8.9 Concluding Remarks

Certain key ideas first introduced in the context of survey sampling in Chapter 7 have recurred in this chapter. We have viewed an estimate as a random variable having a probability distribution called its sampling distribution. In Chapter 7, the estimate was of a parameter, such as the mean, of a finite population; in this chapter, the estimate was of a parameter of a probability distribution. In both cases, characteristics of the sampling distribution, such as the bias and the variance and the large sample approximate form, have been of interest. In both chapters, we studied confidence intervals for the true value of the unknown parameter. The method of propagation of error, or linearization, has been a useful tool in both chapters. These key ideas will be important in other contexts in later chapters as well.

Important concepts and techniques in estimation theory were introduced in this chapter. We discussed two general methods of estimation—the method of moments and the method of maximum likelihood. The latter especially has great general utility in statistics. We developed and applied some approximate distribution theory for maximum likelihood estimates. Other theoretical developments included the concept of efficiency, the Cramér-Rao lower bound, and the concept of sufficiency and some of its consequences.

Bayesian inference was introduced in this chapter. The point of view contrasts rather sharply with that of frequentist inference in that the Bayesian formalism allows uncertainty statements about parameter values to be probabilistic, for example,"After seeing the data, the probability is 95% that $1.8 \leq \theta \leq 6.3$." In frequentist inference, θ is not a random variable, and a statement like this would literally make no sense; it would be replaced by, "A 95% confidence interval for θ is [1.8, 6.3]," perhaps followed by a long convoluted explication of the meaning of a confidence interval. Despite this apparently sharp philosophical difference, Bayesian and frequentist procedures have a great deal in common and typically lead to similar conclusions. Despite the distinction between the two statements above, the statements may well mean essentially the same thing operationally to a practitioner who has analyzed the data. The likelihood function is fundamental for both frequentist and Bayesian inference. In an application, the choice of a model, that is, the choice of a likelihood function, will typically be much more important than whether on subsequently multiplies it by a prior or just maximizes it. This is especially true if flat priors are used; in fact, one might regard a flat prior as a device that allows the likelihood to be treated as a probability density.

In this chapter, we introduced the bootstrap method for assessing the variability of an estimate. Such uses of simulation have become increasingly widespread as computers have become faster and cheaper; the bootstrap as a general method has been developed only quite recently and has rapidly become one of the most important statistical tools. We will see other situations in which the bootstrap is useful in later chapters. Efron and Tibshirani (1993) give an excellent introduction to the theory and applications of the bootstrap.

The context in which we have introduced the bootstrap is often referred to as the **parametric bootstrap.** The nonparametric bootstrap will be introduced in Chapter 10. The parametric bootstrap can be thought about somewhat abstractly in the following