The Philippine Statistician, 1995-1996
Vol. 44-45, Nos. 1-8, pp. 1-25

Sequential Point Estimation of Quantiles

Mercidita T. Navarro and **Adam T. Martinsek**[1]

ABSTRACT

Let $X_1, X_2 \ldots$ be independent observations from a distribution F. Suppose that one wishes to estimate the p^{th} quantile ξ_p of F, subject to a loss function that is a linear combination of the squared error in estimation and the cost of sampling. One may stop after any number of observations n and estimate ξ_p by the sample p^{th} quantile. If $F'(\xi_p) = f(\xi_p)$ is known, the best fixed sample size (*i.e.*, in the sense of minimum risk) can be used. However, if $f(\xi_p)$ is unknown, then the best fixed sample size is also unknown. For this case a stopping rule is proposed. It is shown that under certain smoothness conditions on F and a growth condition on the delay, the sequential procedure derived is asymptotically risk efficient. The results are then extended to a more general loss structure. Estimation of a linear combination of two quantiles, *e.g.*, the interquartile range, is also studied. Results of simulation studies are provided.

KEY WORDS: Asymptotic risk efficiency; Bahadur representation of quantiles; Sequential estimation; Stopping rules; Uniform integrability.

1. INTRODUCTION

Let $X_1, X_2 \ldots$ be independent observations on a distribution function F. Let ξ_p be the p^{th} quantile of F, *i.e.*,

$$\xi_p = F^{-1}(p) = \inf\{x: F(x) \geq p\},\ 0 < p < 1.$$

Given a sample of size n, a natural estimate of ξ_p is the sample p^{th} quantile $\hat{\xi}_{pn} = F_n^{-1}(p)$, where $F_n(\cdot)$ is the empirical distribution function of $X_1, X_2 \ldots, X_n$. In terms of order statistics, $\hat{\xi}_{pn}$ is defined as:

$$\hat{\xi}_{pn} = \begin{cases} X_{n:np} & \text{if } np \text{ is an integer,} \\ X_{[n:np]+1} & \text{otherwise.} \end{cases}$$

where $X_{n:i}$ is the i^{th} order statistic among $X_1, \ldots X_n$ and $[x]$ denotes the integer part of x.

Suppose one wishes to estimate ξ_p by $\hat{\xi}_{pn}$ subject to the loss function

$$L_n = A(\hat{\xi}_{pn} - \xi_p)^2 + n,\ A > 0; \tag{1}$$

[1] M. T. Navarro, Roche Global Development, 3401 Hillview Avenue, Palo Alto, CA 94304-1397, U.S.A. and A. T. Martinsek, Department of Statistics, University of Illinois, 725 South Wright Street, Champaign, IL 61820, U.S.A.

that is, the loss is a linear combination of the squared error in estimation and the cost of sampling, normalized in such a way that the cost per observation is one. For a fixed sample size n (one determined before taking any observations), the risk is

$$\mathcal{R}_n = EL_n = AE(\hat{\xi}_{pn} - \xi_p)^2 + n.$$

Assume in what follows that F is twice differentiable at ξ_p and $F'(\xi_p) = f(\xi_p) > 0$. From Bahadur (1996), $\hat{\xi}_{pn}$ can be expressed as an average of random variables via the empirical distribution function, that is,

$$\hat{\xi}_{pn} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + R_{n,} \tag{2}$$

where $R_n = O(n^{-3/4}(\log n)^{3/4})$ *a.s.* as $n \to \infty$. As will be seen later, the remainder term R_n plays a key role in this study. From Theorem 2 of Duttweiler (1973), $ER_n^2 = O(n^{-3/2})$ as $n \to \infty$. Therefore, the risk is asymptotically up to $O\left(n^{-5/4}\right)$,

$$\mathcal{R}_n \approx A\frac{p(1-p)}{f^2(\xi_p)}\frac{1}{n} + n. \tag{3}$$

Treating n as a continuous variable and differentiating the above function with respect to n, one finds that $\mathcal{R}_n$ is minimized by the best fixed sample size

$$\left[\sqrt{A}\frac{\sqrt{p(1-p)}}{f(\xi_p)}\right] \le n_0 \le \left[\sqrt{A}\frac{\sqrt{p(1-p)}}{f(\xi_p)}\right] + 1, \tag{4}$$

and the minimum risk is

$$\mathcal{R}_{n_0} \approx 2\sqrt{A}\frac{\sqrt{p(1-p)}}{f(\xi_p)} \approx 2n_0.$$

(Notice that $n_0 \approx \sqrt{A}\frac{\sqrt{p(1-p)}}{f(\xi_p)}$, where $\frac{\sqrt{p(1-p)}}{f(\xi_p)}$ is the asymptotic standard deviation of $\sqrt{n}\left(\hat{\xi}_{pn} - \xi_p\right)$.) However, when $f(\xi_p)$ is unknown, no fixed sample size n minimizes $\mathcal{R}_n$ uniformly for all values of $f(\xi_p)$, *i.e.*, the best fixed sample size cannot be calculated. This motivates the need for a sequential estimation procedure.

A number of authors have studied the problem of sequential bounded-length confidence interval estimation for quantiles. Farrell (1966) considered two sequential procedures for constructing bounded-length confidence intervals for any quantile. Swanepoel and Lombard (1978) investigated a variant of Farrell's first procedure, limiting the problem to the class of continuous distribution functions. Geertsema (1970) constructed a sequential confidence interval for the median of a symmetric distribution based on rank tests and established the convergence rate of the coverage probability (as the prescribed length of the interval tends to zero) in Geertsema (1985). Sen and Ghosh (1971) studied the same problem using one-sample rank-order statistics. At present, however, no sequential procedure is yet available for the point estimation of quantiles. In this paper, a stopping rule for estimating any single quantile, under the loss structure described above, is proposed.

Define sequences $\{k_{1n}\}$ and $\{k_{2n}\}$ such that

$$\frac{k_{1n}}{n} = p - \sqrt{\frac{p(1-p)}{n}} \quad \text{and} \quad \frac{k_{2n}}{n} = p + \sqrt{\frac{p(1-p)}{n}}. \tag{5}$$

Without loss of generality, assume in what follows that k_{1n} and k_{2n} are integers. As in (2), Bahadur (1966) provides the following representations for the "central" order statistics $X_{n:k_{1n}}$ and $X_{n:k_{2n}}$:

$$X_{n:k_{1n}} = \xi_p + \frac{p - \sqrt{\frac{p(1-p)}{n}} - F_n(\xi_p)}{f(\xi_p)} + \widetilde{R}_n^1; \tag{6}$$

$$X_{n:k_{2n}} = \xi_p + \frac{p + \sqrt{\frac{p(1-p)}{n}} - F_n(\xi_p)}{f(\xi_p)} + \widetilde{R}_n^2;$$

where $\widetilde{R}_n^i = O\left(n^{-3/4} (\log n)^{(\Delta+1)/2}\right)$ *a.s.* as $n \to \infty$, for any $\Delta \geq ½$, for $i = 1, 2$. An estimator for $f(\xi_p)$ readily obtained from (6) is

$$\hat{f}(\xi_p) = 2\sqrt{\frac{p(1-p)}{n}} \frac{1}{X_{n:k_{2n}} - X_{n:k_{1n}}}, \tag{7}$$

which converges to $f(\xi_p)$ as $n \to \infty$, with probability one. The stopping rule

$$T_A = \inf\left\{n \geq n_A : \frac{\sqrt{p(1-p)}}{\hat{f}(\xi_p)} \leq \frac{n}{\sqrt{A}}\right\} = \inf\left\{n \geq n_A : \frac{n}{4}(X_{n:k_{2n}} - X_{n:k_{1n}})^2 \leq \frac{n^2}{A}\right\}, \tag{8}$$

where n_A is a positive integer which may depend on A, may then be used and ξ_p estimated by $\hat{\xi}_{pT_A}$. The performance of this sequential procedure is summarized in the following theorem.

THEOREM 1. *Suppose $X_1, X_2 \ldots$ are independent observations with a common distribution function F. Assume that F is twice differentiable at ξ_p with $F'(\xi_p) = f(\xi_p) > 0$. Assume* *further that f' exists and is bounded in a neighborhood of ξ_p, and that $E|X_1|^\lambda < \infty$ for some $\lambda > 0$. If T_A is defined by* (8) *and $n_A = cA^\delta$, for some $c > 0$, $\delta \in (0,½)$, then as $A \to \infty$,*

$$\frac{T_A}{n_0} \to 1 \; a.s., \tag{9}$$

$$E\frac{T_A}{n_0} \to 1 \tag{10}$$

and

$$\frac{\mathcal{R}_{T_A}}{\mathcal{R}_{n_0}} = E\left[\frac{A(\hat{\xi}_{pT_A} - \xi_p)^2 + T_A}{2\sqrt{Ap(1-p)}} f(\xi_p)\right] \to 1. \tag{11}$$

In other words, the proposed sequential procedure performs asymptotically as well as the procedure that uses the optimal fixed sample size, under the loss structure (1).

REMARK 1. If one is willing to impose more conditions on the density f, then there are better estimates of $f(\xi_p)$ than the one used here. For example, if f is twice differentiable on the real line and f, f', and f'' are bounded, then the uniform almost sure convergence rate of an appropriate kernel estimate of f is $n^{-1/3}((\log(\log n))^{1/2}$ (see Prakasa Rao (1983), Theorem 2.1.12). The almost sure convergence rate of the estimate used in the current work is $n^{-1/4}(\log n)^{\frac{\Delta+1}{2}}$. However, the estimate is simple, requires fewer assumptions on f, and as Theorem 1 shows, yields asymptomatic efficiency.

Theorem 1 provides an asymptotically risk-efficient procedure for estimating quantiles relative to one particular loss function a linear combination of the squared error in estimation and the cost of sampling. A proof is provided in the next section.

The procedure constructed above can be modified to work for other loss functions, for instance, one that involves the absolute instead of the squared error in estimation. An extension of Theorem 1 is discussed in Section 3. Simulation results are provided in Section 4.

2. PROOF OF THEOREM 1

The proof of Theorem 1 depends on a series of six lemmas, the first of which provides conditions under which positive powers of $A^{-1/2}T_A$ are uniformly integrable.

LEMMA 1. *Suppose F has a derivative on $\left(F^{-1}(p-\varepsilon), F^{-1}(p+\varepsilon)\right)$ and*

$$F'(x) = f(x) \geq f_0 > 0 \text{ for every } x \text{ in } \left(F^{-1}(p-\varepsilon), F^{-1}(p+\varepsilon)\right), \tag{12}$$

for some f_0, $\varepsilon > 0$. Assume also that $E|X_1|^\lambda < \infty$ for some $\lambda > 0$. If T_A is defined by (8), *then for every $s > 0$,*

$$\left\{\left(A^{-1/2}T_A\right)^s, A \geq 1\right\} \textit{ is uniformly integrable.}$$

PROOF. It suffices to show that as $K \to \infty$,

$$P\left(T_A > KA^{1/2}\right) = O\left(K^{-\delta}\right)$$

uniformly in A for all $\delta > 0$. Fix $A > 0$ and let $m = \left[KA^{1/2}\right]$. By the definition of T_A, for any $\delta > 0$,

$$P\left(T_A > KA^{1/2}\right) \leq P\left(X_{m:k_{2m}} - X_{m:k_{1m}} > 2A^{-1/2}m^{1/2}\right) \leq 2^{-\delta}A^{\delta/2}m^{-\delta/2}E\left|X_{m:k_{2m}} - X_{m:k_{1m}}\right|^{\delta}. \tag{13}$$

It follows from the proof of Theorem 2 of Sen (1959) that for every $\delta > 0$, as $m \to \infty$,

$$\sup_{[(p-\varepsilon)m] \leq i \leq [(p+\varepsilon)(m+1)]+1} E\left\{\left|m^{1/2}\left(X_{m:i} - F^{-1}\left(\frac{i}{m+1}\right)\right)\right|^{\delta}\right\} = O(1), \tag{14}$$

for the $\varepsilon > 0$ that satisfies (12). Therefore,

$$E\left\{\left|m^{1/2}\left(X_{m:k_{jm}} - F^{-1}\left(\frac{k_{jm}}{m+1}\right)\right)\right|^{\delta}\right\} = O(1) \text{ as } m \to \infty, \text{ for } j = 1, 2.$$

Now note that by the Mean Value Theorem, there exisits $\zeta \in \left(\frac{k_{1m}}{m+1}, \frac{k_{2m}}{m+1}\right)$ such that

$$\left|F^{-1}\left(\frac{k_{2m}}{m+1}\right) - F^{-1}\left(\frac{k_{1m}}{m+1}\right)\right| = 2\sqrt{\frac{p(1-p)}{m}}\left|\left(F^{-1}\right)'(\zeta)\right|.$$

By (12), $\left(F^{-1}\right)'(\zeta) \leq \frac{1}{f_0} < \infty$ for m sufficiently large, and so as $m \to \infty$,

$$\left|F^{-1}\left(\frac{k_{2m}}{m+1}\right) - F^{-1}\left(\frac{k_{1m}}{m+1}\right)\right|^{\delta} = O\left(m^{-\delta/2}\right).$$

Thus, as $m \to \infty$,

$$E\left\{\left|m^{1/2}\left(X_{m:k_{2m}} - X_{m:k_{1m}}\right)\right|^{\delta}\right\} = O(1), \tag{15}$$

and hence from (13), as $K \to \infty$,

$$P\left(T_A > KA^{1/2}\right) \leq A^{\delta/2}O\left(m^{-\delta}\right) = O\left(A^{\delta/2}A^{-\delta/2}K^{-\delta}\right) = O\left(K^{-\delta}\right)$$

uniformly in A, proving Lemma 1.

The following lemma is patterned after Lemma 1 of Bahadur (1966), a basic element in the derivation of his representations for sample quantiles and central order statistics.

LEMMA 2. *Suppose F is twice differentiable at* ξ_p *with* $F'(\xi_p) = f(\xi_p) > 0$. *Let* $\{a_n\}$ *be a sequence of positive constants such that* $a_n \sim c_0 n^{-\rho}$ *as* $n \to \infty$, *for some* $c_0 > 0$, $\rho \in (\frac{1}{4}, \frac{1}{2})$. *Put*

$$H_{pn} = \sup_{|x| \le a_n} \left| \left[F_n(\xi_p + x) - F_n(\xi_p)\right] - \left[F(\xi_p + x) - F(\xi_p)\right] \right|.$$

Then for every $t > 0$, *as* $n \to \infty$,

$$E\left(\left| n^{(\rho+1/4)} H_{pn} \right|^t \right) = O(1).$$

PROOF. Let $\{b_n\}$ be a sequence of positive integers such that $b_n \sim c_0 n^{1/4}$, $n \to \infty$. For integers $r = -b_n, \ldots, b_n$, put

$$\eta_{r,n} = \xi_p + \frac{a_n}{b_n} r; \quad \alpha_{r,n} = F(\eta_{r+1,n}) - F(\eta_{r,n}),$$

and

$$G_{r,n} = \left| F_n(\eta_{r,n}) - F_n(\xi_p) - \left[F(\eta_{r,n}) - F(\xi_p)\right] \right|.$$

Let

$$K_n = \max \left\{ G_{r,n} : -b_n \le r \le b_n \right\}; \quad \beta_n = \max \left\{ \alpha_{r,n} : -b_n \le r \le b_n - 1 \right\}.$$

By the monotonicity of F_n and F, it follows that

$$H_{pn} \le K_n + \beta_n. \tag{16}$$

Since $\eta_{r+1,n} - \eta_{r,n} = \frac{a_n}{b_n} \sim n^{-(\rho+1/4)}$ as $n \to \infty$, for $-b_n \le r \le b_n - 1$, then by the Mean Value Theorem,

$$\begin{aligned} \alpha_{r,n} &= F(\eta_{r+1,n}) - F(\eta_{r,n}) \\ &\le \sup_{|x| \le a_n} F'(\xi_p + x)\, n^{-(\rho+1/4)}, n \to \infty, \; -b_n \le r \le b_n - 1, \end{aligned}$$

for n sufficiently large. Hence, $\beta_n = O\left(n^{-(\rho+1/4)}\right)$, $n \to \infty$, and so by (16), it suffices to show that for all $t > 0$,

$$E\left(\left| n^{(\rho+1/4)} K_n \right|^t \right) = O(1), n \to \infty.$$

Since $K_n > 0$ *a.s.*, one can write

$$E\left(\left|n^{(\rho+1/4)}K_n\right|^t\right) = t\int_0^\infty x^{t-1}P\left(n^{(\rho+1/4)}K_n \geq x\right)dx$$
$$\leq \text{constant} + t\int_M^\infty x^{t-1}P\left(n^{(\rho+1/4)}K_n \geq x\right)dx$$

for all $M > 0$. The proof is completed by showing that for $x > 0$,

$$P\left(K_n \geq xn^{-(\rho+1/4)}\right) \leq \sum_{r=-b_n}^{b_n} P\left(G_{r,n} \geq xn^{-(\rho+1/4)}\right)$$
$$\leq O\left(n^{1/4}\right)\exp\left(-\lambda n^{-(\rho-1/2)}x\right),\ n \to \infty. \tag{17}$$

The proof of asymptotic risk efficiency involves finding asymptotic bounds for the moments of the remainder terms in the representations (2) and (6). The desired moment bounds are provided in the next two lemmas.

LEMMA 3. *Suppose F is twice differentiable at ξ_p with $F'(\xi_p) = f(\xi_p) > 0,$ and that $E|X_1|^\lambda < \infty$ for some $\lambda > 0$. Let*

$$R_n = \hat{\xi}_{pn} - \xi_p - \frac{p - F_n(\xi_p)}{f(\xi_p)}.$$

If f' exists and is bounded in a neighborhood of ξ_p, then for any $\gamma <$ ¾ and for every $t > 0$,

$$E\left(\left|n^\gamma R_n\right|^t\right) = O(1),\ n \to \infty.$$

REMARK 2. Duttweiler (1973) obtained a sharper bound for the case when t=2,

$$E\left(R_n^2\right) = O\left(n^{-3/2}\right),\ n \to \infty,$$

as noted earlier.

PROOF OF LEMMA 3. Write

$$R_n = \frac{f(\xi_p) - \left(\hat{\xi}_{pn} - \xi_p\right) - F_n\left(\hat{\xi}_{pn}\right) - F_n(\xi_p) + O\left(\frac{1}{n}\right)}{f(\xi_p)},\ n \to \infty.$$

By Taylor's Theorem, if n is sufficiently large, there exists a number ζ between $\hat{\xi}_{pn}$ and ξ_p such that

$$F\left(\hat{\xi}_{pn}\right) - F(\xi_p) = f(\xi_p)\left(\hat{\xi}_{pn} - \xi_p\right) + \frac{f'(\zeta)}{2}\left(\hat{\xi}_{pn} - \xi_p\right)^2,$$

which implies that

$$R_n = \frac{1}{f(\xi_p)}\left\{F(\hat{\xi}_{pn}) - F(\xi_p) - \left[F_n(\hat{\xi}_{pn}) - F_n(\xi_p)\right] + O\left(\frac{1}{n}\right) - \frac{f'(\zeta)}{2}\left(\hat{\xi}_{pn} - \xi_p\right)^2\right\}. \tag{18}$$

Assume $\gamma \in (\frac{1}{2}, \frac{3}{4})$ and set $\rho = \gamma - \frac{1}{4}$. From (14), (18), and Lemma 2, for every $r > 0$, as $n \to \infty$,

$$E\left(\left|\hat{\xi}_{pn} - \xi_p\right|^r\right) = O(n^{-r/2}) + O(1)\left|F^{-1}\left(\frac{i^*}{n+1}\right) - F^{-1}(p)\right|^r, \tag{19}$$

where $[np] \le i^* \le [np] + 1$. It can be shown that

$$E\left(\left|n^{1/2}\left(\hat{\xi}_{pn} - \xi_p\right)\right|^r\right) = O(1), n \to \infty. \tag{20}$$

This implies that as $n \to \infty$, for $\gamma \in (\frac{1}{2}, \frac{3}{4})$ and hence for any $\gamma < \frac{3}{4}$,

$$E\left(\left|n^\gamma R_n\right|^t I_{\{|\hat{\xi}_{pn} - \xi_p| \le n^{-\rho}\}}\right) = O(1).$$

It follows from (20), Markov's inequality and the Marcinkiewicz-Zygmund Inequality (see, for example, Lemma 2.2.2B of Serfling (1980), or Corollary 10.3.2 of Chow and Teicher (1978)) that as $n \to \infty$, for $\gamma < \frac{3}{4}$, if $\rho = \gamma - \frac{1}{4}$,

$$E\left(\left|n^\gamma R_n\right|^t I_{\{|\hat{\xi}_{pn} - \xi_p| > n^{-\rho}\}}\right) = O(1).$$

This completes the proof of Lemma 3.

LEMMA 4. *Suppose F is twice differentiable at ξ_p with $F'(\xi_p) = f(\xi_p) > 0$, and that $E|X_1|^\lambda < \infty$ for some $\lambda > 0$. Define*

$$\widetilde{R}_n^1 = X_{n:k_{1n}} - \xi_p - \frac{p - \sqrt{\frac{p(1-p)}{n}} - F_n(\xi_p)}{f(\xi_p)},$$

$$\widetilde{R}_n^2 = X_{n:k_{2n}} - \xi_p - \frac{p + \sqrt{\frac{p(1-p)}{n}} - F_n(\xi_p)}{f(\xi_p)},$$

and let $\widetilde{R}_n = \widetilde{R}_n^2 - \widetilde{R}_n^1$. If f' exists and is bounded in a neighborhood of ξ_p, then for any $\gamma < \frac{3}{4}$ and for every $t > 0$,

$$E\left(\left|n^{\gamma}\widetilde{R}_n\right|^{t}\right) = O(1),\ n \to \infty.$$

SKETCH OF PROOF. It can be shown that

$$E\left(\left|n^{1/2}\left(X_{n:k_{1n}} - \xi_p\right)\right|^{r}\right) \le O(1),\ n \to \infty, \tag{21}$$

for every $r > 0$, and that

$$E\left(\left|\widetilde{R}_n^1\right|^{t} I_{\left\{\left|X_{n:k_{1n}} - \xi_p\right| > n^{-\rho}\right\}}\right) \le O(n^{-\gamma t}),\ n \to \infty,$$

as in the proof of Lemma 3. This implies that

$$E\left(\left|n^{\gamma}\widetilde{R}_n^1\right|^{t}\right) = O(1)\ \text{ as } n \to \infty,$$

The proof for $\widetilde{R}_n^2$ is similar.

The following provides conditions for uniform integrability of negative powers of $A^{-1/2}\,T_A$.

LEMMA 5. *Suppose that F is twice differentiable at* ξ_p *with* $F'\left(\xi_p\right) = f\left(\xi_p\right) > 0$, *and that* f' *exists and is bounded in a neighborhood of* ξ_p. *Assume further that* $E\left|X_1\right|^{\lambda} < \infty$ *for some* $\lambda > 0$. *If* T_A *is defined by* (8) *and* $n_A = cA^{\delta}$ *for some* $c > 0$, $\delta > 0$, *then for every* $s > 0$,

$$\left\{\left(A^{-1/2}T_A\right)^{-s},\ A \ge 1\right\}\ \textit{is uniformly integrable.}$$

To prove Lemma 5, by Lemma 1 of Chow and Yu (1981), it suffices to prove that for $s > 0$ and some $\gamma \in (0, 1)$,

$$P\left(T_A \le \gamma\sqrt{A}\right) = o\left(A^{-s/2}\right)\ \text{ as } A \to \infty.$$

This can be shown using Lemma 4, or, alternatively, using (7.3.9) of Sen (1981). The proof is omitted.

The following results of Chow and Yu (1981), which deals with uniform integrability of randomly stopped sums, will be applied in the proofs that follow. The reader is referred to the cited paper for the proof.

LEMMA 6. (Chow and Yu (1981)) *Let* $Y_1, Y_2, \ldots$ *be independent random variables with* $EY_n = 0$ *for each* $n \ge 1$. *Assume that for* $s \ge 2$, $\left\{\left|Y_n\right|^{s},\ n \ge 1\right\}$ *is uniformly integrable. Let* $\mathfrak{F}_n$ *be the* σ*-algebra generated by* $\left\{Y_1, Y_2 \ldots, Y_n\right\}$ *for each* $n \ge 1$, $\mathfrak{F}_0 = (\phi, \Omega)$, *and let* $\left\{M(b), b \in B\right\}$ *be*

$\mathfrak{F}_n$-stopping times with $B \subset (0,\infty)$ such that $\left\{\left(b^{-1}M(b)\right)^{s/2},\ b \in B\right\}$ is uniformly integrable. Let $W_n = \sum_{i=1}^{n} Y_i$. Then $\left\{\left|b^{-1/2}W_{M(b)}\right|^s b \in B\right\}$ is uniformly integrable.

PROOF OF THEOREM 1. Define the sequence of random variables

$$y_n = \frac{f^2(\xi_p)}{\hat{f}^2(\xi_p)} = \frac{nf^2(\xi_p)}{4p(1-p)}\left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)^2$$

and let $g(n) = n^2 \dfrac{f^2(\xi_p)}{p(1-p)}$. It is readily seen that $y_n > 0\ a.s.$, and $y_n \to 1\ a.s.$ as $n \to \infty$. Moreover, $g(n) > 0\ \forall n, g(n) \to \infty$ and $\dfrac{g(n)}{g(n-1)} \to 1$ as $n \to \infty$.

Write $T_A = \inf\left\{n \geq n_A : y_n \leq \dfrac{g(n)}{A}\right\}$. By the definition of T_A,

$$y_{T_A} \leq \frac{g(T_A)}{A} \text{ and } y_{T_A-1} > \frac{g(T_A-1)}{A} I_{\{T_A \neq n_A\}}.$$

Since $n_A = o\left(A^{1/2}\right)$ as $A \to \infty$, then $\dfrac{g(T_A)}{A} \to 1\ a.s.$ as $A \to \infty$. But

$$\frac{g(T_A)}{A} = \frac{T_A^{\,2} f^2(\xi_p)}{A\,p(1-p)} = \left(\frac{T_A}{n_0}\right)^2.$$

Therefore, as in Lemma 1 of Chow and Robbins (1965), (9) is obtained. (10) follows immediately from (9) and Lemma 1 (with $s = 1$).

To prove (11), notice that

$$\frac{\mathcal{R}_{T_A}}{\mathcal{R}_{n_0}} = \frac{f(\xi_p)}{2\sqrt{p(1-p)}}\sqrt{A}E\left(\hat{\xi}_{pT_A} - \xi_p\right)^2 + \frac{f(\xi_p)}{2\sqrt{p(1-p)}}\frac{ET_A}{\sqrt{A}}.$$

By (10), it is enough to show that

$$\frac{f(\xi_p)}{\sqrt{p(1-p)}}\sqrt{A}E\left(\hat{\xi}_{pT_A} - \xi_p\right)^2 \to 1 \text{ as } A \to \infty.$$

By the definition of T_A, $T_A \to \infty\ a.s.$ as $A \to \infty$ and so Bahadur's result (2) holds when the sample size n is replaced by the stopping time T_A, *i.e.*,

$$\hat{\xi}_{pT_A} - \xi_p = \frac{p - F_{T_A}(\xi_p)}{f(\xi_p)} + R_{T_A},$$

where $R_{T_A} = O\left(T_A^{-3/4}(\log T_A)^{3/4}\right)$ $a.s.$ as $A \to \infty$. Write

$$\sqrt{A}E\left(\hat{\xi}_p T_A - \xi_p\right)^2 = \sqrt{A}E\left(\frac{\overline{Y}_{T_A} - p}{f(\xi_p)}\right)^2 + E\left(\sqrt{A}R_{T_A}{}^2\right) - 2E\left(\frac{\sqrt{A}\left(\overline{Y}_{T_A} - p\right)}{f(\xi_p)} \cdot R_{T_A}\right), \qquad (21)$$

where $Y_i = I_{\{X_i \le \xi_p\}}$, $i = 1, 2, \ldots$ By Anscombe's Theorem (Anscombe (1952)), in view of the asymptotic normality of $\overline{Y}_n$ and the almost sure convergence of $A^{-1/2}T_A$,

$$\sqrt{T_A}\,\frac{\left(\overline{Y}_{T_A} - p\right)}{\sqrt{p(1-p)}} \xrightarrow{d} N(0,\ 1) \text{ as } A \to \infty.$$

Hence as $A \to \infty$,

$$\frac{\sqrt{A}\left(\overline{Y}_{T_A} - p\right)^2}{\sqrt{p(1-p)}f(\xi_p)} = \left(\frac{\sqrt{T_A}\left(\overline{Y}_{T_A} - p\right)}{\sqrt{p(1-p)}} \cdot \frac{A^{1/4}(p(1-p))^{1/4}}{\sqrt{T_A}\sqrt{f(\xi_p)}}\right)^2$$

converges in distribution to a Chi-squared distributed random variable with one degree of freedom. By Lemma 1, for every $t > 0$, $\left\{\left(A^{-1/2}T_A\right)^t, A \ge 1\right\}$ is uniformly integrable, and so by Lemma 6, for every $t > 0$,

$$\left\{\left|A^{-1/4}\sum_{i=1}^{T_A}(Y_i - p)\right|^t,\ A \ge 1\right\} \text{ is uniformly integrable.} \qquad (22)$$

Furthermore, by Lemma 5, for every $t > 0$,

$$\left\{\left(A^{-1/2}T_A\right)^{-t},\ A \ge 1\right\} \text{ is uniformly integrable.} \qquad (23)$$

Using Hölder's inequality, one obtains from (22) and (23) that

$$\left\{\left|A^{1/4}\left(\overline{Y}_{T_A} - p\right)\right|^t, A \ge 1\right\} \text{ is uniformly integrable,}$$

for every $t > 0$. Thus, the first term in the right side of (21) converges to $\dfrac{\sqrt{p(1-p)}}{f(\xi_p)}$ as $A \to \infty$. It therefore remains to show that

$$E\left(\sqrt{A}R_{T_A}{}^2\right) \to 0 \text{ as } A \to \infty, \qquad (24)$$

for then the rest of the terms in the right side of (21) would vanish as $A \to \infty$. In view of (23), it is enough to show that for some $s > 1$,

$$E\left(\left|T_A R_{T_A}{}^2\right|^s\right) \to 0 \text{ as } A \to \infty. \tag{25}$$

By Lemma 3, as $A \to \infty$, for any $s > 0$,

$$E\left(\left|T_A R_{T_A}{}^2\right|^s\right) \leq O(1)\sum_{n=n_A}^{\infty} n^{-\beta s} P^{(u-1)/u}\left(T_A = n\right), u > 1,$$

where $0 < \beta < ½$. But by Hölder's inequality, for any integer N,

$$\sum_{n=n_A}^{N} n^{-\beta s} P^{(u-1)/u}\left(T_A = n\right) \leq \left(\sum_{n=n_A}^{\infty} n^{-\beta s u}\right)^{1/u}.$$

Taking limits as $N \to \infty$, one obtains, given any $s > 0$,

$$\sum_{n=n_A}^{\infty} n^{-\beta s} P^{(u-1)/u}\left(T_A = n\right) \leq \left(\sum_{n=n_A}^{\infty} n^{-\beta s u}\right)^{1/u} = O\left(n_A{}^{-\beta s + 1/u}\right) = o(1), A \to \infty,$$

by choosing u sufficiently large. Therefore, (25) is proved and hence, (11). This completes the proof of Theorem 1.

3. EXTENSIONS

3.1 Estimating a Single Quantile Using Other Loss Functions

As in the previous section assume that one wishes to estimate the p^{th} quantile of a distribution function F in some optimal fashion. One may wish to estimate ξ_p by $\hat{\xi}_{pn}$, subject to the loss

$$L_n = A\left|\hat{\xi}_{pn} - \xi_{pn}\right|^r + n, A > 0, \tag{26}$$

where r is some positive real number. In this section, a procedure for estimating the p^{th} quantile of F that is asymptotically risk efficient relative to the loss structure (26) will be derived.

Assume that F is twice differentiable at ξ_p with $F'(\xi_p) = f(\xi_p) > 0$.

$$\text{Then} \quad \sqrt{n}\left(\hat{\xi}_{pn} - \xi_p\right) \xrightarrow{d} N\left(0, \frac{p(1-p)}{f^2(\xi_p)}\right) \text{ as } n \to \infty. \tag{27}$$

Furthermore, if f' is bounded in some neighborhood of ξ_p, and $E|X|_1^{\lambda} < \infty$ for some $\lambda > 0$, one can show that

$$\left\{\left|\sqrt{n}\left(\hat{\xi}_{pn} - \xi_p\right)\right|^r, n \geq 1\right\} \text{ is uniformly integrable,} \tag{28}$$

by applying Bahadur's representation of $\hat{\xi}_{pn}$ and Lemma 3, or Sen's (1959) result. Thus, for large n, the risk is approximately

$$\mathcal{R}_n = EL_n \approx AK_r\left(\frac{p(1-p)}{f^2(\xi_p)}\right)^{r/2}\frac{1}{n^{r/2}} + n, \tag{29}$$

where

$$K_r = E|N(0,1)|^r = \sqrt{\frac{2}{\pi}}\int z^r e^{-z^2/2}\,dz = \Gamma\left(\frac{r+1}{2}\right)\frac{2^{r/2}}{\Gamma\left(\frac{1}{2}\right)}. \tag{30}$$

Differentiating (29) with respect to n (treated as a continuous variable), one finds the best fixed sample size

$$\left[\left(A\frac{r}{2}K_r\left(\frac{p(1-p)}{f^2(\xi_p)}\right)^{r/2}\right)^{\frac{2}{r+2}}\right] \le n_0 \le \left[\left(A\frac{r}{2}K_r\left(\frac{p(1-p)}{f^2(\xi_p)}\right)^{r/2}\right)^{\frac{2}{r+2}}\right] + 1. \tag{31}$$

The corresponding minimum risk is

$$\mathrm{R}_{n_0} \approx \left(1+\frac{2}{r}\right)\left(A\frac{r}{2}K_r\left(\frac{p(1-p)}{f^2(\xi_p)}\right)^{r/2}\right)^{\frac{2}{r+2}} \approx \left(\frac{r+2}{r}\right)n_0.$$

But the best fixed sample size is unavailable when $f(\xi_p)$ is unknown. For this case, the following stopping rule is proposed:

$$\begin{aligned} T_A &= \inf\left\{n \ge n_A:\ \left(\frac{\sqrt{n}}{2}\left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)\right)^r \le \frac{2n^{\frac{r+2}{2}}}{rK_r A}\right\} \\ &= \inf\left\{n \ge n_A:\ \left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)^r \le \frac{2^{r+1}n}{rK_r A}\right\}, \end{aligned} \tag{32}$$

where n_A is an integer which may depend on A, and $\{k_{1n}\}$ and $\{k_{2n}\}$ are as defined in (5). ξ_p can then be estimated by $\hat{\xi}_{pT_A}$. The following theorem summarizes the effectiveness of this sequential procedure as compared with the best fixed-sample-size procedure.

THEOREM 2. *Suppose* $X_1, X_2, \cdots$ *are independent observations with a common distribution function* F. *Assume that* F *is twice differentiable at* ξ_p *with* $F'(\xi_p) = f(\xi_p) > 0$, *and that* $E|X_1|^\lambda < \infty$ *for some* $\lambda > 0$. *Assume further that* f' *exists and is bounded in a neighborhood* *of* ξ_p. *If* T_A *is defined by* (32) *and* $n_A = cA^{\delta}$, *for some* $c > 0$, $\delta \in \left(0, \frac{2}{r+2}\right)$, $r > 0$, *then as* $A \to \infty$,

$$\frac{T_A}{n_0} \to 1 \; a.s., \tag{33}$$

$$E\frac{T_A}{n_0} \to 1 \tag{34}$$

and

$$\frac{\mathrm{R}_{T_A}}{\mathrm{R}_{n_0}} = E\left[\frac{A\left|\hat{\xi}_{pT_A} - \xi_p\right|^r + T_A}{\frac{r+2}{r}\left(A\frac{r}{2}K_r\left(\frac{p\,(1-p)}{f^2(\xi_p)}\right)^{r/2}\right)^{\frac{2}{r+2}}}\right] \to 1. \tag{35}$$

Theorem 2 is a generalization of Theorem 1, and is proved using arguments similar to those of the previous section. A sketch of the proof is provided below.

PROOF OF THEOREM 2. As in Section 2, one can show that

$$\frac{T_A}{n_0} \to 1 \; a.s. \text{ as } A \to \infty$$

and for every $s > 0$,

$$\left\{\left(A^{-2/(r+2)}T_A\right)^s, A \geq 1\right\} \text{ is uniformly integrable.} \tag{36}$$

Hence (33) and (34) are easily obtained. To verify (35), in view of (34), it is enough to show that as $A \to \infty$,

$$A^{\frac{r}{r+2}}E\left|\hat{\xi}_{pT_A} - \xi_p\right|^r \to \frac{2}{r}\left(\frac{r}{2}K_r\left(\frac{p\,(1-p)}{f^2\left(\xi_p\right)}\right)^{r/2}\right)^{\frac{2}{r+2}}. \tag{37}$$

Let $Y_i = I_{\{X_i \leq \xi_p\}}$, $i = 1, 2, \cdots$. As above, one can express $\hat{\xi}_{pT_A} - \xi_p = \frac{p - \overline{Y}_{T_A}}{f\left(\xi_p\right)} + R_{T_A}$, where $R_{T_A} = O\left(T_A^{-3/4}\left(\log T_A\right)^{3/4}\right)$ $a.s.$ as $A \to \infty$. Applying (33), Anscombe's Theorem and the Slutsky-Cramér rule (see, for example, Theorem 1.5.4 of Serfling (1980)), one can verify that as $A \to \infty$,

$$\sqrt{T_A}\left(\hat{\xi}_{pT_A}-\xi_p\right)=\sqrt{T_A}\left(\frac{p-\overline{Y}_{T_A}}{f(\xi_p)}+R_{T_A}\right)\xrightarrow{d} N\left(0,\frac{p(1-p)}{f^2(\xi_p)}\right), \tag{38}$$

and hence as $A\to\infty$,

$$A^{\frac{1}{r+2}}\left(\hat{\xi}_{pT_A}-\xi_p\right)\xrightarrow{d}\left(\frac{2p(1-p)}{rK_r f^2(\xi_p)}\right)^{\frac{1}{r+2}} N(0,1). \tag{39}$$

If

$$\left\{\left|A^{\frac{1}{r+2}}\left(\hat{\xi}_{pT_A}-\xi_p\right)\right|^r, A\geq 1\right\} \text{ is uniformly integrable,} \tag{40}$$

then as $A\to\infty$,

$$A^{\frac{r}{r+2}}E\left|\hat{\xi}_{pT_A}-\xi_p\right|^r\to\left(\frac{2p(1-p)}{rK_r f^2(\xi_p)}\right)^{\frac{r}{r+2}}K_r,$$

which is equivalent to (37). Thus, it suffices to show that (40) holds. But

$$\left|A^{\frac{1}{r+2}}\left(\hat{\xi}_{pT_A}-\xi_p\right)\right|^r=\left|A^{\frac{1}{r+2}}\left(\frac{p-\overline{Y}_{T_A}}{f(\xi_p)}+R_{T_A}\right)\right|^r\leq \text{constant}\left\{\left|A^{\frac{1}{r+2}}\left(\frac{p-\overline{Y}_{T_A}}{f(\xi_p)}\right)\right|^r+\left|A^{\frac{1}{r+2}}R_{T_A}\right|^r\right\}.$$

As in Section 2, one can show that

$$E\left(A^{\frac{1}{r+2}}R_{T_A}\right)^r\to 0 \text{ as } A\to\infty. \tag{41}$$

(This is a generalization of (24).) Therefore, it is enough to verify that

$$\left\{\left|A^{\frac{1}{r+2}}\left(p-\overline{Y}_{T_A}\right)\right|^r, A\geq 1\right\} \text{ is uniformly integrable.} \tag{42}$$

By a slight modification of the proof of Lemma 5, one can establish that for all $s>0$,

$$\left\{\left(A^{-\frac{2}{r+2}}T_A\right)^{-s}, A\geq 1\right\} \text{ is uniformly integrable.} \tag{43}$$

Furthermore, using Lemma 6, one can show that for all $s>0$,

$$\left\{\left|A^{-\frac{1}{r+2}}\sum_{i=1}^{T_A}(Y_i-p)\right|^s, A\geq 1\right\} \text{ is uniformly integrable.} \tag{44}$$

A Hölder's inequality argument will then yield (42), finishing the proof.

3.2 Estimating a Linear Combination of Two Quantiles

Let $X_1, X_2\cdots$ be independent observations from a distribution function F. Suppose that one wishes to estimate a linear combination of two quantiles ξ_p and ξ_q,

$$\theta=\alpha\,\xi_p+\beta\,\xi_q, \text{ for some } 0<p,q<1 \text{ and } \alpha,\beta\in\mathbb{R}.$$

For instance, one may be interested in the interquartile range $R^*=\xi_{.75}-\xi_{.25}$, a measure of dispersion, or the midhinge $M^*=\frac{1}{2}(\xi_{.25}+\xi_{.75})$, which is a measure of central tendency. A natural estimate of θ is the corresponding sample value

$$\theta_n=\alpha\,\hat{\xi}_{pn}+\beta\,\hat{\xi}_{qn}.$$

Without loss of generality, assume that $0<p<q<1$. Under certain smoothness conditions on F, if $f(\xi_p)$ and $f(\xi_q)$ are both nonzero,

$$\sqrt{n}\left(\hat{\xi}_{pn}-\xi_p, \hat{\xi}_{qn}-\xi_q\right)\xrightarrow{d} N((0,0),\Sigma) \text{ as } n\to\infty,$$

where

$$\Sigma=\begin{pmatrix}\dfrac{p(1-p)}{f^2(\xi_p)} & \dfrac{p(1-q)}{f(\xi_p)f(\xi_q)}\\[2ex] \dfrac{p(1-q)}{f(\xi_p)f(\xi_q)} & \dfrac{q(1-q)}{f^2(\xi_q)}\end{pmatrix}.$$

Hence, under such conditions,

$$\sqrt{n}(\theta_n-\theta)\xrightarrow{d} N\left(0,\sigma^{*2}\right) \text{ as } n\to\infty, \tag{45}$$

with

$$\sigma^{*2}=\frac{\alpha^2p(1-p)}{f^2(\xi_p)}+\frac{\beta^2q(1-q)}{f^2(\xi_q)}+\frac{2\alpha\beta p(1-q)}{f(\xi_p)f(\xi_q)}. \tag{46}$$

Consider the problem of estimating θ subject to the loss $L_n = A(\theta_n - \theta)^2 + n, A > 0$. Assume that F is twice differentiable at ξ_p and ξ_q, and that $f(\xi_p)$ and $f(\xi_q)$ are both positive. From (28) and (45), for large n, the risk is asymptotically

$$\mathcal{R}_n = EL_n \approx A\frac{\sigma^{*2}}{n} + n, \tag{47}$$

as in (3). The minimum risk occurs (asymptotically) at $n = n_0$ where

$$\left[\sqrt{A}\sigma^*\right] \le n_0 \le \left[\sqrt{A}\sigma^*\right] + 1,$$

and equals $\mathcal{R}_{n_0} \approx 2\sqrt{A}\sigma^* \approx 2n_0$. In practice, however, $f(\xi_p)$, $f(\xi_q)$, and hence σ^* are unknown, making the best fixed sample size unavailable. This motivates the following sequential procedure.

Define sequences $\{k_{1n}\}$ and $\{k_{2n}\}$ as in (5). Similarly, define

$$h_{1n} = nq - \sqrt{nq(1-q)} \text{ and } h_{2n} = nq + \sqrt{nq(1-q)}. \tag{48}$$

Without loss of generality, assume that these are sequences of integers. As in (7), one has the estimates

$$\hat{f}(\xi_p) = 2\sqrt{\frac{p(1-p)}{n}}\frac{1}{X_{n:k_{2n}} - X_{n:k_{1n}}},$$
$$\hat{f}(\xi_q) = 2\sqrt{\frac{q(1-q)}{n}}\frac{1}{X_{n:h_{2n}} - X_{n:h_{1n}}},$$

and hence

$$\begin{aligned}\hat{\sigma}_n^{*2} = {} & \frac{\alpha^2 n}{4}\left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)^2 + \frac{\beta^2 n}{4}\left(X_{n:h_{2n}} - X_{n:h_{1n}}\right)^2 \\ & + \frac{\alpha\beta}{2}\sqrt{\frac{p(1-q)}{q(1-p)}}\, n\left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)\left(X_{n:h_{2n}} - X_{n:h_{1n}}\right).\end{aligned} \tag{49}$$

Therefore, a natural choice for a stopping rule is

$$\begin{aligned}T_A = \inf\Big\{ n \ge n_A : {} & \frac{\alpha^2 n}{4}\left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)^2 + \frac{\beta^2 n}{4}\left(X_{n:h_{2n}} - X_{n:h_{1n}}\right)^2 \\ & + \frac{\alpha\beta}{2}\sqrt{\frac{p(1-q)}{q(1-p)}}\, n\left(X_{n:k_{2n}} - X_{n:k_{1n}}\right)\left(X_{n:h_{2n}} - X_{n:h_{1n}}\right) \le \frac{n^2}{A}\Big\},\end{aligned} \tag{50}$$

where n_A may depend on A.

THEOREM 3. *Suppose $X_1, X_2 \cdots$ are independent observations with a common distribution function F. Assume that F is twice differentiable at ξ_p and ξ_q, with $F'(\xi_p) = f(\xi_p) > 0$ and $F'(\xi_q) = f(\xi_q) < 0$ and that $E|X_1|^\lambda < \infty$ for some $\lambda > 0$. Assume further that f' exists and is bounded in some neighborhoods of ξ_p and ξ_q. If T_A is defined by* (50), *and $n_A = cA^\delta$ for some $c > 0$, $\delta \in \left(0, \frac{1}{2}\right)$, then as $A \to \infty$,*

$$\frac{T_A}{n_0} \to 1 \ a.s., \tag{51}$$

$$E\frac{T_A}{n_0} \to 1 \tag{52}$$

and

$$\frac{\mathcal{R}_{T_A}}{\mathcal{R}_{n_0}} = \frac{E\left[A\left(\theta_{T_A} - \theta\right)^2 + T_A\right]}{2\sqrt{A}\sigma^*} \to 1. \tag{53}$$

In other words, the stopping rule T_A provides asymptotic risk efficiency, as desired.

The following lemmas, which concern uniform integrability of positive and negative powers of $A^{-1/2}T_A$, are used in the proof of Theorem 3.

LEMMA 7. *Suppose F has a derivative on*

$$B_\varepsilon = \left(F^{-1}(p-\varepsilon), F^{-1}(p+\varepsilon)\right) \cup \left(F^{-1}(q-\varepsilon), F^{-1}(q+\varepsilon)\right),$$

and $F'(x) = f(x) \geq f_0 > 0$ for every x in B_ε, for some $f_0, \varepsilon > 0$. Assume further that $E|X_1|^\lambda < \infty$ for some $\lambda > 0$. If T_A is defined by (50), *then for every $s > 0$,*

$$\left\{\left(A^{-1/2}T_A\right)^s, A \geq 1\right\} \text{ is uniformly integrable.}$$

PROOF. It suffices to show that as $K \to \infty$,

$$P\left(T_A > KA^{1/2}\right) = O\left(K^{-\delta}\right)$$

uniformly in A for all $\delta > 0$. Fix $A > 0$ and let $m = \left[KA^{1/2}\right]$. For $r > 0$,

$$P\left(T_A > KA^{1/2}\right) \leq P\left\{\frac{\alpha^2}{4}\left(X_{m:k_{2m}} - X_{m:k_{1m}}\right)^2 + \frac{\beta^2}{4}\left(X_{m:h_{2m}} - X_{m:h_{1m}}\right)^2 + \frac{\alpha\beta}{2}\sqrt{\frac{p(1-q)}{q(1-p)}}\left(X_{m:k_{2m}} - X_{m:k_{1m}}\right)\left(X_{m:h_{2m}} - X_{m:h_{1m}}\right) > \frac{m}{A}\right\} \tag{54}$$

$$\leq\left(\frac{A}{m}\right)^r O(1)\left\{E\left|\frac{\alpha^2}{4}\left(X_{m:k_{2m}}-X_{m:k_{1m}}\right)^2\right|^r + E\left|\frac{\beta^2}{4}\left(X_{m:h_{2m}}-X_{m:h_{1m}}\right)^2\right|^r\right.$$

$$\left. +E\left|\frac{\alpha\beta}{2}\sqrt{\frac{p(1-q)}{q(1-p)}}\left(X_{m:k_{2m}}-X_{m:k_{1m}}\right)\left(X_{m:h_{2m}}-X_{m:h_{1m}}\right)\right|^r\right\}.$$

As in Section 2, one obtains from Sen's (1959) result that as $m\to\infty$,

$$E\left(\left[X_{m:k_{2m}}-X_{m:k_{1m}}\right]^{2r}\right)=O\left(m^{-r}\right) \text{ and } E\left(\left[X_{m:h_{2m}}-X_{m:h_{1m}}\right]^{2r}\right)=O\left(m^{-r}\right)$$

Thus, as $K\to\infty$, $P\left(T_A > KA^{1/2}\right)\leq A^r m^{-r}\, O\left(m^{-r}\right)=O\left(K^{-2r}\right)$ uniformly in *A*, proving the lemma.

LEMMA 8. *Assume that F is twice differentiable at* ξ_p *and* ξ_q, *with* $F'\left(\xi_p\right)=f\left(\xi_p\right)>0$ *and* $F'\left(\xi_q\right)=f\left(\xi_q\right)>0$, *and that* f' *exists and is bounded in some neighborhoods of* ξ_p *and* ξ_q. *Assume further that* $E|X_1|^\lambda<\infty$ *for some* $\lambda>0$. *If* T_A *is defined by* (50), *and* $n_A=cA^\delta$ *for some* $c>0,\ \delta>0$, *then for every* $s>0$,

$$\{(A^{-1/2}T_A)^{-s},\ A\geq 1\} \text{ is uniformly integrable.}$$

PROOF. It suffices to show that for $s>0$ and some $\gamma\in(0,1)$,

$$P\left(T_A\leq\gamma\sqrt{A}\right)=o\left(A^{-s/2}\right) \text{ as } A\to\infty.$$

Notice that

$$P\left(T_A\leq\gamma\sqrt{A}\right)\leq P\left\{\inf_{n_A\leq j\leq\gamma\sqrt{A}}\left(\frac{\alpha^2}{4}\left(X_{j:k_{2j}}-X_{j:k_{1j}}\right)^2+\frac{\beta^2}{4}\left(X_{j:h_{2j}}-X_{j:h_{1j}}\right)^2\right.\right.$$

$$\left.\left.+\frac{\alpha\beta}{2}\sqrt{\frac{p(1-q)}{q(1-p)}}\left(X_{j:k_{2j}}-X_{j:k_{1j}}\right)\left(X_{j:h_{2j}}-X_{j:h_{1j}}\right)\right)\leq\frac{\gamma\sqrt{A}}{A}\right\}.$$

Consider two cases, (*i*) $\alpha\beta>0$, and (*ii*) $\alpha\beta<0$. Under case (*i*),

$$P\left(T_A\leq\gamma\sqrt{A}\right)\leq P\left\{\inf_{n_A\leq j\leq\gamma\sqrt{A}}\frac{\alpha^2}{4}\left(X_{j:k_{2j}}-X_{j:k_{1j}}\right)^2\leq\gamma A^{-1/2}\right\}$$

$$\leq P\left\{\inf_{n_A\leq j\leq\gamma\sqrt{A}}\sqrt{j}\left(X_{j:k_{2j}}-X_{j:k_{1j}}\right)\leq\frac{2}{\alpha}\sqrt{\gamma A^{-1/2}}\sqrt{\gamma A^{1/2}}\right\}$$

$$=P\left\{\inf_{n_A\leq j\leq\gamma\sqrt{A}}\left(\frac{2\sqrt{p(1-p)}}{f\left(\xi_p\right)}+\sqrt{j}\tilde{R}_{pj}\right)\leq\frac{2\gamma}{\alpha}\right\},$$

where for each j, $\widetilde{R}_{pj}$, is the difference of remainder terms associated with the central order statistics $X_{j:k_{1j}}$ and $X_{j:k_{2j}}$ around $\hat{\xi}_{pj}$. Choose γ small enough so that

$$\frac{2\gamma}{\alpha} \le \frac{\sqrt{p(1-p)}}{f(\xi_p)}.$$

Then from Lemma 4, one obtains that

$$\begin{aligned} P\left(T_A \le \gamma\sqrt{A}\right) &\le P\left\{\inf_{n_A \le j \le \gamma\sqrt{A}} \sqrt{j}\widetilde{R}_{pj} \le \frac{-\sqrt{p(1-p)}}{f(\xi_p)}\right\} \\ &\le \sum_{j=n_A}^{\gamma\sqrt{A}} P\left\{\left|\sqrt{j}\widetilde{R}_{pj}\right| \ge \frac{\sqrt{p(1-p)}}{f(\xi_p)}\right\} \\ &\le O(1)\sum_{j=n_A}^{\infty} O\left(j^{-\alpha t}\right), \alpha \in (0, \tfrac{1}{4}) \\ &= O\left(n_A^{\,1-\alpha t}\right), t > \frac{1}{\alpha}, \end{aligned}$$

which is $o\left(A^{-s/2}\right)$ as $A \to \infty$, for sufficiently large t. For case (*ii*), notice that $p < q$ implies that $\sqrt{\frac{p(1-q)}{q(1-p)}} < 1$. Then, since $\alpha\beta < 0$,

$$\begin{aligned} P(T_A \le \gamma\sqrt{A}) &\le P\Bigg\{\inf_{n_A \le j \le \gamma\sqrt{A}} \Bigg(\frac{\alpha^2}{4}\left(X_{j:k_{2j}} - X_{j:k_{1j}}\right)^2 + \frac{\beta^2}{4}\left(X_{j:h_{2j}} - X_{j:h_{1j}}\right)^2 \\ &\qquad + \frac{\alpha\beta}{2}\left(X_{j:k_{2j}} - X_{j:k_{1j}}\right)\left(X_{j:h_{2j}} - X_{j:h_{1j}}\right)\Bigg) \le \gamma A^{-1/2}\Bigg\} \\ &= P\left\{\inf_{n_A \le j \le \gamma\sqrt{A}} \left(\frac{\alpha}{2}\left(X_{j:k_{2j}} - X_{j:k_{1j}}\right) + \frac{\beta}{2}\left(X_{j:h_{2j}} - X_{j:h_{1j}}\right)\right)^2 \le \gamma A^{-1/2}\right\} \\ &\le \left\{P \inf_{n_A \le j \le \gamma\sqrt{A}} \left[\alpha\left(\frac{2\sqrt{p(1-p)}}{f(\xi_p)} + \sqrt{j}\widetilde{R}_{pj}\right) + \beta\left(\frac{2\sqrt{q(1-q)}}{f(\xi_q)} + \sqrt{j}\widetilde{R}_{qj}\right)\right] \le 2\gamma\right\}, \end{aligned}$$

where $\widetilde{R}_{qj}$ is the analogue of $\widetilde{R}_{pj}$ corresponding to $\hat{\xi}_{qj}$. As in case (*i*), choose γ small enough so that for any $t > 0$,

$$P\left(T_A \le \gamma\sqrt{A}\right) \le O(1)\sum_{j=n_A}^{\infty}\left(E\left|\alpha\sqrt{j}\widetilde{R}_{pj}\right|^t + E\left|\beta\sqrt{j}\widetilde{R}_{qj}\right|^t\right).$$

This yields the desired order of magnitude as $A \to \infty$, by applying Lemma 4 and picking t large enough. Therefore, in either case, for every $s > 0$,

$$P\left(T_A \leq \gamma\sqrt{A}\right) = o\left(A^{-s/2}\right) \text{ as } A \to \infty,$$

proving the lemma.

PROOF OF THEOREM 3. As above, (51) follows immediately from almost sure convergence of $\hat{f}(\xi_p)$ and $\hat{f}(\xi_q)$, and Lemma 1 of Chow and Robbins (1965). From Lemma 7, for every $s > 0$,

$$\left\{\left(A^{-1/2}T_A\right)^s, A \geq 1\right\} \text{ is uniformly integrable.} \tag{55}$$

Hence (52) is clear.

To prove (53), it suffices to show that

$$\frac{\sqrt{A}}{\sigma^*}E\left(\theta_{T_A} - \theta\right)^2 \to 1 \; as \; A \to \infty. \tag{56}$$

Let $Y_i = I_{\{X_i \leq \xi_p\}}$ and $Z_i = I_{\{X_i \leq \xi_q\}}, i = 1, 2, \cdots$. For each n, let R_{pn} and R_{qn} be the remainder terms in the Bahadur representations of $\hat{\xi}_{pn}$ and $\hat{\xi}_{qn}$, respectively. As above, since $T_A \to \infty \; a.s.$ as $A \to \infty$, Bahadur's result allows one to write

$$\frac{\sqrt{A}}{\sigma^*}\left(\theta_{T_A} - \theta\right)^2 = \frac{\sqrt{A}}{\sigma^*}\left(\alpha\left(\frac{p - \overline{Y}_{T_A}}{f(\xi_p)} + R_{pT_A}\right) + \beta\left(\frac{q - \overline{Z}_{T_A}}{f(\xi_q)} + R_{pT_A}\right)\right)^2,$$

where each of R_{pT_A} and R_{qT_A} is $O\left(T_A^{-3/4}(\log T_A)^{3/4}\right)$ $a.s.$ as $A \to \infty$. Let

$$W_i = \alpha\left(\frac{p - Y_i}{f(\xi_p)}\right) + \beta\left(\frac{q - Z_i}{f(\xi_q)}\right), i = 1, 2, \cdots.$$

Then

$$\frac{\sqrt{A}}{\sigma^*}E\left(\theta_{T_A} - \theta\right)^2 = \frac{1}{\sigma^*}E\left(\sqrt{A}\;\overline{W}_{T_A}^{\,2}\right) + \frac{1}{\sigma^*}E\left(\sqrt{A}\left(\alpha\, R_{pT_A} + \beta\, R_{qT_A}\right)^2\right)$$
$$+ \frac{2}{\sigma^*}E\left(\sqrt{A}\;\overline{W}_{T_A}\left(\alpha\, R_{pT_A} + \beta\, R_{qT_A}\right)\right).$$

As in Section 2 it will be argued that

$$E\left(\frac{\sqrt{A}\;\overline{W}_{T_A}^{\,2}}{\sigma^*}\right) \to 1 \text{ as } A \to \infty, \tag{57}$$

and that

$$E\frac{\sqrt{A}}{\sigma^*}\left(\alpha\, R_{pT_A} + \beta\, R_{qT_A}\right)^2 \to 0; \tag{58}$$

these are sufficient to prove (56). Write

$$\frac{\sqrt{A}\ \overline{W}_{T_A}^{\ 2}}{\sigma^*} = \left(\frac{\sqrt{T_A}\ \overline{W}_{T_A}}{\sigma^*}\cdot\frac{\sqrt{\sigma^*}\ A^{1/4}}{\sqrt{T_A}}\right)^2.$$

Since as $A \to \infty$,

$$\frac{\sqrt{n}\ \overline{W}_n}{\sigma^*} = \frac{\sqrt{n}}{\sigma^*}\left(\frac{\alpha\left(p-\overline{Y}_n\right)}{f\left(\xi_p\right)} + \frac{\beta\left(q-\overline{Z}_n\right)}{f\left(\xi_q\right)}\right) \xrightarrow{d} N(0,1)$$

and $\dfrac{T_A}{\sqrt{A}\ \sigma^*} \to 1\ a.s.$, by Anscombe's Theorem, $\dfrac{\sqrt{T_A}\ \overline{W}_{T_A}}{\sigma^*} \xrightarrow{d} N(0,1)$ as $A \to \infty$.

Consequently, $\dfrac{\sqrt{A}\ \overline{W}_{T_A}^{\ 2}}{\sigma^*} \xrightarrow{d} \chi_1^{\ 2}$ as $A \to \infty$.

For $s > 0$, write

$$\left|A^{1/4}\ \overline{W}_{T_1}\right|^s = \left|A^{-1/4}\sum_{i=1}^{T_A} W_i\right|^s \cdot \left|A^{-1/2}\ T_A\right|^{-s}.$$

From Lemma 6, in view of (55), $\left\{\left|A^{-1/4}\sum_{i=1}^{T_A} W_i\right|^s, A \geq 1\right\}$ is uniformly integrable, for every $s > 0$. Moreover, from Lemma 8, for every $s > 0$, $\left\{\left(A^{-1/2}\ T_A\right)^{-s}, A \geq 1\right\}$ is uniformly integrable. Therefore, by a Hölder's inequality argument, $\left\{\left|A^{1/4}\ \overline{W}_{T_A}\right|^s, A \geq 1\right\}$ is uniformly integrable, for every $s > 0$. Hence (57) is proved. (58) follows from the fact that

$$E\left(\sqrt{A}\ R_{pT_A}^{\ 2}\right) \to 0 \text{ and } E\left(\sqrt{A}\ R_{qT_A}^{\ 2}\right) \to 0 \text{ as } A \to \infty,$$

which can be verified as in Section 2 (see (24)). This concludes the proof of Theorem 3.

REMARK 3. The sequential procedure derived above can be modified to work for the loss function

$$A\left|\theta_n - \theta\right|^r + n,\ A > 0,\ r > 0.$$

Similar results can be obtained under the assumptions of Theorem 3, with "$\delta \in \left(0, \dfrac{2}{r+2}\right)$" replacing "$\delta \in \left(0, \dfrac{1}{2}\right)$".

4. SIMULATION STUDIES

Monte Carlo studies were conducted to see how the sequential procedure proposed in Section 2 performed under moderate sample sizes. Expected sample sizes and risks were estimated for selected distributions. All estimates were based on 10,000 simulations.

Expected sample sizes and risks were estimated for the following distributions: (a) standard normal, for p = .5, .75; (b) exponential with mean 1 (EXP), for p = .25, .5, .75; (c) double exponential with mean 0 and scale parameter 1 (DEXP), p = .5; and (d) standard Cauchy, for p = .5. The values of A were 400, 1,000, 4,000 and 10,000. The corresponding values of n_A were 10, 15, 25 and 35. The estimates obtained are provided in Table 1. For each quantile, the entries for each value of A are the estimated expected sample size $\hat{E}T_A$, the best fixed sample size n_0, the estimated risk $\mathcal{R}_{T_A}$, the minimum risk $\mathcal{R}_{n_0}$, and the estimated risk efficiency $\mathcal{R}_{n_0} / \mathcal{R}_{T_A}$. Standard errors of $\hat{E}T_A$ and $\mathcal{R}_{T_A}$ are provided in parentheses.

Table 1 shows that, on average, the sequential rule requires fewer observations than the best fixed-sample-size rule. It is clear from the last column of the table that the procedure worked well in general. Although the estimated risk ratios under the exponential model for p = .75, and the double exponential for p = .5, are not as impressive as the others, they are still around 90% for large values of A. An interesting phenomenon is demonstrated by the exponential model for p = .25, when A = 400: the regret $\left(\mathcal{R}_{T_A} - \mathcal{R}_{n_0}\right)$ is significantly less than zero. In this case, the sequential procedure appears to be actually better than its best fixed-sample-size counterpart.

Table 1.
Simulated Expected Sample Sizes, Risks and Risk Efficiencies for Selected Distributions

Distribution	p	A	$\hat{E}T_A$ (s. e.)	n_0	$\mathcal{R}_{T_1}$ (s. e.)	$\mathcal{R}_{n_0}$	$\mathcal{R}_{n_0} / \mathcal{R}_{T_A}$
N(0,1)	.5	400	20.37	25.07	53.14	50.13	.9434
			(.08)		(.48)		
		1,000	32.39	39.63	83.85	79.27	.9453
			(.12)		(.75)		
		4,000	64.77	79.27	170.20	158.53	.9314
			(.24)		(1.62)		
		10,000	104.41	125.33	270.66	250.66	.9261
			(.34)		(2.50)		
N(0,1)	.75	400	21.02	27.25	59.83	54.51	.9111
			(.10)		(.59)		
		1,000	33.77	43.09	94.59	86.18	.9111
			(.14)		(.89)		
		4,000	68.76	86.18	192.85	172.36	.8938
			(.26)		(1.92)		
		10,000	110.77	136.26	297.41	272.53	.9163
			(.39)		(2.79)		

Table 1. (cont.)

Distribution	p	A	$\hat{E}T_A$ (s. e.)	n_0	$\mathcal{R}_{T_A}$ (s. e.)	$\mathcal{R}_{n_0}$	$\mathcal{R}_{n_0} / \mathcal{R}_{T_A}$
EXP	.25	400	11.84 (.03)	11.55	22.47 (.21)	23.09	1.0278[a]
		1,000	18.26 (.05)	18.26	36.35 (.30)	36.51	1.0045
		4,000	33.25 (.10)	36.51	74.09 (.66)	73.03	.9856
		10,000	50.88 (.16)	57.74	119.20 (.99)	115.47	.9687
EXP	.5	400	16.78 (.07)	20.00	40.08 (.33)	40.00	.9979
		1,000	26.09 (.10)	31.62	64.04 (.56)	63.24	.9876
		4,000	51.31 (.19)	63.24	134.11 (1.16)	126.49	.9432
		10,000	82.02 (.29)	100.00	216.73 (1.93)	200.00	.9228
EXP	.75	400	25.76 (.13)	34.64	77.86 (.81)	69.28	.8898
		1,000	42.11 (.20)	54.77	123.21 (1.22)	109.54	.8891
		4,000	87.16 (.35)	109.54	246.58 (2.40)	219.09	.8885
		10,000	142.63 (.50)	173.20	381.17 (3.56)	346.41	.9088
CAUCHY	.5	400	26.37 (.11)	31.42	65.93 (.67)	62.83	.9530
		1,000	41.69 (.16)	49.67	105.12 (1.02)	99.34	.9450
		4,000	82.80 (.29)	99.34	215.84 (2.11)	198.69	.9206
		10,000	133.17 (.42)	157.08	337.46 (3.00)	314.16	.9310
DEXP	.5	400	19.57 (.08)	20.00	45.92 (.45)	40.00	.8710
		1,000	30.17 (.11)	31.62	71.31 (.68)	63.24	.8869
		4,000	57.91 (.21)	63.24	140.61 (1.32)	126.49	.8996
		10,000	90.37 (.30)	100.00	221.07 (2.09)	200.00	.9047

[a] *Estimated regret is significantly less than zero.*

ACKNOWLEDGMENT

This research was partially supported by the National Science Foundation under Grant DMS 88-02556.

REFERENCES

ANSCOMBE, F. J. (1952). Large sample theory of sequential estimation. *Proc. Cambridge Philos. Soc.* 48:600–607.

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* 37:577–580.

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

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales.* New York: Springer-Verlag.

CHOW, Y. S. and YU, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. *Ann. Statist.* 9:184–189.

DUTTWEILER, D. L. (1973). The mean-square error of Bahadur's order-statistic approximation. *Ann. Statist.* 1:446–453.

FARRELL, R. H. (1966). Bounded length confidence intervals for the p-point of a distribution function, III. *Ann. Math. Statist.* 37:586–592.

GEERTSEMA, J. C. (1970). Sequential confidence intervals based on rank tests. *Ann. Math. Statist.* 41:1016–1026.

——— (1985). On the convergence rate of the coverage probability of a sequential confidence interval based on the sign test. *Sequential Analysis.* 4(3):199–212.

NAVARRO, M. T. (1990). *Sequential Estimation of Quantiles and Adaptive Sequential Estimation of Location and Scale Parameters.* Ph.D. Thesis, University of Illinois.

PRAKASA RAO, B. L. S. (1983). *Nonparametric Functional Estimation.* Orlando: Academic Press.

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

——— (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference.* New York: John Wiley & Sons.

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

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* New York: John Wiley & Sons.

SWANEPOEL, J. W. H. and LOMBARD, F. (1978). Fixed width sequential confidence intervals for the median of a distribution. *Commun. Statist.–Theor. Meth.* A7:829–835.