A Generalization of Bayesian Inference*

Arthur P. Dempster

Abstract. Procedures of statistical inference are described which generalize Bayesian inference in specific ways. Probability is used in such a way that in general only bounds may be placed on the probabilities of given events, and probability systems of this kind are suggested both for sample information and for prior information. These systems are then combined using a specified rule. Illustrations are given for inferences about trinomial probabilities, and for inferences about a monotone sequence of binomial $p_i$. Finally, some comments are made on the general class of models which produce upper and lower probabilities, and on the specific models which underlie the suggested inference procedures.

1 Introduction

Reduced to its mathematical essentials, Bayesian inference means starting with a global probability distribution for all relevant variables, observing the values of some of these variables, and quoting the conditional distribution of the remaining variables given the observations. In the generalization of this paper, something less than a global probability distribution is required, while the basic device of conditioning on observed data is retained. Actually, the generalization is more specific. The term Bayesian commonly implies a global probability law given in two parts, first the marginal distribution of a set of parameters, and second a family of conditional distributions of a set of observable variables given potential sets of parameter values. The first part, or prior distribution, summarizes a set of beliefs or state of knowledge in hand before any observations are taken. The second part, or likelihood function, characterizes the information carried by the observations. Specific generalizations are suggested in this paper for both parts of the common Bayesian model, and also for the method of combining the two parts. The

* Read at a Research Methods Meeting of the Society, February 14th, 1968, Professor R.L. Plackett in the Chair.
components of these generalizations are built up gradually in Sect. 2 where they are illustrated on a model for trinomial sampling.

Inferences will be expressed as probabilities of events defined by unknown values, usually unknown parameter values, but sometimes the values of observables not yet observed. It is not possible here to go far into the much-embroiled questions of whether probabilities are or are not objective, are or are not degrees of belief, are or are not frequencies, and so on. But a few remarks may help to set the stage. I feel that the proponents of different specific views of probability generally share more attitudes rooted in the common sense of the subject than they outwardly profess, and that careful analysis renders many of the basic ideas more complementary than contradictory. Definitions in terms of frequencies or equally likely cases do illustrate clearly how reasonably objective probabilities arise in practice, but they fail in themselves to say what probabilities mean or to explain the pervasiveness of the concept of probability in human affairs. Another class of definitions stresses concepts like degree of confidence or degree of belief or degree of knowledge, sometimes in relation to betting rules and sometimes not. These convey the flavour and motivation of the science of probability, but they tend to hide the realities which make it both possible and important for cognizant people to agree when assigning probabilities to uncertain outcomes. The possibility of agreement arises basically from common perceptions of symmetries, such as symmetries among cases counted to provide frequencies, or symmetries which underlie assumptions of exchangeability or of equally likely cases. The importance of agreement may be illustrated by the statistician who expresses his inferences about an unknown parameter value in terms of a set of betting odds. If this statistician accepts any bet proposed at his stated odds, and if he wagers with colleagues who consistently have more information, perhaps in the form of larger samples, then he is sure to suffer disaster in the long run. The moral is that probabilities can scarcely be “fair” for business deals unless both parties have approximately the same probability assessments, presumably based on similar knowledge or information. Likewise, probability inferences can contribute little to public science unless they are as objective as the web of generally accepted fact on which they are based. While knowledge may certainly be personal, the communication of knowledge is one of the most fundamental of human endeavours. Statistical inference can be viewed as the science whose formulations make it possible to communicate partial knowledge in the form of probabilities.

Generalized Bayesian inference seeks to permit improvement on classical Bayesian inference through a complex trade-off of advantages and disadvantages. On the credit side, the requirement of a global probability law is dropped and it becomes possible to work with only those probability assumptions which are based on readily apparent symmetry conditions and are therefore reasonably objective. For example, in a wide class of sampling models, including the trinomial sampling model analysed in Sect. 2, no probabilities are assumed except the familiar and non-controversial representation of a