ISOMORPHISM OF QUADRATIC NORM AND PC ORDERING OF ESTIMATORS ADMITTING FIRST ORDER AN REPRESENTATIONS

by

Pranab Kumar Sen

Department of Biostatistics, University of North Carolina at Chapel Hill, NC.

Institute of Statistics Mimeo Series No. 1892

January 1992

ISOMORPHISM OF QUADRATIC NORM AND PC ORDERING OF ESTIMATORS ADMITTING FIRST ORDER AN REPRESENTATIONS

BY PRANAB KUMAR SEN

University of North Carolina at Chapel Hill

BAN estimators generally admit first order representations which insure them to be also asymptotically Pitman closest. For AN estimators admitting a first order representation, an ordering under the Pitman closeness measure is established and is shown to be isomorphic to the one under the usual quadratic norm. Multiparameter problems are also treated in the same vein.

1. Introduction. BAN estimators are characterized by their asymptotic normality (AN) and bestness (B) judged by minimum asymptotic mean squared error (MSE). These estimators are also known to be asymptotically Pitman closest (PC) within the class of AN estimators admitting a first order representation [Sen (1986a)]. In statistical estimation problems, especially in nonparametric and robustness setups, there are broad families of estimators (e.g., L-, M-, and R-estimators) which meet the AN criterion, admit a first order representation, but may not be BAN in the light of the MSE or the PC criterion. For the MSE or other quadratic risk criteria, transitivity holds and renders a well defined ordering of competing estimators. However, this transitivity may not be taken for granted for the PC measure [viz., Blyth (1972)], and hence, it may be difficult to establish even a partial ordering of competing estimators under the PC measure. This unpleasantness can largely be eliminated if we confine ourselves to estimators admitting first order AN representation.

The main objective of the present study is to focus on such a PC ordering under the usual first order AN representation conditions. Along with the preliminary notions, the main results pertaining to the single parameter case are presented in Section 2. Section 3 deals with their multiparameter extensions. Some general remarks are made in the concluding section.

2. Orderings of AN estimators. Let $(\mathfrak{X}, \mathcal{A}, \mu)$ be a measure space with μ sigma-finite; we take $\mathfrak{X} = E^t$ for some $t \geq 1$ and $\mathcal{A}$ as the Borel field in $\mathfrak{X}$. Let $\{X_i;\ i \geq 1\}$ be a sequence of independent and identically distributed random variables (i.i.d.r.v.), such that X_i takes values in $\mathfrak{X}$ with a probability distribution $P_\theta(dx) = f(x;\ \theta)d\mu(x)$, $x \in x$, $\theta \in \Theta \subset E^p$, for some $p \geq 1$. For simplicity of presentation, we consider first the case of $p = 1$ i.e., θ real valued.

AMS Subject Classifications: 62F10, 62C99

Key Words and Phrases: AN representation; Pitman closeness; quadratic norm; shrinkage estimation; transitivity.

Short Title: PC Ordering of AN Estimators.

Based on a sample $\underset{\sim}{X}^{(n)} = (X_1, \cdots, X_n)$ of size n (≥ 1), let $T_n = T(\underset{\sim}{X}^{(n)})$ be an estimator of θ. T_n is said to admit a first order AN representation (FOANR) if there exists a score function $\phi(x; T, \theta)$, such that

$$(2.1) \qquad T_n - \theta = n^{-1} \sum_{i=1}^{n} \phi(X_i; T, \theta) + R_n ,$$

where $\phi(\cdot)$ may depend on the for of $T(\cdot)$ as well as on θ, and it is so normalized that $E_\theta\, \phi(X_1; T, \theta) = 0$ and

$$(2.2) \qquad 0 < \sigma_T^2 \;=\; E_\theta\, \phi^2(X; T, \theta) < \infty ;$$

$$(2.3) \qquad R_n = o_p(n^{-1/2}) \quad \text{(or in other norms)} .$$

Note that (2.1) through (2.3) imply that as $n \rightarrow \infty$,

$$(2.4) \qquad n^{1/2}(T_n - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_T^2) ,$$

where by the classical Cramér-Rao inequality,

$$(2.5) \qquad \sigma_T^2 \geq \mathcal{I}_\theta^{-1} \;;\; \mathcal{I}_\theta = E_\theta\Big\{\Big((\partial/\partial\theta) \log f(X; \theta)\Big)^2\Big\} .$$

We assume that the Fisher information $\mathcal{I}_\theta$ is finite and positive. The equality sign in (2.5) holds when T_n is a BAN estimator of θ.

Let $\{T_n^{(j)};\, j \in \mathcal{J}\}$ be a class of AN estimators of θ, and we denote by $\sigma_j^2 = \sigma^2_{T^{(j)}}$, $j \in \mathcal{J}$. Then a (partial) ordering of the $T_n^{(j)}$ can be made on the basis of the σ_j^2. We term this an AMSE ordering. Thus,

$$(2.6) \qquad T_n^{(j)} \underset{\text{AMSE}}{>} T_n^{(\ell)} \quad \text{if} \quad \sigma_j^2 \leq \sigma_\ell^2 , \quad \forall\, \theta \in \Theta ,$$

with strict inequality holding for some θ. This ordering may not always exist, especially when the σ_j^2 may also depend on θ. Moreover, in a nonparametric setup, the density $f(x; \theta)$ is allowed to be a member of a class $\mathcal{F}$, so that $\sigma_j^2(f)$ may also depend on f ($\in \mathcal{F}$), and as a result (2.6) may not hold for all $f \in \mathcal{F}$. For the location-scale family of densities, σ_j^2 may not depend on θ but on f, and hence, the ordering in (2.6) may depend on $f \in \mathcal{F}$, and it needs careful study.

If $T_n^{(1)}$ and $T_n^{(j)}$ are two competing estimators of θ, then $T_n^{(1)}$ is said to be closer to θ than $T_n^{(j)}$ in the Pitman (1937) sense, if

$$(2.7) \qquad P_\theta\{\, |T_n^{(1)} - \theta| \leq |T_n^{(j)} - \theta| \,\} \geq 1/2, \quad \forall\, \theta \in \Theta ,$$

with strict inequality holding for some θ. If (2.7) holds for every $T_n^{(j)}$, $j \in \mathcal{J}$, belonging to a class $\mathcal{C}$, then $T_n^{(1)}$ is Pitman closest (PC) within the class $\mathcal{C}$. It is known [viz., Sen (1986a)] that if T_n^* is a BAN estimator (under the MSE criterion) and $\mathcal{C}$ is the class of AN estimators of θ admitting a first order representation, then T_n^* is asymptotically PC within the class $\mathcal{C}$, i.e.,

$$(2.8) \qquad \liminf_{n\rightarrow\infty} P_\theta\{\, |T_n^* - \theta| \leq |T_n - \theta| \,\} \geq \tfrac{1}{2}, \quad \forall\, \theta \in \Theta , \quad \{T_n\} \in \mathcal{C} .$$

An immediate corollary to (2.8) is that all BAN estimators are asymptotically PC-equivalent. Although (2.8) provides a pairwise ordering of a BAN estimator with any other member of the class $\mathcal{C}$, it may not provide an ordering of two arbitrary members, say $T_n^{(1)}$ and $T_n^{(2)}$, when none of them is BAN. This limitation of the PC measure stems from the fact that (2.7) depends on the joint distribution of $T_n^{(1)}$ and $T_n^{(2)}$, and hence, some other characteristic of this joint distribution may play a vital role in this PC ordering. In fact, this also provides a simple explanation why transitivity may not generally hold for the PC criterion.

In view of the fact that $\mathcal{C}$ is the class of AN estimators of θ, the joint distribution of any pair of estimators is asymptotically bivariate normal, and hence, the only other characteristic of this joint distribution not ascribable in the marginals is the covariance (or correlation) function. Thus, it seems very plausible to incorporate such covariance (or correlation) functions in a characterization of PC ordering of two arbitrary estimators belonging to the class $\mathcal{C}$. We say that $T_n^{(1)}$ is asymptotically Pitman closer than $T_n^{(2)}$, denoted by $T_n^{(1)} \underset{APC}{>} T_n^{(2)}$, if (2.7) holds for large n. Again as in (2.6), such an (APC) ordering may not exist universally. However, the interesting feature is that whenever the AMSE ordering and APC ordering exist they are isomorphic. Towards this, we consider the following.

THEOREM 2.1. Within the class $\mathcal{C}$ of AN estimators admitting a first order representation,

$$\{T_n^{(j)} \underset{AMSE}{>} T_n^{(\ell)}\} \Leftrightarrow \{T_n^{(j)} \underset{APC}{>} T_n^{(\ell)}\}, \tag{2.9}$$

for every j, $\ell \in \mathfrak{J}$, for which the ordering is meaningful.

Proof. Note that by (2.1) through (2.4), for every pair (j, ℓ) $\in \mathfrak{J}$,

$$n^{1/2}\begin{pmatrix} T_n^{(j)} - \theta \\ T_n^{(\ell)} - \theta \end{pmatrix} \xrightarrow{\mathfrak{D}} \mathcal{N}_2\left(\underset{\sim}{0}, \begin{pmatrix} \sigma_j^2 & \sigma_{j\ell} \\ \sigma_{j\ell} & \sigma_\ell^2 \end{pmatrix}\right), \tag{2.10}$$

where

$$\sigma_{j\ell} = E_\theta\{\phi(X_1; T^{(j)}, \theta)\phi(X_1; T^{(\ell)}, \theta)\} \tag{2.11}$$

exists whenever both σ_j^2 and σ_ℓ^2 exist. Note that if (U_1, U_2) has a bivariate normal distribution with null mean vector and dispersion matrix $\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$, then letting $V_1 = U_1 - U_2$ and $V_2 = U_1 + U_2$,

$$\begin{aligned} P\{|U_1| \geq |U_2|\} &= P\{U_1^2 - U_2^2 \geq 0\} \\ &= P\{(U_1 - U_2)(U_1 + U_2) \geq 0\} \\ &= P(U_1 - U_2 \geq 0, U_1 + U_2 \geq 0) + P(U_1 - U_2 \leq 0, U_1 + U_2 \leq 0) \\ &= P(V_1 \geq 0, V_2 \geq 0) + P(V_1 \leq 0, V_2 \leq 0), \end{aligned} \tag{2.12}$$

where

$$(V_1, V_2) \sim \mathcal{N}_2\left(0, \begin{pmatrix} \gamma_{11} + \gamma_{22} - 2\gamma_{12} & \gamma_{11} - \gamma_{22} \\ \gamma_{11} - \gamma_{22} & \gamma_{11} + \gamma_{22} + 2\gamma_{12} \end{pmatrix}\right), \tag{2.13}$$

and hence, (2.12) is $\gtreqless \frac{1}{2}$ according as $\gamma_{11} - \gamma_{22}$ is $\gtreqless 0$. Thus, writing $U_1 = T_n^{(\ell)} - \theta$ and $U_2 = T_n^{(j)} - \theta$, and appealing to (2.10) and (2.12), we readily obtain that

$$\{\sigma_j^2 \le \sigma_\ell^2 \,,\ \forall\, \theta \in \Theta\} \Leftrightarrow \{T_n^{(j)} \underset{APC}{>} T_n^{(\ell)}\}\,, \tag{2.14}$$

and hence, by (2.6) and (2.14), we conclude that (2.9) holds. □

Remarks. It may be observed that the isomorphism in (2.9) is solely based on the ordering of the AMSE $\{\sigma_j^2, j \in \mathfrak{J}\}$, and hence, does not depend on any other particular feature of the joint distribution of $(T_n^{(j)}, T_n^{(\ell)})$, not inherent in the marginal ones. This explains the imperceptible dependence of the PCM on the joint distribution of competing estimators belonging to a class $\mathcal{C}$ for which the FOANR in (2.1) holds. It also answers to a philosophical question raised by Savage (1954). Indeed, in the asymptotic theory of statistical estimation (in a regular case), optimal estimators are very much attuned to the classical LAN condition [viz., LeCam (1986)], and an examination of the class $\mathcal{C}$ reveals that the same LAN condition underlies both the AMSE and APC orderings and characterizes their isomorphy. From this standpoint, at least, in an asymptotic setup, there is not basic disagreement between the two approaches, and hence, in a regular case, there is no profound need to abandon one in favor of the other. Of course, in a finite sample setup, possibly for some nonregular cases, it is not difficult to show that either method may perform better than the other. But this feature should not be over emphasized to tilt the preference to either approach.

We have not commented on the plausibility of the partial ordering in (2.9) for notable subclasses of $\mathcal{C}$. We relegate this to the concluding section so that the multiparameter case may also be covered in the same discussion.

3. Multiparameter models. As in Section 2, we consider the model $P_{\underset{\sim}{\theta}}(dx) = f(x;\, \underset{\sim}{\theta})d\mu(x)$ where $\underset{\sim}{\theta} \in \Theta \subset E^p$, for some $p \ge 1$. This includes the classical location-scale model ($p = 2$) in the univariate as well as multivariate setups. In view of the fact that $\underset{\sim}{\theta}$ is a p-vector, we regard $\underset{\sim}{T}_n$ also as a p-variate random element; the X_i may be scalar, vector or even matrix valued r.v.'s depending on the underlying model.

We say that $\underset{\sim}{T}_n$ admits a first order AN representation if there exists a score function (vector) $\underset{\sim}{\phi}(x;\, \underset{\sim}{T},\, \underset{\sim}{\theta})$, such that

$$\underset{\sim}{T}_n - \underset{\sim}{\theta} = n^{-1} \sum_{i=1}^{n} \underset{\sim}{\phi}(X_i;\, \underset{\sim}{T},\, \underset{\sim}{\theta}) + \underset{\sim}{R}_n\,, \tag{3.1}$$

where $\underset{\sim}{\phi}(\cdot)$ may depend on the form of $\underset{\sim}{T}(\cdot)$ and $\underset{\sim}{\theta}$, and is so chosen that

$$E_{\underset{\sim}{\theta}}\, \underset{\sim}{\phi}(X_1;\, \underset{\sim}{T},\, \underset{\sim}{\theta}) = \underset{\sim}{0}\,, \tag{3.2}$$

$$E_{\theta}[\phi(X_1; T, \theta)][\phi(X_1; T, \theta)]' = \Sigma_{\phi} \text{ is p.d. (and finite)} \tag{3.3}$$

and

$$\| R_n \| = o_p(n^{-1/2}) ; \tag{3.4}$$

here $\| x \|$ stands for the Euclidean norm of x. These regularity conditions insure that as $n \to \infty$,

$$n^{1/2}(T_n - \theta) \xrightarrow{\mathfrak{D}} \mathcal{N}_p(0, \Sigma_{\phi}) . \tag{3.5}$$

Let us also denote by

$$I_{\theta} = E_{\theta}\{[(\partial/\partial\theta)\log f(X_1; \theta)][(\partial/\partial\theta)\log f(X_1; \theta)]'\} \tag{3.6}$$

(the per unit Fisher information matrix on θ), so that we have

$$\Sigma_{\theta} - I_{\theta}^{-1} = \text{p.s.d.} \tag{3.7}$$

[viz., Rao (1965, p. 265)]. If (3.7) reduces to a null matrix, T_n is said to be a BAN estimator of θ. There are other bestness criteria based on real valued functions of $\Sigma_{\theta} I_{\theta}$, such as the trace criterion (A-optimality), the determinant criterion (D-optimality) or the largest root criterion (E-optimality), which may be incorporated to induce a (partial) ordering of competing estimators (belonging to the class $\mathcal{C}$), but we shall find it convenient to use the matrix valued function in (3.7) to formulate such an ordering. We define

$$T_n^{(j)} \underset{\text{AMSPE}}{>} T_n^{(\ell)} \quad \text{if} \quad \Sigma_{\phi_\ell} - \Sigma_{\phi_j} = \text{p.s.d.}, \ \forall\, \theta \in \Theta , \tag{3.8}$$

with a nonnull difference for some θ. Here AMSPE stand for the asymptotic mean sum of product error (vectors). Again, such an ordering may not always exist, and in the next section, we shall elaborate this point.

To extend the PCM in the multiparameter case, we may need to replace the simple Euclidean distance [employed in (2.7) – (2.8)] by an appropriate norm. In this respect, we introduce a quadratic norm

$$\| T - \theta \|_Q^2 = (T - \theta)' Q (T - \theta) \tag{3.9}$$

where Q is a suitable p.s.d. matrix. We denote by $\mathcal{Q}$ the class of all such p.s.d. Q. then the definitions in (2.7) and (2.8) extend directly to the multiparameter case if we replace $|T - \theta|$ by $\| T - Q \|_Q$, and the corresponding APC dominance is denoted by APC,Q (to emphasize its possible dependence on Q). If this APC,Q ordering holds for every $Q \in \mathcal{Q}$, we denote it by $\underset{\text{APC},\mathcal{Q}}{>}$.

<u>Theorem 3.1.</u> Within the class $\mathcal{C}$ of AN estimators admitting a first order representation,

$$\{T_n^{(j)} \underset{\text{AMSPE}}{>} T_n^{(\ell)}\} \Leftrightarrow \{T_n^{(j)} \underset{\text{APC},\mathcal{Q}}{>} T_n^{(\ell)}\} , \tag{3.10}$$

for every j, $\ell \in \mathfrak{J}$, for which the ordering is meaningful.

Before we present an outline of the proof of Theorem 3.1, we consider the following lemma (which provides the key step).

Lemma 3.1. Let $(\underset{\sim}{U}, \underset{\sim}{V}) \sim \mathcal{N}_{2p}(\underset{\sim}{0}, \underset{\sim}{\Sigma}^*)$, such that conditionally on $\underset{\sim}{U} = \underset{\sim}{u}$,

$$\underset{\sim}{V} \sim \mathcal{N}_p(-\underset{\sim}{B}\underset{\sim}{u}, \underset{\sim}{\Gamma}) \text{ and } \underset{\sim}{B} \text{ is p.s.d.} \tag{3.11}$$

Then,

$$P\{\underset{\sim}{V}'\underset{\sim}{Q}\underset{\sim}{U} \le 0\} \ge \tfrac{1}{2} \quad \text{for every } \underset{\sim}{Q} \in \mathfrak{Q}\,. \tag{3.12}$$

Proof. Note that

$$\begin{aligned} P\{\underset{\sim}{V}'\underset{\sim}{Q}\underset{\sim}{U} \le 0\} &= P\{(\underset{\sim}{V} + \underset{\sim}{B}\underset{\sim}{U})'\underset{\sim}{Q}\underset{\sim}{V} \le \underset{\sim}{U}'\underset{\sim}{B}'\underset{\sim}{Q}\underset{\sim}{U}\} \\ &\ge P\{(\underset{\sim}{V} + \underset{\sim}{B}\underset{\sim}{U})'\underset{\sim}{Q}\underset{\sim}{U} \le 0\}\,, \end{aligned} \tag{3.13}$$

as $\underset{\sim}{B}$, $\underset{\sim}{Q}$ are both p.s.d., so that $\underset{\sim}{B}'\underset{\sim}{Q}$ is also p.s.d., and hence, $\underset{\sim}{U}'\underset{\sim}{B}'\underset{\sim}{Q}\underset{\sim}{U} \ge 0$ with probability one. Next, we write

$$\begin{aligned} &P\{(\underset{\sim}{V} + \underset{\sim}{B}\underset{\sim}{U})'\underset{\sim}{Q}\underset{\sim}{U} \le 0\} \\ &= E[P\{(\underset{\sim}{V} + \underset{\sim}{B}\underset{\sim}{U})'\underset{\sim}{Q}\underset{\sim}{U} \le 0 \mid \underset{\sim}{U}\}] \end{aligned} \tag{3.14}$$

where (3.11), given $\underset{\sim}{U} = \underset{\sim}{u}$ and $\underset{\sim}{Q}$,

$$(\underset{\sim}{V} + \underset{\sim}{B}\underset{\sim}{U})'\underset{\sim}{Q}\underset{\sim}{U} \sim \mathcal{N}_1(0, \underset{\sim}{u}'\underset{\sim}{Q}'\underset{\sim}{\Gamma}\underset{\sim}{Q}\underset{\sim}{u})\,;\ \underset{\sim}{u}'\underset{\sim}{Q}'\underset{\sim}{\Gamma}\underset{\sim}{Q}\underset{\sim}{u} \ge 0\,, \tag{3.15}$$

so that the right hand side of (3.14) is equal to 1/2, and hence, (3.12) follows from (3.13) – (3.15). □

Let us now return to the proof of Theorem 3.1. Note that by (3.1) through (3.4), for every pair $(j, \ell) \in \mathfrak{J}$,

$$n^{1/2}\begin{pmatrix} \underset{\sim}{T}_n^{(j)} - \underset{\sim}{\theta} \\ \underset{\sim}{T}_n^{(\ell)} - \underset{\sim}{\theta} \end{pmatrix} \xrightarrow{\mathfrak{D}} \mathcal{N}_{2p}\left(\underset{\sim}{0}, \begin{pmatrix} \underset{\sim}{\Sigma}_{\underset{\sim}{\phi}_j} & \underset{\sim}{\Sigma}_{\underset{\sim}{\phi}_j\underset{\sim}{\phi}_\ell} \\ \underset{\sim}{\Sigma}_{\underset{\sim}{\phi}_\ell\underset{\sim}{\phi}_j} & \underset{\sim}{\Sigma}_{\underset{\sim}{\phi}_\ell} \end{pmatrix}\right) \tag{3.16}$$

where the $\underset{\sim}{\Sigma}_{\underset{\sim}{\phi}_j}$, $j \in \mathfrak{J}$ are defined as in (3.3) (for $\underset{\sim}{\phi} = \underset{\sim}{\phi}_j$) and

$$\underset{\sim}{\Sigma}_{\underset{\sim}{\phi}_j\underset{\sim}{\phi}_\ell} = E_{\underset{\sim}{\theta}}[\underset{\sim}{\phi}_j(X_1; \underset{\sim}{T}^{(j)}, \underset{\sim}{\theta})][\underset{\sim}{\phi}_\ell(X_1; \underset{\sim}{T}^{(\ell)}, \underset{\sim}{\theta})]'\,. \tag{3.17}$$

Further note that

$$\begin{aligned} &\{\| \underset{\sim}{T}_n^{(j)} - \underset{\sim}{\theta} \|_{\underset{\sim}{Q}} \le \| \underset{\sim}{T}_n^{(\ell)} - \underset{\sim}{\theta} \|_{\underset{\sim}{Q}}\} \\ &\Leftrightarrow \{(\underset{\sim}{T}_n^{(j)} - \underset{\sim}{T}_n^{(\ell)})'\underset{\sim}{Q}(\underset{\sim}{T}_n^{(\ell)} + \underset{\sim}{T}_n^{(j)} - 2\underset{\sim}{\theta}) \le 0\}\,, \end{aligned} \tag{3.18}$$

so that writing $\underset{\sim}{U} = n^{1/2}(\underset{\sim}{T}_n^{(j)} - \underset{\sim}{T}_n^{(\ell)})$ and $\underset{\sim}{V} = n^{1/2}(\underset{\sim}{T}_n^{(j)} + \underset{\sim}{T}_n^{(\ell)} - 2\underset{\sim}{\theta})$ and using (3.8), it is easy to verify that (3.11) holds, and hence, the proof follows by using (3.12). □

Remarks. It may be noted that although to incorporate a quadratic norm in the definition of the PCM one needs to choose a p.s.d. $\underset{\sim}{Q}$, the ordering of the $\underset{\sim}{T}_n^{(j)}$, $j \in \mathfrak{J}$ under the APC criterion does not depend on $\underset{\sim}{Q}$ ($\in \mathfrak{Q}$). In the single parameter case, Q is a nonnegative constant and it plays no role, and in the multiparameter case too, $\underset{\sim}{Q}$ has no perceptible role. In having the asymptotic Pitman closest characterization of BAN estimators (in a general multivariate setup), Sen (1986a) considered the class $\mathfrak{T}$ of all estimators $\{\underset{\sim}{T}_n\}$ for which

$$(3.19)\qquad n^{1/2}\begin{pmatrix} \underset{\sim}{T}_n - \underset{\sim}{\theta} \\ \hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta} \end{pmatrix} \xrightarrow{\mathfrak{D}} \mathcal{N}_{2p}\left(\underset{\sim}{0}, \begin{pmatrix} \underset{\sim}{\Sigma} & \underset{\sim}{I}_\theta^{-1} \\ \underset{\sim}{I}_\theta^{-1} & \underset{\sim}{I}_\theta^{-1} \end{pmatrix}\right), \quad \underset{\sim}{\Sigma} - \underset{\sim}{I}_\theta^{-1} = \text{p.s.d.},$$

where $\hat{\underset{\sim}{\theta}}_n$ is a BAN estimator of θ, and it was shown there that for $\underset{\sim}{Q} = \underset{\sim}{I}_\theta$, $\hat{\underset{\sim}{\theta}}_n$ is APC within this class. It follows by using Theorem 3.1 that for AN estimators admitting a first order representation, the APC characterization of $\hat{\theta}_n$ remains intact for an arbitrary $\underset{\sim}{Q}$ (p.s.d.). Secondly, as in the single parameter case, the joint distribution of $(\underset{\sim}{T}_n^{(j)}, \underset{\sim}{T}_n^{(\ell)})$ does not have a significant role in this APC ordering; the definition of $\underset{\text{AMSPE}}{>}$ in (3.8) (depending only on the two marginal distributions (of $\underset{\sim}{T}_n^{(j)}$ and $\underset{\sim}{T}_n^{(\ell)}$) suffices in this context. Thus, the LAN condition in a multiparameter setup takes care of both the AMSPE and APC orderings, providing an explanation of the insignificant role of joint distributions in APC orderings and answering to a basic question of Savage (1954) even in a multiparameter model.

There is an important issue relating to multiparameter estimation problems which we discuss briefly in the rest of this section. With respect to a quadratic risk, i.e., $E_\theta\{(\underset{\sim}{T} - \underset{\sim}{\theta})'\underset{\sim}{Q}(\underset{\sim}{T} - \underset{\sim}{\theta})\}$ for a given p.s.d. $\underset{\sim}{Q}$, the usual BAN estimators can be dominated by suitable Stein-rule or shrinkage versions [viz., Sen (1986b)], although such an asymptotic dominance is perceptible only in a Pitman-neighborhood of the assumed pivot. The same phenomenon holds under PMC also [viz., Sen, Kubokawa and Saleh (1989)]. But, such Stein-rule estimators are not AN, and hence, they do not belong to the class $\mathfrak{C}$ (or $\mathfrak{T}$). Therefore, the two orderings (and their isomorphism) considered here for AN estimators may not apply to Stein-rule versions. In this context, the asymptotic risk ordering can easily be studied by standard analysis, but for the APC ordering a more complicated proof may be needed (as the relevant joint distributions are not multinormal even asymptotically).

4. Some general remarks. In a parametric setup, the density f(x; θ) is of assumed form, so that the dominance in (2.6) or (2.7) has to be studied for all $\theta \in \Theta$. In a nonparametric (or semi-parametric) setup, the density f(x; θ) is allowed to be a member of a class (say, $\mathfrak{F}$, and for each $f \in \mathfrak{F}$, one obtains a spectrum over the variation of θ over Θ. Thus, the ordering of competing estimators (belonging to a class $\mathcal{T}$) under either MSE or PC criterion may specifically depend on the underlying f, and generally, no estimator may emerge as best simultaneously for all $f \in \mathcal{F}$. Therefore, it may be of some interest to examine the particular features (i.e., functionals) of $f(\cdot)$ which depict the relative ordering of competing estimators, and such features may also depend on the class ($\mathcal{T}$) of estimators for which the ordering is sought. We illustrate this point with the usual rank based (i.e., R-) estimators of location parameters. A very similar picture holds for M- and L-estimators of location and regression parameters.

Let $X_1, \cdots, X_n$ be n i.i.d.r.v.'s with a pdf $f(x-\theta)$, where f is symmetric about 0 and θ ($\in \Theta \subset E$) is the location parameter. For every $n (\geq 1)$, let $a_n(1) \leq \cdots < a_n(n)$ be a set of scores, defined by

$$a_n(k) = \phi\Big(\frac{k}{n+1}\Big) \quad \text{or} \quad E\phi(U_{n:k}) , \quad 1 \leq k \leq n , \tag{4.1}$$

where $\phi = \{\phi(u),\ 0 < u < 1\}$ is a monotone and square integrable score function and $U_{n:1} < \cdots < U_{n:n}$ are the order statistics of a sample of size n from the uniform (0, 1) distribution. For every real t, let $R^+_{ni}(t)$ be the rank of $|X_i - t|$ among $|X_1 - t|, \cdots, |X_n - t|$, for $i = 1, \ldots, n$. Let then

$$L^\phi_n(t) = \sum_{i=1}^n \operatorname{sign}(X_i - t)\, a_n(R^+_{ni}(t)), \quad t \in E . \tag{4.2}$$

It is known that $L^\phi_n(t)$ is $\searrow$ in t ($\in E$), and $L^\phi_n(\theta)$ is distributed symmetrically around 0 (independently of f). The usual R-estimator of θ based on $L^\phi_n(\cdot)$ is defined by

$$\hat\theta_n(\phi) = \tfrac{1}{2}\{\hat\theta_{n,1}(\phi) + \hat\theta_{n,2}(\phi)\} ; \tag{4.3}$$

$$\hat\theta_{n,1} = \sup\{t\colon L^\phi_n(t) > 0\}, \quad \hat\theta_{n,2}(\phi) = \inf\{t\colon L^\phi_n(t) < 0\} . \tag{4.4}$$

Let us also define $\psi_f = \{\psi_f(u);\ 0 < u < 1\}$ by

$$\psi_f(u) = -f'(F^{-1}(u))/f(F^{-1}(u)), \quad 0 < u < 1 , \tag{4.5}$$

and let $\phi(u) = \phi_0((1+u)/2)$, $0 < u < 1$, where ϕ_0 is skew-symmetric about 1/2, and

$$\gamma_{f\phi} = \gamma(\phi, \psi_f) = \int_0^1 \phi(u)\psi_f(u)du , \tag{4.6}$$

$$A^2_\phi = \int_0^1 \phi^2(u)du = A^2_{\phi_0} , \tag{4.7}$$

and note that by definition

$$A^2_{\psi_f} = \int_0^1 \psi^2_f(u)du = \int_E (f'(x)/f(x))^2 dF(x) = \mathcal{I}(f) \tag{4.8}$$

is the Fisher information which we assume to be finite. Then,

$$n^{1/2}(\hat\theta_n(\phi) - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ A^2_\phi \gamma^{-2}_{f\phi}) \tag{4.9}$$

where

(4.10) $$A_\phi^2 \gamma_{f\phi}^{-2} = [\mathcal{I}(f)\ \rho^2(\phi,\ \psi_f)]^{-1};$$

(4.11) $$\rho^2(\phi_0,\ \psi_f) = \gamma_{f\phi}^2/(A_\phi^2\ A_{\psi f}^2) = <\phi_0,\ \psi_f>^2/(<\phi,\ \phi><\psi_f,\ \psi_f>)\ .$$

The first order AN representation in (2.1) for $\hat{\theta}_n(\phi)$ holds under quite general regularity conditions [viz., Jurečková (1984)], and hence, it follows that the AMSE or APC ordering of R-estimators of θ, for a given f, is dictated by the squared correlation function:

(4.12) $$\rho^2(\phi_0,\ \psi_f)\ ,\quad \phi \in \Phi\ ,$$

where $\Phi = \{\phi\}$ is the class of (monotone) score functions for which the AN representation holds. Note that for $\phi_0 \equiv \psi_f$, $\rho^2(\phi_0,\ \psi_f) = 1$, so that we have a BAN estimator of θ, and hence, the asymptotic Pitman closest characterization holds. Also for pdf's having distinct functional forms, the corresponding ψ_f's do not coincide, and hence, the same ϕ can't lead to a BAN estimator simultaneously for all $f \in \mathcal{F}$. On the other hand, for a $f \in \mathcal{F}$, if $\psi_f \in \Phi$, then the two (partial) orderings in (2.9) exist and are isomorphic. A very similar characterization holds for M- and L-estimators.

Let us consider now the multivariate location model to examine the relative picture. Instead of the simple scalar function in (4.12), we would have here a $p \times p$ matrix for which (3.8) holds. But, as has been discussed at the end of Section 3, such AN estimators can be dominated (either under AMSE or PCM) by appropriate Stein-rule/shrinkage versions. Thus, it may be more pertinent to inquire whether among such Stein-rule versions we may induce an ordering under the PCM? In this respect we may note that even in the case of a multivariate normal distribution with a known dispersion matrix, among the class of Stein-rule estimators of the mean vector, there is no unique estimator which is PC [viz., Sen and Sengupta (1991)]. Thus, a $\underset{PC}{>}$ ordering of Stein-rule estimators may not exist! This drawback stems primarily from the nonlinear (and non-normal) nature of Stein-rule estimators, and even in an asymptotic setup this shortcoming is retained to a greater extent. For this reason, theorem 3.1 may not hold for non-AN estimators (including the shrinkage ones), and hence, we do not attempt to probe into this model here.

REFERENCES

BLYTH, C. (1972). Some probability paradoxes in choice from among random alternatives (with discussion). J. Amer. Statist. Assoc. 67, 366-373.

JUREČKOVÁ, J. (1984). M-, L-, and R-estimators. In Handbook of Statistics, Vol. 4: Nonparametric Methods, (eds., P. R. Krishnaiah and P. K. Sen), North Holland, Amsterdam, pp. 463-485.

LECAM, L. (1986). Asymptotic Methods in Statistical Decision Theory. Springer Verlag, New York.

PITMAN, E. J. G. (1937). The closest estimates of statistical parameters. Proc. Cambridge Phil. Soc.

33, 212-222.

RAO, C. R. (1965). Linear Statistical Inference and its Applications. Wiley, New York.

SAVAGE, L. J. (1954). The Foundations of Statistics. Wiley, New York.

SEN, P. K. (1986a). Are BAN estimators the Pitman closest ones too? Sankhyā, A 48, 51-58.

SEN, P. K. (1986b). On the asymptotic distributional risks of shrinkage and preliminary test versions of maximum likelihood estimators", Sankhyā, A, 48, 354-371.

SEN, P. K., KUBOKAWA, T. and SALEH, A. K. M. E. (1989). The Stein paradox in the sense of the Pitman measure of closeness. Ann. Statist. 17, 1375-1386.

SEN, P. K. and SENGUPTA, D. (1991). On characterizations of Pitman closeness of some shrinkage estimators. Commun. Statist. Theor. Meth. 20, 3551-3580.