

## SECOND ORDER PITMAN CLOSENESS AND PITMAN ADMISSIBILITY

J.K. Ghosh<sup>1</sup>, Pranab Kumar Sen<sup>2</sup> and Rahul Mukerjee<sup>3</sup>

### ABSTRACT

Motivated by the first-order Pitman closeness of best asymptotically normal estimators and some recent developments on higher order asymptotic efficiency of estimators, a second-order asymptotic theory is developed for comparison of estimators under the Pitman closeness criterion. The single and multidimensional parameter cases are studied. The notion of second-order Pitman admissibility is introduced and examined critically.

Key words: Edgeworth expansion; median unbiasedness; parametric orthogonality; Pitman admissibility; posterior median; Stein-type phenomenon.

<sup>1</sup> Purdue University and Indian Statistical Institute, Calcutta.

<sup>2</sup> Departments of Biostatistics and Statistics, University of North Carolina, Chapel Hill.

<sup>3</sup> Indian Institute of Management, Calcutta.

## SECOND-ORDER PITMAN CLOSENESS AND PITMAN ADMISSIBILITY

### 1. Introduction

Best asymptotically normal (BAN) estimators are known to be first-order efficient in the light of conventional quadratic risks as well as the Pitman closeness criteria (see e.g., Sen (1986)); the foundation has been laid down by a first-order asymptotic representation of BAN estimators. The past two decades have witnessed a phenomenal growth of research literature on higher order asymptotic efficiency wherein Edgeworth expansions, bias corrections and asymptotic median unbiasedness have all contributed generously in the accomplished refinements. However, the work is mostly confined to quadratic or related (usually bowl-shaped) risk functions.

The pioneering work of Rao (1981) has led to a revival of interest in recent years in studies on Pitman closeness. Much has been accomplished since then in this area of fruitful research with due emphasis on multiparameter as well as sequential estimation problems; for some comprehensive reviews, we may refer to Rao, Keating and Mason (1986), Sen (1991) and Keating, Mason and Sen (1992), among others. However, very little progress has so far been made beyond the first order asymptotics. Notwithstanding some interesting results recently reported by Severini (1992), it appears that much work remains to be done on higher order asymptotic comparison of estimators with regard to Pitman closeness and with reference to general parametric families.

The Pitman closeness criterion is essentially a measure of pairwise comparisons, and it extends to comparisons within a suitable class of estimators only under additional restrictions such as equivariance, ancillarity of the differences of pairs of estimators in

the class of asymptotically normal laws etc. Although the usual definition of Pitman closeness extends readily to cover the second-order case (see section 2), it has a natural appeal only when the competing estimators are first-order efficient. For this reason, and given the affinity of BAN estimators to the classical maximum likelihood estimate (MLE), in the current study we confine ourselves to a class of estimators which are essentially related to the MLE by small bias corrections. This also enables us to study the second-order Pitman admissibility of estimators within the same class. In this parametric framework, the present work attempts to study the second-order admissibility results in the light of Pitman closeness. The one-parameter case is treated in section 2 while the multiple parameter case is discussed in section 3. Section 4 deals with the case of one-dimensional parameter of interest when there are nuisance parameters. Various examples have been included in the text to illustrate the subtle points in the discourse.

## 2. The one-parameter case

Let  $\{X_i\}$ ,  $i \geq 1$ , be a sequence of independent and identically distributed random variables with common density  $f(x; \theta)$ , where  $\theta$  is an unknown scalar parameter, the parametric space for  $\theta$  being the real line or some open subset thereof. We make the assumptions in Bhattacharya and Ghosh (1978, p. 439) with  $s=3$  (in their notation), and  $f(\cdot; \theta)$  and  $g(\cdot; \theta)$  in their notation interpreted respectively as  $\log f(\cdot; \theta)$  and  $f(\cdot; \theta)$  in our notation. Let  $\hat{\theta}$  ( $\equiv \hat{\theta}_n$ ) be the MLE, defined in the sense of Theorem 3 in Bhattacharya and Ghosh (1978), of  $\theta$  based on  $X_1, \dots, X_n$ , where  $n$  is the sample size. Along the line of Ghosh and Sinha (1981) (see also Pfanzagl and Wefelmeyer (1978)), we consider a class,  $C$ , of estimators of  $\theta$  of the form

$$T_n = \hat{\theta} + n^{-1}Q,$$

where

- (i)  $Q = d(\theta) + o(1)$  under  $\theta$ , over a set with  $P_\theta$ -probability  $1+o(n^{-\frac{1}{2}})$ ,  $d(\cdot)$  being a continuously differentiable function whose functional form is free from  $n$ , and
- (ii) for each positive  $\epsilon$ , free from  $n$ , and each  $\theta$ ,

$$P_\theta(|Q-d(\theta)| > \epsilon) = o(n^{-\frac{1}{2}}).$$

The class  $C$  is quite large. In particular, by Theorem 3 in Bhattacharya and Ghosh (1978), it includes all estimators of the form  $\hat{\theta} + n^{-1}d(\hat{\theta})$ , where  $d(\cdot)$  is continuously differentiable and the functional form of  $d(\cdot)$  is free from  $n$  (cf. Ghosh and Sinha (1981)).

Let  $I = E_\theta\{(\text{dlog } f(X;\theta)/\text{d}\theta)^2\}$  denote the per observation Fisher information at  $\theta$ , which is assumed to be positive for each  $\theta$ . Also, let  $L_{1.1.1} = E_\theta\{(\text{dlog } f(X;\theta)/\text{d}\theta)^3\}$ . Note that both  $I$  and  $L_{1.1.1}$  are functions of  $\theta$ . The following lemma will be useful in the sequel.

Lemma 2.1. Let  $T_n^*$  and  $T_n$  be distinct members of  $C$  such that  $T_n^* = \hat{\theta} + n^{-1}Q^*$ ,  $T_n = \hat{\theta} + n^{-1}Q$ , with  $Q^* = d^*(\theta) + o(1)$ ,  $Q = d(\theta) + o(1)$  under  $\theta$ . Then for each  $\theta$  such that  $d^*(\theta) \neq d(\theta)$ ,

$$P_\theta(|T_n^* - \theta| < |T_n - \theta|) = \frac{1}{2} + \frac{1}{2}(2\pi n)^{-\frac{1}{2}} I^{\frac{1}{2}} \text{sgn}\{d(\theta) - d^*(\theta)\} \{d(\theta) + d^*(\theta) - \frac{1}{3}I^{-2}L_{1.1.1}\} + o(n^{-\frac{1}{2}}).$$

For  $\theta$  such that  $d(\theta) > d^*(\theta)$ , Lemma 2.1 can be proved if one uses conditions (i) and (ii) above to show that

$$P_\theta(|T_n^* - \theta| < |T_n - \theta|) = P_\theta(\xi_n > 0) + o(n^{-\frac{1}{2}}),$$

where  $\xi_n = (nI)^{\frac{1}{2}}(\hat{\theta} - \theta) + \frac{1}{2}n^{-\frac{1}{2}} I^{\frac{1}{2}}\{d(\theta) + d^*(\theta)\}$ , and then employs an Edgeworth expansion for the distribution of  $\xi_n$  under  $\theta$  (cf. the proof of Theorem 3.1 below). The proof is similar for  $\theta$  such that  $d(\theta) < d^*(\theta)$ . Lemma 2.1 is similar to a result in Severini (1992) who compared biased and bias-corrected estimators of a one-dimensional interest

parameter with bias defined in the usual sense of expectation.

Lemma 2.1 provides a simple formula for second-order comparison of two distinct members of  $C$  in terms of Pitman closeness. Let

$$d_0(\theta) = \frac{1}{6}I^{-2}L_{1.1.1}. \quad (2.1)$$

Then it is not hard to deduce the following result from Lemma 2.1.

Theorem 2.1. Let  $T_{on}$  and  $T_n$  be distinct members of  $C$  such that  $T_{on} = \hat{\theta} + n^{-1}Q_0$ ,  $T_n = \hat{\theta} + n^{-1}Q$ , with  $Q_0 = d_0(\theta) + o(1)$ ,  $Q = d(\theta) + o(1)$  under  $\theta$ .

Then for each  $\theta$  such that  $d(\theta) \neq d_0(\theta)$ ,

$$P_{\theta}(|T_{on} - \theta| < |T_n - \theta|) = \frac{1}{2} + \frac{1}{2}(2\pi n)^{-\frac{1}{2}} I^{\frac{1}{2}} |d(\theta) - d_0(\theta)| + o(n^{-\frac{1}{2}}),$$

$$\text{and } \lim_{n \rightarrow \infty} [n^{\frac{1}{2}} \{P_{\theta}(|T_{on} - \theta| < |T_n - \theta|) - \frac{1}{2}\}] > 0.$$

We now discuss the implications of Theorem 2.1 by introducing the notion of second-order Pitman admissibility. An estimator  $T_n = \hat{\theta} + n^{-1}Q$  ( $\in C$ ), with  $Q = d(\theta) + o(1)$  under  $\theta$ , will be called second-order Pitman inadmissible (SOPI) in  $C$  if there exists an estimator  $T_n^* = \hat{\theta} + n^{-1}Q^*$  ( $\in C$ ), with  $Q^* = d^*(\theta) + o(1)$  under  $\theta$  and  $d^*(\theta)$  not identically equal to  $d(\theta)$ , such that  $T_n^*$  is superior to  $T_n$  with regard to second-order Pitman closeness in the sense that defining  $\alpha_{n1}(\theta) = P_{\theta}(|T_n^* - \theta| < |T_n - \theta|) - \frac{1}{2}$  and  $\alpha_{n2}(\theta) = n^{\frac{1}{2}} \alpha_{n1}(\theta)$ , the conditions

- (a)  $\lim_{n \rightarrow \infty} \alpha_{n2}(\theta) \geq 0$  for each  $\theta$  for which the limit exists, and
- (b)  $\lim_{n \rightarrow \infty} \alpha_{n1}(\theta)$  exists and  $\lim_{n \rightarrow \infty} \alpha_{n1}(\theta) \geq 0$  for each  $\theta$  for which  $\lim_{n \rightarrow \infty} \alpha_{n2}(\theta)$  does not exist,

hold, the inequality being strict for some  $\theta$  either in (a) or in (b).

An estimator  $T_n$  ( $\in C$ ) will be called second-order Pitman admissible (SOPA) in  $C$  if it is not SOPI in  $C$ .

An implication of Theorem 2.1 is that the estimator  $T_{on}$  considered there is SOPA in  $C$ . In particular, it follows that the estimator  $\hat{\theta} + n^{-1}d_0(\hat{\theta})$  ( $\in C$ ), where  $d_0(\cdot)$  is given by (2.1), is SOPA in  $C$ .

Note that  $T_{on}$ , considered in Theorem 2.1, is second-order median unbiased in the sense that  $P_{\theta}(T_{on} \geq \theta) = \frac{1}{2} + o(n^{-\frac{1}{2}})$  for each  $\theta$ , as one can prove using an Edgeworth expansion for the distribution of  $(nI)^{\frac{1}{2}}(T_{on} - \theta)$  under  $\theta$ . Hence the second-order Pitman admissibility of  $T_{on}$  is comparable with the exact findings in Ghosh and Sen (1989) who proved, under certain conditions, an optimum property of median unbiased estimators with regard to Pitman closeness. It also follows from Theorem 2.1 that an estimator  $T_n = \hat{\theta} + n^{-1}Q$  ( $\in C$ ), with  $Q = d(\theta) + o(1)$  under  $\theta$  and  $d(\theta) \neq d_0(\theta)$  for each  $\theta$ , will be SOPI in  $C$ , being dominated by  $T_{on}$ . Thus Theorem 2.1 yields one SOPA estimator, namely  $T_{on}$ , and provides a quick way of identifying SOPI estimators.

Remark 1. Under suitable conditions like those in Johnson (1970), together with an assumption regarding the existence of an  $n_0$  such that the posterior distribution of  $\theta$  given  $X_1, \dots, X_{n_0}$  is proper, it can be shown from Theorem 2.1 that the posterior median of  $\theta$  under Jeffreys' prior is SOPA in  $C$  (cf. Welch and Peers (1963)). This frequentist result may be contrasted with the findings in Ghosh and Sen (1991) on properties of the posterior median in terms of posterior Pitman closeness.

Remark 2. The property of  $T_{on}$ , stated in Theorem 2.1, is in fact much stronger than second-order Pitman admissibility. It implies that a rival estimator  $T_n = \hat{\theta} + n^{-1}Q$  ( $\in C$ ), with  $Q = d(\theta) + o(1)$  under  $\theta$  and  $d(\theta)$  not identically equal to  $d_0(\theta)$ , will be inferior to  $T_{on}$ , with regard to second-order Pitman closeness, for each  $\theta$  satisfying  $d(\theta) \neq d_0(\theta)$ . Incidentally, for  $\theta$  such that  $d(\theta) = d_0(\theta)$ , additional regularity conditions (e.g., asymptotic ancillarity) may be required to clearly depict the relative picture.

Remark 3. Under squared error loss, Ghosh and Sinha (1981) characterized second-order admissible (SOA) estimators of the form  $\hat{\theta} + n^{-1}d(\hat{\theta})$ , where  $d(\cdot)$  is continuously differentiable. As many examples, like the following one, indicate, neither a SOPA estimator in our sense is necessarily SOA in their sense nor a SOA estimator in their sense is necessarily SOPA in our sense.

Example 2.1. Let  $f(x; \theta)$  be the univariate normal density with mean  $\theta$  and variance  $\theta^2$ , where  $\theta > 0$ . Then  $I = 3\theta^{-2}$ ,  $L_{1.1.1} = 14\theta^{-3}$ , and by (2.1),  $d_0(\theta) = \frac{7}{27}\theta$ . Let  $T_{on} = \hat{\theta}(1 + \frac{7}{27}n^{-1})$ ,  $T_n = \hat{\theta}(1 - \frac{1}{9}n^{-1})$ . By Theorem 2.1,  $T_{on}$  is SOPA in  $C$  in our sense and  $T_n$  is not so. On the other hand, following the main result in Ghosh and Sinha (1981),  $T_n$  is SOA in their sense while  $T_{on}$  is not.

### 3. The multiparameter case

Consider now a sequence  $\{X_i\}$ ,  $i \geq 1$ , of independent and identically distributed random variables with common density  $f(x; \theta)$ , where  $\theta = (\theta_1, \dots, \theta_p)'$  is an unknown vector parameter, the parametric space for  $\theta$  being  $\mathcal{R}^p$  or some open subset thereof. We make assumptions along the line of those in the last section. Let  $I = ((I_{ij}))$  be the  $p \times p$  per observation Fisher information matrix which is assumed to be positive definite at each  $\theta$ . Let  $I^{-1} = ((I^{ij}))$  and for  $1 \leq i, j, u \leq p$ , let

$$S_{i.j.u} = E_{\theta} \{ (D_i \log f(X; \theta)) (D_j \log f(X; \theta)) (D_u \log f(X; \theta)) \},$$

$$S_{i.ju} = E_{\theta} \{ (D_i \log f(X; \theta)) (D_j D_u \log f(X; \theta)) \},$$

$$S_{iju} = E_{\theta} (D_i D_j D_u \log f(X; \theta)), \quad \bar{S}_{iju} = S_{iju} - S_{i.j.u},$$

where  $D_i$  is the operator of partial differentiation with respect to  $\theta_i$  ( $1 \leq i \leq p$ ). Note that for each  $i, j, u$ ,  $I_{ij}$ ,  $I^{ij}$ ,  $S_{i.j.u}$ ,  $S_{i.ju}$ ,  $S_{iju}$ ,  $\bar{S}_{iju}$  are functions of  $\theta$ .

Let  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_p)'$  be the MLE of  $\theta$  based on  $X_1, \dots, X_n$ , where  $n$  is the sample size. As a natural extension of the class  $C$  considered in the one-parameter case, we consider a class,  $C_p$ , of estimators of  $\theta$  of the form  $T_n = \hat{\theta} + n^{-1}Q$ , where

(i)  $Q = (Q_1, \dots, Q_p)'$  =  $d(\theta) + o(1)$  under  $\theta$ , over a set with  $P_\theta$ -probability  $1 + o(n^{-\frac{1}{2}})$ , with  $d(\theta) = (d_1(\theta), \dots, d_p(\theta))'$ , each component of  $d(\theta)$  being continuously differentiable and having a functional form free from  $n$ , and

(ii) for each positive  $\epsilon$ , free from  $n$ , and each  $\theta$ ,

$$P_\theta(\|Q - d(\theta)\| > \epsilon) = o(n^{-\frac{1}{2}}),$$

where  $\|\cdot\|$  denotes Euclidean norm.

The following result is helpful in comparing estimators in  $C_p$  with regard to (second-order) Pitman closeness.

Theorem 3.1. Let  $T_n^* = \hat{\theta} + n^{-1}Q^*$  and  $T_n = T_n^* + n^{-1}\phi(\hat{\theta})$  be members of  $C_p$ , where  $Q^* = d^*(\theta) + o(1)$ ,  $d^*(\theta) = (d_1^*(\theta), \dots, d_p^*(\theta))'$ ,  $\phi(\theta) = (\phi_1(\theta), \dots, \phi_p(\theta))'$ , and for each  $i$ ,  $d_i^*(\cdot)$ ,  $\phi_i(\cdot)$  are continuously differentiable, with functional forms free from  $n$ , such that the partial derivatives of  $\phi_i(\cdot)$  fulfil the local Lipschitz conditions. Then for each  $\theta$  with  $\phi(\theta) \neq 0$ ,

$$\begin{aligned} & P_\theta \{ (T_n^* - \theta)' I(T_n^* - \theta) < (T_n - \theta)' I(T_n - \theta) \} \\ &= \frac{1}{2} + (2\pi n)^{-\frac{1}{2}} \{ \phi(\theta)' I \phi(\theta) \}^{-3/2} \left[ \frac{1}{2} \{ \phi(\theta)' I \phi(\theta) \}^2 \right. \\ &\quad + \phi(\theta)' I \phi(\theta) \{ \phi(\theta)' I d^*(\theta) + \text{tr}(B(\theta)) \\ &\quad \quad \quad \left. + \sum \sum \sum_{i,j,u=1}^p \phi_i(\theta) I^{ju} (S_{u.ij} + \frac{1}{2} S_{iju}) \right] \\ &\quad - \frac{1}{6} \sum \sum \sum_{i,j,u=1}^p \phi_i(\theta) \phi_j(\theta) \phi_u(\theta) \bar{S}_{iju} - \phi(\theta)' I B(\theta) \phi(\theta) \Big] + o(n^{-\frac{1}{2}}), \end{aligned}$$

where  $B(\theta)$  is a  $p \times p$  matrix with  $(i, j)$ th element  $D_j \phi_i(\theta)$ ,  $1 \leq i, j \leq p$ .

A proof of Theorem 3.1 has been presented in the Appendix. Note that in the multiparameter case Pitman closeness is being defined



using  $I (\equiv I(\theta))$  as a Riemannian metric as was done earlier by Sen (1986) (see also Amari (1985)). It is not hard to see that for  $p=1$  Theorem 3.1 is in agreement with Lemma 2.1. Thus Theorem 3.1 provides a partial generalization of Lemma 2.1 to the multiparameter case. The generalization is partial because it covers pairs of estimators  $(T_n, T_n^*)$ , with  $T_n, T_n^* \in C_p$  and  $T_n - T_n^* = n^{-1}\phi(\hat{\theta})$ , where the components of  $\phi(\theta)$  are continuously differentiable and the partial derivatives of each component of  $\phi(\theta)$  satisfy the local Lipschitz conditions. In the multiparameter case, consideration of such pairs of estimators is helpful since in the proof one requires a stochastic expansion for  $T_n - T_n^*$  up to  $o(n^{-3/2})$  (see (A.4) in the Appendix); in contrast, for  $p=1$ , it can be seen from first principles that an expansion up to that order is not required in the proof of Lemma 2.1. For general  $p$ , with estimators  $T_n, T_n^* \in C_p$  such that  $T_n - T_n^*$  is not of the form  $n^{-1}\phi(\hat{\theta})$ , it should be possible to derive results analogous to Theorem 3.1, proceeding along the line of proof of this theorem, provided sufficient knowledge about  $T_n - T_n^*$  is available so as to allow an expansion for  $T_n - T_n^*$  up to  $o(n^{-3/2})$ .

Notwithstanding what has been mentioned in the last paragraph, Theorem 3.1 serves as an useful tool in the multiparameter case. In particular, as seen in the rest of this section, it helps in settling certain issues of theoretical interest.

One of these issues relates to a possible extension of Theorem 2.1, which is a rather strong result for the case  $p=1$ , to a multiparameter set-up. To be more specific, using  $I$  as a Riemannian metric, in an attempt of extending Theorem 2.1 to a multiparameter set-up, one may wish to investigate whether or not there exists a  $d_o(\theta) = (d_{o1}(\theta), \dots, d_{op}(\theta))'$  such that if  $T_{on} = \hat{\theta} + n^{-1}Q_o$  ( $\in C_p$ ) with  $Q_o =$

$d_0(\theta) + o(1)$  under  $\theta$ , then for every other estimator  $T_n = \hat{\theta} + n^{-1}Q$  ( $\in C_p$ ) with  $Q = d(\theta) + o(1)$  under  $\theta$ , the relation

$$\lim_{n \rightarrow \infty} [n^{\frac{1}{2}} \{P_{\theta}((T_{on} - \theta)' I(T_{on} - \theta) < (T_n - \theta)' I(T_n - \theta)) - \frac{1}{2}\}] > 0 \quad (3.1)$$

holds whenever  $d(\theta) \neq d_0(\theta)$ .

To answer this question, if possible, suppose a  $d_0(\theta)$  as above exists. Let  $T_{on}$  be as defined in the last paragraph and define the estimator  $T_{en}^{(g)} = T_{on} + n^{-1}eg$ , where  $e$  is a non-zero scalar and  $g$  is a non-null  $p \times 1$  vector, both  $e$  and  $g$  being non-random and, evidently, free from  $\theta$ . Then  $eg$  is a  $p \times 1$  non-null vector and (3.1) holds for each  $\theta$  if one takes  $T_n = T_{en}^{(g)}$ . With  $T_n = T_{en}^{(g)}$ , one can now employ Theorem 3.1 to find an expression for the left-hand side of (3.1), note that this is positive for each  $\theta$ , each non-zero choice of  $e$  and each non-null choice of  $g$ , and then make appropriate choices of  $e$  and  $g$  to obtain the relations

$$\bar{S}_{iju} = \frac{1}{3} \{I_{ij} I_{uu}^{-1} \bar{S}_{uuu} + I_{iu} I_{jj}^{-1} \bar{S}_{jjj} + I_{ju} I_{ii}^{-1} \bar{S}_{iii}\}, \quad 1 \leq i, j, u \leq p, \quad (3.2)$$

identically in  $\theta$ , and

$$d_0(\theta) = I^{-1} \chi(\theta), \quad (3.3a)$$

identically in  $\theta$ ,  $\chi(\theta)$  being a  $p \times 1$  vector with  $i$ th ( $1 \leq i \leq p$ ) element

$$\chi_i(\theta) = \frac{1}{6} I_{ii}^{-1} \bar{S}_{iii} - \sum \sum_{j,u=1}^p I^{ju} (S_{u.ij} + \frac{1}{2} S_{iju}). \quad (3.3b)$$

The detailed derivation of (3.2) and (3.3a,b) is being omitted in order to save space.

Thus for the existence of  $d_0(\theta)$  as envisaged above, it is necessary that the underlying model must satisfy (3.2) and that  $d_0(\theta)$  must be as specified by (3.3a,b). For  $p=1$ , (3.2) holds trivially and (3.3a,b) reduce to (2.1). For a model satisfying (3.2), if one takes  $T_n = T_{on} - n^{-1}(\hat{\theta} - \theta_0)$  in (3.1), where  $\theta_0$  is a fixed point in the parametric space, then by Theorem 3.1, (3.1)-(3.3), it follows after some algebra that the relation

$$\frac{1}{2}(\theta - \theta_0)' I(\theta - \theta_0) - (p-1) > 0$$

must hold for each  $\theta \neq \theta_0$ . Since, by assumption, the elements of  $I$  are continuous in  $\theta$ , the impossibility of the above for  $p \geq 2$  follows by allowing  $\theta$  tend to  $\theta_0$ .

The above discussion shows that a strong result like Theorem 2.1 does not hold in the multiparameter case. One may now wish to explore the possibility of deriving weaker admissibility results in the sense of the paragraph following Theorem 2.1. Let SOPA and SOPI estimators in  $C_p$  be defined as in that paragraph with  $\alpha_{n1}(\theta), \alpha_{n2}(\theta)$  redefined as  $\alpha_{n1}(\theta) = P_\theta((T_n^* - \theta)' I(T_n^* - \theta) < (T_n - \theta)' I(T_n - \theta)) - \frac{1}{2}$ ,  $\alpha_{n2}(\theta) = n^{\frac{1}{2}} \alpha_{n1}(\theta)$ . As observed above,  $d_0(\theta)$  given by (3.3a,b) is a natural extension of what is defined by (2.1) in the one-parameter case. Hence one may be interested in checking whether or not an estimator  $T_{on} = \hat{\theta} + n^{-1} Q_0$  ( $\in C_p$ ) with  $Q_0 = d_0(\theta) + o(1)$  under  $\theta$ , where  $d_0(\theta)$  is specified by (3.3a, b), is at least SOPA in  $C_p$ . The answer to this question is also in the negative as the following example, exhibiting a Stein-type phenomenon in the present context, indicates.

Example 3.1. Let  $f(x; \theta)$  represent the  $p$ -variate normal density with mean vector  $\theta = (\theta_1, \dots, \theta_p)'$  and dispersion matrix  $\mathcal{J}_p$ , the  $p \times p$  identity matrix, the parametric space for  $\theta$  being  $\mathcal{R}^p$ . Here  $I = \mathcal{J}_p$ , and for  $1 \leq i, j, u \leq p$ ,  $S_{i.j.u} = S_{i.ju} = S_{iju} = \bar{S}_{iju} = 0$ , identically in  $\theta$ . Hence (3.2) holds and by (3.3a,b),  $d_0(\theta)$  is identically equal to the null vector. The MLE of  $\theta$ , based on  $X_1, \dots, X_n$ , is  $\hat{\theta} = \bar{X}_n$ , the sample mean vector. Consider now estimators

$$T_{on} = \hat{\theta}, \quad T_n^* = \hat{\theta} \{1 - 2n^{-1}(1 + \hat{\theta}' \hat{\theta})^{-1}\}.$$

Both  $T_{on}$  and  $T_n^*$  belong to  $C_p$ . Also,  $T_n^* = \hat{\theta} + n^{-1} Q^*$ ,  $T_{on} = \hat{\theta} + n^{-1} Q_0$ , with  $Q^* = d^*(\theta) + o(1)$ ,  $Q_0 = d_0(\theta) + o(1)$  under  $\theta$ , where  $d_0(\theta) = 0$  and  $d^*(\theta) = -2(1 + \theta' \theta)^{-1} \theta$ . Note that  $T_{on} = T_n^* + n^{-1} \phi(\hat{\theta})$ , where  $\phi(\theta) = 2(1 + \theta' \theta)^{-1} \theta$ .

Hence from Theorem 3.1, it can be seen that for  $\theta \neq 0$ ,

$$\begin{aligned} \lim_{n \rightarrow \infty} [n^{\frac{1}{2}} \{P_{\theta}((T_n^* - \theta)' I(T_n^* - \theta) < (T_{on} - \theta)' I(T_{on} - \theta)) - \frac{1}{2}\}] \\ = (2\pi)^{-\frac{1}{2}} (\theta' \theta)^{-\frac{1}{2}} \{p-1 - \frac{1}{2} \lambda(\theta)(\theta' \theta)\}, \end{aligned} \quad (3.4)$$

where  $\lambda(\theta) = 2(1 + \theta' \theta)^{-1}$ . For  $p \geq 2$ , it is easily seen that the right-hand side of (3.4) (and hence its left-hand side) is positive for

each  $\theta \neq 0$ . Recalling that  $\hat{\theta} = \bar{X}_n$ , one can also check that for  $\theta = 0$ ,

$$P_0((T_n^* - 0)' I(T_n^* - 0) < (T_{on} - 0)' I(T_{on} - 0)) = P_0(\bar{X}_n' \bar{X}_n > n^{-1} - 1) = 1,$$

for each  $n$ . Thus  $T_{on}$  is second-order Pitman inadmissible in  $C_p$ , being dominated by  $T_n^*$ .

Example 3.1 is foreshadowed by the exact results in Sen, Kubokawa and Saleh (1989) who, for  $p \geq 2$ , proved the inadmissibility of  $\bar{X}_n$  as an estimator of the multivariate normal mean vector  $\theta$  (see also Rao, Keating and Mason (1986) in this context) in the sense of Pitman closeness. As shown above, even up to the second order of comparison,  $\bar{X}_n$  is not admissible.

Thus, to summarize, even under the absence of nuisance parameters, the results in section 2, other than Lemma 2.1, do not have extensions to the multiparameter case. It is, however, possible to extend Lemma 2.1 in a partial but useful manner (vide Theorem 3.1).

Before concluding this section, we indicate an extension of Theorem 3.1. Suppose, instead of  $I(\theta)$ , one wishes to use  $\mathcal{M}(\theta)$  ( $\equiv \mathcal{M}$ ) as a Riemannian metric where  $\mathcal{M}(\theta)$  is a  $p \times p$  matrix which is positive definite for each  $\theta$ . Then, under the set-up of Theorem 3.1 and with the same notational system, proceeding along the line of the Appendix, one can show that for each  $\theta$  with  $\theta(\theta) \neq 0$ ,

$$P_{\theta} \{ (T_n^* - \theta)' \mathcal{M}(T_n^* - \theta) < (T_n - \theta)' \mathcal{M}(T_n - \theta) \}$$

$$\begin{aligned}
 &= \frac{1}{2} + (2\pi n)^{-\frac{1}{2}} \{ \tilde{\theta}(\theta)' I \tilde{\theta}(\theta) \}^{-3/2} \left[ \frac{1}{2} \{ \tilde{\theta}(\theta)' I \tilde{\theta}(\theta) \} \{ \theta(\theta)' \mathcal{M} \theta(\theta) \} \right. \\
 &\quad + \tilde{\theta}(\theta)' I \tilde{\theta}(\theta) \{ \theta(\theta)' \mathcal{M} d^*(\theta) + \text{tr} \{ I^{-1} \mathcal{M} B(\theta) \} \\
 &\quad \quad \quad + \sum \sum \sum_{i,j,u=1}^p \tilde{\theta}_i(\theta) I^{ju} (S_{u.ij} + \frac{1}{2} S_{iju}) \} \\
 &\quad \left. - \frac{1}{6} \sum \sum \sum_{i,j,u=1}^p \tilde{\theta}_i(\theta) \tilde{\theta}_j(\theta) \tilde{\theta}_u(\theta) \bar{S}_{iju} - \tilde{\theta}(\theta)' \mathcal{M} B(\theta) \tilde{\theta}(\theta) \right] + o(n^{-\frac{1}{2}})
 \end{aligned}$$

where  $\tilde{\theta}(\theta) = (\tilde{\theta}_1(\theta), \dots, \tilde{\theta}_p(\theta))' = I^{-1} \mathcal{M} \theta(\theta)$ .

#### 4. A case with nuisance parameter(s)

We continue with the set-up of section 3 but consider a situation where the parameter of interest is one-dimensional. Let  $\theta_1$  be the parameter of interest and  $\theta_2, \dots, \theta_p$  be the nuisance parameters. The assumptions are as in the last section. Since the interest parameter is one-dimensional, we suppose that global parametric orthogonality holds, i.e.,  $I_{1j} = 0$  (identically in  $\theta$ ),  $2 \leq j \leq p$  (vide Cox and Reid (1987)). For the sake of notational simplicity, in the rest of this section, we shall consider the case  $p=2$ , i.e., the nuisance parameter will be supposed to be one-dimensional. The treatment for general  $p$  will be exactly similar and only the notational system will get more involved.

Let  $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)'$  be the MLE of  $\theta$  based on a sample of size  $n$ .

As an analogue of the class  $C$  considered in section 2, we consider a class,  $C^*$ , of estimators of  $\theta_1$  of the form  $T_n = \hat{\theta}_1 + n^{-1}Q$ , where

(i)  $Q = d(\theta) + o(1)$  under  $\theta$ , over a set with  $P_\theta$ -probability  $1 + o(n^{-\frac{1}{2}})$ ,  $d(\cdot)$  being a continuously differentiable function whose functional form is free from  $n$ , and

(ii) for each positive  $\epsilon$ , free from  $n$ , and each  $\theta$ ,

$$P_\theta(|Q - d(\theta)| > \epsilon) = o(n^{-\frac{1}{2}}).$$

For  $1 \leq i, j, u \leq 2$ , let  $S_{i.j.u}$  and  $S_{iju}$  be as in section 3. Define

$$d_0(\theta) = \frac{1}{6} I_{11}^{-2} S_{1.1.1} + \frac{1}{2} (I_{11} I_{22})^{-1} S_{122}. \quad (4.1)$$

Then, analogously to Lemma 2.1 and Theorem 2.1 respectively, the following results hold. The proofs are omitted to save space.

**Lemma 4.1.** Let  $T_n^*$  and  $T_n$  be distinct members of  $C^*$  such that  $T_n^* = \hat{\theta}_1 + n^{-1}Q^*$ ,  $T_n = \hat{\theta}_1 + n^{-1}Q$ , with  $Q^* = d^*(\theta) + o(1)$ ,  $Q = d(\theta) + o(1)$  under  $\theta$ . Then for each  $\theta$  such that  $d^*(\theta) \neq d(\theta)$ ,

$$P_{\theta}(|T_n^* - \theta_1| < |T_n - \theta_1|) = \frac{1}{2} + \frac{1}{2}(2\pi n)^{-\frac{1}{2}} I_{11}^{\frac{1}{2}} \text{sgn}\{d(\theta) - d^*(\theta)\} \{d(\theta) + d^*(\theta) - \frac{1}{3} I_{11}^{-2} S_{1.1.1} - (I_{11} I_{22})^{-1} S_{122}\} + o(n^{-\frac{1}{2}}).$$

**Theorem 4.1.** Let  $T_{on}$  and  $T_n$  be distinct members of  $C^*$  such that  $T_{on} = \hat{\theta}_1 + n^{-1}Q_o$ ,  $T_n = \hat{\theta}_1 + n^{-1}Q$ , with  $Q_o = d_o(\theta) + o(1)$ ,  $Q = d(\theta) + o(1)$  under  $\theta$ ,  $d_o(\cdot)$  being given by (4.1). Then for each  $\theta$  such that  $d(\theta) \neq d_o(\theta)$ ,

$$P_{\theta}(|T_{on} - \theta_1| < |T_n - \theta_1|) = \frac{1}{2} + \frac{1}{2}(2\pi n)^{-\frac{1}{2}} I_{11}^{\frac{1}{2}} |d(\theta) - d_o(\theta)| + o(n^{-\frac{1}{2}})$$

and  $\lim_{n \rightarrow \infty} [n^{\frac{1}{2}} \{P_{\theta}(|T_{on} - \theta_1| < |T_n - \theta_1|) - \frac{1}{2}\}] > 0$ .

It can be shown that an estimator  $T_{on}$ , as in Theorem 4.1, is second-order median unbiased, i.e.,  $P_{\theta}(T_{on} \geq \theta_1) = \frac{1}{2} + o(n^{-\frac{1}{2}})$  for each  $\theta$ . Lemma 4.1 is a powerful tool for comparing estimators in  $C^*$ . In the present set-up, defining SOPA and SOPI estimators in  $C^*$  along the line of section 2, it follows from Theorem 4.1 that an estimator  $T_{on}$ , as in the statement of this theorem, is SOPA in  $C^*$ . Also, an estimator  $T_n = \hat{\theta}_1 + n^{-1}Q$  ( $\in C^*$ ), with  $Q = d(\theta) + o(1)$  under  $\theta$  and  $d(\theta) \neq d_o(\theta)$  for each  $\theta$ , will be SOPI in  $C^*$ . In continuation of Remark 1, under suitable conditions, it can be seen from Theorem 4.1 that the posterior median of  $\theta_1$  under a prior with density proportional to  $I_{11}^{\frac{1}{2}}$ , where the constant of proportionality may involve  $\theta_2$  but not  $\theta_1$  (cf. Tibshirani (1989)), will be SOPA in  $C^*$ .

**Example 4.1.** Let  $f(x; \theta)$  represent the univariate normal density with mean  $\theta_2$  and variance  $\theta_1$ , where  $\theta_2 \in \mathcal{R}^1$  and  $\theta_1 > 0$ . Under this parametrization, global parametric orthogonality holds. Here  $I_{11} = \frac{1}{2}\theta_1^{-2}$ ,  $I_{22} = \theta_1^{-1}$ ,  $S_{1.1.1} = \theta_1^{-3}$ ,  $S_{122} = \theta_1^{-2}$ , so that by (4.1),  $d_o(\theta) = \frac{5}{3}\theta_1$ . The MLE

of  $\theta_1$ , based on  $X_1, \dots, X_n$ , is given by  $\hat{\theta}_1 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ . We shall consider the estimators  $T_{on} = \hat{\theta}_1 (1 + \frac{5}{3}n^{-1})$ ,  $T_n = \hat{\theta}_1$ ,  $T_n^* = n\hat{\theta}_1 / (n-1)$  all of which belong to  $C^*$ . Note that  $T_{on}$  is suggested by Theorem 4.1,  $T_n$  is the MLE and  $T_n^*$  is the usual unbiased estimator which can as well be obtained by maximizing the conditional likelihood of Cox and Reid (1987). Clearly,  $T_{on} = \hat{\theta}_1 + n^{-1}Q_0$ ,  $T_n = \hat{\theta}_1 + n^{-1}Q$ ,  $T_n^* = \hat{\theta}_1 + n^{-1}Q^*$ , with  $Q_0 = d_0(\theta) + o(1)$ ,  $Q = d(\theta) + o(1)$ ,  $Q^* = d^*(\theta) + o(1)$ , where  $d_0(\cdot)$  is as stated above,  $d(\theta) = 0$ , and  $d^*(\theta) = \theta_1$ . Clearly,  $T_{on}$  is SOPI in  $C^*$ . Furthermore, as both  $d(\theta)$  and  $d^*(\theta)$  are different from  $d_0(\theta)$  for each  $\theta$ , by Theorem 4.1, both  $T_n$  and  $T_n^*$  are SOPI in  $C^*$ , being dominated by  $T_{on}$ . Next, in order to compare  $T_n$  and  $T_n^*$ , write  $\alpha_{n2}(\theta) = n^{\frac{1}{2}} \{P_{\theta}(|T_n^* - \theta_1| < |T_n - \theta_1|) - \frac{1}{2}\}$ , and from Lemma 4.1 note that  $\lim_{n \rightarrow \infty} \alpha_{n2}(\theta) = \frac{7}{12} \pi^{-\frac{1}{2}} = 0.3291 (> 0)$  for each  $\theta$ . This shows that  $T_n^*$  is superior to  $T_n$  with regard to second-order Pitman closeness. Exact computations are not hard in this example and the exact values of  $\alpha_{n2}(\theta)$  for  $n=3,5,7,9$  can be seen to equal 0.3443, 0.3372, 0.3346, 0.3333 respectively, for each  $\theta$  (cf. Rao (1981)). Since even for small  $n$  the values of  $\alpha_{n2}(\theta)$  are quite close to  $\lim_{n \rightarrow \infty} \alpha_{n2}(\theta)$  ( $= 0.3291$ ), the asymptotic results appear to be reasonably good indicators of the small or moderate sample behaviour of estimators.

Appendix

Proof of Theorem 3.1. Let

$$\Delta_n(\theta) = P_{\theta} \{ (T_n^* - \theta)' I(T_n^* - \theta) < (T_n - \theta)' I(T_n - \theta) \}.$$

Observe that for each  $\theta$  with  $\varnothing(\theta) \neq 0$ ,

$$\Delta_n(\theta) = P_{\theta}(V_n > 0), \tag{A.1}$$

where

$$V_n = \{ \varnothing(\theta)' I \varnothing(\theta) \}^{-\frac{1}{2}} [ \varnothing(\hat{\theta})' I \{ n^{\frac{1}{2}}(T_n^* - \theta) \} + \frac{1}{2} n^{-\frac{1}{2}} \varnothing(\hat{\theta})' I \varnothing(\hat{\theta}) ]. \tag{A.2}$$

Let  $R = (R_1, \dots, R_p)' = n^{\frac{1}{2}} I(\hat{\theta} - \theta)$ . Since for  $1 \leq i \leq p$ ,

$$R_i = H_{1i} + n^{-\frac{1}{2}} [ \sum_{j,u=1}^p I^{ju} H_{2ij} H_{1lu}$$

$$+ \frac{1}{2} \sum \sum \sum \sum_{j,u,s,t=1}^p S_{ijs} I^{ju} I^{st} H_{lu} H_{lt} ] + o(n^{-\frac{1}{2}}),$$

where

$$H_{1i} = n^{-\frac{1}{2}} \sum_{u=1}^n D_i \log f(X_u; \theta), \quad H_{2ij} = n^{-\frac{1}{2}} \sum_{u=1}^n (D_i D_j \log f(X_u; \theta) + I_{ij}),$$

$1 \leq i, j \leq p$ , the approximate cumulants of  $R$  under  $\theta$  are given by

$$k_{1n}(R_i) = n^{-\frac{1}{2}} \sum \sum_{j,u=1}^p I^{ju} (S_{u.ij} + \frac{1}{2} S_{iju}) + o(n^{-\frac{1}{2}}),$$

$$k_{2n}(R_i, R_j) = I_{ij} + o(n^{-\frac{1}{2}}), \quad k_{3n}(R_i, R_j, R_u) = n^{-\frac{1}{2}} \bar{S}_{iju} + o(n^{-\frac{1}{2}}), \quad (A.3)$$

$1 \leq i, j, u \leq p$ . The fourth and higher order cumulants of  $R$  under  $\theta$  are of order  $o(n^{-\frac{1}{2}})$ . Since  $I\{n^{\frac{1}{2}}(T_n^* - \theta)\} = R + n^{-\frac{1}{2}} Id^*(\theta) + o(n^{-\frac{1}{2}})$ , and

$$\varphi(\hat{\theta}) = \varphi(\theta) + n^{-\frac{1}{2}} B(\theta) I^{-1} R + o(n^{-\frac{1}{2}}), \quad (A.4)$$

it follows from (A.2), (A.3) that for each  $\theta$  with  $\varphi(\theta) \neq 0$ , the approximate cumulants of  $V_n$  under  $\theta$  are given by

$$k_{1n}^*(V_n) = n^{-\frac{1}{2}} \{ \varphi(\theta)' I \varphi(\theta) \}^{-\frac{1}{2}} \{ \frac{1}{2} \varphi(\theta)' I \varphi(\theta) + \varphi(\theta)' Id^*(\theta) + \text{tr}(B(\theta))$$

$$+ \sum \sum \sum_{i,j,u=1}^p \varphi_i(\theta) I^{ju} (S_{u.ij} + \frac{1}{2} S_{iju}) \} + o(n^{-\frac{1}{2}}),$$

$$k_{2n}^*(V_n) = 1 + o(n^{-\frac{1}{2}}),$$

$$k_{3n}^*(V_n) = n^{-\frac{1}{2}} \{ \varphi(\theta)' I \varphi(\theta) \}^{-3/2} \{ \sum \sum \sum_{i,j,u=1}^p \varphi_i(\theta) \varphi_j(\theta) \varphi_u(\theta) \bar{S}_{iju}$$

$$+ 6 \varphi(\theta)' I B(\theta) \varphi(\theta) \} + o(n^{-\frac{1}{2}}).$$

The fourth and higher order cumulants of  $V_n$  under  $\theta$  are of order  $o(n^{-\frac{1}{2}})$ . The proof can now be completed using (A.1) and an Edgeworth expansion for the distribution of  $V_n$  under  $\theta$ .

Note that the stochastic expansions used in the above proof are over a set with  $P_\theta$ -probability  $1 + o(n^{-\frac{1}{2}})$ .

### References

- Amari, S. (1985). Differential-Geometrical Methods in Statistics. Springer-Verlag, New York.
- Bhattacharya, R.N. and Ghosh, J.K. (1978). On the validity of formal Edgeworth expansions. Ann. Statist. 6, 434-451.
- Cox, D.R. and Reid, N. (1987). Parametric orthogonality and approximate conditional inference (with discussion). J. Roy. Statist. Soc. B 49, 1-39.



- Ghosh, J.K. and Sinha, B.K. (1981). A necessary and sufficient condition for second order admissibility with applications to Berkson's bioassay problem. *Ann. Statist.* 9, 1334-1338.
- Ghosh, M. and Sen, P.K. (1989). Median unbiasedness and Pitman closeness. *J. Amer. Statist. Assoc.* 84, 1089-1091.
- Ghosh, M. and Sen, P.K. (1991). Bayesian Pitman closeness. *Commun. Statist. - Theor. Meth.* 20, to appear.
- Johnson, R.A. (1970). Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist.* 41, 851-864.
- Keating, J.P., Mason, R.L. and Sen, P.K. (1992). A Unified Discourse on Pitman's Measure of Closeness. Preprint.
- Pfanzagl, J. and Wefelmeyer, W. (1978). A third-order optimum property of the maximum likelihood estimator. *J. Multivariate Anal.* 8, 1-29.
- Rao, C.R. (1981). Some comments on the minimum mean square error as a criterion of estimation. In: *Statistics and Related Topics* (M. Csorgo, D.A. Dawson, J.N.K. Rao and A.K.Md.E. Saleh eds.), pp. 123-143, North-Holland, Amsterdam.
- Rao, C.R., Keating, J.P. and Mason, R.L. (1986). The Pitman nearness criterion and its determination. *Commun. Statist. - Theor. Meth.* 15, 3173-3191.
- Sen, P.K. (1986). Are BAN estimators the Pitman closest ones too? *Sankhyā A* 48, 51-58.
- Sen, P.K. (1991). Pitman closeness of statistical estimators: Latent years and the renaissance. In: *Issues in Statistical Inference: D. Basu Volume* (M. Ghosh and P.K. Pathak eds.), IMS Monograph Lecture Notes Series No. 17, to appear.
- Sen, P.K., Kubokawa, T. and Saleh, A.K.Md.E. (1989). The Stein paradox in the sense of Pitman closeness. *Ann. Statist.* 17, 1375-1384.
- Severini, T. (1992). A comparison of biased and bias-corrected estimators. Preprint.
- Tibshirani, R. (1989). Noninformative priors for one parameter of many. *Biometrika* 76, 604-608.

Welch, B. and Peers, H.W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. J. Roy. Statist. Soc. B 25, 318-329.