ON ϵ-COMPLETE CLASSES OF DECISION FUNCTIONS[1]

By J. Wolfowitz

Columbia University

An example of an "almost subminimax" solution[2] is given in a paper by Hodges and Lehmann ([1], Section 5, Problem 1). Another paper, written independently by Robbins [2], has put much stress on the idea and given a considerable discussion and many examples. Frank and Kiefer [3] have given a prescription for constructing almost subminimax solutions.

Let $\Omega = \{F\}$ be a set of distribution functions F. The statistician has to make one of a set of decisions in a space $D^t = \{d^t\}$. He takes observations in stages (finite subsets) on an infinite sequence of chance variables $X = X_1, X_2, \cdots$, distributed according to an unknown one of the F's. The statistician employs a decision function δ, a rule (which may involve randomization) which tells him when to stop taking observations and what decision to make when he has stopped taking observations. The risk $r(F, \delta)$ of a decision function (d.f.) δ when F is the distribution function of X is the sum of the expected values of the cost function and the loss function. All these ideas are described rigorously and in detail in the book [4] by Wald, whose notation we adopt. Some familiarity with this book and its ideas will be assumed. The brief résumé given in this paragraph was partly for the purpose of recalling some of the important notation.

An almost subminimax d.f. δ^* may be roughly described as follows. Let δ^{**} be, say, an admissible minimax d.f. For all F's in Ω we have $r(F, \delta^*) < r(F, \delta^{**}) + \epsilon$ with ϵ "small" and positive, while for "most" F's in Ω we have $r(F, \delta^{**}) - r(F, \delta^*)$ equal to a "large" positive number.

An important task of the mathematical statistician is to exhibit a complete class of d.f.'s for a problem under consideration; an essentially complete class is even more useful[3]. The choice of d.f. from among the members of an essentially complete class requires additional principles. A possible principle is to choose a minimax d.f. (There may be more than one minimax d.f., and even more than one admissible minimax d.f.) This might be the course of a very conservative statistician whose ignorance of F is complete. The appeal of an almost subminimax d.f. in preference to a minimax d.f. occurs when the statistician considers ϵ small, or the F's for which $r(F, \delta^*) > r(F, \delta^{**})$ rather "unlikely," or both. There will usually be little difficulty in deciding when ϵ is small, but perhaps considerable difficulty in deciding that some "few" F's are "unlikely" and just what is to be done about them.

Let $\epsilon > 0$ be a fixed number. A d.f. δ_1 will be called ϵ-equivalent to the d.f. δ_2 if $| r(F, \delta_1) - r(F, \delta_2) | \leqq \epsilon$ for all F in Ω. (The relation of ϵ-equivalence is obviously not transitive.) A d.f. δ_1 will be said to be ϵ-better than the d.f. δ_2 if δ_1 and δ_2 are not ϵ-equivalent and if $r(F, \delta_1) \leqq r(F, \delta_2) + \epsilon$ for every F in Ω. A

[1] Research under a contract with the Office of Naval Research.

[2] This represents a slight change in nomenclature from that of Robbins' paper [2].

[3] Provided, of course, that it is not complete.