COWLES COMMISSION DISCUSSION PAPER: STATISTICS NO. 371

Consistent Decision Functions and Bayes Solutions*

Roy Radner

February 1, 1952

1. In a previous discussion paper[1] two axioms were proposed governing the relationship between a good decision function and the number of observations taken. One of these axioms (Axiom 6) required, essentially, that if there is an infinite sequence of observations from which, in the limit, the true "state of nature" can be determined with probability one, then a good decision rule should advise taking an arbitrarily large number of observations, provided the costs of observation can be made small enough.[2]

Further thought has made it apparent that any plausibility inherent in this axiom stems from the fact that we desire "consistency" in a statistical decision function, in the sense of a "consistent estimate" or a "consistent test." The purpose of this paper is to give a general definition of consistency and to show a connection between it and Bayes solutions of deci-

---

* Research undertaken by the Cowles Commission for Research in Economics under contract Nonr-358(01), NR 047-006 with the Office of Naval Research.

1. CCDP: Statistics No. 367, "Rational Selection of a Decision Function in an Observational Situation."

2. It is assumed that the true state of nature cannot be determined from a finite number of observations.

sion problems.

Section 2 outlines the main ideas for the finite case. Section 3 indicates extensions to the infinite case which have been made to date. Section 4 discusses the possibility of other theorems and also a different definitive of consistency suggested by E. Lehmann.

No proofs are given in this outline. Firstly, most of the theorems did not originate with me, and secondly it is hoped that the task of extending these notions to the most general cases will be completed in the near future.

2. Given, a sequence of problems Q_n characterized by:

Ω, a finite set of states of nature ω

D, a finite set of terminal decisions d.

$K(\omega, d)$, the gain function. (real valued)

$x_1, \ldots, x_n$, observations whose joint distribution is determined by ω, and is absolutely continuous or discrete.

δ_n, a (decision) function from the space of all n-tuples $(x_1, \ldots, x_n)$ to D.

$\delta = (\delta_1, \delta_2, \ldots$ etc.) a sequence of decision functions.

ω_o, the true state of nature.

<u>Definition 1.</u> δ is <u>consistent</u> $\equiv$

$$\operatorname*{prob\,lim}_{n \to \infty} K(\omega_o, \delta_n(x_1, \ldots, x_n)) = \max_{d \in D} K(\omega_o, d)$$

<u>Definition 2.</u> Let $G_n(\omega | x_1, \ldots, x_n)$ be a probability measure on Ω, which depends upon $x_1, \ldots, x_n$; then

G_n <u>converges</u> <u>upon</u> ω_o

$$\begin{cases} \operatorname*{prob\,lim}_{n \to \infty} G_n(\omega_o | x_1, \ldots, x_n) = 1 \\ \operatorname*{prob\,lim}_{n \to \infty} G_n(\omega | x_1, \ldots, x_n) = 0, \ \omega \neq \omega_o \end{cases}$$

If $\max_{u \in U} f(u)$ is attained at u_0 we will write $u_0 = \max'\text{er}_{u \in U} f(u)$.

Theorem 1.

Let $\delta_n(x_1, \ldots, x_n) \equiv \max'\text{er}_{d \in D} \sum_{\omega \in \Omega} K(\omega, d)\, G_n(\omega | x_1, \ldots, x_n)$

If $G_n(\omega\, x_1, \ldots, x_n)$ converges upon ω_0, then for any function $K(\omega, d)$ $\delta = (\delta_1, \delta_2, \ldots)$ will be consistent.

Definition 3.

$G_n(\omega | x_1, \ldots, x_n)$ is the a posteriori distribution relative to the a priori distribution G_0, given $x_1, \ldots, x_n \equiv$

$$G_n(\omega | x_1, \ldots, x_n) \equiv \frac{f_\omega(x_1, \ldots, x_n)\, G_0(\omega)}{\sum_{\omega \in \Omega} f_\omega(x_1, \ldots, x_n)\, G_0(\omega)}$$

where $f_\omega(x_1, \ldots, x_n)$ is the joint elementary probability law of $x_1, \ldots, x_n$ given ω.

Definition 4. $\hat{\delta}_n$ is Bayes with respect to $G_0 \equiv$

$$\hat{\delta}_n = \max'\text{er}_{\delta_n} \sum_{\Omega} \left[\int K(\omega, \delta_n(x_1, \ldots, x_n))\, f_\omega(x_1, \ldots, x_n)\, dx_1, \ldots, dx_n \right] G_0(\omega).$$

I do not know with whom the next theorem originated, although the substance of it seems to be well known. (cf., for example, Wald, "Statistical Decision Functions", p. 124, theorem 5.1)

Theorem 2.

If G_n is the a posteriori distribution relative to G_0 and

$$\delta_n(x_1, \ldots, x_n) = \max'\text{er}_{d \in D} \sum_{\Omega} K(\omega, d)\, G_n(\omega | x_1, \ldots, x_n)$$

then δ_n is Bayes with respect to G_0.

Definition 5.

G_0 is unprejudiced $\equiv G_0(\omega) > 0$, for every ω.

The next theorem is also well known.

<u>Theorem 3.</u> If $x_1, x_2, \ldots,$ etc. are independent and identically distributed, G_o is unprejudiced and G_n is a posteriori relative to G_o, then G_n coverages upon ω_o.

<u>Theorem 4.</u> If $X_1, X_2, \ldots,$ etc. are independent and identically distributed, $\hat{\delta}_n$ is Bayes with respect to G_o and G_o is unprejudiced, then for any function $K(\omega, d)$, $\delta = (\delta_1, \delta_2, \ldots\ldots)$ will be consistent.

3. <u>Generalizations to the Infinite Case.</u> Let Ω be a topological space with a Borel field containing all the open sets.

<u>Definition 2'.</u> Let $G_n(W \mid x_1, \ldots, x_n)$ be a probability Measure on Ω, which depends on $x_1, \ldots, x_n$; then

G_n <u>coverages upon ω_o</u> $\equiv$

$$\begin{cases} \text{for every open } U \text{ with } \omega_o \in U, \ \underset{n \to \infty}{\text{prob lim}}\ G_n(U \mid x_1, \ldots, x_n) = 1 \\ \text{for every closed } W \text{ with } \omega_o \notin W, \ \underset{n \to \infty}{\text{prob lim}}\ G_n(W \mid x_1, \ldots, x_n) = 0 \end{cases}$$

<u>Theorem 1'.</u> If $K(\omega, d)$ is bounded, and is continuous in ω, <u>uniformly in d</u>, and if there is a sequence $G_n(W \mid x_1, \ldots, x_n)$ converging upon ω_o, then there exists a consistent sequence of decision functions.[3]

The next theorem, a generalization of Theorem 3, is due to Doob.[4]

<u>Theorem 3'.</u> Let $X_1, X_2, \ldots,$ etc. be independent and identically distributed. Let Ω be a Borel set of a complete metric space, such that to different ω's correspond different distributions of the X's. Then if the

---

3. In Definition 1, $\max_{d \in D}$ should now be replaced by $\sup_{d \in D}$.

4. J. L. Doob, "Application of the Theory of Martingales", <u>Colloques Inter-Nationaux du Centre National de la Recherche Scientifique</u>, XIII (1949), pp. 23-27.

a priori measure G_0 assigns non-zero measure to every open set, the sequence of a posteriori measures will converge upon ω_0.

4. Theorem 4 is the best of its kind in the sense that for any prejudiced G_0 there exists a function $K(\omega, d)$ such that the sequence of Bayes decision functions will not be consistent.

However, for any particular $K(\omega, d)$ there may be many prejudiced G_0's which give rise to consistent decisions. This direction needs further study.

E. Lehmann, in comments on CCDP: Statistics No. 367 has independently suggested a slightly different definition of consistency; namely, that as the cost of observations tend to zero, the expected gain of the decision procedure tend to the maximum that could be attained if ω were known. In a certain sense, Lehmann's proposal is more general, in that it is stated in terms of the (basic) concepts of gains and costs only. However, I think that it is useful to put one's finger directly on the important problem of the dependence of the terminal decision on the number of observations. This gives us a different kind of generality in that we obtain results which are independent of the costs of observation.