



In this chapter we discuss some specific examples of probability distributions and their properties. As well as being of interest in their own right, these distributions can form building blocks for more complex models and will be used extensively throughout the book.

One role for the distributions discussed in this chapter is to model the probability distribution $p(\mathbf{x})$ of a random variable \mathbf{x} , given a finite set $\mathbf{x}_1, \dots, \mathbf{x}_N$ of observations. This problem is known as *density estimation*. It should be emphasized that the problem of density estimation is fundamentally ill-posed, because there are infinitely many probability distributions that could have given rise to the observed finite data set. Indeed, any distribution $p(\mathbf{x})$ that is non-zero at each of the data points $\mathbf{x}_1, \dots, \mathbf{x}_N$ is a potential candidate. The issue of choosing an appropriate distribution relates to the problem of model selection, which has already been encountered in the context of polynomial curve fitting and which is a central issue in machine

Section 1.2