The Annals of Probability
1987, Vol. 15, No. 2, 814–823

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

BY JANA JUREČKOVÁ AND PRANAB KUMAR SEN[1]

Charles University 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*. For a general estimator $T_n = T_n(X_1, \dots, X_n)$ of θ, an *asymptotic representation* of the form

$$n^{1/2}(T_n - \theta) = n^{-1/2}\sum_{i=1}^{n} IC_T(X_i - \theta; F) + R_n; \qquad R_n = o_p(1), \tag{1.1}$$

where $IC_T(\cdot)$ stands for the *influence function* of T_n, has been studied by a host of workers [viz., Serfling (1980), Huber (1981), Sen (1981) and the references cited therein]. For a general M-estimator $\hat{\theta}_n$ of θ, we are interested in a representation of this type, supplemented by a more precise characterization of the remainder term R_n. This estimator (corresponding to a fixed scale of F) is defined as a solution of the equation

$$M_n(t) = 0, \tag{1.2}$$

and is a (weakly) consistent estimator of θ; here the estimating function is

$$M_n(t) = \sum_{i=1}^{n} \Psi(X_i - t), \qquad t \in R^1, \quad n \geq 1, \tag{1.3}$$

and $\Psi\colon R^1 \to R^1$ is a function such that

$$\lambda(0) = 0, \quad \text{where } \lambda(t) = \int_{R^1} \Psi(x-t)\,dF(x), \qquad t \in R^1. \tag{1.4}$$

Generally, if F is symmetric about 0 while Ψ is a skew-symmetric and integrable function, then $\lambda(0) = 0$, although for the maximum likelihood estimator (MLE) neither the symmetry of F nor the skew-symmetry of Ψ is needed for $\lambda(0)$ to be equal to 0. In the M-estimation of location, however, the symmetry and the skew-symmetry of Ψ are generally presumed. For nondecreasing Ψ, $\hat{\theta}_n$ may be

Received May 1985; revised February 1986.
[1]Work partially supported by Office of Naval Research Contract No. N00014-83-K-0387.
AMS 1980 *subject classifications*. 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.

written as

$$\hat{\theta}_n = \tfrac{1}{2}(\sup\{t\colon M_n(t) > 0\} + \inf\{t\colon M_n(t) < 0\}). \tag{1.5}$$

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, and we shall not go into these details. The representation in (1.1), supplemented by the order of R_n, has some important applications, including the following: (i) asymptotic relations of different types of estimators (e.g., L-, M- and R-estimators) up to various orders of equivalence [viz., Jurečková (1983, 1985a)]; (ii) rate of approximation of an estimator by its one-step version or by some other form [viz., Jurečková (1983), Jurečková and Sen (1982, 1984) and Janssen, Jurečková and Veraverbeke (1985)]; and (iii) general equivalence results in the sequential case [viz., Jurečková and Sen (1981b, 1982)]. A representation of this type for M-estimators of location was studied by Carroll (1978) for strongly consistent versions of $\hat{\theta}_n$ where Ψ has been assumed to be "smooth." 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). Janssen, Jurečková and Veraverbeke (1985) obtained an analogous result for M-estimators of general parameters (including the MLE and the Pitman estimators).

We may term (1.1) a *first-order asymptotic representation*. The next natural step would be to supplement the order of R_n by the asymptotic distribution (if any) of the same. A result of this type was derived by Kiefer (1967) in the context of the Bahadur (1966) representation of sample quantiles. Asymptotic representations supplemented by asymptotic distributions of the remainder term will be referred to as the *second-order asymptotic* (*distributional*) *representations* (SOADR); typically, the asymptotic distribution of R_n will be nonnormal. For M-estimators of location generated by smooth Ψ-functions, SOADR results have recently been studied by Jurečková (1985b), and we extend these results to possibly discontinuous Ψ-functions. In both the cases, the method used rests on weak convergence of some related M-processes involving a random change of time [viz., Billingsley (1968), Section 17].

For M-estimators, the representation (1.1) 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 regularity 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, $R_n = O_p(n^{-1/4})$. Further, let T_n be an estimator of θ admitting a first-order representation, i.e.,

$$n^{1/2}(T_n - \theta) = n^{-1/2}\sum_{i=1}^{n}\phi(X_i - \theta) + o_p(1), \tag{1.7}$$

for some suitable ϕ, for which

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

Then, for smooth Ψ, generating $\hat{\theta}_n$, Jurečková (1985b) has shown that as $n \to \infty$,

$$(1.9) \qquad n(T_n - \theta) - n\alpha(T_n - \theta)^2 + \gamma^{-1}\{M_n(T_n) - M_n(\theta)\} \to_{\mathscr{D}} \xi_1 \cdot \xi_2,$$

where $\gamma = \int_{R^1}\Psi^{(1)}(x)\,dF(x)$, $\alpha = \int_{R^1}\Psi^{(2)}(x)\,dF(x)/(2\gamma)$ and $\xi = (\xi_1, \xi_2)$ has a bivariate normal distribution. (1.9) also applies for the M-estimator $\hat{\theta}_n$, the least-squares estimator as well as the 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 n) as well as a different type of limiting law arises in this context, although the result is in correspondence with that of Kiefer (1967). Specifically, for an estimator T_n, satisfying (1.7), we have $n^{1/2}(T_n - \theta) \to_{\mathscr{D}} \xi$ and

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

where ξ has a normal distribution with 0 mean,

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

$I[A]$ stands for the indicator function of the set A, and W_1, W_2, are independent copies of a Wiener process on $[0, \infty)$.

Along with the preliminary notions, the main results are presented in Section 2 and their derivations are considered 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 SOADR of one-step versions of M-estimators and, in this context, illustrates the effect of the choice of an initial estimator.

2. A SOADR theorem. To incorporate jump discontinuities in 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 ($\Psi_1^{(1)}$ and $\Psi_1^{(2)}$, 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 $-\infty = a_0 < a_1 < \cdots < a_p < a_{p+1} = +\infty$, such that $\Psi_2(x) = \beta_j$ for $x \in E_j$, $0 \leq j \leq p$; conventionally, we let $\Psi_2(a_j) = (\beta_j + \beta_{j+1})/2$, $j = 1, \ldots, p$. Also, we assume that Ψ is nondecreasing and that $\int_{R^1}\Psi(x)\,dF(x) = 0$.

In the SOADR theorem for M-estimators, the key role is played by the asymptotic behavior of the related M-process: $M_n(\theta + n^{-1/2}t) - M_n(\theta) - E[M_n(\theta + n^{-1/2}t) - M_n(\theta)]$, $t \in T \subset R^1$, where $M_n(\cdot)$ is defined in (1.3). By virtue of (2.1), we may decompose this process as the sum of two processes generated by Ψ_1 and Ψ_2, respectively. Also, it will be convenient to replace the expectation by a part of an appropriate Taylor expansion, and this can be done under quite general regularity conditions on Ψ_1 and Ψ_2 as well as the d.f. F. Concerning the d.f. F, we assume that F possesses an absolutely continuous and symmetric density f having a finite Fisher information

$$I(f) = \int_{R^1} \{ f'(x)/f(x) \}^2 \, dF(x), \quad \text{where } f'(x) = (d/dx)f(x). \tag{2.2}$$

Also, we assume that

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

$$\int_{R^1} \{ \psi_1^{(r)}(x) \}^2 \, dF(x) < \infty, \qquad r = 1,2, \tag{2.4}$$

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

$$\int_{R^1} \{ \psi_1^{(2)}(x \pm t) \}^2 \, dF(x) < K < \infty, \qquad \forall \, |t| < \delta. \tag{2.5}$$

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

$$\gamma_{2r} = \sum_{j=1}^{p} (\beta_j - \beta_{j-1})^r f(a_j), \qquad r = 1,2, \ \gamma = \gamma_{11} + \gamma_{21}, \tag{2.6}$$

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

where $\beta_0 = \beta_{p+1} = 0$ and we assume that

$$\gamma \neq 0, \qquad \gamma_{21} \neq 0 \quad \text{and} \quad \gamma_{22} > 0. \tag{2.8}$$

Then, we have the following:

THEOREM 2.1. *For $\Psi_2 \neq 0$ and any $\{T_n\}$ satisfying* (1.7), *under the assumed regularity conditions, the r.v.*

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

where

$$\xi^* = \kappa \{ I(\xi > 0) W_1(|\xi|) + I(\xi < 0) W_2(|\xi|) \} \tag{2.10}$$

and 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 where*

$$\kappa = (\sigma_\Phi \gamma_{22})^{1/2} \gamma^{-1}. \tag{2.11}$$

Thus, Z_n has the asymptotic distribution

$$P\{\xi^* \le x\} = 2\int_0^\infty \Phi\left(x (\kappa t^{1/2})^{-1} \right) d\Phi(t), \qquad x \in R^1, \tag{2.12}$$

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

In passing, we may note that on letting $\Psi_1 \equiv 0$ and $\Psi_2(x) = p - I[x < 0]$ for some $p \in (0,1)$, we have $\int \Psi(x-t)\,dF(x) = 0$ for $t = F^{-1}(p)$. Replacing then $\Psi(x)$ by $\Psi^*(x) = \Psi(x - F^{-1}(p))$, we have $\gamma_{21} = f(F^{-1}(p)) = \gamma_{22} = \gamma$. Hence, denoting the sample order statistics by $X_{n:\,i}$, $i = 1,\dots,n$, and letting $\sigma_\phi^2 = p(1-p)$, we get the following corollary of Theorem 2.1, which coincides with the Kiefer (1967) result:

COROLLARY. *Suppose that F is twice differentiable at $F^{-1}(p)$ and $f(F^{-1}(p)) > 0$. Then*

$$\begin{aligned} &n^{1/2}\big(X_{n:[np]} - F^{-1}(p)\big) \\ (2.13)\qquad &\quad = n^{-1/2}\{f(F^{-1}(p))\}^{-1}\sum_{i=1}^{n}\big(p - I(X_i \le F^{-1}(p)]\big) + R_n(p), \end{aligned}$$

and

$$\begin{aligned} &\lim_{n\to\infty} P\{n^{1/4} f(F^{-1}(p))R_n(p) \le x\} \\ (2.14)\qquad &\quad = 2\int_0^\infty \Phi\Big(x\{t^2p(1-p)\}^{-1/4}\Big)\,d\Phi(t). \end{aligned}$$

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

(ii) For $T_n = \hat{\theta}_n$ (the M-estimator based on the same score function), 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)–(2.12) hold with $\kappa = (\sigma_0\gamma_{22}/\gamma^3)^{1/2}$.

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

3. The proof of Theorem 2.1. As has been mentioned in Section 2, we proceed to the proof through some related M-processes. Towards this, we consider the process

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

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

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

Also, consider the process $W_n^* = \{W_n^*(t),\ -K \le t \le K\}$, $K > 0$, where

$$(3.3)\qquad W_n^*(t) = \gamma\cdot\gamma_{22}^{-1/2}W_n^0(t), \quad \text{for } t \in [-K, K],$$

and let $W^* = \{W^*(t),\ t \in [-K, K]\}$ be a Gaussian process with $EW^*(t) = 0$, $\forall\ t$, and

$$(3.4)\qquad EW^*(s)W^*(t) = |s| \wedge |t|, \quad \text{if } st > 0 \text{ and } 0 \text{ otherwise}.$$

Then, from Jurečková (1980) (see the corollary on page 69), it follows that W_n^* converges weakly to W^* in the Skorokhod J_1-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 $\{W_n^0(t),\ n^{1/2}(T_n - \theta),\ t \in [-K, K]\}$. Note that by (1.7) and (1.8),

$$n^{1/2}(T_n - \theta) = \sum_{i=1}^{n} U_{ni} + o_p(1), \tag{3.5}$$

where

$$\begin{gathered} U_{ni} = n^{-1/2}\phi(X_i - \theta), \\ EU_{ni} = 0 \quad \text{and} \quad EU_{ni}^2 = n^{-1}\sigma_\phi^2, \quad i = 1, \dots, n. \end{gathered} \tag{3.6}$$

First, we consider the case of $\Psi \equiv \Psi_2$. Then, by the definition of Ψ_2 and (1.3),

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

where

$$\begin{aligned} U_{ni}^*(t) = n^{-1/4} &\sum_{j=1}^{p} (\beta_j - \beta_{j-1}) \\ &\times\Big(\big\{I[t < 0]I\big[a_j + n^{-1/2}t \le X_i - \theta \le a_j\big] \\ &\quad -I[t > 0]I\big[a_j \le X_i - \theta \le a_j + n^{-1/2}t\big]\big\}\Big), \qquad i = 1, \dots, n, \end{aligned} \tag{3.8}$$

are i.i.d.r.v. and n is so large that $a_{j+1} - a_j > n^{-1/2}K$, for $j = 0, 1, \dots, p$. Then

$$\begin{gathered} \operatorname{Var}(U_{ni}^*(t)) = |t|n^{-1}\gamma_{22} + o(n^{-1}), \\ \operatorname{Cov}(U_{ni}, U_{ni}^*(t)) = E(U_{ni}U_{ni}^*(t)) = o(|t|n^{-3/4}), \qquad i = 1, \dots, n. \end{gathered} \tag{3.9}$$

Hence,

$$\operatorname{Cov}\big(n^{-1/4}[M_n(\theta + n^{-1/2}t) - M_n(\theta)],\ n^{1/2}(T_n - \theta)\big) = |t| \cdot o(1) \to 0, \quad \text{as } n \to \infty. \tag{3.10}$$

The convergence of the finite-dimensional distributions of $\{W_n^0, n^{1/2}(T_n - \theta)\}$ to those of $\{(\gamma^{-1}\gamma_{22}^{1/2}W^*, \xi^0)\}$ (where $\xi^0 \sim \mathcal{N}(0, \sigma_\phi^2)$) follows readily from (3.5)–(3.10). Also, $n^{1/2}(T_n - \theta)$ is relatively compact, while the weak convergence of W_n^* in (3.3) insures the relative compactness of W_n^0 as well. Hence, as $n \to \infty$,

$$\{n^{1/2}(T_n - \theta), W_n^0\} \to_{\mathscr{D}} \{\xi^0, \gamma^{-1}\gamma_{22}^{1/2}W^*\}. \tag{3.11}$$

Since, for any $\eta > 0$, there exists a K $(0 < K < \infty)$, such that $P(|\xi^0| \le K) \ge 1 - \eta$, we may use (3.11) and apply the random change of time: $t \to [n^{1/2}(T_n - \theta)]^K$ (where $[Y]^K = Y$ if $|Y| \le K$ and is equal to 0, otherwise). Thus, using the results in Section 17 [cf. (17.5)–(17.9)] of Billingsley (1968), we obtain that as $n \to \infty$,

$$Z_n = W_n^0(n^{1/2}(T_n - \theta)) \to_{\mathscr{D}} \gamma^{-1}\gamma_{22}^{1/2}W^*(\xi^0) = (\sigma_\phi\gamma_{22})^{1/2}\gamma^{-1}W^*(\xi), \tag{3.12}$$

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 the W_j are defined as in the theorem. 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 vanishes, we may define

$$(3.14)\qquad \begin{aligned} W_{n1}^0(t) &= n^{-1/4}\gamma^{-1}\sum_{i=1}^{n}\left[\Psi_1(X_i - \theta - tn^{-1/2}) - \psi_1(X_i-\theta)\right] \\ &\quad + tn^{1/2} - (\gamma_{12}/2\gamma)t^2, \qquad t \in R^1, \end{aligned}$$

as the component of $W_n^0(t)$ corresponding to Ψ_1, and using Lemma 3.1 of Jurečková (1985b), we obtain

$$(3.15)\qquad \sup\left\{|W_{n1}^0(t)|: |t| \le K\right\} = O_p(n^{-1/4}), \quad \text{as } n \to \infty.$$

Thus the weak convergence of $\{n^{1/2}(T_n - \theta), W_n^0\}$ follows from (3.11) and (3.15), and this completes the proof of the theorem. □

4. SOADR for one-step versions of *M*-estimators. By Remark (ii) in Section 2, for $T_n = \hat{\theta}_n$ (the *M*-estimator), we have

$$(4.1)\qquad n^{1/4}R_n = n^{-1/4}\left\{n\left(\hat{\theta}_n - \theta\right) - \gamma^{-1}M_n(\theta)\right\} \to_{\mathscr{D}} \xi_0^*,$$

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

$$(4.2)\qquad \xi_0^* = \kappa^*\{I[\xi>0]W_1(|\xi|) + I[\xi<0]W_2(|\xi|)\},$$

where W_1, W_2 and ξ are defined as in Theorem 2.1 and

$$(4.3)\qquad \kappa^* = \left(\sigma_0\gamma_{22}\gamma^{-3}\right)^{1/2}.$$

It is often convenient to approximate $\hat{\theta}_n$ by its one-step version $\hat{\theta}_n^*$ of the following form: Starting with an initial estimator T_n satisfying $n^{1/2}(T_n - \theta) = O_p(1)$, we put

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

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

$$(4.5)\qquad \hat{\gamma}_n = \left(n^{1/2}(t_2 - t_1)\right)^{-1}\left\{M_n\left(T_n + n^{-1/2}t_2\right) - M_n\left(T_n + n^{-1/2}t_1\right)\right\},$$

where t_1, t_2 $(t_1 < t_2)$ are some arbitrary real numbers; often, we let $t_2 = -t_1 = t$ (>0). The one-step estimators were first considered by Bickel (1975) and later on studied by Jurečková (1983) [see also Janssen, Jurečková and Veraverbeke (1985) for a more general setup]: It was shown that whenever $\Psi_2 \not\equiv 0$,

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

for any $n^{1/2}$-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

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

Towards this, we have the following theorem where the Gaussian process W^* is defined as in (3.4) and the constants κ and σ_ϕ as in (2.11) and (1.8), respectively.

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

$$\begin{aligned} Z_n^* \to \xi_0^* &- (t_2 - t_1)^{-1}\kappa[W^*(\xi + t_1/\sigma_\phi) - W^*(\xi + t_2/\sigma_\phi)] \\ &\times[\xi\sigma_\phi - \gamma^{-1}\xi_0\sigma_0] \\ &= \kappa\big[W^*(\xi) - (t_2 - t_1)^{-1}\{W^*(\xi + t_1/\sigma_\phi) - W^*(\xi + t_2/\sigma_\phi)\} \\ &\qquad\times\{\xi\sigma_\phi - \gamma^{-1}\xi_0\sigma_0\}\big], \end{aligned} \tag{4.8}$$

where (ξ, ξ_0) *has the bivariate normal distribution with null mean vector and dispersion matrix* $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ *with*

$$\rho = \left(\int_{R^1}\phi(x)\Psi(x)\,dF(x)\right)\Big/\left\{\int_{R^1}\phi^2(x)\,dF(x)\int_{R^1}\Psi^2(x)\,dF(x)\right\}^{1/2}. \tag{4.9}$$

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

$$\begin{aligned} V_n &= n^{1/4}(\hat{\gamma}_n - \gamma) \\ &= n^{1/4}\big\{\big[n^{1/4}\big(W_n^0(n^{1/2}(T_n - \theta) + t_1) - W_n^0(n^{1/2}(T_n - \theta) + t_2)\big) \\ &\qquad + \gamma n^{1/2}(t_2 - t_1)\big]n^{-1/2}(t_2 - t_1)^{-1} - \gamma\big\} + o_p(1). \end{aligned} \tag{4.10}$$

Also, by the Slutsky theorem,

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

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

$$\begin{aligned} Z_n^* &= n^{1/4}R_n + \gamma^{-1}n^{1/4}(\gamma\hat{\gamma}_n^{-1} - 1)n^{-1/2}\big[M_n(T_n) - M_n(\hat{\theta}_n)\big] + o_p(1) \\ &= W_n^0(n^{1/2}(T_n - \theta)) + V_n^*\{n^{1/2}(T_n - \hat{\theta}_n)\} + o_p(1) \\ &= W_n^0(n^{1/2}(T_n - \theta)) - (t_2 - t_1)^{-1}\{n^{1/2}(T_n - \theta)\} \\ &\quad\times\big[W_n^0(n^{1/2}(T_n - \theta) + t_1) - W_n^0(n^{1/2}(T_n - \theta) + t_2)\big] + o_p(1). \end{aligned} \tag{4.12}$$

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

$$n^{1/2}(T_n - \hat{\theta}_n) = \sigma_\phi(n^{1/2}(T_n - \theta)/\sigma_\phi) - \gamma^{-1}\sigma_0(n^{1/2}(\hat{\theta}_n - \theta)\gamma/\sigma_0) \tag{4.13}$$

coincides with that of $\xi\sigma_\phi - \gamma^{-1}\xi_0\sigma_0$, where ξ and ξ_0 are defined as before in (4.9).

It follows from the proof of Theorem 2.1 (in Section 3) that the joint (trivariate) distribution of

$$W_n^0\big(n^{1/2}(T_n - \theta)\big), W_n^0\big(n^{1/2}(T_n - \theta) + t_1\big), W_n^0\big(n^{1/2}(T_n - \theta) + t_2\big)$$

converges to that of

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

Moreover, following the lines of the proof of (3.11), we obtain

$$\begin{aligned} &\{n^{1/2}(T_n - \theta), n^{1/2}(\hat{\theta}_n - \theta), \{W_n^0(t), t \in [-K, K]\}\} \\ &\quad \to_{\mathscr{D}} (\xi, \xi_0, \gamma^{-1}\gamma_{22}^{-1/2}W^*), \end{aligned} \tag{4.15}$$

where W^* is defined as before in (3.4). Combining (4.12), (4.13), (4.14) and (4.15), we arrive at (4.8). □

REMARK. (4.8) shows the effect of the choice of the initial estimator T_n as well as of $\hat{\gamma}_n$ in (4.5) (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 $\hat{\theta}_n$ (in the first order), and this happens when $\xi\sigma_\phi - \gamma^{-1}\xi_0\sigma_0 = 0$ with probability 1 or ρ in (4.9) is equal to 1, i.e., $\phi = \Psi$.

Acknowledgments. The authors are grateful to the Associate Editor and the referee for their critical reading of the manuscript and valuable comments which have led to a considerable improvement of the original text.

REFERENCES

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–581.

BICKEL, P. J. (1975). One-step Huber estimates in the linear model. *J. Amer. Statist. Assoc.* **70** 428–433.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

CARROLL, R. J. (1978). On almost sure expansion for *M*-estimates. *Ann. Statist.* **6** 314–318.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

JANSSEN, P., JUREČKOVÁ, J. and VERAVERBEKE, N. (1985). Rate of convergence of one- and two-step *M*-estimators with application to maximum likelihood and Pitman estimators. *Ann. Statist.* **13** 1222–1229.

JUREČKOVÁ, J. (1980). Asymptotic representation of *M*-estimators of location. *Math. Operationsforsch. Statist. Ser. Statist.* **11** 61–73.

JUREČKOVÁ, J. (1983). Robust estimation of location and regression parameters and their second order asymptotic relations. *Transactions of the Ninth Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*. 19–32. Reidel, Dordrecht.

JUREČKOVÁ, J. (1985a). Robust estimators of location and their second order asymptotic relations. In *A Celebration of Statistics: The ISI Centenary Volume* (A. C. Atkinson and S. E. Fienberg, eds.) 377–392. Springer, New York.

JUREČKOVÁ, J. (1985b). Representation of *M*-estimators with the second order asymptotic distribution. *Statist. Dec.* **3** 263–276.

JUREČKOVÁ, J. and SEN, P. K. (1981a). Invariance principles for some stochastic processes relating to *M*-estimators and their role in sequential statistical inference. *Sankhyā Ser. A* **43** 193–210.

JUREČKOVÁ, J. and SEN, P. K. (1981b). Sequential procedures based on M-estimators with discontinuous score functions. *J. Statist. Plann. Inference* **5** 253–266.

JUREČKOVÁ, J. and SEN, P. K. (1982). M-estimators and L-estimators of location: Uniform integrability and asymptotically risk-efficient sequential versions. *Sequential Anal.* **1** 27–56.

JUREČKOVÁ, J. and SEN, P. K. (1984). On adaptive scale-equivariant M-estimators in linear models. *Statist. Dec. Suppl.* **1** 31–46.

KIEFER, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.

SEN, P. K. (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference.* Wiley, New York.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

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

DEPARTMENT OF BIOSTATISTICS 201H
UNIVERSITY OF NORTH CAROLINA,
CHAPEL HILL, NORTH CAROLINA 27514