

NOTE: Cowles Commission Discussion Papers are preliminary materials circulated privately to stimulate private discussion and are not ready for critical comment or appraisal in publications. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

Comments on Statistical Decisions

Leonid Hurwicz

November 29, 1951

(1) Denote by  $F$  the cumulative distribution of a sequence  $X = \{X_1, X_2, \dots\}$  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  $\omega$  be the event ("outcome") that would take place if  $F$  were true and the statistician chose  $a$ , so that  $\omega = \Psi(a, F)$ ,  $a \in A$ ,  $F \in \mathcal{F}^*$ ,  $\omega \in \Omega$ . [The outcome function  $\Psi$  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.]  $\Psi$  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  $\Psi$ . A statistical decision principle may be defined as a functional relation  $\varphi$  assigning a nonempty ("optimal") subset  $\hat{A}$  of  $A$  to each triple  $(x, I_F, \Psi)$  i.e.,

$$(I) \quad \hat{A} = \varphi(x, I_F, \Psi).$$

The choice of  $\varphi$  depends on the statistician's preferences and his concept of "rationality."<sup>1/</sup>

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

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

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

Thus the difficulties with the (negative utility) minimax principle where  $\mathcal{D}_{\mathcal{F}^*}$  is of type (b) in (2) above, have led in some instances to a favorable reconsideration of the Bayesian  $\mathcal{D}_{\mathcal{F}^*}$ , while in others alternative  $\varphi$ 's (with (b) or (c)--type  $\mathcal{D}_{\mathcal{F}^*}$ ) 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).<sup>3/</sup>

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

$$(II) \quad \psi(x, I_{\mathcal{F}}, \Psi) = \varphi(x'', I_{\mathcal{F}}', \Psi) \text{ for all } x'' \text{ and } \Psi.$$

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