ON SEQUENTIAL R-ESTIMATION OF LOCATION IN THE GENERAL BEHRENS-FISHER MODEL

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1439

June 1983

ON SEQUENTIAL R-ESTIMATION OF LOCATION IN THE GENERAL BEHRENS-FISHER MODEL

Pranab Kumar Sen

University of North Carolina, Chapel Hill.

Key Words and Phrases: Asymptotic risk-efficiency; Behrens-Fisher model; cost-function; point estimation; rank estimators; rank order statistics; risk; shift model; uniform integrability.

ABSTRACT

In the context of asymptotically minimum risk (sequential) point estimation of the shift (i.e., difference) of locations of two distributions, sequential analogues of the classical R-estimators are considered. Along with the uniform integrability and moment convergence of the classical (non-sequential) R-estimators, asymptotic risk and distribution of the allied stopping times for the sequential estimators are considered. The generalized form of the Behrens-Fisher problem (relating to the difference of locations of two symmetric distributions of possibly different forms) is presented and desirable sequential R-procedures are studied. In this context, the performance of the two-sample rank estimator is studied, and, side by side, an alternative approach (based on the difference of one-sample R-estimates) is also considered. Allied efficiency results are studied. The choice of (asymptotically) optimal score functions is discussed and the non-optimality of the usual two-sample estimates for the Behrens-Fisher model is studied.

1. INTRODUCTION

Let $\{X_i; i \geq 1\}$ be a sequence of independent and identically distributed random variables (i.i.d.r.v.) with a continuous distribution function (d.f.) F, defined on the real line $E = (-\infty,\infty)$. Also, let $\{Y_j, j \geq 1\}$ be a sequence of i.i.d.r.v. with a continuous d.f. G, defined on E. In the classical two-sample problem, one assumes that

$$G(x) = F(x-\Delta), \quad x \in E, \quad \Delta \text{ real}, \tag{1.1}$$

so that Δ stands for the *shift* (i.e., difference of locations). In the parametric case, the difference of the two sample means is taken as an estimator of Δ, while in the nonparametric case, some two-sample rank statistics may be employed to obtain some robust estimator of Δ -- these are called *R-estimators*. In the context of (asymptotically) *minimum risk point estimation* of Δ (based on an appropriate loss function incorporating the cost of sampling), procedures based on samples of predetermined sizes may not work out, and multi-stage or sequential procedures are advocated. For the one-sample location problem, a detailed discussion of these sequential procedures is given in Sen (1981, Ch. 10, Sec. 6). The first objective of the current investigation is to present sequential R-procedures for the two-sample problem, sketched above. In this setup, one needs to study the uniform integrability and moment-convergence properties of two-sample R-estimators, and these are considered here.

In a somewhat more general setup, one assumes that

$$F(x) = F_0(x-\theta_1), \quad G(x) = G_0(x-\theta_2), \quad x \in E, \quad \Delta = \theta_2 - \theta_1, \tag{1.2}$$

where F_0 and G_0 are both symmetric about 0, θ_1, θ_2 are the respective location parameters, and Δ is the difference of the location parameters. Here, F_0 and G_0 need not be of the same form. In the socalled Behrens-Fisher (BF-) model, one may take $G_0(x) = F_0(\nu x)$, $x \in E$, for some (unknown) $\nu > 0$; in the parametric setup, one additionally assumes that F_0 is a normal d.f. Ghosh and Mukhopadhyay (1979) have considered the BF-model wherein F_0 has not been restricted to the normal family and formulated sequential procedures leading to asymptotically minimum risk point estimates based on the sequence of two sample means and variances. In the non-sequential setup, the robustness of two-sample R-estimates for the generalized BF-model has been studied by Hoyland (1965) and Ramachandramurty (1966), among others. The second objective of the current study is to consider sequential counterparts of these estimates and to focus on their asymptotic risk and related properties.

For the classical shift-model in (1.1), along with the preliminary notions and basic regularity conditions, the R-estimators are introduced in Section 2. The uniform integrability and moment-convergence results for these estimators are studied in Section 3, and, in Section 4, these are incorporated in the study of the asymptotic minimum risk property of the proposed sequential estimates. Properties of these estimates under the general BF-model are then studied in Section 5. Section 6 deals with an alternative form of sequential estimates which behave very similarly to the ones in Section 2 for the model (1.1), but have more intuitive appeal for the general model (1.2). In this context, the one-sample results of Sen (1980) are generalized in a natural manner to the two-sample case, when the general B-F model may hold. The concluding section deals with some general comments on the proposed procedures.

2. SEQUENTIAL R-ESTIMATES FOR THE SHIFT MODEL

For the model (1.1), let $T_n = T(X_1,\dots,X_n;\ Y_1,\dots,Y_n)$ be a suitable estimator of Δ based on samples of sizes n from each of the two distributions. We assume that there exist an n_0 (≥ 1) and a positive, finite number ν^2, such that

$$\nu_n^2 = 2nE(T_n-\Delta)^2 \text{ exists for every } n\geq n_0; \tag{2.1}$$

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

Later on, we shall verify these for the proposed R-estimators. The *loss* in estimating Δ by T_n is taken as

$$L_n(a,c) = a(T_n-\Delta)^2 + c(2n), \quad a>0, \quad c>0, \tag{2.3}$$

so that the *risk* is equal to

$$\rho_n(a,c) = EL_n(a,c) = (2n)^{-1}a\nu_n^2 + 2cn, \ \forall\, n\geq n_0. \tag{2.4}$$

Our objective is to minimize this risk by a proper choice of n (where the total sample size is 2n); this optimal n, say, $n^0(a,c)$, depends on $\{\nu_n^2\}$ as well as on (a,c). Conveniently, in an asymptotic setup, where $c\downarrow 0$ (a fixed), this optimal choice can be derived readily by using (2.2). Without any loss of generality, we set $a = 1$ and $n^0(a,c) = n_c^0$. Then,

$$\lim_{c\downarrow 0}\{2n_c^0/(\nu/\sqrt{c})\} = 1 \quad \text{and} \quad \lim_{c\downarrow 0}\{\rho_{n_c^0}(1,c)/2\nu\sqrt{c}\} = 1. \tag{2.5}$$

For known ν, (2.5) provides the desired solution. When ν is not known, but a sequence $\{\hat{\nu}_n\}$ of consistent estimates $\hat{\nu}_n = \hat{\nu}(X_1,\ldots,X_n;\ Y_1,\ldots,Y_n)$ of ν is available, led by (2.5), we may consider a *stopping number* N_c, $c>0$, by letting

$$N_c = \min\{n\geq n_0:\ 2n\geq c^{-\frac{1}{2}}(\hat{\nu}_n + n^{-h})\} \tag{2.6}$$

(where h (>0) is a suitable constant) and adapt the sequential estimator T_{N_c} based on $X_1,\ldots,X_{N_c};\ Y_1,\ldots,Y_{N_c}$. The risk of estimating Δ by T_{N_c} is then equal to

$$\rho^*(c) = EL_{N_c}(1,c) = 2cEN_c + E(T_{N_c}-\Delta)^2. \tag{2.7}$$

We are primarily interested in showing that for a general class of two-sample rank-based (R)-estimators, we have

$$\lim_{c\downarrow 0}\rho^*(c)/\rho_{n_c^0}(1,c) = 1, \tag{2.8}$$

so that T_{N_c} is *asymptotically* (as $c\downarrow 0$) *risk-efficient*. Also, we shall study the asymptotic behaviour of the stopping number N_c.

First, we consider the following R-estimators of Δ. For every

n (≥ 1) let $R^Y_{n1},\ldots,R^Y_{nn}$ be the ranks of $Y_1,\ldots,Y_n$ among X_1, $X_1,\ldots,X_n$, $Y_1,\ldots,Y_n$. Also, for every m (≥ 1) let $a_m(1),\ldots,a_m(m)$ be a set of *scores* generated by an absolutely continuous, nondecreasing and square integrable score function $\phi = \{\phi(u),\ 0<u<1\}$ in the following way:

$$a_m(k) = E\phi(U_{mk}) \quad \text{or} \quad \phi(EU_{mk}),\quad k = 1,\ldots,m, \tag{2.9}$$

where $U_{m1}<\ldots<U_{mm}$ are the ordered r.v. of a sample of size m from the uniform (0,1) d.f. Without any loss of generality, we may set

$$\overline{\phi} = \int_0^1 \phi(u)du = 0 \quad \text{and} \quad A^2_\phi = \int_0^1 \phi^2(u)du = 1. \tag{2.10}$$

For every real d, let $R^Y_{n1}(d),\ldots,R^Y_{nn}(d)$ be the ranks of Y_1-d, $\ldots,Y_n-d$ among the $2n$ observations $X_1,\ldots,X_n$, $Y_1-d,\ldots,Y_n-d$, and let

$$S_n(d) = n^{-1}\Sigma^n_{i=1} a_{2n}(R^Y_{ni}(d)) - \overline{a}_{2n};\quad \overline{a}_{2n} = \frac{1}{2n}\Sigma^{2n}_{i=1} a_{2n}(i). \tag{2.11}$$

Then $S_n(d)$ is $\searrow$ in d, and we set

$$\hat{\Delta}_{n(R)} = \tfrac{1}{2}(\sup\{d:\ S_n(d)>0 + \inf\{d:\ S_n(d)<0\}). \tag{2.12}$$

$\hat{\Delta}_{n(R)} = T_n$ is the so-called R-estimator of Δ. It is a translation-invariant, consistent and robust estimator of Δ and under quite general regularity conditions, as $n\to\infty$,

$$(2n)^{\frac{1}{2}}(\hat{\Delta}_{n(R)}-\Delta) \sim N(0,\nu^2), \tag{2.13}$$

where $\nu^2 = 4A^2_\phi/\gamma^2 = 4\gamma^{-2}$, and

$$\gamma = \gamma(\phi,\psi) = \int_0^1 \phi(u)\psi(u)du; \tag{2.14}$$

$$\psi(u) = -f'(F^{-1}(u))/f(F^{-1}(u)),\quad 0<u<1; \tag{2.15}$$

it is assumed that the d.f. F has an absolutely continuous density function f with a finite Fisher information: $I(f) = \int (f'/f)^2 df < \infty$. Now ϕ is a specified function, but γ (and hence ν)

depend on the unknown F through ψ. Thus, we need to verify (2.1)-(2.2) first and also to estimate γ (or ν), so that N_c in (2.6) may be defined properly and we may be able to verify (2.8).

Note that under H_0: $\Delta = 0$, $S_n(0)$ has a distribution independent of the underlying F, and hence, for every $\alpha (0<\alpha<1)$, there exists two real values $S_{n,1}$ and $S_{n,U}$ such that $P\{S_{n,L} \leq S_n(0) \leq S_{n,U} | H_0\} = 1 - \alpha_n \ (\geq 1-\alpha)$, where α_n converges to α as n increases. Let then

$$\hat{\Delta}_{n,L} = \sup\{d: \ S_n(d) > S_{n,U}\}, \tag{2.16}$$

$$\hat{\Delta}_{n,U} = \inf\{d: \ S_n(d) < S_{n,L}\}, \tag{2.17}$$

$$\delta_n = \sqrt{2n}(\hat{\Delta}_{n,U} - \hat{\Delta}_{n,L}) \tag{2.18}$$

From the results in Sen (1981), we conclude then

$$\hat{\nu}_n = \hat{\delta}_n/(2\tau_{\alpha/2}) \underset{p}{\rightarrow} \nu, \quad \text{as} \quad n\rightarrow\infty, \tag{2.19}$$

where $\tau_{\alpha/2}$ is the upper $50\alpha\%$ point of the standard normal d.f. For the proposed sequential R-estimates, in (2.6), for $\hat{\nu}_n$, we use the estimator in (2.19). Thus, the proposed sequential R-estimator depends on the stopping rule in (2.6) based on the $\hat{\nu}_n$ in (2.19) and the point-estimation rule in (2.12).

3. MOMENT CONVERGENCE OF R-ESTIMATORS

For the study of (2.1)-(2.2) with possible generalization for the Behrens-Fisher model, first we consider the existence of the moments of $\hat{\Delta}_{n(R)}$, where (1.1) may not hold. Thus, we assume that $X_1,\ldots,X_n$ are i.i.d.r.v. with a d.f. F and $Y_1,\ldots,Y_n$ are i.i.d.r.v. with a d.f. G, where G and F may not satisfy (1.1). We assume that for some $a>0$ (not necessarily ≥ 1), $\int|x|^a dF < \infty$ and $\int|x|^a dG < \infty$.

Let $X_{n:1}<\ldots<X_{n:n}$ and $Y_{n:1}<\ldots<Y_{n:n}$ be the two sets of order statistics corresponding to $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$, respectively. On the scores $a_{2n}(k)$, $1\leq k\leq 2n$, we impose the same conditions as in Section 2, and, without any essential loss of

generality, let $\bar{a}_{2n} = 0$. Then, note that for every n, there exists a k $(=k_n)$ such that

$$\Sigma_{r=k+1}^{n} a_{2n}(r) + \Sigma_{r=2n-k+1}^{2n} a_{2j}(r) < 0, \tag{3.1}$$

$$\Sigma_{r=1}^{k} a_{2n}(r) + \Sigma_{r=n+1}^{2n-k} a_{2n}(r) > 0, \tag{3.2}$$

and further, $0 < \alpha_1 \leq \underline{\lim}\, n^{-1}k_n \leq \overline{\lim}\, n^{-1}k_n \leq \alpha_2 < 1$.

Let us now choose d in (2.11), $d = Y_{n:n-k} - X_{n:k} - 0$. Then

$$Y_j - d = Y_j - Y_{n:n-k} + X_{n:k} + 0 \begin{cases} > X_{n:k} & \text{for } k \text{ values of } j \\ < X_{n:k} & \text{for } (n-k) \text{ values of } j. \end{cases} \tag{3.3}$$

As a result, for $d = Y_{n:n-k} - X_{n:k} + 0$,

$$\begin{aligned} S_n(d) &= n^{-1}\{\text{sum of } (n-k) \text{ of } a_{2n}(1),\ldots,a_{2n}(n) \text{ plus} \\ &\qquad k \text{ of } a_{2n}(n+1),\ldots,a_{2n}(2n)\} \\ &\leq n^{-1}\{\Sigma_{r=k+1}^{n} a_{2n}(r) + \Sigma_{r=2n-k+1}^{2n} a_{2n}(r)\} \\ &< 0, \text{ by (3.1)}. \end{aligned} \tag{3.4}$$

Therefore, by (2.12) and (3.4), we conclude that

$$\hat{\Delta}_{n(R)} \leq Y_{n:n-k} - X_{n:k}. \tag{3.5}$$

Similarly, if we take $d = Y_{n:k} - X_{n:n-k} + 0$, we have by (3.2), $S_n(d) \geq n^{-1}\{\Sigma_{r=1}^{k} a_{2n}(r) + \Sigma_{r=n+1}^{2n-k} a_{2n}(r)\} > 0$, so that

$$\hat{\Delta}_{n(R)} \geq Y_{n:k} - X_{n:n-k}. \tag{3.6}$$

Hence, to show that $E|\hat{\Delta}_{n(R)}|^m < \infty$, for some m (>0) and all $n \geq n_0(m)$, it suffices to show that for F and G admitting finite moments up to the order a (>0), for every $0<\alpha_1<\alpha_2<1$, with $n^{-1}k_n \in (\alpha_1,\alpha_2)$, $E|X_{n:k_n}|^m < \infty$, $E|X_{n:n-k_n+1}|^m < \infty$, $E|Y_{n:k_n}|^m < \infty$ and $E|Y_{n:n-k_n+1}|^m < \infty$, for every $n \geq n_0(m,a)$, and these follow directly from Sen (1959).

Remarks. As regards the existence of the moments of $\hat{\Delta}_{n(k)}$ is concerned, we need not even take the two samples of equal size. A very similar argument works out when there are n_1 first sample and n_2 second sample observations, where $n_1/(n_1+n_2)$ is bounded away from 0 and 1. This case will be faced in the B-F model. This moment existence result extends directly Theorem 2.1 of Sen (1980) to the two-sample case.

Let us now return to the model (1.1) and verify (2.2). For this, we need to impose some additional regularity conditions which are stated below.

For the score function ϕ in (2.9), we assume that there exists a K $(0<K<\infty)$ and $\delta(>\frac{1}{4})$, such that for every $u \in (0,1)$,

$$|(d^r/du^r)\phi(u)| \leq K[u(u-u)]^{-\frac{1}{2}+\delta-r}, \quad r = 0,1,2. \tag{3.7}$$

Also, we assume that F belongs to the class F_0 of all absolutely continuous d.f.'s for which the p.d.f. f and its just derivative f' are bounded almost everywhere (a.e.) and

$$\lim_{x \to \pm\infty} f(x)\phi^{(1)}(F(x)) \text{ are finite.} \tag{3.8}$$

Then, we may virtually repeat the proof of Theorem 2.2 of Sen (1980) wherein we use Theorem A.4.2 of Sen (1981, p. 395) and conclude that if in (3.7), $\delta > (1+\tau)/(4+2\tau)$, for some $\tau>0$, then for every $k < 2(1+\tau)$,

$$\lim_{n\to\infty} E\{(2n)^{k/2}|\hat{\Delta}_{n(k)}-\Delta|^k\} = \nu^k E|Z|^k, \tag{3.9}$$

where Z has a standard normal distribution and $\nu^2 = 4\gamma^{-2}$ is defined by (2.14)-(2.15). As such, if we make use of (2.13), (3.9) and Theorem A.4.2 of Sen (1981), we obtain that under the assumed regularity conditions

$$2n\gamma(\hat{\Delta}_{n(k)}-\Delta) = \Sigma_{i=1}^{n}\{\phi(F(Y_i)) - \phi(F(X_i))\} + R_n^* \tag{3.10}$$

where $R_n^* = O(\sqrt{n})$ a.s., as $n\to\infty$, and

$$E|n^{-\frac{1}{2}}R_n^*|^k \to 0, \quad \forall\ k: \ 0 \leq k<2(1+\tau), \quad \tau>0. \tag{3.11}$$

These results are strengthened versions of the weaker representation results of Jurečková (1969), under more stringent regularity conditions.

4. PROPERTIES OF THE PROPOSED ESTIMATOR FOR THE SHIFT MODEL

By virtue of the results in Section 3, we are in a position to verify (2.8) and consider other asymptotic properties of N_c, for the shift model (1.1).

Theorem 4.1. Under (3.7)-(3.8), *for every* $\delta>\frac{1}{4}$ and for every $h\ (>0)$ *in* (2.6), *as* $c\downarrow 0$, *for the shift model* (1.1),

$$N_c/n_c^0 \xrightarrow{p} 1, \quad E(N_c/n_c^0)^k \to 1, \quad \forall\, k \in [0,1]; \tag{4.1}$$

$$(2n_c^0)^{\frac{1}{2}}(\hat{\Delta}_{N_c}(k)-\Delta)/\nu \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{4.2}$$

where $\nu^2 \equiv 4\gamma^{-2}$.

Theorem 4.2 If in (3.7), $\delta > (1+\tau)/2(2+\tau)$, *where* $\tau > 1 + 2h$ *and* $h\ (>0)$ *is defined as in* (2.6), *then for the shift model* (1.1), *the asymptotic risk efficiency in* (2.8) *holds.*

Proofs. Note that by virtue of (2.16)-(2.18) and Theorem A.6.1 of Sen (1981), we conclude that for every $\tau>0$, there exists an n_0, adequately large, so that for every $\varepsilon>0$,

$$P\{|\hat{\nu}_n-\nu|>\varepsilon\} \le c_\varepsilon n^{-1-\tau}, \quad \forall\, n\ge n_0, \tag{4.3}$$

where $c_\varepsilon\ (<\infty)$ may depend on ε. Also, by virtue of Theorem 4.3.1 of Sen (1981, p. 93), we may conclude that for very $\varepsilon>0$ and $\eta>0$, there exists a δ: $0<\delta<1$ and an $n_0\ (<\infty)$, such that for every $n\ge n_0$,

$$P\{\max\{n^{\frac{1}{2}}|S_m(0)-S_n(0)|:\ n-[\delta n]\le m\le n+[\delta n]\} > \varepsilon|H_0:\ \Delta=0\} < \eta. \tag{4.4}$$

Using (3.10) and (4.4), we conclude that

$$P\{\max\{n^{\frac{1}{2}}|\hat{\Delta}_{m(R)}-\hat{\Delta}_{n(R)}|:\ n-[\delta n]\le m\le n+[\delta n]\} > \varepsilon'|\Delta\} < \eta, \tag{4.5}$$

for every $n \geq n_0$, where $\varepsilon'(>0)$ depends on ε, and $\varepsilon' \to 0$ as $\varepsilon \to 0$. Once these results are obtained, the rest of the proofs of Theorems 4.1 and 4.2 follows virtually the lines of the proofs of Theorems 3.1 and 3.2 of Sen (1980), and hence, we omit the details.

It may be remarked that by definition in (2.6), whenever h is chosen as greater than $\frac{1}{2}$,

$$c^{-\frac{1}{2}}\hat{\nu}_{N_c} \leq 2N_c - c^{-\frac{1}{2}}N_c^{-h}, \tag{4.6}$$

$$c^{-\frac{1}{2}}\hat{\nu}_{N_c-1} > 2(N_c-1) - c^{-\frac{1}{2}}(N_c-1)^{-h}, \tag{4.7}$$

where by (2.5), $n_c^0 \sim \frac{1}{2}\nu c^{-\frac{1}{2}}$ (as $c \downarrow 0$), so that

$$\begin{aligned}(2n_c^0)^{\frac{1}{2}}(\hat{\nu}_{N_c}-\nu)/\nu &\leq (N_c-n_c^0)/\sqrt{n_c^0} \\ &\leq (2n_c^0)^{\frac{1}{2}}(\hat{\nu}_{N_c-1}-\nu)/\nu + c^{-\frac{1}{2}}(N_c-1)^{-h}/\sqrt{n_c^0}\,,\end{aligned} \tag{4.8}$$

where for $h>\frac{1}{2}$, $c^{-\frac{1}{2}}(N_c-1)^{-h}/\sqrt{n_c^0} \sim (2/\nu)\sqrt{n_c^0}\,(N_c-1)^{-h} \overset{p}{\to} 0$, as $c \downarrow 0$, by (4.1), and using a version of (4.5) for the lower as well as the upper confidence limits in (2.16)-(2.18), we may conclude that $(2n_c^0)^{\frac{1}{2}}|\hat{\nu}_{N_c}-\hat{\nu}_{N_c-1}| \overset{p}{\to} 0$, as $c \downarrow 0$. Hence, whenever, for $c \downarrow 0$,

$$(2n_c^0)^{\frac{1}{2}}(\hat{\nu}_{N_c}-\nu)/\nu \underset{\mathcal{D}}{\to} N(0,\beta^2), \tag{4.9}$$

for some β: $0<\beta<\infty$, we conclude from (4.8) that

$$(n_c^0)^{-\frac{1}{2}}(N_c-n_c^0) \underset{\mathcal{D}}{\to} N(0,\beta^2). \tag{4.10}$$

Fortunately, (4.9) follows from (4.1) and the recent results of Hušková (1982), and hence, the asymptotic normality of the stopping time N_c follows from (4.10).

5. PERFORMANCE OF $\Delta_{N(R)}$ FOR THE B-F MODEL

With respect to the generalized Behrens-Fisher model in (1.2), we shall study the performance of the sequential estimator $\hat{\Delta}_{N_c(R)}$ in (2.12) when N_c is defined by (2.6) with $\{\hat{\nu}_n\}$ in (2.19) It follows from the results of Ramachandramurty (1966) that, for the non-sequential case, $\hat{\Delta}_{n(R)}$ in (2.12) is a valid and consistent estimator of $\Delta(=\theta_2-\theta_1)$, and further, as $n\to\infty$,

$$(2n)^{\frac{1}{2}}(\hat{\Delta}_{n(R)}-\Delta) \underset{\mathcal{D}}{\to} N(0,\nu^{*2}), \tag{5.1}$$

where $\nu^* = 2A^*/\gamma^*$,

$$\begin{aligned} A^{*2} = & \iint_{-\infty<x<y<\infty} F_0(x)[1-F_0(y)]\phi^{(1)}(H_0(x))\phi^{(1)}(H_0(y))dG_0(x)dG_0(y) \\ & + \iint_{-\infty<x<y<\infty} G_0(x)[1-G_0(y)]\phi^{(1)}(H_0(x))\phi^{(1)}(H_0(y))dF_0(x)dF_0(y), \end{aligned} \tag{5.2}$$

$$\gamma^* = \int_{-\infty}^{\infty} \phi^{(1)}(H_0(x))f_0(x)g_0(x)dx, \tag{5.3}$$

and $H_0(x) = \frac{1}{2}(F_0(x)+G_0(x))$. Note that in this case, we assume that the score function ϕ is skew-symmetric (i.e., $\phi(u) + \phi(1-u) = 0$, $\forall\ 0<u<1$) -- this was not needed for the shift model in (1.1). The existence of the moments of $\hat{\Delta}_{n(R)}$ has already been established for the B-F model in Section 3. To obtain results analogous to (3.9)-(3.11), we define

$$\begin{aligned} S_n^* = \frac{1}{2n}\sum_{i=1}^{n}\{ & \int_{-\infty}^{\infty}[I(Y_i\le x)-G_0(x)]\phi^{(1)}(H_0(x))dF_0(x) \\ & - \int_{-\infty}^{\infty}[I(X_i\le x)-F_0(x)]\phi^{(1)}(H_0(x))dG_0(x)\}, \end{aligned} \tag{5.4}$$

and note that when (1.2) holds with $\Delta=0$,

$$ES_n^* = 0 \quad \text{and} \quad 2nE(S_n^{*2}) = A^{*2} \tag{5.5}$$

where A^{*2} is defined by (5.2). Further, under (3.7) with $\delta > (1+\tau)/2(2+\tau)$, $\tau>0$, using the fact that $dF_0(x) \le 2dH_0(x)$, $dG_0(x) \le 2dH_0(x)$, we have

$$n^{k/2}E\{|S_n^*|^k|\Delta=0\} < \infty, \quad \forall\ k \le 2(1+\tau) \tag{5.6}$$

Now, $S_n(d)$ in (2.11) is a particular case of a general linear rank statistic, and under the model (1.2), we may virtually repeat the proof of Theorem A.4.2 of Sen (1981, pp. 395-396), and letting (for $k>0$) $J_{nk} = \{d: |d| \le n^{-\frac{1}{2}}(\log n)^k\}$, we obtain that when $\Delta=0$, for every $\delta(<\frac{1}{4})$ and $\varepsilon>0$, there exist a positive integer n_0 and a finite positive constant $K_0(<\infty)$, such that

$$P\{\sup_{d\in J_{nk}} |S_n(d)-S_n(0)+\tfrac{1}{2}d\gamma^*| > \varepsilon/\sqrt{n}\} \le K_0 n^{-s}, \tag{5.7}$$

for every $n \ge n_0$, where $s < (1-2\delta)/2\delta$ and γ^* is defined by (5.3). Also, by the Cauchy-Schwarz inequality, $\forall$ d,

$$n^{\frac{1}{2}}|S_n(d)| \le \sqrt{2n}\,\{\tfrac{1}{n}\Sigma_{i=1}^{2n}[a_{2n}(i)-\bar{a}_{2n}]^2\}^{\frac{1}{2}}$$

$$= O(n^{\frac{1}{2}}), \text{ with probability 1}, \tag{5.8}$$

so that using (5.7), (5.8) and proceeding as in the proof of Theorem 2.2 of Sen (1980), we obtain that under the stated regularity conditions,

$$n\gamma^*(\hat{\Delta}_{n(R)}-\Delta) = 2nS_n(\Delta) + R_n^*, \tag{5.9}$$

where $(R_n^*)/\sqrt{n} \to 0$ a.s., as $n\to\infty$, and for $\delta > (1+\tau)/2(2+\tau)$, $\tau>0$,

$$E|n^{-\frac{1}{2}}R_n^*|^k \to 0, \quad \forall\, k < 2(1+\tau). \tag{5.10}$$

At this stage, we make use of the following result due to Ghosh (1972): For every $\delta = (1+\tau)/2(2+\tau)$, $\tau>0$, there exist a $K_0(<\infty)$ and an $n_0(<\infty)$, such that for every $n \ge n_0$,

$$P\{\sup_x \frac{n^{\frac{1}{2}}|F_{n0}(x)|-F(x)|}{F(x)[1-F(x)]\}^{\frac{1}{2}-\delta}} \ge K_0 \log n\} \le 2n^{-1-\tau} \tag{5.11}$$

where F_{n0} is the empirical d.f. based on $X_1,\ldots,X_n$; a similar result holds for the second sample $Y_1,\ldots,Y_n$ too. If $H_{n0} = \frac{1}{2}[F_{n0}+G_{n0}]$ stands for the combined sample empirical d.f., then under $\Delta=0$, by (5.11), on letting
$A_n = \{x: n^{\frac{1}{2}}|H_{n0}(x)-H_0(x)|/\{H_0(x)[1-H_0(x)]\}^{\frac{1}{2}-\delta} \le K_0 \log n\}$, we may write

$$E(n^{\frac{1}{2}}|S_n(0)-S_n^*|)^k = E\{n^{k/2}|S_n(0)-S_n^*|^k I(A_n)\}$$

$$+ E\{n^{k/2}|S_n(0)-S_n^*|^k I(A_n^c)\}, \qquad (5.12)$$

where by (5.6), (5.8) and (5.11), the second term on the right hand side of (5.12) converges to 0, as $n\to\infty$, whenever $k < 2(1+\tau)$. For the first term on the right hand side of (5.12), we use the Chernoff-Savage integral representation for $S_n(0)$ and by the usual partial integration claim that under the assumed regularity conditions, over the set A_n,

$$n^{\frac{1}{2}}|S_n(0)-S_n^*| \le K_0^* n^{-\frac{1}{2}}(\log n)^2; \quad K_0^* < \infty, \qquad (5.13)$$

so that (5.12) converges to 0, as $n\to\infty$. As a result, by (5.9), (5.10) and (5.12), we conclude that

$$n\gamma^*(\hat{\Delta}_{n(R)}-\Delta) = 2nS_n^* + R_n^{0*} \qquad (5.14)$$

where for $\delta > (1+\tau)/2(2+\tau)$, $\tau>0$,

$$E|n^{-\frac{1}{2}}R_n^{0*}|^k \to 0, \quad \forall\, k < 2(1+\tau). \qquad (5.15)$$

Finally, for S_n^*, involving two sets of i.i.d.r.v.'s, it follows by the well-known moment convergence results that

$$\lim_{n\to\infty}(2n)^{k/2}E(S_n^*)^k = A^{*k}E|Z|^k, \quad \forall\, k < 2(1+\tau), \qquad (5.16)$$

where Z has the standard normal d.f. As such, by (5.14), (5.15) and (5.16), we conclude that

$$\lim_{n\to\infty}(2n)^{k/2}E|\hat{\Delta}_{n(R)}-\Delta|^k = 2^k A^{*k}(\gamma^*)^{-k}E|Z|^k, \quad \forall\, k < 2(1+\tau). \qquad (5.17)$$

Note that for $S_{n,U}$ and $S_{n,L}$, defined in (2.16)-(2.17),

$$\lim_{n\to\infty}\sqrt{2n}\, S_{n,U} = \tau_{\alpha/2}, \quad \lim_{n\to\infty}\sqrt{2n}\, S_{n,L} = -\tau_{\alpha/2}, \qquad (5.18)$$

so that by (2.16)-(2.18) and (5.7), we conclude that under (1.2), $\hat{\delta}_n$ stochastically converges to $4(\gamma^*)^{-1}\tau_{\alpha/2}$, and hence, defining $\hat{\nu}_n$ as in (2.19), we have under (1.2),

$$\hat{\nu}_n \underset{p}{\to} 2/\gamma^* = \nu^*, \quad \text{say, as } n\to\infty. \tag{5.19}$$

As such, keeping (2.6) in mind, we may define

$$n_c^* \sim \tfrac{1}{2}\nu^* c^{-\frac{1}{2}} = (c^{\frac{1}{2}}\gamma^*)^{-1}, \quad \text{as } c \downarrow 0. \tag{5.20}$$

For the model (2.1), n_c^* plays a vital role in the asymptotic theory.

Note that $\{S_n^*\}$ in (5.4) form a reverse martingale (when in (1.2), $\Delta=0$), so that the "uniform continuity in probability" in (4.4) can easily be proved for the S_m^* under (1.2), and using (5.12)-(5.15) [insuring the a.s. convergence of $n^{\frac{1}{2}}|S_n^*-S_n(0)|$ to 0 as $n\to\infty$] along with the above, we conclude that (4.4) also holds for the model (1.2). By virtue of (5.7), (5.9) and (5.10), we have then (4.5) for the model (1.2). The proof of (4.3) wherein ν needs to be replaced by ν^* also follows along the same line as in the proof of (4.3), and hence, parallel to (4.1)-(4.2), we obtain that under the regularity conditions of Theorem 4.1, when (1.2) holds, as $c\downarrow 0$,

$$N_c/n_c^* \underset{p}{\to} 1, \quad E(N_c/n_c^*)^k \to 1, \quad \forall\, k \in [0,1] \tag{5.21}$$

$$(2n_c^*)^{\frac{1}{2}}(\hat{\Delta}_{N_c(R)}-\Delta) \underset{\mathcal{D}}{\to} \mathcal{N}(0,(2A^*/\gamma^*)^2), \tag{5.22}$$

and

$$\rho^*(c) = E(\hat{\Delta}_{N_c(R)}-\Delta)^2 + 2cEN_c \sim 2c^{\frac{1}{2}}(\gamma^*)^{-1}(1+A^{*2}). \tag{5.23}$$

The last expression will be useful in comparing the asymptotic risk of the sequential R-estimator $\hat{\Delta}_{N_c(R)}$ in Section 2, under (1.2), with some alternative ones to be considered in the next section. Note that for F_0 G_0, $\gamma^*=\gamma$ and $A^{*2}=1$, so that (5.23) agrees with (2.5), where $\nu=2/\gamma$.

6. ALTERNATIVE SEQUENTIAL R-ESTIMATORS FOR THE B-F MODEL

The estimator $\hat{\Delta}_{n(R)}$ in (2.12) (or its sequential version) is based on equal sample sizes for the two samples. For the general

B-F model in (1.2), this choice of equal sample sizes is rather counter-intuitive, and may lead to increased risk. For this reason, we consider here another generalization of the procedure in Sen (1980) based on signed-rank statistics.

As in Section 5, we assume that the score function ϕ is skew-symmetric and define $\phi^+(u) = \phi((1+u)/2)$, $0<u<1$. Let then $a_n^+(k)$ be defined by (2.9) with ϕ replaced by ϕ^+, and assume that (2.10) holds, i.e., $A^{+2} = \int_0^1 [\phi^+(u)]^2 du = A_\phi^2 = 1$. Also, let $R_{ni}^{X*}(b)$ (or $R_{ni}^{Y*}(b)$) be the rank of $|X_i - b|$ among $|X_1-b|,\ldots,|X_n-b|$ (or $|Y_i-b|$ among $|Y_1-b|,\ldots,|Y_n-b|$), for $i = 1,\ldots,n$, b real, and let

$$T_n^X(b) = n^{-1}\Sigma_{i=1}^n \mathrm{Sgn}(X_i-b)a_n^+(R_{ni}^{X*}(b)), \quad b \text{ real} \tag{6.1}$$

$$T_n^Y(b) = n^{-1}\Sigma_{i=1}^n \mathrm{Sgn}(Y_i-b)a_n^+(R_{ni}^{Y*}(b)). \tag{6.2}$$

Then, $T_n^X(b)$ and $T_n^Y(b)$ are both $\searrow$ in b, and we set

$$\hat{\theta}_{n,1} = \tfrac{1}{2}(\sup\{b:\ T_n^X(b)>0\} + \inf\{b:\ T_n^X(b)<0\}), \tag{6.3}$$

$$\hat{\theta}_{n,2} = \tfrac{1}{2}(\sup\{b:\ T_n^Y(b)>0\} + \inf\{b:\ T_n^Y(b)<0\}). \tag{6.4}$$

Corresponding to the densities f and g, ψ in (2.15) is denoted by ψ_f and ψ_g, respectively, and γ in (2.14) by γ_f and γ_g, respectively. Then, $\hat{\theta}_{n,1}$ (or $\hat{\theta}_{n,2}$) is a translation-invariant, consistent and robust estimator of θ_1 (or θ_2), and as $n\to\infty$,

$$n^{\frac{1}{2}}(\hat{\theta}_{n1}-\theta_1) \sim N(0,1/\gamma_f^2), \quad n^{\frac{1}{2}}(\hat{\theta}_{n2}-\theta_2) \sim N(0,1/\gamma_g^2) \tag{6.5}$$

Let now the first sample be of size n_1 and the second of size n_2 $(n = n_1 + n_2)$, so that we have the estimator of Δ

$$\hat{\Delta}_{n_1 n_2} = \hat{\theta}_{n_1,1} - \hat{\theta}_{n_2,2} \tag{6.6}$$

and under (2.1)-(2.2), for each estimator, we have for n_1, n_2 adequately large,

$$\rho_{n_1n_2}(c) = E(\hat{\Delta}_{n_1n_2}-\Delta)^2 + c(n_1+n_2)$$

$$= (n_1\gamma_f^2)^{-1} + (n_2\gamma_g^2)^{-1} + c(n_1+n_2) + 0(\tfrac{1}{n}) \tag{6.7}$$

Note that for known γ_f, γ_g, as $c\downarrow 0$, the optimal n_1, n_2 are given by

$$n_1^2\gamma_f^2 \sim n_2^2\gamma_g^2 \sim c^{-1}; \quad n = n_1+n_2 \sim c^{-\frac{1}{2}}(\gamma_f^{-1}+\gamma_g^{-1}). \tag{6.8}$$

Thus, if we let $n_{1c} \sim c^{-\frac{1}{2}}\gamma_f^{-1}$, $n_{2c} \sim c^{-\frac{1}{2}}\gamma_g^{-1}$, then for $c\downarrow 0$,

$$\rho_{n_{1c}n_{2c}}(c) \sim 2c^{\frac{1}{2}}(\gamma_f^{-1}+\gamma_g^{-1}), \tag{6.9}$$

and this represents the asymptotic minimum risk for the estimation based on the difference of individual R-estimates of location. Since γ_f and γ_g are unknown, we proceed as in Sen (1980) and using their estimates, consider the following stopping rule.

Note that under $\theta_1=0$ $(\theta_2=0)$, $T_n^X(0)$ $(T_n^Y(0))$ has a specified distribution, symmetric about 0, and hence, for every α $(0<\alpha<1)$, there exists a positive constant $T_{n,\alpha}$, such that

$$P\{|T_n^X(0)|\geq T_{n,\alpha}|\theta_1=0\} \geq \alpha > P\{|T_n^X(0)|>T_{n,\alpha}|\theta_1=0\}, \tag{6.10}$$

where $n^{\frac{1}{2}}T_{n,\alpha} \to \tau_{\alpha/2}$ as $n\to\infty$, and $T_n^X(0)$ may also be replaced by $T_n^Y(0)$. Defining the $T_n(b)$ as in (6.1)-(6.2), we let

$$\hat{\theta}_{n,L}^X = \sup\{b:\ T_n^X(b)>T_{n,\alpha}\},\quad \hat{\theta}_{n,L}^Y = \sup\{b:\ T_n^Y(b)>T_{n,\alpha}\}, \tag{6.11}$$

$$\hat{\theta}_{n,U}^X = \inf\{b:\ T_n^X(b)<-T_{n,\alpha}\},\quad \hat{\theta}_{n,U}^Y = \inf\{b:\ T_n^Y(b)<-T_{n,\alpha}\}; \tag{6.12}$$

$$V_n^X = (\hat{\theta}_{n,U}^X-\hat{\theta}_{n,L}^X)/2T_{n,\alpha} \sim \sqrt{n}\ (\hat{\theta}_{n,U}^X-\hat{\theta}_{n,L}^X)/2\tau_{\alpha/2}, \tag{6.13}$$

$$V_n^Y = (\hat{\theta}_{n,U}^Y-\hat{\theta}_{n,L}^Y)/2T_{n,\alpha} \sim \sqrt{n}\ (\hat{\theta}_{n,U}^Y-\hat{\theta}_{n,L}^Y)/2\tau_{\alpha/2}. \tag{6.14}$$

Let n_0 (≥ 2) be an initial sample size and h (>0) be some constant, to be defined more formally later on. Define then for every $c>0$

$$N_{1c} = \min\{n \geq n_0 : \; n \geq c^{-\frac{1}{2}}(V_n^X + n^{-h})\}, \tag{6.15}$$

$$N_{2c} = \min\{n \geq n_0 : \; n \geq c^{-\frac{1}{2}}(V_n^Y + n^{-h}) \;. \tag{6.16}$$

Then, the proposed sequential procedure consists in drawing one observation at a time from each of the two populations, until for the first time, one of the stopping numbers N_{1c}, N_{2c} is reached, and then drawing observations sequentially only from the population for which the stopping number has not been reached and terminating when both the stopping numbers are attained. The proposed sequential R-estimator of Δ is then

$$\hat{\Delta}_{N_{1c}N_{2c}} = \hat{\theta}_{N_{1c},1} - \hat{\theta}_{N_{2c},2}, \tag{6.17}$$

where the $\hat{\theta}_{n,1}, \hat{\theta}_{n,2}$ are defined by (6.3)-(6.4). The risk of $\hat{\Delta}_{N_{1c}N_{2c}}$ is

$$\begin{aligned}\rho_c^* &= E(\hat{\Delta}_{N_{1c}N_{2c}} - \Delta)^2 + cE(N_{1c} + N_{2c}) \\ &= E(\hat{\theta}_{N_{1c},1} - \theta_1)^2 + cEN_{1c} + E(\hat{\theta}_{N_{2c},2} - \theta_2)^2 + cEN_{2c} \\ &= \rho_{c,1}^* + \rho_{c,2}^*, \quad \text{say.}\end{aligned} \tag{6.18}$$

Note that with the definitions of the $\hat{\theta}$ and the N_c, Theorems 3.1 and 3.2 of Sen (1980) apply directly to the individual sample estimates and risks, so that we conclude that under the regularity conditions assumed in earlier sections, as $c \downarrow 0$, for $j = 1,2$,

$$N_{jc}/n_{jc} \to 1, \text{ in probability as well as the first mean,} \tag{6.19}$$

$$\rho_{c,1}^* \sim 2c^{\frac{1}{2}}\gamma_f^{-1}, \quad \rho_{c,2}^* \sim 2c^{\frac{1}{2}}\gamma_g^{-1}, \tag{6.20}$$

and the asymptotic normality of $(n_{1c}+n_{2c})^{\frac{1}{2}}(\hat{\Delta}_{N_{1c}N_{2c}} - \Delta)$ holds. Thus, it follows that as $c \downarrow 0$,

$$\rho_c^* / \rho_{n_{1c}n_{2c}}(c) \to 1, \tag{6.21}$$

which establishes the asymptotic risk-efficiency of the proposed sequential procedure for the general B-F model. Note that this asymptotic risk-efficiency is with respect to the opitmal non-sequential procedure based on the same R-estimators with the sample sizes n_{1c}, n_{2c}, if γ_f and γ_g were known.

It may be remarked that if in (1.2), we consider the usual location-scale model, viz., $G_0(x) = F_0(dx)$, for some $d>0$, then $g_0(x) \equiv df_0(dx)$, so that $I(g_0) = d^2 I(f_0)$ and $\gamma_g = d\gamma_f$. Thus, (6.9) is asymptotically equivalent to $2c^{\frac{1}{2}}\gamma_f^{-1}(1+d^{-1})$. In such a case, if we now choose $\phi(u) = k\psi_f(u)$, $0<u<1$, $k\neq 0$, where $\psi_f = \psi$ is defined by (2.15), then by (2.10), $k^2 = [I(f_0)]^{-1}$, and hence, by (2.14), $\gamma_f = [I(f_0)]^{\frac{1}{2}}$. Thus, for this specific choice of ϕ, the asymptotic risk is given by $2c^{\frac{1}{2}}[I(f_0)]^{-\frac{1}{2}}(1+d^{-1})$, which represents the lower bound (information limit) to the risk of any point estimator of Δ, for the location-scale (B-F) model and is attainable (under additional regularity conditions) by the maximum likelihood estimates of θ_1, θ_2. Thus, the proposed sequential procedure, based on the optimal score $\phi \equiv \psi_f$, becomes asymptotically optimal within a bound class of estimation rules. The choice of $\phi \equiv \psi_f$ in this context is in agreement with that of the conventional testing or estimation problems in the non-sequential case [viz., Hájek and Sidák (1967)]. For a given score function ϕ [viz., Wilcoxon: $\phi(u) = \sqrt{12}\,(u-\frac{1}{2})$, sign statistic: $\phi(u) = \mathrm{sgn}(\phi-\frac{1}{2})$, Normal scores: $\phi(u) = \Phi^{-1}(u)$, the inverse of the standard normal d.f.], when f is unknown (and hence, ψ_f is also so), we may write (6.9) as

$$2c^{\frac{1}{2}}\gamma_f^{-1}(1+d^{-1}) = \{2c^{\frac{1}{2}}[I(f_0)]^{-\frac{1}{2}}(1+d^{-1})\}\{I(f_0)/\gamma_f^2\}^{\frac{1}{2}}, \tag{6.22}$$

where the first factor represents the information limit, while by the Rao-Cramér bound, $\gamma_f^2 = \{\lim_{n\to\infty} nE(\hat{\theta}_{n,1}-\theta)^2\}^{-1} \leq I(f_0)$. Thus, $\gamma_f^{-2} I(f_0) \geq 1$, and this explains the effect of non-optimal choice of the score function ϕ, on the asymptotic risk in (6.9), (6.18)-(6.20). Recall that

$$e(\phi,\psi_f) = \gamma_f^2/I(f_0) \tag{6.23}$$

is the Pitman (asymptotic-) efficiency of the conventional rank statistic based on the score function ϕ with respect to the optimal score function ψ_f, and $\rho(\phi,\psi_f) \le 1$, where the equality sign holds when $\phi \equiv k\psi_f$, $(k\neq 0)$. As such, if we define the *asymptotic risk efficiency* by the ratio of the asymptotic risk for the optimal score (ψ_f) and the score ϕ, then, by (6.22) and (6.23), we have this equal to

$$\{e(\phi,\psi_f)\}^{\frac{1}{2}} = \sqrt{\text{Pitman efficiency}}. \tag{6.24}$$

With this identity, the results on the Pitman efficiency, studied in detail elsewhere [viz., Puri and Sen (1971)], provide us with the parallel ones for the location-scale (B-F) model under consideration. Ghosh and Mukhopadhyay (1979) have considered the B-F model and the asymptotically risk efficient sequential point estimation of Δ based on the sample means and variances. It follows that asymptotically the risk of their estimator is $2c^{\frac{1}{2}}(\sigma_X+\sigma_Y)$, where σ_X^2 and σ_Y^2 are the variances for the d.f. F and G, respectively. Note that under the location-scale model, $\sigma_Y^2 = d^{-2}\sigma_X^2$, so that the above reduces to $2c^{\frac{1}{2}}\sigma_X(1+d^{-1})$. Hence, the asymptotic risk-efficiency of the R-estimator with respect to the mean estimator is given by

$$2c^{\frac{1}{2}}(\sigma_X+\sigma_Y)\}/\{2c^{\frac{1}{2}}(\gamma_f^{-1}+\gamma_g^{-1})\}$$
$$= \sigma_X\gamma_f = \sqrt{\text{Pitman efficiency } (R_\phi, t)}, \tag{6.25}$$

where t stands for the Student t-test and R_ϕ for the rank test based on the score function ϕ. In particular, if we use the normal score statistics and the derived estimates, then it is well known that $\sigma_X^2\gamma_f^2 \ge 1$, where the equality sign holds only when F itself is a normal distribution. This explains the asymptotic superiority of the proposed sequential procedure over the conventional parametric procedure.

For the general B-F model in (1.2), we obtain, by (6.9) and (6.23), that as $c \downarrow 0$,

$$\rho_c^* \sim 2c^{\frac{1}{2}}(\gamma_f^{-1}+\gamma_g^{-1}) = 2c^{\frac{1}{2}}\{(e(\phi,\psi_f)I(f))^{-\frac{1}{2}} + (e(\phi,\psi_g)I(g))^{-\frac{1}{2}}\}$$

$$= 2c^{\frac{1}{2}}\{\sigma_X(\sigma_X^2 I(f)e(\phi,\psi_f))^{-\frac{1}{2}} + \sigma_Y(\sigma_Y^2 I(g)e(\phi,\psi_g))^{-\frac{1}{2}}\}, \qquad (6.26)$$

where by the Rao-Cramér inequality, $\sigma_X^2 I(f) \geq 1$ and $\sigma_Y^2 I(g) \geq 1$. Actually, $\sigma_X^2 I(f)\rho(\phi,\psi_f) = \gamma_f^2\sigma_X^2$, $\sigma_Y^2 I(g)\rho(\phi,\psi_g) = \gamma_g^2\sigma_Y^2$, and if we use the normal scores, then $\gamma_f^2\sigma_X^2 \geq 1$, $\sigma_Y^2\gamma_g^2 \geq 1$, where the equality sign holds (in either place) when the d.f. R (or G) is normal. Thus, comparing the first expression in (6.23) and (6.26), we again conclude that for the general B-F model in (1.2), the proposed normal scores prodedure is asymptotically more risk-efficient than the conventional procedure when at least one of F and G is different from a normal d.f.

7. SOME GENERAL REMARKS

We study now the relative performance of the sequential procedure in Section 2 and the alternative one in Section 6, when for both of them, the common score function ϕ is used. First, consider the shift model (1.1). For this model, $A^{*2}=1$ and $\gamma^* = \gamma_f = \gamma_g$. Thus (5.24) reduces to $4c^{\frac{1}{2}}\gamma_f^{-1}$. Also, by (6.18)-(6.20), the asymptotic risk of the sequential estimator in (6.17) reduces to $2c^{\frac{1}{2}}(\gamma_f^{-1}+\gamma_g^{-1}) = 4c^{\frac{1}{2}}\gamma_f^{-1}$. Thus, the two sequential procedures are asymptotically risk equivalent for the shift model (1.1). However, the two-sample approach does not require the symmetry of F, under (1.1), while for the validity of the one sample estimates in Section 6, the symmetry of F is a part of the assumptions for the alternative procedure in Section 6. Hence, for the shift model, the two-sample approach in Section 2 rests on relatively less stringent regularity conditions.

For the general B-F model in (1.2), symmetry of F_0, G_0 is needed for both the procedures. By (5.23) and (6.18)-(6.20), the *asymptotic risk efficiency* of the two-sample procedure in Section 2 relative to the alternative procedure in Section 6, both based on the common ϕ, is equal to

$$\{2c^{\frac{1}{2}}(\gamma_f^{-1}+\gamma_g^{-1})\}/\{2c^{\frac{1}{2}}(\gamma^*)^{-1}(1+A^{*2})\}$$

$$= \{\gamma^*(\gamma_f^{-1}+\gamma_g^{-1})/2\}\{\tfrac{1}{2}(1+A^{*2})\}^{-1}. \qquad (7.1)$$

Note that by the elementary inequalities among the arithmetic mean (A.M.), geometric mean (G.M.) and Harmonic mean (H.M.) of non-negative quantities, we have

$$\tfrac{1}{2}(\gamma_f+\gamma_g) \geq (\gamma_f\gamma_g)^{\frac{1}{2}} \geq \{\tfrac{1}{2}(\gamma_f^{-1}+\gamma_g^{-1})\}^{-1}, \qquad (7.2)$$

where in each place the equality sign holds when $\gamma_f=\gamma_g$. Thus, (7.1) is bounded from below by

$$\{\gamma^*/(\gamma_f\gamma_g)^{\frac{1}{2}}\}\{\tfrac{1}{2}(1+A^{*2})\}^{-1}. \qquad (7.3)$$

For the specific case of the Wilcoxon scores (i.e., $\phi(u) = \sqrt{12}(u-\frac{1}{2})$, $0<u<1$), $\phi^{(1)}\equiv\sqrt{12}$, so that $\gamma^* = \sqrt{12}\int f_0(x)g_0(x)dx = \sqrt{12} < f_0, g_0>$, $\gamma_f = \sqrt{12} < f_0, f_0>$ and $\gamma_g = \sqrt{12} < g_0, g_0>$. Further, $A^{*2} = 12\iint_{x\leq y}\{F_0(x)[1-F_0(y)]dG_0(x)dG_0(y) + G_0(x)[1-G_0(y)]dF_0(x)dF_0(y)\} \leq 3\iint_{x\leq y}\{dG_0(x)dG_0(y) + dF_0(x)dF_0(y)\} = 3$. Hence, (7.3) is bounded from below by

$$\tfrac{1}{2} < f_0, g_0 > /(<f_0,f_0><g_0,g_0>)^{\frac{1}{2}} = \tfrac{1}{2}\alpha(f_0,g_0), \qquad (7.4)$$

where $0\leq\alpha(f_0,g_0)\leq 1$.

To obtain an upper bound for (7.1), we make use of (5.17), for $k=2$, and the Rao-Cramér inequality (along with the fact that $\hat{\Delta}_{n(R)}$ is unbiased for Δ), and obtain that

$$(A^*/\gamma^*)^2 \geq \tfrac{1}{2}(\{I(f_0)\}^{-1} + \{I(g_0)\}^{-1})$$

$$\geq [\tfrac{1}{2}(\{I(f_0)\}^{-\frac{1}{2}} + \{I(g_0)\}^{-\frac{1}{2}})]^2$$

$$= [\tfrac{1}{2}(\gamma_f^{-1}\sqrt{e(\phi,\psi_f)} + \gamma_g^{-1}\sqrt{e(\phi,\psi_g)})]^2, \qquad (7.5)$$

where the last step follows from (6.23), and both the $\rho(\phi,\psi)$ are ≤ 1. As such, if we define

$$e^*(\phi,\psi_f,\psi_g) = \{\gamma^*\tfrac{1}{2}(\gamma_f^{-1}\sqrt{e(\phi,\psi_f)}+\gamma_g^{-1}\sqrt{e(\phi,\psi_g)})/A^*\}^2 \qquad (7.6)$$

and note that by (7.5), $e^*\leq 1$, we have from (7.5) and (7.6), that (7.1) is bounded from above by

$$\{\gamma^*\tfrac{1}{2}(\gamma_f^{-1}+\gamma_g^{-1}) /A^* = \tfrac{1}{2}(\gamma_f^{-1}+\gamma_g^{-1})/(A^*/\gamma^*)$$

$$= \{\tfrac{1}{2}(\gamma_f^{-1}+\gamma_g^{-1})\}\{e^*(\phi,\psi_f,\psi_g)\}^{\frac{1}{2}}\{\tfrac{1}{2}(\gamma_f^{-1}\sqrt{e(\phi,\psi_f)}+\gamma_g^{-1}\sqrt{e(\phi,\psi_g)})\}^{-1}$$

$$= \{e^*(\phi,\psi_f,\psi_g)\}^{\frac{1}{2}}(\gamma_f+\gamma_g)/\{\gamma_f\sqrt{e(\phi,\psi_g)}+\gamma_g\sqrt{e(\phi,\psi_g)}\}$$

$$\leq \{(e^*(\phi,\psi_f,\psi_g)/\min\{e(\phi,\psi_g),e(\phi,\psi_g)\})\}^{\frac{1}{2}} \qquad (7.7)$$

Now, for the location-scale model: $G_0(x) = F_0(dx)$ for some $d>0$, $\gamma_f = d^{-1}\gamma_g$, $e(\phi,\psi_f) = e(\phi,\psi_g)$, and hence, (7.7) reduces to

$$[e^*(\phi,\psi_f,\psi_g)/e(\phi,\psi_f)]^{\frac{1}{2}} \qquad (7.8)$$

Note that $\tfrac{1}{2}(1+A^{*2}) \geq A^*$ and the equality sign holds only when $A^*=1$. Thus, (7.1) is bounded from above by the first line of (7.7) where the equality sign holds when $A^*=1$, while the last equality sign in (7.7) holds for the location-scale model. Hence, (7.8) is an attainable upper bound only for $A^*=1$. In particular, when $\phi\equiv\psi_f$ $(\equiv\psi_g)$, $e(\phi,\psi_f) = 1$, while $e^*(\phi,\psi_f,\psi_g) \leq 1$, and hence, (7.8) is ≤ 1, where the equality sign holds [refer to (7.5)] when $I(f_0) = I(g_0)$ (i.e., $d=1$) and the first equality sign in (7.5) holds. Thus, for the location-scale model, for optimal ϕ, the two-sample approach in Section 2, for $d\neq 1$, does not have the asymptotic risk optimality, which is possessed by the alternative approach in Section 6. Note that for the location-scale model, by (7.5)

$$(A^*/\gamma^*)^2 \geq \tfrac{1}{2}\{I(f_0)\}^{-1}(1+d^{-2})$$

$$= \{e(\phi,\psi_f)/\gamma_f^2\}\tfrac{1}{2}(1+d^{-2}), \quad \text{by (6.23)}, \tag{7.9}$$

so that by (7.6) and (7.9), for the location-scale model,

$$e^*(\phi,\psi_f,\psi_g) \leq \frac{\{e(\phi,\psi_f)/\gamma_f^2\}\{\tfrac{1}{2}(1+d^{-1})\}^2}{\{e(\phi,\psi_f)/\gamma_f^2\}\{\tfrac{1}{2}(1+d^{-2})\}}$$

$$= (\tfrac{1}{2}(1+d^{-1}))^2/(\tfrac{1}{2}(1+d^{-2}))$$

$$= (d+1)^2/\{2d(1+d^2)\}, \tag{7.10}$$

which is equal to 1 for $d=1$ and otherwise is less than one, and it converges to 0 as $d\downarrow 0$ or $d\to\infty$. Thus, (7.8) is bounded from above by

$$[(d+1)^2/\{2d(1+d^2)\}][e(\phi,\psi_f)]^{-\frac{1}{2}} \tag{7.11}$$

and it clearly shows that for every ϕ, for which $e(\phi,\psi_f) > 0$, there exist two values of d, $0<d_0<1<d^0<\infty$, such that for $d \notin [d_0,d^0]$, (7.11) is ≤ 1, indicating the asymptotic risk-inefficiency of the two-sample approach for extreme scale variation.

For the general B-F model, $e^*(\phi,\psi_f,\psi_g)$ is the asymptotic relative efficiency of the rank order estimate $\hat{\Delta}_{n(R)}$ in (2.12) so that (7.7) is comparable to (6.24) or (6.25).

ACKNOWLEDGEMENT

This work was partially supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243-L.

BIBLIOGRAPHY

Ghosh, M. (1972). On the representation of linear functions of order statistics. *Sankhyā, Ser. A* 34, 349-356.

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

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* New York: Academic Press.

Hoyland, A. (1965). Robustness of the Hodges-Lehmann estimate for shift. *Ann. Math. Statist.* **36**, 174-197.

Hušková, M. (1982). On bounded length sequential confidence interval for parameter in regression model based on ranks. In *Coll. Nonparametric Infer., Janos Bolyai Math. Soc.* (ed: B.V. Gnedenks *et al.*) Amsterdam: North Holland, pp.

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

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

Ramachandramurty, P.V. (1966). On some nonparametric estimates for shift in the Behrens-Fisher situation. *Ann. Math. Statist.* **37**, 593-610.

Sen, P.K. (1959). On the moments of sample quantiles. *Calcutta Statist. Assoc. Bull.* **9**, 1-20.

Sen, P.K. (1980). On nonparametric sequential point estimation of location based on general rank order statistics. *Sankhyā, Ser. A*, **42**, 201-220.

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