NOTE ON ESTIMATING ORDERED PARAMETERS

By Ester Samuel

Hebrew University

1. Introduction. We consider the problem of estimating a set of k real valued parameters, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$ where $\theta_i \varepsilon S$, $i = 1, \cdots, k$. Let $\mathbf{X}$ be the (usually vector valued) random variable with values $\mathbf{x}$, the distribution of which depends upon $\boldsymbol{\theta}$ and let $\boldsymbol{\delta} = \boldsymbol{\delta}(\mathbf{X}) = (\delta_1(\mathbf{X}), \cdots, \delta_k(\mathbf{X}))$ be an estimator of $\boldsymbol{\theta}$. Since $\boldsymbol{\theta}$ is known to belong to S^k, the k-fold Cartesian product of S, we shall restrict $\boldsymbol{\delta}$ to belong to S^k with probability one.

We assume that the loss incurred by saying $\boldsymbol{\delta}$ when the parameter is $\boldsymbol{\theta}$ is

$$L(\boldsymbol{\delta}, \boldsymbol{\theta}) = \textstyle\sum_{i=1}^k \phi(|\delta_i - \theta_i|) \tag{1}$$

where $\phi(t)$, $t \geqq 0$, is a monotone increasing function.

The problem described above is usually called an estimation problem only if S is an interval. We shall however not put any restrictions on S except (to avoid trivialities) that it contains at least two elements. Thus, e.g., when S is finite we consider what is usually called a multidecision problem. We shall also allow randomized procedures, but in order not to complicate the notation we shall not introduce a special notation when $\boldsymbol{\delta}$ is randomized. Thus, in what follows, $\boldsymbol{\delta}$ should be interpreted to be the value of the estimator after the randomization experiment has been carried out.

Suppose now that $\boldsymbol{\theta}$ is known to belong to Ω, a subset of S^k. Is it then necessary for $\boldsymbol{\delta}$ to belong to Ω in order for $\boldsymbol{\delta}$ to be admissible? That is, must

$$P(\boldsymbol{\delta}\, \varepsilon\, \Omega; \boldsymbol{\theta}) = 1 \qquad \text{for every } \boldsymbol{\theta}\, \varepsilon\, \Omega \tag{2}$$

in order for $\boldsymbol{\delta}$ to be admissible?

In this generality, the answer is known to be in the negative. Robbins in [2] considers the (nonsequential) compound decision problem where for $i = 1, \cdots, k$ one has observations X_i from a normal population with variance 1 and mean, $\theta_i \varepsilon \{-1, 1\}$, and the X_i's are independent. Thus here $\mathbf{X} = (X_1, \cdots, X_k)$, and S^k contains 2^k points. The only values of $\phi(t)$ of interest here are $\phi(0)$ and $\phi(2)$, which are taken to be 0 and 1 respectively. Suppose it is known that exactly one of the parameters θ_i equals 1 and the $k - 1$ others equal -1. Thus Ω contains the k points having one coordinate $+1$ and the others -1. In [2], p. 138, it is shown that for $k > 2$ the Bayes rule $\boldsymbol{\delta}$ with respect to the *a priori* distribution which assigns equal probability $1/k$ to each element of Ω takes the value $\boldsymbol{\delta} = (-1, \cdots, -1)$ with positive probability under every $\boldsymbol{\theta}\, \varepsilon\, \Omega$, and hence clearly fails to satisfy (2). Since this Bayes rule is essentially unique the rule obtained certainly is admissible for the restricted problem of deciding on $\boldsymbol{\theta}\, \varepsilon\, \Omega$. (This result is actually not too surprising. $\boldsymbol{\delta}(\mathbf{x})$ takes the value $(-1, \cdots, -1)$ when all x_i's are nearly equal. In that case assigning the value $+1$ to some

Received 31 December 1963; revised 2 November 1964.