COWLES COMMISSION DISCUSSION PAPER: STATISTICS NO. 369

Comments on Statistical Decisions

Leonid Hurwicz

November 29, 1951

(1) Denote by F the cumulative distribution of a sequence $X = \{X_1, X_2, \ldots\}$ of random variables and let $\mathcal{F}^*$ be the class of all possible F's. Let A be the set of actions open to the statistician and let o be the event ("outcome") that would take place if F were true and the statistician chose a, so that $o = \psi(a, F)$, $a \in A$, $F \in \mathcal{F}^*$, $o \in O$. [The outcome function ψ is to be interpreted causally; it does not involve the statistician's preferences; Wald's weight function, interpreted, say, as a money penalty to be paid by the statistician (not as negative utility!) is an example of an outcome function.] ψ is assumed known.

The action taken by the statistician will, in general, depend on the observations x, on the ("a priori") information (or beliefs) I_F concerning F held independently of x, and on the nature of ψ. A statistical decision principle may be defined as a functional relation φ assigning a nonempty ("optimal") subset $\hat{A}$ of A to each triple (x, I_F, ψ) i.e.,

$$\hat{A} = \varphi(x, I_F, \psi). \tag{I}$$

The choice of φ depends on the statistician's preferences and his concept of "rationality."[1]

---

1. The more usual procedure is first to form a "risk function" $\rho(I_F, \psi)$ with ρ depending on the statistician's preferences when $\mathcal{J}_{\mathcal{F}^*}$ is of the Bayesian type (i.e., on his "utility function") and then to select an optimality criterion σ in $\hat{D} = \sigma(\rho)$, $\hat{D} \subseteq D$, where the elements d of D are decision functions, i.e., $a = d(x)$.

(2) Denote by $\mathcal{D}_{\mathcal{J}^*}$ the domain of I_F (the class of possible states of a priori information). Two special instances of $\mathcal{D}_{\mathcal{J}^*}$ are well known: (a) the Bayesian case, where each element of $\mathcal{D}_{\mathcal{J}^*}$ is a probability measure ξ on $\mathcal{J}^*$, ξ assumed known; (b) the case where each element of $\mathcal{D}_{\mathcal{J}^*}$ is a nonempty subset $\mathcal{J}$ of $\mathcal{J}^*$, $\mathcal{J}$ assumed known. (c) A more general $\mathcal{D}_{\mathcal{J}^*}$ (with (a) and (b) as special cases) has been suggested[2] where each element of $\mathcal{D}_{\mathcal{J}^*}$ is a set Ξ of ξ's, Ξ assumed known.

(3) The statistician's action choice depends on φ, but also on the $\mathcal{D}_{\mathcal{J}^*}$ selected. When a particular combination (φ, $\mathcal{D}_{\mathcal{J}^*}$) gives what seem "unreasonable" results, the fault may be with φ or $\mathcal{D}_{\mathcal{J}^*}$ or both.

Thus the difficulties with the (negative utility) minimax principle where $\mathcal{D}_{\mathcal{J}^*}$ is of type (b) in (2) above, have led in some instances to a favorable reconsideration of the Bayesian $\mathcal{D}_{\mathcal{J}^*}$, while in others alternative φ's (with (b) or (c)—type $\mathcal{D}_{\mathcal{J}^*}$) have been investigated. Among the latter may be mentioned Savage's minimax "regret" criterion and the class of criteria given at the end of (4), loc. cit. in ft. (2).[3]

(4) Among the various requirements on φ and $\mathcal{D}_{\mathcal{J}^*}$ it might seem proper to expect that if $x = (x', x'')$ then for any (x', I_F) there should exist $I_F' \in \mathcal{D}_{\mathcal{J}^*}$, depending on x' and I_F but not on x'' or Ψ, such that

(II) $\qquad \varphi(x, I_F, \Psi) = \varphi(x'', I_F', \Psi)$ for all x'' and Ψ.

[(II) requires that there should always be a purely a priori equivalent of a mixture of observations and a priori information. This seems related to certain features of Fisher's "fiducial probabilities."]

---

2. "Some Specification Problems and Applications to Econometric Models," Abstract, Econometrica, Vol. 19, No. 3, July, 1951, pp. 343-4, Leonid Hurwicz.

3. A set of postulates defining a somewhat broader class of optimality criteria will be presented elsewhere.

(II) is satisfied in the Bayesian case where I_F' is the a posteriori distribution ξ' of F given x' when the a priori distribution is ξ. Similarly, a class of a posteriori distributions Ξ' generated by a given a priori class Ξ may be used as I_F' with $\mathcal{D}_{\mathcal{J}^*}$ of type (c). A difficulty arises when $\Xi' = \Xi$ as is the case when $\Xi = \Xi_{\mathcal{J}}$ is the class of all ξ's definable on some $\mathcal{J} \subsetneq \mathcal{J}^*$. Whether this is an indication that (II) is unreasonable remains to be seen.