ON PARTIAL A PRIORI INFORMATION IN STATISTICAL INFERENCE

By J. R. Blum and Judah Rosenblatt

University of New Mexico

1. Introduction. Except in rare situations, information concerning the *a priori* distribution of a parameter is likely to be incomplete. Hence the use of a Bayes rule on some systematically produced choice for an *a priori* distribution, as advocated by the Bayesian school, is difficult to justify. This appears to be the case sometimes even if the *a priori* distribution is known fairly accurately (see Theorem 4). Robbins [3] has suggested that attention be paid to the case in which it is known only that the distribution of the parameter is a member of some given family $\mathfrak{J}$ of distributions. In this note we investigate this idea in several specific contexts—in particular the binomial case in which it is known that p is not less than some given p_0, and the case in which the class $\mathfrak{J}$ consists of distributions close to a given one.

Suppose we are given a fixed sample size statistical decision problem, i.e.,

a positive integer k,

a (parameter) set Θ,

an observable random vector $(X_1, \cdots, X_k) \equiv X$ with density f_θ, relative to some given measure m, where $\theta \varepsilon \Theta$ is unknown,

a set $\mathfrak{D}$ of possible decisions D,

a loss function $\mathfrak{L} \geqq 0$ on $\Theta \times \mathfrak{D}$, and

a σ-algebra S of subsets of $\mathfrak{D}$.

If no more information than that listed above is given, then we feel that it is most reasonable to use the minimax criterion, i.e., to use a rule δ which minimizes

$$\sup_{\theta\varepsilon\Theta} E_\theta \mathfrak{L}(\theta, \ \delta[X_1, \cdots, \ X_k]) \ \equiv \ \sup_{\theta\varepsilon\Theta} E_\theta[\textstyle\int_{\mathfrak{D}} \mathfrak{L}(\theta, \ D) \, d\delta(D \mid X_1, \cdots, \ X_k)]$$

where δ is a mapping from real k dimensional space into the set of probability measures over S, provided that such a rule exists.

In contrast to the case in which no further information is available is the situation in which θ is considered to be the value of a random variable governed by a known distribution function F, over an appropriate σ-algebra T of subsets of Θ. Then one naturally attempts to choose δ so as to minimize the average risk

$$\textstyle\int_\Theta E_\theta \mathfrak{L}(\theta, \delta[X_1, \cdots, X_k]) \, dF(\theta) \equiv A(\delta, F).$$

Such a rule is called a Bayes rule relative to F and is usually denoted by δ_F. F itself is called the *a priori* distribution function of θ.

In many problems it is reasonable to assume the existence of an *a priori* distribution function F, but unreasonable to assume perfect knowledge of F. We consider here the problem of decision making when F is known only to be a member of some given class $\mathfrak{J}$.

Received 4 September 1964; revised 11 March 1967.