The Annals of Statistics
1977, Vol. 5, No. 3, 464–472

ASYMPTOTIC RELATIONS OF M-ESTIMATES AND R-ESTIMATES IN LINEAR REGRESSION MODEL

BY JANA JUREČKOVÁ

Charles University

Let $\hat{\mathbf{\Delta}}_M$ be an M-estimator (maximum-likelihood type estimator) and $\hat{\mathbf{\Delta}}_R$ be an R-estimator (rank estimator) of the parameter $\mathbf{\Delta}=(\Delta_1, \cdots, \Delta_p)$ in the linear regression model $X_{Ni}=\sum_{j=1}^{p} \Delta_j c_{ji} + e_i, i = 1, \cdots, N$. The asymptotic distribution of $\hat{\mathbf{\Delta}}_M - \hat{\mathbf{\Delta}}_R$ is derived for p fixed and $N \to \infty$, under some assumptions on the design matrix, on the error distribution F and on the functions generating the respective estimators. The result has several consequences which have an interest of their own; among others, it is shown that to any M-estimator corresponds an R-estimator such that the estimators are asymptotically equivalent, and conversely. A special case when $\hat{\mathbf{\Delta}}_M$ is the maximum likelihood estimator and $\hat{\mathbf{\Delta}}_R$ the R-estimator, both asymptotically efficient for some distribution G, is also considered.

1. Introduction. For $N = 1, 2, \cdots$, let $X_{N1}, \cdots, X_{NN}$ be independent observations such that X_{Ni} has the cdf

$$F(x - \textstyle\sum_{j=1}^{p} \Delta_j^0 c_{ji}), \qquad i = 1, \cdots, N \tag{1.1}$$

where $\mathbf{\Delta}^0 = (\Delta_1^0, \cdots, \Delta_p^0)$ is an unknown parameter and $\mathbf{C}_N = [c_{ji}]_{j=1,\cdots,p}^{i=1,\cdots,N}$ is a given design matrix.

Let us consider the problem of estimating $\mathbf{\Delta}^0$ on the basis of $X_{N1}, \cdots, X_{NN}$ if F is not specified. Besides the classical least squares estimator, other types of estimates were suggested which are less sensitive to the outlying observations and to incorrect assumptions concerning the form of the basic distribution F (see Huber [4], Jaeckel [8], Jurečková [10], Koul [13], Kraft and van Eeden [14], Bickel [1]). In the present paper, we shall study the asymptotic relations of two robust estimates: M-estimates suggested by Huber and R-estimates suggested by Jurečková (or equivalently, R-estimates suggested respectively by Jaeckel and Koul, which are asymptotically equivalent to the estimate of Jurečková).

More precisely, the asymptotic distribution of $\hat{\mathbf{\Delta}}_M - \hat{\mathbf{\Delta}}_R$ is shown to be normal for p fixed and $N \to \infty$, under some assumptions on $\mathbf{C}_N$, on the functions generating the estimates and for the basic distribution F with finite Fisher's information. The cases in which the asymptotic distribution degenerates are of interest. We shall show that to any M-estimate (in the frame of the assumptions) corresponds an R-estimate such that the estimates are asymptotically equivalent in the sense of convergence in probability, and conversely.

An interesting special case happens if $\hat{\mathbf{\Delta}}_M$ is the maximum likelihood estimate corresponding to a distribution G and $\hat{\mathbf{\Delta}}_R$ is a rank estimate which is asymptotic-

Received June 1975; revised September 1976.
AMS 1970 *subject classifications*. Primary 62G05; Secondary 62G35.
Key words and phrases. M-estimate, R-estimate, asymptotically normal distribution.

ally efficient for G in the role of the basic distribution. If G is normal, then $\hat{\mathbf{\Delta}}_M$ and $\hat{\mathbf{\Delta}}_R$ are asymptotically equivalent if and only if F is also normal; a similar proposition holds for G logistic. Both estimates coincide if G is double exponential (the estimates then represent a generalized version of median). If $G \equiv F$, then $\hat{\mathbf{\Delta}}_M$ and $\hat{\mathbf{\Delta}}_R$ are asymptotically efficient.

2. Assumptions and notation.

1°. For $N = 1, 2, \cdots$, let $X_{N1}, \cdots, X_{NN}$ be independent random variables such that X_{Ni} has the distribution function

$$F(x - \textstyle\sum_{j=1}^{p} \Delta_j^0 c_{ji}) \,, \qquad i = 1, \cdots, N \,. \tag{2.1}$$

Also suppose that $f(x) = dF(x)/dx$ exists, is absolutely continuous and has finite Fisher's information, i.e.,

$$I(f) = \int \left(\frac{f'(x)}{f(x)}\right)^2 f(x)\, dx < \infty \,. \tag{2.2}$$

Let us denote

$$\varphi(t, f) = -f'(F^{-1}(t))/f(F^{-1}(t)) \,, \qquad 0 < t < 1 \tag{2.3}$$

where

$$F^{-1}(t) = \inf\{x : F(x) \geqq t\} \,.$$

2°. Let $\mathbf{c}'_{(j)} = (c'_{j1}, \cdots, c'_{jN})$ and $\mathbf{c}''_{(j)} = (c''_{j1}, \cdots, c''_{jN})$ be vectors satisfying

$$\text{(a)} \qquad \textstyle\sum_{i=1}^{N} c'_{ji} = \sum_{i=1}^{N} c''_{ji} = 0 \,, \qquad j = 1, \cdots, p$$

and

$$\mathbf{c}'_{(j)} \cdot (\mathbf{c}'_{(j)})^T \leqq M \,; \qquad \mathbf{c}''_{(j)} \cdot (\mathbf{c}''_{(j)})^T \leqq M \,, \qquad j = 1, \cdots, p;\ N = 1, 2, \cdots \tag{2.4}$$

($M > 0$ is a constant independent of N) where either of the scalar products in (2.4) is either 0 for all but a finite number of N or positive for all but a finite number of N; if $\mathbf{c}'_{(j)} \cdot (\mathbf{c}'_{(j)})^T > 0$ for $N > N'$, then assume

$$\lim_{N\to\infty} \{\max_{1 \leqq i \leqq N} (c'_{ji})^2 [\textstyle\sum_{k=1}^{N} (c'_{jk})^2]^{-1}\} = 0 \tag{2.5}$$

(Noether's condition), and an analogous assumption is to be satisfied for $\mathbf{c}''_{(j)}$, $j = 1, \cdots, p$.

(b) For all pairs $j, h = 1, \cdots, p$ and $i, k = 1, \cdots, N$ $(N = 2, 3, \cdots)$, assume

$$\begin{aligned} (c'_{ji} - c'_{jk})(c'_{hi} - c'_{hk}) &\geqq 0 \\ (c'_{ji} - c'_{jk})(c''_{hi} - c''_{hk}) &\leqq 0 \\ (c''_{ji} - c''_{jk})(c''_{hi} - c''_{hk}) &\geqq 0 \,. \end{aligned} \tag{2.6}$$

Let $\mathbf{C}_N = [c_{ji}]_{j=1,\cdots,p}^{i=1,\cdots,N}$ be a given design matrix with the rows $\mathbf{c}_{(j)}, j = 1, \cdots, p$ and the columns $\mathbf{c}^{(i)}$, $i = 1, \cdots, N$; suppose that, for $N = 1, 2, \cdots$, the rows $\mathbf{c}_{(j)}$ could be decomposed into a sum of vectors $\mathbf{c}'_{(j)}$ and $c''_{(j)}$ satisfying (a) and (b), i.e.,

$$c_{ji} = c'_{ji} + c''_{ji} \,, \qquad i = 1, \cdots, N,\ N = 1, 2, \cdots;\ j = 1, \cdots, p.$$

Moreover, assume that

(c) $\lim_{N\to\infty} \boldsymbol{\Sigma}_N = \boldsymbol{\Sigma}$ exists and $\boldsymbol{\Sigma}$ is a positive definite matrix, where

$$\boldsymbol{\Sigma}_N = \mathbf{C}_N \mathbf{C}_N{}^T . \tag{2.7}$$

3°. Let $\varphi(t)$, $0 < t < 1$, be a nonconstant function which is expressible as a finite sum of monotone functions, square-integrable on (0, 1).

4°. Let $\psi(x)$, $x \in R^1$ be a nonconstant function which is expressible as a finite sum of monotone functions; assume that

$$\int_{R^1} \psi^2(x) f(x) dx < \infty . \tag{2.8}$$

Let us denote

$$\gamma = \int_0^1 \varphi(t)\varphi(\mathrm{t}, f)\, dt , \tag{2.9}$$

$$\beta^2 = \int_0^1 \varphi^2(t)\, dt - \bar{\varphi}^2 , \qquad \bar{\varphi} = \int_0^1 \varphi(t)\, dt , \tag{2.10}$$

$$\omega = - \int_{R^1} \psi(x) f'(x)\, dx \tag{2.11}$$

and

$$\rho^2 = \int_{R^1} \psi^2(x) f(x)\, dx - \left(\int_{R^1} \psi(x) f(x)\, dx\right)^2 . \tag{2.12}$$

For any fixed vector $\boldsymbol{\Delta} = (\Delta_1, \cdots, \Delta_p)$, let us denote

$$\delta_i(\boldsymbol{\Delta}) = X_{Ni} - \sum_{j=1}^p \Delta_j c_{ji} , \qquad i = 1, \cdots, N. \tag{2.13}$$

Let $\hat{\boldsymbol{\Delta}}_M$ be the M-estimate of $\boldsymbol{\Delta}^0$ corresponding to the function ψ, i.e., $\hat{\boldsymbol{\Delta}}_M$ is the solution of the system of equations

$$M_j(\mathbf{X}_N, \boldsymbol{\Delta}) = \sum_{i=1}^N c_{ji} \psi(\delta_i(\boldsymbol{\Delta})) = 0 , \qquad j = 1, \cdots, p, \tag{2.14}$$

with respect to $\boldsymbol{\Delta}$.

Let $\hat{\boldsymbol{\Delta}}_R$ be the rank estimate of Hodges–Lehmann type corresponding to the function φ, suggested by Jurečková in [10]; i.e., $\hat{\boldsymbol{\Delta}}_R$ is any solution of the minimization problem

$$\sum_{j=1}^p \left| \sum_{i=1}^N c_{ji} a_N(R_i{}^\Delta) \right| = \min \tag{2.15}$$

where $R_i{}^\Delta$ is the rank of $\delta_i(\boldsymbol{\Delta})$ among $\delta_1(\boldsymbol{\Delta}), \cdots, \delta_N(\boldsymbol{\Delta})$;

$$R_i{}^\Delta = \sum_{j=1}^N u(\delta_i(\boldsymbol{\Delta}) - \delta_j(\boldsymbol{\Delta})) \tag{2.16}$$

where $u(x) = 1$ if $x \geqq 0$ and $= 0$ if $x < 0$; $a_N(\cdot)$ is the score-function corresponding to φ in the following way:

$$a_N(i) = \varphi\left(\frac{i}{N+1}\right) , \qquad i = 1, \cdots, N. \tag{2.17}$$

3. Asymptotic distribution of $\hat{\boldsymbol{\Delta}}_M - \hat{\boldsymbol{\Delta}}_R$.

THEOREM 3.1. *Under the assumptions* 1°–4° *and under* $\gamma \neq 0$, $\omega \neq 0$, $\hat{\boldsymbol{\Delta}}_M - \hat{\boldsymbol{\Delta}}_R$ *have for* $N \to \infty$ *the asymptotically normal distribution with center* $\mathbf{0}$ *and covariance matrix* $\boldsymbol{\Sigma}$ *inverse multiplied by the scalar*

$$\boldsymbol{\Sigma}^{-1} \cdot \int_0^1 \left[\frac{1}{\omega} (\psi(F^{-1}(t)) - \bar{\psi}) - \frac{1}{\gamma} (\varphi(t) - \bar{\varphi}) \right]^2 dt . \tag{3.1}$$

where

$$\bar{\psi} = \int \psi(x)\, dF(x)\,, \qquad \bar{\varphi} = \int_0^1 \varphi(t)\, dt\,.$$

The proof of Theorem 3.1 is postponed to Section 5.

We shall say that two sequences $\{\mathbf{X}_N\}$ and $\{\mathbf{Y}_N\}$ of random vectors are asymptotically equivalent in probability (denoted $\mathbf{X}_N \sim \mathbf{Y}_N$) if $\|\mathbf{X}_N - \mathbf{Y}_N\| \to_P 0$ as $N \to \infty$.

The following corollaries of Theorem 3.1 have an interest of their own.

COROLLARY 3.1. *Let the assumptions* 1°—4° *be satisfied and let* $\gamma \neq 0$, $\omega \neq 0$. *Then* $\hat{\Delta}_M \sim \hat{\Delta}_R$ *if and only if*

$$\varphi(t) = a\psi(F^{-1}(t)) + b \qquad \text{a.e.} \quad t \in (0, 1) \tag{3.2}$$

for some $a > 0$, $b \in R'$.

REMARK. Jaeckel [7] first expressed a conjecture of a close relation between M and R estimates in the location submodel. Corollary 3.1 essentially answers his conjecture. Moreover, the corollary enables us to carry the asymptotic properties of one type of estimate to the other type. For instance, it follows from Huber [4] and from Corollary 3.1 that M and R estimates have an asymptotically minimax property over the set of asymptotically unbiased estimates in the model of symmetric contamination.

Put

$$\psi(x) = -\frac{g'(x)}{g(x)}\,, \qquad x \in R^1 \tag{3.3}$$

where g is a unimodal density with finite Fisher's information and such that $\int_{R^1} (g'(x)/g(x))^2 f(x)\, dx < \infty$. Then $\hat{\mathbf{\Delta}}_M$ is the maximum likelihood estimate corresponding to g. Similarly, put

$$\varphi(t) = \varphi(t, g)\,, \qquad 0 < t < 1. \tag{3.4}$$

Then $\hat{\mathbf{\Delta}}_R$ corresponding to φ is the rank estimate, asymptotically efficient for g in the role of the basic distribution. It follows from Theorem 3.1 that the asymptotic distribution of $\hat{\mathbf{\Delta}}_M - \hat{\mathbf{\Delta}}_R$ is then normal with center $\mathbf{0}$ and the covariance matrix

$$\mathbf{\Sigma}^{-1} \cdot \int_0^1 \left[\frac{1}{\omega}\left(\frac{g'(F^{-1}(t))}{g(F^{-1}(t))} - \bar{g}\right) + \frac{1}{\gamma}\,\varphi(t, g)\right]^2 dt\,. \tag{3.5}$$

where

$$\bar{g} = \int_{R^1} \frac{g'(x)}{g(x)}\, dF(x)\,.$$

Under (3.3) and (3.4), we have the following corollaries:

COROLLARY 3.2. *Let* ψ *and* φ *satisfy* (3.3) *and* (3.4) *respectively, where* g *is the density of the normal distribution* $N(0, \sigma^2)$, $\sigma^2 > 0$. *Then* $\hat{\mathbf{\Delta}}_M \sim \hat{\mathbf{\Delta}}_R$ *if and only if the basic distribution* f *is normal* $N(a, \lambda^2)$ *for some* $\lambda^2 > 0$, $a \in R^1$.

COROLLARY 3.3. *Let* ψ *and* φ *satisfy* (3.3) *and* (3.4) *respectively with* g *being the logistic density. Then* $\hat{\mathbf{\Delta}}_M \sim \hat{\mathbf{\Delta}}_R$ *if and only if* $f \equiv g$.

REMARK. Let ψ and φ satisfy (3.3) and (3.4) respectively with g being the density of the double-exponential distribution, $g(x) = \frac{1}{2}e^{-|x|}$, $x \in R^1$. Then $\hat{\mathbf{\Delta}}_M$ and $\hat{\mathbf{\Delta}}_R$ coincide for any symmetric basic distribution f.

4. Asymptotic behavior of $M_j(\mathbf{X}_N, \mathbf{\Delta})$. The proof of Theorem 3.1 will utilize an approximation of $M_j(\mathbf{X}_N, \mathbf{\Delta})$ by a linear function of $\mathbf{\Delta}$. A similar linear approximation has been studied by Huber [3] (see his Lemma 3 and its corollary) under a different set of conditions. Although our approach to the asymptotic linearity is quite different from Huber's development, it seems sufficient to outline here only the main ideas of the proof with references to analogous considerations in Jurečková [9] and [10].

THEOREM 4.1. *Under the assumptions* $1°$, $2°$ *and* $4°$,

$$\lim_{N\to\infty} P_{\mathbf{\Delta}^0}\{\max_{\|\mathbf{\Delta}-\mathbf{\Delta}^0\|\leqq K} |M_j(\mathbf{X}_N, \mathbf{\Delta}) - M_j(\mathbf{X}_N, \mathbf{\Delta}^0) + \omega(\mathbf{\Delta} - \mathbf{\Delta}^0)\boldsymbol{\sigma}^{(j)}| \geqq \varepsilon\} = 0 \tag{4.1}$$

holds for any $K > 0$, $\varepsilon > 0$ *and* $j = 1, \cdots, p$; $\boldsymbol{\sigma}^{(j)}$ *is the* jth *column of* $\mathbf{\Sigma}$.

PROOF. We may suppose, without loss of generality, that ψ is nondecreasing and that $\mathbf{\Delta}^0 = \mathbf{0}$. Let us fix $\varepsilon, K > 0$. For a fixed h, $1 \leqq h \leqq p$, denote $A_h = \{\mathbf{\Delta}: \Delta_k = 0 \text{ for } k \neq h, k = 1, \cdots, p\}$. We shall first prove (4.1) for a fixed $\mathbf{\Delta} \in A_h$ and then extend it to any fixed $\mathbf{\Delta} \in R^p$ by contiguity.

From $2°$ and Theorem 2.1 of Hájek–Šidák [2], the densities $\prod_{i=1}^N f(x_i + \Delta_h c_{hi})$ are contiguous with respect to $\prod_{i=1}^N f(x_i)$. Noting this fact and utilizing the same theorem of [2] and the third Le Cam's lemma (see Lemma 6.1.4 of [2]) we get that, for $\Delta \in A_h$, $M_j(\mathbf{X}_N, \mathbf{\Delta})$ is asymptotically normal

$$N(-\Delta_h \omega\sigma_{jh}, \rho^2\sigma_{jj}) \tag{4.2}$$

where σ_{jh} and σ_{jj} are the elements of $\mathbf{\Sigma}$.

Denote

$$\xi(t) = \psi(F^{-1}(t)), \qquad 0 < t < 1 \tag{4.3}$$

and

$$\begin{aligned} \xi^{(m)}(t) &= \xi\left(\frac{1}{m}\right) && \text{if } 0 < t < \frac{1}{m} \\ &= \xi(t) && \text{if } \frac{1}{m} \leqq t \leqq 1 - \frac{1}{m} \\ &= \xi\left(1 - \frac{1}{m}\right) && \text{if } 1 - \frac{1}{m} < t < 1 \end{aligned} \tag{4.4}$$

and further denote $\psi^{(m)}(x) = \xi^{(m)}(F(x))$, $x \in R^1$, $m = 1, 2, \cdots$. Then

$$E[M_j(\mathbf{X}_N, \mathbf{0}) - M_j^{(m)}(\mathbf{X}_N, \mathbf{0})]^2 = \sum_{i=1}^N c_{ji}^2 \int_0^1 [\xi(t) - \xi^{(m)}(t)]^2\, dt \leqq \varepsilon \tag{4.5}$$

holds for $m > m_0$ uniformly in N, where

$$M_j^{(m)}(\mathbf{X}_N, \mathbf{\Delta}) = \sum_{i=1}^N c_{ji}\psi^{(m)}(X_{Ni} - \Delta_h c_{hi}) .$$

The contiguity mentioned above in connection with Lemma 3.5 of Jurečková

[9] then implies that, given an $\eta > 0$, there exists an m_1 such that to any $m > m_1$ corresponds an $N_1(m)$ such that

$$(4.6) \qquad P_0\{|M_j^{(m)}(\mathbf{X}_N, \mathbf{\Delta}) - M_j(\mathbf{X}_N, \mathbf{\Delta})| \geqq \eta\} \leqq \varepsilon \quad \text{for} \quad N > N_1(m) .$$

Further,

$$(4.7) \qquad \operatorname{Var}[M_j^{(m)}(\mathbf{X}_N, \mathbf{\Delta}) - M_j^{(m)}(\mathbf{X}_N, \mathbf{0})] \leqq \sum_{i=1}^N c_{ji}^2 \int [\psi^{(m)}(x - \Delta_h c_{hi}) - \psi^{(m)}(x)]^2 \, dF(x) \to 0$$

as $N \to \infty$ for fixed m in view of bounded Lebesgue's theorem and of 2°. Chebyshev's inequality then implies that, given any $\varepsilon > 0$ and any fixed $m = 2, 3, \cdots$,

$$(4.8) \qquad \lim_{N\to\infty} P_0\{|M_j^{(m)}(\mathbf{X}_N, \Delta) - M_j^{(m)}(\mathbf{X}_N, \mathbf{0}) - EM_j(\mathbf{X}_N, \mathbf{\Delta})| \geqq \varepsilon\} = 0 .$$

The asymptotic normality (4.2) and (4.8) imply that, for any $\mathbf{\Delta} \in A_h$,

$$(4.9) \qquad \lim_{N\to\infty} P_0\{|M_j^{(m)}(\mathbf{X}_N, \Delta) - M_j^{(m)}(\mathbf{X}_N, \mathbf{0}) + \Delta_h \omega^{(m)} \sigma_{jh}| \geqq \varepsilon\} = 0$$

where $\omega^{(m)} = -\int \psi^{(m)}(x) f'(x) \, dx$. It follows from (4.5), (4.6), (4.8) and (4.9) that

$$(4.10) \qquad \lim_{N\to\infty} P_0\{|M_j(\mathbf{X}_N, \mathbf{\Delta}) - M_j(\mathbf{X}_N, \mathbf{0}) + \Delta_h \omega \sigma_{jh}| \geqq \varepsilon\} = 0$$

holds for $\mathbf{\Delta} \in A_h$, $\varepsilon > 0$ and $j = 1, \cdots, p$. (4.10) then extends to any fixed $\mathbf{\Delta} \in R^p$ by the contiguity of the sequences $\prod_{i=1}^N f(x_i + \Delta_k c_{ki})$, $k = 1, \cdots, p$, with respect to $\prod_{i=1}^N f(x_i)$.

It remains to prove that the linear approximations are uniform in $||\mathbf{\Delta}|| \leqq K$. But the proof is quite analogous to the corresponding proof in [10]. Assumption 2° on regression constants and the monotonicity of ψ enable us to decompose M_j into two statistics monotone in the components of $\mathbf{\Delta}$; the monotonicity then implies the uniformity of the linear approximation.

5. Proof of Theorem 3.1. First of all, Theorem 4.1 has an easy corollary, which we write as a lemma.

LEMMA 5.1. *If $\{\mathbf{\Delta}_N^*\}_{N=1}^\infty$ is a sequence of random vectors such that $\mathbf{\Delta}_N^* - \mathbf{\Delta}^0$ are bounded in probability and if the assumptions* 1°, 2° *and* 4° *are satisfied, then*

$$(5.1) \qquad \lim_{N\to\infty} P_{\mathbf{\Delta}^0}\{|M_j(\mathbf{X}_N, \mathbf{\Delta}_N^*) - M_j(\mathbf{X}_N, \mathbf{\Delta}^0) + \omega(\mathbf{\Delta}_N^* - \mathbf{\Delta}^0)\boldsymbol{\sigma}^{(j)}| \geqq \varepsilon\} = 0$$

for $j = 1, \cdots, p$ and any $\varepsilon > 0$.

The proof of Theorem 3.1 is based on an application of Lemma 5.1 to $\Delta_N^* = \hat{\Delta}_M$, so that we need to know if $(\hat{\mathbf{\Delta}}_M - \mathbf{\Delta}^0)$ are bounded in probability. The following lemma shows that this is the case under the assumptions 1°, 2° and 4°. The proof follows the ideas of the proof of Lemma 4.3 of [10].

LEMMA 5.2. *Under the assumptions* 1°, 2° *and* 4°, *to any $\varepsilon > 0$ correspond $K > 0$, $\eta > 0$ and a positive integer N_0 such that*

$$(5.2) \qquad P_{\mathbf{\Delta}^0}\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0|| \geqq K} ||\mathbf{M}(\mathbf{X}_N, \mathbf{\Delta})|| < \eta\} < \varepsilon$$

holds for $N > N_0$, where

$$(5.3) \qquad \mathbf{M}(\mathbf{X}_N, \mathbf{\Delta}) = (M_1(\mathbf{X}_N, \mathbf{\Delta}), \cdots, M_p(\mathbf{X}_N, \mathbf{\Delta})) .$$

PROOF. It follows from Lemma 4.1 that, given any $\varepsilon > 0$, there exist N_0 and $K_0 > 0$ such that

$$P_{\mathbf{\Delta}^0}\{||M(\mathbf{X}_N, \mathbf{\Delta}^0)|| > K_0\} < \tfrac{1}{2}\varepsilon \qquad \text{for} \quad N > N_0 . \tag{5.4}$$

Let K and η be any numbers satisfying

$$K > 2K_0/(\lambda_0\omega) \qquad\qquad \eta < K_0/2 \tag{5.5}$$

where λ_0 is the minimal eigenvalue of $\mathbf{\Sigma}$.

Then, it follows from Theorem 4.1 and from (5.4) that for $N > N_0$

$$P_{\mathbf{\Delta}^0}\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_{j=1}^p (\Delta_j^0 - \Delta_j)M_j(\mathbf{X}_N, \Delta) < \eta_0\} < \varepsilon \tag{5.6}$$

where $\eta_0 = \eta K_0$.

Actually, the left-hand side of (5.6) is less than or equal to

$$\begin{aligned} &P_{\mathbf{\Delta}^0}\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_{j=1}^p (\mathbf{\Delta}_j^0 - \Delta_j)M_j(\mathbf{X}_N, \mathbf{\Delta}) < \eta_0, \\ &\qquad \min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_{j=1}^p (\Delta_j^0 - \Delta_j)[M_j(\mathbf{X}_N, \mathbf{\Delta}^0) \\ &\qquad + \omega(\Delta^0 - \Delta)\boldsymbol{\sigma}^{(j)}] \geqq 2\eta_0\} \\ &\qquad + P_{\mathbf{\Delta}^0}\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_{j=1}^p (\Delta_j^0 - \Delta_j)[M_j(\mathbf{X}_N, \mathbf{\Delta}^0) \\ &\qquad + \omega(\mathbf{\Delta}^0 - \mathbf{\Delta})\boldsymbol{\sigma}^{(j)}] < 2\eta_0\} . \end{aligned} \tag{5.7}$$

The first term of (5.7) is less than or equal to

$$\begin{aligned} &P_{\mathbf{\Delta}^0}\{\max_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_j (\Delta_j^0 - \Delta_j)[M_j(\mathbf{X}_N, \mathbf{\Delta}^0) + \omega(\mathbf{\Delta}^0 - \mathbf{\Delta})\boldsymbol{\sigma}^{(j)} \\ &\quad - M_j(\mathbf{X}_N, \mathbf{\Delta})] \geqq \eta_0\} \\ &\quad \leqq P_{\mathbf{\Delta}^0}\{\max_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_j |M_j(\mathbf{X}_N, \mathbf{\Delta}^0) + \omega(\mathbf{\Delta}^0 - \mathbf{\Delta})\boldsymbol{\sigma}^{(j)} - M_j(\mathbf{X}_N, \mathbf{\Delta})| \\ &\quad \geqq \eta\} \to 0 \qquad \text{as} \quad N \to \infty . \end{aligned} \tag{5.8}$$

The second term of (5.7) is less than or equal to

$$\begin{aligned} &P_{\mathbf{\Delta}^0}\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||=K} \textstyle\sum_{j=1}^p (\Delta_j^0 - \Delta_j)M_j(\mathbf{X}_N, \mathbf{\Delta}^0) + K^2\lambda_0\omega < 2\eta_0\} \\ &\qquad \leqq P_{\mathbf{\Delta}^0}\{-K||\mathbf{M}(\mathbf{X}_N, \mathbf{\Delta}^0)|| < 2\eta_0 - K^2\lambda_0\omega\} \to 0 \qquad \text{as} \quad N \to \infty \end{aligned}$$

so that (5.6) is proved.

Let $\mathbf{\Delta}^1$ be any point such that $||\mathbf{\Delta}^1 - \mathbf{\Delta}^0|| = K$. Put

$$c_i^* = (\mathbf{\Delta}^0 - \mathbf{\Delta}^1)\mathbf{c}^{(i)} , \qquad X_i^* = X_i - \mathbf{\Delta}^0\mathbf{c}^{(i)} , \qquad i = 1, \cdots, N$$

and

$$M(\tau) = \textstyle\sum_{i=1}^N c_i\psi(X_i^* + \tau c_i^*) .$$

Then $M(\tau)$ is nondecreasing in τ, so that

$$\begin{aligned} &\textstyle\sum_{j=1}^p (\Delta_j^0 - \Delta_j^1)M_j(\mathbf{X}_N, \mathbf{\Delta}^0 + \tau(\mathbf{\Delta}^1 - \mathbf{\Delta}^0)) \\ &\qquad = M(\tau) \geqq M(1) = \textstyle\sum_{j=1}^p (\Delta_j^0 - \Delta_j^1)M_j(\mathbf{X}_N, \mathbf{\Delta}^1) \qquad \text{for} \quad \tau \geqq 1 . \end{aligned} \tag{5.9}$$

If $||\mathbf{\Delta} - \mathbf{\Delta}^0|| \geqq K$ and $\mathbf{\Delta}^1 = \mathbf{\Delta}^0 + (K/||\mathbf{\Delta} - \mathbf{\Delta}_0||)(\mathbf{\Delta} - \mathbf{\Delta}^0)$ then $||\mathbf{\Delta}^1 - \mathbf{\Delta}^0|| = K$ and

$\mathbf{\Delta} = \mathbf{\Delta}^0 + \tau(\mathbf{\Delta}^1 - \mathbf{\Delta}^0)$ for $\tau = ||\mathbf{\Delta} - \mathbf{\Delta}^0||/K \geqq 1$. (5.6) and (5.9) then imply that

$$P_{\mathbf{\Delta}^0}\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||\geqq K} ||\mathbf{M}(\mathbf{X}_N, \mathbf{\Delta})|| < \eta\}$$
$$\leqq P_{\mathbf{\Delta}^0}\left\{\min_{||\mathbf{\Delta}-\mathbf{\Delta}^0||\geqq K}\left[\sum_{j=1}^p (\Delta_j^0 - \Delta_j)M_j(\mathbf{X}_N, \mathbf{\Delta})\right]\frac{K}{||\mathbf{\Delta} - \mathbf{\Delta}^0||} < \eta K\right\}$$
$$\leqq P_{\mathbf{\Delta}^0}\{\min_{\mathbf{\Delta}^1:||\mathbf{\Delta}^1-\mathbf{\Delta}^0||=K} \sum_j (\Delta_j^0 - \Delta_j^1)M_j(\mathbf{X}_N, \mathbf{\Delta}^1) < \eta_0\} < \varepsilon \quad \text{for } N > N_0 . \qquad \square$$

PROOF OF THEOREM 3.1. It follows from Lemma 5.1 and Lemma 5.2 that

$$\lim_{N\to\infty} P_{\mathbf{\Delta}^0}\{|M_j(\mathbf{X}_N, \mathbf{\Delta}^0) - \omega(\hat{\mathbf{\Delta}}_M - \mathbf{\Delta}^0)\boldsymbol{\sigma}^{(j)}| \geqq \varepsilon\} = 0 \tag{5.10}$$

holds for $j = 1, \cdots, p$ and any $\varepsilon > 0$; thus

$$\hat{\mathbf{\Delta}}_M \sim \mathbf{\Delta}^0 + (1/\omega)\mathbf{M}(\mathbf{X}_N, \mathbf{\Delta}^0)\,\mathbf{\Sigma}^{-1} . \tag{5.11}$$

On the other hand, it follows from Lemma 4.5 of [10] that

$$\mathbf{\Delta}_R \sim \mathbf{\Delta}^0 + (1/\gamma)\mathbf{S}(\mathbf{X}_N, \mathbf{\Delta}^0)\mathbf{\Sigma}^{-1} \tag{5.12}$$

where

$$\mathbf{S}(\mathbf{X}_N, \mathbf{\Delta}) = (S_1(\mathbf{X}_N, \mathbf{\Delta}), \cdots, S_p(\mathbf{X}_N, \mathbf{\Delta}))$$

and

$$S_j(\mathbf{X}_N, \mathbf{\Delta}) = \sum_{i=1}^N c_{ji}a_N(R_i^{\Delta}) , \qquad j = 1, \cdots, p .$$

(5.11) and (5.12) imply

$$\hat{\mathbf{\Delta}}_M - \hat{\mathbf{\Delta}}_R \sim [(1/\omega)\mathbf{M}(\mathbf{X}_N, \mathbf{\Delta}^0) - (1/\gamma)\mathbf{S}(\mathbf{X}_N, \mathbf{\Delta}^0)]\mathbf{\Sigma}^{-1} . \tag{5.13}$$

Further, it follows from Theorem 5.1.5.a of [2] that

$$\mathbf{S}(\mathbf{X}_N, \mathbf{\Delta}^0) \sim \mathbf{T}_N \tag{5.14}$$

where

$$\mathbf{T}_N = (T_N^{(1)}, \cdots, T_N^{(p)}) \quad \text{and} \quad T_N^{(j)} = \sum_{i=1}^N c_{ji}\varphi[F(\delta_i(\mathbf{\Delta}^0))] . \tag{5.15}$$

(5.13) and (5.15) imply

$$\hat{\mathbf{\Delta}}_M - \hat{\mathbf{\Delta}}_R \sim [(1/\omega)\boldsymbol{\psi}(\boldsymbol{\delta}(\mathbf{\Delta}^0)) - (1/\gamma)\boldsymbol{\varphi}(F(\boldsymbol{\delta}(\mathbf{\Delta}^0))]\mathbf{C}_N^T\mathbf{\Sigma}^{-1} . \tag{5.16}$$

The rest of the proof then follows easily from Theorem 5.1.2 of [2]. $\square$

REFERENCES

[1] BICKEL, P. J. (1973). On some analogues to linear combinations of order statistics in the linear model. *Ann. Statist.* **1** 597-616.

[2] HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academia, Prague.

[3] HUBER, P. J. (1966). The behavior of maximum likelihood estimates under nonstandard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 221-233, Univ. of California Press.

[4] HUBER, P. J. (1969). *Théorie de L'inference Statistique Robuste.* Les presses de l'Université de Montréal.

[5] HUBER, P. J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Ann. Statist.* **1** 799-821.

[6] HUBER, P. J. (1975). Robust methods of estimation of regression coefficients. 2nd Internationale Sommerschule Modellwahl DDR 1975.

[7] JAECKEL, L. A. (1971). Robust estimates of location: Symmetry and asymmetry contamination. *Ann. Math. Statist.* **42** 1020–1034.

[8] JACKEL, L. A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. *Ann. Math. Statist.* **43** 1449–1458.

[9] JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.

[10] JUREČKOVÁ, J. (1971). Nonparametric estimate of regression coefficients. *Ann. Math. Statist.* **42** 1328–1338.

[11] JUREČKOVÁ, J. (1974). Asymptotic relations of least-squares estimate and of two robust estimates of regression parameter vector. To appear in *Trans. Seventh Prague Conf. and European Meeting of Statisticians.*

[12] JUREČKOVÁ, J. (1975). Asymptotic comparison of maximum likelihood and a rank estimate in simple linear regression model. *Comment. Math. Univ. Carolinae* **16** 87–97.

[13] KOUL, H. L. (1971). Asymptotic behavior of a class of confidence regions based on ranks in regression. *Ann. Math. Statist.* **42** 466–476.

[14] KRAFT, C. and VAN EEDEN, C. (1972). Linearized rank estimates and signed-rank estimates for the general linear hypothesis. *Ann. Math. Statist.* **43** 42–47.

[15] LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

DEPARTMENT OF PROBABILITY AND STATISTICS
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
186 00 PRAGUE 8
CZECHOSLOVAKIA