Large deviation efficiency of the maximum likelihood estimator for the Cauchy distribution

Masafumi Akahira
Institute of Mathematics
University of Tsukuba

Abstract

The large deviation (LD) efficiency of estimators was discussed and the maximum likelihood estimator (MLE) was shown to be second order LD efficient for an exponential family of distributions. In this article, for the Cauchy distribution with a location parameter, the MLE is shown to be first order LD efficient, which implies that it is asymptotically Bahadur efficient. However, the MLE is shown to be not second order LD efficient.

Keywords Bahadur efficiency · Saddlepoint approximation · Large deviation efficiency · Maximum likelihood estimator · Cauchy distribution

1 Introduction

From the viewpoint of large deviation, the asymptotic Bahadur efficiency of estimators is well known, and the MLE is shown to be asymptotically Bahadur efficient under suitable regularity conditions (see Bahadur 1967, 1971, Fu 1973). Further, the Bahadur type second order asymptotic efficiency is considered from the Fisher-Rao-Efron approach by Fu (1982) (see Efron 1975).

On the other hand, the LD efficiency of estimators is discussed up to the second order, from a different viewpoint of the Bahadur efficiency (see Akahira 2006, 2010). Its distinctive feature is the direct evaluation of the LD probability of estimators using the saddlepoint approximation. For an exponential family of distributions, the MLE is second order LD efficient in the sense that it has the least LD probability up to the second order, i.e. the order $o(1/n)$ in a class of weakly asymptotically median unbiased estimators, where n is a size of sample.

In order to investigate the influence of the middle part of distributions to the LD probability of estimators, we discussed the LD probability of estimators in the class for flattened distributions in a middle part which do not belong to an exponential family of distributions (see Akahira 2024). In particular, for the flattened normal case in the interval $[-\varepsilon, \varepsilon]$, the sample mean is seen to be asymptotically better than the sample median for both of small and large $\varepsilon(> 0)$ in the sense of the LD probability. For the flattened Laplace case in the interval $[-\varepsilon, \varepsilon]$, the sample median is seen to be asymptotically better than the sample mean for smaller ε, but the sample median is done to be asymptotically worse than the sample mean for bigger ε in the sense of the LD probability.

For the Cauchy distribution with a location parameter θ, it is shown by Bai and Fu (1987) that the MLE $\hat{\theta}_{ML}$ of θ based on a sample of size n is asymptotically Bahadur efficient in the sense of

$$\lim_{a\to 0}\lim_{n\to\infty}\frac{1}{na^2}\log P_{\theta,n}\left\{\left|\hat{\theta}_{ML}-\theta\right|\geq a\right\} = -\frac{I(\theta)}{2},$$

where $I(\theta)$ is the Fisher information amount. Related results are found in Akaoka et al. (2022). In this article we consider the LD efficiency of $\hat{\theta}_{ML}$ using the saddlepoint approximation from the viewpoint of the LD probability. In Section 3 the lower bound for the LD probability of weakly asymptotically median unbiased estimators is obtained and the MLE is shown to be weakly LD efficient in the sense that the MLE attains the lower bound. The result brings that the MLE is asymptotically Bahadur efficient. Further, there arises an interesting problem whether the MLE is second order LD efficient or not. In order to solve it, we obtain the lower bound for the LD probability of weakly asymptotically median unbiased estimators and the LD probability of the MLE, up to the second order, i.e. the order $o(n^{-1})$. As a result the MLE is shown to be not second order LD efficient in Section 4. In the appendix of Section 6, the classical saddlepoint approximations are summarized.

2 Definitions and the lower bound for the LD probability of estimators

Assume that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of independent and identically distributed (i.i.d.) random variables with a probability density function (p.d.f.) $f(x;\theta)$ with respect to a σ-finite measure, where $x \in \mathcal{X}$, $\theta \in$

Θ and Θ is an open interval in the real line $\boldsymbol{R}$. Put $\boldsymbol{X} = (X_1, \cdots, X_n)$. The following definitions are given by Akahira (2006, 2010).

Definition 1 If an estimator $\hat{\theta}_n := \hat{\theta}_n(\boldsymbol{X})$ of θ satisfies

$$P_{\theta,n}\{\hat{\theta}_n \leq \theta\} = \frac{1}{2} + o(1), \quad P_{\theta,n}\{\hat{\theta}_n \geq \theta\} = \frac{1}{2} + o(1),$$

as $n \to \infty$, then $\hat{\theta}_n$ is said to be weakly asymptotically median unbiased (wAMU) for θ.

Let $\mathbb{A}$ be the class of all wAMU estimators of θ.

Definition 2 If there exists $\hat{\theta}_n^* := \hat{\theta}_n^*(\boldsymbol{X})$ in $\mathbb{A}$ such that for any $\hat{\theta}_n \in \mathbb{A}$, any $\theta \in \Theta$ and any $a > 0$

$$\begin{aligned} P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\} &\geq P_{\theta,n}\{|\hat{\theta}_n^* - \theta| > a\}\{1 + o(1)\} \\ &= B_n(a, \theta)\{1 + o(1)\} \qquad \text{(say)} \end{aligned}$$

as $n \to \infty$, then $\hat{\theta}_n^*$ is said to be (first order) large deviation efficient (LDE). If there exists $\tilde{\theta}_n^* = \tilde{\theta}_n^*(\boldsymbol{X})$ in $\mathbb{A}$ such that

$$\lim_{a \to 0} \lim_{n \to \infty} \frac{P_{\theta,n}\{|\tilde{\theta}_n^* - \theta| > a\}}{B_n(a, \theta)} = 1,$$

then $\tilde{\theta}_n^*$ is said to be (first order) weakly large deviation efficient (wLDE).

Definition 3 If there exists $\hat{\theta}_n^{**} := \hat{\theta}_n^{**}(\boldsymbol{X})$ in $\mathbb{A}$ such that for any $\hat{\theta}_n \in \mathbb{A}$, any $\theta \in \Theta$ and any $a (> 0)$

$$\frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{B_n(a, \theta)} \geq 1 + \frac{b(a, \theta)}{n} + o\left(\frac{1}{n}\right) \tag{2.1}$$

as $n \to \infty$ and $\hat{\theta}_n^{**}$ attains the lower bound in (2.1) up to the order $o(n^{-1})$, then $\hat{\theta}_n^{**}$ is said to be second order LDE, where $B_n(a, \theta)$ is given in Definition 2 and $b(a, \theta)$ is a certain constant. If there exists $\tilde{\theta}_n^{**} := \tilde{\theta}_n^{**}(\boldsymbol{X})$ in $\mathbb{A}$ such that

$$\lim_{a \to 0} \lim_{n \to \infty} n \left[\frac{P_{\theta,n}\{|\tilde{\theta}_n^{**} - \theta| > a\}}{B_n(a, \theta)} - 1 - \frac{b(a, \theta)}{n} \right] = 0,$$

then $\tilde{\theta}_n^{**}$ is said to be second order wLDE.

According to Akahira (2006, 2010), we derive the lower bound in a manner of testing hypothesis in the following outline. First we assume that $\{x \mid f(x; \theta) > 0\}$ is independent of θ. Let θ_0 be any fixed in Θ and $a > 0$. Then we consider a problem of testing the hypothesis H: $\theta = \theta_0 + a$ against K: $\theta = \theta_0$, where $\theta_0 + a \in \Theta$. Let $\phi^*(X)$ be the most powerful (MP) test of level $1/2 + o(1)$ as $n \to \infty$. Letting $\hat{\theta}_n$ be any wAMU estimator and

$$A_{\hat{\theta}_n} = \{ \boldsymbol{x} \mid \hat{\theta}_n (\boldsymbol{x}) \le \theta_0 + a \},$$

we see that the indicator $\chi_{A_{\hat{\theta}_n}}$ of $A_{\hat{\theta}_n}$ is a test of level $1/2 + o(1)$ as $n \to \infty$, where $\boldsymbol{x} := (x_1, \cdots, x_n)$. Since

$$E_{\theta_0}[\phi^*] \ge E_{\theta_0}\left[\chi_{A_{\hat{\theta}_n}}\right] = P_{\theta_0,n}\{\hat{\theta}_n \le \theta_0 + a\}$$

for large n, it follows that

$$P_{\theta_0,n}\{\hat{\theta}_n - \theta_0 > a\} \ge 1 - E_{\theta_0}[\phi^*]. \tag{2.2}$$

In order to obtain the lower bound for the tail probability, i.e. the right-hand side of (2.2), it is enough to get the power function of the MP test ϕ^*. Now, it is seen from the fundamental lemma of Neyman-Pearson that a test with rejection region of the type

$$\overline{Z}(\theta_0, a) := \frac{1}{n}\sum_{i=1}^{n} Z_i(\theta_0, a) > c \tag{2.3}$$

is MP, where

$$Z_i(\theta, a) := \log(f(X_i; \theta)/f(X_i; \theta + a)) \quad (i = 1, \cdots, n)$$

and c is a constant chosen so that the asymptotic level of the test is $1/2 + o(1)$ as $n \to \infty$. It is noted that $Z_1(\theta_0, a), \cdots, Z_n(\theta_0, a)$ are i.i.d. Letting $\mu(\theta_0, a) = E_{\theta_0+a}[Z_1(\theta_0, a)]$ and $\sigma^2(\theta_0, a) = V_{\theta_0,a}(Z_1(\theta_0, a))$, we put

$$W_n = \frac{\sqrt{n}}{\sigma(\theta_0, a)}\{\overline{Z}(\theta_0, a) - \mu(\theta_0, a)\},$$

where $\sigma(\theta_0, a) = \sqrt{\sigma^2(\theta_0, a)}$. Since the MP test with the rejection region (2.3) is of asymptotic level $1/2 + o(1)$, i.e.

$$P_{\theta_0+a,n}\left\{W_n \le \frac{\sqrt{n}}{\sigma(\theta_0, a)}\left(c - \mu(\theta_0, a)\right)\right\} = \frac{1}{2} + o(1)$$

as $n \to \infty$, by the central limit theorem we choose $\mu(\theta_0, a)$ as c. Since

$$E_{\theta_0}[\phi^*] = P_{\theta_0,n}\{\overline{Z}(\theta_0, a) > \mu(\theta_0, a)\},$$

it follows from (2.2) that for large n

$$P_{\theta_0,n}\{\hat{\theta}_n - \theta_0 > a\} \ge P_{\theta_0,n}\{\overline{Z}(\theta_0, a) \le \mu(\theta_0, a)\} \tag{2.4}$$

for $a > 0$. In order to obtain the asymptotic expansion of the tail probability of $\overline{Z}(\theta_0, a)$, we apply the saddlepoint approximation to the right-hand side of (2.4) (see Daniels 1954, 1987, Barndorff-Nielsen and Cox 1989, Jensen 1995). In a similar way to the above, considering a problem of testing the hypothesis H: $\theta = \theta_0 - a$ against K: $\theta = \theta_0$, we can derive the lower bound

$$P_{\theta_0,n}\{\overline{Z}\,(\theta_0,-a) \le \mu(\theta_0,-a)\}$$

for the tail probability $P_{\theta_0,n}\{\hat{\theta}_n - \theta_0 < -a\}$ (see Akahira 2024). Since θ_0 is arbitrary, from the above and (6.2) in the appendix of Section 6 later we have the following.

Theorem 1 (Akahira 2024). For any wAMU estimator $\hat{\theta}_n$ of $\theta(\in \Theta)$ and any $a(> 0)$ with $\theta \pm a(\in \Theta)$, it holds that for large n

$$\frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{\frac{1}{\sqrt{2\pi n}}\left(\frac{1}{\sigma}e^{n\mu} + \frac{1}{\tilde{\sigma}}e^{n\tilde{\mu}}\right)} \ge 1 + O\left(\frac{1}{n}\right), \tag{2.5}$$

where $\mu = \mu(\theta, a)$, $\tilde{\mu} = \mu(\theta, -a)$, $\sigma = \sqrt{\sigma^2(\theta, a)}$ *and* $\tilde{\sigma} = \sqrt{\sigma^2(\theta, -a)}$.

In the next section, for the Cauchy distribution we obtain the lower bound for the LD probability of wAMU estimators and show that MLE is LDE.

3 The LD efficiency of the MLE for the Cauchy distribution

Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables according to the Cauchy distribution with p.d.f.

$$f_0(x - \theta) = \frac{1}{\pi(1 + (x - \theta)^2)} \quad (x \in \mathcal{X} = \boldsymbol{R}, \theta \in \Theta = \boldsymbol{R}). \tag{3.1}$$

For each $i = 1, 2, \cdots,$ we have

$$Z_i(\theta, a) = \log\frac{f_0(X_i - \theta)}{f_0(X_i - \theta - a)} = \log\left(1 + \frac{a^2 - 2a(X_i - \theta)}{1 + (X_i - \theta)^2}\right),$$

where $-\infty < a < \infty$. For $|a| < 1/\sqrt{2}$ we have for each $i = 1, 2, \cdots$

$$Z_i(\theta, a) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\left\{\frac{a^2 - 2a(X_i - \theta)}{1 + (X_i - \theta)^2}\right\}^k. \tag{3.2}$$

Then we obtain the lower bound for the LD probability for wAMU estimators as follows.

Theorem 2 For any wAMU estimator $\hat{\theta}_n$ of θ, it holds that

$$\underline{\lim}_{a\to 0}\,\underline{\lim}_{n\to\infty} \frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{2e^{-n((a^2/4)+o(a^2))}/\{\sqrt{\pi n}\,(a + o(a))\}} \ge 1.$$

Proof Letting

$$Y(\theta, a) = \frac{a^2 - 2a(X_1 - \theta)}{1 + (X_1 - \theta)^2},$$

we obtain

$$E_{\theta+a}\,[Y(\theta,a)] = 0, \qquad E_{\theta+a}\,[Y^2(\theta,a)] = \frac{a^2}{2}\,, \qquad E_{\theta+a}\,[Y^3(\theta,a)] = \frac{3a^4}{8}\,,$$

$$E_{\theta+a}\,[Y^4(\theta,a)] = \frac{3}{8}\,a^4 + \frac{5}{16}a^6\,, \qquad E_{\theta+a}\,[Y^5(\theta,a)] = \frac{25}{128}\,a^6 + \frac{35}{128}a^8. \quad (3.3)$$

It follows from (3.2) that for small $a > 0$,

$$\mu(\theta,a) = E_{\theta+a}[Z_1(\theta,a)] = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}E_{\theta+a}\,[Y^k(\theta,a)]$$

$$= -\frac{a^2}{4} + \frac{a^4}{32} + O(a^6)\,, \quad (3.4)$$

$$\sigma^2(\theta\,,a) = E_{\theta+a}[Z_1^2(\theta_0,a)] - \mu^2(\theta,a) = \frac{a^2}{2} - \frac{3a^4}{32} + O(a^6). \quad (3.5)$$

Since $\tilde{\mu} = \mu(\theta,-a)$, $\sigma = \sqrt{\sigma^2(\theta\,,a)}$ and $\tilde{\sigma} = \sqrt{\sigma^2(\theta,-a)}$, substituting (3.4) and (3.5) into (2.5) we have

$$\varliminf_{a\to 0}\varliminf_{n\to\infty}\frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{2e^{-n((a^2/4)+o(a^2))}/\{\sqrt{\pi n}\,(a+o(a))\}} \geq 1,$$

which completes the proof. □

Next, denote the likelihood function by $L(\theta;\boldsymbol{x}) = \prod_{i=1}^{n} f_0(x_i - \theta)$. Letting the MLE $\hat{\theta}_{ML}$ be a root of the likelihood equation $(\partial/\partial\theta)\log L(\theta;\boldsymbol{x}) = 0$ which maximizes $L(\theta;\boldsymbol{x})$. Then $\hat{\theta}_{ML}$ is seen to be wAMU, since the p.d.f. $f_0(x-\theta)$ of (3.1) is symmetric around $x = \theta$. Then we have the following.

Theorem 3 The MLE $\hat{\theta}_{ML}$ is wLDE.

Proof First, for each $i = 1, 2, \cdots,$ we put

$$W_i(a) = \frac{2(X_i - a)}{1 + (X_i - a)^2}.$$

Then we have for small $a(>0)$ and large n,

$$P_{\theta,n}\{\hat{\theta}_{ML} - \theta > a\} = P_{0,n}\{\hat{\theta}_{ML} > a\} = P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial a}\log f_0(X_i - a) > 0\right\}$$

$$= P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{2(X_i - a)}{1+(X_i - a)^2} > 0\right\} = P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n}W_i\,(a) > 0\right\}. \quad (3.6)$$

Then the moment generating function (m.g.f.) of $W_1(a)$ is given by

$$M_{W_1(a)}(t) = E_0\left[\exp\left\{\frac{2t(X_1 - a)}{1+(X_1 - a)^2}\right\}\right]$$

$$= \int_{-\infty}^{\infty} \left\{ \exp\left(\frac{2t(x-a)}{1+(x-a)^2} \right) \right\} \frac{1}{\pi(1+x^2)} dx \,. \tag{3.7}$$

For small $a(> 0)$ and small $|t|$ we have

$$\frac{2t(x-a)}{1+(x-a)^2} = \frac{2tx}{1+x^2} + A_t(x)a + B_t(x)a^2 + \cdots \;, \tag{3.8}$$

where

$$A_t(x) = \frac{4tx^2}{(1+x^2)^2} - \frac{2t}{1+x^2} \,, \tag{3.9}$$

$$B_t(x) = -\frac{6tx}{(1+x^2)^2} + \frac{8tx^3}{(1+x^2)^3} \,. \tag{3.10}$$

From (3.7) and (3.8) we obtain for small $a(> 0)$ and small $|t|$

$$\begin{aligned} M_{W_1(a)}(t) &= \int_{-\infty}^{\infty} \left\{ \exp\left(\frac{2tx}{1+x^2} \right) \right\} \frac{1}{\pi(1+x^2)} dx \\ &\quad + \int_{-\infty}^{\infty} \left\{ \exp\left(\frac{2tx}{1+x^2} \right) \right\} \left\{ A_t(x)a + \left(B_t(x) + \frac{1}{2} A_t^2(x) \right) a^2 \right\} \\ &\qquad \cdot \frac{1}{\pi(1+x^2)} dx + O(a^3). \end{aligned} \tag{3.11}$$

Since for small $|t|$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{1}{\pi(1+x^2)} dx = 1 + \frac{t^2}{4} + \frac{t^4}{64} + O(t^6) \,, \tag{3.12}$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{1}{\pi(1+x^2)^2} dx = \frac{1}{2} + \frac{t^2}{8} + \frac{t^4}{128} + O(t^6) \,,$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{1}{\pi(1+x^2)^3} dx = \frac{3}{8} + \frac{5t^2}{64} + \frac{11t^4}{1536} + O(t^6) \,,$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{x}{\pi(1+x^2)^3} dx = \frac{t}{8} + \frac{t^3}{32} + O(\lceil t \rceil^5) \,,$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{x^2}{\pi(1+x^2)^3} dx = \frac{1}{8} + \frac{3t^2}{64} + \frac{t^4}{1536} + O(t^6) \,,$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{x^2}{\pi(1+x^2)^4} dx = \frac{1}{16} \left(1 + \frac{3}{8} t^2 + \frac{5t^4}{192} \right) + O(t^6) \,,$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{x^3}{\pi(1+x^2)^4} dx = \frac{3t}{64} + \frac{5}{768} t^3 + O(\lceil t \rceil^5) \,,$$

$$\int_{-\infty}^{\infty} \left(\exp\frac{2tx}{1+x^2} \right) \frac{x^4}{\pi(1+x^2)^5} dx = \frac{1}{8} \left(\frac{3}{16} + \frac{5}{64} t^2 + \frac{35}{6144} t^4 \right) + O(t^6) \,,$$

we have by (3.9) and (3.10)

$$\int_{-\infty}^{\infty}\left(\exp\frac{2tx}{1+x^2}\right)A_t\,(x)\frac{1}{\pi(1+x^2)}dx=-\frac{t}{2}-\frac{t^3}{16}-\frac{5t^5}{384}+O(t^6)\,, \tag{3.13}$$

$$\int_{-\infty}^{\infty}\left(\exp\frac{2tx}{1+x^2}\right)B_t\,(x)\frac{1}{\pi(1+x^2)}dx=-\frac{3t^2}{8}-\frac{13t^4}{96}+O(t^6)\,, \tag{3.14}$$

$$\int_{-\infty}^{\infty}\left(\exp\frac{2tx}{1+x^2}\right)A_t^2\,(x)\frac{1}{\pi(1+x^2)}dx\ =\frac{7t^2}{8}+\frac{3t^4}{32}+\frac{107t^6}{3072}+O(\lceil t\rceil^7)\,. \tag{3.15}$$

Substituting (3.12) to (3.15) into (3.11), we have for small $a(>0)$ and small $|t|$

$$M_{W_1(a)}(t)=1-\frac{at}{2}+\frac{1}{4}\left(1+\frac{a^2}{4}\right)t^2+O(\lceil t\rceil^3)\,,$$

which yields the cumulant generating function (c.g.f.)

$$K_{W_1\,(a)}(t):=\log M_{W_1\,(a)}(t)=-\frac{at}{2}+\frac{1}{4}\left(1-\frac{a^2}{4}+a^3-\frac{a^4}{2}\right)t^2\ +O(\lceil t\rceil^3) \tag{3.16}$$

of $Z_1(a)$. Letting $K'_{W_1(a)}(t)=0$, we obtain

$$t=\hat{t}\ =a+O(a^2)\,, \tag{3.17}$$

which yields

$$M^n_{W_1\,(a)}(\hat{t})=e^{-n\left((a^2/4)+O(a^3)\right)} \tag{3.18}$$

and

$$\sigma^2_{W_1\,(a)}(\hat{t}):=K''_{W_1(a)}(\hat{t})=\frac{1}{2}\left(1+O(a^2)\right). \tag{3.19}$$

By the saddlepoint approximation (6.1) in the appendix of Section 6 later and (3.17) to (3.19) we have for small $a(>0)$

$$P_{0,n}\left\{\frac{1}{n}\sum\nolimits_{i=1}^{n}W_i\,(a)>0\right\}=\frac{M^n_{W_1\,(a)}(\hat{t})}{\sqrt{n}\hat{t}\sqrt{\sigma^2_{W_1\,(a)}(\hat{t})}}\left\{\frac{1}{\sqrt{2\pi}}+O\left(\frac{1}{n}\right)\right\}$$

$$=\frac{e^{-n\left((a^2/4)+O(a^3)\right)}}{\sqrt{\pi n}\left(a+O(a^2)\right)}\left\{1+O\left(\frac{1}{n}\right)\right\} \tag{3.20}$$

(see Jensen 1995, Section 2.2 and Akahira 2024, Appendix).

In a similar way to the above, we obtain the LD probability

$$P_{\theta,n}\{\hat{\theta}_{ML}-\theta<-a\}=P_{0,n}\{\hat{\theta}_{ML}<-a\}$$

$$=P_{0,n}\left\{\frac{1}{n}\sum\nolimits_{i=1}^{n}\frac{\partial}{\partial a}\log f_0(X_i+a)<0\right\}$$

$$= P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n} W_i\,(-a) < 0\right\} \tag{3.21}$$

for large n. Letting $K'_{W_1(-a)}(t) = 0$, we obtain for small $a(> 0)$

$$t = \hat{t} = -a + O(a^2)\,, \tag{3.22}$$

which yields

$$M^n_{W_1\,(-a)}(\hat{t}) = e^{-n\left((a^2/4)+O(a^3)\right)}, \tag{3.23}$$

$$\sigma^2_{W_1\,(-a)}(\hat{t}) := K''_{W_1(-a)}(\hat{t}) = \frac{1}{2}\left(1 + O(a^2)\right). \tag{3.24}$$

Then it follows from the saddlepoint approximation (6.2) in the appendix of Section 6 later and (3.22) to (3.24) that for small $a(> 0)$ and large n

$$P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n} W_i\,(-a) < 0\right\} = \frac{M^n_{W_1\,(-a)}(\hat{t})}{\sqrt{n}|\hat{t}|\sqrt{\sigma^2_{W_1\,(-a)}(\hat{t})}}\left\{\frac{1}{\sqrt{2\pi}} + O\left(\frac{1}{n}\right)\right\}$$

$$= \frac{e^{-n\left((a^2/4)+O(a^3)\right)}}{\sqrt{\pi n}\left(a + O(a^2)\right)}\left\{1 + O\left(\frac{1}{n}\right)\right\} \tag{3.25}$$

(see also Akahira 2024, Appendix). From (3.6), (3.20), (3.21) and (3.25) we have for small $a(> 0)$ and large n

$$\frac{P_{\theta,n}\left\{\left|\hat{\theta}_{ML} - \theta\right| > a\right\}}{2e^{-n\left((a^2/4)+O(a^3)\right)}/\left\{\sqrt{\pi n}\,\left(a + O(a^2)\right)\right\}} = 1 + O\left(\frac{1}{n}\right). \tag{3.26}$$

From (3.26) we obtain

$$\lim_{a\to 0}\lim_{n\to\infty}\frac{P_{\theta,n}\left\{\left|\hat{\theta}_{ML} - \theta\right| > a\right\}}{2e^{-n\left((a^2/4)+o(a^2)\right)}/\left\{\sqrt{\pi n}\,\left(a + o(a)\right)\right\}} = 1\,. \tag{3.27}$$

Hence it is seen from Definition 2, Theorem 2 and (3.27) that $\hat{\theta}_{ML}$ is wLDE. □

Remark 1 It is easily seen from Theorem 2 that $\hat{\theta}_{ML}$ is asymptotically Bahadur efficient in the sense of

$$\lim_{a\to 0}\lim_{n\to\infty}\frac{1}{na^2}\log P_{\theta,n}\left\{\left|\hat{\theta}_{ML} - \theta\right| > a\right\} = -\frac{1}{4}\,, \tag{3.28}$$

which is consistent with the result of Bai and Fu (1987). Indeed, it follows from (3.27) that for small $a(> 0)$ and large n

$$\frac{1}{na^2}\log P_{\theta,n}\left\{\left|\hat{\theta}_{ML}-\theta\right|>a\right\}$$

$$=\frac{1}{na^2}\left\{\log 2-n\left(\frac{a^2}{4}+o(a^2)\right)-\frac{1}{2}\log\pi-\frac{1}{2}\log n\right.$$

$$\left.-\log\bigl(a+o(a)\bigr)+\log\left(1+O\left(\frac{1}{n}\right)\right)\right\},$$

which yields (3.28). Here, it is noted that the Fisher information amount $I(\theta)$ is 1/2 in the Cauchy case.

4 Second order LD efficiency

It is interesting to consider the problem whether the MLE is second order LDE or not in the Cauchy case. In a similar way to the theorem in Akahira (2024), the lower bound for the tail probability $P_{\theta,n}\{\hat{\theta}_n-\theta>a\}$ in the class $\mathbb{A}$ is given up to the order $o(n^{-1})$, under the same setup in Section 3 as follows.

Theorem 4 For any $\hat{\theta}_n\in\mathbb{A}$, and any $a>0$, it holds that for large n

$$\frac{P_{\theta,n}\{\hat{\theta}_n-\theta>a\}}{e^{n\mu(\theta,a)}/\sqrt{2\pi n\sigma^2(\theta,a)}}\geq 1-\frac{1}{n\sigma^2(\theta,a)}+\frac{\zeta_3(-1;\theta,a)}{2n\sigma(\theta,a)}$$

$$+\frac{1}{n}\left\{\frac{1}{8}\zeta_4(-1;\theta,a)-\frac{5}{24}\zeta_3^2(-1;\theta,a)\right\}+O\left(\frac{1}{n^2}\right), \quad (4.1)$$

where

$$\zeta_3(-1;\theta,a)=\kappa_{3,\theta+a}\bigl(Z_1(\theta,a)\bigr)/\bigl(\sigma^2(\theta,a)\bigr)^{3/2},$$

$$\zeta_4(-1;\theta,a)=\kappa_{4,\theta+a}\bigl(Z_1(\theta,a)\bigr)/\sigma^4(\theta,a)$$

with third and fourth cumulants $\kappa_{3,\theta+a}\bigl(Z_1(\theta,a)\bigr)$ and $\kappa_{4,\theta+a}\bigl(Z_1(\theta,a)\bigr)$ of $Z_1(\theta,a)$.

The proof is straightforward from (2.4), (6.2) and Remark 3 in the appendix of Section 6 later (see Corollary in Akahira 2024).

Remark 2 Denote by $\Delta(\theta,a)$ the coefficient of order n^{-1} in the right-hand side of (4.1). For any $\hat{\theta}_n\in\mathbb{A}$, and any $a>0$, it holds that for large n

$$\frac{P_{\theta,n}\{\hat{\theta}_n-\theta<-a\}}{e^{n\mu(\theta,-a)}/\sqrt{2\pi n\sigma^2(\theta,-a)}}\geq 1+\frac{1}{n}\Delta(\theta,-a)+O\left(\frac{1}{n^2}\right). \quad (4.2)$$

Let $\mu=\mu(\theta,a)$, $\tilde{\mu}=\mu(\theta,-a)$, $\sigma=\sqrt{\sigma^2(\theta,a)}$ *and* $\tilde{\sigma}=\sqrt{\sigma^2(\theta,-a)}$.

Corollary 1 For any $\hat{\theta}_n \in \mathbb{A}$, and any $a > 0$, it holds that for large n

$$\frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{\frac{1}{\sqrt{2\pi n}}\left(\frac{1}{\sigma}e^{n\mu} + \frac{1}{\tilde{\sigma}}e^{n\tilde{\mu}}\right)}$$

$$\geq 1 + \frac{1}{n}\left(\frac{1}{\sigma}e^{n\mu} + \frac{1}{\tilde{\sigma}}e^{n\tilde{\mu}}\right)^{-1}\left\{\frac{1}{\sigma}e^{n\mu}\Delta(\theta,a) + \frac{1}{\tilde{\sigma}}e^{n\tilde{\mu}}\Delta(\theta,-a)\right\}$$

$$+O\left(\frac{1}{n^2}\right). \tag{4.3}$$

The proof is straightforward from (4.1) and (4.2). Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables according to the Cauchy distribution with p.d.f. (3.1). Then we obtain the lower bound for the LD probability of wAMU estimators as follows.

Theorem 5 For any $\hat{\theta}_n \in \mathbb{A}$, and small $a > 0$,

$$\frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{2e^{-n((a^2/4)+o(a^2))}/\{\sqrt{\pi n}\,(a + o(a))\}} \geq 1 - \frac{1}{n}\left(\frac{3}{16} + \frac{2}{a^2} + o(a)\right) + O\left(\frac{1}{n^2}\right). \tag{4.4}$$

Proof In a similar way to (3.4), we have from (3.3)

$$\kappa_3(\theta,a) := E_{\theta+a}[\{Z_1(\theta,a) - \mu(\theta,a)\}^3] = \frac{3}{16}a^4 + O(a^6),$$

$$\kappa_4(\theta,a) := E_{\theta+a}[\{Z_1(\theta,a) - \mu(\theta,a)\}^4] - 3\sigma^4(\theta,a) = -\frac{3}{8}a^4 + O(a^6),$$

which yields

$$\zeta_3(-1;\theta,a) = \frac{\kappa_3(\theta,a)}{(\sigma^2(\theta,a))^{3/2}} = \frac{3\sqrt{2}}{8}\,a + O(|a|^3), \tag{4.5}$$

$$\zeta_4(-1;\theta,a) = \frac{\kappa_4(\theta,a)}{\sigma^4(\theta,a)} = -\frac{3}{2} + \frac{a^2}{4} + O(a^4) \tag{4.6}$$

from (3.5). Then it follows from (3.4), (3.5), (4.1), (4.5), (4.6) and Remark 2 that for small $a(> 0)$

$$\Delta(\theta,a) = -\frac{1}{\sigma^2(\theta,a)} + \frac{\zeta_3(-1;\theta,a)}{2\sigma(\theta,a)} + \frac{1}{8}\zeta_4(-1;\theta,a) - \frac{5}{24}\zeta_3^2(-1;\theta,a)$$

$$= -\left(\frac{3}{16} + \frac{2}{a^2}\right) + O\left(\frac{a^2}{n}\right)$$

for large n, which yields

$$\Delta(\theta, a) = \Delta(\theta, -a) = -\left(\frac{3}{16} + \frac{2}{a^2}\right) + O\left(\frac{a^2}{n}\right).$$

Since from (3.4) and (3.5)

$$\mu(\theta, a) = \mu(\theta, -a) = -\frac{a^2}{4} + \frac{a^4}{32} + O(a^6),$$

$$\sigma^2(\theta, a) = \sigma^2(\theta, -a) = \frac{a^2}{2} - \frac{3a^4}{32} + O(a^6),$$

for small $a(>0)$, it is seen from (4.3) that

$$\frac{P_{\theta,n}\{|\hat{\theta}_n - \theta| > a\}}{2e^{-n((a^2/4)+o(a^2))}/\{\sqrt{\pi n}\,(a + o(a))\}} \geq 1 - \frac{1}{n}\left(\frac{3}{16} + \frac{2}{a^2} + o(a)\right) + O\left(\frac{1}{n^2}\right)$$

for large n, which yields (4.4). Thus we compete the proof. □

Theorem 6 For small $a > 0$, the LD probability of the MLE $\hat{\theta}_{ML}$ of θ up to the order $o(1/n)$ is given by

$$\frac{P_{\theta,n}\{|\hat{\theta}_{ML} - \theta| > a\}}{2e^{-n((a^2/4)+o(a^2))}/\{\sqrt{\pi n}\,(a + o(a))\}} = 1 + \frac{1}{n}\left(\frac{1}{16} - \frac{2}{a^2} + o(a)\right) + O\left(\frac{1}{a^2 n^2}\right) \quad (4.7)$$

for large n.

Proof Substituting (3.12) to (3.15) into (3.11), we have for small $a(>0)$ and small $|t|$

$$\begin{aligned} M_{W_1(a)}(t) &= 1 - \frac{at}{2} + \frac{1}{4}\left(1 + \frac{a^2}{4}\right)t^2 - \frac{a}{16}t^3 \\ &\quad + \frac{1}{64}\left(1 - \frac{17}{3}a^2\right)t^4 - \frac{5a}{384}t^5 + O(t^6) \\ &= 1 + \Delta_a(t) \quad \text{(say)}, \end{aligned} \quad (4.8)$$

which yields the c.g.f.

$$K_{W_1(a)}(t) = \log M_{W_1(a)}(t) = \log(1 + \Delta_a(t)) \quad (4.9)$$

of $Z_1(a)$. Then we have

$$K'_{W_1(a)}(t) = \frac{\Delta'_a(t)}{1 + \Delta_a(t)}, \quad (4.10)$$

$$K''_{W_1(a)}(t) = \frac{\Delta''_a(t)}{1 + \Delta_a(t)} - \left(K'_{W_1(a)}(t)\right)^2, \quad (4.11)$$

$$K^{(3)}_{W_1(a)}(t) = \frac{d^3}{dt^3} K_{W_1(a)}(t)$$

$$= \frac{\Delta_a^{(3)}(t)}{1+\Delta_a(t)} - \frac{\Delta_a'(t)\Delta_a''(t)}{\left(1+\Delta_a(t)\right)^2} - 2K_{W_1(a)}'(t)K_{W_1(a)}''(t)\,, \tag{4.12}$$

$$K_{W_1(a)}^{(4)}(t) = \frac{d^4}{dt^4}K_{W_1(a)}(t)$$

$$= \frac{\Delta_a^{(4)}(t)}{1+\Delta_a(t)} - \frac{\Delta_a'(t)\Delta_a^{(3)}(t)}{\left(1+\Delta_a(t)\right)^2} - \frac{\Delta_a'(t)\Delta_a^{(3)}(t) + \left(\Delta_a''(t)\right)^2}{\left(1+\Delta_a(t)\right)^2}$$

$$+ \frac{2\left(\Delta_a'(t)\right)^2\Delta_a''(t)}{\left(1+\Delta_a(t)\right)^3} - 2\left(K_{W_1(a)}''(t)\right)^2 - 2K_{W_1(a)}'(t)K_{W_1(a)}^{(3)}(t). \tag{4.13}$$

From (4.8) we obtain for small $a(>0)$ and small $|t|$

$$\Delta_a'(t) = -\frac{a}{2} + \frac{1}{2}\left(1+\frac{a^2}{4}\right)t - \frac{3a}{16}t^2 + \frac{1}{16}\left(1-\frac{17}{3}a^2\right)t^3$$

$$-\frac{25a}{384}t^4 + O(|t|^5), \tag{4.14}$$

$$\Delta_a''(t) = \frac{1}{2}\left(1+\frac{a^2}{4}\right) - \frac{3a}{8}t + \frac{3}{16}\left(1-\frac{17}{3}a^2\right)t^2 - \frac{25a}{96}t^3 + O(t^4), \tag{4.15}$$

$$\Delta_a^{(3)}(t) = -\frac{3a}{8} + \frac{3}{8}\left(1-\frac{17}{3}a^2\right)t - \frac{25a}{32}t^2 + O(|t|^3), \tag{4.16}$$

$$\Delta_a^{(4)}(t) = \frac{3}{8}\left(1-\frac{17}{3}a^2\right) - \frac{25a}{16}t + O(t^2). \tag{4.17}$$

From (4.8), (4.10) and (4.14) we have for small $a(>0)$ and small $|t|$

$$K_{W_1(a)}'(t) = -\frac{a}{2} + \frac{1}{2}\left(1-\frac{a^2}{4}\right)t + \frac{3a}{16}\left(1-\frac{a^2}{6}\right)t^2 - \frac{t^3}{16}\left(1-\frac{28}{3}a^2+\frac{1}{8}a^4\right)$$

$$-\frac{a}{128}t^4\left(\frac{40}{3}+\frac{137}{6}a^2+3a^3+\frac{17}{4}a^4\right) + O(|t|^5).$$

Letting $K_{W_1(a)}'(t) = 0$, we obtain for small $a(>0)$

$$t = \hat{t} = a + O(a^5)\,, \tag{4.18}$$

which yields

$$\Delta_a(\hat{t}) = -\frac{a^2}{4} + O(a^4),\ \ \Delta_a'(\hat{t}) = O(a^5)\,,\ \ \Delta_a''(\hat{t}) = \frac{1}{2} - \frac{a^2}{16} + O(a^4),$$

$$\Delta_a^{(3)}(\hat{t}) = O(a^3), \qquad \Delta_a^{(4)}(t) = \frac{3}{8} + O(a^2)$$

from (4.8) and (4.14) to (4.17). From (4.8) to (4.17) we have for small $a(>0)$

$$K_{W_1(a)}(\hat{t}) = -\frac{a^2}{4} + O(a^4)\,, \tag{4.19}$$

$$K'_{W_1(a)}(\hat{t}) = O(a^5)\,, \tag{4.20}$$

$$K''_{W_1(a)}(\hat{t}) = \frac{1}{2} + \frac{a^2}{16} + O(a^4)\,, \tag{4.21}$$

$$K^{(3)}_{W_1(a)}(\hat{t}) = O(a^3)\,, \tag{4.22}$$

$$K^{(4)}_{W_1(a)}(\hat{t}) = -\frac{3}{8} + O(a^2)\,. \tag{4.23}$$

Then it follows from (3.19), (4.18), (4.20) to (4.23) that for small $a(>0)$,

$$\hat{\lambda} = \sqrt{n}\hat{t}\sqrt{\sigma^2_{W_1(a)}(\hat{t})} = \sqrt{\frac{n}{2}}a\left(1 + \frac{a^2}{16} + O(a^4)\right), \tag{4.24}$$

$$\zeta_3(\hat{t}) = \frac{K^{(3)}_{W_1(a)}(\hat{t})}{\left(\sigma^2_{W_1(a)}(\hat{t})\right)^{3/2}} = O(a^3)\,, \tag{4.25}$$

$$\zeta_4(\hat{t}) = \frac{K^{(4)}_{W_1(a)}(\hat{t})}{\sigma^4_{W_1(a)}(\hat{t})} = -\frac{3}{2} + O(a^2)\,. \tag{4.26}$$

Since $B_i(\cdot)$ $(i = 0, 3, 4, 6)$ are defined in the appendix of Section 6 later, it is seen from (4.24) and Remark 3 in Section 6 that

$$B_0(\hat{\lambda}) = \frac{1}{\sqrt{2\pi}}\left\{1 - \frac{2}{a^2 n} + \frac{1}{4n} + O\left(\frac{a^2}{n}\right)\right\} + O\left(\frac{1}{a^4 n}\right), \tag{4.27}$$

$$B_3(\hat{\lambda}) = -\frac{3}{\sqrt{\pi n}\,a}\left\{1 + \frac{1}{16}a^2 + O(a^4)\right\} + O\left(\frac{1}{a^3 n^{3/2}}\right), \tag{4.28}$$

$$B_4(\hat{\lambda}) = \frac{3}{\sqrt{2\pi}} + O\left(\frac{1}{a^2 n}\right), \tag{4.29}$$

$$B_6(\hat{\lambda}) = -\frac{15}{\sqrt{2\pi}} + O\left(\frac{1}{a^2 n}\right). \tag{4.30}$$

From (4.19), (4.24) to (4.30) and the saddlepoint approximation (6.1) in the appendix of Section 6 later, we have for small $a(>0)$

$$P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n} W_i(a) > 0\right\}$$

$$= \frac{e^{-n\left((a^2/4)+O(a^4)\right)}}{\sqrt{\pi n}\left(a+O(a^3)\right)}\left\{1+\frac{1}{n}\left(\frac{1}{16}-\frac{2}{a^2}+O(a^2)\right)+O\left(\frac{1}{a^2n^2}\right)\right\} \quad (4.31)$$

for large n. Replacing a with $-a$ in the above and using the approximation formula (6.2) in the appendix of Section 6 later, we have for small $a(> 0)$

$$P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n} W_i(-a) < 0\right\}$$

$$= \frac{e^{-n\left((a^2/4)+O(a^4)\right)}}{\sqrt{\pi n}\left(a+O(a^3)\right)}\left\{1+\frac{1}{n}\left(\frac{1}{16}-\frac{2}{a^2}+O(a^2)\right)+O\left(\frac{1}{a^2n^2}\right)\right\}$$

for large n. Then it follows from (3.6), (3.21), (4.31) and (4.32) that

$$P_{\theta,n}\{|\hat{\theta}_{ML}-\theta| > a\} = P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n} W_i(a) > 0\right\} + P_{0,n}\left\{\frac{1}{n}\sum_{i=1}^{n} W_i(-a) < 0\right\}$$

$$= \frac{2e^{-n\left((a^2/4)+o(a^2)\right)}}{\sqrt{\pi n}\left(a+o(a)\right)}\left\{1+\frac{1}{n}\left(\frac{1}{16}-\frac{2}{a^2}+o(a)\right)+O\left(\frac{1}{a^2n^2}\right)\right\},$$

which yields (4.7). Thus we complete the proof. □

From Theorems 5 and 6 it is seen that the MLE $\hat{\theta}_{ML}$ does not attain the lower bound, i.e. the right-hand side of (4.4) up to the order $o(1/n)$, which implies that $\hat{\theta}_{ML}$ is not second order LDE.

5. Concluding remarks

In the problem of estimating a location parameter of the Cauchy distribution, we consider the LD efficiency of the MLE. First, using the saddlepoint approximation we obtain the lower bound for the LD probability of wAMU estimators. Second, the MLE is shown to be LDE in the sense that it attains the lower bound, and it follows that the MLE is asymptotically Bahadur efficient in the sense of (3.28). Further, it is interesting whether the MLE is second order LDE or not. However, it is shown that the MLE is not second order LDE, since the MLE does not attain the lower bound for the LD probability of wAMU estimators up to the order $o(1/n)$.

In the higher order asymptotics evaluated by the concentration probability of estimators around the true parameter, it is known that the MLE is (first order) asymptotically efficient, but not third order asymptotically efficient for the Cauchy distribution (see Akahira and Takeuchi, 1981, pp.110, 111).

6 Appendix

We consider the classical saddlepoint approximation. Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables with p.d.f. $f_0(x)$ w.r.t. a σ-finite measure μ. Let $\mathcal{X}$ be a sample space of X_1. Then the m.g.f. and c.g.f. of X_1 are given by

$$M_{X_1}(t) = E[e^{tX_1}] = \int_{\mathcal{X}} e^{tx} f_0(x)\ d\mu$$

and $K_{X_1}(t) = \log M_{X_1}(t)$, respectively, where

$$t \in \mathcal{T} := \{\, t \mid M_{X_1}(t) < \infty \}.$$

Letting

$$f_t(x) = e^{tx} f_0(x)/M_{X_1}(t), \qquad t \in \mathcal{T}$$

we see that for each $t \in \mathcal{T}$, $f_t(x)$ is p.d.f. (w.r.t. μ). A set of distributions with p.d.f. $f_t(x)$ generates an exponential family of distributions. The mean and variance of X_1 w.r.t. p.d.f. f_t are given by

$$\mu(t) := E_t[X_1] = \frac{d}{dt} K_{X_1}(t),$$

$$\sigma^2(t) := V_t(X_1) = \frac{d^2}{dt^2} K_{X_1}(t),$$

respectively. The mth cumulant of X_1 is also obtained from

$$\kappa_m(t) = \frac{d^m}{dt^m} K_{X_1}(t)$$

for $m \geq 3$. Here, we consider only those values of x which there exists $t = \hat{t}(x)$ such that $\mu(t) = x$, and the upper tail probability $P\{\bar{X} \geq x\}$ of $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$, where $x \geq \mu(0)$. Then it is noted that $\hat{t} \geq 0$. When $x < \mu(0)$, we can consider the lower tail probability $P\{\bar{X} \leq x\}$ by replacing X_i by $-X_i$ for each $i = 1, 2, \cdots$ and note that $\hat{t} < 0$.

The approximation formula to the upper tail probability for $x \geq \mu(0)$ is given by

$$P\{\bar{X} \geq x\} = \frac{\{M_{X_1}(\hat{t})\}^n e^{-n\hat{t}x}}{\sqrt{n}\hat{t}\sigma(\hat{t})}\left[B_0(\hat{\lambda}) + \frac{\zeta_3(\hat{t})}{6\sqrt{n}} B_3(\hat{\lambda}) + \frac{1}{n}\left\{\frac{\zeta_4(\hat{t})}{24} B_4(\hat{\lambda}) + \frac{(\zeta_3(\hat{t}))^2}{72} B_6(\hat{\lambda})\right\} + O\left(\frac{1}{n^2}\right)\right] \qquad (6.1)$$

where $\sigma(t) = \sqrt{\sigma^2(t)}$, $\hat{\lambda} = \sqrt{n}\hat{t}\sigma(\hat{t})$, $\zeta_m(t) = \kappa_m(t)/\{\sigma(t)\}^m$ with $\hat{t} = \hat{t}(x)$, and

$$B_0(\lambda) := \lambda e^{\lambda^2/2}\{1 - \Phi(\lambda)\},$$

$$B_3(\lambda) := -\left\{\lambda^3 B_0(\lambda) - \frac{1}{\sqrt{2\pi}}(\lambda^3 - \lambda)\right\},$$

$$B_4(\lambda) := \lambda^4 B_0(\lambda) - \frac{1}{\sqrt{2\pi}}(\lambda^4 - \lambda^2),$$

$$B_6(\lambda) := \lambda^6 B_0(\lambda) - \frac{1}{\sqrt{2\pi}}(\lambda^6 - \lambda^4 + 3\lambda^2)$$

(see Jensen 1995, Section 2.2). The approximation formula to the lower tail probability for $x < \mu(0)$ is also given by

$$P\{\bar{X} \le x\} = \frac{\{M_{X_1}(\hat{t})\}^n e^{-n\hat{t}x}}{\sqrt{n}|\hat{t}|\sigma(\hat{t})}\left[B_0(\hat{\lambda}) - \frac{\zeta_3(\hat{t})}{6\sqrt{n}}B_3(\hat{\lambda}) + \frac{1}{n}\left\{\frac{\zeta_4(\hat{t})}{24}B_4(\hat{\lambda}) + \frac{(\zeta_3(\hat{t}))^2}{72}B_6(\hat{\lambda})\right\} + O\left(\frac{1}{n^2}\right)\right], \tag{6.2}$$

where $\hat{\lambda} = \sqrt{n}|\hat{t}|\sigma(\hat{t})$.

Remark 3 Mills' ratio is expanded as

$$1 - \Phi(\lambda) = \phi(\lambda)\left\{\frac{1}{\lambda} - \frac{1}{\lambda^3} + \frac{3}{\lambda^5} - \frac{3\cdot 5}{\lambda^7} + \cdots + \frac{(-1)^k}{\lambda^{2k+1}}(1\cdot 3\cdots(2k-1)) + O\left(\frac{1}{\lambda^{2k+3}}\right)\right\}$$

for large λ (see Jensen 1995, page 24), where $\Phi(\cdot)$ and $\phi(\cdot)$ are the c.d.f. and p.d.f. of the standard normal distribution $N(0,1)$. Then

$$B_0(\lambda) = \frac{1}{\sqrt{2\pi}}\left\{1 - \frac{1}{\lambda^2} + \frac{3}{\lambda^4} - \frac{15}{\lambda^6} + O\left(\frac{1}{\lambda^8}\right)\right\},$$

$$B_3(\lambda) = -\frac{1}{\sqrt{2\pi}}\left\{\frac{3}{\lambda} + O\left(\frac{1}{\lambda^3}\right)\right\},$$

$$B_4(\lambda) = \frac{3}{\sqrt{2\pi}} + O\left(\frac{1}{\lambda^2}\right),$$

$$B_6(\lambda) = -\frac{15}{\sqrt{2\pi}} + O\left(\frac{1}{\lambda^2}\right).$$

Acknowledgements

Research was supported by the Research Institute for Mathematical Sciences, an International Joint Usage/Research Center located in Kyoto University and JSPS KAKENHI Grant number JP25K15012.

References

Akahira, M. (2006). Large-deviation efficiency of first and second order. *Student* **5,** 211-219.

Akahira, M. (2010). The first- and second-order large-deviation efficiency for an exponential family and certain curved exponential models. *Commun. Statist. – Theory and Meth.,* **39**, 1387-1403.

Akahira, M. (2024). Large deviation behavior of estimators for flattened distributions in a middle part. *RIMS Kôkyûroku* **2884**, Kyoto University, 105-121.

Akahira, M. and Takeuchi, K. (1981). *Asymptotic Efficiency of Statistical Estimators: Concepts and Higher Order Asymptotic Efficiency*. Lecture Notes in Statistics 7, Springer, New York.

Akaoka, Y., Okamura, K. and Otobe, Y. (2022). Bahadur efficiency of the maximum likelihood estimator and one-step estimator for quasi-arithmetic means of the Cauchy distribution. *Ann. Inst. Statist. Math.,* **74**, 895-923.

Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.,* **38**, 303-324.

Bahadur, R. R. (1971). *Some Limit Theorems in Statistics*.Regional Conference Series in Applied Mathematics 4, SIAM, Philadelphia.

Bai, Z. D. and Fu, J. C. (1987). On the maximum-likelihood estimator for the location parameter of a Cauchy distribution. *Can. J. Statist.,* **15**, 137-146.

Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*. Chapman Hall, London.

Daniels, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.,* **25**, 631-650.

Daniels, H. E. (1987). Tail probability approximations. Int. Statist. Review 55, 37-48.

Efron, B. (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.,* **3**, 1189-1242.

Fu, J. C. (1973). On a theorem of Bahadur on the rate of convergence of point estimators. Daniels, H. E. (1954). Saddlepoint approximations.

Fu, J. C. (1982). Large sample point estimation: A large deviation theory approach. *Ann. Statist.*, **10**, 762-771.

Jensen, J. L. (1995). *Saddlepoint Approximations*. Clarendon Press, Oxford.

Institute of Mathematics
University of Tsukuba
Tsukuba, Ibaraki 305-8571
JAPAN