When does third order efficiency imply fourth order efficiency

Shanti Venetiaan*

November 25, 2013

Abstract

In this article third and fourth order efficiency are studied in the framework of translation equivariant location estimators. We assume $X_1, \ldots, X_n$ i.i.d. $f(\cdot - \theta)$. By recognizing that equality in a special form of the Cauchy-Schwarz inequality leads to a certain dependence of the cumulants of the maximum likelihood estimator (MLE) for $\theta$, it is shown that this MLE is fourth order efficient if the underlying distribution is Gumbel. Contrary to similar results which were previously published this result is not based on symmetry.

1 Introduction

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with common distribution function $F(\cdot - \theta)$, $\theta \in \mathbb{R}$, and with density $f(\cdot - \theta)$, with respect to Lebesgue measure on $(\mathbb{R}, B)$. We assume that the derivative $f'(\cdot - \theta)$ is such that the Fisher information for location is finite, i.e.

\begin{equation}
I(f) = \int \left( \frac{f'}{f} \right) f < \infty
\end{equation}

*Institute of Graduate Studies and Research, Anton de Kom Universiteit van Suriname, Paramaribo, Suriname
holds. We estimate the location parameter $\theta$ by a translation equivariant estimator $T_n = t_n(X_1, \ldots, X_n)$, i.e. $T_n$ satisfies

(1.2) \[ t_n(x_1 + a, \ldots, x_n + a) = t_n(x_1, \ldots, x_n) + a, \quad a, x_1, \ldots, x_n \in \mathbb{R} \]

We denote the distribution function of $a_nT_n$, under $f(\cdot)$ by $G_n(\cdot)$, namely

(1.3) \[ G_n(y) = P_{\theta}(a_n(T_n - \theta) \leq y), \quad y \in \mathbb{R}, \]

and we are interested in constructing confidence intervals for $\theta$ based on $T_n$. The confidence intervals are of the form

(1.4) \[ [T_n - \frac{1}{a_n}G_n^{-1}(v), T_n - \frac{1}{a_n}G_n^{-1}(u)] \]

with $0 < u < v < 1$. It may be shown that $G_n$ has a density (c.f. Theorem 1.1. of Klaassen (1984)). From this follows that $G_n(G_n^{-1}(u)) = u$ for all $u$, $0 < u < 1$, and hence intervals like (1.4) have coverage probability $v - u$. The length of the confidence interval in (1.4) is $\frac{1}{a_n}(G_n^{-1}(v) - G_n^{-1}(u))$ and one may be interested in using the length of a confidence interval as a measure of the performance of the estimator $T_n$. Shorter intervals with the same coverage probability will be considered to be linked to better estimators than estimators related to longer intervals. The ideas mentioned above have already been considered in Klaassen and Venniaan (2010). In that article the confidence interval inequality was introduced which says that under certain regularity conditions,

(1.5) \[ G_n^{-1}(v) - G_n^{-1}(u) \geq \epsilon \]

for all $n \in \mathbb{N}$, $u$, $v$, $0 < u < v < 1$, and $c > 0$ and $\epsilon > 0$ satisfying the relations

(1.6) \[ u = P_0 \left( \prod_{i=1}^{n} f \left( X_i + \frac{\epsilon}{a_n} \right) \geq c \prod_{i=1}^{n} f(X_i) \right) \]

and

(1.7) \[ v = P_{-\epsilon/a_n} \left( \prod_{i=1}^{n} f \left( X_i + \frac{\epsilon}{a_n} \right) \geq c \prod_{i=1}^{n} f(X_i) \right) . \]

Asymptotic expansions were derived for both the left hand side and the right hand side of (1.5) and it can be shown that the inequality in (1.5) may be
used to prove that first order efficiency implies second order efficiency. This was first proved by Pfanzagl (1979) and has after that been studied by scholars like Pfanzagl and Wefelmeyer (1985), and Bickel, Chibisov and van Zwet (1981). In Klaassen and Venetiaan (1994) the spread inequality of Klaassen (1984) was used to prove this phenomenon. For the maximum likelihood estimator which is first order efficient and therefore second order efficient, Klaassen and Venetiaan (2010) also found that locally even third order efficiency is obtained by the MLE and this third order efficiency immediately implies fourth order efficiency, because of the symmetry of the intervals and the even polynomials in the fourth order term. Ghosh (1994) introduced the conjecture that third order efficiency implies fourth order efficiency and this assertion was proved by Akahira (1996) by studying the coverage probability of symmetric confidence intervals.

In this paper we will limit ourselves to the case of taking \( T_n \) to be the MLE and \( a_n = \sqrt{nI(f)} \) in (1.3). In the proof that locally third order efficiency implies fourth order efficiency by Klaassen and Venetiaan (2010) the difference \( G^{-1}_n(u) - G^{-1}_n(v) - \epsilon \) is studied for a special choice for \( u \) and \( v \). If this special choice is not made, studying the difference gives that third order efficiency and even fourth order efficiency may be obtained when a certain dependence among the cumulants of the MLE holds. This dependence of the cumulants may be viewed as equality in a certain case of the Cauchy-Schwarz inequality, and results in

\[
\frac{f'}{f}(X) = \lambda \left( \frac{f''}{f}(X) - \left( \frac{f'}{f}(X) \right)^2 + I(f) \right)
\]

for some \( \lambda \in \mathbb{R} \). Note that \( \frac{f''}{f}(X) - \left( \frac{f'}{f}(X) \right)^2 \) is the derivative of \( \frac{f'}{f}(X) \), so we have a differential equation and its solution is the Gumbel distribution with density \( f(x) = \frac{1}{\beta} \exp\left[ \frac{x-\alpha}{\beta} - \exp\left( \frac{x-\alpha}{\beta} \right) \right] \), for some \( \alpha \in \mathbb{R} \), and \( \beta > 0 \). We will also show that when the underlying distribution is Gumbel, the MLE estimator for location is fourth order efficient. In other words, regardless of the choice of \( u \) and \( v \), an MLE estimator for location is fourth order efficient if the underlying distribution is Gumbel.

In Section 2 we will repeat the theorem on the confidence interval inequality, the theorem concerning the Cornish Fisher expansion for \( G^{-1}_n(u) \) and also repeat the theorem which gives an expansion for the \( \epsilon \) in (1.5). The main result is stated
in Section 3 and its proof may be found in the same section.

## 2 Confidence interval inequality and asymptotic expansions

In this section we mention some results which were published in Klaassen and Venetiaan (2010) and Venetiaan (2010).

**Theorem 2.1** Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with common distribution function $F(\cdot - \theta)$, $\theta \in \mathbb{R}$, and with density $f(\cdot - \theta)$, with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B})$. Let $T_n = t_n(X_1, \ldots, X_n)$ be a translation equivariant estimator for $\theta$ and let $H_n$ be the distribution function of $a_n T_n$, $a_n > 0$, under $f(\cdot)$, i.e.

\[
H_n(y) = P_0(a_n T_n \leq y) = P_\theta(a_n(T_n - \theta) \leq y), \quad y \in \mathbb{R}.
\]

Fix $u$ and $v$ with $0 < u < v < 1$, and assume that there exist $\epsilon > 0$ and $c > 0$ satisfying the relations

\[
u = P_0\left(\prod_{i=1}^{n} f\left(X_i + \frac{\epsilon}{a_n}\right) \geq c \prod_{i=1}^{n} f(X_i)\right)
\]

and

\[
v = P_{-\epsilon/a_n}\left(\prod_{i=1}^{n} f\left(X_i + \frac{\epsilon}{a_n}\right) \geq c \prod_{i=1}^{n} f(X_i)\right).
\]

If $H_n^{-1}(\cdot)$ is strictly increasing at $u$ and continuous at $v$, then

\[
H_n^{-1}(v) - H_n^{-1}(u) \geq \epsilon.
\]

Furthermore we define our notion of asymptotic efficiency. Asymptotic efficiency of $T_n$ means that $H_n$ converges weakly to the standard normal distribution function $\Phi$ as $n \to \infty$,

\[
H_n^{-1}(u) = \inf\{y \in \mathbb{R} : H_n(y) \geq u\} \to \Phi^{-1}(u), \quad 0 < u < 1.
\]
By efficiency of the $j$-th order we mean that the expansions up to and including the $n^{-(j-1)/2}$ term have optimal coefficients. As mentioned earlier, we take $T_n$ to be the maximum likelihood estimator, $a_n = \sqrt{nI(f)}$ in (1.3) which choice leads to asymptotic efficiency of $T_n$. Furthermore in our framework, optimal coefficients for the $n^{-(j-1)/2}$ term of $G_n^{-1}(\cdot)$ will be obtained when $|G_n^{-1}(v) - G_n^{-1}(u) - \epsilon| = o(n^{-(j-1)/2})$.

We now introduce some notation.

$$
\eta_2 = E\psi^2(X_1)/I^2(f), \quad \eta_3 = E\psi^3(X_1)/I^{3/2}(f), \quad \eta_4 = E\psi^4(X_1)/I^2(f),
$$

(2.6)

$$
\eta_5 = E\psi^5(X_1)/I^{5/2}(f), \quad \text{and} \quad \eta_6 = E(\psi_2(X_1)\psi_3(X_1))/I^{5/2}(f)
$$

are defined with

$$
\psi_i(x) = \frac{f^{(i)}(x)}{f(x)}, \quad x \in \mathbb{R},
$$

(2.7)

The next two theorems were previously published in Venetiaan (2010) and Klaassen and Venetiaan (2010) respectively. The first one gives an asymptotic expansion for $G_n^{-1}(\cdot)$ which may be used for the left hand side of (1.5).

**Theorem 2.2** Let $X, X_1, ..., X_n$ be i.i.d. with common density $f(\cdot - \theta_0)$. Let $\rho(\cdot) = -\log f(\cdot)$ satisfy the following conditions.

1. For all $K \subset \mathbb{R}$ compact, $\sup_{\theta \in K} E_0 \rho^2(X) = A < \infty$.
2. $\rho(\cdot)$ is five times differentiable.
3. There exists a function $R(\cdot)$ and a $\delta > 0$ such that, for every $y \in \mathbb{R}$, $|\theta| < \delta$:

$$
|\rho^{(5)}(y) - \rho^{(5)}(y - \theta)| \leq R(y)|\theta| \quad \text{and} \quad E_0 R^{5/2}(X) < \infty.
$$

4. $E_0|\rho^{(i)}(X)|^5 < \infty$ for $i = 1, ..., 5$.

Then

$$
|G_n^{-1}(v) - \tilde{G}_n^{-1}(v)| = o\left(\frac{1}{n^{(j-1)/2}}\right)
$$

(2.8)
holds with
\[ \tilde{G}_n^{-1}(v) = z_v \left[ 1 + \frac{1}{12\sqrt{n}} \eta_3 z_v \right. \
+ \frac{1}{72n} \left\{ (-9 + 12\eta_2 - \eta_3^2 - 5\eta_4)z_v^2 - 9 - 2\eta_3^2 + 3\eta_4 \right\} \
+ \frac{1}{144n\sqrt{n}} \left\{ (6\eta_2\eta_3 - 3\eta_3 - \frac{19}{12}\eta_3^3 - \eta_3\eta_4 + \frac{24}{5}\eta_5 - 18\eta_6)z_v^3 \
+ (12\eta_2\eta_3 - 15\eta_3 - \frac{67}{9}\eta_3^3 + 3\eta_3\eta_4 - \frac{9}{5}\eta_5)z_v \right\} \right] \]
and \( z_v = \Phi^{-1}(v) \). In other words the inverse of the distribution function of \( T_n \) admits the Cornish-Fisher expansion \( \tilde{G}_n^{-1}(\cdot) \).

The next theorem results in an asymptotic expansion for \( \epsilon \) in (1.5).

**Theorem 2.3** Let the conditions of Theorem 2.1 and 2.2 be satisfied. Then for
\[ \tilde{\epsilon} = z_v - z_u + \frac{\eta_3}{12\sqrt{n}} (z_v^2 - z_u^2) \]
\[ + \frac{1}{288n} \left\{ (12\eta_2 + 5\eta_3^2 - 8\eta_4)(z_v^3 - z_u^3) \
+ (36 - 36\eta_2 + 9\eta_3^2 + 12\eta_4)(z_v^2 z_u - z_u^2 z_v) \right\} \]
\[ + \frac{1}{12960n^{3/2}} \left\{ (270\eta_2\eta_3 + 60\eta_3^3 - 180\eta_3\eta_4 + 162\eta_5 - 540\eta_6)(z_v^4 - z_u^4) \
+ (-540\eta_2\eta_3 + 540\eta_3^3 + 270\eta_3^3 - 270\eta_5 + 1080\eta_6)(z_v^3 z_u - z_u^3 z_v) \right\} \]
\[ + \frac{1}{12960n^{3/2}} \left\{ (270\eta_3 - 400\eta_3^3 + 630\eta_3\eta_4 - 162\eta_5)(z_v^2 - z_u^2) \right\}, \]
we have
\[ |\epsilon - \tilde{\epsilon}| = o(1/n\sqrt{n}), \]
where \( z_u = \Phi^{-1}(u) \)

**3 Main result and proof**

In this section we state the main result of this paper.
Theorem 3.1 Let the conditions of Theorem 2.1 and 2.2 hold. Then,

\[ |G_n^{-1}(v) - G_n^{-1}(u) - \epsilon| = o\left(\frac{1}{n^{\frac{1}{2}}}\right), \quad 0 < u < v < 1, \]

if and only if the underlying distribution is Gumbel or \( v = 1 - u \). In other words, in the framework of Theorem 2.1 and Theorem 2.2 the maximum likelihood estimator for location is efficient up to fourth order if and only if the underlying distribution is Gumbel or if \( v = 1 - u \).

In the proof of Theorem 3.1 we will study the difference \( \tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon} \) which equals

\[ -\frac{1}{96n} (12 - 12\eta_2 + 3\eta_3^2 + 4\eta_4)[z_v^3 - z_u^3 + z_u z_v^2 - z_u^2 z_v] + o\left(\frac{1}{n}\right) \]

It’s clear that optimal third order coefficients are obtained for \( \tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) \), i.e. third order efficiency may be obtained when \( 12 - 12\eta_2 + 3\eta_3^2 + 4\eta_4 = 0 \). When we divide both sides of this expression by 3 and take a closer look, we see that this expression is a case of equality in the Cauchy-Schwarz inequality. We may state the following lemma.

Lemma 3.2 Use the framework of the theorems above. Then

\[ \eta_3^2 = 4\eta_2 - \frac{4}{3}\eta_4 - 4 \]

if and only if the underlying distribution is Gumbel.

Proof of Lemma 3.2.

We first mention that

\[ E\psi_1 \psi_2(X_1) = \frac{1}{2} E\psi_1^3(X_1), \quad E\psi_2^2(X_1) = \frac{2}{3} E\psi_1^4(X_1), \]
\[ E\psi_1(X_1) = 0, \quad \text{and} \quad E\psi_2(X_1) = 0. \]

Note that the following form of the Cauchy-Schwarz inequality:

\[ (EYZ)^2 \leq EY^2 EZ^2 \]
with \( Y = \psi_1 \) and \( Z = (\psi'_1 - E\psi'_1) \) results in
\[
\begin{align*}
[E\psi_1(\psi'_1 - E\psi'_1)]^2 &\leq E\psi_1^2E(\psi'_1 - E\psi'_1)^2 \\
[E(\psi_1(\psi_2 - \psi_1^2 + I))]^2 &\leq IE(\psi_2 - \psi_1^2 + I)^2 \\
[E(\psi_1\psi_2 - \psi_1^3 + I\psi_1)]^2 &\leq IE(\psi_2^3 - 2\psi_1^2\psi_2 + 2I\psi_2 + \psi_1^4 - 2I\psi_1^2 + I^2) \\
\frac{[E\psi_1^2]^2}{4} &\leq IE(\psi_2^2 - \frac{\psi_1^4}{3} - I^2)
\end{align*}
\] (3.17)
\[
\frac{\eta_2^2}{4} \leq \eta_2 - \frac{\eta_4}{3} - 1,
\]
where \( \psi_1, \psi'_1 \) and \( I \) are short for \( \psi_1(X_1), \psi'_1(X_1) \), and \( I(f) \) respectively. Note that (3.14) corresponds to equality in the Cauchy Schwarz inequality. It’s well known that equality in the Cauchy-Schwarz inequality (3.16) is obtained if and only if \( Y = \lambda Z \) for some \( \lambda \in \mathbb{R} \). For our case this means that equality in (3.17) is obtained if and only if \( \psi_1 = \lambda((\psi'_1 - E\psi'_1)) \) holds. Now note that this is a differential equation and it’s solution is the density of the Gumbel distribution. \( \Box \)(end of proof of Lemma 3.2).

### 3.1 Proof of Theorem 3.1

Assume that \( |G_n^{-1}(v) - G_n^{-1}(u) - \epsilon| = o(1/n\sqrt{n}) \), then we get
\[
\begin{align*}
|\tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon}| &\leq |\tilde{G}_n^{-1}(v) - G_n^{-1}(v)| + |G_n^{-1}(v) - G_n^{-1}(u)| \\
&\quad + |\epsilon - \tilde{\epsilon}| + |G_n^{-1}(v) - G_n^{-1}(u) - \epsilon| \\
&\leq o\left(\frac{1}{n\sqrt{n}}\right)
\end{align*}
\] (3.18)

by using (2.8), (2.11), and the assumption. As \( \tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon} \) only has terms up to and including a \( 1/n\sqrt{n} \)-term, this means that \( \tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon} = 0 \). In other words, we have \( \eta_3^2 = 4\eta_2 - \frac{4}{3}\eta_4 - 4 \) or we have \( z_u = -z_v \). As Lemma 3.2 states \( \eta_3^2 = 4\eta_2 - \frac{4}{3}\eta_4 - 4 \) only holds when the underlying distribution is Gumbel. The other case \( z_u = -z_v \) means that \( v = 1 - u \).

Now, on the other hand, let \( v = 1 - u \), then \( z_u = -z_v \) and it is easy to see that the \( 1/n \)-term and \( 1/n\sqrt{n} \)-term of \( \tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon} \) vanish. But then,
THIRD AND FOURTH ORDER EFFICIENCY

\[ |G_n^{-1}(v) - G_n^{-1}(u) - \epsilon| \leq |G_n^{-1}(v) - \tilde{G}_n^{-1}(v)| + |G_n^{-1}(u) - \tilde{G}_n^{-1}(u)| + |\epsilon - \tilde{\epsilon}| + |\tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon}| \]

(3.19)\[ = o\left(\frac{1}{n^{1/2}}\right) \]

because of (2.8), (2.11).

At last, we assume that the underlying distribution is Gumbel and we see that the Fisher information and the relevant cumulants are

\[ I(f) = 1, \eta_2 = 5, \eta_3 = -2, \eta_4 = 9, \eta_5 = -44, \text{ and } \eta_6 = -13. \]

(3.20)

Note that substitution of these quantities in \( \tilde{G}_n^{-1}(v) - \tilde{G}_n^{-1}(u) - \tilde{\epsilon} \) results in 0. In the same fashion as in (3.19) it may be shown that \( |G_n^{-1}(v) - G_n^{-1}(u) - \epsilon| = o(1/n^{1/2}) \).

\( \square \) (end of proof of Theorem 3.1).

Acknowledgments. The author thanks Chris Klaassen for pointing out that \( \eta_3^2/4 \leq \eta_2 - \eta_4/3 - 1 \) is a special case of the Cauchy Schwarz inequality and for his feedback with regards to this paper.

References

Akahira, M. (1996). Third order efficiency implies fourth order efficiency: a resolution of the conjecture of J.K. Ghosh. Annals of the Institute of Statistical Mathematics 48:365–380.

Bickel, P.J. Chibisov, D.M. van Zwet, W.R. (1981). On Efficiency of first and second order. International Statistical Review 49:169–175.

Ghosh J. (1994). Higher Order Asymptotics. NSF-CBMS regional conference series Probability and Statistics, Institute of Mathematical Statistics.

Klaassen, C.A.J. (1984). Location estimators and spread. The Annals of Statistics 12:311–321.
Klaassen C.A.J. Venetiaan, S.A. (1994). Spread inequality and efficiency of first and second order. In: Hušková, M. Mandl, P. ed., *Asymptotic Statistics, Proceedings of the Fifth Prague Symposium*:341–348. Heidelberg: Physica-Verlag.

Klaassen C.A.J. Venetiaan, S.A. (2010). Optimizing lengths of confidence intervals: Fourth-order efficiency implies in location models. Communications in Statistics-Theory and methods 39, 8-9:1437–1448

Pfanzagl, J. (1979). First order efficiency implies second order efficiency. In: Jurečková, J. *Contributions to Statistics: J. Hájek Memorial Volume*:157–196. Prague: Academia.

Pfanzagl, J. Wefelmeyer, W. (1985). *Asymptotic Expansions for General Statistical Models, Lecture Notes in Statistics* 31. New York: Springer Verlag.

Venetiaan, S.A. (2010). An expansion for the maximum likelihood estimator of location and its distribution function. Brazilian Journal of Probability and Statistics 23,1: 82–91.