The Annals of Statistics
1984, Vol. 12, No. 2, 533–550

SEQUENTIAL DETERMINATION OF ESTIMATOR AS WELL AS SAMPLE SIZE[1]

BY ADAM T. MARTINSEK

University of Illinois

For certain point and interval estimation problems, sequential procedures are proposed for choosing an estimator (from the class of trimmed means with trimming proportion in a specified range) as well as a sample size on which to base the estimator. It is shown that these procedures, which do not require knowledge of the best trimming proportion or the asymptotic variance of the corresponding trimmed mean, are asymptotically efficient with respect to the procedure that uses the best trimmed mean (in the specified class) and the best fixed sample size for that trimmed mean.

1. Introduction. The problem of sequential estimation of the mean of a sequence of independent, identically distributed (i.i.d.) observations, with unknown variance and loss equal to a linear combination of squared error and sample size, has been considered by (among others) Robbins (1959), Starr (1966), Starr and Woodroofe (1969) and Woodroofe (1977) for the normal case; by Starr and Woodroofe (1972) and Vardi (1979) for the gamma and Poisson cases (respectively); and by Ghosh and Mukhopadhyay (1979), Chow and Yu (1981), Chow and Martinsek (1982) and Martinsek (1983) in the distribution-free case. In all of these papers the sample mean is used to estimate the population mean, and sequential procedures are used solely to determine an appropriate sample size.

It has been shown by Woodroofe (1977) that in the normal case the "regret" due to using a certain stopping rule to determine the sample size when the variance is unknown, rather than the best fixed sample size when the variance is known, approaches ½ as the cost of error becomes infinite. Woodroofe's result has been generalized by Martinsek (1983), who shows that the "regret" can take arbitrarily large negative values as the distribution of the observations varies (even among symmetric distributions). That is, for some (nonnormal) distributions and large cost of error, it is better to use a sequential procedure when the variance is unknown than to use the best fixed sample size when the variance is known. In effect, the sequential procedure does better by being sensitive to characteristics of the distribution other than the variance.

The amount of improvement that can be realized through sequential determination of the sample size, even when one is restricted to using the sample mean as estimator, is impressive. It seems reasonable to try to improve things even further by not restricting attention to the sample mean, but instead to

Received May 1983; revised November 1983.
[1] Research supported by the National Science Foundation under Grant MCS 83-01834.
AMS 1980 *subject classifications.* Primary 62L12; secondary 62F35, 62G05.
Key words and phrases. Sequential point estimation, fixed-width confidence interval, trimmed mean, asymptotic efficiency, adaptive estimation, uniform integrability.

determine a good estimator (and an appropriate sample size for that estimator) sequentially. After all, in general the sample mean is not necessarily the best estimator of the population mean.

In this paper the problem of sequential determination of estimator as well as sample size is considered in the following situation. Suppose that $X_1, X_2, \cdots$ are i.i.d. observations from a population with distribution function $F(x - \theta)$, where F has a density f which is symmetric about zero and θ is an unknown location parameter. For any n, and any $\alpha \in (0, \frac{1}{2})$, the α-trimmed mean $m_n(\alpha)$, based on $X_1, \cdots, X_n$, is defined by

$$m_n(\alpha) = (n - 2[\alpha n])^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} X_{n:i}, \tag{1.1}$$

where $X_{n:i}$ is the ith order statistic among $X_1, \cdots, X_n$ and [ ] denotes greatest integer. It can be shown that if $F^{-1}(\alpha)$ and $F^{-1}(1 - \alpha)$ are uniquely determined, as $n \rightarrow \infty$,

$$n^{1/2}(m_n(\alpha) - \theta) \rightarrow_d N(0, \sigma^2(\alpha)), \tag{1.2}$$

where

$$\sigma^2(\alpha) = (1 - 2\alpha)^{-2}\left\{2 \int_0^{F^{-1}(1-\alpha)} x^2 f(x)\, dx + 2\alpha[F^{-1}(1 - \alpha)]^2\right\}. \tag{1.3}$$

(Cf. Huber, 1981, Theorem 3.2, pages 60–61; and Serfling, 1980, Theorem C, pages 276–277.)

Assume as in Jaeckel (1971) that we are interested in choosing an estimator from the class of α-trimmed means with trimming proportion $\alpha \in (\alpha_0, \alpha_1)$, for some $0 < \alpha_0 < \alpha_1 < \frac{1}{2}$. In the results that follow, α_0 and α_1 may be taken as close to 0 and ½ (resp.) as desired. Assume further that there is a unique value of α, say α^*, which minimizes $\sigma^2(\alpha)$ over all $\alpha \in (\alpha_0, \alpha_1)$. The case when there is more than one such minimum point is discussed in Section 4. For any $\alpha \in (\alpha_0, \alpha_1)$ and $A > 0$, if n is large and we have convergence of the second moments in (1.2),

$$AE[(m_n(\alpha) - \theta)^2] + n \sim A\sigma^2(\alpha)n^{-1} + n, \tag{1.4}$$

the latter expression being minimized over n by a sample size $n_0 \simeq A^{1/2}\sigma(\alpha)$ (in practice, by one of the two integers closest to this number). Therefore, if A is large, so that a large sample size is required to achieve a small risk under the loss function

$$L_n = A(\delta_n - \theta)^2 + n, \tag{1.5}$$

where δ_n is an estimator of θ, the fixed sample size that minimizes (asymptotically) the risk using the α-trimmed mean $m_n(\alpha)$ as estimate is

$$n_0(\alpha) \simeq A^{1/2}\sigma(\alpha), \tag{1.6}$$

with corresponding minimum risk

$$R_{\alpha,n_0(\alpha)} \simeq 2A^{1/2}\sigma(\alpha). \tag{1.7}$$

Since as $A \to \infty$,

$$R_{\alpha,n_0(\alpha)} \sim 2A^{1/2}\sigma(\alpha) \geq 2A^{1/2}\sigma(\alpha^*) \sim R_{\alpha^*,n_0(\alpha^*)}$$

for all $\alpha \in (\alpha_0, \alpha_1)$, it is clearly best (among all trimmed means in the class under consideration) to choose $\alpha = \alpha^*$ and sample size equal to $n_0(\alpha^*)$. However, in practice one does not know either α^* or $\sigma(\alpha^*)$, so that neither the best trimming proportion nor the best fixed sample size for that trimming proportion is available. One would like to construct a sequential procedure for choosing both a trimming proportion and a sample size whose risk will be close to $R_{\alpha^*,n_0(\alpha^*)}$ for large A.

Jaeckel (1971) has considered the problem of determining the best trimming proportion from the data, using the following idea. For each n and $\alpha \in [\alpha_0, \alpha_1]$, one can estimate $\sigma^2(\alpha)$ by

$$\begin{aligned} s_n^2(\alpha) = (1 - 2\alpha)^{-2}\{n^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} (X_{n:i} - m_n(\alpha))^2 \\ + \alpha[X_{n:[\alpha n]+1} - m_n(\alpha)]^2 + \alpha[X_{n:n-[\alpha n]} - m_n(\alpha)]^2\}. \end{aligned} \tag{1.8}$$

For each n, choose a trimming proportion $\hat{\alpha}_n$ that minimizes $s_n^2(\alpha)$ over all $\alpha \in [\alpha_0, \alpha_1]$, i.e., such that

$$s_n^2(\hat{\alpha}_n) = \min_{\alpha_0 \leq \alpha \leq \alpha_1} s_n^2(\alpha). \tag{1.9}$$

Then, under mild conditions on F, as $n \to \infty$,

$$\hat{\alpha}_n \to_P \alpha^* \tag{1.10}$$

and

$$n^{1/2}(m_n(\hat{\alpha}_n) - \theta) \to_d N(0, \sigma^2(\alpha^*)). \tag{1.11}$$

With this as well as the formula for $n_0(\alpha^*)$ in mind, for sequential point estimation of θ with loss function (1.5), define

$$\begin{aligned} T_A &= \inf\{n \geq 2 : [\min_{\alpha_0 \leq \alpha \leq \alpha_1} s_n^2(\alpha)] + n^{-1} \leq A^{-1}n^2\} \\ &= \inf\{n \geq 2 : s_n^2(\hat{\alpha}_n) + n^{-1} \leq A^{-1}n^2\}, \end{aligned} \tag{1.12}$$

and estimate θ by $m_{T_A}(\hat{\alpha}_{T_A})$. It will be shown in Section 3 that this sequential procedure is asymptotically risk efficient with respect to the procedure that uses the optimal trimming proportion α^* and best fixed sample size $n_0(\alpha^*)$, i.e., as $A \to \infty$,

$$R_{\hat{\alpha}_{T_A}, T_A}/R_{\alpha^*,n_0(\alpha^*)} \to 1,$$

under the assumptions that

$$\begin{aligned} &f(x) \geq f_0 > 0 \quad \text{for all} \quad x \in [F^{-1}(\alpha_0 - \varepsilon_0), F^{-1}(1 - \alpha_0 + \varepsilon_0)], \\ &\qquad \text{for some} \quad f_0, \varepsilon_0 > 0 \text{ (so that in particular (1.2) holds} \\ &\qquad \text{for all } \alpha \in [\alpha_0, \alpha_1]); \end{aligned} \tag{1.13}$$

$$f \text{ has a derivative on } [F^{-1}(\alpha_0 - \varepsilon_0), F^{-1}(1 - \alpha_0 + \varepsilon_0)] \text{ that is continuous a.e.;} \tag{1.14}$$

$$\sigma^{2''}(\alpha^*) > 0 \quad (\text{i.e., } \sigma^{2''}(\alpha^*) \neq 0), \quad \text{where } \sigma^{2''}(\alpha^*) \text{ denotes the second derivative of } \sigma^2(\alpha) \text{ at } \alpha^*; \tag{1.15}$$

and

$$E[|X_1|^{\ell}] < \infty \quad \text{for some } \ell > 0. \tag{1.16}$$

In other words, the sequential procedure which estimates θ by $m_{T_A}(\hat{\alpha}_{T_A})$ performs asymptotically as well as the procedure using the optimal fixed sample size for the optimal trimmed mean (in the class of estimators under consideration).

One can also discuss the problem of finding a confidence interval for θ of prescribed width $2d$ and coverage probability $1 - \beta$ $(0 < \beta < 1)$. Based on (1.2), if one uses the confidence interval

$$m_{k_0(\alpha)}(\alpha) \pm d,$$

where

$$k_0(\alpha) = [d^{-2}\sigma^2(\alpha)z_{1-\beta}^2] + 1$$

and $z_{1-\beta}$ satisfies

$$2\Phi(z_{1-\beta}) - 1 = (2\pi)^{-1/2} \int_{-z_{1-\beta}}^{z_{1-\beta}} \exp\left(\frac{-x^2}{2}\right) dx = 1 - \beta,$$

then as $d \to 0$, $k_0(\alpha) \to \infty$ and

$$\begin{aligned} P[\theta \in m_{k_0(\alpha)}(\alpha) \pm d] &\geq P[\theta \in m_{k_0(\alpha)}(\alpha) \pm z_{1-\beta}\, \sigma(\alpha)k_0(\alpha)^{-1/2}] \\ &= P[(k_0(\alpha))^{1/2} \,|\, m_{k_0(\alpha)}(\alpha) - \theta \,| \leq z_{1-\beta}\sigma(\alpha)] \\ &\to 1 - \beta \quad \text{(asymptotic consistency);} \end{aligned}$$

further, $k_0(\alpha)$ is (asymptotically) the smallest fixed sample size that is asymptotically consistent (when the confidence interval is based on an α-trimmed mean). Since as $d \to 0$,

$$k_0(\alpha) \sim d^{-2}\sigma^2(\alpha)z_{1-\beta}^2 \geq d^{-2}\sigma^2(\alpha^*)z_{1-\beta}^2 \sim k_0(\alpha^*)$$

for all $\alpha \in (\alpha_0, \alpha_1)$, the best choice of trimming proportion is α^* (in the sense of achieving, asymptotically, the smallest sample size that gives asymptotic consistency). As in the point estimation case, in practice one does not know α^* or $\sigma^2(\alpha^*)$. However, for

$$\begin{aligned} T_d &= \inf\{n \geq 2 : [\min_{\alpha_0 \leq \alpha \leq \alpha_1} s_n^2(\alpha)] + n^{-1} \leq d^2 z_{1-\beta}^{-2} n\} \\ &= \inf\{n \geq 2 : s_n^2(\hat{\alpha}_n) + n^{-1} \leq d^2 z_{1-\beta}^{-2} n\}, \end{aligned} \tag{1.17}$$

it will be shown in Section 2 that under (1.13)–(1.16), as $d \to 0$,

$$P[\theta \in m_{T_d}(\hat{\alpha}_{T_d}) \pm d] \to 1 - \beta$$

and

$$E(T_d)/k_0(\alpha^*) \to 1.$$

That is, the sequential procedure which determines both trimming proportion and sample size by (1.17) performs asymptotically as well as the best fixed sample size for the best trimmed mean.

It should be mentioned that for the case of symmetric distributions considered here, Sen (1981) and Jurečková and Sen (1982) have investigated the performance of sequential procedures using general M-, L- and R-estimators of location, rather than the sample mean. However, even in their work the (robust) estimator is decided on in advance, and only the sample size is determined sequentially.

It is clear that the stopping rules T_A and T_d defined above require a great deal of computation (the computational scheme described by Jaeckel, 1971, page 1542, is repeated for each n until one stops). This may seem impractical; however, as Efron (1982, pages 2–3) points out, we are now in an "era of cheap and fast computation" in which such procedures are certainly feasible. Theorems 1 and 2 below assert that, in addition to being feasible, these procedures compensate for ignorance of the best estimator (among a certain class of estimators) as well as ignorance of the best fixed sample size for that estimator, at least asymptotically.

2. Interval estimation.

The performance of the sequential procedure for determining estimator as well as sample size in the fixed-width confidence interval problem is given by the following theorem.

THEOREM 1. *Assume $X_1, X_2, \cdots$ are i.i.d. with distribution function $F(x - \theta)$, where F has a density f that is symmetric about zero. Assume further that there is a unique $\alpha^* \in (\alpha_0, \alpha_1)$ (where $0 < \alpha_0 < \alpha_1 < 1/2$) which minimizes $\sigma^2(\alpha)$ defined by* (1.3). *If T_d is defined by* (1.17) *and* (1.13)–(1.16) *hold, then as $d \to 0$,*

$$P\{\theta \in m_{T_d}(\hat{\alpha}_{T_d}) \pm d\} \to 1 - \beta, \tag{2.1}$$

$$T_d/k_0(\alpha^*) \to 1 \quad a.s. \tag{2.2}$$

and

$$E(T_d)/k_0(\alpha^*) \to 1. \tag{2.3}$$

The proof of Theorem 1 depends on a series of three lemmas, the first of which deals with uniform integrability of positive powers of d^2T_d. It clearly suffices to prove Theorem 1 for the case $\theta = 0$, and this will be assumed without loss of generality throughout the rest of this section.

LEMMA 1. *Under the assumptions* (1.13), (1.14) *and* (1.16), *for every $p > 0$,*

$$\{(d^2T_d)^p : d \leq 1\} \quad \textit{is uniformly integrable.}$$

PROOF. For $d \leq 1$ and $\delta > 0$, if $K > 0$ is sufficiently large, by Theorem 4.3 of

Jurečková and Sen (1982),

$$\begin{aligned} P[T_d > Kd^{-2}] &\le P[s^2_{[Kd^{-2}]}(\alpha_0) > [Kd^{-2}]d^2 z^{-2}_{1-\beta} - 1] \\ &\le P[s^2_{[Kd^{-2}]}(\alpha_0) > 2\sigma^2(\alpha_0)] \\ &\le C(\delta)([Kd^{-2}])^{-1-\delta} \le \tilde{C}(\delta)K^{-1-\delta}. \end{aligned}$$

Hence for all $\delta > 0$, as $K \to \infty$,

$$P[T_d > Kd^{-2}] = O(K^{-1-\delta})$$

uniformly in d. This proves the lemma.

The proof of Theorem 1 requires almost sure convergence of $s^2_n(\hat{\alpha}_n)$ to $\sigma^2(\alpha^*)$ (as opposed to the convergence in probability proved by Jaeckel, 1971). Lemma 2 gives this convergence with a rate that will be needed below.

LEMMA 2. *Under the assumptions* (1.13) *and* (1.14), *for every* $\delta > 0$, *as* $n \to \infty$,

$$n^{(1/2)-\delta}\sup_{\alpha_0 \le \alpha \le \alpha_1} |s^2_n(\alpha) - \sigma^2(\alpha)| \to 0 \quad a.s. \tag{2.4}$$

PROOF. For $\alpha \in [\alpha_0, \alpha_1]$, define

$$\begin{aligned} t^2(\alpha) = (1 - 2\alpha)^{-2}\Bigg\{ & n^{-1} \sum_{i=[\alpha n]+1}^{n-[\alpha n]} \left[F^{-1}\left(\frac{i}{n+1}\right)\right]^2 \\ & + \alpha\left(F^{-1}\left[\frac{[\alpha n]+1}{n+1}\right]\right)^2 + \alpha\left(F^{-1}\left[\frac{n-[\alpha n]}{n+1}\right]\right)^2\Bigg\}. \end{aligned} \tag{2.5}$$

As in the proof of Lemma 2 of Jaeckel (1971),

$$\sup_{\alpha_0 \le \alpha \le \alpha_1} |t^2(\alpha) - \sigma^2(\alpha)| = O(n^{-1}) \quad \text{as} \quad n \to \infty. \tag{2.6}$$

Since F^{-1} is bounded on $[\alpha_0 - \varepsilon_0, 1 - \alpha_0 + \varepsilon_0]$,

$$\begin{aligned} &\sup_{\alpha_0 \le \alpha \le \alpha_1}\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |(X_{n:i} - m_n(\alpha))^2 - [F^{-1}(i/(n+1))]^2| \\ &\quad = \sup_{\alpha_0 \le \alpha \le \alpha_1}\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |X_{n:i} - m_n(\alpha) - F^{-1}(i/(n+1))| \\ &\qquad \cdot |X_{n:i} - m_n(\alpha) + F^{-1}(i/(n+1))| \\ &\quad \le \{\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |X_{n:i} - F^{-1}(i/(n+1))| + \sup_{\alpha_0 \le \alpha \le \alpha_1} |m_n(\alpha)|\} \\ &\qquad \cdot \{\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |X_{n:i} - F^{-1}(i/(n+1))| + \sup_{\alpha_0 \le \alpha \le \alpha_1} |m_n(\alpha)| \\ &\qquad + 2\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} F^{-1}(i/(n+1))\} \\ &\quad = O(1)\{\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |X_{n:i} - F^{-1}(i/(n+1))| + \sup_{\alpha_0 \le \alpha \le \alpha_1} |m_n(\alpha)|\} \\ &\qquad + \{\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |X_{n:i} - F^{-1}(i/(n+1))| + \sup_{\alpha_0 \le \alpha \le \alpha_1} |m_n(\alpha)|\}^2. \end{aligned} \tag{2.7}$$

By a slight modification of Theorem 6 of Csörgő and Révész (1978) and the law of the iterated logarithm for the Kiefer process, for every $\delta > 0$, as $n \to \infty$,

$$\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} | X_{n:i} - F^{-1}(i/(n+1)) | =_{\text{a.s.}} o(n^{-(1/2)+\delta}). \tag{2.8}$$

Because $m_n(\alpha)$ is the average of $X_{n:i}$ for $[\alpha n] + 1 \le i \le n - [\alpha n]$, it follows immediately from (2.8) and the symmetry of F that

$$\begin{aligned} &\sup_{\alpha_0 \le \alpha \le \alpha_1} | m_n(\alpha) | \\ &\qquad = \sup_{\alpha_0 \le \alpha \le \alpha_1} | (n - 2[\alpha n])^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} [X_{n:i} - F^{-1}(i/(n+1))] | \\ &\qquad =_{\text{a.s.}} o(n^{-(1/2)+\delta}) \quad \text{as} \quad n \to \infty. \end{aligned} \tag{2.9}$$

By (2.7), (2.8) and (2.9), since $s_n^2(\alpha)$ and $t^2(\alpha)$ are certain averages of the $(X_{n:i} - m_n(\alpha))^2$ and $[F^{-1}(i/(n+1))]^2$ (respectively) and $(1 - 2\alpha)^{-2} \le (1 - 2\alpha_1)^{-2}$, it is straightforward to show that as $n \to \infty$,

$$\sup_{\alpha_0 \le \alpha \le \alpha_1} | s_n^2(\alpha) - t^2(\alpha) | =_{\text{a.s.}} o(n^{-(1/2)+\delta}). \tag{2.10}$$

Combining (2.10) with (2.6) finishes the proof.

LEMMA 3. *Assume* (1.13), (1.14) *and* (1.15). *Then for every* $\delta > 0$, *as* $n \to \infty$,

$$n^{3/4-\delta}(m_n(\hat{\alpha}_n) - m_n(\alpha^*)) \to 0 \quad a.s. \tag{2.11}$$

PROOF. From Lemma 2,

$$s_n^2(\hat{\alpha}_n) \to \sigma^2(\alpha^*) \quad \text{a.s. as} \quad n \to \infty,$$

and hence as in the proof of Lemma 3 of Jaeckel (1971),

$$\hat{\alpha}_n \to \alpha^* \quad \text{a.s.}$$

Moreover, for any $\delta > 0$, as $n \to \infty$,

$$\begin{aligned} 0 &\le \sigma^2(\hat{\alpha}_n) - \sigma^2(\alpha^*) \\ &= \sigma^2(\hat{\alpha}_n) - s_n^2(\hat{\alpha}_n) + s_n^2(\hat{\alpha}_n) - s_n^2(\alpha^*) + s_n^2(\alpha^*) - \sigma^2(\alpha^*) \\ &\le \sigma^2(\hat{\alpha}_n) - s_n^2(\hat{\alpha}_n) + s_n^2(\alpha^*) - \sigma^2(\alpha^*) =_{\text{a.s.}} o(n^{-(1/2)+\delta}). \end{aligned} \tag{2.12}$$

By Taylor's theorem ($\sigma^2(\alpha)$ has continuous second derivative because f has continuous first derivative), for some $\tilde{\alpha}_n$ between α^* and $\hat{\alpha}_n$,

$$\begin{aligned} \sigma^2(\hat{\alpha}_n) - \sigma^2(\alpha^*) &= \sigma^{2\prime}(\alpha^*)(\hat{\alpha}_n - \alpha^*) + \sigma^{2\prime\prime}(\tilde{\alpha}_n)(\hat{\alpha}_n - \alpha^*)^2/2 \\ &= \sigma^{2\prime\prime}(\tilde{\alpha}_n)(\hat{\alpha}_n - \alpha^*)^2/2. \end{aligned} \tag{2.13}$$

Letting $n \to \infty$ in (2.13), since $\sigma^{2\prime\prime}(\tilde{\alpha}_n) \to \sigma^{2\prime\prime}(\alpha^*) > 0$, from (2.12),

$$\begin{aligned} (\hat{\alpha}_n - \alpha^*)^2 &=_{\text{a.s.}} o(n^{-(1/2)+\delta}), \\ \hat{\alpha}_n - \alpha^* &=_{\text{a.s.}} o(n^{-1/4+\delta/2}) \quad \text{as} \quad n \to \infty. \end{aligned} \tag{2.14}$$

Write

$$\begin{aligned}(2.15)\qquad & m_n(\hat{\alpha}_n) - m_n(\alpha^*) \\ & = [(n - 2[\hat{\alpha}_n n])^{-1} - (n - 2[\alpha^* n])^{-1}] \textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} X_{n:i} + (n - 2[\hat{\alpha}_n n])^{-1} V,\end{aligned}$$

where

$$V = \begin{cases} \sum_{i=[\hat{\alpha}_n n]+1}^{[\alpha^* n]} X_{n:i} + \sum_{i=n-[\alpha^* n]+1}^{n-[\hat{\alpha}_n n]} X_{n:i}, & \text{if } [\hat{\alpha}_n n] < [\alpha^* n] \\ 0, & \text{if } [\hat{\alpha}_n n] = [\alpha^* n] \\ -\sum_{i=[\alpha^* n]+1}^{[\hat{\alpha}_n n]} X_{n:i} - \sum_{i=n-[\hat{\alpha}_n n]+1}^{n-[\alpha^* n]} X_{n:i}, & \text{if } [\hat{\alpha}_n n] > [\alpha^* n]. \end{cases}$$

From (2.14), as $n \to \infty$,

$$\begin{aligned}(2.16)\qquad & |(n - 2[\hat{\alpha}_n n])^{-1} - (n - 2[\alpha^* n])^{-1}| \\ & \sim 2|\hat{\alpha}_n - \alpha^*|(1 - 2\hat{\alpha}_n)^{-1}(1 - 2\alpha^*)^{-1} n^{-1} \\ & \le 2(1 - 2\alpha_1)^{-2}|\hat{\alpha}_n - \alpha^*| n^{-1} =_{\text{a.s.}} o(n^{-5/4+\delta/2}).\end{aligned}$$

Also, by symmetry of F and (2.8), as $n \to \infty$,

$$\begin{aligned}(2.17)\qquad |\textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} X_{n:i}| &= |\textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} (X_{n:i} - F^{-1}(i/(n+1)))| \\ &=_{\text{a.s.}} o(n^{1/2+\delta/2}).\end{aligned}$$

It follows from (2.16) and (2.17) that the first term on the right-hand side of (2.15) is (almost surely) $o(n^{-3/4+\delta})$. As for the second term on the right of (2.15), clearly as $n \to \infty$,

$$\begin{aligned}& (n - 2[\hat{\alpha}_n n])^{-1}|V| \\ & \quad \le O(1)(n - 2[\alpha_1 n])^{-1} n |\alpha_n - \alpha^*| \{\sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} |X_{n:i} - F^{-1}(i/(n+1))|\} \\ & \quad =_{\text{a.s.}} o(n^{-1+1-(1/4)+(\delta/2)-(1/2)+(\delta/2)}) =_{\text{a.s.}} o(n^{-(3/4)+\delta}),\end{aligned}$$

finishing the proof.

REMARK. In the proof of Theorem 1 below, it will suffice to have

$$n^{1/2}(m_n(\hat{\alpha}_n) - m_n(\alpha^*)) \to 0 \quad \text{a.s. as} \quad n \to \infty.$$

It therefore would be enough to require, for some $k > 0$, that $\sigma^2(\alpha)$ possess a continuous $(2k)$th derivative with

$$\sigma^{2'}(\alpha^*) = \sigma^{2''}(\alpha^*) = \cdots = \sigma^{2(2k-1)}(\alpha^*) = 0, \quad \sigma^{2(2k)}(\alpha^*) > 0,$$

since this would give a.s. convergence of $\hat{\alpha}_n$ to α^* with rate $n^{-(1/4)k+\delta}$ for every $\delta > 0$. Almost sure convergence of

$$n^{(1/2)+(1/4)k-\delta}(m_n(\hat{\alpha}_n) - m_n(\alpha^*)), \quad \delta > 0,$$

to zero would then follow as in the proof above. Note that some rate of convergence of $\hat{\alpha}_n$ to α^* is required; without it (i.e., with only the almost sure analogue of Jaeckel's Lemma 3), one would have only

$$n^{(1/2)-\delta}(m_n(\hat{\alpha}_n) - m_n(\alpha^*)) \to 0 \quad \text{a.s.,} \quad \text{for every} \quad \delta > 0,$$

which is insufficient in the present case.

PROOF OF THEOREM 1. Since $T_d \geq d^{-1}z_{1-\beta}$, $T_d \to \infty$ a.s. as $d \to 0$. Hence from Lemma 2, as $d \to 0$,

$$\min_{\alpha_0 \leq \alpha \leq \alpha_1} s^2_{T_d}(\alpha) = s^2_{T_d}(\hat{\alpha}_{T_d}) \to \sigma^2(\alpha^*) \quad \text{a.s.,} \tag{2.18}$$

and

$$\min_{\alpha_0 \leq \alpha \leq \alpha_1} s^2_{T_d-1}(\alpha) = s^2_{T_d-1}(\hat{\alpha}_{T_d-1}) \to \sigma^2(\alpha^*) \quad \text{a.s.} \tag{2.19}$$

But

$$s^2_{T_d}(\hat{\alpha}_{T_d}) + T_d^{-1} \leq d^2 z^{-2}_{1-\beta} T_d$$

and

$$s^2_{T_d-1}(\hat{\alpha}_{T_d-1}) + (T_d - 1)^{-1} > d^2 z^{-2}_{1-\beta}(T_d - 1),$$

so from (2.18) and (2.19), as $d \to 0$,

$$d^2 z^{-2}_{1-\beta} T_d \to \sigma^2(\alpha^*) \quad \text{a.s.,}$$

which proves (2.2). (2.3) is now immediate from (2.2) and Lemma 1 (with $p = 1$). As for (2.1), it follows from Jurečková and Sen (1982) that

$$\{n^{1/2}(m_n(\alpha^*)) : n \geq 1\}$$

is uniformly continuous in probability, i.e., for all ε, $\eta > 0$, there exist ν and $c > 0$ such that, for any $n > \nu$,

$$P\{|m_n(\alpha^*) - m_{n'}(\alpha^*)| < \varepsilon n^{-1/2}, \text{ for all } n' \text{ such that } |n' - n| < cn\} > 1 - \eta.$$

From this result and Lemma 3,

$$\{n^{1/2}(m_n(\hat{\alpha}_n)) : n \geq 1\}$$

is uniformly continuous in probability. But then from (1.11), (2.2) and Anscombe (1952), as $d \to 0$,

$$T_d^{1/2}(m_{T_d}(\hat{\alpha}_{T_d})) \to_d N(0, \sigma^2(\alpha^*)). \tag{2.20}$$

Hence, because $d^2 T_d \to z^2_{1-\beta}\sigma^2(\alpha^*)$ a.s.,

$$\begin{aligned} P\{0 \in m_{T_d}(\hat{\alpha}_{T_d}) \pm d\} &= P\{T_d^{1/2} |m_{T_d}(\hat{\alpha}_{T_d})| \leq dT_d^{1/2}\} \\ &\to 2\Phi(z_{1-\beta}\sigma(\alpha^*)/\sigma(\alpha^*)) - 1 = 2\Phi(z_{1-\beta}) - 1 = 1 - \beta, \end{aligned}$$

proving (2.1).

3. Point estimation. The next theorem summarizes the performance of the sequential procedure which estimates θ by $m_{T_A}(\hat{\alpha}_{T_A})$.

THEOREM 2. *Assume* $X_1, X_2, \cdots$ *are i.i.d. with distribution function* $F(x - \theta)$, *where* F *has a density* f *that is symmetric about zero. Assume further that there is a unique* $\alpha^* \in (\alpha_0, \alpha_1)$ (*where* $0 < \alpha_0 < \alpha_1 < 1/2$) *which minimizes* $\sigma^2(\alpha)$ *defined by* (1.3). *For the loss function* (1.5), *if* T_A *is defined by* (1.12) *and* (1.13)—(1.16) *hold, then as* $A \to \infty$,

$$T_A/n_0(\alpha^*) \to 1 \quad a.s., \tag{3.1}$$

$$E(T_A)/n_0(\alpha^*) \to 1 \tag{3.2}$$

and

$$R_{\hat{\alpha}_{T_A}, T_A}/R_{\alpha^*, n_0(\alpha^*)} \to 1, \tag{3.3}$$

where

$$R_{\hat{\alpha}_{T_A}, T_A} = AE[(m_{T_A}(\hat{\alpha}_{T_A}) - \theta)^2] + ET_A.$$

By the same arguments as in Section 2, it can be shown that

$$T_A/n_0(\alpha^*) \to 1 \quad \text{a.s. as} \quad A \to \infty$$

and

$$\{(A^{-1/2}T_A)^p : A \geq 1\} \quad \text{is uniformly integrable for every} \quad p > 0. \tag{3.4}$$

Hence (3.1) and (3.2) are immediate. Moreover, one has the analogue of (2.20) when $\theta = 0$ (which will be assumed without loss of generality throughout the remainder of this section),

$$T_A^{1/2}(m_{T_A}(\hat{\alpha}_{T_A})) \to_d N(0, \sigma^2(\alpha^*)) \quad \text{as} \quad A \to \infty. \tag{3.5}$$

Again, the argument is the same as that leading to (2.20). In addition to (3.5), certain uniform integrability results will be needed. The first of these deals with uniform integrability of negative powers of $A^{-1/2}T_A$. Frequent use will be made of the fact that $T_A \geq A^{1/3}$ (this follows immediately from the definition of T_A).

LEMMA 4. *Under the assumptions* (1.13), (1.14) *and* (1.16), *for every* $q > 0$,

$$\{(A^{-1/2}T_A)^{-q} : A \geq 1\} \quad \textit{is uniformly integrable.} \tag{3.6}$$

PROOF. By Lemma 1 of Chow and Yu (1981), it suffices to show that for every $q > 0$, for some $\gamma \in (0, 1)$, as $A \to \infty$,

$$P(T_A \leq \gamma A^{1/2}) = o(A^{-q/2}).$$

For $n \geq 1$ and $\alpha_0 \leq \alpha \leq \alpha_1$, because $X_{n:[\alpha n]+1} \leq m_n(\alpha_1) \leq X_{n:n-[\alpha n]}$,

$$\begin{aligned} \alpha_1(X_{n:[\alpha_1 n]+1} - m_n(\alpha_1))^2 &+ \alpha_1(X_{n:n-[\alpha_1 n]} - m_n(\alpha_1))^2 \\ &\leq 2\alpha_1(X_{n:[\alpha n]+1} - X_{n:n-[\alpha n]})^2 \\ &= 4\alpha_1[X_{n:[\alpha n]+1} - (X_{n:[\alpha n]+1} + X_{n:n-[\alpha n]})/2]^2 \\ &\quad + 4\alpha_1[X_{n:n-[\alpha n]} - (X_{n:[\alpha n]+1} + X_{n:n-[\alpha n]})/2]^2 \\ &\leq 4\alpha_1(X_{n:[\alpha n]+1} - m_n(\alpha))^2 + 4\alpha_1(X_{n:n-[\alpha n]} - m_n(\alpha))^2 \\ &\leq 4\alpha_1\alpha^{-1}(1 - 2\alpha)^2 s_n^2(\alpha). \end{aligned} \tag{3.7}$$

Also, for $\alpha_0 \leq \alpha \leq \alpha_1$, clearly

$$\sum_{i=[\alpha_1 n]+1}^{n-[\alpha_1 n]} (X_{n:i} - m_n(\alpha_1))^2 \leq \sum_{i=[\alpha n]+1}^{n-[\alpha n]} (X_{n:i} - m_n(\alpha))^2. \tag{3.8}$$

It follows from (3.7) and (3.8) that for $n \geq 1$ and $\alpha \in [\alpha_0, \alpha_1]$,

$$s_n^2(\alpha_1) \leq (1 - 2\alpha_1)^{-2}(1 - 2\alpha_0)^2(1 + 4\alpha_1\alpha_0^{-1})s_n^2(\alpha) = K_{\alpha_0,\alpha_1} s_n^2(\alpha). \tag{3.9}$$

By the definition of T_A and (3.9), if $\gamma \in (0, 1)$ is close enough to zero so that $\gamma^2 K_{\alpha_0,\alpha_1} \leq \sigma^2(\alpha_1)/2$,

$$\begin{aligned} P(T_A \leq \gamma A^{1/2}) &\leq P\{\min_{A^{1/3} \leq j \leq \gamma A^{1/2}} j^{-2}(\min_{\alpha_0 \leq \alpha \leq \alpha_1} s_j^2(\alpha)) \leq A^{-1}\} \\ &\leq P\{\min_{A^{1/3} \leq j \leq \gamma A^{1/2}} s_j^2(\alpha_1) \leq A^{-1}(\gamma^2 A) K_{\alpha_0,\alpha_1}\} \\ &\leq \sum_{j=[A^{1/3}]}^{[\gamma A^{1/2}]+1} P(s_j^2(\alpha_1) \leq \sigma^2(\alpha_1)/2). \end{aligned} \tag{3.10}$$

By Theorem 4.3 of Jurečková and Sen (1982), for $\delta > (3/2)(q + 1) - 1$, there exist $C(\delta) \in (0, \infty)$ and a positive integer $M(\delta)$ such that for $j \geq M(\delta)$,

$$P(|s_j^2(\alpha_1) - \sigma^2(\alpha_1)| > \sigma^2(\alpha_1)/2) \leq C(\delta) j^{-1-\delta}. \tag{3.11}$$

Hence if $A \geq (M(\delta))^3$, by (3.10) and (3.11),

$$\begin{aligned} P(T_A \leq \gamma A^{1/2}) &\leq C(\delta) \sum_{j=[A^{1/3}]}^{[\gamma A^{1/2}]+1} j^{-1-\delta} \leq C(\delta)([\gamma A^{1/2}] + 1)A^{-(1+\delta)/3} \\ &\leq \tilde{C}(\delta)\gamma A^{1/2-(1+\delta)/3}. \end{aligned}$$

It follows that for every $q > 0$, as $A \to \infty$,

$$P(T_A \leq \gamma A^{1/2}) = o(A^{-q/2}),$$

proving the lemma.

The next two lemmas together will give the moment convergence corresponding to (3.5).

Lemma 5. *Assume* (1.13), (1.14) *and* (1.16). *For every* $p > 0$,

$$\{|T_A^{-1/2} \textstyle\sum_{i=[\alpha^* T_A]+1}^{T_A - [\alpha^* T_A]} X_{T_A:i}|^p : A \geq 1\} \quad \textit{is uniformly integrable.}$$

PROOF. Define $Y_i = (X_i \vee F^{-1}(\alpha^*)) \wedge F^{-1}(1 - \alpha^*)$. The Y_i are independent, bounded, and have mean zero, so by (3.4) and Lemma 5 of Chow and Yu (1981), for all $p > 0$,

$$\{|A^{-1/4} \textstyle\sum_{i=1}^{T_A} Y_i|^p : A \geq 1\} \text{ is uniformly integrable.} \tag{3.12}$$

Furthermore, from Lemma 4 and (3.12), by a Hölder's inequality argument, for all $p > 0$,

$$\{|T_A^{-1/2} \textstyle\sum_{i=1}^{T_A} Y_i|^p : A \geq 1\} \text{ is uniformly integrable.} \tag{3.13}$$

Consider

$$\textstyle\sum_{i=[\alpha^* T_A]+1}^{T_A - [\alpha^* T_A]} X_{T_A:i} - \sum_{i=1}^{T_A} Y_i.$$

Clearly, this difference consists of a random number $M_{T_A,\ell}$ of terms whose absolute value is no larger than

$$|F^{-1}(\alpha^*) - X_{T_A:[\alpha^* T_A]+1}|,$$

plus a random number $M_{T_A,u}$ of terms whose absolute value is no larger than

$$|F^{-1}(1 - \alpha^*) - X_{T_A:T_A - [\alpha^* T_A]}|.$$

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

$$\sup_{[\alpha_0 n] \leq i \leq n - [\alpha_0 n]} E\{|n^{1/2}(X_{n:i} - F^{-1}(i/(n+1)))|^q\} = O(1) \tag{3.14}$$

(cf. Lemma 4.2 of Jurečková and Sen, 1982). Define $M_{n,\ell}$ to be the number of X_i, among the first n, lying between $F^{-1}(\alpha^*)$ and $X_{n:[\alpha^* n]+1}$, and $M_{n,u}$ to be the number lying between $F^{-1}(1 - \alpha^*)$ and $X_{n:n-[\alpha^* n]}$. For $\delta' \in (0, 1)$, on the set

$$\{M_{n,\ell} > n^{1-\delta'}\} \cap \{X_{n:[\alpha^* n]+1} < F^{-1}(\alpha^*)\},$$

by the Mean Value Theorem we have for some $\tilde{\alpha}$ between α^* and

$$([\alpha^* n] + n^{1-\delta'})/(n+1)$$

(assuming $n \geq (1 + \alpha^*)^{1/(1-\delta')}$, so that $\alpha^* \leq ([\alpha^* n] + n^{1-\delta'})/(n+1)$),

$$\begin{aligned} &n^{1/2} f(F^{-1}(\tilde{\alpha})) n^{-\delta'} \\ &\quad = n^{1/2}\{F^{-1}(([\alpha^* n] + n^{1-\delta'})/(n+1)) - F^{-1}(\alpha^*)\} \\ &\quad \leq n^{1/2}\{F^{-1}(([\alpha^* n] + n^{1-\delta'})/(n+1)) - X_{n:[[\alpha^* n]+n^{1-\delta'}]}\}. \end{aligned} \tag{3.15}$$

It follows from (1.13), (3.14) and (3.15) that for $\delta' < \frac{1}{2}$, as $n \to \infty$,

$$P[\{M_{n,\ell} > n^{1-\delta'}\} \cap \{X_{n:[\alpha^* n]+1} < F^{-1}(\alpha^*)\}] = o(n^{-q}) \tag{3.16}$$

for all $q > 0$. By a similar argument

$$P[\{M_{n,\ell} > n^{1-\delta'}\} \cap \{X_{n:[\alpha^* n]+1} \geq F^{-1}(\alpha^*)\}] = o(n^{-q}) \quad \text{for all} \quad q > 0. \tag{3.17}$$

Hence, since $M_{n,\ell} \le n$, for all $q > 0$,

$$
\begin{aligned}
E[(n^{\delta'-1}M_{n,\ell})^q] &\le 1 + n^{\delta' q}P(M_{n,\ell} > n^{1-\delta'}) = 1 + o(n^{(\delta'-1)q}) \\
&= O(1) \quad \text{as} \quad n \to \infty.
\end{aligned} \tag{3.18}
$$

For any $r > 0$, if A is sufficiently large and $s > 2$, taking $q = rs$ in (3.18) and applying (3.4) and Hölder's inequality,

$$
\begin{aligned}
E[(T_A^{\delta'-1}M_{T_A,\ell})^r] &= \textstyle\sum_{j=[A^{1/3}]}^{\infty} E[(j^{\delta'-1}M_{j,\ell})^r I_{\{T_A=j\}}] \\
&\le \textstyle\sum_{j=[A^{1/3}]}^{\infty} E^{1/s}[(j^{\delta'-1}M_{j,\ell})^{rs}]P^{(s-1)/s}(T_A = j) \\
&\le O(1) \textstyle\sum_{j=[A^{1/3}]}^{\infty} P^{(s-1)/s}(T_A \ge j) \\
&\le O(A^{1/2}) + A^{(s-1)/s} \textstyle\sum_{j=[A^{1/2}]}^{\infty} j^{-2(s-1)/s}E^{(s-1)/s}(A^{-1}T_A^2) \\
&= O(A^{1/2}) + O(A^{(s-1)/s}A^{-(s-1)/s+1/2}) \\
&= O(A^{1/2}) \quad \text{as} \quad A \to \infty.
\end{aligned} \tag{3.19}
$$

By an argument similar to that in (3.19), using (3.14) instead of (3.18), for every $r > 0$,

$$
E\{ | T_A^{1/2}(X_{T_A:[\alpha^* T_A]+1} - F^{-1}(\alpha^*)) |^r\} = O(A^{1/2}) \tag{3.20}
$$

as $A \to \infty$. Therefore, for any $\delta' \in (0, \frac{1}{2})$, for all $p > 2/(3\delta')$, from (3.19), (3.20) and Lemma 4, by Hölder's inequality, as $A \to \infty$,

$$
\begin{aligned}
&E\{ | T_A^{-1/2}M_{T_A,\ell}(F^{-1}(\alpha^*) - X_{T_A:[\alpha^* T_A]+1}) |^p\} \\
&\qquad = E\{ | T_A^{-\delta'} T_A^{\delta'-1} M_{T_A,\ell} T_A^{1/2}(F^{-1}(\alpha^*) - X_{T_A:[\alpha^* T_A]+1}) |^p\} \\
&\qquad \le E^{1/3}(T_A^{-3p\delta'})E^{1/3}\{(T_A^{\delta'-1}M_{T_A,\ell})^{3p}\} \\
&\qquad\qquad \cdot E^{1/3}\{ | T_A^{1/2}(F^{-1}(\alpha^*) - X_{T_A:[\alpha^* T_A]+1}) |^{3p}\} \\
&\qquad = O(A^{1/3})E^{1/3}(T_A^{-3p\delta'}) \\
&\qquad = O(A^{-(p\delta'/2)+1/3}) = o(1).
\end{aligned} \tag{3.21}
$$

Similarly, it can be shown that for p sufficiently large,

$$
E\{ | T_A^{-1/2}M_{T_A,u}(F^{-1}(1 - \alpha^*) - X_{T_A:T_A-[\alpha^* T_A]}) |^p\} = o(1) \quad \text{as} \quad A \to \infty. \tag{3.22}
$$

It follows from (3.13), (3.21) and (3.22) that for p sufficiently large (and hence for all $p > 0$),

$$
\{ | T_A^{-1/2} \textstyle\sum_{i=[\alpha^* T_A]+1}^{T_A-[\alpha^* T_A]} X_{T_A:i} |^p : A \ge 1\} \quad \text{is uniformly integrable,}
$$

proving the Lemma.

LEMMA 6. *Under the assumptions* (1.13)–(1.16), *for every $p > 0$, as $A \to \infty$*

$$
E\{ | T_A^{1/2}(m_{T_A}(\hat\alpha_{T_A}) - m_{T_A}(\alpha^*)) |^p\} \to 0. \tag{3.23}
$$

PROOF. It suffices to prove this for p sufficiently large. Proceeding as in (2.9), using (3.14) and Jensen's inequality, for $p \geq 1$, as $n \to \infty$,

$$\begin{aligned} &\sup_{\alpha_0 \le \alpha \le \alpha_1} E\{ | n^{1/2} \hat{m}_n(\alpha) |^p \} \\ &= \sup_{\alpha_0 \le \alpha \le \alpha_1} E\{ | (n - 2[\alpha n])^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} n^{1/2}(X_{n:i} - F^{-1}(i/(n+1))) |^p \} \\ &\le \sup_{\alpha_0 \le \alpha \le \alpha_1} E\{ (n - 2[\alpha n])^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} | n^{1/2}(X_{n:i} - F^{-1}(i/(n+1))) |^p \} \\ &\le \sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} E\{ | n^{1/2}(X_{n:i} - F^{-1}(i/(n+1))) |^p \} = O(1). \end{aligned} \tag{3.24}$$

Also, as in (2.7), using (3.14), (3.24) and Hölder's inequality, for $p \geq 1$, as $n \to \infty$,

$$\sup_{\alpha_0 \le \alpha \le \alpha_1} \sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} E\{ n^{p/2} | (X_{n:i} - \hat{m}_n(\alpha))^2 - [F^{-1}(i/(n+1))]^2 |^p \} = O(1). \tag{3.25}$$

Therefore, by Jensen's inequality, for $p \geq 1$, as $n \to \infty$,

$$\begin{aligned} &\sup_{\alpha_0 \le \alpha \le \alpha_1} E\{ n^{p/2} | n^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} (X_{n:i} - m_n(\alpha))^2 - n^{-1} \sum_{i=[\alpha n]+1}^{n-[\alpha n]} [F^{-1}(i/(n+1))]^2 |^p \} \\ &\le O(1) \sup_{\alpha_0 \le \alpha \le \alpha_1} n^{-1} \textstyle\sum_{i=[\alpha n]+1}^{n-[\alpha n]} E\{ n^{p/2} | (X_{n:i} - m_n(\alpha))^2 - [F^{-1}(i/(n+1))]^2 |^p \} \\ &\le O(1) \sup_{\alpha_0 \le \alpha \le \alpha_1} \sup_{[\alpha_0 n] \le i \le n-[\alpha_0 n]} E\{ n^{p/2} | (X_{n:i} - m_n(\alpha))^2 - [F^{-1}(i/(n+1))]^2 |^p \} \\ &= O(1). \end{aligned} \tag{3.26}$$

Similarly, from (3.25), for $p \geq 1$, as $n \to \infty$,

$$\sup_{\alpha_0 \le \alpha \le \alpha_1} E\left\{ \left| \alpha (X_{n:[\alpha n]+1} - m_n(\alpha))^2 - \alpha \left[F^{-1}\left(\frac{[\alpha n]+1}{n+1} \right) \right]^2 \right|^p \right\} = O(n^{-p/2}) \tag{3.27}$$

and

$$\sup_{\alpha_0 \le \alpha \le \alpha_1} E\left\{ \left| \alpha (X_{n:n-[\alpha n]} - m_n(\alpha))^2 - \alpha \left[F^{-1}\left(\frac{n-[\alpha n]}{n+1} \right) \right]^2 \right|^p \right\} = O(n^{-p/2}). \tag{3.28}$$

It follows from (3.26)–(3.28) that for all $p \geq 1$, as $n \to \infty$,

$$\sup_{\alpha_0 \le \alpha \le \alpha_1} E\{ | s_n^2(\alpha) - t^2(\alpha) |^p \} = O(n^{-p/2}),$$

where $t^2(\alpha)$ is defined by (2.5). Since

$$\begin{aligned} &\sup_{\alpha_0 \le \alpha \le \alpha_1} | t^2(\alpha) - \sigma^2(\alpha) | = O(n^{-1}), \\ &\sup_{\alpha_0 \le \alpha \le \alpha_1} E\{ | s_n^2(\alpha) - \sigma^2(\alpha) |^p \} = O(n^{-p/2}) \quad \text{for all} \quad p \ge 1. \end{aligned} \tag{3.29}$$

It follows immediately from (3.29), (2.12) and (2.13) that as $n \to \infty$,

$$E\{ | n^{1/4}(\hat{\alpha}_n - \alpha^*) |^p \} = O(1) \quad \text{for all} \quad p \ge 1. \tag{3.30}$$

As in the proof of Lemma 3, write

$$\begin{aligned} & m_n(\hat{\alpha}_n) - m_n(\alpha^*) \\ (3.31) \qquad & = \{(n - 2[\hat{\alpha}_n n])^{-1} - (n - 2[\alpha^* n])^{-1}\} \textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} (X_{n:i} - F^{-1}(i/(n+1))) \\ & \quad + (n - 2[\hat{\alpha}_n n])^{-1} V. \end{aligned}$$

By (3.14), (3.30), Jensen's inequality and Hölder's inequality, for any $p \ge 1$, as $n \to \infty$,

$$\begin{aligned} & E\{n^{p/2} \,|\, \{(n - 2[\hat{\alpha}_n n])^{-1} - (n - 2[\alpha^* n])^{-1}\} \textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} (X_{n:i} - F^{-1}(i/(n+1))) \,|^p\} \\ & \quad \le O(1) E\{|\hat{\alpha}_n - \alpha^*|^p \,|\, (n - 2[\alpha^* n])^{-1} \textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} n^{1/2}(X_{n:i} - F^{-1}(i/(n+1))) \,|^p\} \\ (3.32) \qquad & \quad \le O(1) E^{1/2}\{|\hat{\alpha}_n - \alpha^*|^{2p}\} (n - 2[\alpha^* n])^{-1/2} \\ & \qquad \cdot \left(\textstyle\sum_{i=[\alpha^* n]+1}^{n-[\alpha^* n]} E\{|n^{1/2}(X_{n:i} - F^{-1}(i/(n+1)))|^{2p}\}\right)^{1/2} \\ & \quad \le O(n^{-p/4}) (\sup_{[\alpha_0 n] \le i \le n - [\alpha_0 n]} E\{|n^{1/2}(X_{n:i} - F^{-1}(i/(n+1)))|^{2p}\})^{1/2} = O(n^{-p/4}). \end{aligned}$$

Similarly, one can show

$$E\{|(n - 2[\hat{\alpha}_n n])^{-1} V|^p\} = O(n^{-3p/4}). \tag{3.33}$$

Therefore, from (3.31), (3.32) and (3.33), for all $p \ge 1$, as $n \to \infty$,

$$\begin{aligned} (3.34) \qquad & E\{|n^{1/2}(m_n(\hat{\alpha}_n) - m_n(\alpha^*))|^p\} = O(n^{-p/4}), \\ & E\{|n^{3/4}(m_n(\hat{\alpha}_n) - m_n(\alpha^*))|^p\} = O(1). \end{aligned}$$

Proceeding exactly as in (3.19), using (3.34) instead of (3.18), for all $p \ge 1$, as $A \to \infty$,

$$E\{|T_A^{3/4}(m_{T_A}(\hat{\alpha}_{T_A}) - m_{T_A}(\alpha^*))|^p\} = O(A^{1/2}), \tag{3.35}$$

and hence from Lemma 4, if $p > 4$, as $A \to \infty$, by Hölder's inequality,

$$\begin{aligned} E\{|T_A^{1/2}(m_{T_A}(\hat{\alpha}_{T_A}) - m_{T_A}(\alpha^*))|^p\} & = O(A^{1/2}) E^{1/2}(T_A^{-p/2}) \\ & = O(A^{1/2 - p/8}) = o(1), \end{aligned}$$

finishing the proof of Lemma 6.

PROOF OF THEOREM 2. As remarked above, the proofs of (3.1) and (3.2) are exactly analogous to those for (2.2) and (2.3) of Theorem 1. To prove (3.3), write

$$R_{\hat{\alpha}_{T_A}, T_A} / R_{\alpha^*, n_0(\alpha^*)} \sim \{A E[(m_{T_A}(\hat{\alpha}_{T_A}))^2] + E T_A\} / (2A^{1/2}\sigma(\alpha^*)). \tag{3.36}$$

From (3.2), as $A \to \infty$,

$$E T_A / (2A^{1/2}\sigma(\alpha^*)) \to 1/2,$$

and it suffices to show

$$A^{1/2} E[(m_{T_A}(\hat{\alpha}_{T_A}))^2] / \sigma(\alpha^*) \to 1.$$

From (3.1) and (3.5), as $A \to \infty$,

$$A^{1/4}\sigma^{-1/2}(\alpha^*)(m_{T_A}(\hat{\alpha}_{T_A})) = (A^{1/4}\sigma^{1/2}(\alpha^*)T_A^{-1/2})\{T_A^{1/2}(m_{T_A}(\hat{\alpha}_{T_A}))/\sigma(\alpha^*)\} \to_d N(0, 1). \tag{3.37}$$

Hence it is enough to show convergence of the second moments in (3.37). But uniform integrability of all positive powers of

$$T_A^{1/2}(m_{T_A}(\alpha^*))$$

follows from Lemma 5 and $(1 - 2\alpha^*)^{-1} \le (1 - 2\alpha_1)^{-1}$; one then has immediately the uniform integrability of all positive powers of

$$T_A^{1/2}(m_{T_A}(\hat{\alpha}_{T_A})),$$

via Lemma 6. Finally, from Lemma 4 and Hölder's inequality, all positive powers of

$$A^{1/4}\sigma^{-1/2}(\alpha^*)(m_{T_A}(\hat{\alpha}_{T_A})) = (A^{1/4}\sigma^{1/2}(\alpha^*)T_A^{-1/2})\{T_A^{1/2}(m_{T_A}(\hat{\alpha}_{T_A}))/\sigma(\alpha^*)\}$$

are uniformly integrable. In particular, the second moments in (3.37) converge to $E(\chi_1^2) = 1$, completing the proof.

4. Further remarks. One can extend the results of this paper to the case when there is more than one $\alpha \in (\alpha_0, \alpha_1)$ which minimizes $\sigma^2(\alpha)$, using a suggestion of Jaeckel (1971). Define

$$\alpha^* = \inf\{\alpha' \in (\alpha_0, \alpha_1) : \sigma^2(\alpha') = \min_{\alpha_0<\alpha<\alpha_1}\sigma^2(\alpha)\},$$

and let c_n be any sequence of positive constants such that $c_n \to 0$ and $n^{1/2-\varepsilon}c_n \to \infty$ as $n \to \infty$, for some $\varepsilon > 0$. For each n, choose $\hat{\alpha}_n$ to be the smallest $\alpha' \in (\alpha_0, \alpha_1)$ for which

$$s_n^2(\alpha') \le \{\min_{\alpha_0\le\alpha\le\alpha_1}s_n^2(\alpha)\}(1 + c_n).$$

Then $\hat{\alpha}_n \to \alpha^*$ a.s. as $n \to \infty$, and if we define

$$T_d = \inf\{n \ge 2 : s_n^2(\hat{\alpha}_n) + n^{-1} \le d^2 z_{1-\beta}^{-2} n\}$$

and

$$T_A = \inf\{n \ge 2 : s_n^2(\hat{\alpha}_n) + n^{-1} \le A^{-1}n^2\},$$

the conclusions of Theorems 1 and 2 still hold. The proofs given in Sections 2 and 3 go through with slight modifications. For example, in the proof of Lemma 1 one should now choose K large enough so that

$$[Kd^{-2}]d^2 z_{1-\beta}^{-2} - 1 > 2\sigma^2(\alpha_0)(1 + c_n);$$

in the proof of Lemma 3, $\hat{\alpha}_n$ and $m_n(\hat{\alpha}_n) - m_n(\alpha^*)$ will converge (to α^* and 0, respectively) at somewhat slower rates; similarly, the convergence rates in the proof of Lemma 6 will be somewhat slower. The proofs of Lemmas 2 and 5 will be unchanged, while Lemma 4 will be even easier (since the new procedure takes more observations).

There are a number of distributions, among them the Cauchy, the logistic, the t distribution with various degrees of freedom, and the Tukey contamination models, for which the asymptotic variances of the trimmed means are minimized over all $\alpha \in [0, \frac{1}{2}]$ by some $\alpha^* \in (0, \frac{1}{2})$. Therefore, if one takes α_0 and α_1 close to 0 and $\frac{1}{2}$, respectively, one can in many cases achieve performance that is asymptotically the same as for the optimal trimming proportion using the optimal fixed sample size (among the class of all trimmed means). Moreover, by taking $\alpha_0 \simeq 0$ and $\alpha_1 \simeq \frac{1}{2}$, one would hope to do reasonably well even when the best choice of trimming proportion is really 0 or $\frac{1}{2}$ (i.e., when the sample mean or sample median is the best trimmed mean).

It would be of interest to develop fully efficient adaptive sequential estimation procedures, along the lines of Beran (1974), Sacks (1975) and Stone (1975) for the nonsequential case. Instead of determining sequentially an estimator and sample size which are asymptotically efficient with respect to a particular class of estimators (the trimmed means with trimming proportions in a certain range), it would be nice to have asymptotic efficiency with respect to the best fixed sample size for a fully efficient estimator (one whose asymptotic variance is equal to the inverse of the Fisher information). Such a procedure might be based on estimates of the variance of a (nonsequential) fully efficient adaptive estimator, made at each stage until a certain inequality obtained. The procedure would then stop and estimate θ.

Finally, the results of this paper apply only to the case when the distribution of the observations is symmetric about the parameter θ. For asymmetric distributions it also makes sense (maybe even more sense) to select the estimator as well as the sample size sequentially: one would hope that the data would give a good idea of the direction and magnitude of the skewness. Unfortunately, adaptation (or some extension of it) for asymmetric distributions is a very difficult problem, one that has not yet been satisfactorily solved even in the nonsequential setting.

REFERENCES

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

BERAN, R. (1974). Asymptotically efficient adaptive rank estimates in location models. *Ann. Statist.* **2** 63–74.

BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671.

CHOW, Y. S. and MARTINSEK, A. T. (1982). Bounded regret of a sequential procedure for estimation of the mean. *Ann. Statist.* **10** 909–914.

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.

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

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans.* Society for Industrial and Applied Mathematics, Philadelphia.

GHOSH, M. and MUKHOPADHYAY, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Commun. Statist.—Theor. Math.* **A8** 637–651.

HUBER, P. J. (1981). *Robust Statistics.* Wiley, New York.

JAECKEL, L. A. (1971). Some flexible estimates of location. *Ann. Math. Statist.* **42** 1540–1552.

JUREČKOVÁ, J. and SEN, P. K. (1982). M-estimators and L-estimators of location: uniform integrability and asymptotic risk-efficient sequential versions. *Commun. Statist.—Sequential Analysis* **1** 27–56.

MARTINSEK, A. T. (1983). Second order approximation to the risk of a sequential procedure. *Ann. Statist.* **11** 827–836.

ROBBINS, H. E. (1959). Sequential estimation of the mean of a normal population. *Probability and Statistics* (*Harald Cramér Volume*). Almquist, Uppsala.

SACKS, J. (1975). An asymptotically efficient sequence of estimators of a location parameter. *Ann. Statist.* **3** 285–298.

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

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

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

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

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

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

STEIN, C. (1956). Efficient nonparametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 187–196. University of California Press.

STONE, C. (1975). Adaptive maximum likelihood estimation of a location parameter. *Ann. Statist.* **3** 267–284.

VARDI, Y. (1979). Asymptotic optimal sequential estimation: the Poisson case. *Ann. Statist.* **7** 1040–1051.

WOODROOFE, M. (1977). Second order approximations for sequential point and interval estimation. *Ann. Statist.* **5** 984–995.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
273 ALTGELD HALL
1409 W. GREEN STREET
URBANA, ILLINOIS 61801