The Annals of Statistics
1980, Vol. 8, No. 3, 506–521

# SECOND ORDER EFFICIENCY OF THE MLE WITH RESPECT TO ANY BOUNDED BOWL-SHAPED LOSS FUNCTION

By J. K. Ghosh[1], B. K. Sinha[1] and H. S. Wieand

*University of Pittsburgh*

Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables, each having density $f(x, \theta_0)$ where $\{f(x, \theta)\}$ is a family of densities with respect to a dominating measure $\mu$. Suppose $n^{\frac{1}{2}}(\hat{\theta} - \theta)$ and $n^{\frac{1}{2}}(T - \theta)$, where $\hat{\theta}$ is the mle and $T$ is any other efficient estimate, have Edgeworth expansions up to $o(n^{-1})$ uniformly in a compact neighbourhood of $\theta_0$. Then (under certain regularity conditions) one can choose a function $c(\theta)$ such that $\hat{\theta}' = \hat{\theta} + c(\hat{\theta})/n$ satisfies

$$P_{\theta_0}\{-x_1 < n^{\frac{1}{2}}(\hat{\theta}' - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\}$$
$$> P_{\theta_0}\{-x_1 < n^{\frac{1}{2}}(T - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\} + o(n^{-1}),$$

for all $x_1, x_2 > 0$. This result implies the second order efficiency of the mle with respect to any bounded loss function $L_n(\theta, a) = h(n^{\frac{1}{2}}(a - \theta))$, which is bowl-shaped i.e., whose minimum value is zero at $a - \theta = 0$ and which increases as $|a - \theta|$ increases. This answers a question raised by C. R. Rao (Discussion on Professor Efron's paper).

**1. Introduction.** After the pioneering work of Fisher (1925) and Rao (1961, 1962, 1963), second order efficiency has been discussed recently by Ghosh and Subramanyam (1974) and Efron (1975) for curved exponential families. As pointed out by Ghosh (1975) the approach and results of these authors are different from those of Pfanzagl (1973); results of Akahira (1975) seem to be similar to those of Pfanzagl (1973). The spirit of Pfanzagl (1975) seems nearer to ours.

The main motivation for the present work came from a problem posed by Rao (1975): Does the second order efficiency of the mle continue to hold if one has more general loss functions than the squared error loss? Under certain assumptions we prove the following result (Theorem 1):

*There exists a function $c(\theta)$ such that $\hat{\theta}' = \hat{\theta} + c(\hat{\theta})/n$ satisfies*

$$P_{\theta_0}\{-x_1 < n^{\frac{1}{2}}(\hat{\theta}' - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\}$$
$$> P_{\theta_0}\{-x_1 < n^{\frac{1}{2}}(T - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\} + o(n^{-1})$$

*for all $x_1, x_2 > 0$ (at least one being positive) where $\hat{\theta}$ is the* mle *and $T$ any other efficient estimate.*

---

Received June 1977; revised October 1978.

[1]On leave from the Indian Statistical Institute, Calcutta, India.

*AMS* 1970 *subject classifications.* Primary 62B10, 62F20.

*Key words and phrases.* Maximum likelihood estimate, second order efficiency, Edgeworth expansion, Bayes test, Rao-Blackwell theorem, bowl-shaped loss function.

This provides, as proved in the Corollary, an affirmative answer to the above question for any bounded loss function $L_n(\theta, a) = h(n^{\frac{1}{2}}(a - \theta))$ which is bowl-shaped i.e., whose minimum value is zero at $a - \theta = 0$ and which increases as $|a - \theta|$ increases.

Our technique of proof is borrowed from that of Bahadur (1964), Rao (1963) and Weiss and Wolfowitz (1966). We associate with each (efficient) estimate $T$ a natural test of $H_1$: $\theta = \theta_0 - r_1/n^{\frac{1}{2}}$ vs. $H_2$: $\theta = \theta_0 + r_2/n^{\frac{1}{2}}$ and compare the performance with that of the Bayes test under a suitably chosen prior. It turns out that a test associated with the mle is Bayes up to $o(n^{-1})$. This fact is at the root of the inequality asserted in Theorem 1. It was pointed out in Ghosh and Subramanyam (1974) that a similar Bayes property explains the second order efficiency with respect to the squared error loss.

For curved exponential families dominated by the Lebesgue measure, a direct calculation shows $n^{\frac{1}{2}}(T - \theta)$ and $n^{\frac{1}{2}}(\hat{\theta}' - \theta)$ have the same first four cumulants up to $o(n^{-\frac{1}{2}})$; this means, as pointed out by Ghosh (1975), that Theorem 6 of Pfanzagl (1973) will fail to discriminate between these estimates. It will be shown elsewhere that when we consider terms up to $o(n^{-1})$, the second cumulant of $n^{\frac{1}{2}}(\hat{\theta}' - \theta)$ is smaller but the other three of the first four cumulants remain the same for both estimates. This leads to a straightforward proof of Theorem 1. It is also possible to show that $\hat{\theta}'$ can be interpreted as the conditional expectation of $T$ given $\hat{\theta}$ up to $o(n^{-1})$. This provides a Rao-Blackwell type justification of second order efficiency. We shall deal with these aspects elsewhere.

A major drawback of the present results is that they don't apply to the discrete case. In particular the curved multinomials which were the main source of inspiration to Fisher and Rao have to be excluded.

A proof of Theorem 1 in the vector parameter case where we replace the interval $(\theta - x_1/n^{\frac{1}{2}}$ to $\theta + x_2/n^{\frac{1}{2}})$ by $\theta + C/n^{\frac{1}{2}}$ (where $C$ is a convex set with the origin as an interior point) is under investigation. Pfanzagl and Wefelmeyer (1978) have proved this result for symmetric convex sets under stronger conditions on the estimates.

We present in Section 2 the notations, assumptions and the statement of the theorem. Section 3 contains a sketch of the proof of the theorem, with statements but not proofs of the necessary auxiliary results. Section 4 is devoted to these proofs.

**2. Main results.** Let $X_1, X_2, \cdots$ be a sequence of i.i.d. rv's distributed as $f(x, \theta) \in \mathcal{F} = \{f(x, \theta) : \theta \in \Theta\}$, a family of densities with respect to a dominating measure $\mu$. We shall require the following assumptions to hold throughout this paper.

ASSUMPTION 1. $\Theta$ is an open interval of the real line.

ASSUMPTION 2. $f(x, \theta)$ is jointly measurable in $x$ and $\theta$.

ASSUMPTION 3. For each $x$, $f(x, \theta)$ admits partial derivatives up to fifth order with respect to $\theta$ and these are continuous in $\theta$ for $\theta \in \Theta$.

ASSUMPTION 4. The measures corresponding to $\{f(x, \theta), \theta \in \Theta\}$ are mutually absolutely continuous.

ASSUMPTION 5. For all $\theta \in \Theta$,

$$E_\theta|\log f(X, \theta)| < \infty,$$

$$0 < I(\theta) = -E_\theta\left[\frac{\partial^2 \log f(X, \theta)}{\partial \theta^2}\right].$$

ASSUMPTION 6. For each $\theta_0 \in \Theta$, there exists a compact neighbourhood $\Theta_0 \ni \theta_0$ and functions $G(x)$, $H(x)$ satisfying

$$\left|\frac{\partial^i}{\partial \theta^i} \log f(x, \theta)\right| < G(x), \qquad i = 1, 2, 3, 4$$
$$< H(x) \qquad i = 5$$

for $\theta \in \Theta_0$ and

$$\operatorname{Sup}_{\theta \in \Theta_0} E_\theta\left[G(X)^4\right] < \infty, \quad \operatorname{Sup}_{\theta \in \Theta_0} E_\theta\left[H(X)\right] < \infty.$$

ASSUMPTION 7. For each $\theta_0 \in \Theta$, there exists a compact neighbourhood $\Theta_0 \ni \theta_0$ such that $n^{-\frac{1}{2}}\{L_2(\theta) + nI(\theta)\}/(\mu_{2,2}(\theta))^{\frac{1}{2}}$ and $n^{-\frac{1}{2}}\{L_3(\theta) - n\mu_{3,1}(\theta)\}/(\mu_{3,2}(\theta))^{\frac{1}{2}}$ under $\theta$ have Edgeworth expansions up to $o(n^{-1})$ uniformly for $\theta \in \Theta_0$ where $L_i(\theta) = \partial^i/\partial\theta^i \log l(\theta)$, $i = 1, 2, 3, 4$ with $l(\theta) = \prod_1^n f(x_i, \theta)$, the likelihood at $\theta$ and $n\mu_{3,1}(\theta) = E_\theta\{L_3(\theta)\}$, $n\mu_{3,2}(\theta) = E_\theta\{L_3(\theta) - n\mu_{3,1}(\theta)\}^2$, $n\mu_{2,2}(\theta) = E_\theta\{L_2(\theta) + nI(\theta)\}^2$.

ASSUMPTION 8. For any fixed $\theta_0 \in \Theta$, $n^{\frac{1}{2}}(\hat\theta - \theta)(I(\theta))^{\frac{1}{2}}$ where $\hat\theta$ is the mle has an Edgeworth expansion up to $o(n^{-1})$ uniformly in a compact neighbourhood $\Theta_0$ of $\theta_0$ in the following form:

$$P_\theta\left\{n^{\frac{1}{2}}(\hat\theta - \theta)(I(\theta))^{\frac{1}{2}} < x\right\} = \Phi(x) + \frac{\Phi_{1,x}(\theta)}{n^{\frac{1}{2}}} + \frac{\Phi_{2,x}(\theta)}{n} + o(n^{-1}) \tag{2.1}$$

where

$$\Phi(x) = \textstyle\int_{-\infty}^x e^{-t^2/2}\,dt/(2\pi)^{\frac{1}{2}}, \quad \phi(x) = \Phi'(x),$$

$$\Phi_{1,x}(\theta) = \int_{-\infty}^x \left\{K_{11}(\theta)H_1(z) + \frac{K_{21}(\theta)}{2}H_2(z) + \frac{K_{31}(\theta)}{6}H_3(z)\right\}\phi(z)\,dz,$$

$$\Phi_{2,x}(\theta) = \int_{-\infty}^x \left[K_{12}(\theta)H_1(z) + \frac{K_{22}(\theta)}{2}H_2(z) + \frac{K_{32}(\theta)}{6}H_3(z)\right.$$
$$+ \frac{K_{41}(\theta)}{24}H_4(z) + \frac{1}{2}\left\{K_{11}^2(\theta)H_2(z) + \frac{K_{21}^2(\theta)}{4}H_4(z)\right.$$

$$+\frac{K_{31}^2(\theta)}{36}H_6(z) + K_{11}(\theta)K_{21}(\theta)H_3(z) + \tfrac{1}{3}K_{11}(\theta)K_{31}(\theta)H_4(z)$$

$$+\tfrac{1}{6}K_{21}(\theta)K_{31}(\theta)H_5(z)\Big\}\Big]\phi(z)\,dz,$$

$$H_p(x)\phi(x) = \left(-\frac{d}{dx}\right)^p\phi(x)$$

and the term $o(n^{-1})$ in (2.1) is uniform in $\theta \in \Theta_0$.

REMARK. Assumption 7 holds if Assumption 6 holds and $Z_j = \partial^j \log f(X,\theta)/\partial\theta^j$, $j = 2, 3$ satisfy a uniform version of Cramer's condition like $c^*$ of Pfanzagl (1973, page 1012). A more general sufficient condition is given in Bhattacharya and Ghosh (1976). Conditions under which Assumption 8 is valid are given in Pfanzagl (1973), Chibisov (1973a, 1973b), Gusev (1976), and Bhattacharya and Ghosh (1976). The various $K_{ij}(\theta)$'s in the Edgeworth expansion come from formal expansions of the cumulants of $n^{\frac{1}{2}}(\hat{\theta} - \theta)(I(\theta))^{\frac{1}{2}}$ in powers of $n^{-\frac{1}{2}}$. Specifically, we have used the following formal expansion for the $r$th cumulant $K_r(\theta)$:

$$\begin{aligned}
K_r(\theta) &= \frac{K_{r,1}(\theta)}{n^{\frac{1}{2}}} + \frac{K_{r,2}(\theta)}{n} + o(n^{-1}), && r = 1, 3\\
&= 1 + \frac{K_{r,1}(\theta)}{n^{\frac{1}{2}}} + \frac{K_{r,2}(\theta)}{n} + o(n^{-1}), && r = 2\\
&= \frac{K_{r,1}(\theta)}{n} + o(n^{-1}), && r = 4\\
&= o(n^{-1}), && r > 5.
\end{aligned}$$

The main result of the paper is the following.

THEOREM 1. *In addition to Assumptions 1–8, assume that $T$ is an efficient estimate and for any fixed $\theta_0 \in \Theta$, $n^{\frac{1}{2}}(T - \theta)(I(\theta))^{\frac{1}{2}}$ has an Edgeworth expansion up to $o(n^{-1})$ uniformly for $\theta \in \Theta_0$, analogous to (2.1) with the $K$'s denoted as $\bar{K}$'s. Then there exists a function $c(\theta)$ such that $\hat{\theta}' = \hat{\theta} + c(\hat{\theta})/n$ satisfies*

(a) $$P_{\theta_0}\{-x < n^{\frac{1}{2}}(\hat{\theta}' - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x\}$$
$$> P_{\theta_0}\{-x < n^{\frac{1}{2}}(T - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x\} + o(n^{-1})$$
*for all $x > 0$;*

(b) $$P_{\theta_0}\{-x_1 < n^{\frac{1}{2}}(\hat{\theta}' - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\}$$
$$> P_{\theta_0}\{-x_1 < n^{\frac{1}{2}}(T - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\} - o(n^{-1})$$

*for all $x_1, x_2 > 0$ (at least one of $x_1, x_2$ being positive) provided the condition $K_{12}(\theta_0) = K_{32}(\theta_0) = \bar{K}_{12}(\theta_0) = \bar{K}_{32}(\theta_0) = 0$ is satisfied in case $K_{21}(\theta_0) = \bar{K}_{21}(\theta_0)$.*

REMARK. It will be proved elsewhere that the condition in part (b) is satisfied for curved exponential families if $T$ is Fisher-consistent and thrice continuously differentiable.

**3. Sketch of proof of the theorem.** In this section we will provide a sketch of the proof of the theorem. In the sequel we state also the necessary auxiliary results whose proofs will be given in the next section.

For fixed $\theta_0 \in \Theta$ consider the problem of finding the Bayes test $\delta$ for the two-decision problem of accepting $H_1$: $\theta = \theta_0 - r_1/n^{\frac{1}{2}}$ or $H_2$: $\theta = \theta_0 + r_2/n^{\frac{1}{2}}$ on the basis of $\mathbf{X} = (X_1, \cdots, X_n)$. Here $r_1, r_2 > 0$ with at least one being positive. Assume $0 - 1$ loss function and let

$$C_1 = \exp(-r_2^2 I(\theta_0)/2)/\{\exp(-r_2^2 I(\theta_0)/2) + \exp(-r_1^2 I(\theta_0)/2)\}$$

$$C_2 = \exp(-r_1^2 I(\theta_0)/2)/\{\exp(-r_2^2 I(\theta_0)/2) + \exp(-r_1^2 I(\theta_0)/2)\}$$

denote the respective prior probability of $H_1$ and $H_2$ being true. The Bayes test $\delta$ accepts $H_1(H_2)$ if $\psi(\theta_0) > (<) 0$ where

$$\psi(d) = \log\frac{C_1}{C_2} + L\left(d - \frac{r_1}{n^{\frac{1}{2}}}\right) - L\left(d + \frac{r_2}{n^{\frac{1}{2}}}\right). \tag{3.1}$$

We now demonstrate that a test based on the mle $\hat{\theta}$ is nearly Bayes. Towards this end, consider the class $\mathcal{C}$ of decision rules which may be described by

$$\text{Accept } H_1(H_2) \text{ if } T(\mathbf{x}) < (>) \theta_0$$

for some measurable function $T(\mathbf{x})$ of the data $\mathbf{x} = (x_1, \cdots, x_n)$. The following result will be proved in the next section.

PROPOSITION 1. *For large $n$, with probability $1 - o(n^{-1})$ under both $H_1$ and $H_2$, the Bayes test $\delta$ belongs to $\mathcal{C}$ with $T(\mathbf{x}) = d^*(\mathbf{x})$ (a unique solution of $\psi(d) = 0$ in $A_n$ defined in the next section).*

It will be seen that $d^*(\mathbf{x})$ does not differ much from the mle $\hat{\theta}$. In fact, we will show $d^*(\mathbf{x})$ satisfies for large $n$

$$|d^*(\mathbf{x}) - \hat{\theta}| = 0\left(\frac{(\log n)^3}{n}\right) \qquad \text{for } \mathbf{x} \in \tilde{C}_n \tag{3.2}$$

with

$$P_{H_i}(\tilde{C}_n) = 1 - o(n^{-1}),\ i = 1, 2, \text{ for a suitably chosen } \tilde{C}_n.$$

Now under $0 - 1$ loss, the Bayes risk $R(T)$ of the rule in $\mathcal{C}$ based on $T(\mathbf{x})$ can be calculated as

$$R(T) = C_1 P_{H_1}(T(\mathbf{X}) > \theta_0) + C_2 P_{H_2}(T(\mathbf{X}) < \theta_0) \tag{3.3}$$

$$= \int \frac{C_1 l_1 I(T(\mathbf{x}) > \theta_0) + C_2 l_2 I(T(\mathbf{x}) < \theta_0)}{C_1 l_1 + C_2 l_2}\, dP$$

where $l_1 = l(\theta_0 - r_1/n^{\frac{1}{2}})$, $l_2 = l(\theta_0 + r_2/n^{\frac{1}{2}})$, $P = C_1P_1 + C_2P_2$ with $P_i$ = the probability measure under $H_i$, $i = 1, 2$. $I(T(\mathbf{X}) > \theta_0)$ denotes the indicator of the event that $T(\mathbf{X}) > \theta_0$ and $I(T(\mathbf{X}) < \theta_0)$ is the indicator of the complementary event. Since $I(T(\mathbf{X}) > \theta_0) = 1 - I(T(\mathbf{X}) < \theta_0)$,

$$R(T) = \int \frac{C_1l_1 - C_2l_2}{C_1l_1 + C_2l_2} I(T(\mathbf{x}) > \theta_0)\, dP + \int \frac{C_2l_2}{C_1l_1 + C_2l_2}\, dP.$$

If $T^*(\mathbf{x})$ is another measurable function of $\mathbf{x}$, the corresponding rule has risk $R(T^*)$ which obeys a similar formula and the difference between the two risks obeys

$$\begin{aligned}|R(T) - R(T^*)| &= \left|\int \frac{C_1l_1 - C_2l_2}{C_1l_1 + C_2l_2}\left(I(T(\mathbf{x}) > \theta_0) - I(T^*(\mathbf{x}) > \theta_0)\right) dP\right| \\ &< \int \left|\frac{C_1l_1 - C_2l_2}{C_1l_1 + C_2l_2}\right| I((T(\mathbf{x}) - \theta_0)(T^*(\mathbf{x}) - \theta_0) < 0)\, dP \qquad (3.4) \\ &= \int_{D_n} \left|\frac{e^{\psi(\theta_0)} - 1}{e^{\psi(\theta_0)} + 1}\right| dP,\end{aligned}$$

where $\psi$ is defined in (3.1), $D_n$ is the region where $T(\mathbf{x})$ and $T^*(\mathbf{x})$ lie on opposite sides of $\theta_0$. Using (3.2), (3.4) and some other related results, we will prove in Section 4:

LEMMA 3.

$$|R(d^*) - R(\hat{\theta})| = o(n^{-\frac{1}{2}}).$$

This result implies that the test in $\mathcal{C}$ based on the mle $\hat{\theta}$ is Bayes up to $o(n^{-\frac{1}{2}})$. Considering now a test in $\mathcal{C}$ based on an efficient estimate $T(\mathbf{x})$, it follows that $R(T) - R(\hat{\theta}) > o(n^{-\frac{1}{2}})$ for all $r_1, r_2 > 0$ (at least one being positive) or equivalently, using the expression of $R(T)$ as given in the first equality in (3.3),

$$\begin{aligned}(3.5)\quad & C_1P_{\theta_0-(r_1/n^{\frac{1}{2}})}(\hat{\theta} < \theta_0) + C_2P_{\theta_0+(r_2/n^{\frac{1}{2}})}(\hat{\theta} > \theta_0) \\ &> C_1P_{\theta_0-(r_1/n^{\frac{1}{2}})}(T < \theta_0) + C_2P_{\theta_0+(r_2/n^{\frac{1}{2}})}(T > \theta_0) + o(n^{-\frac{1}{2}})\end{aligned}$$

for all $r_1, r_2 > 0$ (at least one being positive). The idea now is to evaluate the probabilities involved in (3.5) using Assumption 8 and also the assumption in Theorem 1 regarding Edgeworth expansions of $\hat{\theta}$ and $T$ and to conclude that certain relations regarding the cumulants of the two distributions hold in view of (3.5). Towards this end, note that straightforward computations yield the following:

$$\begin{aligned}P_{\theta_0-r_1/n^{\frac{1}{2}}}\{\hat{\theta} < \theta_0\} &= \Phi\left(r_1\left(I\left(\theta_0 - r_1/n^{\frac{1}{2}}\right)\right)^{\frac{1}{2}}\right) + \Phi_{1,\, r_1(I(\theta_0))^{\frac{1}{2}}}(\theta_0)/n^{\frac{1}{2}} \\ &\quad + \frac{1}{n}\left\{\Phi_{2,\, r_1(I(\theta_0))^{\frac{1}{2}}}(\theta_0) - \left[\frac{r_1^2}{2}\cdot\frac{I'(\theta_0)}{(I(\theta_0))^{\frac{1}{2}}} K_{11}(\theta_0) H_1\left(r_1(I(\theta_0))^{\frac{1}{2}}\right)\right.\right.\end{aligned}$$

$$+\frac{r_1^2}{4}\cdot\frac{I'(\theta_0)}{(I(\theta_0))^{\frac{1}{2}}}K_{21}(\theta_0)H_2\left(r_1(I(\theta_0))^{\frac{1}{2}}\right) \tag{3.6}$$

$$+\frac{r_1^2}{12}\cdot\frac{I'(\theta_0)}{(I(\theta_0))^{\frac{1}{2}}}K_{31}(\theta_0)H_3\left(r_1(I(\theta_0))^{\frac{1}{2}}\right)\Bigg]\phi\left(r_1(I(\theta_0))^{\frac{1}{2}}\right)$$

$$-\int_{-\infty}^{r_1(I(\theta_0))^{\frac{1}{2}}}\left(r_1K'_{11}(\theta_0)H_1(z)+\frac{r_1}{2}K'_{21}(\theta_0)H_2(z)\right.$$

$$\left.+\frac{r_1}{6}K'_{31}(\theta_0)H_3(z)\right)\phi(z)\,dz\Big\}+o(n^{-1}),$$

and

$$P_{\theta_0+r_2/n^{\frac{1}{2}}}\{\hat\theta>\theta_0\}=1-\Phi\left(-r_2\left(I\left(\theta_0+r_2/n^{\frac{1}{2}}\right)\right)^{\frac{1}{2}}\right)+\Phi_{-1,r_2(I(\theta_0))^{\frac{1}{2}}}(\theta_0)/n^{\frac{1}{2}}$$

$$+\frac{1}{n}\Bigg[\Phi_{-2,r_2(I(\theta_0))^{\frac{1}{2}}}(\theta_0)-\Bigg\{\frac{r_1^2}{2}K_{11}(\theta_0)H_1\left(-r_2(I(\theta_0))^{\frac{1}{2}}\right)$$

$$+\frac{r_1^2}{4}K_{21}(\theta_0)H_2\left(-r_2(I(\theta_0))^{\frac{1}{2}}\right)+\frac{r_1^2}{12}K_{31}(\theta_0)H_3\left(-r_2(I(\theta_0))^{\frac{1}{2}}\right)\Bigg\} \tag{3.7}$$

$$\cdot\frac{I'(\theta_0)}{(I(\theta_0))^{\frac{1}{2}}}\cdot\phi\left(-r_2(I(\theta_0))^{\frac{1}{2}}\right)$$

$$+\int_{-r_2(I(\theta_0))^{\frac{1}{2}}}^{\infty}\left(r_2K'_{11}(\theta_0)H_1(z)+\frac{r_2}{2}K'_{21}(\theta_0)H_2(z)\right.$$

$$\left.+\frac{r_2}{6}K'_{31}(\theta_0)H_3(z)\right)\phi(z)\,dz\Bigg]+o(n^{-1})$$

where

$$\Phi_{-i,x}(\theta)=-\Phi_{i,-x}(\theta),\qquad i=1,2.$$

Using (3.6), (3.7) and similar expressions with $K_{ij}$'s replaced by $\bar K_{ij}$'s, which hold for the probabilities for $T$, (3.5) implies

$$\exp\left[-\frac{r_2^2}{2}I(\theta_0)\right]\int_{-\infty}^{r_1(I(\theta_0))^{\frac{1}{2}}}(K_{11}(\theta_0)H_1(z)+K_{21}(\theta_0)H_2(z)/2$$

$$+K_{31}(\theta_0)H_3(z)/6)\phi(z)\,dz$$

$$+\exp\left[-\frac{r_1^2}{2}I(\theta_0)\right]\int_{-r_2(I(\theta_0))^{\frac{1}{2}}}^{\infty}(K_{11}(\theta_0)H_1(z)+K_{21}(\theta_0)H_2(z)/2$$

$$+K_{31}(\theta_0)H_3(z)/6)\phi(z)\,dz$$

$$>\exp\left[-\frac{r_2^2}{2}I(\theta_0)\right]\int_{-\infty}^{r_1(I(\theta_0))^{\frac{1}{2}}}\left(\bar K_{11}(\theta_0)H_1(z)+\bar K_{21}(\theta_0)H_2(z)/2\right. \tag{3.8}$$

$$\left.+\bar K_{31}(\theta_0)H_3(z)/6\right)\phi(z)\,dz$$

$$+\exp\left[-\frac{r_1^2}{2}I(\theta_0)\right]\int_{-r_2(I(\theta_0))^{\frac{1}{2}}}^{\infty}\left(\bar K_{11}(\theta_0)H_1(z)+\bar K_{21}(\theta_0)H_2(z)/2\right.$$

$$\left.+\bar K_{31}(\theta_0)H_3(z)/6\right)\phi(z)\,dz$$

which after some simplification reduces to

$$\left(K_{21}(\theta_0) - \bar{K}_{21}(\theta_0)\right) + (I(\theta_0))^{\frac{1}{2}}\left(K_{31}(\theta_0) - \bar{K}_{31}(\theta_0)\right)(r_1 - r_2)/3 < 0$$

for all $r_1, r_2 > 0$ (at least one being positive), yielding the relations

$$K_{31}(\theta_0) = \bar{K}_{31}(\theta_0), \quad K_{21}(\theta_0) < \bar{K}_{21}(\theta_0). \tag{3.9}$$

We shall now use these relations to prove Theorem 1. We need one more lemma.

LEMMA 5. *Let* $\hat{\hat{\theta}} = \hat{\theta} - g(\hat{\theta})/n$ *where* $g(.)$ *is a twice differentiable function. For* $\theta \in \Theta_0$, $P_\theta\{n^{\frac{1}{2}}(\hat{\hat{\theta}} - \theta)(I(\theta))^{\frac{1}{2}} < x\}$ *is given by* (2.1) *with* $\Phi_{1,x}(\theta)$, $\Phi_{2,x}(\theta)$ *as given in Assumption* 8 *except that* $K_{11}(\theta)$ *and* $K_{22}(\theta)$ *are to be replaced by* $\mathrm{K}_{11}(\theta) = K_{11}(\theta) - g(\theta)(I(\theta))^{\frac{1}{2}}$ *and* $\mathrm{K}_{22}(\theta) = K_{22}(\theta) - 2g'(\theta)$ *respectively.*

Define now $c(\theta) = \{\bar{K}_{11}(\theta) - K_{11}(\theta)\}/(I(\theta))^{\frac{1}{2}}$ and $\hat{\theta}' = \hat{\theta} + c(\hat{\theta})/n$. It then follows from Lemma 5 that $K'_{11}(\theta) = \bar{K}_{11}(\theta)$ where we have written $K'_{ij}$'s for $K_{ij}$'s of Lemma 5. This implies, in view of (2.1)

$$\begin{aligned} &P_{\theta_0}\left\{-x_1 < n^{\frac{1}{2}}(\hat{\theta}' - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\right\} \\ &\qquad - P_{\theta_0}\left\{-x_1 < n^{\frac{1}{2}}(T - \theta_0)(I(\theta_0))^{\frac{1}{2}} < x_2\right\} \\ &\qquad = \left(K_{21}(\theta_0) - \bar{K}_{21}(\theta_0)\right)/2n^{\frac{1}{2}}.\textstyle\int_{-x_1}^{x_2} H_2(z)\phi(z)\,dz + o(n^{-\frac{1}{2}}). \end{aligned} \tag{3.10}$$

We now consider two cases in the light of (3.9).

CASE (i). $K_{21}(\theta_0) < \bar{K}_{21}(\theta_0)$. This, of course, proves both parts of the theorem since $\int_{-x_1}^{x_2} H_2(z)\phi(z)\,dz < 0$ for all $x_1, x_2 > 0$ (at least one being positive).

CASE (ii). $K_{21}(\theta_0) = \bar{K}_{21}(\theta_0)$. In this case the Edgeworth expansions of $n^{\frac{1}{2}}(\hat{\theta}' - \theta_0)(I(\theta_0))^{\frac{1}{2}}$ and $n^{\frac{1}{2}}(T - \theta_0)(I(\theta_0))^{\frac{1}{2}}$, under $\theta_0$, agree up to $o(n^{-\frac{1}{2}})$. This means that the result in Lemma 3 is not adequate and one needs to strengthen it. For this purpose, considering the case $r_1 = r_2 = r$ (say) $> 0$, we prove in section 4 that, for large $n$

$$|d^* - \bar{d}| = o(n^{-1}) \text{ for } \mathbf{x} \in \tilde{C}_n \tag{3.11}$$

with $P_{H_i}(\tilde{C}_n) = 1 - o(n^{-1})$, $i = 1, 2$ for a suitably chosen $\tilde{C}_n$, where $\bar{d} = \hat{\theta}' + \psi_o(\hat{\theta}', r)/n$ with

$$\psi_0(\theta, r) = \frac{\mu_{3.1}(\theta)}{-I(\theta)} \cdot \frac{1}{12r} - c(\theta).$$

(3.11) implies the following result whose proof is similar to that of Lemma 3 and will be omitted.

LEMMA 4.

$$|R(d^*) - R(\bar{d})| = o(n^{-1}).$$

Lemma 4 implies that (when $r_1 = r_2 = r > 0$) the test in $\mathcal{C}$ based on $\bar{d}$ is Bayes up to $o(n^{-1})$. Comparing now the Bayes risks of the two tests in $\mathcal{C}$ corresponding to the estimates $\bar{d}$ and $\bar{T} = T + \psi_0(T, r)/n$, we get the following which is similar to (3.5). Of course, here $C_1 = C_2 = \frac{1}{2}$.

$$(3.12)\quad P_{\theta_0 - \frac{r}{n^{1/2}}}\left(\bar{d} < \theta_0\right) + P_{\theta_0 + \frac{r}{n^{1/2}}}\left(\bar{d} > \theta_0\right)$$

$$> P_{\theta_0 - \frac{r}{B(n)^{1/2}}}\left(\bar{T} < \theta_0\right) + P_{\theta_0 + \frac{r}{n^{1/2}}}\left(\bar{T} > \theta_0\right) + o(n^{-1})$$

for all $r > 0$. However, repeated use of Lemma 5 along with the conditions $K_{31}(\theta_0) = \bar{K}_{31}(\theta_0)$, $K_{21}(\theta_0) = \bar{K}_{21}(\theta_0)$ and $K'_{11}(\theta_0) = \bar{K}_{11}(\theta_0)$ shows that (3.12) implies

$$(3.13)\quad \left(K'_{22}(\theta_0) - \bar{K}_{22}(\theta_0)\right)\left(-\int_{r(I(\theta_0))^{1/2}}^{\infty} H_2(z)\phi(z)\,dz\right)$$

$$+ \left(K_{41}(\theta_0) - \bar{K}_{41}(\theta_0)\right)\left(-\int_{r(I(\theta_0))^{1/2}}^{\infty} H_4(z)\phi(z)\,dz\right)/12 > o(n^{-1})$$

for all $r > 0$, proving part (a) of the Theorem, by (2.1).

To prove part (b) of the Theorem, note that the l.h.s. of (3.10), under the assumed condition $K_{12}(\theta_0) = K_{32}(\theta_0) = \bar{K}_{12}(\theta_0) = \bar{K}_{32}(\theta_0) = 0$, is equal to

$$\frac{1}{n}\int_{-x_1}^{x_2}\Big[K'_{22}(\theta_0)H_2(z)/2 - \bar{K}_{22}(\theta_0)H_2(z)/2$$

$$+ \left\{K_{41}(\theta_0) - \bar{K}_{41}(\theta_0)\right\}H_4(z)/24\Big]\phi(z)\,dz + o(n^{-1})$$

$$(3.14)\quad = \frac{1}{n}\Big[\left\{K'_{22}(\theta_0) - \bar{K}_{22}(\theta_0)\right\}/2.\left\{-\int_{x_2}^{\infty}H_2(z)\phi(z)\,dz - \int_{x_1}^{\infty}H_2(z)\phi(z)\,dz\right\}$$

$$+ \left\{K_{41}(\theta_0) - \bar{K}_{41}(\theta_0)\right\}/24.\left\{-\int_{x_2}^{\infty}K_4(z)\phi(z)\,dz\right.$$

$$\left.-\int_{x_1}^{\infty}H_4(z)\phi(z)\,dz\right\}\Big] + o(n^{-1})$$

$$> o(n^{-1}),$$

by (3.13).

To complete the proof we have to check Proposition 1, (3.2), (3.11) and the lemmas. This will be done in the next section.

COROLLARY. *The* mle *is second order efficient with respect to any bounded (bowl-shaped) loss function* $L_n(\theta, a) = h(n^{\frac{1}{2}}(a - \theta))$ *whose minimum value is zero at* $\theta - a = 0$ *and which increases with* $|a - \theta|$.

PROOF. The proof follows from the fact that

$$E_\theta\{L_n(\theta, T)\} = \int_0^\infty\left[1 - P_\theta\left\{n^{\frac{1}{2}}(T - \theta) < y\right\}\right]dh(y)$$

$$-\int_{-\infty}^{0}P_\theta\left\{n^{\frac{1}{2}}(T - \theta) < y\right\}dh(y)$$

and that a similar expression holds for $E_\theta\{L_n(\theta, \hat{\theta}')\}$.

**4. Proofs of lemmas and other results.** In this section we provide proofs of Proposition 1, (3.2), (3.11) and Lemmas 3 and 5. Towards this end, define for some $\alpha > 0$, $c > 0$, $\varepsilon$, $c_1$, $c_2$, $M_1$, $M_2$ to be chosen later,

$$A_n = \left\{ d : \hat{\theta} - \frac{\alpha}{n^{\frac{1}{2}}} < d < \hat{\theta} + \frac{\alpha}{n^{\frac{1}{2}}} \right\},$$

$$B_n = \left\{ d : \hat{\theta} - \frac{c \log n}{n^{\frac{1}{2}}} < d < \hat{\theta} + \frac{c \log n}{n^{\frac{1}{2}}} \right\},$$

$$C_{1n} = \left\{ \mathbf{X} : |\hat{\theta} - \theta_0| < \frac{c \log n}{n^{\frac{1}{2}}} \right\}, \quad C_{2n} = \left\{ \mathbf{X} : \sup_{|\theta - \theta_0| < \varepsilon} \left| \frac{L_3(\theta)}{n} \right| < M_1 \right\},$$

$$C_{3n} = \left\{ \mathbf{X} : \sup_{|\theta - \theta_0| < \varepsilon} \left| \frac{L_4(\theta)}{n} \right| < M_2 \right\}, \quad C_{4n} = \left\{ \mathbf{X} : \left| \frac{L_2(\hat{\theta})}{n} + I(\theta_0) \right| < \frac{c_1 \log n}{n^{\frac{1}{2}}} \right\}$$

$$C_{5n} = \left\{ \mathbf{X} : \left| \frac{L_3(\hat{\theta})}{n} - \mu_{3,1}(\theta_0) \right| < \frac{c_2 \log n}{n^{\frac{1}{2}}} \right\}$$

and

$$\tilde{C}_n = C_{1n} \cap C_{2n} \cap C_{3n} \cap C_{4n} \cap C_{5n}. \tag{4.1}$$

We shall first prove Proposition 1, (3.2), and (3.11). For that we need

LEMMA 1. *For large $n$, $P_{H_i}(\mathbf{X} \in \tilde{C}_n) = 1 - o(n^{-1})$, $i = 1, 2$.*

LEMMA 2. *If $\mathbf{X} \in \tilde{C}_n$, $\psi(d) \uparrow$ in $B_n$ and $\psi(d) = 0$ has a unique solution $d^*$ in $A_n$.*

(The proofs of these lemmas will be given later.)

Hence, if $\mathbf{X} \in \tilde{C}_n$, then $\theta_0 \in B_n$ and so by Lemmas 1 and 2 the proof of Proposition 1 follows.

To prove (3.2), note from (3.1) that expanding $L(.)$ around $\hat{\theta}$, we get

$$\begin{aligned}
\psi(d) &= \log \frac{C_1}{C_2} - \frac{r_2^2 - r_1^2}{2} \cdot \frac{L_2(\hat{\theta})}{n} - \frac{r_1 + r_2}{n^{\frac{1}{2}}}(d - \hat{\theta}) L_2(\hat{\theta}) \\
&\quad + \frac{\left(d - \frac{r_1}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3}{3!} L_3(\bar{\theta}) - \frac{\left(d + \frac{r_2}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3}{3!} L_3(\bar{\bar{\theta}}) \\
&= -\frac{r_2^2 - r_1^2}{2} \left\{ \frac{L_2(\hat{\theta})}{n} + I(\theta_0) \right\} - \frac{r_1 + r_2}{n^{\frac{1}{2}}}(d - \hat{\theta}) L_2(\hat{\theta}) \\
&\quad + \frac{\left(d - \frac{r_1}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3}{3!} L_3(\bar{\theta}) - \frac{\left(d + \frac{r_2}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3}{3!} L_3(\bar{\bar{\theta}})
\end{aligned} \tag{4.2}$$

where $\bar{\theta}$ lies between $d - r_1/n^{\frac{1}{2}}$ and $\hat{\theta}$ and $\bar{\bar{\theta}}$ between $d + r_2/n^{\frac{1}{2}}$ and $\hat{\theta}$. For

$\mathbf{X} \in \tilde{C}_n$, $d^*$ therefore satisfies

$$d^* - \hat{\theta} = \left[\left(\frac{r_1^2 - r_2^2}{2}\right)\left\{\frac{L_2(\hat{\theta})}{n} + I(\theta_0)\right\} + \left(d^* - \frac{r_1}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3 L_3(\bar{\theta})/3!\right.$$
$$\left. - \left(d^* + \frac{r_2}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3 L_3(\bar{\bar{\theta}})/3!\right] \Big/ n^{\frac{1}{2}}(r_1 + r_2)\left\{\frac{L_2(\hat{\theta})}{n}\right\}$$
$$= 0((\log n)^3/n),$$

proving (3.2).

To prove (3.11), note that under the condition $r_1 = r_2 = r$ (so that $C_1 = C_2$) (4.2) implies

(4.3)
$$\frac{2r}{n^{\frac{1}{2}}}(d^* - \hat{\theta})L_2(\hat{\theta}) + \left(d^* + \frac{r}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3 \frac{L_3(\bar{\bar{\theta}})}{3!} - \left(d^* - \frac{r}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3 \frac{L_3(\bar{\theta})}{3!} = 0$$

implying, for $\mathbf{X} \in \tilde{C}_n$

$$d^* - \hat{\theta} = \frac{n^{\frac{1}{2}}}{12rL_2(\hat{\theta})}\left[\left(d^* - \frac{r}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3 L_3(\bar{\theta}) - \left(d^* + \frac{r}{n^{\frac{1}{2}}} - \hat{\theta}\right)^3 L_3(\bar{\bar{\theta}})\right]$$
$$= \frac{1}{n} \cdot \frac{L_3(\hat{\theta})}{L_2(\hat{\theta})} \cdot \frac{1}{12r} + o(n^{-2})$$
$$= \frac{\psi(\hat{\theta}, r)}{n} + 0\left(\frac{\log n}{n^{\frac{3}{2}}}\right)$$

where

$$\psi(\theta, r) = \frac{\mu_{3,1}(\theta)}{-I(\theta)} \cdot \frac{1}{12r}.$$

The second line follows since $L_3(\bar{\theta}) = L_3(\hat{\theta}) + o(1)$, $L_3(\bar{\bar{\theta}}) = L_3(\hat{\theta}) + o(1)$ and $|d^* - \hat{\theta}| < \alpha/n^{\frac{1}{2}}$ by Lemma 2. This proves (3.11).

It remains now to complete the proofs of the various lemmas. Before proving the lemmas, we first establish the following auxiliary result which is similar to the uniform strong law of large numbers.

LEMMA 0. *Let $C$ be a compact interval and let $u(x, t)$ be a real-valued function measurable in $x$ for each $t \in C$ and continuous in $t$ for each $x$. Let $H(x)$ and $A(x)$ satisfy*

(4.4) $$|u(x, t)| < H(x) \qquad \text{for all } t \in C,$$
$$\sup_{\theta \in \Theta_0} \int H^4(x)\, dF_\theta(x) < \infty$$

*and*

(4.5) $$|u(x, t)| - u(x, t')| < |t - t'|A(x),$$
$$\sup_{\theta \in \Theta_0} \int A(x)\, dF_\theta(x) < \infty.$$

*Then if the $X_i$'s are* i.i.d. $\{F_\theta, \theta \in \Theta_0\}$,

$$P_\theta\left[\sup_{t\in C}\left|n^{-1}\Sigma_1^n u(X_i, t) - \int u(x, t)\, dF_\theta(x)\right| < \varepsilon\right] > 1 - \frac{K^*(\varepsilon)}{n^2}, \forall\theta \in \Theta_0,$$

*where $\varepsilon$ is any positive quantity and $0 < K^*(\varepsilon) < \infty$ is a constant depending only on $\varepsilon$ and independent of $\theta$.*

PROOF. Consider a finite division of $C$ in the form $C = \cup_{j=1}^{K(\varepsilon)} C_j$ where $K(\varepsilon)$ will be specified later. Let $t_j$ be the midpoint of $C_j$, $j = 1, \cdots, K(\varepsilon)$ and put $\mu_\theta(t) = \int u(x, t)\, dF_\theta(x)$. Then

$$P_\theta\left\{\sup_{t\in C}\left|n^{-1}\Sigma_1^n u(X_i, t) - \int u(x, t)\, dF_\theta(x)\right| > \varepsilon\right\}$$
$$< \Sigma_{j=1}^{K(\varepsilon)} P_\theta\left\{\sup_{t\in C_j}\left|n^{-1}\Sigma_1^n u(X_i, t) - \mu_\theta(t) - n^{-1}\Sigma_1^n u(X_i, t_j) + \mu_\theta(t_j)\right| > \frac{\varepsilon}{2}\right\}$$
$$+ \Sigma_{j=1}^{K(\varepsilon)} P_\theta\left\{\left|n^{-1}\Sigma_1^n u(X_i, t_j) - \mu_\theta(t_j)\right| > \frac{\varepsilon}{2}\right\}.$$

The first term is

$$\Sigma_{j=1}^{K(\varepsilon)} P_\theta\left\{n^{-1}\Sigma_1^n \sup_{t\in C_j}\left|u(X_i, t) - u(X_i, t_j) - \mu_\theta(t) + \mu_\theta(t_j)\right| > \frac{\varepsilon}{2}\right\}.$$

Let

$$W_j(X_i) = \sup_{t\in C_j}\left|u(X_i, t) - u(X_i, t_j) - \mu_\theta(t) + \mu_\theta(t_j)\right|$$

and

$$v_\theta(t_j) = E_\theta\{W_j(X_i)\} \geq 0, j = 1, \cdots, K(\varepsilon).$$

Note that by the assumption (4.5) of the lemma

$$W_j(X_i) \leq \sup_{t\in C_j}|t - t_j|\{A(X_i) + E_\theta A(X_i)\}$$
$$\Rightarrow v_\theta(t_j) \leq \sup_{t\in C_j}|t - t_j| 2E_\theta A(X_i)$$
$$\leq 2\left[\sup_{t\in C_j}|t - t_j|\right]\left[\sup_{\theta\in\Theta_0}\{E_\theta A(X_i)\}\right]$$

$\Rightarrow$ given any $\varepsilon > 0$, $\varepsilon/2 > \delta > 0$, we can choose $C_j$'s suitably so that $0 \leq v_\theta(t_j) < \varepsilon/2 - \delta$, $\forall\theta \in \Theta_0$, $\forall j$. This shows that $K(\varepsilon)$ is independent of $\theta$.

Hence using Chebyshev's inequality with the fourth moment and bounding the moments about mean of $W_j(X_i)$ by condition (4.4)

$$P_\theta\left\{n^{-1}\Sigma_1^n W_j(X_i) > \frac{\varepsilon}{2}\right\} \leq P_\theta\left\{\left|n^{-1}\Sigma_1^n W_j(X_i) - v_\theta(t_j)\right| > \delta\right\}$$
$$\leq \frac{nE_\theta\{W_j(X_i) - v_\theta(t_j)\}^4 + {}^nC_2\left(E_\theta\{W_j(X_i) - v_\theta(t_j)\}^2\right)^2}{n^4\delta^4}$$
$$\leq \frac{C_1(\delta, \varepsilon)}{n^2},$$

for some $0 < C_1(\delta, \varepsilon) < \infty$, which is independent of $\theta$ and $j = 1, \cdots, K(\varepsilon)$.

yielding

$$\text{First term} < \frac{K(\varepsilon)C_1(\delta)}{n^2}.$$

Similarly by Chebyshev's inequality with fourth moment and condition (4.4),

$$\text{Second term} = \Sigma_1^{K(\varepsilon)} P_\theta\left\{|n^{-1}\Sigma_1^n u(X_i, t_j) - \mu_\theta(t_j)| > \frac{\varepsilon}{2}\right\}$$

$$< \Sigma_1^{K(\varepsilon)} \frac{E_\theta\{n^{-1}\Sigma_1^n u(X_i, t_i) - \mu_\theta(t_j)\}^4}{(\varepsilon/2)^4} < \frac{K(\varepsilon)C_2(\varepsilon)}{n^2}$$

for some $0 < C_2(\varepsilon) < \infty$. □

PROOF OF LEMMA 1. Let $\theta = \theta_0 + d/n^{\frac{1}{2}}$, $d = -r_1$ or $r_2$. Then

$$P_\theta\{C_{1n}\} = P_\theta\{-c(\log n)(I(\theta_0))^{\frac{1}{2}} + 0(1) < n^{\frac{1}{2}}(\hat{\theta} - \theta)(I(\theta))^{\frac{1}{2}}$$

$$< c(\log n)(I(\theta_0))^{\frac{1}{2}} + 0(1)\} \tag{4.6}$$

$$> 1 - o(n^{-1}),$$

for a suitable $c$, by assumption 8.

Let $\Theta_0$ be a compact neighbourhood of $\theta_0$. Application of Lemma 0 shows that

$$P_\theta\left[\sup_{|t-\theta_0|<\varepsilon}\left|\frac{L_3(t)}{n} - \mu_{3,1}(t)\right| < \varepsilon'\right] > 1 - \frac{K_1(\varepsilon')}{n^2} \quad \text{for any } \varepsilon' > 0$$

for all $\theta \in \Theta_0$, by assumption 6, implying

$$P_\theta\{C_{2n}\} > 1 - \frac{K_1}{n^2} \tag{4.7}$$

for a suitable $M_1 < \infty$. Similarly one gets

$$P_\theta\{C_{3n}\} > 1 - \frac{K_2}{n^2} \tag{4.8}$$

for some $M_2 < \infty$ and all $\theta \in \Theta_0$.

Finally, by assumption 7,

$$P_\theta\left[\left((\mu_{2,2}(\theta))^{\frac{1}{2}}\right)^{-1}\left|\frac{L_2(\theta)}{n} + I(\theta)\right| < \frac{c_1^*\log n}{n^{\frac{1}{2}}}\right] > 1 - o(n^{-1})$$

uniformly for $\theta \in \Theta_0$ for a suitable $c_1^*$. Define

$$\tilde{D}_n = \left\{\mathbf{X} : \left|\frac{L_2(\theta)}{n} + I(\theta)\right| < \frac{c_1^*\log n}{n^{\frac{1}{2}}}\right\}$$

so that $P_\theta(\tilde{D}_n) > 1 - o(n^{-1})$ uniformly for $\theta \in \Theta_0$. If $\theta = \theta_0 + d/n^{\frac{1}{2}}$, one gets by Taylor's expansion

$$P_\theta\{C_{4n}\} > P_\theta\{C_{1n} \cap C_{2n} \cap \tilde{D}_n\} > 1 - o(n^{-1}) \tag{4.9}$$

for a suitable $c_1$. Similarly one gets for $\theta = \theta_0 + d/n^{\frac{1}{2}}$,

$$P_\theta\{C_{5n}\} > 1 - o(n^{-1}) \tag{4.10}$$

for a suitable $c_2$. Combining the inequalities for all the $P(C_{in})$'s one gets the lemma. □

PROOF OF LEMMA 2. If $\tilde{C}_n$ occurs and if $n$ is big enough and if $\varepsilon$, $M_1$ and $c_1$ were suitably chosen, $L(\theta)$ is strictly concave in $|\theta - \theta_0| < \varepsilon$. Hence for any $0 < \delta_i < \varepsilon/2$, $i = 1, 2$, $\delta_1 \neq \delta_2$, $L(\theta - \delta_1) - L(\theta + \delta_2)$ is increasing in $|\theta - \theta_0| < \varepsilon/2$. And if $n$ is large enough $B_n \subseteq [\theta_0 - \varepsilon/2, \theta_0 + \varepsilon/2]$ so $\psi$ increases in $B_n$.

Now consider $\psi(\hat{\theta} \pm \alpha/n^{\frac{1}{2}})$. Certainly

$$\psi(d) = \log\frac{C_1}{C_2} - L_1\left(d + \frac{\beta}{n^{\frac{1}{2}}}\right)\left(\frac{r_1 + r_2}{n^{\frac{1}{2}}}\right)$$

$$= (r_1^2 - r_2^2)\frac{I(\theta_0)}{2} - L_1\left(d + \frac{\beta}{n^{\frac{1}{2}}}\right)\left(\frac{r_1 + r_2}{n^{\frac{1}{2}}}\right)$$

where $|\beta| < \max(r_1, r_2)$. But

$$\frac{1}{n^{\frac{1}{2}}} L_1\left(\hat{\theta} + \frac{\tau}{n^{\frac{1}{2}}}\right) = \frac{L_2(\hat{\theta} + \tau\gamma)\tau}{n}$$

for $|\gamma| < 1$, which, for all $n$ large enough, is approximately $-I(\theta_0)\tau$, and thus $\psi(\hat{\theta} + \tau/n^{\frac{1}{2}}) \approx I(\theta_0)/2(r_1 - r_2 + 2\tau)$. So if $\alpha$ is sufficiently large, $\psi(\hat{\theta} - \alpha/n^{\frac{1}{2}}) < 0$ and $\psi(\hat{\theta} + \alpha/n^{\frac{1}{2}}) > 0$. □

PROOF OF LEMMA 3. Define $D_{1n} = \{\mathbf{X} : d^* < \theta_0 < \hat{\theta}\}$, $D_{2n} = \{\mathbf{X} : d^* > \theta_0 > \hat{\theta}\}$, $D_n = D_{1n} \cup D_{2n}$ and note that if $\tilde{C}_n$ and $D_n$ both occur,

$$|\psi(\theta_0)| = 0\left(\frac{(\log n)^3}{n^{\frac{1}{2}}}\right)$$

and then

$$|e^{\psi(\theta_0)} - 1|/|e^{\psi(\theta_0)} + 1| = 0\big((\log n)^3/n^{\frac{1}{2}}\big).$$

But

$$P_{H_i}(\tilde{C}_n \cap D_n) = 0\left(\frac{\log n}{n^{\frac{1}{2}}}\right) \quad \text{by assumption 8, } i = 1, 2.$$

So

$$|R(d^*) - R(\hat{\theta})| = 0\left(\frac{(\log n)^4}{n}\right) = o(n^{-\frac{1}{2}}).$$ □

PROOF OF LEMMA 5.

$$P_\theta\left\{n^{\frac{1}{2}}\left(\underline{\hat\theta} - \theta\right)(I(\theta))^{\frac{1}{2}} < x\right\}$$

$$= P_\theta\left\{n^{\frac{1}{2}}\left(\hat\theta - \theta - \frac{g(\theta)}{n} - \frac{(\hat\theta-\theta)g'(\theta)}{n} - \frac{(\hat\theta-\theta)^2}{2!}\frac{g''(\bar\theta)}{n}\right)(I(\theta))^{\frac{1}{2}} < x\right\}$$

$$\left(\text{where } \bar\theta \in (\hat\theta, \theta)\right)$$

$$= P_\theta\left\{n^{\frac{1}{2}}(\hat\theta-\theta)\left(1 - \frac{g'(\theta)}{n} - \frac{(\hat\theta-\theta)g''(\bar\theta)}{n}\right)(I(\theta))^{\frac{1}{2}} < x + \frac{g(\theta)(I(\theta))^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right\}$$

$$= P_\theta\left\{n^{\frac{1}{2}}(\hat\theta-\theta)(I(\theta))^{\frac{1}{2}} < \left(x + \frac{g(\theta)(I(\theta))^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right)\left(1 - \frac{g'(\theta)}{n} - \frac{(\hat\theta-\theta)g''(\bar\theta)}{n}\right)^{-1}\right\}$$

$$= P_\theta\left\{n^{\frac{1}{2}}(\hat\theta-\theta)(I(\theta))^{\frac{1}{2}} < \left(x + \frac{g(\theta)(I(\theta))^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right)\left(1 - \frac{g'(\theta)}{n} - \frac{(\hat\theta-\theta)g''(\bar\theta)}{n}\right)^{-1},\right.$$

$$\left.|\hat\theta-\theta| < \frac{c\log n}{n^{\frac{1}{2}}}\right\} + o(n^{-1})$$

$$= P_\theta\left\{n^{\frac{1}{2}}(\hat\theta-\theta)(I(\theta))^{\frac{1}{2}} < x + \frac{xg'(\theta)}{n} + \frac{g(\theta)(I(\theta))^{\frac{1}{2}}}{n^{\frac{1}{2}}} + o(n^{-1})\right\} + o(n^{-1})$$

$$= P_\theta\left\{n^{\frac{1}{2}}(\hat\theta-\theta)(I(\theta))^{\frac{1}{2}} < x + \frac{g(\theta)(I(\theta))^{\frac{1}{2}}}{n^{\frac{1}{2}}} + \frac{xg'(\theta)}{n}\right\} + o(n^{-1})$$

$$= \Phi\left(x + \frac{g(\theta)(I(\theta))^{\frac{1}{2}}}{n^{\frac{1}{2}}} + \frac{xg'(\theta)}{n}\right) + \frac{\Phi_{1,\,x+g(\theta)I(\theta)/n^{\frac{1}{2}}+xg'(\theta)/n}(\theta)}{n^{\frac{1}{2}}}$$

$$+ \frac{\Phi_{2,\,x+g(\theta)(I(\theta))^{\frac{1}{2}}/n^{\frac{1}{2}}+xg'(\theta)/n}(\theta)}{n} + o(n^{-1})$$

which after straightforward simplification reduces to

$$\Phi(x) + \frac{1}{n^{\frac{1}{2}}}\int_{-\infty}^{x}\left\{\left(K_{11}(\theta) - g(\theta)(I(\theta))^{\frac{1}{2}}\right)H_1(z) + \frac{K_{21}(\theta)}{2}H_2(z) + \frac{K_{31}(\theta)}{6}H_3(z)\right\}\phi(z)dz$$

$$+\frac{1}{n}\left[\int_{-\infty}^{x}\left[\left\{K_{12}(\theta)H_1(z) + \frac{K_{22}(\theta) - 2g'(\theta)}{2}H_2(z) + \frac{K_{32}(\theta)}{6}H_3(z) + \frac{K_{41}(\theta)}{24}H_4(z)\right\}\right.\right.$$

$$+\tfrac{1}{2}\left\{\left(K_{11}(\theta) - g(\theta)(I(\theta))^{\frac{1}{2}}\right)^2 H_2(z) + \frac{K_{21}^2(\theta)}{4}H_4(z) + \frac{K_{31}^2(\theta)}{36}H_6(z)\right.$$

$$+ \left(K_{11}(\theta) - g(\theta)(I(\theta))^{\frac{1}{2}}\right)K_{21}(\theta)H_3(z) + \frac{\left(K_{11}(\theta) - g(\theta)(I(\theta))^{\frac{1}{2}}\right)K_{31}(\theta)}{3}H_4(z)$$

$$\left.\left.\left.+ \frac{K_{21}(\theta)K_{31}(\theta)}{6}H_5(z)\right\}\right]\phi(z)\,dz\right] + o(n^{-1}),$$

thereby proving the lemma. □

**Acknowledgments.** Our sincere thanks are due to the referee for valuable comments leading to a substantial improvement in the presentation of the paper, including a correction to one of our crucial remarks in the earlier version.

## REFERENCES

[1] AKAHIRA, M. (1975). A note on the second order asymptotic efficiency of estimators in an autoregressive process. *Rep. Univ. Electro-Comm.* 26 143–149.

[2] BAHADUR, R. R. (1964). On Fisher's bound for asymptotic variances. *Ann. Math. Statist.* 35 1545–1552.

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

[4] CHIBISOV, D. M. (1973a). An asymptotic expansion for a class of estimators containing maximum likelihood estimators. *Theor. Probability Appl.*, 18 295–303.

[5] CHIBISOV, D. M. (1973b). An asymptotic expansion for the distribution of sums of a special form with an application to minimum contrast estimators. *Theor. Probability Appl.* 18 649–661.

[6] EFRON, BRADLEY (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.* 3 1189–1242.

[7] FISHER, R. A. (1925). Theory of statistical estimation. *Proc. Cambridge Phil. Soc.* 22 700–725.

[8] GHOSH, J. K. and SUBRAMANYAM, K. (1974). Second order efficiency of maximum likelihood estimators. *Sankhya* (*A*) 36 325–358.

[9] GHOSH, J. K. (1975). Discussion on Professor Efron's paper. *Ann. Statist.* 3 1224–1226.

[10] GUSEV, S. I. (1976). Asymptotic expansions associated with some statistical estimators in the smooth case II. Expansions of moments and distributions. *Theory Probability Appl.* 21 14–33.

[11] PFANZAGL, J. (1973). Asymptotically optimum estimation and test procedures. In *Proc. Prague Symp. Asymptotic Statist., Vol* I (J. Hajek, ed.) pages 201–272. Charles University, Prague.

[12] PFANZAGL, J. (1975). On asymptotically complete classes. In *Statistical Inference and Related Topics. Proc. Summer Res. Inst. Statist. Inference Stochastic Processes, Vol.* 2 (M. L. Puri, ed.), pages 1–43. Academic Press, New York.

[13] PFANZAGL, J. and WEFELMEYER, W. (1978). A Third Order Optimum Property of the Maximum Likelihood Estimator. *J. Multivariate Anal.* 8 1–29.

[14] RAO, C. R. (1961). Asymptotic efficiency and limiting information. *Proc. Fourth Berkeley Symp. Math. Statist. Probability.* 1 531–546.

[15] RAO, C. R. (1962). Efficient estimates and optimum inference procedures in large samples, with discussion. *J. Roy. Statist. Soc. Ser. B* 24, 46–63.

[16] RAO, C. R. (1963). Criteria of estimation in large samples. *Sankhya Ser. A* 25 189–206.

[17] RAO, C. R. (1975). Discussion on Professor Efron's paper. *Ann. Statist.* 3 1218–1219.

[18] WEISS, L. and WOLFOWITZ, J. (1966). Generalized maximum likelihood estimators. *Teor. Verojatnost.* 11 68–93.

J. K. GHOSH
B. K. SINHA
STATISTICS-MATHEMATICS DIVISION
INDIAN STATISTICAL INSTITUTE
203, BARRACKPORE TRUNK ROAD
CALCUTTA 700035, INDIA

H. S. WIEAND
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15213