M-ESTIMATORS AND L-ESTIMATORS OF LOCATION: UNIFORM INTEGRABILITY AND ASYMPTOTIC RISK-EFFICIENT SEQUENTIAL VERSIONS

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. 1280

APRIL 1980

M-ESTIMATORS AND L-ESTIMATORS OF LOCATION: UNIFORM INTEGRABILITY AND ASYMPTOTIC RISK-EFFICIENT SEQUENTIAL VERSIONS

by

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

and

PRANAB KUMAR SEN[1]
Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514, U.S.A.

ABSTRACT

Sequential M- and L-estimators of location minimizing the risk asymptotically as the cost of one observation tends to 0 are constructed. Their asymptotic risk efficiencies are shown to coincide with the asymptotic efficiencies of the respective non-sequential estimators; this enables to construct the asymptotically minimax sequential M- and L-estimators in the model of contaminacy. The asymptotic distributions of the stopping times are derived for both types of estimators. The theorems on uniform integrability and moment convergence of (non-sequential) M- and L-estimators, developed as the main tools for the proofs, have an interest of their own.

AMS 1970 *Subject Classifications:* Primary 62L12; Secondary 60F25

Key Words & Phrases: M-estimator, L-estimator, asymptotic risk-efficiency, sequential point estimator, moment convergence.

---

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

1. INTRODUCTION

Nonparametric sequential point estimation of location has received considerable attention during the recent past. Ghosh and Mukhopadhyay (1979) and Chow and Yu (1980) have considered asymptotically risk-efficient sequential point estimation of the mean of a population based on the sequence of sample means and variances, while Sen and Ghosh (1980) have extended the theory to general U-statistics. Sen (1980) has considered the problem of estimating the location of symmetric (but unknown) distribution based on a general class of *rank-order* (or so called R-) *estimators* and established the asymptotic risk-efficiency of the proposed sequential procedure. In the classical non-sequential case, the R-estimators form one of the three main groups of robust competitors of classical estimation procedures; the other two major groups are formed by M-*estimators* and L-*estimators* [viz., Huber (1973, 1977)]. The theory of asymptotically risk-efficient (sequential) point estimation based on a broad class of M- and L-estimators is developed in the current paper. *Uniform integrability* and *moment-convergence* properties of these M- and L-estimators play a fundamental role in this context.

Along with the preliminary notions, the proposed sequential point estimation procedures are outlined in Section 2. Section 3 is devoted to the study of uniform integrability and moment convergence of the M-estimators. Parallel results for the L-estimators are considered in Section 4. These results are then applied in the proofs of main theorems of Section 5 concerning the properties of the proposed sequential procedures. In particular, the Section 5 deals with the asymptotic risk-efficiency and with the asymptotic normality of the allied *stopping times*. Similarly as in the case of R-estimators [Sen (1980)], it is shown that the asymptotic risk efficiencies of sequential estimators coincide with the asymptotic efficiencies of their non-sequential versions. This among others enables to extend the *asymptotic minimax properties* of M- and L-estimators in the model of contaminacy to the sequential case.

2. THE PROPOSED SEQUENTIAL PROCEDURES

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed random variables (i.i.d.r.v.) with distribution function (d.f.) $F_\theta(x) = F(x-\theta)$, $x \in R = (-\infty,\infty)$, where F (unknown) is symmetric about 0 and θ is the unknown location parameter to be estimated. Let T_n be a suitable estimator of θ based on $X_1,\ldots,X_n$ and assume that

$$\nu_n^2 = nE(T_n - \theta)^2 \text{ exists for all } n \geq n_0, \tag{2.1}$$

for some $n_0(\geq 1)$ and

$$\nu_n^2 \longrightarrow \nu^2 \quad \text{as} \quad n \to \infty, \quad 0 < \nu < \infty. \tag{2.2}$$

We conceive the loss (in estimating θ by T_n)

$$Q_n(a, c) = a(T_n - \theta)^2 + cn, \tag{2.3}$$

where a and c are positive constants. Then the *risk* is

$$\lambda_n(a, c) = EQ_n(a, c) = n^{-1}a\nu_n^2 + cn. \tag{2.4}$$

We like to minimize (2.4) by a proper choice of n. The optimal choice of n generally depends on the unknown F, for any fixed c as well as asymptotically as $c \downarrow 0$. In this asymptotic case, the optimal choice of n is $n_0(c)$, where

$$n_0(c) \sim b\nu, \quad b = (a/c)^{\frac{1}{2}} \quad \text{and} \quad \lambda_{n_0(c)}(a, c) \sim 2\nu\sqrt{ac} \tag{2.5}$$

where $q(c) \sim r(c)$ denotes that $\lim_{c\downarrow 0} q(c)/r(c) = 1$. This suggests the following procedure: Let $\{\hat{\nu}_n\}$ be a sequence of estimates of ν and let n' be an initial sample size (≥ 2) and $h(>0)$ be an arbitrary constant. Define a *stopping number* $N(=N_c)$ by

$$N_c = \min\{n \geq n': n \geq b(\hat{\nu}_n + n^{-h})\},\ c > 0 \tag{2.6}$$

and consider the sequential point estimator T_{N_c} based on $X_1, \ldots, X_{N_c}$. The risk of estimating θ by T_{N_c} is then

$$\lambda^*(a, c) = aE(T_{N_c} - \theta)^2 + cEN_c. \tag{2.7}$$

We are primarily interested in showing that

$$\lambda^*(a, c)/\lambda_{n_0(c)}(a, c) \longrightarrow 1 \quad \text{as} \quad c \downarrow 0, \tag{2.8}$$

which means that the sequential procedure is asymptotically (as $c \downarrow 0$) equally risk-efficient as the optimal non-sequential one, if ν were known.

The convergence (2.8) has been studied by more authors [referred to in Section 1] in the case that $\{T_n\}$ is either the sample mean, U-statistic or some case of R-estimator. In the current paper, we shall show that (2.8) holds for a broad classes of M-estimators and L-estimators (i.e., the estimators of maximum-likelihood type and of linear combination of order statistics type, respectively).

An M-*estimator* M_n of θ is a solution of the equation

$$S_n(t) = \sum_{i=1}^{n} \psi(X_i - t) = 0 \tag{2.9}$$

with respect to t, where ψ is some nondecreasing score function (so that $S_n(t)$ is $\searrow$ in t). More precisely, M_n is defined by

$$M_n = (M_n^* + M_n^{**})/2, \tag{2.10}$$

where

$$M_n^* = \sup\{t: S_n(t) > 0\} \quad \text{and} \quad M_n^{**} = \inf\{t: S_n(t) < 0\}. \tag{2.11}$$

Under suitable regularity conditions on ψ and on F [viz., Huber (1964)], to be specified later on,

$$\mathcal{L}\{n^{\frac{1}{2}}(M_n - \theta)\} \longrightarrow \mathcal{N}(0, \nu_{(M)}^2) \text{ as } n \to \infty \tag{2.12}$$

where

$$\nu^2_{(M)} = \sigma^2_{(M)}/\gamma_{(M)}, \quad \sigma^2_{(M)} = \int_{-\infty}^{\infty} \psi^2(x)dF(x), \tag{2.13}$$

$$\gamma_{(M)} = \gamma(\psi, F) = \int_{-\infty}^{\infty} \{-f'(x)/f(x)\}\psi(x)dF(x) \ (>0) \tag{2.14}$$

and $f'(x) = \frac{d}{dx}f(x) = \frac{d^2}{dx^2}F(x)$ is assumed to exist almost everywhere. In Section 3, we shall show that (2.1) and (2.2) hold for M-estimators generated by a class of bounded ψ-functions. In this case, we shall estimate $\nu^2_{(M)}$ as follows. Let

$$s^2_{n(M)} = n^{-1}\sum_{i=1}^{n}\psi^2(X_i - M_n), \ n \geq 1, \tag{2.15}$$

let Φ be the standard normal d.f. and let $\Phi(-\tau_\varepsilon) = \varepsilon$, $0 < \varepsilon < 1$. Put

$$\begin{aligned} M_n^- &= \sup\{t: n^{-\frac{1}{2}}S_n(t) > \tau_{\alpha/2}s_{n(M)}\} \\ M_n^+ &= \inf\{t: n^{-\frac{1}{2}}S_n(t) < -\tau_{\alpha/2}s_{n(M)}\} \end{aligned} \tag{2.16}$$

$$d_{n(M)} = M_n^+ - M_n^-(\geq 0), \tag{2.17}$$

where $\alpha (0 < \alpha < 1)$ is some preassigned number. Then, it follows from Jurečková (1977) that as n increases,

$$\hat{\nu}_{n(M)} = \sqrt{n}\, d_{n(M)}/2\tau_{\alpha/2} \xrightarrow{p} \nu_{(M)} = \sigma_{(M)}/\gamma_{(M)}; \tag{2.18}$$

in fact, stronger convergence properties of $\nu_{n(M)}$ have been studied by Jurečková and Sen (1980 a, b). The stopping number defined by (2.6), corresponding to $\{\hat{\nu}_n\} = \{\hat{\nu}_{n(M)}\}$ is denoted by $N_c^{(M)}$, so that

$$N_c^{(M)} = \min\{n \geq n': n \geq b(\hat{\nu}_{n(M)} + n^{-h})\} \tag{2.19}$$

and we shall show in Section 5 that (2.8) holds for $\{M_{N_c^{(M)}}\}$.

The L-*estimator* L_n of location θ is typically of the form

$$L_n = \sum_{i=1}^{n} c_{ni}X_{n,i} \tag{2.20}$$

where $X_{n,1} \leq \cdots \leq X_{n,n}$ are the order statistics corresponding to $X_1, \ldots, X_n$ and

$$c_{ni} = c_{n,n-i+1} \geq 0, \ \forall\ 1 \leq i \leq n, \quad \text{and} \quad \textstyle\sum_{i=1}^{n} c_{ni} = 1. \tag{2.21}$$

Denote

$$J_n(t) = n \cdot c_{ni} \ \text{for} \ \frac{i-1}{n} < t \leq \frac{i}{n}\ ,\ i = 1, \ldots, n \tag{2.22}$$

and suppose that

$$J_n(t) \to J(t) \ \text{a.s.},\ t \in (0, 1),\ J(1-t) = J(t) \geq 0,\ \int_0^1 J(t)dt = 1. \tag{2.23}$$

Then, under some regularity conditions [viz., Huber (1969)],

$$L(n^{\frac{1}{2}}(L_n - \theta)) \to N(0, \nu^2_{(L)}) \ \text{as} \ n \to \infty, \tag{2.24}$$

where

$$\nu^2_{(L)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F(x \wedge y) - F(x)F(y)]J(F(x))J(F(y))dxdy\ . \tag{2.25}$$

We proceed to estimate $\nu^2_{(L)}$ by

$$\hat{\nu}^2_{n(L)} = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} c_{ni}c_{nj}\{n(i \wedge j) - ij\}(X_{n,i+1} - X_{n,i})(X_{n,j+1} - X_{n,j}), \tag{2.26}$$

where under suitable regularity conditions [viz., Sen (1978)]

$$\hat{\nu}_{n(L)} \to \nu_L \ \text{a.s.} \ \text{as} \ n \to \infty\ ; \tag{2.27}$$

asymptotic normality results pertaining to the $\nu_{n(L)}$ are due to Gardiner and Sen (1979). The stopping number, defined by (2.6), for $\{\hat{\nu}_n\} = \{\hat{\nu}_{n(L)}\}$, is denoted by $N_c^{(L)}$, so that

$$N_c^{(L)} = \min\{n \geq n' :\ n \geq b(\hat{\nu}_{n(L)} + n^{-h})\}. \tag{2.28}$$

We shall show in Section 5 that (2.8) holds for $\{L_{N_c^{(L)}}\}$ corresponding to a class of J-functions which vanish outside of a compact subinterval of (0, 1).

In the remaining of this section, we state the basic regularity conditions on F, ψ and J, pertaining to our study. Sections 3 and 4 are devoted to the study of the uniform integrability and the moment convergence of $\{M_n\}$ and $\{L_n\}$, respectively; these results are used in Section 5 in the proofs of (2.8) for the corresponding sequential procedures.

Assumptions on F: We assume that F has an absolutely continuous density f such that $f(x) = f(-x)$, $\forall\, x \in R$ and

$$f(x) \text{ is } \searrow \text{ in } x \text{ for } x \geq 0. \tag{2.29}$$

Moreover, F is supposed to have finite Fisher information, i.e.,

$$I(F) = \int_{-\infty}^{\infty} \{f'(x)/f(x)\}^2 dF(x) < \infty \tag{2.30}$$

and we assume that there exists a positive number ℓ (not necessarily an integer or ≥ 1), such that

$$E|X_1|^{\ell} = \int_{-\infty}^{\infty} |x|^{\ell} dF(x) < \infty . \tag{2.31}$$

Assumptions on ψ: We assume that ψ is nondecreasing and skew-symmetric, i.e.,

$$\psi(x) = -\psi(-x) \text{ is } \nearrow \text{ in } x \in R^+ = [0, \infty), \tag{2.32}$$

and that there exists a positive number k such that

$$\psi(x) = \psi(k) \cdot \text{sign}\, x \quad \text{for} \quad |x| \geq k. \tag{2.33}$$

Moreover, suppose that ψ could be decomposed in the absolutely continuous and step components, i.e.,

$$\psi(x) = \psi_1(x) + \psi_2(x) \;\; \forall\, x \in R \tag{2.34}$$

where $\psi_1(x)$ is absolutely continuous [inside (-k, k)] and ψ_2 is pure step-function having a finite number of jumps inside (-k, k); we denote the jump-points by a_j, $1 \leq j \leq m$, and let $\psi_2(x) = \beta_j$ for $a_{j-1} < x < a_j$, $1 \leq j \leq m+1$, where $a_0 = -k$ and $a_{m+1} = k$. Then the constant $\gamma_{(M)}$ defined in (2.14) is equal to

$$\gamma_{(M)} = \int_{-\infty}^{\infty} \psi_1'(x) dF(x) + \textstyle\sum_{j=1}^{m} (\beta_j - \beta_{j-1}) f(a_j) > 0. \tag{2.35}$$

Put

$$c_\ell(x) = |x|^\ell F(x)[1 - F(x)],\ x \in R;\ c_\ell^* = \sup_{x\in R} c_\ell(x) \tag{2.36}$$

where ℓ is given by (2.31). Then

$$c_\ell^* < \infty,\quad \lim_{x\to\pm\infty} c_\ell(x) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \{F(x)(1 - F(x))\}^b dx < \infty \quad \forall\ b > \frac{1}{\ell} > 0. \tag{2.37}$$

<u>Assumptions on</u> $\{c_{ni}: 1 \le i \le n\}$ <u>and on</u> J: We assume that $c_{ni} = c_{n,n-i+1} \ge 0$, $1 \le i \le n$; $\sum_{i=1}^n c_{ni} = 1$ and that there exist an α_0 $(0 < \alpha_0 < \frac{1}{2})$ and a sequence $\{k_n\}$, $k_n > 0$ such that

$$c_{ni} = c_{n,n-i+1} = 0 \quad \text{for} \quad i \le k_n \quad \text{where} \quad \frac{k_n}{n} \longrightarrow \alpha_0 \quad \text{as} \quad n \to \infty. \tag{2.38}$$

Moreover, denoting

$$J_n(t) = n \cdot c_{ni} \text{ for } \frac{i-1}{n} < t \le \frac{i}{n},\quad i = 1, \dots, n \tag{2.39}$$

we assume that

$$\lim_{n\to\infty} J_n(t) = J(t) \quad \text{a.s.},\ t \in [0, 1] \tag{2.40}$$

where the function $J(t)$ has bounded variation on $[0, 1]$ and

$$J(t) = J(1 - t) \ge 0,\quad t \in [0, 1],\ \int_0^1 J(t)dt = 1. \tag{2.41}$$

In the context of L-estimators, the assumption of finite Fisher information may be replaced by a weaker assumption

$$\sup_{F^{-1}(\alpha_0) \le x \le F^{-1}(1-\alpha_0)} |f'(x)| < \infty\ . \tag{2.42}$$

3. MOMENT CONVERGENCE OF M-ESTIMATORS

Uniform integrability and some moment-convergence properties of $\{n^{\frac{1}{2}}(M_n - \theta)\}$ are studied here. The following lemmas are needed in the sequel but they have an interest of their own.

LEMMA 3.1. *Under the regularity conditions on* ψ *and* F *of Section 2, for every* $c_1 > 0$ *and* $0 < t < \sqrt{n}\, c_1$,

$$P_\theta\{\sqrt{n}|M_n - \theta| > t\} = P_0\{\sqrt{n}|M_n| > t\} \le 2e^{-c_2 t^2} , \tag{3.1}$$

where

$$c_2 \ge 2[f(k + c_1)]^2 > 0 . \tag{3.2}$$

PROOF. Note that for every $t > 0$,

$$P_0\{\sqrt{n}|M_n| > t\} = P_0\{\sqrt{n}\, M_n > t\} + P_0\{\sqrt{n}\, M_n < -t\} = 2P_0\{\sqrt{n}\, M_n > t\}, \tag{3.3}$$

where, by (2.9) - (2.11),

$$\begin{aligned} P_0\{\sqrt{n}\, M_n > t\} &\le P_0\{n^{-1} S_n(t/\sqrt{n}) \ge 0\} \\ &= P_0\{n^{-1} S_n(t/\sqrt{n}) - \mu_n(t) \ge -\mu_n(t)\}, \end{aligned} \tag{3.4}$$

and

$$\begin{aligned}
&- \mu_n(t) = -E_0 n^{-1} S_n(t/\sqrt{n}) = -E_0\psi(X_1 - (t/\sqrt{n})) \\
&= E_0[\psi(X_1) - \psi(X_1 - (t/\sqrt{n}))] = \int_{-\infty}^{\infty} \psi(x)\, d[F(x) - F(x + (t/\sqrt{n}))] \\
&= \int_{-\infty}^{\infty} [F(x + (t/\sqrt{n})) - F(x)]\, d\psi(x) \\
&= \int_{-k}^{k} [F(x + (t/\sqrt{n})) - F(x)]\, d\psi(x) \\
&= \int_{0}^{k} [F(x + (t/\sqrt{n})) - F(x) + F(-x + (t/\sqrt{n})) - F(-x)]\, d\psi(x) \\
&\qquad\qquad [\text{as } \psi(x) + \psi(-x) = 0,\ \forall\ x] \\
&= \int_{0}^{k} [F(x + (t/\sqrt{n})) - F(x - (t/\sqrt{n}))]\, d\psi(x) \qquad [\text{as } F(x) + F(-x) = 1,\ \forall\ x] \\
&= (2t/\sqrt{n}) \int_{0}^{k} f(x + (\theta t/\sqrt{n}))\, d\psi(x) \qquad [\text{where } |\theta| < 1] \\
&\ge (2t/\sqrt{n}) f(k + (t/\sqrt{n}))[\psi(k) - \psi(0)] \qquad [\text{as } f(x) \text{ is } \downarrow \text{ in } x,\ x \ge 0] \\
&\ge (2t/\sqrt{n}) f(k + c_1)\psi(k),\ \forall\ t \in (0, c_1\sqrt{n}] \qquad [\text{as } \psi(0) = 0] \\
&= (2c_2)^{\frac{1}{2}}\psi(k)(t/\sqrt{n}) .
\end{aligned} \tag{3.5}$$

Therefore, by (3.3) - (3.5), for every $0 < t < c_1\sqrt{n}$,

$$P_0\{\sqrt{n}|M_n| > t\} \le 2P_0\{\tfrac{1}{n}\textstyle\sum_{i=1}^n Z_{ni} \ge (2c_2)^{\frac{1}{2}}\psi(k)(t/\sqrt{n})\} \tag{3.6}$$

where

$$Z_{ni} = \psi(X_i - \frac{t}{\sqrt{n}}) - E_0\psi(X_i - \frac{t}{\sqrt{n}}), \quad i = 1,\ldots,n \tag{3.7}$$

are independent r.v. with mean 0, bounded by $2\psi(k)$ with probability 1. Hence, using Theorem 2 of Hoeffding (1963) on r.v. (3.7), the desired result follows from (3.6). Q.E.D.

LEMMA 3.2. *Under the regularity conditions on* ψ *and* F *of Section* 2, *for every* $t > 0$,

$$P_0\{\sqrt{n}|M_n| > t\} \le 2n\binom{n-1}{m-\delta_n}\int_{F(-k+\frac{t}{\sqrt{n}})}^1 u^m(1-u)^{n-m-\delta_n}du, \tag{3.8}$$

where

$$m = \left[\frac{n+1}{2}\right] \text{ and } \delta_n = 1 \text{ for } n = 2m,\ \delta_n = 0 \text{ for } n = 2m+1,\ m \ge 1. \tag{3.9}$$

PROOF. We consider only the case $n = 2m$; the proof for $n = 2m+1$ is analogous. Note that for every $t > 0$,

$$P_0\{\sqrt{n}|M_n| > t\} = 2P_0\{\sqrt{n}\,M_n > t\} \le 2P_0\{X_{n,m+1} \ge -k + (t/\sqrt{n})\} \tag{3.10}$$

where $X_{n,1} \le \cdots \le X_{n,n}$ are the order statistics corresponding to $X_1,\ldots,X_n$. Since the right hand side of (3.10) equals to that of (3.8) the proof of (3.8) is complete. Q.E.D.

For any $a \in [0, 1]$, put

$$\rho(a) = 4a(1-a), \quad \text{so that} \quad 0 \le \rho(a) \le 1. \tag{3.11}$$

LEMMA 3.3. *For every* $n(\ge 1)$ *and* $a > \frac{1}{2}$,

$$2n\binom{n-1}{m-\delta_n}\int_a^1 u^m(1-u)^{n-m-\delta_n}du \le 2(\rho(a))^n \tag{3.12}$$

where m, δ_n *and* $\rho(a)$ *are defined by* (3.9) *and* (3.11).

PROOF. We shall again prove (3.12) for $n = 2m$ only; the proof for

$n = 2m+1$ is analogous. Note that by repeated partial integration, the left-hand side of (3.12) reduces to

$$2\sum_{i=0}^{m-1}\binom{n}{i}a^{i}(1-a)^{n-i} \le 2P\left\{B(n,a) \le \frac{n}{2}\right\}$$
$$\le 2P\left\{\frac{1}{n}B(n,a) - a \le \frac{1}{2} - a\right\}, \qquad (3.13)$$

where $B(n, a)$ is a binomial r.v. with parameters (n, a). Since $a > \frac{1}{2}$, by using Theorem 1 of Hoeffding (1963), we may bound the right hand side of (3.13) by $2[\rho(a)]^{n}$. Q.E.D.

We are now in a position to prove the main theorems of this Section.

THEOREM 3.1. *For every* $r > 0$, *there exists an* $n_r(<\infty)$ *such that, under the regularity conditions* (2.29) - (2.34),

$$E_0\{n^{r/2}|M_n|^r\} < \infty \text{, uniformly in } n \ge n_r. \qquad (3.14)$$

PROOF. Let $c_1 > k > 0$, where k is defined by (2.33). Note that

$$E_0\{n^{r/2}|M_n|^r\} = \int_0^\infty rt^{r-1}P_0\{\sqrt{n}|M_n| > t\}dt$$
$$= \left(\int_0^{c_1\sqrt{n}} + \int_{c_1\sqrt{n}}^\infty\right)rt^{r-1}P_0\{\sqrt{n}|M_n| > t\}dt = I_{n1} + I_{n2}, \text{ say.} \qquad (3.15)$$

Then, by Lemma 3.1,

$$I_{n1} \le 2r\int_0^{c_1\sqrt{n}} e^{-c_2t^2}t^{r-1}dt \le 2r\int_0^\infty e^{-c_2t^2}t^{r-1}dt < \infty \qquad (3.16)$$

uniformly in $n = 1,2,\dots$. On the other hand, if we let

$$n_r = \left[\frac{r}{\ell}\right] + 1, \text{ where } \ell \text{ is defined by (2.31)} \qquad (3.17)$$

use (2.37) and Lemmas 3.2 and 3.3, we get

$$
\begin{aligned}
I_{n2} &\le 2r\int_{c_1\sqrt{n}}^{\infty} t^{r-1}[\rho(F(-k+(t\sqrt{n})))]^n dt \\
&\le 2rn^{r/2}\int_{c_1}^{\infty} u^{r-1}[\rho(F(-k+u))]^n du \\
&\le 2r(c_{\ell}^{*})^{(r-1)/\ell}\left\{n^{r/2}[\rho(F(-k+c_1))]^{n-n_r-b}\right\}\int_{c_1-k}^{\infty}[4F(y)(1-F(y))]^b dy
\end{aligned}
\tag{3.18}
$$

where $c_{\ell}^{*} < \infty$ is given by (2.37) and b is any number satisfying $b > 1/\ell$. Since $c_1 > k$, it holds $F(c_1 - k) > F(0) = \frac{1}{2}$, so that $\rho(F(c_1 - k)) < 1$ and hence, $n^{r/2}\,(F(c_2 - k))^{n-n_r-b}$ is uniformly bounded for $n \ge n_r$ and converges to 0 as $n \to \infty$. Finally, $\int_{c_1-k}^{\infty} (4F(y)(1-F(y)))^b dy < \infty$ by (2.36) and (2.37). Hence, $I_{n2} < \infty$ uniformly in $n \ge n_r$ and it converges to 0 as $n \to \infty$. This completes the proof of the theorem.

LEMMA 3.4. *Under the regularity conditions on* ψ *and* F,

$$
\sqrt{n}(M_n - \theta) - \frac{1}{\gamma_{(M)}\sqrt{n}}\sum_{i=1}^{n}\psi(X_i - \theta) = 0_p(n^{-\frac{1}{4}}) \tag{3.19}
$$

where $\gamma_{(M)}$ *is defined by* (2.14).

PROOF. The Lemma was proved in Jurečková (1980) [Theorem 3.3].

LEMMA 3.5. *Under the regularity conditions on* ψ *and* F *of Section* 2, *the sequence* $|n^{-\frac{1}{2}}\sum_{i=1}^{n}\psi(X_i - \theta|^{2r}$ *is uniformly integrable for* $n = 1,2,\ldots$ *and*

$$
E_{\theta}|n^{-\frac{1}{2}}\sum_{i=1}^{n}\psi(X_i - \theta)|^{2r} \longrightarrow \sigma_{(M)}^{2r}\frac{(2r)!}{r!\,2^r}, \quad r = 1,\ 2,\ldots \tag{3.20}
$$

as $n \to \infty$, *where* $\sigma_{(M)}$ *is defined by* (2.13).

PROOF. Since $E_{\theta}\psi(X_1 - \theta) = 0$ and $|\psi(y)| \le \psi(k) < \infty$, $\forall\ y \in R$, moments of all orders of $\psi(X_1 - \theta)$ exist. Hence, the result follows directly from the moment convergence result of von Bahr (1965).

THEOREM 3.2. *Under the regularity conditions* (2.29) - (2.34),

$$\lim_{n\to\infty} E_\theta(\sqrt{n}|M_n - \theta|)^{2r} = (\sigma_{(M)}/\gamma_{(M)})^{2r}(2r)!/(2^r r!) \tag{3.21}$$

holds for $r = 1,2,\ldots$.

PROOF. It follows directly from the uniform integrability in Theorem 3.1, from (3.19) and from Lemma 3.5. In fact, we need not confine ourselves to even integer $2r$. Since, by the Jensen inequality, $|n^{-\frac{1}{2}}\sum_{i=1}^{n}\psi(X_i - \theta)|^r$ is uniformly integrable for any $r' \varepsilon [2r-2, 2r]$, we may prove on parallel lines that as $n\to\infty$,

$$E_\theta(\sqrt{n}|M_n - \theta|)^r \longrightarrow (\sigma_{(M)}/\gamma_{(M)})^r \int_{-\infty}^{\infty} |z|^r d\Phi(z) \tag{3.22}$$

for any fixed real $r > 0$.

LEMMA 3.6. *For any* $\varepsilon > 0$ *and* $\delta > 0$, *there exist positive constants* c and n_0 *such that*

$$P\{|s_n^2 - \sigma_0^2| > \varepsilon\} \le cn^{-1-\delta}, \ \forall\ n \ge n_0. \tag{3.23}$$

PROOF. Let us define

$$s_n^{02} = n^{-1}\sum_{i=1}^{n}\psi^2(X_i - \theta),\ n \ge 1. \tag{3.24}$$

Since $\psi^2(X_i - \theta)$, $i = 1,\ldots,n$, are i.i.d. bounded valued r.v., by Theorem 1 of Hoeffding (1963), for every $\varepsilon > 0$ there exists an $\eta > 0$ such that

$$P\{|s_n^{02} - \sigma_0^2| > \varepsilon/2\} \le 2e^{-2n\eta},\ \forall\ n \ge 1. \tag{3.25}$$

Again, by virtue of Theorem 3.1,

$$P\{|M_n| > \tfrac{1}{2}\varepsilon\} = P\{\sqrt{n}|M_n| > \tfrac{1}{2}\sqrt{n}\,\varepsilon\} \le cn^{-1-\delta},\quad \forall\ n \ge n_0\ . \tag{3.26}$$

By (2.32) - (2.34), ψ^2 is of bounded variation on R, so that

$$\psi^2(y) = \psi_1^*(y) + \psi_2^*(y)\ , \tag{3.27}$$

where ψ_1^*, ψ_2^* are nonnegative and ψ_1^* is $\downarrow$ while ψ_2^* is $\nearrow$ in $y \varepsilon R$. Hence,

$$\sup_{|t|\le\frac{1}{2}\varepsilon} |\psi^2(y+t) - \psi^2(y)| \le |\psi_1^*(y - \tfrac{1}{2}\varepsilon) - \psi_1^*(y + \tfrac{1}{2}\varepsilon)| + |\psi_2^*(y + \tfrac{1}{2}\varepsilon) - \psi_2^*(y - \tfrac{1}{2}\varepsilon)|. \tag{3.28}$$

Since, by (2.33), ψ_1^* and ψ_2^* are bounded, by using (3.28) and the Markov inequality, we obtain that for every $\varepsilon' > 0$,

$$P\left\{ \sup_{|t| \le \frac{1}{2}\varepsilon} \left| \frac{1}{n}\sum_{i=1}^{n} \psi^2(X_i + t) - \frac{1}{n}\sum \psi^2(X_i) \right| > \frac{1}{2}\varepsilon' \right\} \le c' n^{-1-\delta} \quad \forall\ n \ge n_0 . \tag{3.29}$$

(3.23) then follows from (2.13), (2.15), (3.24), (3.25), (3.26) and (3.29). Q.E.D.

THEOREM 3.3. *Under the regularity conditions of* (2.29) - (2.34), *for every* $\varepsilon > 0$ *and* $\delta > 0$, *there exist a* C, $0 < C < \infty$ *and an* n_0 $(< \infty)$ *such that*

$$P\{ |\hat{\nu}_{n(M)} - \nu_{(M)}| > \varepsilon \} \le C n^{-1-\delta} \quad \forall\ n \ge n_0 . \tag{3.30}$$

PROOF. Since ψ is nondecreasing, exploiting technique of Jurečková (1969), we obtain that for every fixed $K(< \infty)$ and $\varepsilon > 0$, there exist an $m(= m_{K\varepsilon})$ and a set of points $(-K \le)\ t_1 < \cdots < t_m (\le K)$ such that

$$\begin{aligned} &\sup_{|t| \le K} \left| n^{-\frac{1}{2}}\{ S_n(n^{-\frac{1}{2}}t) - S_n(0)\} + t\gamma_{(M)} \right| \\ &\le \max_{0 \le j \le m} \left| n^{-\frac{1}{2}}\{ S_n(n^{-\frac{1}{2}}t_j) - S_n(0)\} + t_j\gamma_{(M)} \right| + \varepsilon/2 \\ &\le \max_{1 \le j \le m} \left| n^{-\frac{1}{2}}\{ S_n(n^{-\frac{1}{2}}t_j) - S_n(0)\} - \sqrt{n}\,\mu_n(t_j) \right| \\ &+ \max_{1 \le j \le m} \left| \sqrt{n}\,\mu_n(t_j) + t_j\gamma_{(M)} \right| + \varepsilon/2, \end{aligned} \tag{3.31}$$

where $\mu_j(t)$ is defined by (3.5) and

$$\sup_{|t| \le k} \left| \sqrt{n}\,\mu_n(t) + t\gamma_{(M)} \right| \longrightarrow 0 \quad \text{as} \quad n \to \infty . \tag{3.32}$$

The random variables $Z_{ni}^{(j)} = \psi(X_i - n^{-\frac{1}{2}}t_j) - \psi(X_i) - E\psi(X_i - n^{-\frac{1}{2}}t_j)$, $i = 1, \ldots, n$ are independent for any fixed $j (= 1, 2, \ldots, m)$ and $EZ_{ni}^{(j)} = 0$, $E(Z_{ni}^{(j)})^2 = O(n^{-\frac{1}{2}})$.

Proceeding as in Theorem 3.1 of Jurečková and Sen (1980b), we claim that under (2.32) - (2.34),

$$E[Z_{ni}^{(j)}]^{2q} = O(n^{-q/2}) \text{ for } q=1,2,\ldots . \tag{3.33}$$

Thus, by Markov inequality,

$$\begin{aligned} &P\{\max_{1\le j\le m} |n^{-\frac{1}{2}}S_n(n^{-\frac{1}{2}}t_j) - S_n(0) - \sqrt{n}\,\mu_n(t_j)| > \varepsilon'\} \\ &\le \textstyle\sum_{j=1}^{m} P\{|n^{-\frac{1}{2}}S_n(n^{-\frac{1}{2}}t_j) - S_n(0) - \sqrt{n}\,\mu_n(t_j)| > \varepsilon'\} = O(n^{-q/2}) \end{aligned} \tag{3.34}$$

where, given δ, we may choose q so large that $q/2 \ge 1+\delta$. From (3.31), (3.32) and (3.34), we obtain that for every $\varepsilon > 0$, $\delta > 0$,

$$P\{\sup_{|t|\le K} |n^{-\frac{1}{2}}[S_n(n^{-\frac{1}{2}}t) - S_n(0)] + t_{\gamma(M)}| > \varepsilon\} \le cn^{-1-\delta}, \quad \forall\ n \ge n_0 \tag{3.35}$$

and (3.31) follows readily from (3.23), (2.15) - (2.18) and (3.35).

Q.E.D.

4. MOMENT CONVERGENCE OF L-ESTIMATORS

First, we consider the following theorem on a.s. representation of L_n.

THEOREM 4.1. *Let* L_n *be an* L-*estimator of the form* (2.20) *with* c_{ni}, $1 \le i \le n$, *satisfying* (2.21) *and* (2.38). *Let* d.f. F *have the absolutely continuous symmetric density satisfying* (2.29) *and* (2.41). *Then, for any* $n \ge n_0$, *there exists a sequence* $\{Y_{ni}\}_{i=1}^{n+1}$ *of i.i.d. random variables with standard normal distribution such that*

$$|n^{\frac{1}{2}}(L_n - \theta) - (n+1)^{-\frac{1}{2}} \sum_{j=1}^{n+1} a_{nj}Y_{nj}|_{a.s.} = O(n^{-\frac{1}{2}}\log n) \quad \textit{as} \quad n \to \infty , \tag{4.1}$$

where

$$a_{nj} = \sum_{i=j}^{n} b_{ni} - \textstyle\sum_{i=1}^{n+1} \frac{1}{n+1} b_{ni}, \quad b_{ni} = c_{ni}/f(F^{-1}(\frac{i}{n+1})), \ 1 \le i \le n. \tag{4.2}$$

PROOF. We may put $\theta = 0$ without loss of generality. Note that by (2.29) and (2.41), for every $\beta \varepsilon (0, \frac{1}{2})$,

$$\sup_{F^{-1}(\beta)<x<F^{-1}(1-\beta)} \{F(x)[1-F(x)]\cdot|f'(x)/f^2(x)|\} \le [4/f^2(F^{-1}(\beta))] \sup_{F^{-1}(\beta)\le x\le F^{-1}(1-\beta)} |f'(x)| \le \gamma_\beta < \infty \tag{4.3}$$

where $\gamma_\beta (>0)$ may depend on β. As such, the condition (3.2) in Theorem 6 of Cxörgő and Révész (1978) holds for the case of $F^{-1}(\beta) \le x \le F^{-1}(1-\beta)$, and hence, we may virtually repeat the proof of their Theorem 3 [using our (4.3) instead of their more stringent (3.2)] and claim that for every $n (\ge n_0)$, there exists a Brownian Bridge $\{B_n(t): 0 \le t \le 1\}$ such that

$$\sup_{F^{-1}(\beta)\le x\le F^{-1}(1-\beta)} |f(x)q_n(x) - B_n(F(x))| \underset{a.s.}{=} O(n^{-\frac{1}{2}}\log n) \tag{4.4}$$

where

$$q_n(x) = n^{\frac{1}{2}}[X_{n,i} - F^{-1}(F(x))] \text{ for } \frac{i-1}{n} < F(x) \le \frac{i}{n},\ 1 \le i \le n. \tag{4.5}$$

By (4.4), (4.5), we have for $n \to \infty$

$$\max_{k_n+1\le i\le n-k_n} |\sqrt{n}[X_{n,i} - F^{-1}(\tfrac{i}{n+1})] - B(i/(n+1))f(F^{-1}(i/(n+1)))| \underset{a.s.}{=} O(n^{-\frac{1}{2}}\log n), \tag{4.6}$$

so that by (2.20), (2.21) and (2.38),

$$|\sqrt{n}\, L_n - \textstyle\sum_{i=1}^n b_{ni} B_n(i/(n+1))| \underset{a.s.}{=} O(n^{-\frac{1}{2}}\log n) \tag{4.7}$$

with b_{ni} given by (4.2).

$\{W_n(t) = (t+1)B_n(\frac{t}{t+1}): t \in R^+\}$ is a standard Wiener process on R^+, thus there exist $\{Y_{ni}\}_{i=1}^{n+1}$ of i.i.d. random variables with the standard normal distribution such that $W_n(m) = \sum_{i=1}^m Y_{ni}$, $m = 12, \ldots$.

Therefore,

$$\begin{aligned}\sqrt{n+1}\ B_n(\tfrac{i}{n+1}) &= \sqrt{n+1}[W_n(\tfrac{i}{n+1}) - \tfrac{i}{n+1}W_n(1)] \\ &= W_n(i) - \tfrac{i}{n+1}W_n(n+1) = \sum_{j=1}^{i} Y_{nj} - \tfrac{i}{n+1}\sum_{j=1}^{n+1} Y_{nj}, \qquad (4.8)\\ &\qquad i = 1,2,\ldots,n,\end{aligned}$$

so that

$$\sum_{i=1}^{n} B_n\left(\frac{i}{n+1}\right) = \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n} b_{ni}\left[\sum_{j=1}^{i} Y_{nj} + \frac{i}{n+1}\sum_{j=1}^{n+1} Y_{nj}\right] = \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} a_{nj} \cdot Y_{nj}. \tag{4.9}$$

(4.1) then follows from (4.7) and (4.9). Q.E.D.

LEMMA 4.1. *Under* (2.21), (2.38) - (2.41) *and* (2.42), *it holds for any positive integer* r,

$$\lim_{n\to\infty} E\left[\frac{1}{\sqrt{n+1}}\sum_{j=1}^{n+1} a_{nj} \cdot Y_{nj}\right]^{2r} = \nu_{(L)}^{2r} \cdot \frac{(2r)!}{r!\,2^r} \tag{4.10}$$

where $\nu_{(L)}^2$ *is defined by* (2.25).

PROOF. Note that $(n+1)^{-\frac{1}{2}}\sum_{j=1}^{n+1} a_{nj} \cdot Y_{nj}$ have a normal distribution with mean zero and variance

$$(n+1)^{-1}\sum_{j=1}^{n+1} a_{ni}^2 = \nu_{Ln}^{*2}, \text{ say.} \tag{4.11}$$

To prove (4.10), it thus suffices to show that

$$\lim_{n\to\infty} \nu_{Ln}^{*2} = \nu_{(L)}^2, \tag{4.12}$$

but it readily follows from (4.2), (2.21), (2.38) - (2.40).

LEMMA 4.2. *Under the assumptions of Therorem* 4.1, *for any positive integer* r, *there exists a* $C_r > 0$ *and an integer* n_r *such that*

$$E_0(\sqrt{n}\, L_n)^{2r} \le C_r < \infty \quad \forall\ n \ge n_r. \tag{4.13}$$

PROOF. Regarding (2.20) and the fact that $\sum_{i=1}^{n} c_{ni} F^{-1}\left(\frac{i}{n+1}\right) = 0$, we get by Jensen inequality that under $\theta = 0$,

$$(\sqrt{n}|L_n - \theta|)^{2r} = (\sqrt{n}|L_n|)^{2r} \le \left(\sum_{i=1}^{n} c_{ni}\sqrt{n}|X_{n,i} - F^{-1}\left(\frac{i}{n+1}\right)|\right)^{2r}$$
$$\le \sum_{i=k_n+1}^{n-k_n} c_{ni}(\sqrt{n}|X_{ni} - F^{-1}\left(\frac{i}{n+1}\right)|)^{2r} \tag{4.14}$$

so that

$$E_0(\sqrt{n}|L_n|)^{2r} \le \sum_{i=k_n+1}^{n-k_n} c_{ni} E(\sqrt{n}|X_{ni} - F^{-1}\left(\frac{i}{n+1}\right)|)^{2r} < C_r^* < \infty$$

holds for $n \ge n_r$, as it follows from Theorem 2 of Sen (1959).

THEOREM 4.2. *Let* L_n *be an* L-*estimator of the form* (2.20) *with the coefficients satisfying* (2.21), (2.38) - (2.41). *Let* d.f. F *have the absolutely continuous symmetric density satisfying* (2.29) *and* (2.42). *Then, for every positive integer* r,

$$\lim_{n\to\infty} E_\theta[\sqrt{n}(L_n - \theta)]^{2r} = \nu_{(L)}^{2r}\frac{(2r)!}{2^r r!} . \tag{4.15}$$

PROOF. It follows directly from Theorem 4.1, Lemma 4.1 and Lemma 4.2.

Let $\nu^2_{(L)}$ be the asymptotic variance (2.25) of $\sqrt{n}(L_n - \theta)$ and let $\hat{\nu}^2_{n(L)}$ be its estimator (2.26). Then we shall need the following

THEOREM 4.3. *Under the assumptions of Theorem* 4.2, *to any* $\varepsilon > 0$ *and* $\delta > 0$, *there exist* $C > 0$ *and* n_0 *such that, for* $n \geq n_0$,

$$P\{|\hat{\nu}^2_{n(L)} - \nu^2_{(L)}| > \varepsilon\} \leq Cn^{-1-\delta} . \tag{4.16}$$

PROOF. Let F_n be the empirical d.f. of $X_1,\ldots,X_n$. Then by (2.25), (2.26), (2.39) and (2.40),

$$\hat{\nu}^2_{n(L)} - \nu^2_{(L)} = \iint_{R^2}\Big\{[F_n(x\wedge y) - F_n(x)F_n(y)]J_n(F_n(x))J_n(F_n(y)) - [F(x\wedge y) - F(x)F(y)]J(F(x)J(f(y))\Big\}dxdy. \tag{4.17}$$

Now, for every $\eta > 0$,

$$P\{\sup_{x\in R}|F_n(x) - F(x)| > \eta\} \leq 2e^{-2n\eta^2}, \ \forall\ n \geq 1, \tag{4.18}$$

$$P\{X_{n,k_n+1} < F^{-1}(\alpha_0) - \eta\} = P\{X_{n,n-k_n} > F^{-1}(1-\alpha_0) + \eta\} \leq [\rho(\eta)]^n, \ \forall\ n \geq 1, \tag{4.19}$$

where $0 < \rho(\eta) < 1$. Also, excepting at countably many points (with Lebesgue measure 0), $J(t)$ has a derivative (with respect to t) inside $[\alpha_0, 1-\alpha_0]$ and $\sup\{|J(t)|: \alpha_0 \leq t \leq 1-\alpha_0\} \leq K < \infty$. Thus, by (4.19) and the fact that $c_{ni} - c_{n,n-i+1} = 0$, $\forall\ i \leq k_n$, with probability $\geq 1 - 2[\rho(\eta)]^n$, we may replace the domain R^2 in (4.17) by $\{F^{-1}(\alpha_0) - \eta,\ F^{-1}(1-\alpha_0) + \eta\}^2$, while on this compact region, (4.18), (2.39) - (2.40) and the boundedness of $|J(t)|$ lead us to the desired result when $J(t)$ is continuous inside $[\alpha_0, 1-\alpha_0]$. Next, let us suppose that $J(t)$ has only a finite number of saltus on $[\alpha_0, 1-\alpha_0]$. Excluding small neighborhoods of these saltus points, repeating the above proof and finally using the boundedness of $|J(t)|$ for these neighborhoods, the proof follows. Finally, $J(t)$ is a function of bounded variation, and hence, for any $\eta > 0$, $J(t)$ can have only a finite number of points of discontinuities at which its saltus is greater than η. Therefore, the proof for the case of a finite number of points of discontinuity extends to that of a countable number. Q.E.D.

5. PROPERTIES OF SEQUENTIAL M- AND L-ESTIMATORS

Let $X_1, X_2, \ldots$ be i.i.d. random variables distributed according to the d.f. $F(x-\theta)$ such that F is symmetric and satisfies regularity conditions (2.29) - (2.31). Unless otherwise stated, T_n will denote either M-estimator generated by the ψ-function satisfying (2.32) - (2.35) or the L-estimator with the coefficients satisfying (2.21), (2.38) - (2.41). ν^2 will denote the asymptotic variance of $\sqrt{n}(T_n - \theta)$ and $\hat{\nu}_n^2$ its estimator (2.15) or (2.26), respectively. Let N_c be the stopping variable defined in (2.6) and let T_{N_c} be the estimator based on $X_1, \ldots, X_{N_c}$.

THEOREM 5.1. *Under regularity conditions of Secton* 2, *for any* $h > 0$ (*in* (2.6)),

$$N_c/n_0(c) \xrightarrow{p} 1 \quad \text{and} \quad E(N_c/n_0(c)) \longrightarrow 1 \text{ as } c \downarrow 0 \tag{5.1}$$

where $n_0(c)$ *is defined by* (2.5);

$$\sqrt{n_0(c)}(T_{N_c} - \theta)/\nu \xrightarrow{\mathcal{D}} N(0, 1) \tag{5.2}$$

and

$$\lim_{c \downarrow 0}\{\lambda^*(a, c)/\lambda_{n_0(c)}(a, c)\} = 1 \tag{5.3}$$

where $\lambda^*(a, c)$ *and* $\lambda_{n_0(c)}$ *are given by* (2.7) *and* (2.4), *respectively.*

PROOF. Put $b = \left(\frac{a}{c}\right)^{\frac{1}{2}}$, so that $b \to \infty$ as $c \to 0$. Then, by (2.6),

$$N_c \ge b^{1/(1+h)} \text{ with probability 1.} \qquad (5.4)$$

For every $c > 0$ and ε: $0 < \varepsilon < 1$, put

$$n_c^* = [b^{1/(1+h)}], \quad n_{1c} = [n_0(c)(1-\varepsilon)] \text{ and } n_{2c} = [n_0(c)(1+\varepsilon)] \qquad (5.5)$$

where we choose c so small that $n' \le n_c^* < n_{1c} < n_0(c) < n_{2c}$. Then, by (2.6), Theorem 3.3 and Theorem 4.3 (on noting that $n/b \le \nu(1-\varepsilon)$, $\forall\ n \le n_{1c}$), as $c \downarrow 0$,

$$\begin{aligned} P\{N_c \le n_{1c}\} &= P\{\hat{\nu}_n \le n/b, \text{ for some } n\text{: } n_c^* \le n \le n_{1c}\} \\ &\le P\{\hat{\nu}_n \le \nu(1-\varepsilon), \text{ for some } n\text{: } n_c^* \le n \le n_{1c}\} \qquad (5.6) \\ &\le \textstyle\sum_{n=n_c^*}^{n_{1c}} P\{|\hat{\nu}_n - \nu| > \varepsilon\nu\} = 0[(n_c^*)^{-\delta}] = 0(c^{\delta/(2(1+h))}) \to 0. \end{aligned}$$

Similarly, noting that $n/b \ge \nu(1+\varepsilon)$, $\forall\ n \ge n_{2c}$, we have for $n \ge n_{2c}$,

$$\begin{aligned} P(N_c \ge n) &= P\{m < b(\hat{\nu}_m + m^{-h}),\ \forall\ n' \le m \le n\} \\ &\le P\{\hat{\nu}_n - \nu < \eta\} \quad (\text{where } \eta > 0) \qquad (5.7) \\ &\le P\{|\hat{\nu}_n - \nu| > \eta\} = 0(n^{-1-\delta}) \text{ by Theorems 3.3 and 3.4.} \end{aligned}$$

Then (5.6) and (5.7) imply that $N_c/n_0(c) \xrightarrow{p} 1$ as $c \downarrow 0$. Moreover, if $n \le n_{1c}$, then $n/n_0(c) < 1 - \varepsilon$ and, by (5.7), $E\{N_c \cdot I(N_c > n_{2c})\}/n_0(c) \longrightarrow 0$. as $c \downarrow 0$, so that $(EN_c)/n_0(c) \longrightarrow 1$ as $c \downarrow 0$. This proves (5.1).

Lemma 3.4 and Theorem 4.1 ensure that the distribution of $\{n^{\frac{1}{2}}(T_n - \theta)\}$ is asymptotically normal, $\mathcal{N}(0, \nu^2)$ as well as the "uniform continuity in probability" of this sequence in the sense of Anscombe (1952). Hence, (5.2) follows from Anscombe (1952) theorem.

Finally, to prove (5.3), we may follow the ideas of the proof of Theorem 3.2 of Sen (1980), where the Lemmas 3.1 - 3.6, 4.1, 4.2 and Theorem 3.1 - 3.3, 4.1 - 4.3 of Sections 3 and 4 provide the analogous tools to apply the same technique in the current context. Hence, for intended brevity, the details are omitted. Q.E.D.

It has been proved by Jurečková and Sen (1980 a, b) that

$$L\{n^d(\hat{\nu}_{n(M)} - \nu_{(M)})/\beta\} \longrightarrow N(0, 1) \quad \text{as} \quad n \to \infty \tag{5.8}$$

and

$$\sup_{m: |m-n| < \delta n} \{n^d |\hat{\nu}_{m(M)} - \nu_{n(M)}|\} \xrightarrow{p} 0 \quad \text{as} \quad \delta \downarrow 0 \tag{5.9}$$

where $\nu_{(M)}$ and $\nu_{n(M)}$ are given by (2.13) and (2.18), respectively, $\psi = \psi_1 + \psi_2$ with ψ_1 being the absolutely continuous component and ψ_2 the step-function component and where $d = \frac{1}{2}$ if $\psi_2 \equiv 0$ and $d = \frac{1}{4}$ if $\psi_2 \not\equiv 0$, and

$$\beta^2 = \begin{cases} [(\sigma_2^2/4\sigma^2_{(M)}) + \nu^2\sigma_1^2 - (\zeta/\gamma_1)]/\sigma_0^2 & \cdots \text{ if } d = \frac{1}{2} \\ \sum_{j=1}^{m} (\beta_j - \beta_{j-1})^2 f(a_j) & \cdots \text{ if } d = \frac{1}{4} \end{cases} \tag{5.10}$$

where

$$\sigma_2^2 = \int_{-\infty}^{\infty} \psi^4(x)dF(x) - \sigma^4_{(M)}, \quad \sigma_1^2 = \int_{-\infty}^{\infty} (\psi'(x))^2 dF(x) - \gamma^{(2)}_{(M)} \tag{5.11}$$

$$\zeta = \int_{-\infty}^{\infty} \psi^2(x)\psi'(x)dF(x) - \sigma^2_{(M)}\gamma_{(M)} \tag{5.12}$$

(in the case $\psi \equiv \psi_1$, where $\sigma^{(2)}_{(M)}$ and $\gamma_{(M)}$ are given by (2.13) and (2.14), respectively), and $a_1, \ldots, a_m$ are the jump-points of ψ_2 with jumps $(\beta_j - \beta_{j-1})$, $1 \le j \le m$.

THEOREM 5.2. *If the constant* h *in* (2.6) *satisfies* $h > d$, *then, under the regularity conditions on* F *and* ψ *of Section* 2, *as* $c \downarrow 0$,

$$L\{(n_0(c))^d[(N_c^{(M)}/n_0(c)) - 1]\} \to N(0, \beta^2/\nu^2). \tag{5.13}$$

PROOF. Note that by (2.6), whenever $N_c > n'$,

$$b\hat{\nu}_{N_c} \le N_c \le b(\hat{\nu}_{N_c - 1} + (N_c - 1)^{-h}) \tag{5.14}$$

so that if we put $n_{0c} = [b\nu] + 1$, we get from (2.5) and (5.14)

$$b(\nu_{N_c} - \nu) \leq N_c - n_{0c} \leq b(\hat{\nu}_{N_c - 1} - \nu) + b(N_c - 1)^{-h}, \tag{5.15}$$

whenever $N_c > n'$. Now, by Theorem 5.1, $N_c/n_{0c} \xrightarrow{p} 1$, while for $d < h$, $n_{0c}^{d}(N_c - 1)^{-h} \longrightarrow 0$ as $c \downarrow 0$, and

$$b^{-1} n_{0c}^{d} \sim \nu n_{0c}^{-1+d}, \quad \text{as} \quad c \downarrow 0. \tag{5.16}$$

Thus, regarding that $n_0(c) \sim n_{0c}$, (5.13) follows from (5.8), (5.9), (5.15) and (5.16). Q.E.D.

REMARK. Theorem 5.2 shows that, if $\psi_2 \equiv 0$, the rate of convergence to the asymptotic normal distribution in (5.13) is faster than in the case $\psi_2 \equiv 0$.

Let us now consider the asymptotic distribution of the stopping variable corresponding to the L-estimator. Gardiner and Sen (1979) proved

$$L\{n^{\frac{1}{2}}(\hat{\nu}_{n(L)} - \nu_{(L)}\} \longrightarrow N(0, \kappa^2/(4\nu_{(L)}^2)) \quad \text{as} \quad n \to \infty \tag{5.17}$$

and

$$\sup_{m: |m-n| < \delta n} \{n^{\frac{1}{2}} | \hat{\nu}_{m(L)} - \hat{\nu}_{n(L)} | \} \xrightarrow{p} 0 \quad \text{as} \quad \delta \downarrow 0 \tag{5.18}$$

where ν_L^2 and $\hat{\nu}_{n(L)}^2$ are given by (2.25) and (2.26), respectively, and

$$\kappa^2 = \int_0^1 \int_0^1 (s \wedge t - st) L_0(s) L_0(t) dF^{-1}(s) dF^{-1}(t) \tag{5.19}$$

with

$$L_0(t) = L_1(t) J'_{(1)}(t) - L_1(1-t) J'_{(1)}(1-t), \quad 0 \leq t \leq 1 \tag{5.20}$$

and

$$J_{(1)}(t) = t \cdot J(t), \quad L_1(t) = 2\int_0^{1-t} uJ(u) dF^{-1}(u), \quad 0 \leq t \leq 1. \tag{5.21}$$

THEOREM 5.3. *If the constant* h *in* (2.6) *satisfies* $h > \frac{1}{2}$, *then, under the regularity conditions on* F *and* J *of Section* 2, *as* $c \downarrow 0$,

$$L\{\sqrt{n_0(c)}[(N_c^{(L)}/n_0(c)) - 1]\} \longrightarrow N(0, (\kappa^2/4\nu_{(L)}^4)). \tag{5.22}$$

PROOF. (5.22) follows from (5.17) and (5.18) similarly as in the proof of Theorem 5.2.

6. ASYMPTOTIC MINIMAX PROPERTY OF SEQUENTIAL M- AND L-ESTIMATORS

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables distributed according to d.f. $F(x-\theta)$ where F is symmetric but generally unknown; F is only supposed to belong to an appropriate neighborhood $\mathcal{F}$ of a given d.f. G. Suppose that the loss incurred in estimating θ by T_n is given by (2.3). Let $\mathcal{T}$ denotes the set of sequences $\{T_n\}$ of translation equivariant estimators, asymptotically normally distributed and such that the minimum asymptotic risk $\lambda_{n_0(c)}(a, c)$ in (2.5) exists and satisfies

$$\lim_{c \downarrow 0} [\lambda^2_{n_0(c)}/(4ac)] = \nu^2(T_n, F) \quad \forall\, F \in \mathcal{F} \tag{6.1}$$

where $\nu^2(T_n, F)$ is the asymptotic variance of $\sqrt{n}(T_n - \theta)$ if $F(x-\theta)$ is the underlying d.f., and for which there exists sequential point estimation procedure T_{N_c} with the risk satisfying

$$\lim_{c \downarrow 0} [\lambda^*(c)/\lambda_{n_0(c)}(a, c)] = 1. \tag{6.2}$$

Then we may consider the limit

$$e(T_n, F) = \lim_{c \downarrow 0} \frac{\sqrt{c}}{\lambda^*(c)} \tag{6.3}$$

as a measure of efficiency of the sequential point estimator T_{N_c} if F is the underlying distribution.

Similarly as in the non-sequential estimation procedures, for an appropriate family $\mathcal{F}$ of distributions, there may exist an M- or L-estimator providing the saddle-point of the function (6.3) over $\mathcal{T} \times \mathcal{F}$. We shall formulate such result for the case that $\mathcal{F}$ represents the contaminated distribution G; it is an extension of the Huber's (1964) result to the sequential case.

Let $\mathcal{F}_1^*$ be the set of distribution functions

$$\mathcal{F}_1^* = \{F = (1-\varepsilon)G + \varepsilon H: \ H \in \mathcal{H}\} \tag{6.4}$$

where $\varepsilon \in [0, 1)$ is a fixed number, G is a symmetric d.f. which has twice continuously differentiable density g, g is strongly unimodal, $I(G) < \infty$ and $\int |x|^{\ell} dG(x) < \infty$ for some $\ell > 0$, while $\mathcal{H}$ is the set of all absolutely continuous symmetric d.f.'s with $\int |x|^{\ell} dH(x) < \infty$ for some $\ell > 0$. Then, for the M- and L-estimators considered in this paper, (6.1) - (6.2) hold for $F \in \mathcal{F}_1^*$. (6.1) - (6.2) also hold for the sample mean provided we assume that $F \in \mathcal{F}_2^*$, where $\mathcal{F}_2^* = \{F = (1-\varepsilon)G + \varepsilon H: \ H \in \mathcal{H}_2^*\}$ and $\mathcal{H}_2^* = \{H \in \mathcal{H}: \int x^2 dH(x) < \infty\}$ while G satisfies the same condition as in (6.4). Further, for translation-equivariant U-statistics, (6.1) - (6.2) hold [c.f., Sen and Ghosh (1980)] for $F \in \mathcal{F}_3^* = \{F: (1-\varepsilon)G + \varepsilon H: \ H \in \mathcal{H}_3^*\}$ where $\mathcal{H}_3^*$ is the class of all d.f.'s for which the kernel (generating the U-statistics) has finite ℓ-th absolute moment for some $\ell > 2$). Finally, (6.1) - (6.2) hold for a general class of R-estimators [c.f. Sen (1980)] with absolutely continuous (but possibly unbounded) score function, provided $F \in \mathcal{F}_4^*$ where $\mathcal{F}_4^*$ is a sub-class of $\mathcal{F}_1^*$ for which $\sup_x f(x)(F(x)[1-F(x)])^{-s} < \infty$, for some $s > 1/6$. It is easy to verify that the intersection of $\mathcal{F}_1^*$ $\mathcal{F}_2^*$, $\mathcal{F}_3^*$ and $\mathcal{F}_4^*$ is a non-null set.

THEOREM 6.1. *Let* $\mathcal{F}_1^*$ *be the set of distributions defined in* (6.4) *and let* $\mathcal{T}_1$ *be the set of sequential estimators satisfying* (6.1) *and* (6.2) *for* $F \in \mathcal{F}_1^*$. *Then there exist a sequential* M*-estimator* $T_{N_c}^{(1)}$ *and a sequential* L*-estimator* $T_{N_c}^{(2)}$ *and* $F_0 \in \mathcal{F}_1^*$ *such that*

$$e(T_{N_c}, F_0) \le e(T^{(i)}_{N_c}, F_0) = [I(F_0)]^{\frac{1}{2}} \le e(T^{(i)}_{N_c}, F) \tag{6.5}$$

holds for $i = 1,2$, $\forall\ T_{N_c} \in \mathcal{T}_1$ *and* $F \in \mathcal{F}^*_1$.

PROOF. It follows from Huber (1964) that the asymptotic variances of M-estimators with the ψ-functions satisfying (2.32) - (2.34) have a saddle-point which corresponds to the density

$$f_0(x) = \begin{cases} (1-\varepsilon)g(k)e^{q(x+k)}, & x < -k \\ (1-\varepsilon)g(x), & -k \le x \le k \\ (1-\varepsilon)g(k)e^{-q(x+k)}, & k \le x \end{cases} \tag{6.6}$$

and the corresponding minimax M-estimator is the maximum-likelihood estimator corresponding to f_0, i.e.,

$$\psi_0(x) = -\frac{f_0'(x)}{f_0(x)}, \quad x \in R, \quad \text{i.e.,} \tag{6.7}$$

$$\psi_0(x) = \begin{cases} -q & \cdots\ x < -k \\ -\dfrac{g'(x)}{g(x)} & \cdots\ -k < x < k \\ q & \cdots\ k < x \end{cases} \tag{6.8}$$

where $q = -\frac{g'(k)}{g(k)}$ and k is related to ε according

$$1/(1-\varepsilon) = \int_{-k}^{k} g(t)dt + 2g(k)/k\ . \tag{6.9}$$

It follows from Theorem 5.1 that the sequential M-estimator $T^{(1)}_{N_c^{(M)}}$ corresponding to ψ_0 and to the stopping rule (2.19) is the solution of (6.5).

Moreover, it follows from Jaeckel (1971) that the L-estimator $T^{(2)}_n$ corresponding to the weight function,

$$J_0(t) = [I(F_0)]^{-1} \cdot \psi_0'(F_0^{-1}(t)), \quad 0 \le t \le 1 \tag{6.10}$$

is asymptotically equivalent to $T_n^{(1)}$. The sequential L-estimator $T^{(2)}_{N_c(L)}$ with the stopping rule $N_c(L)$ given by (2.28) is then an alternative solution of (6.5).

REFERENCES

ANSCOMBE, F.J. (1952). Large sample theory of sequential estimation. *Proc. Camb. Phil. Soc.* **48**, 600-607.

CHOW, Y.S. and YU, K.F. (1980). The performance of a sequential procedure for the estimation of the mean. (Unpublished).

CSÖRGÖ, M. and RÉVÉSZ, P. (1978). Strong approximations of the quantile process. *Ann. Statist.* **6**, 882-894.

GARDINER, J.C. and SEN, P.K. (1979). Asymptotic normality of a variance estimator of a linear combination of a function of order statistics. *Zeit. Wahrsch. Verw. Geb.* **50**, 205-221.

GHOSH, M. and MUKHOPADHYAY, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Comm. Statist.* A**8**, 637-652.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *Jour. Amer. Statist. Assoc.* **58**, 13-29.

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

HUBER, P.J. (1969). Théorie de l'inférence statistique robuste. *Séminaire de mathématiques supérieures.* Université de Montréal.

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

HUBER, P.J. (1977). Robust methods of estimations of regression coefficients. *Opert. Statist. Ser. Stat.* **8**, 41-53.

JAECKEL, L.A. (1971). Robust estimates of location: Symmetry and asymmetric contamination. *Ann. Math. Statist.* **42**, 1020-1034.

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

JUREČKOVÁ, J. (1977). Asymptotic relations of M-estimates and R-estimates in linear regression model. *Ann. Statist.* **5**, 464-672.

JUREČKOVÁ, J. (1980). Asymptotic representation of M-estimators of location. *Opert. Statist. Ser. Stat.* **11**, (in press).

JUREČKOVÁ, J. and SEN, P.K. (1980a). Invariance principles for some stochastic processes relating to M-estimators and their role in sequential statistical inference. *Sankhya, Series A* (under consideration for publication).

JUREČKOVÁ, J. and SEN, P.K. (1980b). Sequential procedures based on M-estimators with discontinuous score functions, *Institute of Statistics Mimeo Series No. 1279.* The University of North Carolina, Chapel Hill, NC.

SEN, P.K. (1959). On the moments of sample quantiles. *Calcutta Statist. Assoc. Bull.* **9**, 1-19.

SEN, P.K. (1978). An invariance principle for linear combinations of order statistics. *Zeit. Wahrsch. Verw. Geb.* **42**, 327-340.

SEN, P.K. (1980). On nonparametric sequential point estimation of location based on general rank order statistics. *Sankhyā, Ser. A* **42** (to appear).

SEN, P.K. and GHOSH, M. (1980). Sequential point estimation of estimable parameters based on U-statistics (under consideration for publication).

von BAHR, B. (1965). On the convergence of moments in the central limit theorem. *Ann. Math. Statist.* **36**, 808-818.