SEQUENTIAL PROCEDURES BASED ON M-ESTIMATORS WITH DISCONTINUOUS SCORE FUNCTIONS

by

JANA JUREČKOVÁ
Department of Probability & Statistics
Charles University
Prague 8, Czechoslovakia

and

PRANAB KUMAR SEN
Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514, U.S.A.

Institute of Statistics Mimeo Series No. 1279

APRIL 1980

SEQUENTIAL PROCEDURES BASED ON M-ESTIMATORS WITH DISCONTINUOUS SCORE FUNCTIONS

JANA JUREČKOVÁ[1)] and PRANAB KUMAR SEN[2)]

Department of Probability & Statistics
Charles University
Prague 8, Czechoslovakia

Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514, U.S.A.

ABSTRACT

Bounded-length sequential confidence intervals and sequential tests for regression parameter based on M-estimators are extended to the case where the score-functions generating the M-estimators have jump-discontinuities. In the context of the asymptotic normality of the stopping variable, for the confidence interval problem, it is observed that the jump-discontinuities induce a slower rate of convergence. The proofs of the main theorems rest on the weak convergence of some related processes, which is also studied.

AMS Subject Classification: 62L10, 62L12, 60F17.

Key Words & Phrases: Bounded-length confidence intervals, jump-discontinuity, M-estimator, regression (location) parameter, score functions, stopping variables, sequential test, weak convergence.

[1)] This work was done while this author was visiting the University of North Carolina, Chapel Hill, NC, U.S.A.

[2)] Work of this author was partially supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243 from the National Institutes of Health.

1. INTRODUCTION

Let $\{X_i, i \geq 1\}$ be a sequence of independent random variables (r.v.) with continuous distribution functions (d.f.) $\{F_i, i \geq 1\}$ where

$$F_i(x) = F(x - \Delta c_i), \ i \geq 1, \quad x \in E = (-\infty, \infty), \tag{1.1}$$

the c_i are *known (regression) constants* (not all equal to 0), Δ is an *unknown parameter* and the unknown d.f. F belongs to some class $\mathcal{F}$ of d.f.'s. If all the c_i are equal to 1, (1.1) reduces to the classical *location model*, while, for the c_i not all equal, this relates to the so called *regression model*. We intend to construct (i) a *confidence interval* for Δ having the *coverage probability* $1 - \alpha$ and *width* $\leq 2d$ and (ii) a test for

$$H_0: \Delta = \Delta_0 (\text{specified}) \text{ vs. } H_1: \Delta = \Delta_1 (\text{specified}) > \Delta_0, \tag{1.2}$$

having type I and type II errors bounded by α_1 and α_2, respectively, where $d(> 0)$, $\alpha_1 (0 < \alpha_1 < 1)$ and $\alpha_2 (0 < \alpha_2 < 1)$ are all preassigned numbers.

For unspecified F, for either of these problems, no *fixed-sample size procedure* exists, but under an *asymptotic* (as $d \downarrow 0$ or as $\Delta_1 - \Delta_0 \downarrow 0$) *set-up*, a possible solution is to draw observations *sequentially* with respect to some appropriate *stopping rules*. For the confidence interval problem, such procedures based on the *least squares* and *rank order estimators* are due to Chow and Robbins (1965), Gleser (1965), Geertsema (1968), Sen and Ghosh (1971) and Ghosh and Sen (1972), among others. For the sequential testing problem, rank based procedures are due to Sen and Ghosh (1974) and Ghosh and Sen (1976, 1977), among others. For the location model, a sequential confidence interval based on *M-estimators* is due to Carroll (1977), while Jurečková and Sen (1980) have studied the general regression model, both the confidence interval and the testing problem, under less restrictive regularity conditions. In both these papers, the *score-function* (ψ) [generating the M-estimators] is assumed to be absolutely continuous and some extra regularity conditions are needed for the *asymptotic normality* of the *stopping numbers*.

M-estimators corresponding to score-functions having *jump-discontinuities* constitute an important class of estimators (the sample median belongs to the same). The object of the current investigation is to study the asymptotic theory of sequential confidence intervals and tests for Δ based on M-estimators when the generating score functions have jump-discontinuities. For the location model, Jurečková (1980) has observed some qualitative difference in the *rate of convergence* of M-estimators corresponding to absolutely continuous ψ and ψ having jump-discontinuities. As will be seen later on, the same feature is generally true for the regression model and analogous qualitative difference appears in the rates of convergence of the stopping times.

Along with the preliminary notions, the sequential confidence interval procedure is presented in Section 2. Section 3 deals with some *invariance principles* relating to M-estimators corresponding to ψ having jump-discontinuities and, in Section 4, these are incorporated in the proofs of the main theorems of Section 2. The last section is devoted to the study of the asymptotic theory of sequential tests for Δ based on M-estimators, where, also, the results of Section 3 are utilized in the derivation of the main results.

2. THE SEQUENTIAL CONFIDENCE INTERVAL FOR Δ

Based on $X_1,\ldots,X_n$, an *M-estimator* $\hat{\Delta}_n$ of Δ [in (1.1)] [c.f. Huber (1973), Jurečková (1977), and others] is a solution of

$$S_n(t) = \sum_{i=1}^{n} c_i \psi(X_i - tc_i) \ (=0), \tag{2.1}$$

where ψ is some *score-function*. We assume that

$$\psi(x) = \psi_1(x) + \psi_2(x), \quad x \in E, \tag{2.2}$$

where ψ_1 is absolutely continuous on any bounded interval in E, it possesses first and second order derivatives ($\psi_1^{(1)}$ and $\psi_1^{(2)}$, respectively) almost everywhere (a.e.), and ψ_2 is a step-function, defined as follows. For some positive integer p, assume that there exist open-intervals $E_j = (a_j, a_{j+1})$, $j = 0,\dots,p$ with $a_0 = -\infty < a_1 < \cdots < a_p < a_{p+1} = +\infty$, such that

$$\psi_2(x) = \beta_j \quad \text{for} \quad x \in E_j, \quad 0 \le j \le p, \tag{2.3}$$

where the β_j are real numbers (not all equal). Conventionally, we let

$$\psi_2(a_j) = \tfrac{1}{2}(\beta_{j-1} + \beta_j), \quad \text{for} \quad j = 1,\dots,p \ . \tag{2.4}$$

Also, we assume that both ψ_1, ψ_2 (and hence, ψ) are nondecreasing and skew-symmetric, so that

$$\psi_j(-x) = -\psi_j(x), \ \forall\ x \in E,\ j = 1,\ 2. \tag{2.5}$$

The d.f. F in (1.1) is assumed to be symmetric, so that

$$F(x) + F(-x) = 1, \ \forall\ x \in E. \tag{2.6}$$

Note that by (2.5) and (2.6),

$$\int_{-\infty}^{\infty} \psi_j(x)dF(x) = 0, \quad j = 1,\ 2, \tag{2.7}$$

so that by (2.1) and (2.7),

$$E_\Delta S_n(\Delta) = 0, \ \forall\ n \ge 1, \tag{2.8}$$

where E_Δ stands for the expectation when Δ holds. Further, for every $n (\ge 1)$,

$$S_n(t) \quad \text{is} \searrow \text{in} \quad t \in E. \tag{2.9}$$

Moreover, we assume that F possesses an absolutely continuous probability density function (p.d.f.) f having a finite Fisher information

$$I(f) = \int_{-\infty}^{\infty} \{f'(x)/f(x)\}^2 dF(x) \quad (< \infty), \tag{2.10}$$

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

$$\text{(i)} \quad \gamma_{1r} = \int_{-\infty}^{\infty} \psi_1^{(r)}(x)dF(x) \quad \text{exists, for } r = 1, 2, \tag{2.11}$$

$$\text{(ii)} \quad \int_{-\infty}^{\infty} [\psi_1^{(1)}(x)]^2 dF(x) < \infty, \tag{2.12}$$

$$\text{and (iii)} \quad \lim_{t\to 0} \int_{-\infty}^{\infty} \{\psi_1^{(1)}(x+t) - \psi_1^{(1)}(x)\}^2 dF(x) = 0. \tag{2.13}$$

Note that (2.10) insures that $\int |f'(x)|dx < \infty$. However, we need a slightly more stringent condition that for each $j(=1,\ldots,p)$, in some neighborhood of a_j, f' is bounded and

$$\gamma_{2r} = \sum_{j=1}^{p} (\beta_j - \beta_{j-1})^r f(a_j) > 0, \quad \text{for } r = 1, 2. \tag{2.14}$$

Finally, we assume that

$$0 < \sigma_0^2 < \int_{-\infty}^{\infty} \psi^2(x)dF(x) < \infty ; \tag{2.15}$$

and that as $n \to \infty$,

$$\left(\max_{1\le k\le n} c_k^2\right)/C_n^2 \to 0, \quad \text{where} \quad C_n^2 = \sum_{i=1}^{n} c_i^2 . \tag{2.16}$$

<u>Remark</u>. If $\psi \equiv \psi_2$ (i.e., $\psi_1 \equiv 0$) then the assumptions (2.10) - (2.13) are not needed; in this case (2.14) and the boundedness of f' in some neighborhoods of the a_j suffice. Moreover, if ψ_1 is constant outside a bounded interval $[a, b] \subset E$ (most of the functions suggested for M-estimators are of this type), then, for (2.11) - (2.13), it suffices to assume that $\psi_1^{(2)}$ is bounded inside (a, b).

Regarding (2.8) and (2.9), $\hat{\Delta}_n$ (in (2.1)) may be formally written as $\hat{\Delta}_n = (\hat{\Delta}_n^* + \hat{\Delta}_n^{**})/2$, where

$$\hat{\Delta}_n^* = \sup\{t: S_n(t) > 0\} \quad \text{and} \quad \hat{\Delta}_n^{**} = \inf\{t: S_n(t) < 0\}. \tag{2.17}$$

Moreover, we note that by (2.8), (2.15) and (2.16),

$$L(C_n^{-1} S_n(\Delta)) \longrightarrow N(0, \sigma_0^2) \quad \text{as} \quad n \to \infty. \tag{2.18}$$

σ_0^2 given by (2.15) depends on the unknown distribution function F. We shall estimate σ_0^2 by the sequence

$$s_n^2 = \int_{-\infty}^{\infty} \psi^2(x) d\hat{F}_n(x), \quad n \geq 1 \tag{2.19}$$

with

$$\hat{F}_n(x) = n^{-1} \sum_{i=1}^n u(x - X_i - \hat{\Delta}_n c_i), \; x \in E; \; n = 1, 2, \ldots , \tag{2.20}$$

where $u(t) = 1$ if $t \geq 0$ and $u(t) = 0$ otherwise. Finally, denote

$$\hat{\Delta}_n^- = \sup\{t: S_n(t) > C_n s_n \tau_{\alpha/2}\} \tag{2.21}$$

$$\hat{\Delta}_n^+ = \inf\{t: S_n(t) < -C_n s_n \tau_{\alpha/2}\} \tag{2.22}$$

and

$$I_n = (\hat{\Delta}_n^-, \hat{\Delta}_n^+), \quad L_n = \hat{\Delta}_n^+ - \hat{\Delta}_n^- \tag{2.23}$$

where Φ is the standard normal d.f. and $\Phi(\tau_{\alpha/2}) = 1 - \frac{\alpha}{2}$, $0 < \alpha < 1$. Notice that $L_n \geq 0 \;\; \forall\, n \geq 1$ [by (2.9) and (2.21) - (2.22)].

The proposed sequential confidence interval for Δ is then I_{N_d} corresponding to the *stopping variable*,

$$N_d = \inf\{n \geq n_0: L_n \leq 2d\} \tag{2.24}$$

where n_0 is an initial sample size $(n_0 \geq 2)$ and $d > 0$. It follows from (2.24) that $L_{N_d} \leq 2d$. The following theorems give the asymptotic properties of N_d and of I_{N_d} as $d \downarrow 0$. Define

$$n_d = \inf\{n \geq n_0: dC_n \geq \tau_{\alpha/2}\sigma_0/\gamma\}, \quad \gamma = \gamma_n + \gamma_{21} (> 0). \tag{2.25}$$

Theorem 2.1. *Under assumptions made above, for every (fixed)* $d > 0$, $P_\Delta(N_d < \infty) = 1$, $\lim_{d \downarrow 0} N_d = +\infty$ a.s., *and, as* $d \downarrow 0$,

$$C_{n_d}^{-1} C_{N_d} \longrightarrow 1 \quad \textit{in probability},$$

$$P_\Delta(\Delta \varepsilon I_{N_d}) \longrightarrow 1 - \alpha. \tag{2.27}$$

Corollary 2.1. *If, in addition to the hypothesis of Theorem 2.1, for some* $a < 1 < b$ *and every* $t \varepsilon [a, b]$,

$$\lim_{n\to\infty} C_{[nt]}^2 / C_n^2 = \rho(t) \quad \textit{exists and is} \uparrow \textit{in} \quad t, \tag{2.28}$$

then (2.26) *may also be replaced by: as* $d \downarrow 0$,

$$n_d^{-1} N_d \longrightarrow 1 \quad \textit{in probability}. \tag{2.29}$$

Note that for the location and the equispaced regression models, (2.28) holds for $\rho(t) = t$ and t^3, respectively. To prove the asymptotic distribution of the stopping variable N_d, we need the following assumption on the sequence $\{c_n\}_{n=1}^{\infty}$:

$$\lim_{n\to\infty} \kappa_n = \kappa(\varepsilon[1, \infty)) \quad \text{and} \quad \lim_{n\to\infty} \max_{1 \le k \le n} c_k^4 / C_n^{**} = 0 \tag{2.30}$$

where

$$\kappa_n^2 = \sqrt{n}\, C_n^* / C_n^3, \quad C_n^3 = \textstyle\sum_{i=1}^n |c_i|^3 \text{ and } C_n^{**} = \sum_{i=1}^n c_i^4 . \tag{2.31}$$

Notice that $\kappa_n \ge 1 \;\; \forall\; n \ge 1$.

Finally, we assume that

$$\int_{-\infty}^{\infty} \psi^4(x)\, dF(x) < \infty. \tag{2.32}$$

Theorem 2.2. *Under* (2.30), (2.32) *and the assumptions of* Theorem 2.1,

$$\lim_{d \downarrow 0} P_\Delta\{n_d^{\frac{1}{4}}(dC_{N_d} - \theta)/\theta \le y\} = \Phi((2\theta/\gamma_{22})^{\frac{1}{2}} y\gamma/\kappa) \tag{2.33}$$

holds for all $y \varepsilon E$, *where* $\theta = \tau_{\alpha/2}\sigma_0/\gamma$ *and* γ_{22} *is defined by* (2.14).

Corollary 2.2. *If, in addition to the hypothesis of Theorem 2.2,*

$$(C_n^{-1}C_m - 1)(n^{-1}m - 1) \longrightarrow \beta \quad as \quad n \to \infty \tag{2.34}$$

holds for some β, $0 < \beta < \infty$, *whenever* $|(m/n) - 1| < \varepsilon_n$ *for any* $\{\varepsilon_n\}$: $\lim_{n\to\infty} \varepsilon_n = 0$, *then, for every* $y \in E$,

$$\lim_{d\downarrow 0} P_\Delta\{n_d^{\frac{1}{4}}(n_d^{-1}N_d - 1) \le y\} = \Phi((2\theta/\gamma_{22})^{\frac{1}{2}}\beta\gamma/\kappa). \tag{2.35}$$

The proofs of the theorems and corollaries are presented in Section 4. It may be noted that (2.34) holds for the simple location and the equispaced regression models with $\beta = 1$ and $\frac{3}{2}$, respectively; (2.30) also holds in these cases.

The convergence properties in (2.26) - (2.29) hold without change even if $\psi_2 \equiv 0$. However, Theorem 2.2 does not extend to the case $\psi_2 \equiv 0$ (i.e. $\psi = \psi_1$): while in the case $\psi \equiv \psi_1$ has $n_d^{\frac{1}{2}}\left(\frac{N_d}{n_d} - 1\right)$ a nondegenerate asymptotic distribution as $d \downarrow 0$ (see Jurečková and Sen (1980)), in the current case, where $\psi_2 \not\equiv 0$, we have $n_d^{\frac{1}{4}}\left(\frac{N_d}{n_d} - 1\right)$ asymptotically normal as $d \downarrow 0$. Hence, jump-discontinuities in the score function lead to slower rates of convergence of the stopping numbers.

3. SOME INVARIANCE PRINCIPLES

The results of this section will provide the main tool for the proofs of Theorems 2.1 and 2.2.

Let $\{c_{ni}: i \ge 0, n \ge 0\}$ and $\{d_{ni}: i \ge 0, n \ge 0\}$ (with $c_{n0} = d_{n0} = 0$ $\forall\, n \ge 0$) be two triangular arrays of real numbers and let $\{k_n\}$ be any sequence of positive integers such that $k_n \to \infty$ as $n \to \infty$ and that

$$D_n^2 = \sum_{j\le k_n} d_{nj}^2 \le D^2 < \infty \quad \forall n \tag{3.1}$$

and

$$\lim_{n\to\infty}\left\{\max_{1\le k\le k_n} |c_{nk}^r||d_{nk}^s| \Big/ \textstyle\sum_{j\le k_n} |c_{nj}^r d_{nj}^s|\right\} = 0 \tag{3.2}$$

for every $r=0, 2$ and $s=0, 1, 2$. Moreover, denote

$$C_{nk}^2 = \sum_{j\le k} c_{nj}^2, \quad A_{nk}^2 = \sum_{j\le k} c_{nj}^2 d_{nj}^2, \quad B_{nk}^2 = \sum_{j\le k} c_{nj}^2 |d_{nj}| \tag{3.3}$$

for $k \ge 0$, $n \ge 0$, and write

$$C_n^2 = c_{nk_n}^2, \quad A_n^2 = A_{nk_n}^2, \quad B_n^2 = B_{nk_n}^2. \tag{3.4}$$

It follows from (3.2) that

$$A_n/B_n \to 0 \text{ and } B_n/C_n \to 0 \text{ as } n \to \infty. \tag{3.5}$$

Let $X_1, X_2, \ldots$ be i.i.d. random variables distributed according to d.f. F and let F, ψ satisfy the regularity conditions of Section 2. Consider the process

$$W_n(s, t) = \{[Q_{n,n(s)}(t) - EQ_{n,n(s)}(t)]/(B_n\sqrt{\gamma_{22}}): \quad (s, t) \in K^*\} \tag{3.6}$$

where

$$K^* = \{(s, t): 0 \le s \le 1, |t| \le k^*\} (= [0, 1] \times (-k^*, k^*)) \tag{3.7}$$

is a compact set in E^2,

$$n(s) = \max\{k: B_{nk}^2 \le sB_n^2\}, \; 0 \le s \le 1, \tag{3.8}$$

$$Q_{nk}(t) = \sum_{j\le k} c_{nj}[\psi(X_j - td_{nj}) - \psi(X_j)], \; k \ge 0, \; t \in E \tag{3.9}$$

and γ_{22} is defined in (2.14); suppose that $\gamma_{22} > 0$ (i.e., $\psi_2 \not\equiv 0$). Then $W_n(s, t)$ belongs to $D[K^*]$ for every n. The following theorem states the weak convergence of W_n to some two-parameter Gaussian process.

<u>Theorem 3.1</u>. *Let* $\{W_n(s, t): (s, t) \in K^*\}$ *be the process defined in* (3.6) - (3.9). *Then, under* (3.1), (3.2) *and for* ψ *and* F *satisfying the regularity conditions of Section 2,* W_n *converges weakly to the Gaussian process* $\{W(s, t): (s, t) \in K^*\}$ *such that* $W(s, t) = 0$ $\forall\ (s, t) \in K^*$ *and*

$$EW(s, t) \cdot W(s', t') = (s \wedge s') \cdot g(t, t'); \; \forall (s,t), (s', t') \in K^* \tag{3.10}$$

where

$$g(t, t') = \begin{cases} |t| \wedge |t'| & \text{if } \; t \cdot t' > 0 \\ 0 & \textit{otherwise.} \end{cases} \tag{3.11}$$

Proof:: Write $Q_{nk}(t) = Q_{nk}^{(1)}(t) + Q_{nk}^{(2)}(t)$ where $Q_{nk}^{(i)}(t)$ is generated by ψ_i according to (3.9), $i = 1, 2$. Denote

$$W_n^{(1)}(s, t) = \{[Q_{n,n(s)}^{(1)}(t) - EQ_{n,n(s)}^{(1)}(t)]/(\sigma_1 A_n) : (s, t) \varepsilon K^*\} \tag{3.12}$$

where

$$\sigma_1^2 = \int_{-\infty}^{\infty} (\psi_1^{(1)}(x))^2 dF(x) - \gamma_{11}^2 \quad (< \infty, \text{by } (2.12)). \tag{3.13}$$

Then

$$W_n(s, t) = (\sigma_1/\sqrt{\gamma_{22}})(A_n/B_n) W_n^{(1)}(s, t) + W_n^{(2)}(s, t) \tag{3.14}$$

where $W_n^{(2)}$ is defined by (3.6) and (3.9) with $\psi = \psi_2$. It follows from Theorem 2.1 of Jurečková and Sen (1980) that

$$\sup_{(s,t)\varepsilon K^*} |W_n^{(1)}(s, t)| = 0_p(1) \tag{3.15}$$

so that by (3.5), (3.11) and (3.15),

$$\sup_{(s,t)\varepsilon K^*} |W_n(s, t) - W_n^{(2)}(s, t)| \xrightarrow{p} 0 \text{ as } n \to \infty. \tag{3.16}$$

Thus, the asymptotic behavior of W_n is solely determined by $W_n^{(2)}$ generated by the discontinuous component ψ_2. Hence, to prove the theorem, it suffices to show that (a) the finite dimensional distributions of $W_n^{(2)}$ converge to those of W and (b) $\{W_n^{(2)}\}$ is tight.

It follows from the definition of ψ_2 that

$$W_n^{(2)}(s, t) = \frac{1}{B_n\sqrt{\gamma_{22}}} \sum_{i \le n(s)} c_{ni} \sum_{j=1}^{p} (\beta_j - \beta_{j-1}) [I(X_i < a_j < X_i - td_{ni}) - I(X_i - td_{ni} < a_j < X_i)]. \tag{3.17}$$

where $I(A)$ stands for the indicator function of the set A. Note that

$$E[I(X_i < a_j < X_i + |td_{ni}|)]^k = E[I(X_i - |td_{ni}| < a_j < X_i)]^k + o(|td_{ni}|)$$
$$= f(a_j)|td_{ni}| + o(|td_{ni}|), \; i \geq 1, \; k \geq 1, \quad (3.18)$$

$$E[I(X_i < a_j < X_i + |td_{ni}|) \cdot I(X_i < a_j < X_i + |t'd_{ni}|)]$$
$$= f(a_j)(|t| \wedge |t'|)|d_{ni}| \quad (3.19)$$

and

$$E[I(X_i - |td_{ni}| < a_j < X_i) \cdot I(X_i < a_j < X_i + |t'd_{ni}|)] = 0, \; i \geq 1. \quad (3.20)$$

(3.18) - (3.20) imply

$$\lim_{n\to\infty} EW^{(2)}(s, t)\, W_n^{(2)}(s', t') = (s \wedge s')g(t, t') \quad (3.21)$$

where $g(t, t')$ is defined by (3.11); thus the limiting covariances of $W_n^{(2)}$ coincide with those of W.

Moreover, any linear combination of $\{W_n^{(2)}(s_r, t_r) : 1 \leq r \leq m\}$ can be expressed through (3.17) as a sum of independent random variables and the convergence of finite dimensional distributions of $W_n^{(2)}$ then follows from the classical central limit theorem.

To prove the *tightness* of $\{W_n^{(2)}\}$, for any pair of blocks $B = \{(s_1, t_1), (s_2, t_2)\}$, $B' = \{(s_2, t_1), (s_3, t_2)\}$ such that $s_1 < s_2 < s_3$, and $t_1 < t_2$, consider the increments

$$W_n^{(2)}(B) = W_n^{(2)}(s_2, t_2) - W_n^{(2)}(s_2, t_1) - W_n^{(2)}(s_1, t_2) + W_n^{(2)}(s_1, t_2); \quad (3.22)$$

with $W_n^{(2)}(B')$ defined analogously. It follows from (3.18) - (3.20)

$$\overline{\lim_{n\to\infty}}\, E\{[W_n^{(2)}(B)]^2 [W_n^{(2)}(B')]^2\} \leq K[\lambda(B \cup B')]^2 \quad (3.23)$$

where $\lambda(A)$ denotes the Lebesque measure of A and K is a finite positive constant, independent of $(s_1, s_2, s_3, t_1, t_2)$. Analogous inequality holds for increments in $B = \{(s_1, t_1), (s_2, t_2)\}$, $B' = \{(s_1, t_2), (s_2, t_3)\}$, $t_1 < t_2 < t_3$. The tightness of $\{W_n^{(2)}\}$ then follows from the results of Bickel and Wichura (1971). Theorem 3.1

is proved.

Corollary 3.1. *Let*

$$W_n^*(s, t) = [Q_{n,n(s)}(t) + \gamma t \sum_{i \le n(s)} c_{ni} d_{ni}]/B_n \sqrt{\gamma_{22}} \tag{3.24}$$

where, $\gamma = \gamma_{11} + \gamma_{21}$ *is given by* (2.11) *and* (2.14), *i.e.,*

$$\gamma = \int_{-\infty}^{\infty} \psi^{(1)}(x) dF(x) + \sum_{j=1}^{p} (\beta_j - \beta_{j-1}) f(a_j). \tag{3.25}$$

Then, under the assumptions of Theorem 3.1, the process $\{W_n^*(s, t):$ $(s, t) \in K^*\}$ *converges weakly to the process* $W = \{W(s,t): (s,t) \in K^*\}$.

Proof. (3.24) follows from (3.1) - (3.3) and from (3.5).

Corollary 3.2. *Let* $\hat{\Delta}_n$ *be the M-estimator defined by* (2.17) *with the sequence* $\{c_n\}_{n=1}^{\infty}$ *satisfying* (2.30). *Then, under* $\psi_2 \not\equiv 0$, *it holds*

$$C_n(\hat{\Delta}_n - \Delta) - \frac{1}{\gamma C_n} \sum_{j=1}^{n} c_j \psi(X_j - \Delta c_j) = O_p(n^{-1/4}). \tag{3.26}$$

Remark.1. *If* $\psi_2 \equiv 0$, *the order on the right-hand side of* (3.26) *is* $O_p(n^{-1/2})$, *as it follows from Jurečková and Sen* (1980).

Theorem 3.2. *Let* $X_1, X_2, \ldots$ *be i.i.d. random variables distributed according to d.f.* F. *Suppose that* F *and* ψ *satisfy the regularity conditions of Section 2. Then, under* (2.1) - (2.5) *and* (2.10) - (2.16),

$$\sup\{|M_{nk}(t) + tC_{k\gamma}|: \ k \le n \ \text{and} \ |t| \le k^*\} \xrightarrow{p} 0 \tag{3.27}$$

as $n \to \infty$, *where*

$$M_{nk}(t) = \{\sum_{i \le k} c_i [\psi(X_i - tC_n^{-1} c_i) - \psi(X_i)]\}/(\sigma_0 C_n) \tag{3.28}$$

and c_i, C_n, σ_0 *are all defined as in Section 2.*

Proof. Let $\mathcal{B}_k$ be the σ-field generated by $\{X_i, i \le k\}$, $k \ge 1$ and $\mathcal{B}_0$ be the trivial σ-field. Then, for every $n \ge 1$, $\{(M_{nk}(t) - EM_{nk}(t), -k^* \le t \le k^*), \mathcal{B}_k; \ k \le n\}$ is a martingale (process), so that

$$\left\{\sup_{-k^*\le t\le k^*} |M_{nk}(t) - EM_{nk}(t)| > \varepsilon,\ \mathcal{B}_k;\ k \le n\right\} \tag{3.29}$$

is a nonnegative submartingale. Therefore, by the Kolmogorov inequality (for submartingales), for every $\varepsilon > 0$,

$$P\{\sup_{k\le n}\ \sup_{-k^*\le t\le k^*} |M_{nk}(t) - EM_{nk}(t)| > \varepsilon\}$$

$$\le \frac{1}{\varepsilon} E\{\sup_{-k^*\le t\le k^*} |M_{nn}(t) - EM_{nn}(t)|\}. \tag{3.30}$$

Regarding the monotonicity of ψ, the supremum on the right-hand side of (3.30) could be approximated by a supremum over a finite set of points and it follows from the assumptions that $\lim_{n\to\infty} E|M_{nn}(t) - EM_{nn}(t)| = 0$ uniformly in $[-k^*, k^*]$. Finally, the expectation term $EM_{nn}(t)$ may be replaced by $-t\gamma C_k$ similarly as in Corollary 3.1. Q.E.D.

To prove the Theorem 2.2 of Section 2, we shall still need the following result concerning the error of the approximation of σ_0^2 by s_n^2:

Lemma 3.1. *Let* s_n^2 *be defined by* (2.19) *with the sequence* $\{c_n\}_{n=1}^{\infty}$ *satisfying* (2.30); *let for any* T, $0 < T < \infty$,

$$n_T = \max\{k:\ C_k^2 \le TC_n^2\}. \tag{3.31}$$

Then

$$\max_{n_\varepsilon \le k \le n_T} n^{\frac{1}{2}} |(s_k^2 - \sigma_0^2) - (s_k^{02} - \sigma_0^2)| \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty \tag{3.32}$$

holds for any $\varepsilon > 0$,

$$s_n^{02} = \frac{1}{n} \sum_{i=1}^{n} \psi^2(X_i - \Delta c_i). \tag{3.33}$$

Remark 2. *Note that by the Kintchine law of large numbers,*

$$s_n^{02} \to \theta_0^2 \quad \text{a.s.} \quad \textit{as} \quad n \to \infty. \tag{3.34}$$

Proof of Lemma 3.1. Theorem 3.2 implies

$$\max_{n_\varepsilon \le k \le n_T} \frac{1}{C_n} |C_k^2(\hat{\Delta}_k - \Delta) - (1/\gamma)\sum_{i=1}^n c_i \psi(X_i - \Delta c_i)| \xrightarrow{p} 0 \tag{3.35}$$

as $n \to \infty$. The sequence $\left\{\frac{1}{\gamma C_n} \sum_{i \le k} c_i \psi(X_i - \Delta c_i): \quad k \le n_T\right\}$ is tight by Donsker theorem and thus it follows from (3.35) that for every $\varepsilon > 0$ and $\eta > 0$ there exist an n_0 and $\delta(> 0)$ such that

$$P\{\max_{n_{1-\delta} \le \ \le n_{1+\delta}} |C_n(\hat{\Delta}_\varepsilon - \hat{\Delta}_n)| > \varepsilon\} < \eta \quad \text{for} \quad n \ge n_0. \tag{3.36}$$

Moreover, by the same technique as in the proof of (3.27), we obtain

that

$$\sup\{|M^*_{nk}(t)|: \ k \le n, \quad |t| \le k^*\} \xrightarrow{p} 0, \text{ as } n \to \infty, \tag{3.37}$$

where

$$M^*_{nk}(t) = n^{-\frac{1}{2}} \sum_{i \le k} [\psi^2(X_i - \Delta c_i) - \psi^2(X_i - (\Delta + C_n^{-1} t) c_i)]. \tag{3.38}$$

Lemma then follows from (3.36) and (3.37).

Remark 3. *It follows from Theorem 3.1 that, if* $s > 0$ *and* $t_2 - t_1 \ne 0$,

$$[W_n(s, t_2) - W_n(s, t_1)]/\sqrt{s|t_2 - t_1} \xrightarrow{\mathcal{D}} \xi \tag{3.39}$$

where ξ *has the standard normal distribution. Under some assumptions,* (3.39) *extends to the case when* t_1, t_2, s *are replaced by sequences of random variables. More precisely, if* s_n, t_{n1}, t_{n2} *are random variables such that*

$$s_n \xrightarrow{p} s(\varepsilon(0, 1)) \quad \textit{and} \quad |t_{n1} - t_{n2}| \xrightarrow{p} \tau(> 0) \tag{3.40}$$

for some constants s *and* τ, *then*

$$[W_n(s_n, t_{n2}) - W_n(s_n, t_{n1})]/\sqrt{s_n|t_{n2} - t_{n1}|} \xrightarrow{\mathcal{D}} \xi. \tag{3.41}$$

We shall recall to this result in the next section.

4. PROOFS OF THEOREMS 2.1 AND 2.2

Proof of Theorem 2.1

For any fixed $d > 0$ and any Δ, by (2.24),

$$P_\Delta(N_d > n) = P_\Delta\{L_m > 2d, \quad \forall\, m \in (n_0, n]\}$$

$$\leq P_\Delta(L_n > 2d) = P_\Delta(C_n L_n > 2dC_n) \ . \tag{4.1}$$

Now, it follows from (2.21) - (2.23), Theorem 3.2 and Lemma 3.1, that

$$L_n C_n \xrightarrow{p} 2\theta = 2\tau_{\alpha/2}\sigma_0/\gamma \quad \text{as} \quad n \to \infty, \tag{4.2}$$

while, by (2.16), $\lim_{n\to\infty} 2dC_n = \infty$ and thus $\lim_{n\to\infty} P\,(N_d > n) = 0$.

Moreover, it follows from the definition (2.24) of N_d that N_d is $\downarrow$ in $d(> 0)$; this together with the fact that $L_n > 0$ with probability 1 implies that $N_d \to \infty$ a.s. as $d \downarrow 0$.

To prove (2.26), define $n_{d\varepsilon}^{(i)}$, $i = 1, 2$, by

$$n_{d\varepsilon}^{(i)} = \max\{k: \ C_k^2 \leq (1 + (-1)^i\varepsilon)C_{n_d}\}, \quad i = 1, 2 \tag{4.3}$$

where n_d is defined by (2.25). Then

$$P(C_{n_d}^{-1} C_{N_d} > 1 + \varepsilon) = P(N_d \geq n_{d_\varepsilon}^{(2)})$$

$$\leq P(L_{n_{d\varepsilon}^{(2)}-1} \geq 2d) = P\left\{C_{n_{d\varepsilon}^{(2)}-1} L_{n_{d\varepsilon}^{(2)}-1} \geq 2dC_{n_{d\varepsilon}^{(2)}-1}\right\} \tag{4.4}$$

where by (4.2), $C_{n_{d\varepsilon}^{(2)}-1} L_{n_{d\varepsilon}^{(2)}-1} \xrightarrow{p} 2\theta$ as $d \downarrow 0$, while by (2.25) and (4.3), $2dC_{n_{d\varepsilon}^{(2)}-1} \longrightarrow 2\theta(1 + \varepsilon) > 2\theta$. Hence, (4.4) converges to 0 as $d \downarrow 0$. Similarly, $P(C_{n_d}^{-1} C_{N_d} \leq 1 - \varepsilon) \longrightarrow 0$ as $d \downarrow 0$. This proves (2.26).

By virtue of Theorem 3.2, Lemma 3.1, (2.21) and (2.26), it follows that

$$\lim_{d\downarrow 0} P_\Delta\{C_{N_d}(\hat{\Delta}_{N_d}^{-} - \Delta) + \theta \leq x\} = \Phi(\frac{x}{\sigma_0}), \ x \in E \tag{4.5}$$

and

$$C_{N_d} L_{N_d} \xrightarrow{p} 2\theta \quad \text{as} \quad d \downarrow 0. \tag{4.6}$$

Thus, (2.27) follows from (4.5) and (4.6).

To prove Corollary 2.1, we may note that (2.26) and (2.28) imply (2.29). This completes the proof of Theorem 2.1 and Corollary 2.1.

Proof of Theorem 2.2

It follows from the definition of N_d,

$$P_\Delta\{n_d^{\frac{1}{4}}(dC_{N_d} - \theta)/\theta \le y\} \le P_\Delta\{n_d^{\frac{1}{4}}(L_{N_d} C_{N_d} - 2\theta)/\theta \le 2y\} \tag{4.7}$$

and

$$P_\Delta\{n_d^{\frac{1}{4}}(dC_{N_d-1} - \theta)/\theta \ge y\} \le P_\Delta\{n_d^{\frac{1}{4}}(L_{N_d-1} C_{N_d-1} - 2\theta)/\theta \ge 2y\}. \tag{4.8}$$

Hence, to prove (2.33), it suffices to show that as $d \downarrow 0$,

$$P_\Delta\{n_d^{\frac{1}{4}}(L_{n_d} C_{n_d} - 2\theta)/(2\theta) \le y\} \longrightarrow \Phi((2\theta/\gamma_{22})^{\frac{1}{2}} y\gamma/\kappa) \tag{4.9}$$

and

$$\max\{|n^{\frac{1}{2}}(L_n C_n - L_m C_m)| : |C_m^2/C_n^2 - 1| < \delta\} = o_p(1). \tag{4.10}$$

Put $c_{ni} = c_i$, $d_{ni} = \dfrac{c_i}{C_n}$, $1 \le i \le n$. Then, regarding (3.20) and (3.21), we have

$$\sqrt{n}\, B_n^2/C_n^2 = \sqrt{n}\, C_n^*/C_n^3 = \kappa_n^2 \longrightarrow \kappa^2 \quad \text{as} \quad n \to \infty \tag{4.11}$$

where

$$B_n^2 = \sum_{i=1}^n |c_i|^3/C_n. \tag{4.12}$$

It follows from Corollary 3.1 and from Remark 3 that

$$L\{(B_{n_d}\sqrt{\gamma_{22}})^{-1}(\textstyle\sum_{j \le n_d} c_i[\psi(X_i - \hat{\Delta}_{n_d}^- c_i) - \psi(X_i - \hat{\Delta}_{n_d}^+ c_i)] - \gamma C_{n_d}^2 L_{n_d})\} \to N(0, 2\theta). \tag{4.13}$$

By virtue of Lemma 3.1, (2.21) - (2.22), (2.25), (4.11) and (4.12), we see that (4.13) implies (4.9). Moreover, (4.10) follows from Lemma 3.1 and from the tightness property of the process $W_n^*(s, t)$ in (3.24). Finally, (2.34) along with (2.33) ensures (2.35). Hence the proof of Theorem 2.2 and Corollary 2.2 is complete.

5. SEQUENTIAL TEST FOR Δ BASED ON M-ESTIMATORS

Consider again the single regression model (1.1) and the testing problem of (1.2). Without loss of generality, we may take

$$H_0:\ \Delta = 0 \quad \text{vs.} \quad H_1:\ \Delta = \delta > 0, \quad \delta \text{ specified.} \tag{5.1}$$

Along the same lines as in Jurečková and Sen (1980), we consider the following sequential test. Define s_n^2 as in (2.19), L_n as in (2.33) and put

$$\hat{\nu}_n = C_n L_n / 2\tau_{\alpha/2} \left(\xrightarrow{p} \nu = \sigma_0/\gamma \text{ by (4.2)}\right). \tag{5.2}$$

Let α_1 and α_2 be the prescribed probabilities of the errors of the type I and type II, respectively, $0 < \alpha_1 + \alpha_2 < 1$. Let A and B be two numbers such that $0 < \frac{\alpha_2}{1-\alpha_1} \le B < 1 < A \le \frac{1-\alpha_2}{\alpha_1} < \infty$ and put $a = \log A$, $b = \log B$ (so that $b < 0 < a$). We start with an initial sample of size $n_0 = n_0(\delta)$ and continue sampling as long as

$$b s_n \hat{\nu}_n < \delta S_n(\tfrac{1}{2}\delta) < a_n s_n \hat{\nu}_n. \tag{5.3}$$

Thus, the stopping number $N(=N(\delta))$ is defined by

$$N(\delta) = \min\{n \ge n_0(\delta):\ \delta S_n(\tfrac{1}{2}\delta)/(s_n \hat{\nu}_n) \notin (b, a)\} \tag{5.4}$$

and we stop sampling when $N(\delta)$ observations are made with the acceptance or rejection of H_0 according as

$$\delta S_{N(\delta)}(\tfrac{\delta}{2})/(s_{N_d} \hat{\nu}_{N_d}) \quad \text{is} \quad \le b \quad \text{or} \quad \ge a. \tag{5.5}$$

Jurečková and Sen (1980) studied the properties of the test in the case of an absolutely continuous function, i.e., when $\psi \equiv \psi_1$. They proved that the testing procedure terminates with probability 1 and derived the limiting OC function as $\delta \downarrow 0$.

Unlike in the case of sequential confidence interval, it could be shown that the presence of jump-discontinuities of the score function does not change the limiting OC function of the sequential test (excepting the contribution of γ_{21} in γ). The proofs of (i) the finiteness of the sequential procedure and (ii) Theorem 5.1 of Jurečková and Sen (1980) readily extend to the case of $\psi = \psi_1 + \psi_2$ when the theorems of Section 3 of the current paper are incorporated. Hence, we omit the details.

REFERENCES

[1] BICKEL, P.J. and WICHURA, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42**, 1656-1670.

[2] CAROLL, R.J. (1977). On the asymptotic normality of stopping times based on robust estimators. *Sankhyā, Ser A.* **39**, 355-377.

[3] CHOW, Y.S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36**, 457-462.

[4] GEERTSEMA, J.C. (1968). Sequential confidence intervals based on rank tests. *Ann. Math. Statist.* **39**, 1016-1026.

[5] GHOSH, M. and Sen, P.K. (1972). On bounded-length confidence interval for the regression coefficient based on a class of rank statistics. *Sankhyā, Ser. A.* **34**, 33-52.

[6] GHOSH, M. and SEN, P.K. (1976). Asymptotic theory of sequential tests based on linear functions of order statistics. In *Essays in Prob. & Statist.* (Ogawa Vol.), Ed. S. Ikeda et al. 480-498.

[7] GHOSH, M. and SEN, P.K. (1977). Sequential rank tests for regression. *Sankhyā, Ser A.* **39**, 45-62.

[8] GLESER, L.J. (1965). On the asymptotic theory of fixed size sequential confidence bounds for linear regression parameters. *Ann. Math. Statist.* **36**, 463-467.

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

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

[11] JUREČKOVÁ, J. (1973). Central limit theorem for Wilcoxon rank statistics process. *Ann. Statist.* **1**, 1046-1060.

[12] JUREČKOVÁ, J. (1977). Asymptotic relations of M-estimates and R-estimates in linear regression model. *Ann. Statist.* **5**, 464-472.

[13] JUREČKOVÁ, J. (1979). Bounded-length sequential confidence intervals for regression and location parameters. *Proc. 2nd Prague Conf. on Asymptotic Statistics.* 239-250.

[14] JUREČKOVÁ, J. (1980). Asymptotic representation of M-estimators of location. *Operationsforschung und Statistik, Ser Statist.*

[15] JUREČKOVÁ, J. and SEN, P.K. (1980). Invariance principles for some stochastic processes relating to M-estimators and their role in sequential statistical inference. *Sankhyā, Ser. A.*

[16] SEN, P.K. and GHOSH, M. (1971). On bounded length sequential confidence intervals based on one-sample rank order statistics. *Ann. Math. Statist.* **42**, 189-203.

[17] SEN, P.K. and GHOSH, M. (1974). Sequential rank tests for location. *Ann. Statist.* **2**, 540-552.