Chapter 1

Axiomatic Probability Theory

1.1 Axioms of probability

In the previous chapter, we have introduced sample spaces. In a mathematically concrete sense, we are now going to measure the probability of events.

Definition 1 *A measure is a function $\mu : \mathcal{S} \to \mathbb{R} : S \mapsto \mu(S)$ that maps elements from a set of sets $\mathcal{S}$ (formally a σ-algebra) to real numbers. Such a measure has the following properties:*

1. $\mu(S) \geq 0$ *for all* $S \in \mathcal{S}$
2. $\mu\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} \mu\left(S_i\right)$

The second property ensures that we can compute the measure of the union of two or more sets by first computing the measure of each set individually and then simply adding them up. We have already seen an example of a measure, namely the counting measure $|\cdot|$ that counts the elements in a set.

There is one measure, however, that is going to be the star of the rest of this script, namely the **probability measure**.

Definition 2 *A probability measure $\mathbb{P} : \mathcal{A} \to \mathbb{R} : A \mapsto \mathbb{P}(A)$ from an event space $\mathcal{A}$ associated with a sample space Ω to the real numbers has the following properties.*

1. $\mathbb{P}(A) \geq 0$ *for all* $A \in \mathcal{A}$

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

3. $\mathbb{P}(\Omega) = 1$

Notice that we only added item 3 to the general definition of a measure. This has some interesting consequences. For one, since $\mathbb{P}(\Omega) = 1$ we know that no event can have probability greater than 1. In fact, we even know that all events that are a proper subsets of Ω must have probability strictly less than 1. Moreover, we have

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) \tag{1.1}$$

which implies that $\mathbb{P}(\emptyset) = 0$ always holds. We can therefore safely exclude the empty set from our event spaces.

We have already discussed uniform probability in the previous chapter. Now we can formally explain what we meant by that. The uniform probability measure is the measure $\mathbb{P}$ such that $\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}$ for all $\omega \in \Omega$. This is where the distinction between sample and event spaces becomes important. We cannot measure the elements of a sample space, only the elements of an event space! Recall our convention that we will always assume that $\mathcal{A} = \mathcal{P}(\Omega)$ which obviously contains a singleton for each element in Ω. Thus the uniform probability measure is indeed well-defined. Thus, whenever we talk about *uniform probability* we either mean the uniform probability measure or, more often, the real value $\frac{1}{|\Omega|}$ to which this measure uniformly evaluates.