11.1 Introduction

In introductory statistics courses, one has to know why the (univariate) normal distribution is important—especially that the random variables that occur in many situations are approximately normally distributed and that it arises in theoretical work as an approximation to the distribution of many statistics, such as averages of independent random variables. More or less, the same reasons apply to the bivariate normal distribution. “But the prime stimulus has undoubtedly arisen from the strange tractability of the normal model: a facility of manipulation which is absent when we consider almost any other multivariate data-generating mechanism.”—Barnett (1979). We may also note the following views expressed by different authors:

- “In multivariate analysis, the only distribution leading to tractable inference is the multivariate normal”—Mardia (1985).
- “The only type of bivariate distribution with which most of us feel familiar (other than the joint distribution of a pair of independent random variables) is the bivariate normal distribution”—Anscombe (1981, p. 305).
- “But who has ever seen a multivariate normal sample?” asks Barnett (1979) rhetorically, and then goes on to present, without any conscious bias in their selection, three bivariate datasets from the published literature that all turn out to be grossly non-normal.
- “The only sure defense against a successful disproof of the assumption of multivariate normality is to abstain from collecting, or presenting, too much data!”—wording adapted from Burnaby (1966, p. 109).

The origins of the bivariate normal are found in the first half of the nineteenth century in the work of Laplace, Plana, Gauss, and Bravais. Seal (1967) and Lancaster (1972) have given accounts of these developments. The latter pointed out that the early authors derived the bivariate normal as the joint distribution of the linear forms of independently distributed normal variables but did not define a coefficient of correlation; the distribution was used as a
basis for a theory of measurement error. Francis Galton (b.1822, d.1911), in analyzing the measurements of the heights of parents and their adult children, studied the structure of a bivariate normal density function. He observed that the marginal distributions of the data were normal and the contours of equal frequency were ellipses. He was the first to recognize the need for a measure of correlation in bivariate data. Since his time, the growth in the use of the bivariate normal has been enormous, so as to produce the comments already quoted.

In Section 11.2, we present some basic formulas and properties of the bivariate normal distribution. In Section 11.3, different methods of deriving bivariate normal distributions are mentioned. Some well-known characterizations of the bivariate normal distributions are listed in Section 11.4. Distributions, moments, and other properties of order statistics arising from a bivariate normal distribution are discussed in Section 11.5. While some available illustrations of the bivariate normal are described in Section 11.6, relationships to some other distributions are mentioned in Section 11.7. Next, the estimation of parameters of the bivariate normal distribution is discussed in Section 11.8. In Section 11.9, some other interesting properties of the bivariate normal distribution are briefly mentioned. Some specialized fields in which the bivariate normal model is applied in interesting ways are listed in Section 11.10, while common applications of the bivariate normal distribution are mentioned in Section 11.11. In Section 11.12, different computational methods and algorithms that are available for computing the bivariate normal distribution function are discussed. Many different test procedures and graphical methods are available for assessing the validity of the bivariate normal distribution, and these are detailed in Section 11.13. Distributions with normal conditionals and bivariate skew-normal distributions are described in Sections 11.14 and 11.15, respectively. Some univariate transformations on a bivariate normal random vector and the resulting distributions are described in Section 11.16. In Section 11.17, the truncated bivariate normal distribution and its properties are presented. The bivariate normal mixture distributions and related issues are described in Section 11.18. In Section 11.19, some bivariate non-normal distributions with normal marginals are presented. Finally, in Section 11.21, the bivariate inverse Gaussian distribution and its properties are discussed.

For further information, interested readers may refer to Johnson and Kotz (1972, Chapter 36), Kotz et al. (2000, Chapter 46), Kendall and Stuart (1977, Chapter 15; 1979, Chapters 18 and 26), and Patel and Read (1982, Chapter 10).