A SECOND-ORDER ASYMPTOTIC DISTRIBUTIONAL REPRESENTATION M-ESTIMATORS WITH DISCONTINUOUS SCORE FUNCTIONS

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1485

May 1985

A SECOND-ORDER ASYMPTOTIC DISTRIBUTIONAL REPRESENTATION OF M-ESTIMATORS WITH DISCONTINUOUS SCORE FUNCTIONS

By Jana Jurečková and Pranab Kumar Sen[1]
Charles University, Prague, Czechoslovakia,
and University of North Carolina, Chapel Hill

For a nondecreasing score function having finitely many jump-discontinuities, a representation of M-estimators with the second order asymptotic distribution is established, and the result is also extended to one-step versions of M-estimators.

1. Introduction. Let $\{X_i;\ i \geq 1\}$ be a sequence of independent and identically distributed random variables (i.i.d.r.v.) with a distribution function (d.f.) $F(x - \theta)$, where θ is an unknown location parameter. Let $\Psi: R^1 \to R^1$ be a function such that for

(1.1) $$\lambda(t) = \int_{-\infty}^{\infty} \Psi(x-t)\, dF(x), \quad t \in R^1, \quad \lambda(0) = 0.$$

Consider an estimating function

(1.2) $$M_n(t) = \Sigma_{i=1}^{n} \Psi(x_i - t), \quad t \in R^1,\ n \geq 1.$$

AMS Subject Classification: 60F17, 62E20, 62F35, 62G05
Key Words and Phrases: Jump discontinuity, M-estimator, one-step version of M-estimator, random change of time, weak convergence of M-processes.

1) Work of this author was partially supported by the Office of Naval Research, Contract No. N00014-83-K-0387.

The allied M-estimator $\hat{\theta}_n$ of θ (corresponding to a fixed scale of F) is defined as a solution of the equation

$$(1.3) \qquad M_n(t) = 0;$$

under quite general regularity conditions, $\hat{\theta}_n$ is a (weakly) consistent estimator of θ. For nondecreasing Ψ, $\hat{\theta}_n$ may be written as

$$(1.4) \qquad \hat{\theta}_n = \frac{1}{2}(\sup\{t: M_n(t) > 0\} + \inf\{t: M_n(t) < 0\}).$$

The existence of $\hat{\theta}_n$ in general as well as the boundedness (in probability) of $n^{1/2}|\hat{\theta}_n - \theta|$ has been studied by a host of workers. For a general estimator $T_n = T_n(x_1,\ldots,x_n)$ of θ, denoting by $IF_T(\cdot)$ the influence function of T_n, an asymptotic representation of the form

$$(1.5) \qquad n^{1/2}(T_n-\theta) = n^{-1/2} \Sigma_{i=1}^n IF_T(x_i-\theta) + R_n; \quad R_n = o_p(1),$$

is of primary interest, and has been studied by many authors (viz., Serfling (1980), Huber (1981) and the references cited therein). A representation of this type for M-estimators of location, supplemented by the order of R_n, was studied by Carroll (1978) for strongly consistent versions of $\hat{\theta}_n$ and for smooth Ψ-functions. Jurečková (1980) derived the exact orders of R_n for smooth as well as discontinuous Ψ-functions; this result was extended to the regression model by Jurečková and Sen (1981a,b). Jurečková, Janssen and Veraverbeke (1985) obtained an analogous result for M-estimators of general parameters (including the maximum likelihood estimator and the Pitman estimator).

The representation in (1.5), supplemented by the order of R_n, has important applications. Firstly, asymptotic relations of different types of estimators (e.g., L-, M-, R-estimators), up to different orders of equivalence, can be fruitfully studied with the aid of (1.5). We may refer to Jurečková (1983) for some of these developments. Secondly, in M-estimation theory, generally, $\hat{\theta}_n$ is not scale-equivariant, and hence, either one-step versions or some other modifications (see, for example, Jurečková and Sen (1984)) are usually adapted to suit the purpose. Approximating an estimator by its one-step version or some other form, may also be effectively studied by incorporating (1.5) for such a T_n. Actually, in sequential analysis (or in related areas), for such a representation, one needs to be more precise on the order of R_n (see for example, Jurečková and Sen (1982)), and this constitutes an important area of research too. There are numerous other applications of such a representation.

The next natural step in representations of type (1.5) (which we term as the first order asymptotic representations) would be to supplement the order of R_n by the asymptotic distribution (if any) of the same. A result of this kind was derived by Kiefer (1967) in the context of Bahadur's (1966) representation of sample quantiles. Asymptotic representations supplemented by asymptotic distribution of remainder term will be referred to as the second order asymptotic (distributional) representations; the asymptotic distributions of R_n will be typically nonnormal. The second order asymptotic representation of M-estimator of location generated by a smooth Ψ-function was recently derived by Jurečková (1985), who used the method of random change of time in the invariance principle connected with M-estimator.

Our aim will be to extend this result to possibly discontinuous Ψ-functions.

More precisely, the representation (1.5) for M-estimator $\hat{\theta}_n$ takes on the form

$$n^{1/2}(\hat{\theta}_n - \theta) = n^{-1/2}\gamma^{-1} M_n(\theta) + R_n \tag{1.6}$$

under quite general conditions, provided $\gamma = \gamma(\Psi, F) \neq 0$. If Ψ is smooth (i.e., twice differentiable almost everywhere (a.e.), then $R_n = O_p(n^{-1/2})$ while, for Ψ admitting finitely many jump discontinuities, $R_n = O_p(n^{-1/4})$. Let further T_n be an estimator of θ admitting a first order representation i.e.,

$$n^{1/2}(T_n - \theta) = n^{-1/2} \Sigma_{i=1}^{n} \gamma(x_i - \theta) + o_p(1) \tag{1.7}$$

where

$$\int_{R^1} \phi(x)\, dF(x) = 0 \text{ and } 0 < \sigma_\phi^2 = \int_{R^1} \phi^2(x)\, dF(x) < \infty. \tag{1.8}$$

Then, for smooth Ψ, as $n \to \infty$,

$$n(T_n - \theta) - n\alpha(T_n - \theta)^2 + \gamma^{-1}\{M_n(T_n) - M_n(\theta)\} \xrightarrow{\mathcal{D}} \xi_1 \cdot \xi_2 \tag{1.9}$$

with $\alpha = \int \Psi(x)\, dF(x)/2\gamma$ and where $\xi = (\xi_1, \xi_2)$ has a bivariate normal distribution (see Jurečková (1985)). (1.9) applies also for $\hat{\theta}_n$, for the least-square estimator (LSE) as well as for the maximum likelihood estimator

(MLE) in the role of T_n under general regularity conditions. However, it breaks down when Ψ is not smooth.

The primary objective of the present study is to focus on such a second order representation in the case where Ψ admits of jump-discontinuities. A different normalizing rate (in μ) as well as a different type of limiting law arises in this context comparing with (1.9). However, the result is in correspondence with that of Kiefer (1967). Specifically, for an estimator T_n, satisfying (1.7), we have $n^{1/2}(T_n - \theta) \xrightarrow{\mathcal{D}} \xi$ and

$$(1.10) \qquad n^{-1/4}\{n(T_n-\theta) - \gamma^{-1}[M_n(T_n) - M_n(\theta)]\} \to \xi^*$$

where, letting I[A] be the indicator function of the set A,

$$(1.11) \qquad \xi^* = I[\xi>0]\; W_1(|\xi|) + I[\xi<0]\; W_2(|\xi|),$$

ξ has a normal distribution with zero mean and W_1, W_2 are independent copies of Wiener process on $[0,\infty)$.

Along with the preliminary notions, the main results are presented in Section 2 and they are derived in Section 3. The method used there is based on a random change of time in certain invariance principles for M-statistics, studied earlier by Jurečková (1980) and Jurečková and Sen (1981a, b). The last section deals with the second order behavior of the one-step version of M-estimators and, in this context, illustrates the effect of the choice of an initial estimator.

2. A second order representation theorem. Concerning the score function Ψ, we assume that

$$(2.1) \qquad \Psi(x) = \Psi_1(x) + \Psi_2(x), \qquad x \in R^1$$

where Ψ_1 is absolutely continuous on any bounded interval in R^1 and it possesses first and second derivatives ($\Psi^{(1)}$ and $\Psi^{(2)}$, respectively) a.e., and Ψ_2 is a step-function. Specifically, we assume that for some p (≥ 1), there exist real numbers β_j and open intervals $E_j=(a_j, a_{j+1})$, $j=0,1,\ldots p$, where $-\infty = a_0 < a_1 < \ldots < a_0 < a_{p+1} = \infty$, such that $\Psi_2(x)=\beta_j$ for $x \in E_j$, $0 \leq j \leq p$; conventionally, we let $\Psi_2(a_j) = \frac{1}{2}(\beta_j + \beta_{j-1})$, $1 \leq j \leq p$. Also, we assume that Ψ is increasing and that $\int_{R^1} \Psi(x)\, dF(x)=0$.

Concerning the d.f. F, we assume that F possesses an absolutely continuous and symmetric density f having a finite Fisher information

$$(2.2) \qquad I(f) = \int_{R^1} (f'(x)/f(x))^2\, dF(x) \quad (<\infty)$$

where $f'(x) = (d/dx)\, f(x)$. Also, we assume that

$$(2.3) \qquad \gamma_{1\nu} = \int_{R^1} \Psi_1^{(\nu)}(x)\, dF(x) \qquad \text{exists} \quad (\nu = 1,2),$$

$$(2.4) \qquad \int_{R^1} (\Psi_1^{(\nu)}(x))^2\, dF(x) < \infty,$$

and, either, Ψ_1 is constant outside a fixed interval $[-k_1,k]$, $k > 0$, or, there exist positive and finite numbers σ and K such that

(2.5) $$\int_{R^1} (\Psi_1^{(2)}(x \pm t))^2 \, dF(x) < \infty \quad \forall |t| < \sigma.$$

Further, we assume that f' is bounded and continuous in a neighbourhood of a_j, $j=1,\dots,p$.

Denote

(2.6) $$\gamma_{2\nu} = \Sigma_{j=1}^{p} (\beta_j - \beta_{j-1})^{\nu} f(a_j), \quad \nu = 1,2,; \quad \gamma = \gamma_{11} + \gamma_{21},$$

(2.7) $$\gamma_2^0 = \Sigma_{j=1}^{p} (\beta_j - \beta_{j-1}) f'(a_j); \quad \gamma^* = (\gamma_2^0 + \gamma_{12})/2\gamma$$

where we assumed

(2.8) $$\gamma \neq 0, \quad \gamma_{22} > 0.$$

Then, we have the following

THEOREM 2.1. <u>Under the assumed regularity conditions, provided</u> $\Psi_2 \neq 0$, <u>for any</u> $\{T_n\}$ <u>satisfying</u> (1.7), <u>the</u> r.v.

(2.9) $$Z_n = n^{-1/4}\{\gamma^{-1}[M_n(T_n) - M_n(\theta)] + n(T_n - \theta) + n\gamma^*(T_n - \theta)^2\}$$

<u>converges in law to</u>

(2.10) $$\xi^* = H\{I[\xi > 0]\, W_1(|\xi|) + I[\xi < 0]\, W_2(|\xi|)\}$$

where W_1, W_2 *are independent copies of a standard Wiener process on* $[0,\infty)$ *and* ξ *has a standard normal distribution, independently of* W_1, W_2 *and*

(2.11) $$H = (\sigma_\phi \, \gamma_{22})^{1/2} \, \gamma^{-1};$$

i.e., Z_n *has an asymptotic* d.f.

(2.12) $$P(\xi^* \le x) = 2 \int_0^\infty \Phi(H\, x\, t^{-1/2}) \, d\phi(t)$$

where Φ *is the standard normal* d.f. *If, however,* $\Psi_2 \equiv 0$ *in* (2.1), *then* (1.9) *holds*.

Remarks. (i) For (1.9) to hold, we need that $\Psi_2 \equiv 0$ although, for (2.9) - (2.12), it is not necessary to assume that $\Psi_1 \equiv 0$.

(ii) For $T_n = \hat{\theta}_n$ we have $\sigma_\phi^2 = \gamma^{-2}\sigma_0^2$ with $\sigma_0^2 = \int_{R^1} \Psi^2(x)\, dF(x)$, so that (2.9) - (2.11) holds with $H = (\sigma_0 \, \gamma_{22}/\gamma^{-3})^{1/2}$.

(iii) The quadratic term in (2.9) may be omitted as $n^{3/4}(T_n - \theta)^2 = 0_p(n^{-1/4}) = 0_p(1)$.

(iv) In particular, let $0 < p < 1$ and put $\Psi_1 \equiv 0$ and $\Psi_2(x) = p - I[x < 0]$; then $\int \Psi(x - t)\, dF(x) - 0$ for $t - F^{-1}(p)$. Replacing $\Psi(x)$ by $\Psi^*(x) = \Psi(x - F^{-1})(p))$, we have $\gamma_{21} = f(F^{-1}(p)) = \gamma_{22} = \gamma$, $\sigma_0^2 = p(1 - p)$; hence we arrive at Kiefer's (1967) result.

3. The proof of Theorem 2.1. Consider the process

$$(3.1)\qquad W_n^0(t) = n^{-1/4}\{\gamma^{-1}[M_n(\theta + n^{-1/2}t) - M_n(\theta)] + n^{1/2}t - \gamma^* t^2\}, \qquad t \in R^1$$

where γ and γ^* are defined as in Section 2. By (2.9),

$$(3.2)\qquad Z_n = W_n^0(n^{1/2}(T_n - \theta)).$$

By Jurečková (1980) (see Corollary on p. 69), the process $W_n^* = \{W_n^*(t), |t| \le k\}$, where

$$(3.3)\qquad W_n^*(t) = \gamma \cdot \gamma_{22}^{-1/2}\, W_n^0(t), \qquad -K \le t \le K$$

converges to Gaussian process $W^* = \{W^*(t), |t| \le k\}$ with $EW^*(t) = 0$ $\forall\ t \in [-K, K]$ and

$$(3.4)\qquad EW^*(s)\, W^*(t) = \begin{cases} 0, & st \le 0 \\ |s| \wedge |t|, & st > 0, \end{cases}$$

in the Scorokhod topology on $D[-K, K]$ for any $K > 0$ as $n \to \infty$. Hence, to prove Theorem 2.1, we may make use of a random change of time (i.e., $t \to n^{1/2}(T_n - \theta)$) and use (3.2), (3.3) and (3.4). Towards this, we consider the weak convergence of $\{(W_n^0(t), \sqrt{n}(T_n - \theta)), |t| \le K\}$. Note that by (1.7) and (1.8)

(3.5) $$\sqrt{n}(T_n - \theta) = \sum_{i=1}^{n} U_{ni} + o_p(1)$$

where

(3.6) $$U_{ni} = n^{-1/2}\ \phi(x_i - \theta);\ E\ U_{ni} = 0,\ E\ U_{ni}^2 = n^{-1}\sigma_\phi^2,$$

$i = 1,\ldots,n.$

Let us first consider the case $\Psi \equiv \Psi_2$. Then, by (1.2) and by the definition of Ψ_2;

(3.7) $$n^{-1/4}\{M_n(\theta + n^{-1/2}) - M_n(\theta)\} = \sum_{i=1}^{n} U_{ni}^{*}\ (t)$$

where

$$U_{ni}^{*}(t) = n^{-1/4} \sum_{j=1}^{p} (\beta_j - \beta_{j-1})\{[I[t<0]\ I[a_j + n^{-1/2}t \le X_i - \theta \le a_j] -$$

(3.8) $$I[t>0]\ I[a_j \le X_j - \theta \le a_j + n^{-1/2}t]\}, \qquad i=1,\ldots,n$$

are i.i.d. random variables and n is so large that $a_{j+1} - a_j > n^{-1/2}K$, $0 \le j \le 0$.

Then

(3.9) $$\text{Var}\ (U_{ji}^{*}(t)) = |t|n^{-1}\ \gamma_{22} + o(n^{-1}) \quad \text{and}$$

$$\text{Cov}\ (U_{ni},\ U_{ni}^{*}(t)) = E(U_{ni} \cdot U_{ni}^{*}(t)) = 0(|t|n^{-5/4})$$

hence

$$(3.10) \qquad \mathrm{Cov}(n^{-1/4}[M_n(\theta + n^{-1/2}t) - M_n(\theta)], \quad n^{1/2}(T_n - \theta))$$

$$= 0(|t|\ n^{-1/4}) \to 0 \quad \text{as} \quad n \to \infty.$$

(3.5) - (3.10) readily leads to the convergence of the finite dimensional distributions of $\{W_n^0, \sqrt{n}(T_n - \theta)\}$ to those of $\{(\gamma^{-1}\gamma_{22}^{1/2}\ W^*, \xi^0)\}$ where $\xi^0 \sim \mathcal{N}(0, \sigma_\phi^2)$. Also, $\sqrt{n}(T_n - \theta)$ is relatively compact, while $\{W_n^0(t), |t| \leq K\}$ is relatively compact due to the weak convergence of W_n^* of (3.3). Thus, as $n \to \infty$,

$$(3.11) \quad \{(W_n^0(t), \sqrt{n}(T_n - \theta)), |t| \leq K|\} \overset{\mathcal{D}}{\to} (\gamma^{-1}\gamma_{22}^{1/2}\ W^*, \ \xi^0)\}.$$

Since, for any $\eta > 0$, there exists a $K(0 < K < \infty)$ such that $P(|\xi^0| \leq K) \geq 1 - \eta$, we may use (3.11) and apply the <u>random change of time</u>: $t \to [n^{1/2}(T_n - \theta)]^K$ (where $Y^K = Y$ if $|Y| \leq K$ and $Y^K = 0$ otherwise). Using the results in Section 17 of Billingsley (1968), we obtain that as $n \to \infty$,

$$(3.12)\ Z_n = W_n^0(\sqrt{n}(T_n - \theta)) \overset{\mathcal{D}}{\to} \gamma^{-1}\gamma_{22}^{1/2}\ W^*(\xi^0) = (\sigma_\phi \gamma_{22})^{1/2}\ \gamma^{-1}\ W^*(\xi)$$

where $\xi = \xi^0/\sigma_\phi \sim \mathcal{N}(0, 1)$. It is easy to show that

$$(3.13) \qquad W^*(t) = I[t>0]\ W_1(|t|) + I[t<0]\ W_2(|t|), \qquad t \in R^1$$

where W_1 and W_2 are two independent copies of a standard Wiener process on $[0,\infty)$. This completes the proof for $\Psi \equiv \Psi_2$. The proof for $\Psi \equiv \Psi_1$ is contained in Jurečková (1985).

If $\Psi = \Psi_1 + \Psi_2$ where none of the components vanish, it follows from Lemma 3.1 of Jurecková (1985) that

$$\sup_{|t| \leq K} |W^0_{n1}(t)| = 0_p(n^{-1/4}) \quad \text{as} \quad n \to \infty \tag{3.14}$$

where

$$W^0_{n1}(t) = n^{-1/4} \gamma^{-1} \sum_{i=1}^{n} [\Psi_1(X_i - \theta - n^{-1/2}) - \Psi_1(X_i - \theta)]$$

$$+ n^{1/2} t - (\gamma_{12}/2\gamma)\ t^2\}, \qquad t \in R^1 \tag{3.15}$$

is the component of $W^0_n(t)$ corresponding to Ψ_1. Hence, the weak convergence of $(W^0_n(t), \sqrt{n}(T_n - \theta))$ follows that of the component corresponding to Ψ_2.

Q.E.D.

4. Representation of one-step versions of M-estimators. Applying Theorem 2.1 to $T_n - \hat{\theta}_n$ of (1.4), we get that

$$n^{1/4} R_n = n^{-1/4} \{n(\hat{\theta}_n - \theta) - \gamma^{-1} M_n(\theta)\} \xrightarrow{\mathcal{D}} \xi^*_0 \tag{4.1}$$

where R_n is the remainder term in the representation (1.6) and

$$\xi^*_0 = H^* \{I[\xi>0]\ W_1(|\xi|) + I[\xi>0]\ W_2(|\xi|)\} \tag{4.2}$$

where W_1 and W_2 are independent copies of a standard Wiener process and ξ has the standard normal distribution and

$$H^* = (\sigma_0 \gamma_{22})^{1/2} \gamma^{-3/2}. \tag{4.5}$$

It is often convenient to approximate $\hat{\theta}_n$ by its one-step version $\hat{\theta}_n^*$ of the following form: starting from an initial estimate T_n satisfying $\sqrt{n}(T_n - \theta) = O_p(1)$, we put

$$\hat{\theta}_n^* = \begin{cases} T_n, & \text{if } \hat{\gamma}_n = 0 \\ T_n + (n\hat{\gamma}_n)^{-1} M_n(T_n), & \text{if } \hat{\gamma}_n \neq 0 \end{cases} \tag{4.4}$$

where $\hat{\gamma}_n$ is a consistent estimator of γ, and may be taken as

$$\hat{\gamma}_n = \{M_n(T_n + n^{-1/2} t_1) - M_n(T_n + n^{-1/2})\}/(\sqrt{n}(t_2 - t_1)) \tag{4.5}$$

with t_1, t_2 ($t_1 < t_2$) being some arbitrary real numbers; often, we let $t_1 = -t_2 = t(> 0)$. The one-stop M-estimators were first considered by Bickel (1975) and later on studied by Jurečková (1982) (see also Janssen, Jurečková and Veraverbeke (1985) for a more general setup) who showed that, provided $\Psi_2 \neq 0$,

$$|\hat{\gamma}_n^{-1} \gamma - 1| = O_p(n^{-1/4}) \quad \text{and} \quad |\hat{\theta}_n^* - \hat{\theta}_n| = O_p(n^{-3/4}) \tag{4.6}$$

for any $\sqrt{n}$-consistent initial estimator T_n. Hence, the asymptotic distribution of $\hat{\theta}_n^*$ coincides with that of $\hat{\theta}_n$ and the effect of the choice of T_n may appear only in the second order asymptotic properties. Hence, parallel to (4.1), we are interested in the limit distribution (if any) of

$$(4.7)\qquad Z_n^* - n^{-1/4}\{n(\hat{\theta}_n^* - \theta) - \hat{\gamma}^{-1} M_n(\theta)\}.$$

This is given in the following theorem.

THEOREM 4.1 Assume that the conditions of Theorem 2.1 are satisfied. Let $\hat{\theta}_n^*$ be the one-step version of $\hat{\theta}_n$ given in (4.4) with $\hat{\gamma}_n$ of (4.5) and with T_n satisfying (1.7) and (1.8). Then

$$Z_n^* \overset{\mathcal{D}}{\to} \xi_0^* - (t_2-t_1)^{-1}\, \mathcal{H}[W^*(\xi+t_1\sigma_\phi^{-1}) - W^*(\xi+t_2\sigma_\phi^{-1})]$$

$$\cdot\, [\sigma_\phi\, \xi - \gamma_1^{-1}\sigma_0\xi_0]$$

$$(4.8)\qquad = \mathcal{H}\{W^*(\xi) - (t_2-t_1)^{-1}\, [W^*(\xi+\frac{t_1}{\sigma_\phi}) - W^*(\xi+\frac{t_2}{\sigma_\phi})][\sigma_\phi\, \xi - \frac{\sigma_0}{\gamma}\xi_0]$$

where W^* is the Gaussian process of (3.4), $\mathcal{H}$ is defined in (2.11) and the vector $\xi = (\xi, \xi_0)$ has the bivariate normal distribution $\mathcal{N}_2(\tilde{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$ with

$$(4.9)\quad \rho = \int_{-\infty}^{\infty} \phi(x)\, \Psi(x)\, dF(x) \cdot \{\int_{-\infty}^{\infty} \phi^2(x)\, dF(x) \cdot \int_{-\infty}^{\infty} \Psi^2(x)\, dF(x)\}^{-1/2}.$$

Proof. It follows from (3.1) and (4.5) that

$$V_n = n^{1/4}(\hat{\gamma}_n-\gamma) = n^{1/4}\{[\gamma n^{1/4}(W_n^0(\sqrt{n}(T_n-\theta) + t_1) - W_n^0(\sqrt{n}(T_n-\theta) + t_2)$$

$$+ \gamma n^{1/2}(t_2-t_1)]\ n^{-1/2}(t_2-t_1)^{-1} - \gamma\} + o_p(1). \tag{4.10}$$

Also, by the Slutsky theorem,

$$V_n^* = n^{1/4}(\hat{\gamma}\gamma^{-1} - 1) = -\gamma^{-1}\ V_n + O_p(n^{-1/4}). \tag{4.11}$$

Further, by (3.1) and (4.4), we have

$$Z_n^* = n^{1/4}R_n + \gamma^{-1}\ n^{1/4}(\gamma\hat{\gamma}_n^{-1} - 1)\ n^{-1/2}[M_n(T_n) - M_n(\hat{\theta}_n)] + o_p(1)$$

$$= W_n^0(\sqrt{n}(T_n-\theta)) + V_n^*\ n^{1/2}(T_n-\hat{\theta}_n) + o_p(1)$$

$$= W_n^0(\sqrt{n}(T_n-\theta)) - (t_2-t_1)^{-1}[W_n(\sqrt{n}(T_n-\theta) + t_1)$$

$$- W_n^0(\sqrt{n}(T_n-\theta) + t_2)]\ n^{1/2}(T_n-\hat{\theta}_n). \tag{4.12}$$

Moreover, utilizing (1.6) and (1.7), we obtain that the limiting distribution of

$$\sqrt{n}(T_n-\hat{\theta}_n) = \sigma_\phi(\sqrt{n}(T_n-\theta)/\sigma_\phi) - \gamma^{-1}\sigma_0(\sqrt{n}(\hat{\theta}_n-\theta)\,\gamma/\sigma_0) \tag{4.13}$$

coincides with that of $\sigma_\phi\ \xi - \gamma^{-1}\sigma_0\xi_0$ where $(\xi,\ \xi_0)$ has the bivariate normal distribution with zero expectation and correlation matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ with ρ of (4.9).

It follows from the proof of Theorem 2.1 that the joint (trivariate) distribution of $(W_n^0(\sqrt{n}(T_n-\theta)),\ W_n^0(\sqrt{n}(T_n-\theta) + t_1),\ W_n^0(\sqrt{n}(T_n-\theta) + t_2)$ converges to that of

$$(4.14) \qquad \gamma^{-1}\gamma_{22}^{1/2}\ (W^*(\sigma_\phi\ \xi),\ W^*(\sigma_\phi\ \xi + t_1),\ W^*(\sigma_\phi\ \xi + t_2)).$$

Moreover, following the lines of the proof of (3.11), we get that

$$(4.15)\ \{(W_n^0(t),\ \sqrt{n}(T_n-\theta),\ \sqrt{n}(\hat{\theta}_n-\theta)),\ |t| \le K\} \overset{\mathcal{D}}{\to} (\gamma^{-1}\gamma_{22}^{-1/2}\ W^*,\ \xi,\ \xi_0).$$

Combining (4.12), (4.13), (4.14) and (4.15), we arrive at (4.8).

Q.E.D.

<u>Remark</u>. (4.8) shows the effect of the choice of T_n as well as of $\hat{\gamma}_n$ (more precisely, of t_1 and t_2). It follows from (4.8) that the limiting distribution of Z_n^* coincides with that of $n^{1/4}R_n$ (i.e., $\hat{\theta}_n$ coincides with $\hat{\theta}_n$ up to the second order term) if and only if T_n is asymptotically equivalent to $\hat{\theta}_n$ (when $\sigma_\phi\ \xi - \frac{\sigma_0}{\gamma}\ \xi_0 = 0$ with probability 1).

References

[1] Bahadur, R.R. (1968). A note on quantiles in large samples. Ann. Math. Statist. 37, 577-581.

[2] Bickel, P.J. (1975). One-step Huber estimates in the linear model. J. Amer. Statist. Assoc. 70, 428-433.

[3] Billingsley, P. (1968). Convergence of Probability Measures. J. Wiley.

[4] Carroll, R.J. (1978). On almost sure expansions for M-estimates. Ann. Statist. 6, 314-318.

[5] Huber, P.J. (1981). Robust Statistics. J. Wiley, New York.

[6] Janssen, P., Jurečková, J. and Veraberbeke (1985). Rate of convergence on one- and two-step M-estimators with application to maximum likelihood and Pitman estimators. Ann. Statist. 13, No. 3.

[7] Jurečková, J. (1980). Asymptotic representation of M-estimators of location. Math. Operationsforsch. Statist., Series Statistics 11, 61-73.

[8] Jurečková, J. (1983). Robust estimation of location and regression parameters and their second order asymptotic relations. Trans. 9th Prague Conf. on Inform. Theory, Statist. Decis. Functions and Random Processes, 19-32. Reidel, Dordrecht.

[9] Jurečková, J. (1985). Representation of M-estimators with the second order asymptotic distribution. Statistics & Decisions, 3, in press.

[10] Jurečková, J. and Sen, P.K. (1981a). Invariance principles for some stochastic processes relating to M-estimators and their role in sequential statistical inference. Sanehyā A, 43, 190-210.

[11] Jurečková, J. and Sen, P.K. (1981b). Sequential procedures based on M-estimators with discontinuous score functions. Jour. Statist. Plan. Inf. 5, 253-266.

[12] Jurečková, J. and Sen, P.K. (1982). M-estimators and L-estimators of location: Uniform integrability and asymptotic risk-efficient sequential versions. Sequen. Anal. 1, 27-56.

[13] Jurečková, J. and Sen, P.K. (1984). On adaptive scale-equivariant M-estimators in linear models. Stat. & Dec. 2, Suppl. 31-46.

[14] Kiefer, J. (1967). On Bahadur's representation of sample quantiles. Ann. Math. Statist. 38, 1323-1342.

[15] Serfling, R.J. (1980). Approximation Theorems of Mathematical Statistics. J. Wiley, New York.