THE USE OF PRIOR PROBABILITY DISTRIBUTIONS IN STATISTICAL INFERENCE AND DECISIONS

D. V. LINDLEY
UNIVERSITY OF CAMBRIDGE

1. Introduction

The object of this paper is to discuss some of the general points that arise in common statistical problems when a prior probability distribution is used. Most current statistical thinking does not use such a distribution but we feel that some illumination and understanding can be gained through its use, and, in our more enthusiastic moments, we even feel that only a completely Bayesian attitude towards statistical thinking is coherent and practical. The purpose of this paper is not to present propaganda in favor of the Bayesian attitude but merely to explore the consequences of it.

The first two sections of the paper deal with the large-sample problem where the effect of the prior distribution is small and the inferences and decisions made using a Bayesian approach correspond fairly closely with current practice and are similar to those given by Le Cam [7] using the usual approach. If we emphasize any differences that do exist it is because they are of interest: the most important point is the similarity. In the final sections of the paper the problem of small samples is discussed. Here the prior distribution is important and much of the discussion centers around the choice of it. Some tentative rules are suggested for the choice and comparison is made with the work of Jeffreys. The discussion here is fragmentary and incomplete: but it is hoped that the arguments, unsatisfactory as they are, will stimulate others to produce better ones.

Throughout this paper only the problem of the analysis of an experiment is discussed. Design problems are not considered.

2. Large-sample problems of the estimation type

We first formulate the mathematical model following, apart from the prior distribution, the usual pattern. A *sample space* $\mathfrak{X}$ of points x supports a σ-algebra $\mathcal{B}$ of sets X. For each point $\boldsymbol{\theta}$ of a *parameter space* Θ there is defined a probability measure over $(\mathfrak{X}, \mathcal{B})$. It will be assumed that each of these measures is dominated by a σ-finite measure over $(\mathfrak{X}, \mathcal{B})$; from which it follows that there exist probabil-

Present address: University College of Wales, Aberystwyth.