SECOND-ORDER PITMAN CLOSENESS AND PITMAN ADMISSIBILITY

BY JAYANTA K. GHOSH,[1] PRANAB K. SEN AND RAHUL MUKERJEE[2]

Indian Statistical Institute, University of North Carolina and Indian Institute of Management

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, covering both the cases without and with nuisance parameters. The notion of second-order Pitman admissibility is also developed.

1. Introduction. Best asymptotically normal (BAN) estimators are known to be first-order efficient in the light of conventional quadratic risk as well as the Pitman closeness criterion (PCC), and an asymptotic first-order representation of estimators plays a vital role in this context [see, Keating, Mason and Sen (1993), Chapter 6]. 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 made significant contributions toward the accomplished unifications, although the work is mostly confined to quadratic or related (e.g., bowl-shaped) risk functions. The pioneering work of Rao (1981) has led to a revival of interest in recent years in studies on the PCC. In a logically integrated form, a systematic and detailed account of the advantages and disadvantages of the PCC compared to the classical measures based on risk functions is contained in the recent work of Keating, Mason and Sen (1993). The earlier chapters of this monograph deal with the genesis of the PCC along with the related anomalies and controversies, while the last two chapters are devoted to characterizations of Pitman closest estimators (for various parametric families) and unification of the PCC with the conventional decision-theoretic measures in a simple asymptotic framework. However, very little progress has so far been made beyond the first-order asymptotics. Even in a conventional decision-theoretic setup, in the context of higher-order asymptotics, it is not uncommon [see, Ghosh and Sinha (1981)] to confine attention only to a (smaller) class of estimators which is in a sense asymptotically second-order complete, and the recent noteworthy

Received January 1992; revised May 1993.

[1]Also at Purdue University, West Lafayette, Indiana.

[2]Supported by a grant from the Center for Management and Development Studies, Indian Institute of Management, Calcutta.

AMS 1991 *subject classifications.* 62C15, 62F12.

Key words and phrases. Asymptotics, bias, Edgeworth expansion, median unbiasedness, parametric orthogonality, Pitman admissibility, posterior median.

work of Severini (1991, 1992) also pertains to a similar class (where transitivity, equivariance and other desirable properties in a decision-theoretic formulation may not be that important, and may not even hold, in general). To us, the use of the PCC in the context of this higher-order optimality of estimators seems to be an interesting supplement to the usual decision-theoretic formulation for the following reasons. It shows that there is an intuitively reasonable way of comparing estimators which may not preserve transitivity so that transitivity may not be as fundamental as we usually assume it to be. This seems to be the view of Blyth and Pathak (1985) and is shared by others. Moreover, it points to the importance of the joint distribution of estimators which is also ignored in the usual decision-theoretic formulations. In this context, it seems to be of some interest to compare from this point of view members of a class of estimators which is asymptotically second-order complete in a certain sense [cf. Ghosh and Sinha (1981) and Ghosh (1994)]. It therfore appears to us that much work remains to be done on higher-order asymptotic comparisons of estimators with regard to the PCC and with reference to general parametric families, and the current study pertains to this general objective.

The PCC, essentially a measure of pairwise comparisons, extends to comparisons within a suitable class of estimators only under additional restrictions such as equivariance (with respect to suitable groups of transformations), ancillarity (of the difference of pairs of estimators in the class) or asymptotic first-order representation (yielding the asymptotic normality) and so on. As mentioned before, the usual definition of Pitman closeness extends in a natural way to cover the second-order case (see Section 2), but it has a natural appeal only when the competing estimators are first-order efficient, that is, they are BAN in a general sense. For this reason, and given the affinity of BAN estimators to the classical maximum likelihood estimators (MLE's), in the current study we confine ourselves to the class of estimators which essentially adhere 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 light of the PCC. We confine ourselves to the case of a single parameter of interest, although the results in Section 3 pertain to a more general case where there are some nuisance parameters.

2. The one-parameter case. Let $\{X_i;\, i \geq 1\}$ be a sequence of independent and identically distributed (i.i.d.) random variables (r.v.'s) with a distribution function (d.f.) F admitting a density function $f(x;\theta)$ with respect to some sigma-finite measure μ, where θ is an unknown scalar parameter; Θ, the parameter space for θ, is the real line $\mathbb{R}$ or some open subset of $\mathbb{R}$. We adhere to the Assumptions in Bhattacharya and Ghosh (1978) with $s = 3$ (in their notation) and with $f(\cdot;\theta)$ and $g(\cdot;\theta)$ in their notation interpreted, respectively, as $\log f(\cdot;\theta)$ and $f(\cdot;\theta)$ in our notation. Let $\hat{\theta}\,(=\hat{\theta}_n)$ be the MLE of θ based on $X_1,\dots,X_n$ (where n is the sample size), defined in the sense of Theorem 3 of Bhattacharya and Ghosh (1978). Then, along the lines of Ghosh and Sinha (1981) [see also

Pfanzagl and Wefelmeyer (1978)], we consider a class $\mathcal{C}$ of estimators of the form

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

where the following hold:

1. $Q = d(\theta) + o_p(1)$ under θ, $d(\cdot)$ being continuously differentiable with a functional form independent of n;
2. for every positive ε, free from n, and every $\theta \subset \Theta$,

$$P_\theta\{|Q - d(\theta)| > \varepsilon\} = o(n^{-1/2}) \quad \text{as } n \to \infty.$$

The class $\mathcal{C}$ is quite large. In particular [see, Ghosh and Sinha (1981)], by Theorem 3 of Bhattacharya and Ghosh (1978), it includes all estimators of the form $\widehat{\theta} + n^{-1}d(\widehat{\theta})$, where $d(\cdot)$ satisfies condition 1.

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

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

$$P_\theta\Big\{|T_n^* - \theta| < |T_n - \theta|\Big\} = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)(2\pi n)^{-1/2}I^{1/2}\,\mathrm{sgn}\{d(\theta) - d^*(\theta)\} \times \left\{d(\theta) + d^*(\theta) - \tfrac{1}{3}I^{-2}L_{1.1.1}\right\} + o(n^{-1/2}).$$

Lemma 2.1 is similar to a result in Severini (1992), who considered biased and bias-corrected estimators of a one-dimensional interest parameter with bias defined in the usual sense of expectation. For θ such that $d(\theta) > d^*(\theta)$, this lemma can be proved if one notes that, by conditions 1 and 2 (pertaining to the class $\mathcal{C}$), $P_\theta\{Q < Q^*\} = o(n^{-1/2})$, and hence,

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

where $\xi_n = (nI)^{1/2}(\widehat{\theta}_n - \theta) + (1/2)(n^{-1}I)^{1/2}\{d(\theta) + d^*(\theta)\}$, and then employs an Edgeworth expansion for the distribution of ξ_n under θ. A similar proof holds for $d(\theta) < d^*(\theta)$.

Lemma 2.1 does not cover the case of θ such that $d(\theta) = d^*(\theta)$ since then neither $P_\theta\{Q < Q^*\}$ nor $P_\theta\{Q > Q^*\}$ may be $o(n^{-1/2})$, in general. For such θ, it may be possible to discriminate between T_n and T_n^* even at the first order of comparison. For example, under the univariate normal model with unknown mean θ ($\in R$) and variance 1 (known), let $T_n^* = \widehat{\theta}$ and $T_n = \widehat{\theta} + n^{-1}\widehat{\theta}^2 + n^{-3/2}\beta\widehat{\theta}$, where β is a constant which is free from n. Then $d^*(\theta) = 0$, $d(\theta) = \theta^2$ and $d(\theta) = d^*(\theta)$ only when $\theta = 0$. At $\theta = 0$, it can be shown that $\lim_{n\to\infty} P_0\{|T_n^* - 0|$

$< |T_n - \theta|\}$ can assume any value between 0 and 1 depending on the choice of β. Anyway, in the sequel we will be comparing estimators with distinguishable stochastic expansions up to $o_p(n^{-1})$ [i.e., with distinct $d(\cdot)$]; for this purpose, Lemma 2.1 plays a vital role.

THEOREM 2.1. *Let T_{0n} and T_n be distinct members of $\mathcal{C}$ such that $T_{0n} = \hat{\theta} + n^{-1}Q_0$ and $T_n = \hat{\theta} + n^{-1}Q$ with $Q = d(\theta) + o_p(1)$, $Q_0 = d_0(\theta) + o_p(1)$, under θ, and*

$$d_0(\theta) = \left(\tfrac{1}{6}\right) I^{-2} L_{1,1,1}. \tag{2.1}$$

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

$$P_\theta\left\{|T_{0n} - \theta| < |T_n - \theta|\right\} = \tfrac{1}{2} + \left(\tfrac{1}{2}\right)(2\pi n)^{-1/2} I^{1/2} |d(\theta) - d_0(\theta)| + o\left(n^{-1/2}\right), \tag{2.2}$$

so that

$$\lim_{n\to\infty}\left[n^{1/2}\left\{P_\theta\left\{|T_{0n} - \theta| < |T_n - \theta|\right\} - 1/2\right\}\right] = \left(\tfrac{1}{2}\right)\left(I/2\pi\right)^{1/2} |d(\theta) - d_0(\theta)| > 0. \tag{2.3}$$

The proof is a direct consequence of Lemma 2.1, and hence is omitted.

Let us discuss the implications of Theorem 2.1 by introducing the notion of second-order Pitman admissibility. An estimator $T_n = \hat{\theta} + n^{-1}Q$ ($\in \mathcal{C}$), with $Q = d(\theta) + o_p(1)$, under θ, will be called second-order Pitman inadmissible (SOPI) in $\mathcal{C}$ if there exists an estimator $T_n^* = \hat{\theta} + n^{-1}Q^*$ ($\in \mathcal{C}$), with $Q^* = d^*(\theta) + o_p(1)$ under θ and $d^*(\theta)$ not identically equal to $d(\theta)$, such that T_n^* is superior to T_n with regard to the second-order Pitman closeness in the following sense. Let $\alpha_{n1}(\theta) = P_\theta\{|T_n^* - \theta| < |T_n - \theta|\} - \frac{1}{2}$ and $\alpha_{n2}(\theta) = n^{1/2}\alpha_{n1}(\theta)$. Then (a) $\lim_{n\to\infty} \alpha_{n2}(\theta) \geq 0$, for each θ for which the limit exists, and (b) $\lim_{n\to\infty} \alpha_{n1}(\theta)$ exists and $\lim_{n\to\infty} \alpha_{n1}(\theta) \geq 0$, for each θ for which $\lim_{n\to\infty} \alpha_{n2}(\theta)$ does not exist, the inequality being strict for some θ ($\in \Theta$) either in (a) or (b).

An estimator T_n ($\in \mathcal{C}$) will be called second-order Pitman admissible (SOPA) in $\mathcal{C}$ if it is not SOPI in $\mathcal{C}$. An implication of Theorem 2.1 is that the estimator T_{0n} considered there is SOPA in $\mathcal{C}$; as a referee suggests, T_{0n} can as well be interpreted as representing a subclass of $\mathcal{C}$ consisting of estimators having the expansion $\hat{\theta} + n^{-1}[d_0(\theta) + o_p(1)]$, under θ, where $d_0(\theta)$ is given by (2.1). In particular, it follows that the estimator $\hat{\theta} + n^{-1}d_0(\hat{\theta})$ ($\in \mathcal{C}$) is SOPA in $\mathcal{C}$. Note that T_{0n}, considered in Theorem 2.1, is second-order median unbiased in the sense that $P_\theta\{T_{0n} \geq \theta\} = \frac{1}{2} + o(n^{-1/2})$ for every $\theta \in \Theta$, as one can prove by using an Edgeworth expansion for the distribution of $(nI)^{1/2}(T_{0n} - \theta)$ under θ. Hence, the second-order Pitman admissibility of T_{0n} is comparable with the exact findings in Ghosh and Sen (1989), who proved, under certain conditions, an optimal property of median unbiased estimators with regard to Pitman closeness. It also follows from Theorem 2.1 that an estimator $T_n = \hat{\theta} + n^{-1}Q$ ($\in \mathcal{C}$), with $Q = d(\theta) + o_p(1)$ under θ, and $d(\theta) \neq d_0(\theta)$, for each θ, will be SOPI in $\mathcal{C}$.

Thus, Theorem 2.1 yields a class of SOPA estimators, namely, that represented by T_{0n}, 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 θ given $X_1, \dots, X_{n_0}$ is proper, it can be shown from Theorem 2.1 that the posterior median of θ under the Jeffreys prior is SOPA in $\mathbb{C}$. [c.f. 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_{0n} depicted in Theorem 2.1 is in fact much stronger than the second-order Pitman admissibility. It implies that a rival estimator $T_n = \widehat{\theta} = n^{-1}Q$ ($\in \mathbb{C}$), with $Q = d(\theta) + o_p(1)$ under θ and $d(\theta)$ not identically equal to $d_0(\theta)$, will be inferior to T_{0n}, with regard to second-order Pitman closeness, for each θ satisfying $d(\theta) \neq d_0(\theta)$. Incidentally, for θ such that $d(\theta) = d_0(\theta)$, additional regularity conditions (e.g., asymptotic ancillarity) may be required to depict clearly the relative picture.

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

EXAMPLE 2.1. Let $f(x;\theta)$ be the univariate normal density with mean θ and variance θ^2, where $\theta \in \mathbb{R}^+$. Then $I = 3\theta^{-2}, L_{1\,1,1} = 14\theta^{-3}$ and, by (2.1), $d_0(\theta) = (7/27)\theta$. Let $T_{0n} = \widetilde{\theta}(1 + 7/27n)$ and $T_n = \widetilde{\theta}(1 - 1/9n)$. By Theorem 2.1, T_{0n} is SOPA in $\mathbb{C}$ while T_n is not so. On the other hand, proceeding as in Ghosh and Sinha (1981), we obtain that T_n is SOAE in their sense, while T_{0n} is not.

REMARK 4. A multiparameter extension of Theorem 2.1 can be formulated along the same line. However, in such a case, a quadratic norm involves a nonnegative definite (nnd) matrix, and, in general, the dominance results depend on the choice of such a matrix. Sen (1986) incorporated the Fisher information matrix in this formulation (albeit in a first-order setup), and recently Sen (1994) has shown that this result extends generally to a larger class of loss functions. It seems quite natural to formulate analogous second-order properties, and we would like to pursue the same in a follow-up study. In the multiparameter case, one needs to take into account the classical Stein phenomenon [see, Sen, Kubokawa and Saleh (1989)], and the picture becomes more complex.

3. A case with nuisance parameter(s). We proceed now to consider a more general case where the density $f(\cdot;\cdot)$ involves some nuisance parameters (in addition to the parameter of interest). For the sake of notational simplicity, we describe the results with a one-dimensional nuisance parameter (along with a one-dimensional parameter of interest). The treatment for a multidimensional nuisance parameter will be exactly similar, and only the notational system will become more involved.

Consider a sequence $\{X_i;\, i \geq 1\}$ of i.i.d. r.v.'s with a density $f(x;\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\theta_1, \theta_2)'$, θ_1 is the parameter of interest and θ_2 is the nuisance parameter. It is assumed that $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^2$, and other regularity assumptions are very similar to those in Section 2. Let us formulate the per-observation Fisher information matrix at $\boldsymbol{\theta}$ as

$$E_{\theta}\Big\{(\partial/\partial\boldsymbol{\theta})\log f(X;\boldsymbol{\theta})\cdot(\partial/\partial\boldsymbol{\theta}')\log f(X;\boldsymbol{\theta})\Big\} = \boldsymbol{I} = \big((I_{ij})\big)_{i,j=1,2}$$

and assume that $\boldsymbol{I}$ is positive definite at each $\boldsymbol{\theta} \subset \Theta$. Since θ_1 is scalar, we may suppose (without any loss of generality) that global parametric orthogonality holds, that is, $I_{12} = 0 = I_{21}$, for all $\boldsymbol{\theta} \in \Theta$ [see, Cox and Reid (1987)]. Let

$$S_{1.1.1} = E_{\boldsymbol{\theta}}\Big\{\big(D_1 \log f(X;\boldsymbol{\theta})\big)^3\Big\} \quad \text{and} \quad S_{122} = E_{\boldsymbol{\theta}}\Big\{\big(D_1 D_2^2 \log f(X;\boldsymbol{\theta})\big)\Big\},$$

where D_i is the operator of partial differentiation with respect to θ_i, $i = 1, 2$. Note that $I_{11}, I_{22}, S_{1.1.1}$ and S_{122} are all functions of $\boldsymbol{\theta}$.

Based on a sample $X_1, \dots, X_n$ of size n, let $\widehat{\boldsymbol{\theta}} = (\widehat{\theta}_1, \widehat{\theta}_2)'$ be the MLE of $\boldsymbol{\theta}$. As an analogue of the class $\mathbb{C}$ considered in Section 2, we consider here a class $\mathbb{C}^*$ of estimators of θ_1 of the form $T_n = \widehat{\theta}_1 + n^{-1}Q$, where the following hold:

1. $Q = d(\boldsymbol{\theta}) + o_p(1)$ under $\boldsymbol{\theta}$, $d(\cdot)$ being a continuously differentiable function whose functional form is free from n;
2. for each positive ε, free from n, and each $\boldsymbol{\theta}$,

$$P_{\boldsymbol{\theta}}\{|Q - d(\boldsymbol{\theta})| > \varepsilon\} = o(n^{-1/2}) \quad \text{as } n \to \infty.$$

Define then [analogous to (2.1)]

$$d_0(\boldsymbol{\theta}) = \big(6I_{11}^2\big)^{-1}S_{1.1.1} + (2I_{11}I_{22})^{-1}S_{122}. \tag{3.1}$$

Then, analogously to Lemma 2.1 and Theorem 2.1, respectively, we present the following results. The proofs are omitted to save space. They involve an Edgeworth expansion and rest on computations of certain higher-order cumulants.

LEMMA 3.1. *Let T_n^* and T_n be distinct members of $\mathbb{C}^*$, such that $T_n^* = \widehat{\theta}_1 + n^{-1}Q^*$ and $T_n = \widehat{\theta}_1 + n^{-1}Q$, with $Q^* = d^*(\boldsymbol{\theta}) + o_p(1)$ and $Q = d(\boldsymbol{\theta}) + o_p(1)$, under $\boldsymbol{\theta}$. Then, for each $\boldsymbol{\theta}$, such that $d(\boldsymbol{\theta}) \neq d^*(\boldsymbol{\theta})$,*

$$\begin{aligned} P_{\boldsymbol{\theta}}\Big\{\big|T_n^* - \theta_1\big| < \big|T_n - \theta_1\big|\Big\} = {} & 1/2 + (I_{11}/8\pi n)^{1/2}\operatorname{sgn}\{d(\boldsymbol{\theta}) - d^*(\boldsymbol{\theta})\} \\ & \times \{d(\boldsymbol{\theta}) + d^*(\boldsymbol{\theta}) - 2d_0(\boldsymbol{\theta})\} + o(n^{-1/2}), \end{aligned} \tag{3.2}$$

where $d_0(\boldsymbol{\theta})$ is defined by (3.1).

THEOREM 3.1. *Let T_{0n} and T_n be distinct members of $\mathcal{C}^*$ such that $T_{0n} = \widehat{\theta}_1 + n^{-1}Q_0$ and $T_n = \widehat{\theta}_1 + n^{-1}Q$ with $Q_0 = d_0(\boldsymbol{\theta}) + o_p(1)$ and $Q = d(\boldsymbol{\theta}) + o_p(1)$, under $\boldsymbol{\theta}$, where $d_0(\boldsymbol{\theta})$ is given by* (3.1). *Then, for each $\boldsymbol{\theta}$ such that $d(\boldsymbol{\theta}) \neq d_0(\boldsymbol{\theta})$,*

$$(3.3)\quad P_{\boldsymbol{\theta}}\{|T_{0n} - \theta_1| < |T_n - \theta_1|\} = 1/2 + (I_{11}/8\pi n)^{1/2}|d(\boldsymbol{\theta}) - d_0(\boldsymbol{\theta})| + o(n^{-1/2}),$$

so that

$$(3.4)\quad \begin{aligned}\lim_{n\to\infty} &\left[n^{1/2}P_{\boldsymbol{\theta}}\{|T_{0n} - \theta_1| < |T_n - \theta_1|\} - 1/2\right]\\ &= (I_{11}/8\pi)^{1/2}|d(\boldsymbol{\theta}) - d_0(\boldsymbol{\theta})| > 0.\end{aligned}$$

It can be shown that an estimator T_{0n}, as introduced in Theorem 3.1, is second-order median unbiased, that is, $P_{\boldsymbol{\theta}}\{T_{0n} > \theta_1\} = 1/2 + o(n^{-1/2})$ for each $\boldsymbol{\theta}$. Lemma 3.1 is a powerful tool for comparing estimators in $\mathcal{C}^*$. In the present setup, defining SOPA and SOPI estimators in $\mathcal{C}^*$ along the lines of Section 2, it follows from Theorem 3.1 that an estimator T_{0n}, as introduced in the theorem, is SOPA in $\mathcal{C}^*$. Also, an estimator $T_n = \widehat{\theta}_1 + n^{-1}Q (\in \mathcal{C}^*)$, with $Q = d(\boldsymbol{\theta}) + o_p(1)$ under $\boldsymbol{\theta}$, and $d(\boldsymbol{\theta}) \neq d_0(\boldsymbol{\theta})$, for each $\boldsymbol{\theta} \in \Theta$, will be SOPI in $\mathcal{C}^*$. In continuation of Remark 1 (in Section 2), under suitable conditions, it can be shown by using Theorem 3.1 that the posterior median of θ_1 under a prior with density proportional to $I_{11}^{1/2}$ [where the constant of proportionality may involve θ_2 but not θ_1; cf. Tibshirani (1989)], will be SOPA in $\mathcal{C}^*$.

EXAMPLE 3.1. Let $f(x;\boldsymbol{\theta})$ represent the univariate normal density with mean θ_2 and variance θ_1, so that $\boldsymbol{\theta} \in \mathbb{R} \times \mathbb{R}^+$. Under this parameterization, the global parametric orthogonality, as mentioned before, holds. Here $I_{11} = (2\theta_1^2)^{-1}$, $I_{22} = \theta_1^{-1}$, $S_{1.1.1} = \theta_1^{-3}$ and $S_{122} = \theta_1^{-2}$, so that, by (3.1), $d_0(\boldsymbol{\theta}) = (5/3)\theta_1$. The MLE of θ_1 based on $X_1, \ldots, X_n$ is given by $\widehat{\theta}_1 = n^{-1}\Sigma_{i=1}^n (X_i - \overline{X}_n)^2$, where $\overline{X}_n = n^{-1}\Sigma_{i=1}^n X_i$ is the MLE of θ_2. Allied to the MLE $T_n = \widehat{\theta}_1$ is the usual unbiased estimator $T_n^* = n(n-1)^{-1}T_n = \widehat{\theta}_1(1 + (n-1)^{-1})$, whereas the median unbiased estimator of θ_1 is given by $T_{0n} = n(m_{n-1}^{-1})T_n$, where m_{n-1} is the median of the central chi-squared distribution with $n-1$ degrees of freedom. It is well known that, for large n,

$$(3.5)\quad m_n = n - 2/3 + (32/405n) + O(n^{-2}),$$

so that we may even replace m_{n-1} by $n - 5/3$, and define $T_{0n} = \widehat{\theta}_1(1 + 5/3n)$. Note that all these estimators belong to the class $\mathcal{C}^*$, and T_{0n} is the Pitman closest one in the sense of Ghosh and Sen (1989, 1991). Moreover, the unbiased estimator T_n^* can as well be obtained by maximizing the conditional likelihood of Cox and Reid (1987). We have then $T_{0n} = \widehat{\theta}_1 + n^{-1}Q_0$, $T_n = \widehat{\theta}_1 + n^{-1}Q$ and $T_n^* = \widehat{\theta}_1 + n^{-1}Q^*$ with $Q_0 = d_0(\boldsymbol{\theta}) + o_p(1)$, $Q = d(\boldsymbol{\theta}) = 0$ and $Q^* = d^*(\boldsymbol{\theta}) + o_p(1)$, where $d_0(\boldsymbol{\theta}) = (5/3)\theta_1$ and $d^*(\boldsymbol{\theta}) = \theta_1$. Moreover, both $d(\boldsymbol{\theta})$ and $d^*(\boldsymbol{\theta})$ are different from each other and from $d_0(\boldsymbol{\theta})$, for each $\boldsymbol{\theta}$. As such, we conclude from Theorem 3.1 that T_{0n} is SOPA in $\mathcal{C}^*$ and both T_n and T_n^* are SOPI in $\mathcal{C}^*$ and are dominated by T_{0n} (under the second-order PCC). Next, in order to compare the MLE T_n and the unbiased version T_n^* (under PCC), we write

$$\alpha_{n2}(\theta) = n^{1/2}\Big\{P_\theta\Big\{|T_n^* - \theta_1| < |T_n - \theta_1|\Big\} - 1/2\Big\},$$

and from Lemma 3.1 we obtain that

$$\lim_{n\to\infty} \alpha_{n2}(\theta) = 7/(12\sqrt{\pi}) = 0.3291 \quad (> 0) \quad \text{for each } \theta. \tag{3.6}$$

This shows that T_n^* is superior to T_n with respect to the second-order Pitman closeness. Exact computations are not hard in this example, and the exact values of $\alpha_{n2}(\theta)$ for $n = 3, 5, 7$ and 9 can be seen to be equal to 0.3443, 0.3372, 0.3346 and 0.3333, respectively, for each θ [cf. Rao (1981)]. Since even for small values of n the values of $\alpha_{n2}(\theta)$ are quite close to the asymptotic value in (3.6), the asymptotic results in Theorem 3.1 appear to be reasonably good indicators of the small to moderate sample behavior of the estimators in this specific case.

In general, each of our results is based on a second-order Edgeworth expansion which is used to approximate the probability that a certain random variable is less than 0. Since Edgeworth expansions tend to be very accurate in this central part of the range, this may explain why, as observed in the last example, results for small to moderate samples have a tendency essentially to follow the pattern expected from asymptotic considerations.

Acknowledgment. The authors are grateful to the reviewers for their helpful comments on the manuscript.

REFERENCES

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of formal Edgeworth expansions. *Ann. Statist.* **6** 434–451.

BLYTH, C. R. and PATHAK, P. K. (1985). Does an estimator's distribution suffice? In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.) **1** 45–52. Wadsworth, Monterey, CA.

COX, D. R. and REID, N. (1987). Parametric orthogonality and approximate conditional inference (with discussion). *J. Roy. Statist. Soc. Ser. B* **49** 1–39.

GHOSH J. K. (1994). *Higher Order Asymptotics*. IMS, Hayward, CA.

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. *Comm. Statist. Theory Methods* **20** 3659–3678.

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. (1993). *Pitman's Measure of Closeness: A Comparison of Statistical Estimators*. SIAM, Philadelphia.

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. Csörgő, D. A. Dawson, J. N. K. Rao and A. K. M. E. Saleh, eds.) 123–143. North-Holland, Amsterdam.

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

SEN P. K. (1992). Pitman closeness of statistical estimators: latent years and the renaissance. In *Current Issues in Statistical Inference: Essays in Honor of D. Basu* (M. Ghosh and P. K. Pathak, eds.) 52–74. IMS, Hayward, CA.

SEN, P. K (1994). Isomorphism of quadratic norm and PC ordering of estimators admitting first order AN representations. *Sankhyā Ser. A* **56**. To appear.

SEN, P. K., KUBOKAWA, T. and SALEH, A. K. M. E. (1989). The Stein paradox in the sense of Pitman closeness. *Ann. Statist.* **17** 1375–1384.

SEVERINI, T. A. (1991). A comparison of the maximum likelihood estimator and the posterior mean in the single parameter case. *J. Amer. Statist. Assoc.* **86** 997–1000.

SEVERINI, T. A. (1992). A comparison of biased and bias-corrected estimators. Unpublished manuscript.

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. Ser. B* **25** 318–329.

JAYANTA K. GHOSH
DIVISION OF MATHEMATICS
AND THEORETICAL STATISTICS
INDIAN STATISTICAL INSTITUTE
203 B.T. ROAD
CALCUTTA 700 035
INDIA

PRANAB K. SEN
DEPARTMENTS OF STATISTICS
AND BIOSTATISTICS
UNIVERSITY OF
NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27599-7400

RAHUL MUKERJEE
INDIAN INSTITUTE OF MANAGEMENT
JOKA, DIAMOND HARBOUR ROAD
POST OFFICE BOX NO. 16757,
ALIPORE POST OFFICE
CALCUTTA 700 027
INDIA