SEQUENTIAL POINT ESTIMATION OF ESTIMABLE PARAMETERS
BASED ON U-STATISTICS

by

Pranab Kumar Sen
Department of Biostatistics
University of North Carolina, Chapel Hill

and

Malay Ghosh
Iowa State University

Institute of Statistics Mimeo Series No. 1236

June 1979

SEQUENTIAL POINT ESTIMATION OF ESTIMABLE PARAMETERS BASED ON U-STATISTICS

by

PRANAB KUMAR SEN[1]
University of North Carolina

MALAY GHOSH[2]
Iowa State University

ABSTRACT

Asymptotically risk-efficient sequential point estimation of regular functionals of distribution functions based on U-statistics is considered under appropriate regularity conditions. Some auxiliary results on U-statistics are also considered in this context.

AMS Subject Classification: 62L12, 62L15, 62G99.

Key Words & Phrases: Asymptotic normality, asymptotic risk-efficiency, estimable parameter, risk function, sequential estimation, stopping times, U-statistics.

---

[1]Work supported partially by the National Institutes of Health, Contract No. NIH-NHLBI-71-2243 from the National Heart, Lung and Blood Institute.

[2]Work supported by the Army Research Office, Durham Grant Number DAAG29-76-G-0057.

1. INTRODUCTION

Robbins (1959) initiated the study of the *sequential point estimation* of the mean of a normal distribution and this was later extended by Starr (1966) and Starr and Woodroofe (1969). Sequential point estimation of the scale parameter of a gamma distribution was considered by Starr and Woodroofe (1972), while the case of the multinormal mean vector was treated in Ghosh, Sinha and Mukhopadhyay (1976). The relevant properties of the normal and gamma distributions were fully exploited in the above papers.

Recently, Ghosh and Mukhopadhyay (1979) have proposed a sequential procedure for the point estimation of the mean of an unspecified distribution (admitting finite eight moments) and established its *asymptotic risk-efficiency* (to be defined in Section 2). Their *nonparametric appraoch* provides clues for further generalizations embracing a broader class of statistics and requiring less stringent conditions.

The object of the present investigation is to study nonparametric sequential point estimation of an *estimable parameter* based on U-statistics. In this context, the moment-condition of Ghosh and Mukhopadhyay (1979) is relaxed considerably and their results are extended to a broad class of U-statistics. Along with the preliminary notions, the main theorems are presented in Section 2. Two relevant theorems on U-statistics are also considered in this section. Section 3 is devoted to the proofs of the main theorems. Section 4 deals with some generalizations of these theorems along with some general remarks. The Appendix deals with the proofs of the theorems on U-statistics.

2. THE MAIN RESULTS

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed random variables (i.i.d.r.v.) with a distribution function (d.f.) F defined on the real q-plane R^q, for some $q \geq 1$. Let $\phi(X_1,\ldots,X_m)$, symmetric in its $m(\geq 1)$ arguments, be a Borel measurable *kernel* of *degree* m and consider the *estimable parameter* (a functional of the d.f. F)

$$\theta(F) = E\phi(X_1,\ldots,X_m)$$
$$= \int_{R^{qm}}\cdots\int \phi(x_1,\ldots,x_m)dF(x_1)\cdots dF(x_m), \quad F \in \mathcal{F}, \tag{2.1}$$

where $\mathcal{F} = \{F: |\theta(F)| < \infty\}$. Then, for $n \geq m$, the U-statistic U_n, corresponding to $\theta(F)$, is defined by [c.f. Hoeffding (1948)]

$$U_n = \binom{n}{m}^{-1} \sum_{C_{n,m}} \phi(X_{i_1},\ldots,X_{i_m}); \quad C_{n,m} = \{i_1,\ldots,i_m: 1 \leq i_1 < \cdots < i_m \leq n\} \tag{2.2}$$

Note that U_n is symmetric in $X_1,\ldots,X_n$ and is unbiased for $\theta(F)$. Let then $\phi_d(x_1,\ldots,x_d) = E\phi(x_1,\ldots,x_d, X_{d+1},\ldots,X_m)$, $0 \leq d \leq m$ and let

$$\zeta_d = E\phi_d^2(X_1,\ldots,X_d) - \theta^2(F), \quad 0 \leq d \leq m \quad (\zeta_0 = 0). \tag{2.3}$$

Then, whenever $n \geq m$ and $E\phi^2 < \infty$

$$\sigma_n^2 = \mathrm{Var}(U_n) = \binom{n}{m}^{-1} \sum_{d=1}^{m} \binom{m}{d}\binom{n-m}{m-d}\zeta_d . \tag{2.4}$$

Note that by the reverse martingale property of $\{U_n, n \geq m\}$ $\sigma_n^2 - \sigma_{n+1}^2 = \mathrm{Var}(U_n - U_{n+1}) \geq 0$, so that σ_n^2 is $\downarrow$ in $n(\geq m)$.

To motivate the sequential procedure, suppose that the *loss* incurred in estimating $\theta(F)$ by U_n is

$$L_n = a[U_n - \theta(F)]^2 + cn; \quad a > 0, \quad c > 0, \tag{2.5}$$

where a and c (cost per unit sample) are specified constants. The object is the minimize the *risk* (for given a, c)

$$R_c(n;\ a,\ F) = EL_n = a\sigma_n^2 + cn, \tag{2.6}$$

by a proper choice of n. Towards this, we have the following

Lemma 2.1. For every $a > 0$, $c > 0$ *and* $m \geq 1$, *whenever* $E\phi^2 < \infty$

$R_c(n;\ a,\ F)$ *is a convex function of* n.

The proof of the lemma is given in the Appendix. Note that by Lemma 2.1, there exists an n_c^* $(= n^*(a,\ c;\ F))$, such that

$$\begin{aligned} R_c(n_c^*;\ a,\ F) &= \min_n R_c(n;\ a,\ F) \\ &= a\sigma_{n_c^*}^2 + cn_c^*\ , \end{aligned} \tag{2.7}$$

where in (2.7), the minimization is restricted to integers $n \geq m$ and thereby, n_c^* is also an integer $(\geq m)$, though it need not be unique (there may be two consecutive values of n_c^* for which (2.7) holds).

From (2.4), (2.6) and (2.7), it follows that n_c^* depends on a, c, m as well as ζ_d, $1 \leq d \leq m$, where the later parameters are all (unknown) functionals of the (unspecified) d.f. F. Thus, in the absence of knowledge of these ζ_d, $1 \leq d \leq m$, no fixed sample size minimizes the risk simultaneously for all ζ_d, $1 \leq d \leq m$, and hence, a sequential procedure may be desirable to achieve this goal.

We assume that $\theta(F)$ is *stationary of order* 0 [viz. Hoeffding (1948)], so that

$$0 < \zeta_1 \leq \frac{1}{m}\zeta_m < \infty\ . \tag{2.8}$$

Note that by (2.4), (2.8) and Theorem 5.2 of Hoeffding (1948),

$$\sigma_n^2 = m^2 n^{-1}\zeta_1 + \xi_{(n)} \quad \text{where} \quad \xi_{(n)} = 0(n^{-2}) \tag{2.9}$$

and $n\xi_{(n)} \downarrow 0$ as $n \uparrow \infty$. Suppose that in (2.6), we neglect (for σ_n^2) the contribution of $\xi_{(n)}$ and, then in (2.7), we denote the resulting solution by n_c^o, so that we have for small c,

$$n_c^o \sim (c^{-1}am^2\zeta_1)^{\frac{1}{2}} \text{ and } R_c(n_c^o;\ a,\ F) \sim 2cn_c^o\ , \tag{2.10}$$

where $g(c) \sim h(c)$ means that $g(c)/h(c) \to 1$ as $c \to 0$. Using (2.4) and (2.7), it is possible to write $R_c(n_c^*;\ a,\ F) = \sum_{i=1}^{\infty} d_i(n_c^*)^{-i} + cn_c^*$, where $d_1 = am^2\zeta_1$ and $\sum_{i=1}^{\infty} id_i(n_c^*)^{-i-1} \sim c$, so that $c(n_c^*)^2 \sim a^2m^2\zeta_1$. Hence,

$$\lim_{c\downarrow 0} n_c^*/n_c^o = 1 \text{ and } \lim_{c\downarrow 0} R_c(n_c^*;\ a,\ F)/R_c(n_c^o;\ a,\ F) = 1, \tag{2.11}$$

whenever (2.8) holds. Hence, in the sequel, we shall occasionally interchange n_c^* and n_c^o for the convenience of our manipulations.

For the proposed sequential procedure, we proceed to estimate ζ_1 first. As in Sen (1960, 1977), we let $U_{n-1}^{(i)}$ be the U-statistic based on $(X_1,\ldots,X_{i-1},\ X_{i+1},\ldots,X_n)$, for $i = 1,\ldots,n\ (\geq m+1)$. Let then

$$s_n^2 = (n-1)^{-1}\sum_{i=1}^{n}[U_{n-1}^{(i)} - U_n]^2\ ,\quad n \geq m+1\ . \tag{2.12}$$

Whenever $\zeta_m < \infty$, s_n^2 is a (strongly) consistent estimator of $m^2\zeta_1$ $(= \lim_{n\to\infty} n\sigma_n^2)$. Motivated by (2.7), (2.5), (2.10) and (2.11), we propose the following sequential procedure:

Let $n_0\ (\geq m+1)$ be an initial sample size and define the *stopping number* $N_c\ (= N_c(a))$ by

$$N_c = \min\{n \geq n_0:\ n \geq (a/c)^{\frac{1}{2}}(s_n + n^{-\gamma})\}, \tag{2.13}$$

where $\gamma\ (> 0)$ is a suitable constant, to be defined later on. Our proposed (sequential point) estimator of $\theta(F)$ is U_{N_c} and the *risk* for the proposed procedure is

$$R_c^*(a) = EL_{N_c} = aE\{U_{N_c} - \theta(F)\}^2 + cEN_c. \tag{2.14}$$

The main theorem of the paper is the following

Theorem 1. If (i) $\theta(F)$ *is stationary of order zero,* (ii) $E|\phi|^{4+\delta} < \infty$ *for some* $\delta > 0$ *and* (iii) *in* (2.13), $\gamma \in (0, (2+\delta)/4)$, *then*

$$\lim_{c \downarrow 0} R_c^*(a)/R_c(n_c^*;\ a,\ F) = 1\ . \tag{2.15}$$

The proof of the theorem is deferred to Section 3. It may be remarked that (2.15) [in the sense of Starr (1966)] asserts that the risk involved in the sequential procedure is asymptotically (as $c \downarrow 0$) equivalent to the risk involved in the corresponding "optimal" fixed-sample size procedure and hence, the sequential procedure is *asymptotically risk-efficient,* for all F satisfying (i) and (ii) of Theorem 1. Also, it may be mentioned that Ghosh and Mukhopadhyay (1979) have considered the case of the population mean (which corresponds to $m = 1$) and obtain (2.15) under $E|\phi|^{\delta} < \infty$ and assuming that in (2.13) $\gamma \in (0, \frac{1}{4})$. In our present setup, $m\ (\geq 1)$ is arbitrary, $\gamma \in (0, (2+\delta)/4)$ (note $(2+\delta)/4) > \frac{1}{2}$) and we need that $E|\phi|^{4+\delta} < \infty$ for some $\delta > 0$. The relaxation of the regularity conditions is achieved here by using some reverse martingale properties of $\{U_n\}$ and the components of $\{s_n^2\}$. Further, results weaker than (2.15) can be obtained even without assuming that $E|\phi|^{4+\delta} < \infty$ for some $\delta > 0$. In fact, we have the following:

Theorem 2. Under (2.8), *and for* $E|\phi|^{2+\delta} < \infty$ *for some* $\delta > 0$,

$$\lim_{c \downarrow 0} E(N_c/n_c^*)^k = 1,\ \forall\ k \in [0,\ 1], \tag{2.16}$$

$$[U_{N_c} - \theta(F)]/\sigma_{n_c^*} \xrightarrow{\mathcal{D}} N(0,\ 1), \quad as \quad c \downarrow 0\ , \tag{2.17}$$

and in (2.17), $\sigma_{n_c^*}$ *may also be replaced by* $m(\zeta_1/N_c)^{\frac{1}{2}}$.

It is of natural interest to study the asymptotic distribution of N_c (if it exists). We shall see later on that under $E\phi^4 < \infty$,

$$V(s_n^2) = \nu^2/n + O(n^{-2}), \quad \forall\ n \geq 2m \tag{2.18}$$

where ν^2 depends on F. Then, we have the following

Theorem 3. If (i) $E|\phi|^{4+\delta} < \infty$ *for some* $\delta > 0$ *and* (ii) *in* (2.13), $\frac{1}{2} < \gamma < (2+\delta)/4$, *then as* $c \downarrow 0$

$$2m^2\zeta_1(N_c - n_c^o)/(\nu^2 n_c^o)^{\frac{1}{2}} \xrightarrow{\mathcal{D}} N(0, 1)\ . \tag{2.19}$$

The proofs of these theorems are presented in Section 3.

For an integer $k(\geq 1)$, moment-inequality for U-statistics have been considered by Funk (1970) and Grams and Serfling (1973), while Sen (1974) studied the L^p-convergence of U-statistics, when $p \geq 1$. In the following lemma, we derive a moment in equality for U-statistisc valid for any power bigger than 1.

Lemma 2.2. Assume that $E|\phi|^r < \infty$ *for some* $r > 1$. *Then, there exists a positive constant* K_r $(< \infty)$, *such that*

$$E|U_n - \theta(F)|^r \leq K_r n^{-s}, \quad \forall\ n \geq m \tag{2.20}$$

where

$$s = \begin{cases} r-1, & \text{if } 1 < r \leq 2; \\ \min(r-1, k), & \text{if } 2(k-1) < r \leq 2k;\ k \geq 2, \end{cases} \tag{2.21}$$

and K_r *does not depend on* n.

The proof of the lemma is considered in the Appendix. In the remaining of this section, we consider the representation of s_n^2 in (2.12) in terms of a set of U-statistics, due to Sproule (1969). For each $d(= 0, 1, \ldots, m)$, let

$$\phi^{(\bullet d)}(X_1,\ldots,X_{2m-d}) = \left\{\frac{(2m-d)!}{d!((m-d)!)^2}\right\}^{-1} \sum^{(d)} \phi(X_{\alpha_1},\ldots,X_{\alpha_m})\phi(X_{\beta_1},\ldots,X_{\beta_m}) \tag{2.22}$$

where the summation $\sum^{(d)}$ extends over all combinations of (distinct) $\alpha_1,\ldots,\alpha_m$ $(\beta_1,\ldots,\beta_m)$ from $(1,\ldots,2m-d)$ with exactly d of the β_j being common with the α_i. Let then (for $n \ge 2m$)

$$U_n^{(d)} = \binom{n}{2m-d}^{-1} \sum_{C_{n,2m-d}} \phi^{(d)}(X_{i_1},\ldots,X_{i_{2m-d}}), \quad 0 \le d \le m. \tag{2.23}$$

Then, by (2.12), (2.22) and (2.23), we have (by some routine steps)

$$s_n^2 = m^2(U_n^{(1)} - U_n^{(0)}) + \sum_{d=0}^{m} e_{nd} U_n^{(d)}, \tag{2.24}$$

where, for some positive constants K_1 and K_2 (independent of n),

$$K_1 n^{-1} \le |e_{nd}| \le K_2 n^{-1}, \ \forall\ 0 \le d \le m, \quad n \ge 2m; \tag{2.25}$$

$$E[U_n^{(1)} - U_n^{(0)}] = \zeta_1. \tag{2.26}$$

This representation plays a vital role in the proof of the main theorems.

3. PROOFS OF THE MAIN THEOREMS

First, we consider the following lemma, which is crucial in the proofs to follow.

Lemma 3.1. If $E|\phi|^{2r} < \infty$, *for some* $r \ge 1$ *and* (2.8) *holds, then, for every* $\varepsilon \in (0, 1)$,

$$P\{N_c \le n_c^*(1-\varepsilon)\} = O(c^{s/2(1+\gamma)}), \quad \text{as} \quad c \downarrow 0, \tag{3.1}$$

where s *is defined by* (2.21).

PROOF: Note that by (2.24) - (2.26),

$$s_n^2 - m^2\zeta_1 = m^2(U_n^{(1)} - U_n^{(0)} - \zeta_1) + \sum_{d=1}^{m} e_{nd} U_n^{(d)}, \tag{3.2}$$

and, by (2.13), $N_c \geq b^{1/(1+\gamma)}$, with probability 1, where $b = (a/c)^{\frac{1}{2}}$. Let then $n_{1c} = [b^{1/(1+\gamma)}]$ and $n_{2c} = n_c^*(1-\varepsilon)$; choose c so small that $n_{1c} \leq n_{2c}$ (otherwise, there is no need to prove (3.1)). Then, by (2.13),

$$\begin{aligned} P\{N_c \leq n_c^*(1-\varepsilon)\} &= P\{N_c \leq n_{2c}\} \leq P\{s_n < b^{-1}n, \text{ for some } n_{1c} \leq n \leq n_{2c}\} \\ &\leq P\{s_n^2 \leq b^{-2}n_{2c}^2 \text{ for some } n_{1c} \leq n \leq n_{2c}\} \\ &\leq P\{s_n^2 - m^2\zeta_1 \leq m^2\zeta_1\{(1-\varepsilon)^2 - 1\}, \text{ for some } n_{1c} \leq n \leq n_{2c}\} \\ &\leq P\{|s_n^2 - m^2\zeta_1|/m^2\zeta_1 \geq \varepsilon(2-\varepsilon), \text{ for some } n_{1c} \leq n \leq n_{2c}\} . \end{aligned} \tag{3.3}$$

By (2.25), (3.2) and (3.3), we have

$$\begin{aligned} P\{N_c \leq n_c^*(1-\varepsilon)\} &\leq P\{\max_{n_{1c}\leq n\leq n_{2c}} \left|U_n^{(1)} - U_n^{(0)} - \zeta_1\right| \geq \varepsilon\zeta_1\} \\ &\quad + P\{\max_{n_{1c}\leq n\leq n_{2c}} \textstyle\sum_{d=0}^m |U_n^{(d)}| \geq Kn_{1c}\} \\ &\leq P\{\max_{n_{1c}\leq n\leq n_{2c}} \left|U_n^{(1)} - U_n^{(0)} - \zeta_1\right| \geq \varepsilon\zeta_1\} + \\ &\quad \textstyle\sum_{d=0}^m P\{\max_{n_{1c}\leq n\leq n_{2c}} \left|U_n^{(d)}\right| \geq Kn_{1c}\} , \end{aligned} \tag{3.4}$$

where K (>0) does not depend on c (but depends on ε).

Let F_n be the σ-field generated by the ordered collection of $X_1,\ldots,X_n$ and by X_{n+j}, $j \geq 1$ (so that F_n is nondecreasing in n). Then, $\{U_n, F_n; n \geq m\}$ and $\{U_n^{(d)}, F_n; n \geq 2m-d\}$, for every $d = 0, 1,\ldots,m$ are reverse martingales, and hence $\{U_n^{(1)} - U_n^{(0)} - \zeta_1, F_n; n \geq 2m\}$ is also a reverse martingale, so that for $n_{1c} \geq 2m$, by the Kolmogorov-Hájek-Rényi-Chow inequality for reverse martingales and our Lemma 2.2,

$$P\{\max_{n_{1c}\le n\le n_{2c}} |U_n^{(1)} - U_n^{(0)} - \zeta_1| > \varepsilon\zeta_1\}$$

$$\le E|U_{n_{1c}}^{(1)} - U_{n_{1c}}^{(0)} - \zeta_1|^r/(\varepsilon\zeta_1)^r \le K_r n_{1c}^{-s}, \tag{3.5}$$

where s is defined by (2.21), and,

$$\sum_{d=0}^{m} P\{\max_{n_{1c}\le n\le n_{2c}} |U_n^{(d)}| \ge Kn_{1c}\}$$

$$\le (m+1)K_r n_{1c}^{-r}\left\{\max_{0\le d\le m} E|U_{n_{1c}}^{(d)}|^r\right\}, \tag{3.6}$$

where $E|\phi|^{2r} < \infty \implies E|U_n^{(d)}|^r < \infty$, $\forall\ 0 \le d \le m$, $n \ge 2m$; (3.1) then follows from (3.4), (3.5) and (3.6) after noting that by (2.21), $s < r$ and $n_{1c} = [b^{1/(1+\gamma)}] = 0(c^{1/2(1+\gamma)})$ as $c \downarrow 0$. Q.E.D.

Now, by virtue of (2.24) - (2.26) and Lemma 2.2, we have for $E|\phi|^{2r} < \infty$ for some $r \ge 1$,

$$E|s_n^2 - m^2\zeta_1|^r \le K_r n^{-s}, \tag{3.7}$$

where s is defined in (2.21). Using (3.7) one can follow the lines of proof in part (d) of the lemma of Ghosh and Mukhopadhyay (1979) to conclude that

$$E(N_c/n_c^*)^k \to 1 \quad \text{as} \quad c \downarrow 0,\ \forall\ k < s, \tag{3.8}$$

where s is defined in (2.21). In particular, if in (3.7), we let $r = 1 + \delta/2$, $\delta > 0$, then $s > 1$, so that (3.8) holds for every $0 \le k \le 1$. This proves (2.16). Moreover, by (2.13),

$$bs_{N_c} \le N_c \le n_0 + b(s_{N_c-1} + (N_c - 1)^{-\gamma}) \tag{3.9}$$

where $b^2 = a/c$. Since by the Convergence Theorem for reverse martingales, $U_n^{(d)}$, $0 \le d \le m$ all converge a.s. to their expectations as $n \to \infty$, by (3.2), we claim that $E\phi^2 < \infty \implies s_n^2 \to m^2\zeta_1$ a.s., as

$n \to \infty$. Hence, dividing all sides of (3.9) by n_c^* and letting $c \downarrow 0$ (i.e., $n_c^* \to \infty$), we obtain that

$$N_c/n_c^* \to 1 \text{ a.s., as } c \downarrow 0; \tag{3.10}$$

(2.17) follows then by using (3.10) and the results in Section 5 of Miller and Sen (1972). In fact, for (2.17), $E\phi^2 < \infty$ suffices. This completes the proof of Theorem 2.

We proceed now to prove Theorem 1. First, note that by (2.16) $(EN_c)/n_c^* \to 1$ as $c \downarrow 0$. Hence, by virtue of (2.6) and (2.14), to prove (2.15), it suffices to show that

$$\lim_{c\downarrow 0} aE\{U_{N_c} - \theta(F)\}^2/(cn_c^*) = 1. \tag{3.11}$$

Let us write $E\{U_{N_c} - \theta(F)\}^2 = E\{U_{n_c^*} - \theta(F)\}^2 + E\{U_{N_c} - U_{n_c^*}\}^2 +$ $2E(U_{n_c^*} - \theta(F))(U_{N_c} - U_{n_c^*})$ and note that by (2.9) and (2.11), $E\{U_{n_c^*} - \theta(F)\}^2 = (m^2\zeta_1/n_c^*) + O(n_c^{*-2})$, so that

$$\lim_{c\downarrow 0} aE\{U_{n_c^*} - \theta(F)\}^2/(cn_c^*) = 1 . \tag{3.12}$$

Hence, to prove (3.11), it suffices to show that

$$\lim_{c\downarrow 0} n_c^* E\{U_{N_c} - U_{n_c^*}\}^2 = 0 . \tag{3.13}$$

Using the definition of the ϕ_d, prior to (2.2), we may write

$$U_n = mU_n^{(1)} + U_n^*; \ \forall\ n \geq m; \tag{3.14}$$

$$U_n^{(1)} = n^{-1}\textstyle\sum_{i=1}^n \phi_1(X_i), \quad EU_n^* = 0; \tag{3.15}$$

$$EU_n^{*2} = C_1 n^{-2} \text{ and } EU_n^{*4} \leq C_2 n^{-4}, \ \forall\ n \geq m, \tag{3.16}$$

where C_1 and C_2 are positive and finite constants, independent of n. Also, we have by (3.14),

$$n_c^*[U_{N_c} - U_{n_c^*}]^2 \leq 2m^2 n_c^*[U_{N_c}^{(1)} - U_{n_c^*}^{(1)}]^2 + 4n_c^* U_{n_c^*}^{*2} + 4n_c^* U_{N_c}^{*2} \tag{3.17}$$

where by (3.16),

$$En_c^* U_{n_c^*}^{*2} \le C_1 n_c^{*-1} \to 0 \quad \text{as} \quad c \downarrow 0 \ . \tag{3.18}$$

Further,

$$\begin{aligned} n_c^* U_{N_c}^{*2} &= n_c^* \sum_{k \ge n_{1c}} I(N_c = k) U_k^{*2} \\ &\le n_c^* \sum_{n_{1c} \le k \le n_{2c}} I(N_c = k) U_k^{*2} + \left(\sup_{n > n_{2c}} U_n^{*2}\right) I(N_c > n_{2c}) \end{aligned} \tag{3.19}$$

where n_{1c} and n_{2c} are defined after (3.2). Now

$$\begin{aligned} &E\{n_c^* \sum_{n_{1c} \le k \le n_{2c}} U_k^{*2} I(N_c = k)\} \\ &\le n_c^* \sum_{n_{1c} \le k \le n_{2c}} \sqrt{P(N_c \le k)}\ E(U_k^{*4})^{\frac{1}{2}} \\ &\le n_c^* \sqrt{P(N_c \le n_{2c})} \sum_{n_{1c} \le k \le n_{2c}} E(U^{*4})^{\frac{1}{2}} \\ &\longrightarrow 0 \quad \text{as} \quad c \to 0, \quad \text{by (3.1), (3.19) and definition } n_{1c}. \end{aligned} \tag{3.20}$$

Also, by the Doob maximal inequality for (reverse) submartingales,

$$\begin{aligned} &n_c^* E\{ \sup_{n > n_{2c}} U_n^{*2} I(N_c > n_{2c})\} \\ &\le n_c^* E\{ \sup_{n > n_{2c}} U_n^{*2}\} \le n_c^* 4E(U_{n_{2c+1}}^{*2}) \to 0 \quad \text{as} \quad c \downarrow 0 \end{aligned} \tag{3.21}$$

where the last step follows from (3.16) and the definition of n_{2c}.

Hence, it suffices to show that

$$\lim_{c \downarrow 0} n_c^* E\{U_{N_c}^{(1)} - U_{n_c^*}^{(1)}\}^2 = 0. \tag{3.22}$$

Now $U_n^{(1)}$ is a sample mean for all $n \ge m$. If follows from Anscombe's (1952) result and (3.13) that

$$(n_c^*)^{\frac{1}{2}} (U_{N_c}^{(1)} - U_{n_c^*}^{(1)}) \xrightarrow{P} 0 \quad \text{as} \quad c \downarrow 0 \ . \tag{3.23}$$

We follow then the line of proof of Ghosh and Mukhopadhyay (1979) [in view of our Lemma 2.2, which is stronger than the moment inequality of Grams and Serfling (1973) (restricted to integer power), their eighth moment condition is not needed here]. and obtain that

$$\{n_c^*(U_{N_c}^{(1)} - U_{n_c^*}^{(1)})^2\} \text{ is uniformly integrable in } c \le c_0, \tag{3.24}$$

for some $c_0 \downarrow 0$. From (3.23) and (3.24), we conclude that (3.22) holds and the proof of Theorem 1 is now complete.

To prove Theorem 3, first, note that from Sproule's (1974) Theorem, one gets

$$N_c^{\frac{1}{2}}(s_{N_c}^2 - m^2\zeta_1)/\nu \xrightarrow{\mathcal{D}} N(0, 1), \text{ as } c \downarrow 0, \tag{3.25}$$

where $s_{N_c}^2$ can also be replaced by $s_{N_c-1}^2$ in (3.25). Hence using the Mann-Wald theorem, we obtain that

$$2N_c^{\frac{1}{2}}(s_{N_c} - m\zeta_1^{\frac{1}{2}})/(\nu/m\zeta_1^{\frac{1}{2}}) \xrightarrow{\mathcal{D}} N(0, 1), \text{ as } c \downarrow 0, \tag{3.26}$$

where in (3.26), also, s_{N_c} may be replaced by s_{N_c-1}. From (3.26) and the definition of the stopping time in (2.13), one finds that the sufficient condition in Theorem 3 of Ghosh and Mukhopadhyay (1979) holds with $a = m\zeta_1^{\frac{1}{2}}$ and $b = \nu/(2m\zeta_1^{\frac{1}{2}})$. A direct appeal to this theorem now yields (2.19). Q.E.D.

4. SOME ADDITIONAL REMARKS

Ghosh, Sinha and Mukhopadhyay (1976) have considered sequential point estimation of the multinormal mean vector with unknown covariance matrix. If $\underset{\sim}{X}_i$, $i \ge 1$ are i.i.d. random $p(\ge 1)$ - vectors with $E\underset{\sim}{X}_1 = \mu$ and $V(\underset{\sim}{X}_1) = \underset{\sim}{\Sigma}$, positive definite (p.d.), then assuming the loss function, based on $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$, to be of the form

$$L_n = (\bar{\underset{\sim}{X}}_n - \underset{\sim}{\mu})'\underset{\sim}{A}(\bar{\underset{\sim}{X}}_n - \underset{\sim}{\mu}) + cn \tag{4.1}$$

where $\underset{\sim}{A}$ is a known p.d. *weight matrix,* c is the cost per unit sample and $\bar{\underset{\sim}{X}}_n = n^{-1}\sum_{i=1}^{n} X_i$, $n \geq 1$, the risk function is given by

$$R_n = n^{-1}\mathrm{Tr}(\underset{\sim}{A}\underset{\sim}{\Sigma}) + cn, \tag{4.2}$$

so that for known $\underset{\sim}{\Sigma}$, the risk is minimized at

$$n_0 = \{c^{-1}\mathrm{Tr}(\underset{\sim}{A}\underset{\sim}{\Sigma})\}^{\frac{1}{2}}. \tag{4.3}$$

For unknown $\underset{\sim}{\Sigma}$, by analogy to (2.13), we define the *stopping time*

N = smallest positive integer $n(\geq 2)$ for which

$$n \geq c^{-\frac{1}{2}}\{\mathrm{Tr}(\underset{\sim}{A}\underset{\sim}{S}_n) + n^{-\gamma}\}, \tag{4.4}$$

where $\underset{\sim}{S}_n = (n-1)^{-1}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{X}_n)(\underset{\sim}{X}_i - X_n)'$, $n \geq 2$. Now, $\mathrm{Tr}(\underset{\sim}{A}\underset{\sim}{S}_n)$ is a U-statistic and the results of the previous sections apply to yield an "asymptotically risk efficient" sequential procedure. In this context, the multinormality of the $\underset{\sim}{X}_i$ is no longer needed.

In the context of jackknifing, Sen (1977) has considered a class of smooth functions of U-statistics. Under his assumed boundedness conditions of the first and second order (partial) derivatives and the conditions of Cramér (1946, p. 353) the results of Sections 2 and 3 can also be extended to such functions of U-statistics.

Finally, Robbins (1959), while considering the case of the normal mean, proposed a slightly different loss function, namely,

$$L_n = a|U_n - \theta(F)| + cn. \tag{4.5}$$

Our asymptotically risk efficient procedure also holds for such a loss function, provided in (2.6) through (2.10), we make the necessary modifications. Note that as $n \to \infty$,

$$E\{n^{1/2}|U_n - \theta(F)|\} \longrightarrow (2/\pi)^{1/2} m\zeta_1^{1/2}\,, \tag{4.6}$$

so that the optimal value n_c^*, in this case, is given by

$$n_c^* \sim (a^2m^2\zeta_1/2\pi c^2)^{1/3} \quad \text{as} \quad c \downarrow 0\,, \tag{4.7}$$

and analogous to (2.13), we define the stopping number by

$$N_c = \min\{n \geq n_0 \colon \; n \geq (a/c)^{2/3}(2\pi)^{-1/3}(s_n^2 + n^{-\gamma})^{1/3}\} \tag{4.8}$$

where γ and n_0 are defined as in Section 2. With these modifications, the theorems in Section 2 extend directly to the case of the loss function defined by (4.5). A similar case holds for $L_n = a|U_n - \theta(F)|^b + cn$ for some $b > 0$, where for $b \geq 2$, we need to replace condition (ii) of Theorem 1 by $E|\phi|^{(2+\delta)b} < \infty$ for some $\delta > 0$.

5. APPENDIX

<u>Proof of Lemma 2.1</u>

Note that by (2.6),

$$\Delta R_c(n;\, a,\, F) = R_c(n+1;\, a,\, F) - R_c(n;\, F) = c - a(\sigma_n^2 - \sigma_{n+1}^2), \tag{5.1}$$

$$\begin{aligned} \Delta^2 R_c(n;\, a,\, F) &= R_c(n+2;\, a,\, F) - 2R_c(n+1;\, a,\, F) + R_c(n;\, a,\, F) \\ &= a(\sigma_{n+2}^2 - 2\sigma_{n+1}^2 + \sigma_n^2), \end{aligned} \tag{5.2}$$

where $\sigma_{n+1}^2 \leq \sigma_n^2$, $\forall\, n \geq m$. Hence, it suffices to show that for all $n \geq m$, $\Delta^2 R_c(n;\, a,\, F) \geq 0$. For this, define as in Hoeffding (1948),

$$\delta_d = \sum_{i=0}^{d} (-1)^i \binom{d}{i} \zeta_{d-i}, \quad 0 \leq d \leq m, \tag{5.3}$$

so that $\delta_0 = \zeta_0 = 0$. Then

$$\zeta_d = \sum_{i=0}^{d} \binom{d}{i} \delta_{d-i} = \sum_{i=0}^{d} \binom{d}{d-i} \delta_i\,. \tag{5.4}$$

From (2.4), (5.3) and (5.4), we have by some standard steps

$$\sigma_n^2 = \binom{n}{m}^{-1} \sum_{i=1}^{m} \binom{m}{i} \binom{n-i}{m-i} \delta_i \tag{5.5}$$

so that by (5.2) and (5.5), we have

$$\Delta^2 R_c(n;\ a,\ F) = \sum_{i=1}^{m} \binom{m}{i} \binom{n}{m}^{-1} \binom{n-i}{m-i} \delta_i \left[\frac{(n-i+1)(n-i+2)}{(n+1)(n+2)} - 2\,\frac{n-i+1}{n+1} + 1 \right]$$

$$\geq 0, \tag{5.6}$$

as $(n-i+1)(n-i+2) - 2(n+2)(n-i+1) + (n+1)(n+2) = i^2 + 1 > 0,\ \forall\ i \geq 1$

and by Lemma 5.1 of Hoeffding (1948), $\delta_k \geq 0,\ \forall\ k = 0, 1, \ldots, m$. Q.E.D.

Proof of Lemma 2.2

Let $n^o = [n/m]$, $n \geq m$ and let

$$T_n = (n^o)^{-1} \sum_{r=1}^{n^o} \phi(X_{(r-1)m+1}, \ldots, X_{rm}). \tag{5.7}$$

Then, $ET_n = \theta(F),\ \forall\ n \geq m$ and defining $\mathcal{F}_n$ as in after (3.4),

$$E(T_n | \mathcal{F}_n) = U_n,\ \forall\ n \geq m, \tag{5.8}$$

so that by the Jensen inequality for conditional expectations,

$$E|U_n - \theta(F)|^r \leq E|T_n - \theta(F)|^r, \text{ for any } r \geq 1 . \tag{5.9}$$

On the other hand, T_n is an average of n^o i.i.d.r.v.'s, and hence, Theorem 3 of Sen (1970) applies to the right hand side of (5.9) and this yields (2.20) and (2.21). Q.E.D.

As noted already, the above generalizes and strengthens the results of Funk (1970) and Grams and Serfling (1973), where they needed r to be a positive integer. Also, our method of proof is elementary and quite different from the earlier ones.

REFERENCES

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

[2] CRAMER, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press, New Jersey.

[3] FUNK, G.M. (1970). The probabilities of moderate deviations of U-statistics and excessive deviations of Kolmogorov-Smirnov and Kuiper statistics. *Ph.D. dissertation, Michigan State University.*

[4] GHOSH, M., SINHA, B.K. and MUKHOPADHYAY, N. (1976). Multivariate sequential point estimation. *J. Mult. Anal.*, 6, 281-294.

[5] GHOSH, M., and MUKHOPADHYAY, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Comm. Statist.*, 8, to appear in the July issue.

[6] GRAMS, W.F. and SERFLING, R.J. (1973). Convergence rates for U-statistics and related statistics. *Ann. Statist.*, 1, 153-160.

[7] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19, 293-325.

[8] MILLER, R.G., JR. and SEN, P.K. (1972). Weak convergence of U-statistics and von Mises' differentiable statistical functions. *Ann. Math. Statist.*, 43, 31-41.

[9] ROBBINS, H. (1959). Sequential estimation of the mean of a normal population. *Probability and Statistics* (H. Cramér Volume), Almquist and Wiksell, Uppsala, 235-245.

[10] SEN, P.K. (1960). On some convergence properties of U-statistics. *Cal. Statist. Assoc. Bull.*, 10, 1-18.

[11] SEN, P.K. (1970). On some convergence properties of one-sample rank order statistics. *Ann. Math. Statist.*, 41, 2140-2143.

[12] SEN, P.K. (1974). On L^p-convergence of U-statistics. *Ann. Inst. Statist. Math.*, 26, 55-60.

[13] SEN, P.K. (1977). Some invariance principles relating to Jackknifing and their role in sequential analysis. *Ann. Statist.*, 5, 316-329.

[14] SPROULE, R.N. (1969). A sequential fixed width confidence interval for the mean of a U-statistic. *Ph.D. dissertation, U.N.C., Chapel Hill.*

[15] SPROULE, R.N. (1974). Asymptotic properties of U-statistics. *Trans. Amer. Math. Soc.*, 199, 55-64.

[16] STARR, N. (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean. *Ann. Math. Statist.*, 37, 1173-1185.

[17] STARR, N. and WOODROOFE, M. (1969). Remarks on sequential point estimation. *Proc. Nat. Acad. Sci., U.S.A.*, 63, 285-288.

[18] STARR, N. and WOODROOFE, M. (1972). Further remarks on sequential point estimation: the exponential case. *Ann. Math. Statist.*, 43, 1147-1154.