Preface

The estimation of probabilities is of practical and philosophical interest, and can be difficult when the sample is small, difficult enough in fact so that the literature is not as extensive as one might expect, although some of it is ancient. This monograph contains a review of much of the immediately relevant literature known to me, but the main emphasis is on methods that are new or have not been written up in a connected manner.

The difficulties become clear when it is realized that we estimate probabilities every minute of the day, at least implicitly, and that how we do this is unknown. When this problem is solved, a potential pathway to artificial intelligence will be cleared, apart from easier applications, such as to character recognition and medical diagnosis. The work described in this monograph is only a fraction of all that is required for these purposes. A more complete treatment would involve much discussion of problems of classification. For this purpose, another monograph will be necessary.

We shall be concerned with methods for the estimation of probabilities from "effectively small" samples, and with some implications of these methods for tests of significance for multinomial samples and for contingency tables. Most of the techniques described depend on a modern Bayesian approach. The method mentioned in Chapter 8, for multinomial sampling when the number of categories is large, as in the sampling of species or of vocabulary, originated in a suggestion made by A. M. Turing in 1941, but was not published until 1953. Herbert

Robbins pointed out, when I described the method in a recent lecture, that it is an example of what he now calls the "empirical Bayes method." The earlier chapters make much use of a hierarchy of three types of probabilities, usually to be roughly interpreted as physical, logical, and subjective, although I have called them Types I, II, and III in order to decrease controversy. (They *could* all be physical and they *could* all be subjective.) In my opinion, this is a necessary improvement of the Bayesian methods that are usually used. In Chapter 9 the emphasis shifts to the maximum-entropy principle and multidimensional contingency tables.

Fuller abstracts are given in the Introduction (Chapter 1) and Summary (Chapter 10).

It is a pleasure to acknowledge the facilities provided by the Communications Research Division of the Institute for Defense Analyses, and especially the expert mathematical typing of Mrs. Euthie Anthony, the use of the electronic computer, and the help of Mr. G. J. Mitchell in extracting bugs from my FORTRAN programs. I am also much indebted to Dr. Oscar Rothaus and Dr. Lloyd Welch for literally invaluable help in connection with Appendices D and E, respectively.

Princeton, New Jersey IRVING JOHN GOOD
April, 1964