DECISION THEORY FOR SOME NONPARAMETRIC MODELS

KJELL A. DOKSUM
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction and summary

The problem considered in this paper is that of obtaining optimal decision rules when a parametric form of the distribution of the observations is not known exactly. Thus we assume that the underlying distribution function F of the X_i in the random sample $X = (X_1, \cdots, X_n)$ is in a class Ω of distribution functions, and Ω is not indexed in a natural way by a parameter θ in m dimensional Euclidean space R^m. Let $R(F, d)$ denote the risk of the decision rule $d = d(X)$ when F is the true distribution. Minimax procedures that minimize the maximum risk sup $\{R(F, d); F \in \Omega\}$ have been obtained in special cases by Hoeffding [8], Ruist [14], Huber [9], [10], [11], and Doksum [3]. In particular, Huber was able to show that if Ω is the class of all distributions in a neighborhood of a normal distribution, then the minimax procedures are based on statistics that are, approximately, trimmed means. Most stringent procedures that minimize the maximum shortcoming $\sup_F \{R(F, d) - \inf_d R(F, d)\}$ have been considered by Schaafsma [15].

Another approach would be to define a probability (weight function) P on Ω and then minimize the average (Bayes) risk $\int_\Omega R(F, d)\, P(dF)$, thereby obtaining what is called the Bayes solution. This approach has been taken by Kraft and van Eeden [13], Ferguson [6], and Antoniak [1], who were able to obtain explicit Bayes solutions for some probabilities P. Their work is closely related to the work of Fabius [5], who considered properties of posterior distributions for a class of probability measures P that essentially contains those of Kraft and van Eeden and of Ferguson. Fabius' work in turn is related to that of Freedman [7], who considered properties of Bayes procedures in the case where the X_i are discrete random variables. The relationship between these papers will be discussed further in Section 5.

In this paper, we introduce a criterion which involves minimizing a quantity between the maximum risk and the average risk. This criterion is appropriate when the probability P on Ω is not fully specified, but only the distribution of $F(t_1), \cdots, F(t_k)$ is known for some $t_1 < \cdots < t_k$. Thus past records may

Prepared with the partial support of the National Science Foundation under Grant GP-15283, and with the partial support of the Office of Naval Research under Contract N00014-69-A-0200-1036. Completed while the author was at the University of Oslo.