ON SEQUENTIAL NONPARAMETRIC ESTIMATION OF MULTIVARIATE LOCATION

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1452

October 1983

ON SEQUENTIAL NONPARAMETRIC ESTIMATION OF MULTIVARIATE LOCATION

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514 USA

For the multivariate one-sample location model (relating to a diagonally symmetric distribution), sequential nonparametric (point as well as interval) estimators based on appropriate rank statistics are considered and their asymptotic properties studied. In this context, asymptotic risk-efficiency of the proposed estimators and asymptotic normality of the associated stopping times are established. These results rest on moment-convergence of the permutation dispersion matrix to its population counterpart; these are studied as well. A brief review of some other robust procedures is made along with.

INTRODUCTION

Let $\{\underset{\sim}{X}_i; i>1\}$ be a sequence of independent and identically distributed (i.i.d.) random vectors (r.v.) with a continuous distribution function (d.f.) $F(.;\underset{\sim}{\theta})$, defined on E^p, for some $p\geq 1$, where $\underset{\sim}{\theta}\epsilon\Theta\subset E^p$. It is assumed that

$$F(\underset{\sim}{x};\underset{\sim}{\theta}) = F(\underset{\sim}{x}-\underset{\sim}{\theta}), \quad x\epsilon E^p \quad , \tag{1}$$

where the p marginal d.f.s of F, denoted by $F_{[1]},\ldots,F_{[p]}$, respectively, are all symmetric about 0; in fact, F is assumed to be diagonally symmetric about the origin. Thus, $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_p)'$ is the vector of location parameters, and we are interested in the estimation of $\underset{\sim}{\theta}$. Based on a sample $(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n)$ of size n, let $\underset{\sim}{T}_n = (T_{n1},\ldots,T_{np})'$ be an estimator of $\underset{\sim}{\theta}$. The <u>loss</u> incurred due to estimating $\underset{\sim}{\theta}$ by $\underset{\sim}{T}_n$ is denoted by

$$L(n,c;\underset{\sim}{T}_n,\underset{\sim}{\theta}) = \rho(\underset{\sim}{T}_n,\underset{\sim}{\theta}) + cn \;, \quad (c>0) \tag{2}$$

where ρ: $E^p\times E^p \rightarrow [0,\infty)$ is a suitable metric and c is the cost of sampling per unit observation. Some commonly adapted forms of $\rho(.)$ are the following:

$$\begin{aligned} \rho(\underset{\sim}{a},\underset{\sim}{b}) &= \max\{|a_j-b_j|:\ 1\leq j\leq p\} && \text{(MAX-norm)} \\ \rho(\underset{\sim}{a},\underset{\sim}{b}) &= p^{-1}\Sigma_{j=1}^{p}|a_j-b_j| && \text{(MAD-norm)} \\ \rho(\underset{\sim}{a},\underset{\sim}{b}) &= (\underset{\sim}{a}-\underset{\sim}{b})'\underset{\sim}{Q}(\underset{\sim}{a}-\underset{\sim}{b}) && \text{(QUAD-norm)} \end{aligned} \tag{3}$$

where $\underset{\sim}{Q}$ is some given positive semi-definite (p.s.d.) matrix. Corresponding to (2), the risk in estimating $\underset{\sim}{\theta}$ by $\underset{\sim}{T}_n$ is given by

$$\lambda_F(n,c;\underset{\sim}{\theta}) = EL(n,c;\underset{\sim}{T}_n,\underset{\sim}{\theta}) = E\rho(\underset{\sim}{T}_n,\underset{\sim}{\theta}) + cn \quad . \tag{4}$$

Note that the risk in (4) depends on F through the distribution of $\underset{\sim}{T}_n$ and the metric $\rho(.)$. If T_n is translation-invariant (as will be the case in our subsequent analysis), then $\delta_n(F) = E\rho(\underset{\sim}{T}_n,\underset{\sim}{\theta})$ and $\lambda_F(n,c) = \lambda_F(n,c;\underset{\sim}{\theta})$ do not depend on $\underset{\sim}{\theta}$. Suppose now that there exists a positive integer n_0, such that $\delta_n(F)$ exists for every $n \geq n_0$ and $\delta_n(F) \downarrow 0$ as $n \to \infty$. For every $c>0$, we may define then

$$n_c^o = \min\{m \geq n_0: \ \lambda_F(m,c) = \inf_n \lambda_F(n,c)\} \quad . \tag{5}$$

Then, $\underset{\sim}{T}_{n_c^o}$ is the minimum risk estimator (MRE) of $\underset{\sim}{\theta}$. Note that $\delta_n(F)$ generally depends on F (through some other functionals of F), so that when F is not completely specified (as is the case in practice), no single n_c^o may lead to the MRE simultaneously for all $F \in \mathcal{F}$, a class of d.f.'s. However, suitable sequential procedures may be incorporated to achieve this goal in an asymptotic setup where c is made to converge to 0. Towards this, we assume that for some $q(>0)$, as n increases,

$$n^q \delta_n(F) \to \delta(F): \ 0<\delta(F)<\infty \quad , \tag{6}$$

and, further, there exists a sequence $\{d_n\}$ of consistent estimators of $\delta(F)$. [For the first two cases in (3), q is typically equal to ½, while for the last one, it is equal to 1.] Note that for $c \downarrow 0$, (4) may be rewritten as

$$cn + n^{-q}\delta(F) + o(n^{-q}) \quad , \tag{7}$$

so that asymptotically, as $c \downarrow 0$, we have

$$n_c^o \sim [(c^{-1}q\delta(F))^{1/(1+q)}] \quad , \tag{8}$$

$$\lambda_F^o(c) = \lambda_F(n_c^o,c) \sim \{c^q(q+q^{-q})\delta(F)\}^{1/(1+q)} \tag{9}$$

Motivated by (8) and the stochastic convergence of $\{d_n\}$ (to $\delta(F)$), for every $c(>0)$, we may define a stopping variable

$$N_c = \inf\{n \geq n_0: \ n^{1+q} \geq c^{-1}q(d_n+n^{-h})\} \quad , \tag{10}$$

where $h(>0)$ is a suitable constant, to be chosen later on.

Based on the stopping rule N_c, we consider the (sequential) point estimator $\underset{\sim}{T}_{N_c}$ of $\underset{\sim}{\theta}$, defined for every $c>0$, and we denote the corresponding risk by

$$\lambda_F^*(c) = E\rho(\underset{\sim}{T}_{N_c},\underset{\sim}{\theta}) + cEN_c \ , \quad c>0 \quad . \tag{11}$$

We say that $\underset{\sim}{T}_{N_c}$ is an asymptotically MRE of $\underset{\sim}{\theta}$, if

$$\lim_{c \downarrow 0}\{\lambda_F^*(c)/\lambda_F^o(c)\} = 1 \ , \quad \forall \ F \in \mathcal{F} \tag{12}$$

In the literature, this is referred to as the (first order) asymptotic risk-efficiency (A.R.E.) property. Besides this A.R.E., we shall also study the following results: as $c \downarrow 0$,

$$N_c/n_c^o \to 1 \text{ , in the first mean,} \tag{13}$$

$$(n_c^o)^{\frac{1}{2}}(\underset{\sim}{T}_{N_c} - \underset{\sim}{\theta}) \sim \mathcal{N}(\underset{\sim}{0}, \underset{\sim}{\Gamma}) \quad , \tag{14}$$

where

$$\underset{\sim}{\Gamma} = \lim_{n\to\infty} E\{n(\underset{\sim}{T}_n - \underset{\sim}{\theta})(\underset{\sim}{T}_n - \underset{\sim}{\theta})'\} \quad . \tag{15}$$

[$\underset{\sim}{\Gamma}$ is also the dispersion matrix of the asymptotic multi-normal d.f. of $n^{\frac{1}{2}}(\underset{\sim}{T}_n - \underset{\sim}{\theta})$.] Under some additional regularity conditions, we will also have the asymptotic normality of the stopping time i.e.,

$$(n_c^o)^{-\frac{1}{2}}(N_c - n_c^o) \sim \mathcal{N}(0, \gamma^2) \text{ , as } c \downarrow 0 \quad , \tag{16}$$

where $\gamma(0<\gamma<\infty)$ is a suitable constant (depending on q and F).

The second problem is to provide a confidence set $I_n(\subset E^p)$, based on a sample of size n, such that

$$P(\underset{\sim}{\theta} \in I_n | \underset{\sim}{\theta}) \geq 1-\alpha(0<\alpha<1) \quad , \tag{17}$$

$$\text{maximum diameter of } I_n \leq 2d(d>0) \quad , \tag{18}$$

where α and d are preassigned. For this problem too, fixed sample size procedures may not work out (for all $F \in \mathcal{F}$), and hence, one may take recourse a sequential scheme [viz., Chapter 10 of Sen (1981)], where (17)-(18) hold in an asymptotic setup: $d \downarrow 0$. Asymptotic properties of the stopping times are studied here.

In this paper, we specifically consider the case of R-estimators of location; some general remarks pertaining to some other robust (viz., M- and L-) estimators of location are made in the concluding section. Asymptotic properties of these multivariate R-estimators are studied in Section 2. Section 3 contains some asymptotic results on the permutation dispersion matrix which are needed in Section 4 in the derivation of the asymptotic results in (12, (13), (14) and (16). Section 5 deals with the confidence set problem. Throughout the paper, we consider the genuine multivariate case (i.e., $p \geq 2$). The univariate case has already been studied in detail in Chapter 10 of Sen (1981) and Jurečková and Sen (1982), among other places. The last section deals with some other robust estimators of $\underset{\sim}{\theta}$ (e.g., M- and L-estimators), and only the general setup is discussed.

MULTIVARIATE R-ESTIMATORS: ASYMPTOTIC PROPERTIES

For each j $(= 1,\dots,p)$, let $\phi_j^* = \{\phi_j^*(u), 0<u<1\}$ be a non-decreasing, non-constant, skew-symmetric and square-integrable score-function, and let

$$\phi_j = \{\phi_j(u) = \phi_j^*((1+u)/2), 0<u<1\} \text{ , } 1 \leq j \leq p \quad . \tag{19}$$

For a sample $\underset{\sim}{X}_i = (X_{i1},\ldots,X_{ip})'$, $1 \le i \le n$, of size n, we then introduce a set of <u>scores</u> by letting

$$a_{nj}(i) = E\phi_j(U_{ni}) \quad \text{or} \quad \phi_j(i/(n+1)) \ , \quad 1 \le i \le n;\ j = 1,\ldots,p \ , \tag{20}$$

where $U_{n1} < \ldots < U_{nn}$ are the ordered r.v.'s of a sample of size n from the uniform (0,1) d.f. Further, for every real b, let $R^+_{nij}(b)$ be the rank of $|X_{ij}-b|$ among $|X_{1j}-b|\ldots,|X_{nj}-b|$, for $i = 1,\ldots,n$ and $j = 1,\ldots,p$. Let $\underset{\sim}{S}_n(\underset{\sim}{b}) = (S_{n1}(b_1),\ldots,S_{np}(b_p))'$ be defined by

$$S_{nj}(b_j) = \Sigma_{i=1}^n \operatorname{sgn}(X_{ij}-b_j)a_{nj}(R^+_{nij}(b_j)) \ , \quad 1 \le j \le p \ . \tag{21}$$

Note that for each j (= $1,\ldots,p$), $S_{nj}(b)$ is $\searrow$ in b, and $S_{nj}(\theta_j)$ has a distribution symmetric about 0. Let then

$$\hat\theta_{nj} = \tfrac{1}{2}(\sup\{b: \ S_{nj}(b)>0\} + \inf\{b: \ S_{nj}(b)<0\}) \ , \tag{22}$$

for $j = 1,2,\ldots,p$, and let

$$\underset{\sim}{\hat\theta}_n = (\hat\theta_{n1},\ldots,\hat\theta_{np})' \ . \tag{23}$$

Then $\underset{\sim}{\hat\theta}_n$ is a translation-invariant, median-unbiased, robust and consistent estimator of $\underset{\sim}{\theta}$ [viz., Chapter 6 of Puri and Sen (1971)]. Let $\underset{\sim}{B} = \operatorname{Diag}(B_1,\ldots,B_p)$ be defined by letting

$$B_j = \int_{-\infty}^{\infty} (d/dx)\phi_j^*(F_{[j]}(x))dF_{[j]}(x) \ , \quad 1 \le j \le p \ . \tag{24}$$

Further, let $\underset{\sim}{\nu} = ((\nu_{jj'}))$ be defined by letting

$$\nu_{jj'} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_j^*(F_{[j]}(x))\phi_{j'}^*(F_{[j']}(y))dF_{[jj']}(x,y) \ , \tag{25}$$

for $j,j' = 1,\ldots,p$, where $F_{[jj']}$ is the bivariate d.f. of $(X_{ij}-\theta_j, X_{ij'}-\theta_{j'})$, for $j \ne j' = 1,\ldots,p$. Note that the ν_{jj} do not depend on F, but $\nu_{jj'}$, $j \ne j'$, are dependent on F through the joint d.f. $F_{[jj']}$, $j \ne j'$. Finally, let

$$\underset{\sim}{\Gamma} = ((\gamma_{jj'})) = \underset{\sim}{B}^{-1}\underset{\sim}{\nu}\underset{\sim}{B}^{-1} = ((\nu_{jj'}/B_jB_{j'})) \ . \tag{26}$$

Then, under quite general conditions [cf. Puri and Sen (1971, Ch. 6)], we have for $n\to\infty$,

$$n^{\frac{1}{2}}(\underset{\sim}{\hat\theta}_n - \underset{\sim}{\theta}) \sim \mathcal{N}(\underset{\sim}{0},\underset{\sim}{\Gamma}) \ . \tag{27}$$

Note that the asymptotic normality in (27) does not necessarily ensure (15). If, however, for some $a>0$ (not necessarily ≥ 1), $E||\underset{\sim}{X}_i||^a < \infty$, then, it follows from Theorem 2.1 of Sen (1980b) that there exists a positive integer n_a (depending on a), such that $E||\underset{\sim}{\hat\theta}_n - \underset{\sim}{\theta}||^2 < \infty$, $\forall\ n \ge n_a$. For our purpose, we assume that for every j (= $1,\ldots,p$), for some δ ($0 \le \delta < \frac{1}{4}$) and K ($0<K<\infty$),

$$|(d^r/du^r)\phi_j(u)| \le K(1-u)^{-\delta-r} \ , \quad 0<u<1;\ r = 0,1,2, \tag{28}$$

and $F_{[j]}$ has an absolutely continuous density function $f_{[j]}$ (with a first derivative $f'_{[j]}$ bounded almost everywhere), such that

$$\sup\{f_{[j]}(x)[F_{[j]}(x)\{1-F_{[j]}(x)\}]^{\delta-\eta}\} < \infty \ , \quad \text{for some} \quad \eta>0 \ , \tag{29}$$

for $j = 1,\ldots,p$. We write $\delta = (4+2\tau)^{-1}$, $\tau>0$. Then, by a direct multivariate extension of Theorem 2.2 of Sen (1980b), we may conclude that for every $k<2(1+\tau)$, $\overline{\lim} E\{n^{k/2}||\hat{\theta}_n-\theta||^k\}<\infty$. Note that throughout the paper $||\cdot||$ stands for the max-norm. Then (15) holds under (28)-(29). Actually, the following representation [a coordinatewise extension of (2.49) of Sen (1980b)] is the key to the above results:

$$n(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}) = \underset{\sim}{B}^{-1}\underset{\sim}{S}_n(\underset{\sim}{\theta}) + \underset{\sim}{\xi}_n \tag{30}$$

where, under (28)-(29), as $n\to\infty$,

$$n^{-\frac{1}{2}}||\underset{\sim}{\xi}_n|| \to 0 \ \text{ a.s. as well as in the kth mean,} \tag{31}$$

for every $k < 2(1+\tau)$. Further, $\{\underset{\sim}{S}_n(\underset{\sim}{\theta}), n\geq 1\}$ is a martingale sequence [see Sen and Ghosh (1971)], so that $\{||\underset{\sim}{S}_m(\underset{\sim}{\theta})-\underset{\sim}{S}_n(\underset{\sim}{\theta})||: \ m\geq n\}$ is a sub-martingale. Hence using the Kolmogorov maximal-inequality along with the fact that for every $m\geq n$, $E(\underset{\sim}{S}_m(\underset{\sim}{\theta})-\underset{\sim}{S}_n(\underset{\sim}{\theta}))(\underset{\sim}{S}_m(\underset{\sim}{\theta})-\underset{\sim}{S}_n(\underset{\sim}{\theta}))' = (m-n)\underset{\sim}{\nu}_n$, where $\underset{\sim}{\nu}_n\to\underset{\sim}{\nu}$ as $n\to\infty$, we obtain by some standard steps that for every $\varepsilon>0$ and $\eta>0$, there exist an n^* and an $\varepsilon^*(>0)$, such that for every $n\geq n^*$,

$$P\{\max_{m:\ |m-n|\leq\varepsilon^* n} n^{-\frac{1}{2}}||\underset{\sim}{S}_m(\underset{\sim}{\theta})-\underset{\sim}{S}_n(\underset{\sim}{\theta})||>\varepsilon\} < \eta \ . \tag{32}$$

Consequently, by using (30), (31) and (32), we obtain that

$$P\{\max_{m:\ |m-n|\leq\varepsilon^* n} n^{\frac{1}{2}}||\hat{\underset{\sim}{\theta}}_m-\hat{\underset{\sim}{\theta}}_n||>\varepsilon\} < \eta \ , \tag{33}$$

for every $n\geq n^{**}$. Clearly, if $\{N_n\}$ be any sequence of positive integer valued r.v.'s, such that $n^{-1}N_n\to 1$, in probability, as $n\to\infty$, then by (27) and (33), as $n\to\infty$,

$$n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_{N_n}-\underset{\sim}{\theta}) \sim N(\underset{\sim}{0},\underset{\sim}{\Gamma}) \ . \tag{34}$$

Defining $\underset{\sim}{B}$ as in (24), we let

$$\omega_n = \sup\{n^{-\frac{1}{2}}||\underset{\sim}{S}_n(\underset{\sim}{b})-\underset{\sim}{S}_n(\underset{\sim}{0})+n\underset{\sim}{B}\underset{\sim}{b}||: \ ||\underset{\sim}{b}||<n^{-\frac{1}{2}}\log n\} \ . \tag{35}$$

Then, as a direct multivariate generalization of (2.37) of Sen (1980b), we obtain that for $\delta\leq(4+2\tau)^{-1}$, $\tau>0$, there exist two positive constants c_1,c_2 and an integer n_o^*, such that

$$P\{\omega_n>n^{-c_1}(\log n)|\underset{\sim}{\theta}=\underset{\sim}{0}\}\leq c_2 n^{-1-\tau} \ , \quad \forall\ n\geq n_o^* \ . \tag{36}$$

This last result yields suitable estimates of $\underset{\sim}{B}$ along with their rates of convergence. Note that under $\theta_j = 0$, $S_{nj}(0)$ has a completely specified distribution, symmetric about 0, and hence, for every $\alpha(0<\alpha<1)$ and n, there exists a $C_{n,\alpha}^{(j)}$, such that $P\{|S_{nj}(0)|\geq C_{n,\alpha}^{(j)}|\theta_j=0\}\geq\alpha>P\{|S_{nj}(0)|>C_{n,\alpha}^{(j)}\}$, for $j = 1,\ldots,p$, where $n^{-\frac{1}{2}}C_{n,\alpha}^{(j)}\to\tau_{\alpha/2}\nu_{jj}^{\frac{1}{2}}$, $1\leq j\leq p$, $\tau_{\alpha/2}$ being the upper $50\alpha\%$ point of the standard normal distribution. Let then $\hat{\theta}_{L,n}^{(j)} = \sup\{b: \ S_{nj}(b)>C_{n,\alpha}^{(j)}\}$, $\hat{\theta}_{U,n}^{(j)} = \inf\{b: \ S_{nj}(b)<-C_{n,\alpha}^{(j)}\}$ and let

$$B_{nj} = 2C_{n,\alpha}^{(j)}/\{n(\hat{\theta}_{U,n}^{(j)} - \hat{\theta}_{L,n}^{(j)})\} \ , \quad j = 1,\ldots,p \ . \tag{37}$$

Then, $\underset{\sim}{B}_n = \mathrm{Diag}(B_{n1},\ldots,B_{np})$ is a translation-invariant, robust and consistent estimator of $\underset{\sim}{B}$. By (36), (37) and some standard manipulations, we obtain that for every $\varepsilon>0$ and $\delta \leq (4+2\tau)^{-1}$, $\tau>0$, there exist a constant C $(0<c<\infty)$ and an integer n_0, such that for every $n \geq n_0$,

$$P\{||\underset{\sim}{B}_n - \underset{\sim}{B}|| > \varepsilon\} \leq cn^{-1-\tau} \ . \tag{38}$$

Some other properties of $\underset{\sim}{B}_n$ are studied in the next section.

RANK BASED ESTIMATORS OF $\underset{\sim}{\Gamma}$: ASYMPTOTIC PROPERTIES

In (20), we replace the ϕ_j by ϕ_j^* and denote the resulting scores by $a_{nj}^*(i)$, $1 \leq i \leq n$, $1 \leq j \leq p$. Let then R_{nij} be the rank of X_{ij} among $X_{1j},\ldots,X_{nj}$, for $i = 1,\ldots,n$ and $j = 1,\ldots,p$. Define $\underset{\sim}{V}_n = ((v_{njj'}))$ by letting

$$v_{njj'} = n^{-1}\Sigma_{i=1}^{n} a_{nj}^*(R_{nij})a_{nj'}^*(R_{nij'}) \ , \quad j,j' = 1,\ldots,p \ . \tag{39}$$

Then $\underset{\sim}{V}_n$ is a translation-invariant, robust and consistent estimator of $\underset{\sim}{\nu}$, defined by (25) [See Puri and Sen (1971, Ch. 5)]. Defining the B_{nj} by (37), we let then

$$\hat{\underset{\sim}{\Gamma}}_n = \underset{\sim}{B}_n^{-1}\underset{\sim}{V}_n B_n^{-1} = ((\hat{\gamma}_{njj'})) = ((v_{njj'}/B_{nj}B_{nj'})) \ . \tag{40}$$

To study the asymptotic properties of $\hat{\underset{\sim}{\Gamma}}_n$, consider first an asymptotic representation for $\underset{\sim}{V}_n$. By (28)-(29) and Theorem 7.5.1 of Sen (1981), we obtain that for $\delta \leq (4+2\tau)^{-1}$, $\tau>0$,

$$\underset{\sim}{V}_n = n^{-1}\Sigma_{i=1}^{n}\underset{\sim}{W}_i + \underset{\sim}{\xi}_n^* \tag{41}$$

where the $\underset{\sim}{W}_i$ are i.i.d. random matrices with $E\underset{\sim}{W}_i = \underset{\sim}{\nu}$ and $E||\underset{\sim}{W}_i||^k < \infty$, $\forall\, k \in [0,2+\tau)$, and further $||\underset{\sim}{\xi}_n^*|| = O(n^{-\frac{1}{2}-\eta})$ a.s., as $n\to\infty$, for some $\eta>0$. Actually, by the same proof it follows that for every $\varepsilon>0$, there exist a finite c_1 and an integer n_0, such that for every $n \geq n_0$,

$$P\{||\underset{\sim}{\xi}_n^*|| > \varepsilon\} \geq c_1 n^{-1-\tau/2} \ . \tag{42}$$

Also, for the i.i.d. random matrices $\{\underset{\sim}{W}_i\}$, by the Markov inequality,

$$P\{||n^{-1}\Sigma_{i=1}^{n}\underset{\sim}{W}_i - \underset{\sim}{\nu}|| > \varepsilon\} \leq \varepsilon^{-k}E||n^{-1}\Sigma_{i=1}^{n}\underset{\sim}{W}_i - \underset{\sim}{\nu}||^k \leq c_2(\varepsilon)n^{-1-\tau/2} \ ; \quad c_2(\varepsilon)<\infty \ . \tag{43}$$

As such, from (38), (40), (41), (42) and (43), we obtain that for n adequately large, for every $\varepsilon>0$,

$$P\{||\hat{\underset{\sim}{\Gamma}}_n - \underset{\sim}{\Gamma}|| > \varepsilon\} \leq c_3 n^{-1-\tau/2} \ ; \quad c_3<\infty \ . \tag{44}$$

Let us now write $\underset{\sim}{U}_n = n^{\frac{1}{2}}(\underset{\sim}{V}_n - \underset{\sim}{\nu})$ and $\underset{\sim}{Y}_n = n^{\frac{1}{2}}(\underset{\sim}{B}_n^{-1} - \underset{\sim}{B}^{-1})$. Then, by (41) and the classical central limit theorem (on the $\underset{\sim}{W}_i$),

$$\underset{\sim}{U}_n \sim N(\underset{\sim}{0},\underset{\sim}{\Omega}_1) \ , \quad \text{as} \quad n\to\infty \ , \tag{45}$$

where $\underset{\sim}{\Omega}_1$ is the dispersion matrix of the $\underset{\sim}{W}_i$. The asymptotic normality results on $\underset{\sim}{Y}_n$ demands some extra regularity conditions. We assume that in (28), $|(d^2/du^2)\phi_j(u)| \leq K(1-u)^{-2}$, $0<u<1$ (where $K<\infty$) and in (29), we assume that for every $\eta>0$, there exists an $\varepsilon>0$, such that

$$\int_{-\infty}^{\infty} \{F_{[j]}(x)[1-F_{[j]}(x)]\}^{-1+\eta} dF_{[j]}(x \underline{+} \varepsilon) < \infty \;, \tag{46}$$

for every $j = 1,\ldots,p$. Then, we may use the second order linearity results of Hušková (1982), as adapted to the one-sample case, and conlcude that for $\underset{\sim}{B}_n^{-1}$, a representation similar to that in (41) holds, where the remainder term is $o(n^{-\frac{1}{2}})$ a.s., as $n\to\infty$. This representation ensures both the asymptotic multinormality of $n^{\frac{1}{2}}(\underset{\sim}{B}_n^{-1}-\underset{\sim}{B}^{-1})$ and the "uniform continuity in probability" of the $n^{\frac{1}{2}}(\underset{\sim}{B}_n^{-1}-\underset{\sim}{B}^{-1})$. Hence, under these additional regularity conditions,

$$n^{\frac{1}{2}}(\underset{\sim}{B}_n^{-1}-\underset{\sim}{B}^{-1}) \sim N(\underset{\sim}{0},\underset{\sim}{\Omega}_2) \tag{47}$$

for some p.s.d. $\underset{\sim}{\Omega}_2$. Writing $\underset{\sim}{B}_n^{-1} = \underset{\sim}{B}^{-1} + n^{-\frac{1}{2}}\underset{\sim}{Y}_n$ and $\underset{\sim}{V}_n = \underset{\sim}{\nu} + n^{-\frac{1}{2}}\underset{\sim}{U}_n$, we obtain on using (40), (45) and (47) that under the regularity conditions mentioned above

$$n^{\frac{1}{2}}(\hat{\underset{\sim}{\Gamma}}_n-\underset{\sim}{\Gamma}) = 2\underset{\sim}{Y}_n\underset{\sim}{\nu}\underset{\sim}{B}^{-1} + \underset{\sim}{B}^{-1}\underset{\sim}{U}_n\underset{\sim}{B}^{-1} + \underset{\sim}{O}_p(n^{-\frac{1}{2}}) \sim N(\underset{\sim}{0},\Omega^*) \;;\; \Omega^* \text{ p.s.d.} \tag{48}$$

Also, the asymptotic "uniform continuity in probability" result:

$$P\{\max_{m:\,|m-n|\leq\varepsilon' n} n^{\frac{1}{2}}||\hat{\underset{\sim}{\Gamma}}_m-\hat{\underset{\sim}{\Gamma}}_n|| > \varepsilon\} < \eta \;,\quad \forall\; n\geq n_0(\varepsilon,\eta) \tag{49}$$

follows from the above results, so that (48) extends as well to a sequence $\{N_n\}$ of integer valued random variables, where $n^{-1}N_n \to 1$, in probability, as $n\to\infty$.

ASYMPTOTIC PROPERTIES OF N_c AND $\hat{\underset{\sim}{\theta}}_{N_c}$

The asymptotic normality of the R-estimator has already been considered in (27). Also, after (29), the moment-convergence result in (15) has been studied. In the sequential case, the study of (11), (12), (13), (14) and (16) depends on the form of the metric $\rho(\cdot)$ in (2); we assume that the following holds:

(a) For $\delta_n = E\rho(\hat{\underset{\sim}{\theta}}_n,\underset{\sim}{\theta})$, (6) holds with $\delta(F)$ depending on F only through $\underset{\sim}{\Gamma}$ (which is a functional of F), i.e., $\delta(F) = \delta(\underset{\sim}{\Gamma})$, and further, $\delta(\underset{\sim}{\Gamma})$ satisfies the (Lipschitz-type) condition that

$$|\delta(\underset{\sim}{\Gamma}_1)-\delta(\underset{\sim}{\Gamma}_2)| \leq K_0||\underset{\sim}{\Gamma}_1-\underset{\sim}{\Gamma}_2||, \quad \forall\; \underset{\sim}{\Gamma}_1,\underset{\sim}{\Gamma}_2 \;, \tag{50}$$

where K_0 is independent of the $\underset{\sim}{\Gamma}_j$ and $||\cdot||$ is the max-norm.

(b) With q defined as in (6), there exists a finite, positive K_1, such that whenever the expectations exist,

$$E\{\rho(\hat{\underset{\sim}{\theta}}_n,\underset{\sim}{\theta})\}^r \leq K_1 E||\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}||^{2rq} \;,\quad (r\geq 1). \tag{51}$$

Note that (50)-(51) hold for each of the three metrics in (3). Also, note that in (9), we take $d_n = \delta(\hat{\underset{\sim}{\Gamma}}_n)$. The Lipschitz-type condition (50) along with (44) ensure that for n adequately large, for every $\varepsilon>0$,

$$P\{|d_n-\delta(F)|>\varepsilon\} \le c_4 n^{-1-\tau/2} \; ; \quad c_4<\infty \quad . \tag{52}$$

Now, by virtue of (44) and (52), we may virtually repeat the proof of the first part of Theorem 3.1 of Sen (1980b) and conclude that (13) holds. Further, (13) and (34) ensure (14). To establish the A.R.E. property in (12), we require to establish some uniform integrability properties, and, for these, we need to put some restrain on h in (10) and δ in (28). We let

$$\delta < (4+2\tau)^{-1} \; , \quad \tau > 1 + 2h \; , \quad h>0 \quad , \tag{53}$$

so that $\delta<1/6$. Note that by (13), $\lim_{c\downarrow 0} EN_c/n_c^o = 1$. Hence, by virtue of (4) and (11), it suffices to show that

$$\lim_{c\downarrow 0}\{E\rho(\hat{\underset{\sim}{\theta}}_{N_c},\underset{\sim}{\theta})/E\rho(\hat{\underset{\sim}{\theta}}_{n_c^o},\underset{\sim}{\theta})\} = 1 \quad , \tag{54}$$

or equivalently,

$$\lim_{c\downarrow 0}(n_c^o)^q E\rho(\hat{\underset{\sim}{\theta}}_{N_c},\underset{\sim}{\theta}) = \delta(F) = \delta(\underset{\sim}{\Gamma}) \quad . \tag{55}$$

By virtue of Theorem 2.2 of Sen (1980b), whenever, for some $a>0$, $E||\underset{\sim}{X}||^a<\infty$, under (28), (29) and (53),

$$\overline{\lim}\; E\{(n^{\frac{1}{2}}||\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}||)^k\} < \infty \; , \quad \forall\; k < 2(1+\tau) \quad . \tag{56}$$

Therefore, by (51) and (56),

$$\overline{\lim}\; \{n^{kq}E(\rho(\hat{\underset{\sim}{\theta}}_n,\underset{\sim}{\theta}))^k\} < \infty \; , \quad \forall\; k<1+\tau \quad . \tag{57}$$

With this, we may again follow the line of attack of the proof of Theorem 3.2 of Sen (1980b), employing the Hölder inequality for the two tails and the uniform integrability of $||\underset{\sim}{S}_n(\underset{\sim}{\theta})||^k$ $(k<2+\tau))$ for the central part. For intended brevity, the details are therefore omitted.

Finally, to prove (16), we note that by (10),

$$N_c \ge n_c^* = [(c^{-1}q)^{1/(1+q+h)}] \; , \quad \text{with probability 1} \quad , \tag{58}$$

where $n_c^*\to\infty$ as $c\downarrow 0$. Also, by (10), whenever, $N_c>n_c^*$,

$$c^{-1}q\delta(\hat{\underset{\sim}{\Gamma}}_{N_c}) \le N_c^{1+q} \; , \quad (N_c-1)^{q+1} < c^{-1}q(\delta(\hat{\underset{\sim}{\Gamma}}_{N_c-1}) + (N_c-1)^{-h}) \quad , \tag{59}$$

where by (8) and (59), taking $(n_c^o)^{q+1} = [c^{-1}q\delta(F)] + 1$,

$$(N_c/n_c^o)^{q+1}-1 \ge \{\delta(\hat{\underset{\sim}{\Gamma}}_{N_c}) - \delta(\underset{\sim}{\Gamma})\}/\delta(\underset{\sim}{\Gamma}) \quad , \tag{60}$$

$$((N_c-1)/n_c^o)^{q+1}-1 < \{\delta(\hat{\underset{\sim}{\Gamma}}_{N_c-1}) - \delta(\underset{\sim}{\Gamma})\}/\delta(\underset{\sim}{\Gamma}) + c^{-1}q(N_c-1)^{-h}(n_c^o)^{-q-1} \quad . \tag{61}$$

Note that for $h>\frac{1}{2}$, by (8) and (13), as $c\downarrow 0$,

$$c^{-1}q(N_c-1)^{-h}(n_c^o)^{-q-1} = O_p((n_c^o)^{-h}) = o_p((n_c^o)^{-\frac{1}{2}}) \quad , \tag{62}$$

so that from (13), (60), (61) and (62), as $c\downarrow 0$,

$$(n_c^o)^{\frac{1}{2}}\{(N_c/n_c^o)^{q+1}-1\} \sim (n_c^o)^{\frac{1}{2}}(\delta(\hat{\Gamma}_{N_c}) - \delta(\Gamma))/\delta(\Gamma) + o_p(1) \quad . \tag{63}$$

Further, by virtue of (49) and (50), under (13), as $c\downarrow 0$,

$$(n_c^o)^{\frac{1}{2}}\{\delta(\hat{\Gamma}_{N_c})-\delta(\Gamma)\} \sim (n_c^o)^{\frac{1}{2}}\{\delta(\hat{\Gamma}_{n_c^o})-\delta(\Gamma)\} \quad . \tag{64}$$

Hence, whenever $h>\frac{1}{2}$ [in (10)] and $n^{\frac{1}{2}}(\delta(\hat{\Gamma}_n)-\delta(\Gamma))$ is asymptotically normal, by (63) and (64), $(n_c^o)^{\frac{1}{2}}\{(N_c/n_c^o)^{q+1}-1\}$ is also asymptotically normal. Finally, by the Slutsky theorem, $(n_c^o)^{\frac{1}{2}}\{(N_c/n_c^o)^{q+1}-1\}$ asymptotically normal implies that $(n_c^o)^{\frac{1}{2}}\{N_c/n_c-1\}$ is also asymptotically normal (with a different asymptotic variance), and this proves (16). Note that for (64), one needs the extra regularity conditions due to Hušková (1982); for other results these are not needed. In any case, her regularity conditions are not very restrictive, and hold for the common types of score functions; we shall comment more on these in the last section.

SEQUENTIAL CONFIDENCE SETS

Note that defining ν as in (25) and $S_n(\theta)$ as in earlier, under θ,

$$n^{-1}(S_n(\theta))'\nu^{-1}(S_n(\theta)) \sim \chi_p^2 \quad , \tag{65}$$

while

$$\max_{1\le j\le p}\{n^{-1}S_{nj}^2(\theta_j)/\nu_{jj}\} \le n^{-1}(S_n(\theta))'\nu^{-1}(S_n(\theta)) \quad . \tag{66}$$

Hence, if $\chi_{p,\alpha}^2$ stands for the upper $100\alpha\%$ point of the chi-square distribution with p degrees of freedom, we have from (65) and (66),

$$P\{\max_{1\le j\le p} n^{-1}S_{nj}^2(\theta_j)/\nu_{jj}\le\chi_{p,\alpha}^2|\theta\} \ge 1-\alpha \quad , \tag{67}$$

when n is large. Following the steps after (36), we define then

$$\theta_{L,n}^{*(j)} = \sup\{b:\ S_{nj}(b)>n^{\frac{1}{2}}\nu_{jj}^{\frac{1}{2}}\chi_{p,\alpha}\} \quad , \tag{68}$$

$$\theta_{U,n}^{*(j)} = \inf\{b:\ S_{nj}(b)<-n^{-\frac{1}{2}}\nu_{jj}^{\frac{1}{2}}\chi_{p,\alpha}\} \quad , \tag{69}$$

for $j = 1,\ldots,p$. For every d (>0), consider then a stopping time

$$N_d^* = \inf\{n\ge n_0:\ \max_{1\le j\le p}(\theta_{U,n}^{*(j)}-\theta_{L,n}^{*(j)}) \le 2d\} \quad , \tag{70}$$

and the proposed (sequential) confidence set is

$$I_{N_d^*} = \{\theta:\ \theta_{L,N_d^*}^{*(j)} \le \theta_j \le \theta_{U,N_d^*}^{*(j)},\ 1\le j\le p\} \quad . \tag{71}$$

If we define, for every $d>0$,

$$n_d^* = \inf\{n\ge n_0:\ n \ge d^{-2}\chi_{p,\alpha}^2(\max_{1\le j\le p}\gamma_{jj})\} \quad , \tag{72}$$

then our basic concern is to verify (17) for $I_{N_d^*}$ in (71) (when $d\downarrow 0$), and to show that $N_d^*/n_d^*\to 1$, in probability, as $d\downarrow 0$. These results follow directly by using (35)-(36) [where we let $b = n^{\frac{1}{2}}(\theta_{L,n}^*-\theta)$ and $n^{\frac{1}{2}}(\theta_{U,n}^*-\theta)$], along the same line as in the univariate case [See Sen and Ghosh (1971)], and hence, the details are omitted.

We may also note that

$$\begin{aligned}&\sup\{n^{\frac{1}{2}}|\underset{\sim}{\ell}'(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta})|(\underset{\sim}{\ell}'\underset{\sim}{\ell})^{-\frac{1}{2}}:\ \underset{\sim}{\ell}\neq\underset{\sim}{0}\}\\ &=\sup\{n^{\frac{1}{2}}|\underset{\sim}{\ell}'\underset{\sim}{\Gamma}^{\frac{1}{2}}\underset{\sim}{\Gamma}^{-\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta})|(\underset{\sim}{\ell}'\underset{\sim}{\ell})^{-\frac{1}{2}}:\ \underset{\sim}{\ell}\neq\underset{\sim}{0}\} \qquad (73)\\ &\leq(ch_1(\underset{\sim}{\Gamma})\{n(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta})'\underset{\sim}{\Gamma}^{-1}(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta})\})^{\frac{1}{2}}\end{aligned}$$

where ch_1 stands for the largest characteristic root. Now, by (30)-(31), $n(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta})'\underset{\sim}{\Gamma}^{-1}(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}) = n^{-1}(\underset{\sim}{S}_n(\underset{\sim}{\theta}))'\underset{\sim}{\nu}^{-1}(\underset{\sim}{S}_n(\underset{\sim}{\theta})) + o(1)$ a.s., as $n\to\infty$. Hence, by (65) and (73), for large n

$$P\{\sup_{\underset{\sim}{\ell}\neq\underset{\sim}{0}}(\underset{\sim}{\ell}'\underset{\sim}{\ell})^{-\frac{1}{2}}|\underset{\sim}{\ell}'(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta})|\leq n^{-\frac{1}{2}}ch_1^{\frac{1}{2}}(\underset{\sim}{\Gamma})\chi_{p,\alpha}|\underset{\sim}{\theta}\}\approx 1-\alpha\ . \qquad (74)$$

Motivated by (74) and (44), we may define a stopping time N_d^o (for every $d>0$), by letting

$$N_d^o=\inf\{n\geq n_o:\ n\geq ch_1(\hat{\underset{\sim}{\Gamma}}_n)\chi^2_{p,\alpha}d^{-2}\}\ . \qquad (75)$$

In this case, if we define, for every $d>0$,

$$n_d^o=\min\{n\geq n_o:\ n\geq ch_1(\underset{\sim}{\Gamma})\chi^2_{p,\alpha}d^{-2}\}\ , \qquad (76)$$

then by (44),

$$N_d^o/n_d^o\to 1\ ,\quad \text{a.s., as}\quad d\downarrow 0\ , \qquad (77)$$

while (17) follows from (74), (76) and (33). The same method of proof of the asymptotic normality of the stopping time as in (58)-(64) works out here; we need (48)-(49) in this context, and in turn, these require the more stringent regularity conditions due to Hušková (1982).

SOME GENERAL REMARKS

In this paper, only the case of sequential R-estimators has been treated in detail. In the univariate case, Jurečková and Sen (1982) have considered sequential M- and L- estimators. Multivariate generalizations of these estimators follow on parallel lines. The coordinatewise uniform integrability of these estimators, established in Jurečková and Sen (1982) goes to the vector case as well. For the M-estimators, based on the vector $\underset{\sim}{\psi}$ of score functions, the sample (aligned) mean (score-) product matrix can be used to estimate the true mean product matrix, and under the assumption that $E||\underset{\sim}{\psi}||^r<\infty$, for some $r>4$, results parallel to (41)-(42) hold for this covariance matrix too. The estimator of the diagonal matrix with the elements $\int\psi_j'(x)dF_{[j]}(x)$, $1\leq j\leq p$ is the (diagonal) matrix with the marginal estimators, considered in Jurečková and Sen (1982), and hence, the convergence properties [parallel to (38) and (47)] hold under the same regularity conditions. In view of the fact that the score function $\underset{\sim}{\psi}$ are taken to be essentially bounded, the extra regularity conditions of Hušková (1982) are not needed here. Similarly, for the L-estimators, the (univariate) decomposition of Gardiner and Sen (1979) goes through in the multivariate case, and this yields the asymptotic normality as well as the rates of convergence of the estimated covariance matrices of such L-estimators. In this context, the jackknife procedure for obtaining this estimated covariance matrix also work out well; we may refer to Sen (1982) where the close connection of these estimators are discussed in detail.

We have considered only the first order A.R.E. property. Much more elaborate analysis will be needed to pursue higher order A.R.E. results. These are posed as open problems. We conclude this section with a closer examination of the Hušková-condition (46). It suffices to consider the right hand tail

$$G(a,d) = \int_a^\infty \{1-F(x)\}^{-1+\eta} dF(x-d) \ , \quad d>0,\ \eta>0,\ a>0 \ . \tag{78}$$

[Note that $G(a,-d) < G(a,0) = \eta^{-1}\{1-F(a)\}^{\eta} < \infty$, $\forall\ \eta>0,\ d>0$, while the lower tail will follow by symmetry.] We write

$$1 - F(x) = \exp\{-H(x)\} \quad \text{and} \quad h(x) = (d/dx)H(x) \ , \tag{79}$$

so that $h(x)$ is the <u>failure rate</u> (FR) and $H(x)$ is increasing in x with $H(\infty) = \infty$. Then, we have

$$G(z,d) = \int_{a-d}^\infty \exp\{(1-\eta)H(x+d)\} d(1 - \exp\{-H(x)\}) \ . \tag{80}$$

Hence, to verify (46), it suffices to show that for every $\eta>0$, there exists a d_η (>0), such that

$$\limsup_{x\to\infty}\{H(x+d)/H(x)\} < (1-\eta)^{-1} \ , \quad \forall\ 0<d\le d_\eta \ . \tag{81}$$

Towards this, we note that if

$$\limsup_{x\to\infty} h(x)/H(x) \le c < \infty \ , \tag{82}$$

then for all x (sufficiently large), $H(x+d)/H(x) \le \exp(cd)$, for every $d>0$. Consequently, choosing d_η (>0) such that $\exp(-cd_\eta) > 1-\eta$, (81) follows from (82). Hence, we may verify (82) as well. Note that $H(x)\uparrow\infty$ as $x\uparrow\infty$, so that for the entire class of d.f.'s with bounded or non-increasing FR, (82) holds trivially (with $c = 0$). For distributions with increasing failure rates (IFR), note that to verify (82), it suffices to show (by integration) that

$$\limsup_{x\to\infty}\{x^{-1}\log(-\log\{1-F(x)\})\} \le c < \infty \ , \tag{83}$$

and this general condition holds for almost all IFR d.f.'s (including some of the extreme value d.f.'s). Thus, (46) [via (82)-(83)] may not be regarded at all very restrictive. Finally, note that the condition that $|(d^2/du^2)\phi_j(u)| \le K(1-u)^{-2}$, $\forall\ 0<u\le 1$, holds for all the commonly adapted score functions (including particularly the normal scores and the Wilcoxon scores).

REFERENCES

[1] Gardiner, J.C. and Sen, P.K., Asymptotic normality of a variance estimator of a linear combination of a function of order statistics, Zeit. Wahrsch. Verw. Geb. 50 (1979), 205-221.

[2] Hušková, M., On bounded length sequential confidence interval for parameter in regression model based on ranks, Coll. Math. Soc. János Bolyai, 32: Nonparamet. Statist. Inf. (1982), 435-463.

[3] Hušková, M. and Jurečková, J., Second order asymptotic relations of M-estimators and R-estimators in two-sample location model, Jour. Statist. Plan. Infer. 5 (1981), 309-328.

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

[5] Puri, M.L. and Sen, P.K., Nonparametric Methods in Multivariate Analysis (John Wiley, New York, 1971).

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

[7] Sen, P.K., On almost sure linearity theorems for signed rank order statistics, Ann. Statist. 8 (1980), 313-321.

[8] Sen, P.K., On nonparametric sequential point estimation of location based on general rank order statistics, Sankhyā, Ser. A 42 (1980), 201-219.

[9] Sen, P.K., Sequential Nonparametrics: Invariance Principles and Statistical Inference (John Wiley, New York, 1981).

[10] Sen, P.K., Jackknife L-estimators: Affine structure and asymptotics, Institute of Statistics, University of North Carolina, Mimeo Report No. 1415.

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