Sankhyā : The Indian Journal of Statistics
1979, Volume **41**, Series A, Pts. 1 and 2, pp. 48–58.

RELATIONSHIP BETWEEN BAYES, CLASSICAL AND DECISION THEORETIC SUFFICIENCY

By K. K. ROY and R. V. RAMAMOORTHI
Indian Statistical Institute

SUMMARY. Three notions of sufficiency, Bayes, classical and decision theoretic have been considered in the literature. These three notions are equivalent when the statistical structure is dominated, In this paper relationship between the three notions is investigated in the undominated case with particular attention to the case when the σ-fields are countably generated.

0. Introduction

Suppose $(X, \mathcal{A})$ is a measurable space carrying a family of probability measures $\{P_\theta : \theta \in \Theta\}$. Though relevant only in a later section we shall throughout assume that Θ is equipped with a σ-field $\mathcal{C}$ and that for all A in $\mathcal{A}$ $\theta \to P_\theta(A)$ is $\mathcal{C}$-measurable. There are three approaches to the concept of sufficiency of a sub σ-field $\mathcal{B}$ of $\mathcal{A}$.

(i) *Classical :* There is a conditional probability on $\mathcal{A}$ given $\mathcal{B}$ independent of θ in Θ.

(ii) *Decision theoretic :* Given any decision problem and any decision rule δ therein, there is a $\mathcal{B}$-measurable decision rule δ' equivalent to δ.

(iii) *Bayesian :* Given any prior ξ on $(\Theta, \mathcal{C})$, the posterior on Θ given $\mathcal{A}$ is the same as the posterior given $\mathcal{B}$.

These concepts are defined more precisely in the next section. We shall refer to classical sufficiency simply as Sufficiency and to (ii) and (iii) as D-Sufficiency and Bayes Sufficiency respectively.

The three notions are equivalent when $\{P_\theta : \theta \in \Theta\}$ is dominated by a σ-finite measure. Burkholder's Example (1961) of a non-sufficient σ-field containing a sufficient σ-field shows that neither (ii) nor (iii) is equivalent to (i). Blackwell conjectured to us that when the spaces $(X, \mathcal{A})$ and $(\Theta, \mathcal{C})$ are standard Borel and $\mathcal{B}$ is countably generated (i), (ii) and (iii) would be equivalent even if $\{P_\theta : \theta \in \Theta\}$ is undominated. As a first step towards settling the conjecture, in this paper we study the relationship between the three definitions when the σ-fields considered are all countably generated. Interest in countably generated σ-fields stems from the fact that these are precisely the σ-fields generated by Borel measurable real valued statistics.

1. Relationship between D-sufficiency and sufficiency

We begin by giving precise definitions of Sufficiency and D-Sufficiency.

Definition : $\mathscr{B} \subset \mathscr{A}$ is *Sufficient* for $(X, \mathscr{A}, P_\theta : \theta \epsilon \Theta)$ if given any bounded real valued $\mathscr{A}$ measurable function f, there is a $\mathscr{B}$-measurable function f^* such that f^* is a version of $E_\theta(f \mid \mathscr{B})$ for all $\theta \epsilon \Theta$.

Let (A, σ) be a set A equipped with a σ-field σ. We shall refer to (A, σ) as Action Space. By a decision rule $\delta(\cdot\,, \cdot)$ we mean a stochastic kernel from $(X, \mathscr{A})$ to (A, σ) i.e., for every x, $\delta(x, .)$ is a probability measure on σ and for every $E \epsilon \sigma$ $\delta(., E)$ is $\mathscr{A}$-measurable as a function of x. A decision rule $\delta(.\,, .)$ is said to be $\mathscr{B}$-measurable if for all $E \epsilon \sigma$, $\delta(x, E)$ as a function of x is $\mathscr{B}$-measurable.

Definition : A sub σ-field $\mathscr{B}$ of $\mathscr{A}$ is *D-Sufficient* if given any action space (A, σ) and a decision rule $\delta(., .)$ there is a -measurable decision rule $\delta'(.\,, .)$ such that for all $E \epsilon \sigma$ and $\theta \epsilon \Theta$

$$\int_X \delta(x, E)\, dP_\theta = \int_X \delta'(x, E)\, dP_\theta \qquad \ldots \quad (1.1)$$

Proposition 1.1 : *$\mathscr{B}$ is D-Sufficient for $(X, \mathscr{A}, P_\theta : \theta \epsilon \Theta)$ iff there is a $\mathscr{B}$-measurable stochastic kernel $Q(.\,, .)$ from $(X, \mathscr{B})$ to $(X, \mathscr{A})$ such that for $A \epsilon \mathscr{A}$ and $\theta \epsilon \Theta$*

$$\int_X Q(x, A)\, dP_\theta = P_\theta(A). \qquad \ldots \quad (1.2)$$

Proof: The 'if' part is trivial. For the 'only if' part choose (A, σ) to be $(X, \mathscr{A})$ and as δ the decision rule $\delta(x, E) = I_E(x)$.

We now state the main theorem of this section. [Q. E. D.]

Theorem 1 : *Suppose $\mathscr{B}$ is D-sufficient for $(X, \mathscr{A}, P_\theta : \theta \epsilon \Theta)$ then $\mathscr{B}$-contains a sufficient σ-field.*

Proof: Let $Q(.\,, .)$ be a $\mathscr{B}$-measurable kernel satisfying (1.2). For each bounded $\mathscr{A}$-measurable function f define

$$T f(x) = \int f(y)\, Q(x, dy).$$

Associate with each bounded $\mathscr{A}$-measurable function f a $\mathscr{B}$-measurable function f^* as follows

$$f^*(x) = \begin{cases} \lim\limits_{n \to \infty} \dfrac{1}{n} \sum\limits_{k=1}^{n} T^k f(x) & \text{when the limit exists} \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathcal{B}_0 = \sigma\{f^* : f\text{-bounded } \mathcal{A}\text{-measurable}\}$.

$\mathcal{B}_0 \subset \mathcal{B}$ and we shall show that $\mathcal{B}_0$ is sufficient. By Hopf's ergodic theorem (Neveu, 1965) for all $\theta \in \Theta$

$$f^*(x) = E_\theta(f \mid \mathcal{B}_\theta)[P_\theta] \quad \dots \quad (1.3)$$

where

$$\mathcal{B}_\theta = \{A \in \mathcal{A} : TI_A = I_A\,[P_\theta]\}.$$

By (1.3) $\mathcal{B}_0 = \mathcal{B}_\theta\,[P_\theta]$. Hence by (1.3)

$$f^*(x) = E_\theta(f \mid \mathcal{B}_0)\,[P_\theta]$$

establishing sufficiency of $\mathcal{B}_0$. [Q. E. D.].

Theorem 2 : *If $\mathcal{B}$ is countably generated and D-sufficient then $\mathcal{B}$ is itself sufficient.*

Proof: By Theorem 5 of Burkholder (1961) any countably generated σ-field containing a sufficient σ-field is itself sufficient. [Q. E. D.]

The following corollary is an immediate consequence of Theorem 1.

Corollary : *If $\mathcal{B}$ is D-sufficient then $\mathcal{B}$ is also Bayes sufficient.*

Weaker forms of D-Sufficiency can be obtained by considering restricted classes of action spaces such as

(D_1) compact metric action spaces

(D_2) finite action spaces

(D_3) 2-point action spaces.

When the sample space is standard Borel D_1 would be equivalent to D. D_3 is known in the literature as 'Test Sufficiency'. We do not know the relationship between D_1, D_2, D_3 in the undominated case.

2. Bayes sufficiency and classical sufficiency

As before $(\Theta, \mathcal{C})$ is a measurable space and $P_\theta(\cdot)$ is a stochastic kernel from $(\Theta, \mathcal{C})$ to $(X, \mathcal{A})$. $\hat{\Theta}$ stands for all probability measures on $(\Theta, \mathcal{C})$. For each probability measure ξ in $\hat{\Theta}$ we denote by λ_ξ the probability measure on $(X \times \Theta, \mathcal{A} \times \mathcal{C})$ defined by

$$\lambda_\xi(A \times C) = \int_C P_\theta(A)\, d\xi(\theta).$$

We shall denote by λ_ξ' the marginal of λ_ξ on $(X, \mathcal{A})$. We shall denote by $X \times \mathcal{C}$ the σ-field containing all sets of the form $X \times C$ for C in $\mathcal{C}$. $\mathcal{A} \times \Theta$ and $\mathcal{B} \times \Theta$ are similarly defined. For a function f on X, by $\bar{f}$ we shall mean the function on $X \times \Theta$ defined by $\bar{f}(x, \theta) = f(x) \;\forall x, \theta$.

Definition: A sub-σ-field $\mathcal{B}$ of $\mathcal{A}$ is said to be *Bayes Sufficient* for $(X, \mathcal{A}, P_\theta : \theta \in \Theta)$ if for all C in $\mathcal{C}$ and for all ξ in $\hat{\Theta}$.

$$E_{\lambda_\xi}(I_{X \times C} \mid \mathcal{B} \times \Theta) = E_{\lambda_\xi}(I_{X \times C} \mid \mathcal{A} \times \Theta).$$

Proposition 2.1 : *The following are equivalent :*

(i) *$\mathcal{B}$ is Bayes sufficient for $(X, \mathcal{A}, P_\theta : \theta \in \Theta)$.*

(ii) *The sub-σ-fields $\mathcal{A}\Theta \times$ and $X \times \mathcal{C}$ are conditionally independent given $\mathcal{B} \times \Theta$ on the probability sapce $(X \times \Theta, \mathcal{A} \times \mathcal{C}, \lambda_\xi)$ for each ξ in $\hat{\Theta}$.*

(iii) *For every bounded $\mathcal{A}$-measurable function f on X there is a $\mathcal{B}$-measurable function f^* such that*

$$\bar{f}^* = E_{\lambda_\xi}(\bar{f} \mid \mathcal{B} \times \mathcal{C}) \text{ for each } \xi \in \hat{\Theta}.$$

Proof: Immediate from Proposition 25.3A of Loeve (1955, p. 351). [Q. E. D.]

We had already remarked that in the undominated case Bayes sufficiency does not in general imply sufficiency. In this section we address ourselves to the situation when the σ-fields under consideration are all countably generated. We first give an example to show that the assumption of countable generation is not enough to ensure the implication.

Example 2.1 :

$X = \Theta = [0, 1]$.

D : a non-Borel universally measurable subset of [0, 1].

$\mathcal{B}$: Borel σ-field on [0, 1].

$\mathcal{A} = \mathcal{C}$: σ-field generated by $\{\mathcal{B}, D\}$.

$P_\theta(\mathcal{A}) = I_A(\theta)$: i.e., P_θ is the measure degenerate at θ.

Now given ξ in Θ there is a B_ξ in $\mathcal{B}$ such that $\xi(B_\xi) = 1$ and $\mathcal{B}$ is clearly sufficient for $(X, \mathcal{A}, P_\theta : \theta \in B_\xi)$. Bayes sufficiency of $\mathcal{B}$ now follows from Proposition 2.2. But $\mathcal{B}$ is not sufficient.

Remarks: In the above example $\mathcal{B}$ is far from being sufficient. It is easy to see, by considering $I_D(x)$, that $\mathcal{B}$ is not even test sufficient. In particular in the testing problem $H_0 : \theta \in D$ against $H_1 : \theta \notin D$ with 0-1 loss function the decision rule

$$\delta(x) = \begin{cases} \text{accept } H_0 \text{ if } x \text{ belongs to } D \\ \text{accept } H_1 \text{ if } x \text{ belongs to } D^c \end{cases}$$

has a risk function identically equal to zero. In terms of risk function every $\mathcal{B}$ measurable decision rule is worse than δ and in this context use of $\mathcal{B}$ measurable rules would be unsatisfactory. On the other hand if one were concerned with only Bayes risks then restriction to $\mathcal{B}$-measurable decision rules will not entail any additional loss.

In what follows we investigate the equivalance of these two notions under certain additional assumptions. However we are unable to decide the restrictiveness of these assumptions even in the case when the spaces Θ and X are both standard Borel.

Proposition 2.2 : *Suppose $\mathcal{A}$ and $\mathcal{B}$ are countably generated. Then $\mathcal{B}$ is Bayes sufficient iff for every $\xi \in \hat{\Theta}$ there is a set E_ξ in $\mathcal{C}$ of ξ-measure* 1 *such that $\mathcal{B}$ is sufficient for $(X, \mathcal{A}, P_\theta : \theta \in E_\xi)$.*

Proof: 'If part'.

Given ξ since there is an E_ξ of ξ-measure 1 such that $\mathcal{B}$ is sufficient for $(X, \mathcal{A}, P_\theta : \theta \in E_\xi)$, for any bounded $\mathcal{A}$-measurable function f choose an f^*, $\mathcal{B}$-measurable such that

$$f^* = E_\theta(f \mid \mathcal{B}) \quad \theta \in E_\xi.$$

Now for $C \in \mathcal{C}$

$$\int_C \int_B \bar{f}^* \, d\lambda_\xi = \int_C \int_B \bar{f}^* \, dP_\theta d\,\xi(\theta) = \int_{C \cap E_\xi} \int_B f^* dP_\theta d\xi(\theta)$$

$$= \int_{C \cap E_\xi} \int_B f \, dP_\theta \, d\xi(\theta) = \int_C \int_B \bar{f} \, d\lambda_\xi.$$

'Only if part'.

Fix $\xi \in \hat{\Theta}$. Given f bounded $\mathcal{A}$-measurable, there is by Bayes sufficiency of $\mathcal{B}$, a $\mathcal{B}$-measurable f^* such that

$$\bar{f}^* = E_{\lambda_\xi}(\bar{f} \mid \mathcal{B} \times \mathcal{C}).$$

Now $\int_C \int_B f^* \, dP_\theta \, d\xi(\theta) = \int_C \int_B f \, dP_\theta \, d\xi(\theta)$ for all C in $\mathscr{C}$. Therefore $\int_B f^* \, dP_\theta = \int_B f \, dP_\theta$ a.e. ξ for each B in $\mathscr{B}$. By running B through a countable subalgebra generating $\mathscr{B}$

$$\int_B f^* dP_\theta = \int_B f \, dP_\theta \text{ for } B \in \mathscr{B} \text{ outside a } \xi \text{ null set } N_\xi.$$

Thus

$$f^* = E_\theta(f \mid \mathscr{B}). \quad \theta \notin N_\xi.$$

Now taking a countable union of null sets with f running through indicators of sets in a countable algebra generating $\mathscr{A}$, the proposition is proved. [Q. E. D.]

Proposition 2.3 : *Assume that $\mathscr{B}$ is countably generated. Let f be a bounded $\mathscr{A}$-measurable function. There is then a version of $E_\theta(f \mid \mathscr{B})$ which is jointly measurable in (x, θ) with respect to $(\mathscr{B} \times \mathscr{C})$.*

Proof: Let $\{B_1, B_2, \ldots\}$ generate $\mathscr{B}$.

$\mathscr{B}_n = \sigma(B_1, B_2, \ldots, B_n)$, let $B_n^1 \ldots B_n^{k(n)}$ be atoms of $\mathscr{B}_n$. For $A \in \mathscr{A}$

$$f_\theta^n(x) = \sum_{i=1}^{k(n)} \frac{P_\theta(A \cap B_n^i)}{P_\theta(B_n^i)} I_{B_n^i}(x)$$

is a version of $E_\theta(I_A \mid \mathscr{B}_n)$ which is jointly measurable with respect to $(\mathscr{B}_n \times \mathscr{C})$.

Define

$$f_\theta^*(x) = \begin{cases} \lim_n f_\theta^n(x) & \text{whenever the limit esists} \\ 0 & \text{otherwise.} \end{cases}$$

Since $\{f_\theta^n(x) : n \geqslant 1\}$ forms a martingale for each θ and $\mathscr{B}_n \uparrow \mathscr{B}$ it is easy to see that $f_\theta^*(x) = E_\theta(I_A \mid \mathscr{B})$. The proof can now be completed by considering simple functions and their limits. [Q. E. D.]

Definition : $\{(\Theta, \mathscr{C})(X, \mathscr{A}, P_\theta : \theta \in \Theta)\}$ is said to be 'weakly coherent' if for any bounded $\mathscr{A} \times \mathscr{C}$-measurable function $f_\theta(x)$ satisfying (*)

$$\left[\begin{array}{l} \forall \xi \in \hat{\Theta} \quad (E_\xi, \ \xi(E_\xi) = 1, \text{ and } f_\xi : \mathscr{A}\text{-measurable}) \\ \text{such that for } \quad \theta \in E_\xi f_\xi(x) = f_\theta(x) \ [P_\theta] \end{array} \right] \quad \ldots \quad (*)$$

there is an $\mathscr{A}$-measurable function f^* such that

$$f^*(x) = f_\theta(x) \ [P_\theta] \quad \text{for all } \theta \text{ in } \Theta.$$

A discussion on weak coherence will be deferred to the next section. Here we shall investigate the effect of weak coherence on Bayes Sufficiency.

In what follows $\mathcal{N}_\xi$ will denote the set of $\mathcal{A}$-measurable λ^1_ξ null sets and

$$\mathcal{N} = \bigcap_{\xi \in \hat{\Theta}} \mathcal{N}_\theta = \bigcap_{\theta \in \hat{\Theta}} \mathcal{N}_\xi .$$

Theorem 3 : *Assume that the experiment is weakly coherent. If $\mathcal{B}$ is Bayes sufficient then $\hat{\mathcal{B}} = \bigcap_{\xi \in \hat{\Theta}} \mathcal{B} \vee \mathcal{N}_\xi$ is sufficient. Consequently if $\hat{\mathcal{B}} = \mathcal{B} \vee \mathcal{N}$, then $\mathcal{B}$ is itself sufficient.*

Proof: Let f be bounded $\mathcal{A}$-measurable. Choose a jointly measurable version $f^*_\theta(x)$ of $E_\theta(f \mid \mathcal{B})$. Since $\mathcal{B}$ is Bayes sufficient, by Proposition 2.2, $f^*_\theta(x)$ satisfies (*). Now by weak coherence there is an $\mathcal{A}$-measurable function f^* such that $f^*(x) = f_\theta(x) [P_\theta]$, for $\theta \in \Theta$. We shall complete the proof by showing that $f^*(x)$ is $\mathcal{B} \vee \mathcal{N}_\xi$ measurable for each $\xi \in \hat{\Theta}$.

$$E = \{x : f^*(x) \neq f_\xi(x)\}$$

$$\begin{aligned} \lambda^1_\xi(E) &= \int_\Theta P_\theta(E)\, d\xi(\theta) \\ &= \int_{E_\xi} P_\theta(E)\, d\xi(\theta) + \int_{E^c_\xi} P_\theta(E)\, d\xi(\theta) = 0. \end{aligned}$$ [Q. E. D.]

Theorem 4 : *Assume*

(i) *$(X, \mathcal{A})$ is standard Borel ;*

(ii) *$\{(\Theta, \mathcal{C})\ (X, \mathcal{A}, P_\theta, \theta \in \Theta)\}$ is weakly coherent ;*

(iii) *$\mathcal{N} = \{\phi\}$.*

If further $\mathcal{B}$ is countably generated and Bayes sufficient then $\mathcal{B}$ is itself sufficient.

Proof: By Theorem 3 it is enough to show

$$\hat{\mathcal{B}} = \bigcap_\xi \mathcal{B} \vee \mathcal{N}_\xi = \mathcal{B}.$$

Since $(X, \mathcal{A})$ is Standard Borel and $\mathcal{B}$ is countably generated it suffices to show, (see Blackwell, 1955), that every set $B \in \mathcal{B}$ is a union of $\mathcal{B}$ atoms.

Suppose B_0 is a $\mathcal{B}$-atom such that $B_0 \cap B \neq \phi$ and $B_0 \cap B^c \neq \phi$. Since $\mathcal{N} = \{\phi\}$ there is θ_1 and θ_2 such that $P_{\theta_1}(B_0 \cap B) > 0$ and $P_{\theta_2}(B_0 \cap B^c) > 0$. Let ξ_0 give mass $\frac{1}{2}$ to each of θ_1 and θ_2. Then $B \notin \mathcal{B} \vee \mathcal{N}_{\xi_0}$ for, any set E in $\mathcal{B}$ must either contain B_0 or be disjoint with it and in both the cases $\lambda_{\xi_0}(E \Delta B) > 0$. And this proves the Theorem. [Q. E. D.]

3. Coherence, weak coherence and measurable coherence

The concept of coherent statistical structure is introduced by Hasegawa and Perlman (1974). The original idea is due to Pitcher (1965) who introduces compact statistical structures and generalises results in sufficiency for dominated structures. Compactness may be shown to be equivalent to coherence (see Ghosh, Morimoto and Yamada (1978).

Weak coherence differs from coherence in the following two aspects :

(a) restriction to jointly measurable function $f_\theta(x)$

(b) requirement in (*) for all priors on $(\Theta, \mathcal{C})$ rather than discrete priors only.

If Θ is equipped with a natural σ-field then in the context of Bayes sufficiency and also in view of Proposition 2.3, requirement given by (a) is very natural. Using the requirement (a) we define a concept stronger than weak coherence as follows.

Definition : A statistical structure $(X, \mathcal{A}, P_\theta, \theta \in \Theta, \Theta, \mathcal{C})$ is called measurably coherent if for any bounded $\mathcal{A} \times \mathcal{C}$-measurable function satisfying the following restriction (**)

$$\left[\begin{array}{l} \text{for all pairs } \theta_1, \theta_2 \text{ in } \Theta \text{ there is an } \mathcal{A}\text{-measurable function} \\ f_{\theta_1, \theta_2}(\cdot) \text{ such that } f_{\theta_1, \theta_2}(x) = f_{\theta_i}(x)[P_{\theta_i}] \text{ for } i = 1, 2, \end{array}\right] \quad (**)$$

there is an $\mathcal{A}$-measurable function $f^*(\cdot)$ such that $f^*(x) = f_\theta(x)[P_\theta]$, for all θ in Θ.

Trivially measurable coherence implies weak coherence but the converse is not always true as is seen in the following example.

Example 3.1 :

$$X = \Theta = [0, 1]$$

$$\mathcal{C} = \mathcal{A} = \text{Borel } \sigma\text{-field}$$

$$P_\theta(A) = \frac{1}{2}\lambda(A) + \frac{1}{2} I_A(\theta) \quad \theta \in [0, 1],\ A \in \mathcal{A}$$

where λ is Lebesgue measure on $\mathcal{A}$.

It is easy to see that the above structure is not measurably coherent by considering

$$f_\theta(x) = \begin{cases} 1 & \text{if } \theta = x \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand let $f_\theta(x)$ be a jointly measurable function on $\Theta \times X$ satisfying (*). Define

$$f^*(x) = f_x(x) \quad \text{for all } x \text{ in } X$$

then f^* is $\mathscr{A}$-measurable and to check that for any θ in Θ, $f^*(x) = f_\theta(x)$ $[P_\theta]$ one takes a prior $\xi = \frac{1}{2}\lambda + \frac{1}{2}\ \delta_\theta$ where δ_θ is the probability measure concentrated at θ and one uses the condition (*) for ξ. Hence this structure is weakly coherent but not measurably coherent.

It is also easy to see that coherence with appropriate σ-field on Θ implies measurable coherence and hence weak coherence. Thus if $\{P_\theta, \theta \in \Theta\}$ is dominated by a σ-finite measure the statistical structure is measurably coherent, being already coherent. Further coherence with countably generated $\mathscr{A}$ would entail $\{P_\theta, \theta \in \Theta\}$ to be dominated by a σ-finite measure (see Rogge, 1972). However many undominated structures are measurably (a fortiori weakly) coherent even if $\mathscr{A}$ is countably generated.

Below we shall exhibit a class of undominated structures which are measurably coherent.

Let us assume that

(i) Each P_θ, θ in Θ is discrete

(ii) $\mathscr{N} = \bigcap_\theta \mathscr{N}_\theta = \{\phi\}$.

Definition : Say that $\{(X, \mathscr{A}, P_\theta : \theta \in \Theta), (\Theta, \mathscr{C})\}$ admits a measurable estimator for θ if there is a measurable function g from X to Θ such that $P_{g(x)}\{x\} > 0$.

Such a g will be referred to as a measurable estimator of θ.

Theorem 5 : *If $\{(\Theta, \mathscr{C}), (X, \mathscr{A}, P_\theta : \theta \in \Theta\}$ admits a measurable estimator then it is measurably coherent.*

Proof : Let g be a measurable estimator. Then given any $f_\theta(x)$ jointly measurable satisfying (**) define f^* as

$$f^*(x) = f_{g(x)}(x).$$

It is easy to see that $f^*(x) = f_\theta(x)\,[P_\theta]$ for all θ. [Q. E. D.]

Remarks : (1) Measurable estimators of θ are loosely speaking measurable versions of good estimators of θ. For instance if a measurable version of the maximum likelihood estimate for θ exists, then the MLE itself would be a measurable estimator.

(2) Assume $P_\theta(x)$ is jointly measurable in θ and x; so that the set $\{(x, \theta) : P_\theta(x) > 0\}$ is measurable in $X \times \Theta$. For each x in X look at $A_x = \{\theta : P_\theta(x) > 0\}$. A_x's are all measurable sets and that they are all non-empty is ensured by requiring $\mathcal{N}$ to be empty. The problem of obtaining measurable estimators is then one of measurable selection of points from $\{A_x : x \in X\}$. Various theorems on the existence of such selectors are available when the underlying spaces are Polish (Wagner, 1977).

While the existence of measurable selectors ensures the weak coherence of statistical structures, we do not know the validity of the converse in the Polish case. Below we shall give an example of a statistical structure not admitting a measurable estimator. It should be noted that in the example Θ and X are both Polish, P_θ's are discrete further $\theta_1 \neq \theta_2$ implies $P_{\theta_1} \neq P_{\theta_2}$. This example is due to B. V. Rao.

Example 3.2 :

$$\Theta = X = [0, 1].$$

$$\mathcal{C} = \mathcal{A} : \text{ Borel } \sigma\text{-field.}$$

Let D be a Borel subset of $[0, 1] \times [0, 1]$ not containing a graph (Blackwell, 1969) such that (a) $\pi_1 D = [0, 1]$ where π_1 denotes the projection to the 1-st coordinate (b) D does not intersect the diagonal. By the Borel isomorphism theorem there is a 1–1, measurable map ϕ from $\Theta = [0, 1]$ onto D. Let $\phi = (\phi_1, \phi_2)$ be such an isomorphism. For $\theta \in [0, 1]$ define P_θ as the measure giving mass $\frac{1}{3}$ to $\phi_1(\theta)$ and $\frac{2}{3}$ to $\phi_2(\theta)$.

We shall show that the above statistical structure does not admit a measurable estimator.

For, suppose g is a measurable estimator

$$A = \{x : \phi_1(g(x)) = x\}, \quad B = \{x : \phi_2(g(x)) = x\}$$

$A \cup B = [0, 1]$, since for all x, $P_{g(x)}(x) > 0$ and $A \cap B = \phi$ since D does not intersect the diagonal.

Then the graph of h, defined by

$$h = \phi_2 \circ g I_A + \theta_1 \circ g I_B$$

is contained in D.

It is easy to construct examples of non-weakly coherent structures where the spaces X, Θ are not standard Borel. We do not know of any non-weakly coherent statistical structure in the standard Borel case. We are unable to check whether example 3.2 is weakly coherent or otherwise.

ACKNOWLEDGEMENTS

Discussions with Professors D. Blackwell and A. Maitra stimulated our interest in these problems. Professor B. V. Rao has been our adviser in matters set theoretic and Section 3 owes much to him. Professor J. K. Ghosh has helped us throughout the preparation of this paper in more ways than one. We like to thank all of them for their help.

REFERENCES

BLACKELL, D. (1955): On a class of probability spaces. *Proc. Third Berkeley Symp. Math. Statist. and Prob.* **III** 1-6.

BLACKWELL, D. (1969): A Borel set not containing graph. *Ann. Math. Statist.* **39**, 1345-1347.

BURKHOLDER, D. L. (1961): Sufficiency in the undominated case. *Ann. Math. Statist.* **32**, 1191-1200.

GHOSH, J. K. MORIMOTO, H. and YAMADA, (1978): S. Neyman factorization and minimality of pairwise sufficient subfields. To appear

HASEGAWA M. and PERLMAN, M. D. (1974): On existence of minimal sufficient subfield. *Ann. Statist*, **32**, 1049-1055.

HASEGAWA, M. and YAMADA, S. (1975): Correction to "On the existence of minimal sufficient subfield". *Ann. Statist*, **3**, 1371-1372.

LOEVE, M. (1955): *Probability Theory*, D. Van Nostrand Company Inc.

NEVEU, J. (1965): *Mathematical Foundations of the Calculus of Probability*, Holden Day, London.

PITCHER, T. S. (1965): A more general property than domination for sets of probability measures. *Pacific. J. Math.*, **15**, 597-611.

ROGGE, L. (1972): Compactness and domination. *Manuscripta Math*, **7**, 299-306.

WAGNER, D. (1977): Survey of measurable selection theorems. *SIAM. Journal of Control and Optimization.* **15**, No. 5, 859-903.

Paper received: September, 1978.
Revised: February, 1979.