THIRD ORDER EFFICIENCY OF THE MLE: A COUNTEREXAMPLE (U)

JUN 82  J K GHOSH, B K SINHA

UNCLASSIFIED  TR-82-13

AFOSR-TR-82-0860

NL
THIRD ORDER EFFICIENCY OF THF MLE - A COUNTEREXAMPLE

J.K. Ghosh*
Indian Statistical Institute

and

B.K. Sinha
University of Pittsburgh

June 1982
Technical Report No. 82-13

This work is sponsored by the Air Force Office of Scientific Research under Contracts F49620-79-C-0161 and F49629-82-K-0001. Reproduction in whole or in part is permitted for any purpose of the United States Government.

*Jointly issued as a Technical Report at the Indian Statistical Institute, Calcutta, India.
The authors give an example of a curved exponential where the maximum likelihood estimate is not third order efficient either in the sense of Fisher-Rao or Rao.
THIRD ORDER EFFICIENCY OF THE MLE
- A COUNTEREXAMPLE

J.K. Ghosh *
Indian Statistical Institute

and

B.K. Sinha
University of Pittsburgh

ABSTRACT

We give an example of a curved exponential where the maximum likelihood estimate is not third order efficient either in the sense of Fisher-Rao or Rao.

This work is sponsored by the Air Force Office of Scientific Research under Contracts F49620-79-C-0161 and F49629-82-K-001. Reproduction in whole or in part is permitted for any purpose of the United States Government.

*Jointly issued as a Technical Report at the Indian Statistical Institute, Calcutta, India.
1. INTRODUCTION

That the maximum likelihood estimate (m.l.e.) is second order efficient has been proved under various conditions by Ghosh and Subramanyam (1974), Efron (1975), Ghosh, Sinha and Subramanyam (1979), Ghosh, Sinha and Wieand (1980), Pfanzagl and Wefelmeyer (1976, 1979) and Takechi and Akahira (1978). In the last three references this property is referred to as third order efficiency; this use of the term should not be confused with ours, which will be formally introduced in the next section. The purpose of this note is to produce an example of what Efron (1975) calls curved exponentials where the efficiency of the mle fails in the class of Fisher consistent estimates when we take into consideration third order terms, (namely terms of $O(n^{-3})$), in the expression for the mean square of the error of estimates concerned or terms of order $O(n^{-1})$ in the expression for the loss of information, vide (3.1) and (4.1). On the whole we have conformed to the notations and conventions of Ghosh and Subramanyam (1974) and our definition of third order efficiency is in analogy with that of second order efficiency presented there, but this note can be read independently of it. All our computations are a straightforward application of the so called delta method and may be justified as in Ghosh and Subramanyam (1974) or Ghosh, Sinha and Subramanyam (1979); no further reference will be made to this aspect.

In Section 2 the counterexample is introduced and the mle expanded. Lack of third order efficiency is established in Sections 3 and 4.
2. THE COUNTEREXAMPLE

Consider a one-parameter trivariate density

\[ f_\theta(x_1, x_2, x_3) = \exp\left(-\frac{1}{2} (x_1-\theta)^2 - \frac{1}{2}(x_2-\theta^2)^2 - \frac{x_3^2}{2}\right) \]

\[ (Ax_1^2 x_2^2 + Bx_3^2 + 2C x_1 x_2 x_3) \]  

\[ (A(1+\theta^2)(1+\theta^4)+B)^{-1}(2\pi)^{-3/2} \]

\[-\infty < x_i < \infty, \ i = 1, 2, 3 \]

\[-\infty < \theta < \infty \]

\[ A > 0, \ B > 0, \ AB > C^2 \]

with \( A, B, C \) independent of \( \theta \).

Let \( \{X_{1i}, X_{2i}, X_{3i}\} \) be \( n \) i.i.d copies of \( X_1, X_2, X_3 \). Then the likelihood \( L(\theta) \) is the product of \( n \) terms exemplified by (2.1). Hence the derivatives of log likelihood are

\[ Z_n = \frac{1}{n} \frac{d}{d\theta} \log L \]

\[ W_n = \frac{1}{n} \frac{d^2}{d\theta^2} \log L \]

\[ V_n = \frac{1}{n} \frac{d^3}{d\theta^3} \log L \]

\[ U_n = \frac{1}{n} \frac{d^4}{d\theta^4} \log L \]

\[ = \frac{1}{n} \frac{d^4}{d\theta^4} \log L \]

\[ = n(\theta), \ \text{say.} \]
Clearly
\[ I(\theta) = E_\theta \left( -\frac{1}{n} \frac{d^2 \log L}{d\theta^2} \right) = -2 E_\theta (\bar{X}_n^2 - \theta^2) + 4\theta^2 + \frac{A(2 + 12\theta^2 + 30\theta^4)}{(A+B) + A(\theta^2 + \theta^4 + \theta^6)} - \frac{A^2(2\theta^2 + 4\theta^3 + 6\theta^5)^2}{(A+B) + A(\theta^2 + \theta^4 + \theta^6)^2}. \]

To compute \( I(\theta) \) we need
\[ E_\theta (\bar{X}_n^2 - \theta^2) = \frac{2A\theta^2(1 + \theta^2)}{(A+B) + A(\theta^2 + \theta^4 + \theta^6)}. \]

Let \( W_n = W_n^* + I(\theta) \) so that \( E_\theta W_n = 0 \).

We shall now get an expansion of the mle \( \hat{\theta} \). Clearly
\[ 0 = Z_n + (\hat{\theta} - \theta)(W_n - I(\theta)) + \frac{1}{2}(\hat{\theta} - \theta)^2 V_n^* + \frac{1}{3}(\hat{\theta} - \theta)^3 U_n^* + \text{smaller terms}. \]

So writing,
\[ \hat{\theta} - \theta = \frac{A_1}{\sqrt{n}} + \frac{A_2}{n} + \frac{A_3}{n^{3/2}} + \ldots. \quad (2.2) \]

and substituting in the previous equation we get
\[ Z_n - \frac{A_1}{\sqrt{n}} I(\theta) = 0 \]
\[ \frac{A_1}{\sqrt{n}} W_n - \frac{A_2}{n} I(\theta) + \frac{1}{2} \frac{A_1}{n} V_n^* = 0 \]
\[ \frac{A_2}{n} W_n - \frac{A_2}{n} I(\theta) + \frac{1}{\sqrt{n}} \frac{A_1}{n} V_n^* + \frac{1}{6} \frac{A_1}{\sqrt{n}} U_n^* = 0 \]

From these we determine \( A_1, A_2, A_3 \) and substitute in (2.2) yielding
\[ \hat{\theta} - \theta = \frac{Z_n}{I(\theta)} + \left\{ \frac{Z_n W_n}{I^2(\theta)} \right\} + \frac{Z_n V_n^*}{I^3(\theta)} + \left\{ \frac{Z_n W_n}{I^2(\theta)} + \frac{Z_n^2 V_n^*}{I^3(\theta)} \right\} \left\{ \frac{W_n}{I(\theta)} + \frac{Z_n V_n^*}{I^2(\theta)} \right\} + \frac{1}{\theta} \frac{U_n^* Z_n^3}{I^4(\theta)} + \ldots. \quad (2.3) \]
Let
\[ Q_2(\theta) = \frac{Z_n W_n}{I''(\theta)} + \frac{Z_n^2 V_n^*}{I^3(\theta)} \] (2.4)
and
\[ Q_3(\theta) = \left( \frac{Z_n W_n}{I''(\theta)} + \frac{Z_n^2 V_n^*}{I^3(\theta)} \right) \left( \frac{W_n}{I'(\theta)} + \frac{Z_n V_n^*}{I^2(\theta)} + \frac{1}{6} \frac{U_n^* Z_n}{I^4(\theta)} \right) \] (2.5)
So
\[ \hat{\theta} - \theta = \frac{Z_n}{I'(\theta)} + Q_2(\theta) + Q_3(\theta) + o_p(n^{-3/2}). \] (2.6)

3. LACK OF THIRD ORDER EFFICIENCY IN THE SENSE OF RAO

Define
\[ T' = \hat{\theta} + \delta^* \{ \bar{X}_3 - \pi_3(\hat{\theta}) \}^3 \]
where
\[ \pi_3(\theta) = E_{\theta} \{ \bar{X}_3 \} , \]
and \( \delta^* \) is a constant to be suitably chosen. Then like \( \hat{\theta} \), \( T' \) is Fisher consistent. Let
\[ E_{\theta} \{ \bar{X}_3 - \pi_3(\hat{\theta}) \}^3 = \xi(\theta)/n^2 + o(n^{-2}). \]
Then \( T' \) adjusted to have same bias as the mle (up to \( o(n^{-2}) \)) is
\[ T = \hat{\theta} + \delta^* \{ \{ \bar{X}_3 - \pi_3(\hat{\theta}) \}^3 - \xi(T')/n^2 \} . \]

We shall show
\[ E_{\theta=0}(\hat{\theta} - \theta)^2 > E_{\theta=0}(T - \theta)^2 \] (3.1)
if we neglect terms of order \( o(n^{-3}) \). Of course a similar inequality
\[
E_{\theta=0}(\hat{\theta}'-\theta)^2 > E_{\theta=0}(T'-\theta)^2
\]
holds if instead of adjusting \( T' \) we adjust \( \hat{\theta} \) to \( \hat{\theta}' \) to ensure same bias as \( T' \) up to \( o(n^{-2}) \). We shall refer to (3.1) as lack of third order efficiency in the sense of Rao.

For \( \theta = 0 \),
\[
Z_n = \bar{X}_1, \ W_n = 2\bar{X}_2, \ V_n^* = 0,
\]
\[
U_n^* = -12(2A^2+4AB+B^2)/(A+B)^2 \tag{3.2}
\]
\[
I(0) = (3A+B)/(A+B),
\]
\[
\pi_3(0) = 0, \ \pi_3'(0) = 0, \ E_{\hat{\theta}=0}(\bar{X}_3^3) = 0.
\]

To calculate \( \xi(\theta) \), we use
\[
(\bar{X}_3 - \pi_3(\hat{\theta}))^3 = \{ (\bar{X}_3 - \pi_3(\hat{\theta}))^3 - 3(\bar{X}_3 - \pi_3(\hat{\theta}))^2(\hat{\theta}-\theta)(\pi_3'(\theta)) \}
+ 3(\bar{X}_3 - \pi_3(\hat{\theta}))(\hat{\theta}-\theta)^2(\pi_3'(\theta))^2 - (\hat{\theta}-\theta)^3(\pi_3''(\theta))^3] + \text{smaller terms}
\]
which, using (3.2), gives for \( \theta = 0 \)
\[
(\bar{X}_3 - \pi_3(\hat{\theta}))^3 = \bar{X}_3^3 + o_p(n^{-2})
\]
and so \( \xi(0) = 0 \).

Observe
\[
E_{\theta=0}(T-\theta)^2 = E_{\theta=0}(\hat{\theta}-\theta)^2 + 2\delta^2 E_{\theta=0}[(\bar{X}_3 - \pi_3(\hat{\theta}))^3 - \xi(T')/n^2]^2
\]
\[
+ 2\delta^2 E_{\theta=0}(\hat{\theta}-\theta)((\bar{X}_3 - \pi_3(\hat{\theta}))^3 - \xi(T')/n^2) \tag{3.3}
\]
We first note that the second term on the RHS of (3.3) is (using (3.2))

\[ E_{\theta=0}(X_3^6) + o(n^{-3}) \]

\[ = \frac{15}{n^3} (A+3B)^3 + o(n^{-3}) \]  

(3.4)

For the third term on the RHS of (3.3), we have,

\[ E_{\theta=0}(\hat{\theta} - \theta)((\bar{X}_3 - \pi_3(\hat{\theta}))^3 - \xi(T')/n^2) \]

\[ = E_{\theta=0}(\hat{\theta} - \theta)((\bar{X}_3 - \pi_3(\hat{\theta}))^3 - \xi(T')/n^2) + o(n^{-3}) \]  

(3.5)

We observe, after some straightforward calculations,

\[ E_{\theta=0} X_1 \bar{X}_2 \bar{X}_3 = o(n^{-3}) \]

\[ E_{\theta=0} \bar{X}_1^3 \bar{X}_3 = o(n^{-3}) \]  

(3.6)

\[ E_{\theta=0} \bar{X}_1 X_2 \bar{X}_3 = \frac{A+3B}{(A+B)^2} \cdot \frac{6C}{n^3} + o(n^3) \]
Using (3.4), (3.5) and (3.6) we get from (3.3),

\[ E_{\theta=0}(T-\theta)^2 = E_{\theta=0}(\hat{\theta}-\theta)^2 + (\delta^* 215(A+3B)^3 + \frac{4\delta^* A+3B}{12 I(0)^{\frac{1}{2}}} \cdot 6C) \frac{1}{n^3} + o(n^{-3}). \]

Recall that \( I(0) = (3A+B)/(A+B) \) so that the coefficient of \( n^{-3} \) above is

\[ 15\delta^* 2(A+3B)^3 + 4\delta^* \frac{(A+3B)}{(3A+B)^2} 6C \]

which can be made < 0 if, say,

\[ A = B = 1, \quad \delta^* < 0 \text{ and } 20\delta^* + C > 0. \]

This proves (3.1).

4. LACK OF THIRD ORDER EFFICIENCY IN THE SENSE OF FISHER-RAO

The third order loss of information for \( \hat{\theta} \) may be defined as follows. Let

\[ V^\theta_n = \frac{d \log L}{d \theta} - \alpha \sqrt{n} - \beta n(\hat{\theta}-\theta) - \gamma n(\hat{\theta}-\theta)^2 - \delta n(\hat{\theta}-\theta)^3. \]

Let \( E_n(\hat{\theta},\theta) \) be the variance of above expanded up to \( o(n^{-1}) \) and then minimized with respect to the coefficients \( \beta, \gamma, \delta \). Of course \( \alpha \) does not play any role in this. The values of \( \beta, \gamma, \delta \) obtained in this way will depend on \( n \) and will be denoted as \( \beta_n \) etc.

Similarly we define

\[ V^T_n(\theta) = \frac{d \log L}{d \theta} - \alpha' \sqrt{n} - \beta' n(T-\theta) - \gamma' n(T-\theta)^2 - \delta' n(T-\theta)^3 \]

and then \( E_n(T,\theta) \) as above. We shall show
We sketch a proof.

Using (2.6) and (3.1) one can show that for the present purpose of evaluating $E_3$ up to $o(n^{-1})$, we may take

$$

\hat{\theta}_n = I(\theta) + \frac{\delta_1}{\sqrt{n}} + \frac{\delta_2}{n}

$$

$$

\gamma_n = \gamma_0 + \frac{\gamma_1}{\sqrt{n}} + \frac{\gamma_2}{n}

$$

(4.2)

(where $\gamma_0$ is in fact the coefficient $\lambda$ arising from considerations of second order efficiency, vide Ghosh and Subramanyam (1974)).

The following facts will be needed:

$$

\text{Cov}_{\theta=0}\{\bar{X}_1, \bar{X}_2, \bar{X}_3\} = o(n^{-3})

$$

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2, \bar{X}_3\} = \text{E}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2, \bar{X}_3\} = \text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2, \bar{X}_3\} = o(n^{-3})

$$

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_3\} = \text{E}_{\theta=0}\{\bar{X}_1^2, \bar{X}_3\} = \text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_3\} = o(n^{-3})

$$

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2^3\} = \text{E}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2^3\} = \text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2^3\} = o(n^{-3})

$$

(4.3)

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2, \bar{X}_3\} = o(n^{-3})

$$

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_3\} = o(n^{-3})

$$

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_2^3\} = o(n^{-3})

$$

$$

\text{Cov}_{\theta=0}\{\bar{X}_1^2, \bar{X}_3\} = 0

$$

Now at $\theta = 0$ (with $A = B = 1$ and hence $I = 2$)
\[
\hat{\theta} \sim \mathcal{N}(\theta, \frac{2\bar{x}_1 \bar{x}_2}{n} + \frac{4\bar{x}_1^2}{n} - \frac{U^* \bar{x}_3^2}{61^4}) \cdot \\
\text{Hence}
\]
\[
V_n^T(\hat{\theta} = 0) = V_n(\theta = 0) - \delta \cdot \bar{x}_3^2 \cdot \text{smaller terms} \quad (4.4)
\]

and
\[
V_n^T(\hat{\theta} = 0) = V_n^T(\theta = 0) - \delta \cdot \bar{x}_3^2 \cdot \text{smaller terms} \quad (4.5)
\]

Using (4.3), (4.4) and (4.5) we can prove
\[
\text{Var}_n(\hat{\theta} = 0) = E(\hat{\theta} = 0) \cdot \text{Var}_n(\theta = 0) \cdot \text{Var}_n^T(\theta = 0) = 0(n^{-1})
\]

The sum of the last two terms on the RHS of (4.6) can be made < 0 by taking \( \delta^* < 0 \), \( 20\delta^* + C > 0 \). Now (4.1) follows from (4.6).

5. CONCLUDING REMARKS

A close inspection of the proof will reveal that the reason for lack of third order efficiency is due to the non-zero covariance between \( (Q_2 + Q_3) \) and \( (T - \hat{\theta}) \). This covariance is likely
to be non-zero in most cases but except in specially constructed simple examples like the present one checking this would involve prohibitive calculations.

It would be interesting to show that not only the mle but no other estimate can possess third order efficiency.
REFERENCES


