notas
de
matemática

LOCATION ESTIMATORS BASED ON LINEAR COMBINATIONS OF MODIFIED ORDER STATISTICS

Víctor J. Yohai   Ricardo A. Maronna

N° 31   1975

UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE CIENCIAS EXACTAS - DEPARTAMENTO DE MATEMATICA

LOCATION ESTIMATORS BASED ON LINEAR COMBINATIONS OF MODIFIED ORDER STATISTICS

Víctor J. Yohai Ricardo A. Maronna

Edición previa

UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE CIENCIAS EXACTAS
Departamento de Matemática

La Plata, República Argentina 1975

1 - Introduction. In this paper we introduce a class of location estimators which seemingly has not received much attention up to now. Let $X_1, \ldots, X_n$ be i.i.d. random variables with common distribution function $F(x - \theta)$, where $F(x)$ has a symmetric density $f(x)$. Consider a non-negative and nonincreasing function $a(t)$ defined in $[0,1]$, such that $\int_0^1 a(u)\, du = 1$. For each $t \in R$ we define the t-order statistics as: $X_{in}(t) = X_k$ if there exist exactly $(i - 1)$ X_j's such that $|X_j - t| < |X_k - t|$; and we define the new variable

$$T_n(t) = n^{-1} \sum_{i=1}^{n} a(i/n)\, X_{in}(t).$$

Suppose we have an initial estimator $\tilde{\theta}$ of θ. It seems reasonable to measure the reliability of each observation X_i according to its distance to $\tilde{\theta}$; thus we may construct a new estimator of θ as a weighted average, where the observations close to $\tilde{\theta}$ receive larger weights, i.e., as $T_n(\tilde{\theta})$.

A natural possibility is to look for the estimator $\hat{\theta}$ which is invariant for the tranformation $T_n(.)$, i.e., which satisfies

$$(1.1) \qquad T_n(\hat{\theta}) = \hat{\theta}.$$

We shall call the estimator defined by (1.1) a "fixed-point" estimate, to distinguish it from $T_n(\tilde{\theta})$ which we shall call a "one-step" estimate.

It can be verified that the value of θ which minimizes $\sum_{i=1}^{n} (X_{in}(\theta) - \theta)^2 \, a(i/n)$ satisfies (1.1).

The case in which $a(u) = (1 - 2\alpha)^{-1} \, I(u \leq 1 - 2\alpha)$ may be considered as a kind of "trimming" around the estimator itself. Gnanadesikan and Kettenring [2] considered a similar proposal in the context of multivariate data analysis. More recently Shorack [8] studied a class of estimators which contains the one-step version of these trimming estimates.

In section 2 we shall prove the asymptotic normality of the estimates defined by (1.1), we shall calculate their asymptotic variances and we shall show that they are equivalent to a suitable M-estimator (Huber [4]) defined as a solution θ of the equation $\sum_{i=1}^{n} \psi(X_i - \theta) = 0$, where

(1.2) $$\psi(x) = x \, a(F(|x|) - F(-|x|)).$$

Similar results are obtained for the one-step estimates, in section 3.

In section 4 we obtain by Monte Carlo methods the variances of some fixed-point estimators and of their one-step versions, for the Normal and Cauchy distributions. The resuts seem to indicate that it is not possible to find robust estimates within this class

2 - Asymptotic distributions. For notational convenience, define for all $t \in R$

(2.1) $$L_n(t) = n^{1/2}(T_n(t) - t)$$

$$(2.2)\qquad M_n(t) = n^{-1/2} \sum_{i=1}^{n} \psi(X_i - t)$$

where ψ is given by (1.2). We may write (1.1) as

$$(2.3)\qquad L_n(\hat{\theta}) = 0.$$

We shall need the following assumptions:

A1) The distribution function F has a symmetric density f.

A2) $I(f) = \int_{-\infty}^{\infty} (f'(x)/f(x))^2 f(x)\,dx < \infty$.

B1) $\int_0^1 a(u)\,du = 1$, $a(u) \geq 0$.

B2) $a(u)$ is monotone nonincreasing

B3) There exists $\alpha \in (0,1)$ such that $a(u) = 0$ for $1 - \alpha < u \leq 1$.

B4) The function a has a finite number of discontinuities.

Let ψ be defined by (1.2). The following "assumption" is in fact a consequence of B2, B3 and B4:

C1) $\psi = \psi^+ - \psi^-$, where ψ^+ and ψ^- are monotone nondecreasing and

$$\int_{-\infty}^{+\infty} [\psi^{\pm}(x+h) - \psi^{\pm}(x-h)]^2\, dF(x) = o(1) \text{ for } h \to 0.$$

We shall also need:

C2) There exist coefficients $A(\psi^{\pm}, F)$ such that

$$\int_{-\infty}^{+\infty} (\psi^{\pm}(x-h) - \psi^{\pm}(x))dF(x) = h\, A(\psi^{\pm},F) + o(h) \quad (h \to 0).$$

We define $A(\psi, F) = A(\psi^{+}, F) - A(\psi^{-}, F)$.

To study the asymptotic behavior of the solution $\hat{\theta}$ of (1.1) we shall need the following linear expansion of $L_n(t)$ and $T_n(t)$:

THEOREM 2.1: Assume conditions A1, A2, B1, B2, B3, B4, C1 and C2. Then for any $K > 0$ and $\theta = 0$ we have

(2.4) $\quad p\lim_{n\to\infty} \sup_{|t|\le K} |L_n(n^{-1/2}t) - L_n(0) - t\,A(\psi,F)| = 0.$

(2.5) $\quad p\lim_{n\to\infty} \sup_{|t|\le K} |L_n(n^{-1/2}t) - M_n(0) - t\,A(\psi,F)| = 0.$

We start by proving some auxiliary lemmas.

LEMMA 2.2: The sequence $R_n(t) = L_n(t) - M_n(t)$ tends to zero in probability for each $t \in R$.

PROOF: There is no loss of generality in taking $\theta = 0$, so that the X_i's have a symmetric distribution. To simplify notation put $F_+(x) = F(|x|) - F(-|x|)$, $U_{(i)} = F_+(|X_{in}(0)|)$ (which are the order statistics of a uniform distribution), $S_i = \mathrm{sgn}\, X_{in}(0)$ and call G the inverse function of F_+. Then

(2.6) $\quad R_n(0) = n^{-1/2} \sum_{i=1}^{n} G(U_{(i)})\,[a(i/n) - a(U_{(i)})]\, S_i.$

Put $G_1(u) = G(u)\, I\{u < 1 - \alpha/2\}$ and $R_n = B_n + C_n$, where B_n is the expression (2.6) with G replaced by G_1;

We shall prove that B_n and C_n tend to zero in probability.

The symmetry of the X_i's implies that their signs are mutually independent, and independent of their absolute values, and hence of the $U_{(i)}$'s, so that

$$E(B_n^2) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} E\{G_1(U_{(i)})\, G_1(U_{(j)})[a(i/n) - a(U_{(i)})]$$

$$[a(j/n) - a(U_{(j)})]\}\, E(S_i\, S_j) =$$

$$= n^{-1} \sum_{i=1}^{n} E\{[G_1(U_{(i)})]^2\, [a(i/n) - a(U_{(i)})]^2\}$$

By B4, given ε, there exists a set $D \subseteq (0,1)$ which is a finite union of open intervals, each one containing a discontinuity point of a, and having total Lebesgue measure $< \varepsilon$. Then

$$E(B_n^2) \leq [G(1-\alpha/2)]^2\, \{\varepsilon.4 \sup a^2 + n^{-1} \sum_{i/n \notin D} E[a(i/n)-a(U_{(i)})]^2\}$$

and since a is uniformly continuous outside D, the second term in brackets tends to zero by the Glivenko-Cantelli Theorem.

At the same time, by B3:

$$|C_n| \leq n^{-1/2} \sum_{i=1}^{n} G(U_{(i)})\, a(i/n)\, I\{U_{(i)} > 1-\alpha/2;\ i/n < 1-\alpha\}.$$

Putting $c = \sup\{G(u + \alpha/2)\, a(u) | u < 1-\alpha/2\} < \infty$, we have

$$P\{\sup_{1\le i\le n} G(U_{(i)})\, a(i/n) > c\} \le P\{\sup_{1\le i\le n} |U_{(i)} - i/n| > \alpha/2\}$$

which tends to zero by the Glivenko-Cantelli Theorem. Besides

$$E\{n^{-1/2} \sum_{i=1}^{n} I[|U_{(i)} - i/n| > \alpha/2]\}$$

$$\le 4\, \alpha^{-2}\, n^{-1/2} \sum_{i=1}^{n} E(|U_{(i)} - i/n|^2) = 0(n^{-1/2}).$$

LEMMA 2.3: For any t we have

$$\operatorname*{p\,lim}_{n\to\infty} [L_n(n^{-1/2}\, t) - M_n(n^{-1/2}\, t)] = 0$$

PROOF: The distribution of $R_n\,(n^{-1/2}\, t)$ when $\theta = 0$, is the same as that of $R_n(0)$ when $\theta = -\, n^{-1/2}\, t$. But since for any $t, \theta_n = -n^{-1/2}t$ is a sequence of alternatives contiguous to $\theta = 0$ (see Chapter 6 of Hájek [3]) we obtain, applying Lemma 2.2, that

$$\operatorname*{p\,lim}_{n\to\infty} R_n(0) = 0 \text{ when } \theta_n = -n^{-1/2}t.$$

The following result was proved by Bickel [1]:

LEMMA 2.4: <u>If ψ is any function verifying C1 and C2, then for any</u> $K > 0$

$$p \lim_{n\to\infty} \sup_{|t|\leq K} [M_n(n^{-1/2}t) - M_n(0) - t\,A(\psi,F)] = 0.$$

PROOF of THEOREM 2.1: Let $H_n(t) = n^{1/2}\,T_n(t)$, so that

$$L_n(t) = H_n(t) - n^{-1/2}\,t \sum_{i=1}^{n} a(i/n)$$

and hence

$$(2.7) \quad L_n(n^{-1/2}\,t) = H_n(n^{-1/2}\,t) - t\,(1 + u_n)$$

where u_n is a numerical sequence such that

$$(2.8) \quad \lim_{n\to\infty} u_n = 0.$$

Then from Lemmas 2.3 and 2.4 we obtain

$$(2.9) \quad p \lim_{n\to\infty}{}^{*} [H_n(n^{-1/2}t) - H_n(0) - t\,A^{\circ}(\psi,F)] = 0$$

where

$$(2.10) \quad A^{\circ}(\psi,F) = A(\psi,F) + 1.$$

Now B2 implies that $H_n(n^{-1/2}t)$ is a monotone nondecreasing function of t, so that arguing as in the proof of Theorem 2.1 in Jureckova [7] we can show that the convergence in (2.9) is uniform for bounded t, so that for each $K > 0$ we have

$$(2.11) \qquad p \lim_{n\to\infty} \sup_{|t|\le K} |H_n(n^{-1/2}t) - H_n(0) - t A^{*}(\psi, F)| = 0$$

and using (2.7), (2.8) and (2.10) we obtain

$$(2.12) \qquad p \lim_{n\to\infty} \sup_{|t|\le K} |L_n(n^{-1/2}t) - L_n(0) - t A(\psi, F)| = 0$$

which proves (2.4); (2.5) follows from this last formula and Lemma 2.2.

Now we are ready to prove the asymptotic normality of solutions of (1.1).

THEOREM 2.5: <u>Suppose that</u> $\hat{\theta}_n$ <u>is a sequence of solutions of (1.1) such that</u> $n^{1/2}(\hat{\theta}_n - \theta)$ <u>is bounded in probability. Then</u> $n^{1/2}(\hat{\theta}_n - \theta)$ <u>has an asymptotic normal distribution with mean 0 and variance</u> $B(\psi,F)/A(\psi,F)^2$, <u>where</u>

$$(2.13) \qquad B(\psi, F) = \int_{-\infty}^{+\infty} \psi(x)^2 \, dF(x).$$

PROOF: Since the solutions of (1.1) are translation invariant, we can assume that $\theta = 0$. Then (2.5) implies

$$(2.14) \qquad L_n(n^{-1/2}t) = M_n(0) + A(\psi,F)\, t + h_n(t)$$

where for each $K > 0$

$$(2.15) \qquad \operatorname{p\,lim}_{n\to\infty} \sup_{|t|\leq K} |h_n(t)| = 0$$

Then since $L_n(\hat{\theta}_n) = 0$ we have

$$n^{1/2}\,\hat{\theta}_n = -M_n(0)/A(\psi,F) - h_n(n^{1/2}\,\hat{\theta}_n)$$

Now, since $n^{1/2}\,\hat{\theta}_n$ is bounded in probability, (2.15) implies that $\operatorname{p\,lim}_{n\to\infty} h_n(n^{1/2}\,\hat{\theta}_n) = 0$, and hence

$$(2.16) \qquad \operatorname{p\,lim}_{n\to\infty} [n^{1/2}\,\hat{\theta}_n - M_n(0)/A(\psi,F)] = 0.$$

We know by the Central Limit Theorem that the distribution of $M_n(0)$ converges to a normal law with mean 0 and variance $B(\psi,F)$; hence from (2.16) we obtain the desired result.

REMARK: According to the former Theorem, a sequence $\hat{\theta}_n$ of solutions of (1.1) such that $n^{1/2}(\hat{\theta}_n - \theta)$ is bounded in probability, has the same limit distribution as the M-estimator corresponding to the psi-function (1.2). Let us point out that (1.1) does not in general possess a unique solution. We have not succeded in proving that any sequence of solutions $\hat{\theta}_n$ verifies the required boundedness condition. However, we shall show that it is possible to choose a sequence of solutions satisfying the boundedness condition.

THEOREM 2.6: _Assume that_ $A(\psi,F) \neq 0$ _and that_ $\tilde{\theta}_n$ _is a sequence of estimators such that_ $n^{1/2}(\tilde{\theta}_n - \theta)$ _is bounded in probability. Let_ $\hat{\theta}_n$ _be the solution of_ (1.1) _nearest to_ $\tilde{\theta}_n$. _Then_ $n^{1/2}(\hat{\theta}_n - \theta)$ _is bounded in probability_.

PROOF: We can again assume $\theta = 0$. Take any $\varepsilon > 0$. Since $M_n(0)$ converges in law to a normal distribution, we can find K such that

$$P\{|M_n(0)| \leq K\} \geq 1 - \varepsilon/3 \text{ for all } n.$$

Let $h_n(t)$ be defined as in (2.14); then according to (2.15) there exists n_0 such that

$$P\,(\sup_{|t| \leq K_1} |h_n(t)| \leq 1) \geq 1 - \varepsilon/3 \text{ for all } n \geq n_0$$

where $K_1 = (2K+1)/A(\psi,F)$.

Put

$$A_n = \{|M_n(0)| \leq K\} \cap \{\sup_{|t| \leq K_1} |h_n(t)| \leq 1\}.$$

Then

$$(2.17) \quad P(A_n) \geq 1 - (2/3)\varepsilon \qquad \text{for } n \geq n_0.$$

But using (2.14) we have in A_n

$$(2.18) \quad L_n(-n^{-1/2} K_1) \geq -K + 2K + 1 - 1 \geq K > 0$$

and similarly

$$(2.19) \quad L_n(n^{-1/2} K_1) \leq -K < 0.$$

Then since the function $L_n(t)$ is piecewise continuous and jumps only upwards, we obtain from (2.18) and (2.19)

$$(2.20) \quad A_n \subseteq \{\text{there exist } \hat{t}_n \text{ such that } L_n(\hat{t}_n) = 0 \text{ and } |n^{1/2}\hat{t}_n| \leq K_1\}$$

We can also find K_2 such that

$$(2.21) \quad P(|n^{1/2} \tilde{\theta}_n| \leq K_2) \geq 1 - \varepsilon/3.$$

Put

$$(2.22) \quad B_n = A_n \cap \{|n^{1/2} \tilde{\theta}_n| \leq K_2\}.$$

Then from (2.22) and (2.17) we have

(2.23) $P(B_n) \geq 1 - \varepsilon$ for $n \geq n_0$

and from (2.20) and (2.22) we have

$$B_n \subseteq \{|n^{1/2}\hat{\theta}_n| \leq K_1 + K_2\}$$

so that from (2.23) we get

$$P(|n^{1/2}\hat{\theta}_n| \leq K_1 + K_2) \geq 1 - \varepsilon \quad \text{for } n \geq n_0.$$

3 - "One-step" estimates. In this section we consider the asymptotic behavior of $T_n(\tilde{\theta})$ when $\tilde{\theta}$ is an M-estimate.

THEOREM 3.1: Let $\tilde{\theta}_n$ be an M-estimator corresponding to a function ψ_0 satisfying C1 and C2, and assume that $n^{1/2}(\tilde{\theta}_n - \theta)$ converges in distribution to a normal law (see Huber [4]). Assume also the same conditions as in Theorem 2.1. Then $n^{1/2}(T_n(\tilde{\theta}_n) - \theta)$ converges in distribution to a normal law with mean 0 and variance given by

$$(3.1)\quad B(\psi,F) + (B(\psi_0,F)\ [(A(\psi,F)+1)/A(\psi_0,F)]^2 -$$

$$- 2\ B(\psi,\psi_0,F)\ (A(\psi,F)+1)/A(\psi_0,F)$$

<u>where</u> $B(\psi,F)$ <u>is defined in</u> (2.13) <u>and</u>

$$(3.2)\quad B(\psi,\psi_0,F) = \int_{-\infty}^{\infty} \psi(x)\psi_0(x)\ dF(x).$$

<u>PROOF</u>: Assume again that $\theta = 0$. Using (2.9), (2.10), Lemma 2.2 and the assumption that $n^{1/2}(\tilde{\theta}_n - \theta)$ converges in distribution to a normal law, we obtain, recalling the definition of H_n at the beginning of the proof of Theorem 2.1

$$(3.3)\quad \operatorname*{p\,lim}_{n\to\infty} [n^{1/2}\ T_n(\tilde{\theta}_n) - M_n(0) - n^{1/2}\tilde{\theta}_n(A(\psi,F)+1)] = 0.$$

Put

$$M_{n0}(t) = n^{-1/2} \sum_{i=1}^{n} \psi_0(X_i - t).$$

Since ψ_0 satisfies C1 and C2, and since $n^{1/2}\ \tilde{\theta}_n$ is bounded in probability, we obtain from Lemma 2.3

$$\operatorname*{p\,lim}_{n\to\infty} [M_{n0}(\tilde{\theta}_n) - M_{n0}(0) - n^{1/2}\tilde{\theta}_n\ A(\psi_0,F)] = 0.$$

Then, since $M_{n0}(\tilde{\theta}_n) = 0$ we have

(3.4) $$\underset{n\to\infty}{p\lim}\,[n^{1/2}\tilde{\theta}_n - M_{n0}(0)/A(\psi_0,F)] = 0.$$

Hence, from (3.3) and (3.4) we obtain

(3.5) $$\underset{n\to\infty}{p\lim}\,[n^{1/2}\, T_n(\tilde{\theta}_n) - M_n(0) - M_{n0}(0)(A(\psi,F)+1)/A(\psi_0,F)] = 0.$$

But the joint asymptotic distribution of $(M_n(0), M_{n0}(0))$ is bivariate normal with zero means and covariance matrix

$$\begin{pmatrix} B(\psi,F) & B(\psi,\psi_0,F) \\ B(\psi,\psi_0,F) & B(\psi_0,F) \end{pmatrix}$$

Hence $M_n(0) + M_{n0}(0)\,(A(\psi,F)+1)/A(\psi_0,F)$ converges in distribution to a normal law with mean 0 and variance given by (3.1). Then by (3.5) the theorem is proved.

__Remark__: If $\bar{\theta}_n$ is an L-estimator, i.e. $\bar{\theta}_n = n^{-1} \sum\ h(i/n)\, X_{(i)}$ -where the $X_{(i)}$'s are the sample order statistics- then under general conditions $\bar{\theta}_n$ is asymptotically equivalent to the M-estimator $\tilde{\theta}_n$ corresponding to the function ψ_0 given by $\psi_0'(x) = h(F(x))$ (see Jaeckel [6]). Then it is easy to prove that under the same conditions as in Theorem 3.1, $n^{1/2}\ (T_n(\bar{\theta}_n) - \theta)$ has the same asymptotic distribution as $n^{1/2}(T_n(\tilde{\theta}_n) - \theta)$.

4 - Two particular families. We have computed the asymptotic variances and studied by Monte Carlo methods the small-sample behavior for two particular classes of a-functions. We define for $0 \leq \alpha \leq \beta \leq 1$

$$a_{\alpha\beta}(u) = \begin{cases} 2/(\alpha+\beta) & \text{if } u \leq \alpha \\ (2/(\alpha+\beta))((\beta-u)/(\beta-\alpha)) & \text{if } \alpha \leq u \leq \beta \\ 0 & \text{if } \beta \leq u \leq 1 \end{cases}$$

This family contains "trimming" as the special case $\alpha = \beta$.

Let $\psi_{ABC}(x)$ be the psi-function corresponding to Hampel's M-estimator defined in (Huber (5), page 1064). Let $a_{ABC}(u)$ be the a-function which for the distribution N(0,1) is equivalent to the psi-function $\psi(x) = \psi_{ABC}(x/q)$, where q is the 0.75-point of N(0,1).

For some members of these families (both "fixed-point" and "one-step") we computed the variances σ_n^2 for samples of size n = 20 and n = 40 for the Normal and Cauchy distributions, by means of a Monte Carlo simulation.

The computation of the fixed-point estimate for finite samples was performed by taking the sample median as an initial estimate, and then iteratively applying the function T_n defined in section 1. If this procedure converges, it must do so in a finite number of steps; in practice, we have always observed it to converge rather quickly.

As it can be seen from these results, it does not seem possible to find estimates of this class which are robust in the sense that they are efficient for both the normal and log-tailed distributions.

TABLE 1

VARIANCES OF $n^{1/2}\hat{\theta}_n$

	$n = \infty$		$n = 20$		$n = 40$	
	N(0,1)	Cauchy	N(0,1)	Cauchy	N(0,1)	Cauchy
$a_{\alpha\beta}$						
$\alpha = 0.80$, $\beta = 0.95$						
Fixed-point	2.856	3.205	1.60	5.51	1.79	3.74
One-step			1.51	5.11	1.57	3.44
$\alpha = 0.90$, $\beta = 0.95$						
Fixed-point	1.506	5.562	1.26	>13	1.35	8.36
One-step			1.26	>13	1.33	7.99
a_{ABC}						
A = 2.5, B = 4.5, C = 9.5						
Fixed-point						
One-step	1.025		1.00	>50		
			1.00	>50		
A = 1.2, B = 3.5, C = 8.0						
Fixed-point	1.166		1.11	>20		
One-step			1.15	>20		

REFERENCES

[1] BICKEL, P. (1971). A note on approximate (M) estimates in the linear model. To appear in J. Amer. Stat. Assoc.

[2] GNANADESIKAN, R. and KETTENRING, J.R. (1972). Robust estimates. residuals, and outlier rejection with multivariate data. Biometrics 28, 81-124.

[3] HAJEK, J. and SIDAK, Z. (1967). Theory of Rank Tests. Academic Press, New York.

[4] HUBER, P. (1964) Robust estimation of a location parameter. Ann. Math. Statist. 35, 73-101.

[5] HUBER, P. (1972) Robust statistics. Ann. Math. Statist. 43, 1041-1067

[6] JAECKEL, L.A. (1971). Some flexible estimates of location Ann. Math. Statist. 42, 1020-1034

[7] JURECKOVA, J. (1969) Aymptotic linearity of a rank statistic in regression parameter. Ann. Math. Statist. 40, 1889-1900

[8] SHORACK, G. (1974). Random means. Ann. Math. Statist 2, 661-675.