Commentationes Mathematicae Universitatis Carolinae

Jana Jurečková

Finite-sample comparison of L-estimators of location

Commentationes Mathematicae Universitatis Carolinae, Vol. 20 (1979), No. 3, 507--518

Persistent URL: http://dml.cz/dmlcz/105947

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE
20,3 (1979)

FINITE-SAMPLE COMPARISON OF L-ESTIMATORS OF LOCATION

Jana JUREČKOVÁ

Abstract: Let $X_1,\dots,X_n$ be a random sample from a population with the density $f(x-\Theta)$ such that f is symmetric and positive. The efficiency of estimator T_n of Θ based on $X_1,\dots,X_n$ with respect to extreme deviations from Θ is established. The sample mean $\bar{X}_n$ is then compared with other L-estimators with respect to this efficiency. The comparison among others yields a surprising result that, from this point of view, the sample mean is more efficient than the sample median even for the double-exponential and logistic distributions. On the other hand, each trimming brings an improvement with respect to $\bar{X}_n$ in the case of Cauchy distribution.

Key words: Tails of a probability distribution, L-estimator, efficiency of an estimator.

AMS: 62F10, 62G05

---

1. Introduction. Let $X_1,X_2,\dots$ be a sequence of independent random variables, identically distributed according to an absolutely continuous distribution function $F(x-\Theta)$ with positive density $f(x-\Theta)$ such that $f(-x) = f(x)$, $x \in R^1$. For any fixed n, let $T_n = T_n(X_1,\dots,X_n)$ be an estimator of Θ based on $X_1,\dots,X_n$. Different measures of performance of T_n have been suggested and investigated. Besides the classical mean-square-error approach, the probability

$$P_\Theta(|T_n-\Theta| > a) \tag{1.1}$$

of the absolute error of the estimator exceeding a fixed number $a>0$, or the relative measure

$$(1.2) \qquad \max\{P_\theta(T_n < \theta - a),\ P_\theta(T_n > \theta + a)\}$$

have been considered in several contexts. For instance, Huber [4] has shown that the M-estimator is the translation equivariant estimator which minimizes the inaccuracy (1.2) for any finite n.

If the sequence $\{T_n\}$ is consistent for θ, the inaccuracy (1.1) tends to 0 as $n \to \infty$. Bahadur ([1],[2]) proposed

$$(1.3) \qquad \lim_{n\to\infty}\{-\frac{1}{n}\ \ell n\ P_\theta(|T_n-\theta|>a)\} = e^*$$

for a fixed a, $a>0$, as a measure of asymptotic performance of $\{T_n\}$, if the limit exists. Bahadur [1] and Fu [3] gave an upper bound for e^* for consistent sequences of estimators. Sievers [8] evaluated the limits e^* and their upper bounds for several estimators and several distribution shapes. From this point of view, he found the sample median less efficient than the sample mean not only for normal but also for logistic distribution. He observed a similar feature even in the case of double-exponential distribution unless the values a were small.

Jurečková [5] considered the measure of performance (1.1) locally for small values a and n fixed. She found, for symmetric and unimodal populations, the maximum likelihood estimator T_n^* locally efficient in the sense that, for any other median unbiased estimator T_n, there exists an $a_0>0$ such that

$$P_\theta(|T_n^*-\theta|>a) \le P_\theta(|T_n-\theta|>a) \text{ for } 0<a<a_0.$$

She has also suggested an L-estimator (linear combination of order statistics) which has good local properties.

Generally, there is no estimator which minimizes (1.1) for all $a>0$. If we wish to find estimators which are not too sensitive to the gross-errors, we feel that the local aspect is unappropriate. Also, not any fixed value a seems to be reasonably preferable. On the other hand, if we want to be secure against the extreme deviations of T_n from the true parameter value, we may consider the inaccuracy (1.1) asymptotically as $a \to \infty$ and n is fixed. The inaccuracy (1.1) will tend to 0 as $a \to \infty$ for any reasonable estimator; thus, we shall consider

$$(1.4) \qquad \lim_{a\to\infty} \left\{ - \frac{1}{q(a)} \ln P_\theta(|T_n - \theta| > a) \right\}$$

where q(a) is a proper measure of the tails of f, being defined below, as a measure of performance of T_n, if the limit exists.

Considering the L-estimators from this point of view, we obtain results which may be surprising but which are consistent with Sievers' results. More precisely, we shall show that, for any distribution with the tails of the form $q(a) = b \cdot a^r$; $b>0$, $r \geq 1$ (it involves normal, logistic and double-exponential distributions), the sample mean is more efficient than any other L-estimator which puts weights 0 on the extreme observations. The situation is similar even for $0<r<1$ but then we are able to find the sample mean more efficient only for n up to some value $n_0 = n_0(r)$.

On the other hand, for Cauchy distribution which is typically heavy-tailed ($q(a) = \ln a$), any estimator which puts zero weights on the extreme observations is undisputably better than the sample mean. The situation is not yet clear for slightly less heavy-tailed distributions with $q(a) = m \cdot \ln a$ (e.g. t-distribution with m degrees of freedom).

2. Bounds for efficiency of an L-estimator with respect to extreme deviations

Definition 2.1. Let f be any symmetric density such that $f(x) > 0$, $x \in R^1$. We shall call any positive function $q(a)$, $a > 0$, such that $q(a) \uparrow \infty$ as $a \to \infty$ and that

$$\lim_{a \to \infty} \left\{ - \frac{1}{q(a)} \ln \int_{|x| > a} f(x)\, dx \right\} = 1 \tag{2.1}$$

the logarithmic measure of tails of the distribution with the density f.

Definition 2.2. Let $X_1, X_2, \ldots$ be a sequence of independent random variables, identically distributed according to a density $f(x-\theta)$, $\theta \in R^1$, where $f(-x) = f(x) > 0$, $x \in R^1$, with the logarithmic measure of tails $q(a)$. Let $T_n = T_n(X_1, \ldots, X_n)$ and $T_n' = T_n'(X_1, \ldots, X_n)$ be two translation equivariant estimators of θ based on $X_1, \ldots, X_n$. We say that T_n is more efficient than T_n' with respect to extreme deviations (shortly T_n is more efficient than T_n') if it holds

$$\begin{aligned} &\underline{\lim}_{a \to \infty} \left\{ - \frac{1}{q(a)} \ln P_\theta(|T_n| > a) \right\} \geq \\ &\geq \overline{\lim}_{a \to \infty} \left\{ - \frac{1}{q(a)} \ln P_0(|T_n'| > a) \right\} \end{aligned} \tag{2.2}$$

where the probability P_0 corresponds to $\theta = 0$.

In the subsequent text, we shall be interested in comparison of the sample mean $\bar{X}_n$ with other L-estimators with respect to Definition 2.2. Let us remind that T_n is called an L-estimator of θ if it has the form

$$T_n = \sum_{i=1}^{n} c_i \, X_n^{(i)} \tag{2.3}$$

where $X_n^{(1)} \leq \ldots \leq X_n^{(n)}$ are the order statistics corresponding to $X_1,\ldots,X_n$ and $c_i \geq 0$, $c_i = c_{n-i+1}$, $i=1,\ldots,n$, and $\sum_{i=1}^{n} c_i = 1$. The following theorem provides lower and upper bounds for the efficiency of an L-estimator.

Theorem 2.1. Let $X_1, X_2, \ldots$ be a sequence of independent random variables, identically distributed according to a positive density $f(x-\theta)$ such that f is symmetric and has logarithmic measure of tails $q(a)$. Let $T_n = \sum_{i=1}^{n} c_i \, X_n^{(i)}$ be an L-estimator of θ based on $X_1,\ldots,X_n$. Put $c_0 = c_{n+1} = 0$ and assume that $c_i = c_{n-i+1} = 0$ for $i=0,1,\ldots,k$ where $0 \leq k < \frac{n}{2}$. Then, it holds for any $\theta \in R^1$

$$\begin{aligned} k+1 &\leq \varliminf_{a\to\infty} \left\{ - \frac{\ln P_\theta(|T_n-\theta|>a)}{q(a)} \right\} \leq \\ &\leq \varlimsup_{a\to\infty} \left\{ - \frac{\ln P_\theta(|T_n-\theta|>a)}{q(a)} \right\} \leq n-k. \end{aligned} \tag{2.4}$$

Proof. Noting that T_n is translation equivariant, we may put $\theta = 0$ without loss of generality. We have

$$\begin{aligned} P_0(|T_n|>a) &= P_0(T_n > a) + P_0(T_n < -a) \leq \\ &\leq P_0(X_n^{(n-k)} > a) + P_0(X_n^{(k+1)} < -a) = \end{aligned}$$

$$= 2n \binom{n-1}{k} \int_0^{1-F(a)} t^k (1-t)^{n-k-1}\, dt \le 2^{-k} \binom{n}{k+1} [2(1-F(a))]^{k+1}$$

(F is the distribution function corresponding to f), so that

$$\underline{\lim}_{a\to\infty} \left\{ - \frac{\ell n\, P_0(|T_n| > a)}{q(a)} \right\} \ge$$

$$\ge \underline{\lim}_{a\to\infty} \left\{ - \frac{\ell n [2^{-k} \binom{n}{k+1}]}{q(a)} - \frac{(k+1)\,\ell n [2(1-F(a))]}{q(a)} \right\} = k+1.$$

Considering the third inequality in (2.4), we have

$$P_0(|T_n| > a) \ge P_0(X_n^{(k+1)} > a) + P_0(X_n^{(n-k)} < -a) =$$

$$= 2n \binom{n-1}{k} \int_0^{1-F(a)} t^{n-k-1} (1-t)^k\, dt \ge$$

$$\ge 2^{-n+k+1} \binom{n}{k} (F(a))^k [2(1-F(a))]^{n-k}$$

so that

$$\overline{\lim}_{a\to\infty} \left\{ - \frac{n\, P_0(|T_n| > a)}{q(a)} \right\} \le$$

$$\le \overline{\lim}_{a\to\infty} \left\{ - \frac{\ell n [2^{-n+k+1} \binom{n}{k}]}{q(a)} - \frac{k\, \ell n\, F(a)}{q(a)} - \frac{(n-k)\,\ell n [2(1-F(a))]}{q(a)} \right\} =$$

$$= n-k.$$

<u>Corollary 2.1</u>. Let $\tilde{X}_n$ be the sample median based on $X_1, \ldots, X_n$. Then, under the assumptions of Theorem 2.1,

$$(2.5) \qquad \frac{n}{2} \le \underline{\lim}_{a\to\infty} \left\{ - \frac{\ell n\, P_\theta(|\tilde{X}_n - \theta| > a)}{q(a)} \right\} \le$$

$$\le \overline{\lim}_{a\to\infty} \left\{ - \frac{\ell n\, P_\theta(|\tilde{X}_n - \theta| > a)}{q(a)} \right\} \le \frac{n}{2} + 1 \text{ for n even,}$$

and

$$(2.6) \qquad \lim_{a\to\infty} \left\{ - \frac{\ell n\, P_\theta(|\tilde{X}_n - \theta| > a)}{q(a)} \right\} = \frac{n+1}{2} \text{ for n odd.}$$

3. Comparison of the sample mean with other L-estimators. Let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean based on $X_1,\ldots,X_n$. It follows from Theorem 2.1 that

$$
(3.1)\qquad
\begin{aligned}
1 \leq \underline{\lim}_{a\to\infty} \left\{ -\frac{\ell n\, P_\theta(|\overline{X}_n-\theta|>a)}{q(a)} \right\} \leq \\
\leq \overline{\lim}_{a\to\infty} \left\{ -\frac{\ell n\, P_\theta(|\overline{X}_n-\theta|>a)}{q(a)} \right\} \leq n.
\end{aligned}
$$

We shall show that the upper bound in (3.1) is attained for the distributions with $q(a) = b\, a^r$, $b>0$, $r\geq 1$, for any positive integer n. It further implies that, for such distributions, $\overline{X}_n$ is more efficient than any trimmed L-estimator.

Theorem 3.1. Let $X_1,\ldots,X_n$ be a random sample from a distribution with positive density $f(x-\theta)$ such that $f(-x)=f(x)$, $x\in R^1$ and $q(a)=b\, a^r$, $b>0$, $r\geq 1$. Then

$$
(3.2)\qquad \lim_{a\to\infty} \left\{ -\frac{\ell n\, P_\theta(|\overline{X}_n-\theta|>a)}{q(a)} \right\} = n.
$$

holds for $\theta\in R^1$ and for any positive integer n.

The theorem will be proved with the aid of the following lemma.

Lemma 3.1. Let $X_1,\ldots,X_n$ be a sample from a distribution with the density $f(x-\theta)$ where f is positive and symmetric and has logarithmic measure of tails $q(a)$. Let T_n be any translation equivariant estimator of θ and d_n any positive number such that

$$
(3.3)\qquad E_0[\exp\{d_n\, q(|T_n|)\}] < \infty .
$$

Then

(3.4) $$\varliminf_{a\to\infty}\left\{-\frac{\ln P_\theta(|T_n-\theta|>a)}{q(a)}\right\}\geq d_n$$

holds for $\theta\in R^1$.

Proof of Lemma 3.1. Again, we may put $\theta = 0$. Markov's inequality implies

(3.5) $$P_0(|T_n|>a)\leq\frac{E_0[\exp\{d_n\, q(|T_n|)\}]}{\exp\{d_n\, q(a)\}}$$

so that

$$\varliminf_{a\to\infty}\left\{-\frac{\ln P_0(|T_n|>a)}{q(a)}\right\}\geq\varliminf_{a\to\infty}\left\{-\frac{\ln E_0[\exp\{d_n\, q(|T_n|)\}]}{q(a)}+\right.$$
$$\left.+\,d_n\right\}=d_n.$$

Proof of Theorem 3.1. Let $q(a) = b\,a^r$, $b>0$, $r\geq 1$. Then to any ε, $0<\varepsilon<1$, there exists $K_\varepsilon>0$ such that, for all $a\geq K_\varepsilon$

(3.6) $$1-\exp\{-(1-\tfrac{\varepsilon}{2})ba^r\}\leq 2F(a)-1\leq 1-\exp\{-(1+\tfrac{\varepsilon}{2})ba^r\}.$$

Let L_ε be the largest number of the interval $[0,K_\varepsilon]$ satisfying

(3.7) $$2\,F(L_\varepsilon)-1 = 1-\exp\{-(1-\tfrac{\varepsilon}{2})b\,L_\varepsilon^r\}.$$

We could easily see that such number always exists. Consider the distribution function

(3.8) $$G_\varepsilon(x)=\begin{cases}0 & \dots\quad x\leq 0\\ 2\,F(x)-1 & \dots\quad 0<x\leq L_\varepsilon\\ 1-\exp\{-(1-\tfrac{\varepsilon}{2})b\,x^r\} & \dots\quad L_\varepsilon<x\end{cases}$$

Then (3.6) and (3.7) imply that $G_\varepsilon(x)$ is continuous and that

(3.9) $$2\,F(x)-1\geq G_\varepsilon(x)\ \text{for}\ x>0.$$

Put $d_n = (1-\varepsilon)n$. Then, using Hölder's inequality, we get

$$(3.10)\qquad E_0[\exp\{d_n\, q(|\bar{X}_n|)\}] \leq E_0[\exp\{(1-\varepsilon)b \sum_{i=1}^{n} |X_i|^r\}] = $$
$$= (E_0 \exp\{(1-\varepsilon)b\, |X_1|^r\})^n$$

and we get from (3.6) and (3.9) by integration by parts that

$$E_0[\exp\{(1-\varepsilon)b|X_1|^r\}] \leq \int_0^\infty \exp\{(1-\varepsilon)b\, x^r\} dG_\varepsilon(x) =$$
$$(3.11)\qquad = 2\int_0^{L_\varepsilon} \exp\{(1-\varepsilon)b\, x^r\} dF(x) +$$
$$+ b\, r(1-\tfrac{\varepsilon}{2}) \int_{L_\varepsilon}^\infty \exp\{-\tfrac{\varepsilon}{2} b\, x^r\} x^{r-1}\, dx < \infty .$$

Thus, it follows from Lemma 3.1 that

$$(3.12)\qquad \varliminf_{a\to\infty} \left\{- \frac{\ln P_\Theta(|\bar{X}_n| > a)}{q(a)}\right\} \geq (1-\varepsilon)n$$

holds for any ε, $0 < \varepsilon < 1$; this implies (3.2).

Remark 1. The set of probability distributions with the tails of the form $q(a) = b\, a^r$; $b > 0$, $r \geq 1$, contains among others normal, logistic and double exponential distributions.

Corollary 3.1. Let $X_1,\ldots,X_n$ be a sample from any distribution with a positive density $f(x-\Theta)$ such that f is symmetric and has logarithmic measure of tails $q(a) = b\, a^r$; $b > 0$, $r \geq 1$. Then, for any positive integer n, the sample mean $\bar{X}_n$ is more efficient than any L-estimator which puts zero weights on $X_n^{(1)}$ and $X_n^{(n)}$.

Theorem 3.2. Let $X_1, X_2, \ldots$ be a sequence of independent random variables, identically distributed according to a density $f(x-\Theta)$ such that $f(-x) = f(x) > 0$, $x \in R^1$ and which has logarithmic measure of tails $q(a) = b\, a^r$; $b>0$, $0<r<1$. Let $T_n^{(k)}$ be an L-estimator, $T_n^{(k)} = \sum_{i=1}^{n} c_i\, X_n^{(i)}$ with $c_i = c_{n-i+1} = 0$ for $i=0,1,\dots,k$, where $c_0 = c_{n+1} = 0$ and $1 \leq k < \frac{n}{2}$. Then $\overline{X}_n$ is more efficient than $T_n^{(k)}$ if it holds

$$n(1-n^{r-1}) < k < \frac{n}{2}. \tag{3.12}$$

<u>Proof</u>. Let $q(a) = b\, a^r$, $b>0$, $0<r<1$. Then, to any $\varepsilon > 0$, there exists $K'_\varepsilon > 0$ such that (3.6) holds for all $a \geq K'_\varepsilon$. Let L'_ε be the largest number of $[0, K'_\varepsilon]$ satisfying (3.7). Then

$$2\, F(x) - 1 \geq G_\varepsilon(x) \tag{3.13}$$

holds for $x>0$ where $G_\varepsilon(x)$ is defined by (3.8) with K_ε, L_ε replaced by K'_ε, L'_ε, respectively. Put $d_n = (1-\varepsilon)n^r$. Then, we get using c_r-inequality (see Loève [7]) that

$$\begin{aligned} E_0[\exp\{d_n\, q(|\overline{X}_n|)\}] &\leq E_0[\exp\{(1-\varepsilon)b \sum_{i=1}^{n} |X_i|^r\}] = \\ &= (E_0[\exp\{(1-\varepsilon)b\,|X_1|^r\}])^n \end{aligned} \tag{3.14}$$

and it follows from (3.13)

$$E_0[\exp\{(1-\varepsilon)b\,|X_1|^r\}] \leq \int_0^\infty \exp\{(1-\varepsilon)b\, x^r\} dG_\varepsilon(x) =$$

$$= 2\int_0^{L'_\varepsilon} \exp\{(1-\varepsilon)b\, x^r\} dF(x) + (1-\tfrac{\varepsilon}{2})br \int_{L'_\varepsilon}^\infty \exp\{-\tfrac{\varepsilon}{2}\, b\, x^r\}\, x^{r-1}\, dx < \infty$$

and we get by Lemma 3.1 that

$$\varliminf_{a\to\infty} \left\{ - \frac{\ell n\, P_\theta(|\overline{X}_n - \theta| > a)}{q(a)} \right\} \geq n^r \text{ for } \theta \in R^1 \tag{3.15}$$

and this together with Theorem 2.1 implies the proposition of the theorem.

<u>Remark 2</u>. The method of Theorems 2.1 and 3.1 does not

provide a comparison of $\bar{X}_n$ and other L-estimators based on the sample of size n from a distribution with the tails of the type $q(a) = m \cdot \ln a$, $m>1$, m integer (for instance, it could be easily found that t-distribution with m degrees of freedom belongs to this type). On the other hand, if $X_1,\ldots$ $\ldots,X_n$ is a sample from Cauchy distribution which corresponds to $q(a) = \ln a$, then $\bar{X}_n$ is also distributed according to the same Cauchy distribution, so that

$$\lim_{a\to\infty}\left\{-\frac{\ln P_\theta(|\bar{X}_n-\theta|>a)}{q(a)}\right\} = 1$$

holds for $\theta \in R^1$ in this case and Theorem 2.1 implies that any L-estimator which puts zero weights on $X_n^{(1)}$ and $X_n^{(n)}$ is more efficient than $\bar{X}_n$.

References

[1] R.R. BAHADUR: Rates of convergence of estimates and test statistics, Ann. Math. Statist. 38(1967), 303-324.

[2] R.R. BAHADUR: Some Limit Theorems in Statistics, SIAM, Philadelphia (1971).

[3] J.C. FU: The rate of convergence of point estimators, Ann. Statist. 3(1975), 234-240.

[4] P.J. HUBER: Robust confidence limits, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 10(1968), 269-278.

[5] J. JUREČKOVÁ: Locally optimal estimates of location, Comment. Math. Univ. Carolinae 18(1977), 599-610.

[6] E.L. LEHMANN: Testing Statistical Hypotheses, J. Wiley (1959).

[7] M. LOÈVE: Probability Theory, Moscow (1962)(Russian translation).

[8] G.L. SIEVERS: Estimation of location: A large deviation comparison, Ann. Statist. 6(1978), 610-618.

Matematicko-fyzikální fakulta
Universita Karlova
Sokolovská 83, 18600 Praha 8
Československo

(Oblatum 21.2. 1979)