B Mathematical Prerequisites

The beginner... should not be discouraged if... he finds that he does not have the prerequisites for reading the prerequisites.

P. Halmos[1]

In this chapter, we introduce mathematical results that might not be known to all readers, but which are sufficiently standard that they not be put into the actual chapters.

This exposition is almost certainly incomplete, and some readers will inevitably happen upon terms in the book that are unknown to them, yet not explained here. Consequently, we also give some further references.

B.1 Probability

B.1.1 Probability Spaces

Let us start with some basic notions of probability theory. For further detail, we refer to [77, 165, 561]. We do not try to be rigorous; instead, we endeavor to give some intuition and explain how these concepts are related to our present interests.

Domain

Assume we are given a nonempty set $\mathcal{X}$, called the *domain* or *universe*. We refer to the elements x of $\mathcal{X}$ as *patterns*. The patterns are generated by a stochastic source. For instance, they could be handwritten digits, which are subject to fluctuations in their generation best modelled probabilistically. In the terms of probability theory, each pattern x is considered the outcome of a *random experiment*.

We would next like to assign probabilities to the patterns. We naively think of a probability as being the limiting frequency of a pattern; in other words, how often, relative to the number of trials, a certain pattern x comes up in a random experiment, if we repeat this experiment infinitely often?

Event

Probability

It turns out to be convenient to be slightly more general, and to talk about the probability of *sets* of possible outcomes; that is, subsets C of $\mathcal{X}$ called *events*. We denote the *probability* that the outcome of the experiment lies in C by

1. Quoted after [429].