the study of **probability theory**. When we do not know the probabilistic mechanism governing the process, and we must estimate parameters from data, this is **statistics**.

Example:

**Probability:** Previous studies showed that the drug was 80% effective. Then we can anticipate that for a study on 100 patients, in average 80 will be cured and at least 65 will be cured with 99.99% chances.

**Statistics:** Observe that 78/100 patients were cured. We (will be able to) conclude that we are 95% confident that for other studies the drug will be effective on between 69.88% and 86.11% of patients.

Notice how in the probability case, we are given the likelihoods of something happening. While in the case of statistics we take a population sample and use this to infer something about our data.

## 1.1 Definition of Probability

If  $n$  is the number of trials of an experiment (such as the number of flips of a coin), one might define the probability of an event  $E$  (such as the outcome we get a heads) by

$$P(E) = \lim_{n \rightarrow \infty} \frac{\text{Number of times } E \text{ occurs}}{n} \quad (1.1)$$

However this is not an acceptable mathematical definition of probability.

Even though probability distributions existed, a formal definition for probability lacked until 1933, when Andrey N. Kolmogorov set three axioms of probability in *Foundations of the Calculus of Probabilities*. Which was defined as the following:

Suppose that a random experiment has associated with it a sample space  $S$ . A probability is a numerically valued function that assigns a number  $P(A)$  to every event  $A$  so that the following axioms hold:

1.  $P(A) \geq 0$
2.  $P(S) = 1$
3. if  $A_1, A_2, \dots$  is a sequence of mutually exclusive events (that is, a sequence in which  $A_i A_j = \phi$  for any  $i \neq j$ ), then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad (1.2)$$

The first axiom states that a probability of any event occurring must be greater or equal to 0 (i.e. non-negative). By the second axiom we also have that the probability of any event occurring cannot be greater than 1, in fact, the probability that at least some event will occur in the sample space  $S$  is 1. And finally the third axiom states that for mutually exclusive events  $A$  and  $B$  then the probability that both occur is  $P(A \cup B) = P(A) + P(B)$ .

## 1.2 Conditional Probability

Sometimes partial information about an event can help us when calculating probabilities. For instance, if a family has two children but we are told at least one is a girl, what is the probability that the family has two girls? Using event  $A$  to denote the event we have two girls and event  $B$  to denote that at least one is a girl. We have the following:

$$\begin{aligned}
P(\text{two girls given we have at least one girl}) &= \\
&= P(A \text{ given } B) \quad (1.3) \\
&= P(A \mid B)
\end{aligned}$$

The definition for conditional probability is: If  $A$  and  $B$  are any two events, then the conditional probability of  $A$  given  $B$  denoted by  $P(A|B)$  is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad (1.4)$$

Note that  $\mid$  is used to denote the word "given". And  $\cap$  is the set intersection, where both events  $A$  and  $B$  occur.

Therefore in our example of two girls given we have at least one girl in a family with two children, the probability is  $1/3$ . It may help to visualise the set of possible outcomes  $S$ .

$$S = \{BB, BG, GB, GG\} \quad (1.5)$$

Then we can clearly see that,  $P(A \mid B) = \frac{1/4}{3/4} = \frac{1}{3}$

### 1.2.1 Independence

If an event occurring does not change the probability of another event, then such events are said to be independent. Two events  $A$  and  $B$  are said to be independent if

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

which is the equivalent of saying

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

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

### 1.2.2 Bayes' Rule

Given our formula for conditional probability defined above, we can rearrange for  $P(A \cap B)$  and use substitution as follows:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \mid A)P(A)}{P(B)}$$