## 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.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON HOUSTON, TEXAS





# CHARACTERIZATIONS OF LINEAR SUFFICIENT STATISTICS

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

REPORT #65 OCTOBER 1977



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]

We develop necessary and sufficient conditions that a surjective bounded linear operator $T$ from a Banach space $X$ to a Banach space $Y$ be a sufficient statistic for a dominated family of probability measures defined on the Borel sets of $X$. We give applications of these results that characterize linear sufficient statistics for families of the exponential type, including as special cases the Wishart and multivariate normal distributions. The latter result is used to establish precisely which procedures for sampling from a normal population have the property that the sample mean is a sufficient statistic.

---

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

1. Introduction: Let $T$ be a surjective measureable transformation from the measureable space $(X,A)$ to the measureable space $(Y,B)$, and let $\mathcal{D}$ be a set of totally finite measures on $A$. Following Halmos and Savage [2], we say that $T$ is a sufficient statistic relative to $\mathcal{D}$ if for each $E \in A$ there exists a measureable function $P(E|\cdot) : (Y,B) \to R$ (the real numbers) such that for each $F \in B$, $\mu \in \mathcal{D}$

$$\mu(E \cap T^{-1}(F)) = \int_F P(E|y)d\mu T^{-1}(y) .$$

In another nonequivalent definition of a sufficient statistic given by Lehmann and Scheffe´[3], $B$ is always taken to be $B_T$, the largest $\sigma$-field on $Y$ consistent with the measureability of $T$. Bahadur [1] discusses the relationship between these two definitions at length.

In this paper our particular concern is that of developing necessary and sufficient conditions that a surjective bounded linear operator $T$ from a Banach space $X$ to a Banach space $Y$ be a sufficient statistic, where $A$ and $B$ are the respective Borel fields of $X$ and $Y$. Our first theorem shows that under a very natural condition the aforementioned definitions of sufficiency are equivalent. Specifically, the condition is that $\ker T = \{x \in X | Tx = \theta\}$ be complemented in $X$; that is, for some closed subspace $S$ of $X$, $X = \ker T \oplus S$. (For example, if $X$ is a Hilbert space, take $S = (\ker T)^{\perp}$.) As a corollary we obtain a simple characterization of sufficient linear statistics for

dominated sets of measures. In Theorem 2 we replace the condition that $\ker T$ be complemented with conditions on the density functions corresponding to a dominated set $\mathcal{D}$. Finally, we give applications of these results that characterize linear sufficient statistics for families of the exponential type, including as special cases the Wishart and multivariate normal distributions. The latter result is used to establish precisely which procedures for sampling from a normal population have the property that the sample mean is a sufficient statistic. This generalizes the classical result that the sample mean is sufficient for independent samples. The final result deals with the connection between linear sufficient statistics and the Gauss-Markov theorem.

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 $M(W)$. We will write $\mu << \nu$ for the relation of absolute continuity and $d\mu/d\nu$ for the equivalence class of Radon-Nikodym derivatives of $\mu$ with respect to $\nu$. For the definitions of a dominated set of measures, equivalent sets of measures, and their connection with $\sigma$-finite measures defined on $\mathcal{B}(W)$, we refer the reader to Halmos and Savage [2].

2. Principal Results: Our first theorem shows that if $\ker T$ is complemented in $S$ then, the two definitions of sufficiency described in the introduction are equivalent.

Theorem 1: Let $X$ and $Y$ be Banach spaces, let $A = \mathcal{B}(X)$ and let $T$ be a surjective bounded linear operator from $X$ to $Y$ such that

ker T is complemented in X . Then $\mathcal{B}_T = \mathcal{B}(Y)$ .

Proof: Since T is Borel measureable, it suffices to show that $\mathcal{B}_T \subset \mathcal{B}(Y)$ . Let S be a closed subspace of X such that $X = \ker T \oplus S$ . If $F \in \mathcal{B}_T$ , then $T^{-1}(F) \in \mathcal{B}(X)$ and if $\hat{T}$ denotes the restriction of T to S , then $\hat{T}^{-1}(F) = T^{-1}(F) \cap S \in \mathcal{B}(X)$ . It follows that $\hat{T}^{-1}(F) \in \mathcal{B}(S)$ , and since $\hat{T}$ is a topological isomorphism, $F = \hat{T}\hat{T}^{-1}(F) \in \mathcal{B}(Y)$ .

Henceforth, we will assume that X and Y are Banach spaces; $\mathcal{A} = \mathcal{B}(X)$ , $\mathcal{B} = \mathcal{B}(Y)$ and $T:(X,\mathcal{A}) \to (Y,\mathcal{B})$ is a surjective bounded linear operator. According to [2, Lemma 7], for a dominated collection of measures $\mathcal{D} \subset M(X)$ a measure $\lambda$ , equivalent to $\mathcal{D}$ , can be defined by

$$\lambda(E) \equiv \sum_{i=1}^{\infty} a_i \mu_i(E)$$

where $\{\mu_i\}_{i=1}^{\infty}$ is a countable subset of $\mathcal{D}$ which is equivalent to $\mathcal{D}$ and $\sum_{i=1}^{\infty} a_i \mu_i(X) < \infty$ . Obviously, if $\mathcal{D}$ is homogeneous, we can take $\lambda \in \mathcal{D}$ . Combining the results of Theorem 1 with those of Lemma 2 and Theorem 1 of [2], we have:

Theorem 2: If ker T is complemented in X , then T is sufficient for $\mathcal{D}$ if and only if for each $\mu \in \mathcal{D}$ there exists a real valued function $g_\mu$ on Y such that $g_{\mu 0} T \in d\mu/d\lambda$ .

Proof: By Theorem 1 of [2], T is sufficient if and only if for each $\mu \in \mathcal{D}$ there exists a real valued Borel measureable function $g_\mu$ on Y such that $g_{\mu 0} T \in d\mu/d\lambda$ . Since ker T is complemented in X , $\mathcal{B}(Y) = \mathcal{B}_T$ and each real valued function $g_\mu$ such that

$g_{\mu 0} T$ is Borel measureable on $X$ must be Borel measureable on $Y$.

In all that follows $\delta g(x;z)$ will denote the Gateaux differential of the function $g$ at $x$ in the direction of $z$.

Corollary 1: If $\ker T$ is complemented in $X$, then $T$ is sufficient for $\mathcal{D}$ if and only if for each $\mu \in \mathcal{D}$ there exists $f_\mu \in d\mu/d\lambda$ such that $x \in X$ and $y \in \ker T$ implies $\delta f_\mu(x;y) = 0$.

Proof: If $T$ is sufficient, then for each $\mu \in \mathcal{D}$ there exists $g_\mu : Y \to R$ such that $f_\mu = g_{\mu 0} T \in d\mu/d\lambda$. It follows immediately that $\delta f_\mu(x;y) = 0$ for each $x \in X$, $y \in \ker T$.

If $f_\mu \in d\mu/d\lambda$ and $\delta f_\mu(x;y) = 0$ for $\mu \in \mathcal{D}$, $x \in X$, $y \in \ker T$, then $f_\mu(x+y) = f_\mu(x)$ for each $x \in X$, $y \in \ker T$. For $z \in Y$ define $g_\mu(z) = f_\mu(x)$ where $z = Tx$. Then $g_\mu$ is well defined and $f_\mu = g_{\mu 0} T$. Hence, $T$ is sufficient.

The next theorem concerns a replacement of the complemented kernel condition whenever there is a continuous Radon-Nikodym derivative $f_\mu \in d\mu/d\lambda$ for each $\mu \in \mathcal{D}$.

Theorem 3: Let $V \subset X$ be an open set such that $\lambda(X \sim V) = 0$ and let $\lambda(U) > 0$ for each nonempty open subset $U$ of $V$. Suppose $\lambda(B+y) = 0$ whenever $B \subset V$, $\lambda(B) = 0$ and $y \in \ker T$. For each $\mu \in \mathcal{D}$, let $f_\mu \in d\mu/d\lambda$ be continuous on $V$. Then $T$ is sufficient if and only if $f_\mu(x) = f_\mu(z)$ whenever $x, z \in V$ and $Tx = Tz$.

Proof: If $T$ is a sufficient statistic, then there exists $g_\mu \in d\mu/d\lambda$ such that $g_\mu(x) = g_\mu(z)$ whenever $x, z \in V$, $Tx = Tz$. Let $\mu \in \mathcal{D}$ and $y \in \ker T$ be fixed. The set

$$U = \{x \in V \cap (V-y) \mid f_\mu(x) \neq f_\mu(x+y)\}$$

is an open subset of $V$ contained in $B \cup (B-y)$, where

$$B = \{x \in V \mid f_\mu(x) \neq g_\mu(x)\} .$$

Since $\lambda(B) = 0$, it follows from the hypothesis that $\lambda(U) = 0$ and hence, $U = \emptyset$. Thus $f_\mu(x) = f_\mu(x+y)$ whenever $x, x+y \in V$.

Conversely, suppose $f_\mu(x) = f_\mu(z)$ for $\mu \in \mathcal{D}$, $x, z \in V$ whenever $Tx = Tz$. The function $g_\mu : T(V) \to R$ defined by $g_\mu(Tx) = f_\mu(x)$ for $x \in V$ is well defined on $T(V)$. Since $f_\mu$ is continuous on $V$, $f_\mu = g_\mu \circ T$ on $V$, and $T$ is an open mapping, it follows that $g_\mu$ is continuous on the open set $T(V)$. For $y \notin T(V)$ define $g_\mu(y) = 0$. Then $g_\mu$ is Borel measureable on $Y$ and $f_\mu = g_\mu \circ T$. Thus $T$ is sufficient for $\mathcal{D}$.

The proof of the following corollary is clear and will be omitted.

Corollary 2: If, in addition to the hypotheses of Theorem 4, the set $V$ is convex, then $T$ is sufficient for $\mathcal{D}$ if and only if $\delta f_\mu(x;y) = 0$ for each $\mu \in \mathcal{D}$, $x \in V$, $y \in \ker T$.

3. Exponential Families: Let $X$ and $Y$ be Banach spaces, $(H, \langle \cdot | \cdot \rangle)$ a Hilbert space and $\nu$ a $\sigma$-finite measure on $B(X)$ such that $\nu(X \sim V) = 0$ for some nonempty open convex set $V \subset X$ for which $\nu(U) > 0$ for each nonempty open set $U \subset V$. Let $\mathcal{D} = \{\mu_\gamma\}$, $\gamma \in \Gamma$ be a family of probability measures having exponential densities $f_\gamma(x) = c(\gamma)h(x) \exp \langle Q(\gamma) | t(x) \rangle \in d\mu_\gamma / d\nu$ where $c(\gamma) > 0$, $h(x) > 0$ on $V$ a.e.$(\nu)$, $t : X \to H$ is continuous

and Gateaux differentiable on $V$, and $Q:\Gamma \to H$.

Theorem 4. Let $T:X \to Y$ be linear, bounded, surjective and $\nu(B+y) = 0$ whenever $B \in \mathcal{B}(X)$, $B \subset V$, $\nu(B) = 0$ and $y \in \ker T$. If $\beta \in \Gamma$, $T$ is a sufficient statistic for the exponential family $\mathcal{D}$ if and only if $\langle Q(\gamma) - Q(\beta) | \delta t(x;y) \rangle = 0$ for each $\gamma \in \Gamma$, $x \in X$ and $y \in \ker T$.

Proof: Under the stated assumptions $\mathcal{D}$ is homogeneous and thus $\lambda$ may be taken to be an arbitrary element, say $\mu_\beta$, of $\mathcal{D}$. Applying Corollary 2, $T$ is sufficient for $\mathcal{D}$ if and only if $\delta g_{\gamma,\beta}(x;y) = \theta$ for each $\gamma \in \Gamma$, $x \in V$ $y \in \ker T$, where

$$g_{\gamma,\beta}(x) = \frac{c(\gamma)}{c(\beta)} \exp \{ \langle Q(\gamma) - Q(\beta) | t(x) \rangle \} .$$

This is equivalent to $\langle Q(\gamma) - Q(\beta) | \delta t(x;y) \rangle = 0$ for each $\gamma \in \Gamma$, $x \in V$, $y \in \ker T$.

4. Applications. Let $S$ denote the symmetric $n \times n$ matrices, $\Gamma$ a subset of the positive definite elements of $S$ and $\mathcal{D}$ a family of Wishart probability measures with $m \geq n$ degrees of freedom having densities

$$f_\gamma(S) = c(\gamma)|S|^{(m-n-1)/2} \exp \{ -\tfrac{1}{2} \operatorname{tr} (\gamma^{-1} S) \} .$$

Theorem 5. If $\beta \in \Gamma$ and $T:S \to \text{range } (T)$ is linear, then $T$ is a sufficient statistic for the Wishart family $\mathcal{D}$ if and only if $\operatorname{tr}[(\gamma^{-1} - \beta^{-1})K] = 0$ for each $\gamma \in \Gamma$ and $K \in \ker T$.

Proof. The preliminary conditions of Theorem 4. are satisfied with $\gamma$ = Lebesgue measure on $S$ and the obvious identifications of $c(\gamma)$

and $h(S)$ . Let $H$ equal $S$ with $\langle A|B\rangle \equiv tr(AB)$ , $t(S) = S$ and $Q(\gamma) = -\gamma^{-1}/2$ . Observe that $\delta t(S;F) = F$ and apply Theorem 4.

Remark: Theorem 5. implies that there is a nontrivial linear sufficient statistic if and only if there exists a linear manifold $M \subsetneq S$ such that $\gamma^{-1} \in M$ for each $\gamma \in \Gamma$ .

We will now apply these results to normal families of probability measures. In Theorem 6. we will state set theoretical, algebraic and geometrical conditions, each equivalent to the condition that $T$ be a linear sufficient statistic for a family $\mathcal{D} = \{P_\gamma\}$ , $\gamma \in \Gamma$ of normal n-variate probability measures having densities, with respect to Lebesgue measure on $R^n$ ,

$$p_\gamma(x) = (2\pi)^{-n/2}|\Omega_\gamma|^{-1/2} \exp\left[-\frac{1}{2}(x-\eta_\gamma)'\Omega_\gamma^{-1}(x-\eta_\gamma)\right]$$

We will assume that for some $\beta \in \Gamma$ , $\eta_\beta = \theta$ and $\Omega_\beta = I$ . This requirement imposes no loss of generality since for any $\beta \in \Gamma$ there exists a non singular matrix $M_\beta$ for which $M_\beta \Omega_\beta M_\beta' = I$ and a change of coordinate system defined by the transformation $x \to M_\beta(x-\eta_\beta)$ allows one to recover the sufficient statistic in the original coordinate system.

Theorem 6. If $T:R^n \to R^k$ is a linear transformation of rank $k$ and $\mathcal{D} = \{P_\gamma\}$ , $\gamma \in \Gamma$ is an arbitrary family of n-variate normal probability measures such that for some $\beta \in \Gamma$ , $\eta_\beta = \theta$ and $\Omega_\beta = I$ then the following conditions are equivalent:

(1) $T$ is sufficient for $\mathcal{P} = \{P_\gamma\}$, $\gamma \in \Gamma$.

(2) $\ker T \subset \bigcap_{\gamma\in\Gamma} [\ker(\Omega_\gamma - I) \cap [\eta_\gamma]^\perp]$

(3) For each $\gamma \in \Gamma$,

(a) $T^+T\eta_\gamma = \eta_\gamma$

(b) $T^+T(\Omega_\gamma - I) = \Omega_\gamma - I$

where the notation $(\cdot)^+$ denotes the generalized inverse of $(\cdot)$.

Proof: To see that (1) $\rightarrow$ (2) observe that the preliminary conditions of Theorem 4. are satisfied with $\nu$ = Lebesgue measure on $X = R^n$. Make the obvious identifications for $c(\gamma)$ and $h(x)$. Let $M_n$ denote the $n \times n$ real matrices and define $Q:\Gamma \rightarrow H = M_n \times R^n \times M_n$, $t:X \rightarrow H$ and $\langle\cdot|\cdot\rangle$ on $H$, respectively, by $Q(\gamma) = (-\Omega_\gamma^{-1}/2,\ \Omega_\gamma^{-1}\eta_\gamma,\ -\Omega_\gamma^{-1}\eta_\gamma\eta_\gamma'/2)$, $t(x) = (xx', x, I)$ and $\langle(A_1,w_1,B_1)|(A_2,w_2,B_2)\rangle = \mathrm{tr}(A_1'A_2) + w_1'w_2 + \mathrm{tr}(B_1'B_2)$.

Since $Q$, $t$ and $\langle\cdot|\cdot\rangle$ satisfy the remaining hypotheses of Theorem 4. and $\delta t(x;z) = (xz' + z x',\ z, \theta)$ for each $x, z \in R^n$, it follows that for each $\gamma \in \Gamma$;

$$\ker T \subset \{y \in R^n : x'(\Omega_\gamma^{-1} - I)y - y'\Omega_\gamma^{-1}\eta_\gamma = 0,\ x \in R^n\}$$
$$= \ker(\Omega_\gamma^{-1} - I) \cap [\Omega_\gamma^{-1}\eta_\gamma]^\perp = \ker(\Omega_\gamma - I) \cap [\eta_\gamma]^\perp.$$

To see that (2) $\rightarrow$ (3) note that $T^+T$ is the orthogonal projection on range $(T') = (\ker T)^\perp$. Since $\eta_\gamma \in (\ker T)^\perp$, (3a) holds. Furthermore, $\ker T^+T = \ker T \subset \ker(\Omega_\gamma - I)$ implies range $(\Omega_\gamma - I) \subset$ range $(T^+T)$ and hence that $T^+T(\Omega_\gamma - I) = (\Omega_\gamma - I)$ which is (3b).

In order to see that (3) $\rightarrow$ (1) recall the definition of $Q(\gamma)$, $t(x)$ and the fact that $\delta t(x;z) = (xz' + z x',\ z,\ \theta)$.

We need only show that $x'(\Omega_\gamma - I)y - \eta_\gamma' y = 0$ for each $\gamma \in \Gamma$, $x \in X$ and $y \in \ker T$. Using (3b) and symmetry together with (3a) it follows that

$$x'(\Omega_\gamma - I)y - \eta_\gamma' y = x'(\Omega_\gamma - I)T^+(Ty) - \eta_\gamma' T^+(Ty) = 0 .$$

We state the following corollary without proof.

Corollary 3. Under the hypotheses of Theorem 6., there exists a $k \times n$ rank $k$ sufficient statistic for $\{P_\gamma\}$, $\gamma \in \Gamma$ if and only if there exists a rank $k$ orthogonal projection $P$ on $R^n$ such that (a) $P\eta_\gamma = \eta_\gamma$ and (b) $P(\Omega_\gamma - I) = \Omega_\gamma - I$ for each $\gamma \in \Gamma$. Moreover, any $k \times n$ rank $k$ matrix such that $T^+T = P$ is a sufficient statistic for $\{P_\gamma\}$, $\gamma \in \Gamma$.

Corollary 4. If $\Gamma = \{0, 1, \cdots, m-1\}$, $\eta_0 = \theta$, $\Omega_0 = I$ and $B \equiv [\eta_1 | \eta_2 | \cdots | \eta_{m-1} | \Omega_1 - I | \Omega_2 - I | \cdots | \Omega_{m-1} - I]$ then $T$ is a linear sufficient statistic for the finite family $\{P_\gamma\}$, $\gamma \in \Gamma$ of $n$-variate normal probability measures if and only if range $(T') \subset$ range $(B)$. Moreover, $k = \text{rank } B$ is the smallest integer for which there exists a $k \times n$ sufficient statistic for $\{P_\gamma\}$, $\gamma \in \Gamma$.

Proof: The equivalent condition is an immediate consequence of Theorem 6. The minimality statement follows from the fact that if $T$ is a $p \times n$ rank $p$ sufficient statistic then $T^+TB = B$, hence, $T^+TBB^+ = BB^+$. It follows that range $(BB^+) \subset$ range $(T^+T)$ and, since $(BB^+)B = B$, $BB^+$ satisfies Theorem 6.(3) so that $k = p$.

Example 1. Let $x_1, x_2, \cdots, x_n, \cdots$ be a sequence of univariate $N(\mu,\sigma)$ variables such that the joint density of $x_1, x_2, \cdots, x_n$ is $N(\mu\xi_n, \Omega_n)$ where $\xi_n' = (1,1,\cdots,1)$. Let $\{P_\mu\}$, $\mu \in R$ be the family of probability measures having densities $N(\mu\xi_n, \Omega_n)$ and $T \neq 0$ a $1 \times n$ matrix.

Observe that $T$ is sufficient for $\{P_\mu\}$, $\mu \in R$ if and only if $T\Omega_n^{1/2}$ is sufficient for the family of probability measures $\{\hat{P}_\mu\}$, $\mu \in R$ having densities $N(\mu\Omega_n^{-1/2}\xi_n, I)$ and, according to Theorem 6., that this is equivalent to the condition that $\ker T\,\Omega_n^{1/2} \subset [\Omega_n^{-1/2}\xi_n]^{\perp}$. This is equivalent to $\xi_n' = \alpha_n T\Omega_n$ for some scalar $\alpha_n$. A simple calculation shows that $\alpha_n = n(T\Omega_n\xi_n)^{-1}$ so that the statistic $T$ is sufficient for $\{P_\mu\}$, $\mu \in R$ if and only if $T = [(T\Omega_n\xi_n)^{-1}\xi_n'\Omega_n^{-1}]/n$. In particular, note that $T = \hat{T} \equiv (\xi_n'\Omega_n^{-1}\xi_n)^{-1}\xi_n'\Omega_n^{-1}$ is sufficient for $\{P_\mu\}$, $\mu \in R$ and the $\hat{T}(x_1,\cdots,x_n)'$ is an unbiased estimate of $\mu$ for each integer $n$. This generalizes the classical result that the sample mean is a sufficient statistic for $\mu$ when the samples $x_1, x_2, \cdots$ are independent.

Further note that if $T \equiv \xi_n'/n$ (the statistic $T$ for the sample mean) is a sufficient statistic for $\{P_\mu\}$, $\mu \in R$ for each integer $n$, the column sums (row sums) of $\Omega_n$ are identically $\alpha_n = (\xi_n'\Omega_n\xi)/n$. A routine induction argument shows that, in the latter case, Cov $(x_i, x_j)$ = constant for $i, j = 1,2,\cdots$, $i \neq j$.

Example 2. Let $y = W\gamma + \varepsilon$, where $W$ is a fixed $m \times n$ matrix of rank $n$ and $\varepsilon \sim N(0,I)$. According to the Gauss-Markov theorem, the minimum variance unbiased linear estimate of $\gamma$ is $\hat{\gamma} = (W'W)^{-1}W'y$

Let $T = (W'W)^{-1}W'$ and observe that for $\gamma \in R^n$, $T'(TT')^{-1}T\,W\gamma = W\gamma$ and, since $T'(TT')^{-}T = T^{+}T$, Theorem 6. implies $T$ is a sufficient statistic for the set of probability measures $(P_\gamma)$, $\gamma \in R^n$ having densities $N(W\gamma, I)$.

On the other hand, if $\hat{T}$ is a sufficient linear statistic for $\{P_\gamma\}$, $\gamma \in R^n$ such that $\hat{T}y$ is an unbiased estimate of $\gamma$ then, since $\hat{T}W = I$, $\hat{T}$ has rank $n$. Corollary 4 implies that $n$ is the smallest integer for which there exists a linear $n \times m$ sufficient statistic for $\{P_\gamma\}$, $\gamma \in R^n$. Moreover, $\hat{T} = B(W'W)^{-1}W'$ for some nonsingular $n \times n$ matrix $B$. Since $\hat{T}W = I$, $\hat{T} = (W'W)^{-1}W'$.

Since $\hat{\gamma} = Ty$, the Gauss-Markov estimate of $\gamma$ may be characterized as the unique linear sufficient statistic $T$ for $\{P_\gamma\}$, $\gamma \in R^n$ for which $Ty$ is an unbiased estimate of $\gamma$.

## REFERENCES


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

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

3. Lehmann and Scheffe´ (1950) Completeness, similar regions and unbiased estimation. Sankhya. Part I, 10 305-340.