General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

NASA CR-
151233

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON HOUSTON, TEXAS

(NASA-CR-151233) CHARACTERIZATIONS OF LINEAR SUFFICIENT STATISTICS (Houston Univ.) 15 p HC A02/MF A01 CSCL 12A

N77-19824

Unclas
G3/65 20569

CHARACTERIZATIONS OF LINEAR SUFFICIENT STATISTICS

BY B.C. PETERS, JR., R. REDNER,
AND H.P. DECELL, JR.

REPORT #59 AUGUST 1976

MAR 1977
RECEIVED
NASA STI FACILITY
INPUT BRANCH

PREPARED FOR
EARTH OBSERVATION DIVISION, JSC
UNDER
CONTRACT NAS-9-15000

HOUSTON, TEXAS 77004

Characterizations of Linear Sufficient Statistics

by

B. Charles Peters, Jr.[1], Richard Redner,[1]
and Henry P. Decell, Jr.[1]

University of Houston

August, 1976

Report #59
NAS-9-15000

Characterizations of Linear Sufficient Statistics

By B. Charles Peters, Jr.[1], Richard Redner,[1]
and Henry P. Decell, Jr.[1]

University of Houston

We develop a necessary and sufficient condition that there exist a continous linear sufficient statistic T for a dominated collection of totally finite measures defined on the Borel field generated by the open sets of a Banach space X. In particular, corollary necessary and sufficient conditions that there exist a rank k linear sufficient statistic T for any finite collection of probability measures having n-variate normal densitites are given. In this case a simple calculation, involving only the population means and covariances, determines the smallest integer k for which there exists a rank k linear sufficient statistic T (as well as an associated statistic T itself).

[1]This author was partially supported by NASA Contract NAS-9-15000 with the University of Houston during the preparation of this work.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

1. Introduction. If W is a Banach space, $\mathcal{B}(W)$ will denote the Borel field generated by the open sets of W. The totally finite measures defined on $\mathcal{B}(W)$ will be denoted by $\mathcal{M}(W)$. For $\mu, \lambda \in \mathcal{M}(W)$ we will write $\mu << \lambda$ provided $B \in \mathcal{B}(W)$ and $\lambda(B) = 0$ implies $\mu(B) = 0$. Whenever $\mu << \lambda$, $[d\mu/d\lambda]$ will denote the equivalence class of Radon-Nikodym derivatives of μ with respect to [2] [3]. If $\mathcal{D} \subset \mathcal{M}(W)$, $\mathcal{D}$ will be called a dominated (by λ) set of measures provided there exists $\lambda \in \mathcal{M}(W)$ (λ not necessarily in $\mathcal{D}$) such that $\mu \in \mathcal{D}$ implies $\mu << \lambda$. We will call $\mathcal{D} \subset \mathcal{M}(W)$ equivalent to λ ($\mathcal{D} \equiv \lambda$) provided $\mathcal{D}$ is dominated by λ and $\mu(B) = 0$ for each $\mu \in \mathcal{D}$ implies $\lambda(B) = 0$.

If X and Y are Banach spaces and $T: X \to Y$ then, following the notation in [3], we write $f(\epsilon)T^{-1}(\mathcal{B}(Y))$ provided $f: X \to R$ (= Reals) and f is $(T^{-1}(\mathcal{B}(Y)), \mathcal{B}(R))$ - measurable (as well as $(\mathcal{B}(X), \mathcal{B}(R))$ - measurable).

In [3], Halmos and Savage develop an approach to sufficient statistics. Their results provide an alternate definition, within a very general mathematical framework, of statistical sufficiency for dominated sets of measures. This alternate definition is particularly suitable to the development of the results in this paper. We will require the statement (Theorem 1.) of the alternate definition in the setting of Banach spaces.

In all that follows X and Y will be Banach spaces, T a linear continuous mapping of X onto Y, and $\mathcal{D} \subset \mathcal{M}(X)$ a dominated set of measures.

Theorem 1. (Halmos-Savage [3]) A necessary and sufficient condition that T be a sufficient statistic for $\mathcal{D}$ is that there exist $\lambda \in \mathcal{M}(X)$ such

that $\mathcal{D} \equiv \lambda$ and $g_\mu \in [d\mu/d\lambda]$ such that $g_\mu(\varepsilon)T^{-1}(\mathcal{B}(Y))$ for each $\mu \in \mathcal{D}$.

In this paper our particular concern will be that of developing necessary and sufficient conditions that a linear continuous mapping T of X onto Y be a sufficient statistic for a dominated set of measures $\mathcal{D} \subset \mathcal{M}(X)$.

In Theorem 2. we will require an additional condition on T which, to the best of our knowledge, is generally unavoidable . We will require that the kernel of T (= ker T) be <u>complemented</u>, in the sense that there exists a closed subspace S of X such that $X = \ker T \oplus S$ (e.g., if X is a Hilbert space, take $S = (\ker T)^{\perp}$).

In Theorem 4. we will show that the condition $X = \ker T \oplus S$ may be relaxed whenever $[d\mu/d\lambda]$ contains a continuous representative. The results we develop are finally used to establish necessary and sufficient conditions that a linear statistic $B:R^n \to R^k (k \leq n)$ be sufficient for a finite collection of probability measures having n-variate normal densities.

2. <u>Principal Results</u>. In all that follows we will assume that X and Y are Banach spaces, $T:X \to Y$ is a linear continuous mapping of X <u>onto</u> Y, and $\mathcal{D} \subset \mathcal{M}(X)$ is a dominated set of measures.

Theorem 2. Let $X = \ker T \oplus S$ for some closed subspace of X. A necessary and sufficient condition that T be a sufficient statistic for $\mathcal{D}$ is that there exist $\lambda \in \mathcal{M}(X)$ such that $\mathcal{D} \equiv \lambda$ and,

$$\ker T \subset \{y : g_\mu(x + y) = g_\mu(x),\ x \in X\}$$

for each $\mu \in \mathcal{D}$ and some $g_\mu \in [d\mu/d\lambda]$.

Proof. If T is a sufficient statistic for $\mathcal{D}$ and $\mu \in \mathcal{D}$ then there exists (Theorem 1 $\lambda \equiv \mathcal{D}$ and $g_\mu \in [d\mu/d\lambda]$ such that $g_\mu(\epsilon)T^{-1}(\mathcal{B}(Y)$. Suppose $y \in \ker T$ and, without loss of generality, there exists $x_0 \in X$ such that $g_\mu(x_0 + y) < g_\mu(x_0)$. Choose $r \in R$ such that $g_\mu(x_0 + y) < r < g_\mu(x_0)$. Since $g_\mu^{-1}(-\infty,r)$ and $g_\mu^{-1}(r,\infty)$ are elements of $\mathcal{B}(X)$ and $g_\mu(\epsilon)T^{-1}(\mathcal{B}(Y))$ it follows that there exist B_1 and $B_2 \in \mathcal{B}(Y)$ such that $x_0 + y \in g^{-1}(-\infty,r) = T^{-1}(B_1)$ and $x_0 \in g^{-1}(r,\infty) = T^{-1}(B_2)$. Now, since T is linear and $y \in \ker T$, $T(x_0) \in B_1 \cap B_2 = \phi$, which is absurd.

Conversely, suppose $\mathcal{D} \equiv \lambda$, $\mu \in \mathcal{D}$ and $\ker T \subset \{y: g_\mu(x + y) = g_\mu(x), x \in X\}$ for some $g_\mu \in [d\mu/d\lambda]$. We need only show (according to Theorem 1) that $g_\mu(\epsilon)T^{-1}(\mathcal{B}(Y)$. It will only be necessary to show that for $r \in R$ there exists $B_r \in \mathcal{B}(Y)$ such that $g_\mu^{-1}(-\infty,r) = T^{-1}(B_r)$. We will show first that $g_\mu^{-1}(-\infty,r) = T^{-1}\, T(g_\mu^{-1}(-\infty,r) \cap S)$ and then that $B_r \equiv T(g_\mu^{-1}(-\infty,r) \cap S) \in \mathcal{B}(Y)$.

If $x \in T^{-1}(T(g_\mu^{-1}(-\infty,r) \cap S)$ then $T(x) \in T(g_\mu^{-1}(-\infty,r) \cap S)$ and hence $T(x) = T(z)$ for some $z \in g_\mu^{-1}(-\infty,r) \cap S$. Since T is linear $x - z \in \ker T$ so that $g_\mu(x) = g_\mu(x - z + z) = g_\mu(z) < r$ and $x \in g_\mu^{-1}(-\infty,r)$.

If $x \in g_\mu^{-1}(-\infty,r)$ then, since $X = \ker T \oplus S$, $x = k + s$ for $k \in \ker T$ and $s \in S$. It follows that $T(x) = T(s)$, $s - x \in \ker T$, $g_\mu(s) = g_\mu(s - x + x) = g_\mu(x) < r$, $s \in g_\mu^{-1}(-\infty,r)$, $T(x) = T(s) \in T(g_\mu^{-1}(-\infty,r) \cap S)$ and, finally, that $x \in T^{-1}(T(g_\mu^{-1}(-\infty,r) \cap S))$.

We now show that $T(g_\mu^{-1}(-\infty,r) \cap S) \in \mathcal{B}(Y)$. Let $T_S: S \to Y$ be the restriction of T to S and observe that T_S is a one to one continuous

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

mapping of the Banach space S onto the Banach space Y. Since T_S satisfies the hypothesis of the open mapping theorem T_S is a homeomorphism of S onto Y. Since such mappings take elements of $\mathcal{B}(S)$ into elements of $\mathcal{B}(Y)$ and g_μ is measurable, $g_\mu^{-1}(-\infty,r) \cap S \in \mathcal{B}(X) \cap S = \mathcal{B}(S)$. It follows that $T(g_\mu^{-1}(-\infty,r) \cap S) = T_S(g_\mu^{-1}(-\infty,r) \cap S) \in \mathcal{B}(Y)$ and the proof of the theorem is complete.

Theorem 3. Let $\mathcal{D} \equiv \lambda$, $\lambda(B) = \lambda(B - y)$ for each $y \in \ker T$ and $B \in \mathcal{B}(X)$ such that $\lambda(B) = 0$, $\lambda(C) > 0$ for each non-empty open subset C of X and let $[d\mu/d\lambda]$ contain a continuous representative element f_μ for each $\mu \in \mathcal{D}$.

A necessary and sufficient condition that T be a sufficient statistic for $\mathcal{D}$ is that

$$\ker T \subset \{y : f_\mu(y + x) = f_\mu(x),\ x \in X\}$$

Proof: In order to see that the condition is sufficient we need only show (according to Theorem 1.) that $f_\mu(\varepsilon)T^{-1}(\mathcal{B}(Y))$, or equivalently, if $r \in R$ that $f_\mu^{-1}(-\infty,r) = T^{-1}(B_r)$ for some $B_r \in \mathcal{B}(Y)$. In fact, since T is an open mapping and f_μ is continuous, $T(f^{-1}(-\infty,r)) \in \mathcal{B}(Y)$. We take $B_r \equiv T(f^{-1}(-\infty,r))$ and conclude the argument by showing that $f_\mu^{-1}(-\infty,r) = T^{-1}T(f_\mu^{-1}(-\infty,r))$. We clearly need only establish that $T^{-1}T(f_\mu^{-1}(-\infty,r)) \subset f_\mu^{-1}(-\infty,r)$. If $x \in T^{-1}T(f_\mu^{-1}(-\infty,r))$ then $T(x) = T(z)$ for some $z \in f_\mu^{-1}(-\infty,r)$. Since $x - z \in \ker T$ it follows that $f_\mu(x) = f_\mu(x - z + z) = f_\mu(z) < r$ and hence that $x \in f_\mu^{-1}(-\infty,r)$.

In order to prove the necessity of the condition, recall the proof of the necessity of the condition in Theorem 2. and observe that the hypothesis $X = \ker T \oplus S$ for some closed subspace S of X was not essential. We may conclude that if $\mu \in \mathcal{D}$ there exists $g_\mu \in [d\mu/d\lambda]$ such that $\ker T \subset \{y : g_\mu(y + x) = g_\mu(x),\ x \in X\}$ and $f_\mu = g_\mu$ except on a set $B \in \mathcal{B}(X)$ such that $\lambda(B) = 0$.

Fix $y \in \ker T$. Since $\{x : f_\mu(y + x) \neq g_\mu(y + x)\} = B - y$ and $\lambda(B - y) = \lambda(B) = 0$, we may conclude that $f_\mu(x) = f_\mu(y + x)$ except on $C = B \cup (B - y)$ and $\lambda(C) = 0$. Moreover, since the mapping $x \to y + x$ is a homeomorphism of X onto X and f_μ is continuous, C is an open subset of X. According to the hypothesis, $\lambda(C) = 0$ and C open imply C is empty so that $f_\mu(y + x) = f_\mu(x)$ for each $x \in X$.

3. <u>Normal Families</u>. In what follows we will assume that $\mathcal{D} = \{P_i\}_{i=0}^{m-1}$ is a family of m probability measures defined on $\mathcal{B}(R^n)$ having normal densities

$$p_i(x) = (2\pi)^{-n/2}|\Omega_i|^{-1/2} \exp[-\tfrac{1}{2}(x - \eta_i)^T\Omega_i^{-1}(x - \eta_i]; \quad i = 0, 1, \ldots, m-1.$$

where η_i and Ω_i are known and Ω_i is symmetric and positive definite.

We will derive necessary and sufficient conditions that a $k \times n$ matrix B $(k \leq n)$ mapping R^n <u>onto</u> R^k (i.e., rank $(B) = k$) be a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$. We first prove a Lemma.

Lemma 1. If $1 \leq i \leq m - 1$ and $f_i(x) = p_i(x)/p_0(x)$ then

$$\{y : f_i(y + x) = f_i(x),\ x \in X\} = \ker(\Omega_i^{-1} - \Omega_0^{-1}) \cap \{\Omega_i^{-1}\eta_i - \Omega_0^{-1}\eta_0\}^\perp .$$

Proof: Fix $y \in R^n$. After a little matrix algebra (which we will omit) we find that $f_i(y + x) = f_i(x)$ for each $x \in R^n$ if and only if

$$2x^T(\Omega_i^{-1} - \Omega_0^{-1})y - 2y^T(\Omega_i^{-1}\eta_i - \Omega_0^{-1}\eta_0) + y^T(\Omega_i^{-1} - \Omega_0^{-1})y = 0$$

for each $x \in R^n$. For $x = -y/2$ we see that $y^T(\Omega_i^{-1}\eta_i - \Omega_0^{-1}\eta_0) = 0$ so that $y \in \{\Omega_i^{-1}\eta_0 - \Omega_0^{-1}\eta_0\}^\perp$. In addition, it follows that $2x^T(\Omega_i^{-1} - \Omega_0^{-1})y + y^T(\Omega_i^{-1} - \Omega_0^{-1})y = 0$ and, writing $x = (z - y)/2$, that $z^T(\Omega_i^{-1} - \Omega_0^{-1})y = 0$ for each $z \in X$. This clearly implies $(\Omega_i^{-1} - \Omega_0^{-1})y = \Theta$ so that $y \in \ker(\Omega_i^{-1} - \Omega_0^{-1})$. The remaining containment follows easily.

Theorem 4. A necessary and sufficient condition that a $k \times n$ rank k matrix B be a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$ is that

$$\ker B \subset \bigcap_{i=1}^{m-1} [\ker(\Omega_i^{-1} - \Omega_0^{-1}) \cap \{\Omega_i^{-1}\eta_i - \Omega_0^{-1}\eta_0\}^\perp] .$$

Proof: Since the preliminary conditions of Theorem 3 are clearly satisfied for $\lambda = P_0$, Lemma 1 insures the necessity and sufficiency of the condition.

Theorem 5. A necessary and sufficient condition that a $k \times n$ rank k matrix B be a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$ is that, for $j = 1, \ldots, m - 1$,

(a) $\Omega_j B^T(B\Omega_j B^T)^{-1} = \Omega_0 B^T(B\Omega_0 B^T)^{-1}$

(b) $\eta_j - \Omega_j B^T(B\Omega_j B^T)^{-1}B\eta_j = \eta_0 - \Omega_0 B^T(B\Omega_0 B^T)^{-1}B\eta_0$

(c) $\Omega_j - \Omega_j B^T(B\Omega_j B^T)^{-1}B\Omega_j = \Omega_0 - \Omega_0 B^T(B\Omega_0 B^T)^{-1}B\Omega_0$.

Proof: Let $(x|y) = x^T y$ and $(x|y)_i = x^T \Omega_i^{-1} y$ $i = 0, 1, \ldots, m-1$. For $S \subset R^n$, $S^\perp$ and $S^{\perp i}$ will denote, respectively, the orthogonal complements of S relative to the inner products $(\cdot|\cdot)$ and $(\cdot|\cdot)_i$.

If A is an $n \times n$ matrix A^{*i} will denote the adjoint of A relative to the inner product $(\cdot|\cdot)_i$ on R^n. If A is a $k \times n$ matrix A^{*i} will denote the adjoint of A relative to the inner products $(\cdot|\cdot)_i$ on R^n and $(\cdot|\cdot)$ on R^k. It follows that $B^{*i} = \Omega_i B^T$.

If B is a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$ then, according to Theorem 3., $\ker B \subset \ker(\Omega_j^{-1} - \Omega_0^{-1})$; $j = 1, \ldots, m-1$ and hence $(\ker B)^{\perp j} = (\ker B)^{\perp 0}$. Since this implies range (B^{*j}) = range (B^{*0}) we have that $B^{*0}(BB^{*0})^{-1}BB^{*j} = B^{*j}$ and hence that $\Omega_j B^T(B\Omega_j B^T)^{-1} = \Omega_0 B^T(B\Omega_0 B^T)^{-1}$ which is (a).

Now let $Q = \Omega_0 B^T(B\Omega_0 B^T)^{-1}B$ and observe that $Q^{*j} = Q = Q^2$ for $j = 1, \ldots, m-1$. It follows that $\ker Q = \ker B \subset \ker(\Omega_j^{-1} - \Omega_0^{-1})$ and that $Q(\Omega_j^{-1} - \Omega_0^{-1})^{*0} = (\Omega_j^{-1} - \Omega_0)^{*0}$ and hence that $Q(\Omega_j - \Omega_0) = \Omega_j - \Omega_0$ which, recalling the definition of Q, is equivalent to (c).

Since $\ker(\Omega_j^{-1} - \Omega_0^{-1}) \cap (\Omega_j^{-1}\eta_n - \Omega_0^{-1}\eta_0) \subset (\eta_j - \eta_0)^{\perp j}$ and $\eta_j - \eta_0 \in (\ker B)^{\perp j}$ = range (B^{*j}) = range (Q), it follows that $Q(\eta_j - \eta_0) = \eta_j - \eta_0$ which, recalling the definiton of Q, is equivalent to (b).

Since all of the preceeding arguments are reversible, (a), (b) and (c) imply B is a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$, completing the proof of the theorem.

In the next theorem we will use the fact that there exists a non singular matrix M such that $M\Omega_0 M^T = I$ and hence that the affine transform-

ation $x \to Mx - \eta_0$ provides a change of variables that allows (without loss of generality or the ability to recover the sufficient statistic relative to the original variables) one to assume that $\eta_0 = 0$ and $\Omega_0 = I$.

Theorem 6. If $\eta_0 = 0$ and $\Omega_0 = I$ then a necessary and sufficient condition that a $k \times n$ rank k matrix B be sufficient for $\{P_i\}_{i=0}^{m-1}$ is that there exist a rank k orthogonal projection Q such that, for $i = 1, \ldots, m-1$,

$$(I - Q)[\eta_1|\eta_2| \ldots |\eta_{m-1}|\ \Omega_1 - I\ |\Omega_2 - I| \ldots |\Omega_{m-1} - I] = Z$$

where Z is the $n \times (n+1)(m-1)$ zero matrix.

Proof: If B is a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$, we may assume without loss of generality that $BB^T = I$ since B is a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$ if and only if KB is a sufficient statistic for each nonsingular $k \times k$ matrix K. One may indeed choose K such that $KBB^TK^T = (KB)(KB)^T = I$.

For $i = 1, \ldots, m-1$ Theorem 5. implies that

$$\Omega_i B^T(B\Omega_i B^T)^{-1} = I\ B^T(B\ I B^T)^{-1} = B^T$$

so that

$$(B\Omega_i B^T)^{-1} = B\Omega_i^{-1}B^T \quad \text{and} \quad \Omega_i B^T(B\Omega_i B^T)^{-1}B = B^TB\ .$$

Right multiplication of the latter equation by $\Omega_i B^T B$ will establish that

$$\Omega_i B^T B = B^T B\Omega_i B^T B$$

from whence it follows, using symmetry, that

$$\Omega_i B^T B = B^T B\Omega_i \quad .$$

Since $\eta_1 = \Theta$ and $\Omega_1 = I$, Theorem 5 further implies

$$\eta_1 - B^T B = \Theta$$

and

$$\Omega_1 - B^T B \Omega_1 = I - B^T B$$

Since $BB^T = I$, it follows that $B^T = B^+$ (where $(\cdot)^+$ denotes the generalized inverse of $(\cdot)$) and hence that $Q \equiv B^T B = B^+ B$ is the orthogonal projection on the range of B^T [5]. Clearly Q has rank k and we conclude that

$$(I - Q)\eta_1 = \Theta$$

and

$$(I - Q)(\Omega_1 - I) = Z$$

and the condition follows. Conversely, if the conditon holds let B be any $k \times n$ rank k matrix such that range (B^T) = range (Q). Clearly $B^+ B = Q$, $BB^+ = I$ and $B^+ = B^T$. Using the symmetry of $I - Q$ and $\Omega_1 - I$ we conclude that

$$\Omega_1 B^T B = B^T B \Omega_1$$

and hence that

$$Q = B^+ B = B^+ B \Omega_1 B^T (B \Omega_1 B^T)^{-1} B = \Omega_1 B^+ B B^T (B \Omega_1 B^T)^{-1} B$$
$$= \Omega_1 B^T (B \Omega_1 B^T)^{-1} B \ .$$

In addition,

$$\Omega_1 B^T (B \Omega_1 B^T)^{-1} = B^T$$

The obvious substitution for Q guarantees the satisfaction of the conditions of Theorem 5.

Definition 1. We will say that a rank k orthogonal projection Q <u>generates a sufficient statistic for</u> $\{P_i\}_{i=0}^{m-1}$ provided Q satisfies the condition in Theorem 6.

Corollary 1. If $M = [\eta_1 | \eta_2 | \ldots | \eta_{m-1} | \Omega_1 - I | \ldots | \Omega_{m-1} - I]$ then

a) $Q = MM^+$ generates a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$

and

b) $k = \text{rank}\,(MM^+) \equiv \text{tr}\,(MM^+)$ is the smallest integer for which there exists a rank k orthogonal projection generating a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$.

Proof: Let k be the smallest integer for which there exists a rank k orthogonal projection P generating a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$.

According to the definition of M, $(I - P)M = Z$ so that $PM = M$ and $PMM^+ = MM^+$. Since $(I - MM^+)M = Z$, MM^+ generates a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$. However, $PMM^+ = MM^+$ implies that range $(MM^+) \subset$ range (P) so that the minimality of k and the fact that MM^+ is an orthogonal projection imply that range (MM^+) = range (P) and hence that $MM^+ = P$.

Corollary 2. If B is a sufficient statistic for $\{P_i\}_{i=0}^{m-1}$ then

$$(B\Omega_i B^T)^{-1} = B\Omega_i^{-1}B^T \quad i = 0, 1, \ldots, m-1 .$$

Proof: The conclusion is an immediate consequence of line 6 in the proof of Theorem 6.

4. Concluding Remarks. Theorems 4 and 5, although not so stated, are valid for arbitrary families of n-variate normal probability measures. Corollary 1. formally gives the construction for a sufficient statistic for finite families of n-variate normal probability measures solely in terms of the known parameters that determine the densities. In fact, if k=rank (M) (=rank MM^+) then any rank k matrix B for which range (B)=range (M) is a sufficient statistic for the family. Moreover, in terms of the dimension of the range of a sufficient statistic, k=rank M is the smallest integer for which there exists a sufficient statistic.

Several open questions concerning the "appropriate" definition of a "almost" sufficient statistic using the characterizations given in Theorems 4. and 5. will be the subject of a later paper. In this connection the results of Le Cam [4], although the approach is different, should be of significant value.

5. Acknowledgement. The authors would like to express there sincere appreciation to Professor H. Elton Lacey for his comments.

REFERENCES

1. Anderson, T. W. (1958) An Introduction to Multivariate Statistical Analysis. Wiley, New York.

2. Bahadur, R. R. (1954) Sufficiency and statistical decision function. Ann. Mathe. Statist. 25 423-463.

3. Halmos, P. R. and Savage, L. J. (1949) Application of the Radon-Nikodym theorem to the theory of sufficient statistics. Ann. Math. Statist. 20 225-241.

4. Le Cam, L. (1964) Sufficiency and approximate sufficiency. Ann. Math. Statist. 35 1419-1455.

5. Rao, C. R. and Mitra, S. K. (1971) Generalized Inverse of Matrices and its Applications. Wiley, New York.