The Annals of Statistics
1975, Vol. 3, No. 2, 285-298

AN ASYMPTOTICALLY EFFICIENT SEQUENCE OF ESTIMATORS OF A LOCATION PARAMETER[1]

BY JEROME SACKS

Northwestern University

An asymptotically efficient sequence of estimators for the location parameter of an (unknown) symmetric density f is given under weak conditions on the tails of f. The estimators are related to the usual linear combination of order statistics except that the unknown f has to be estimated and only some of the order statistics are used.

1. Introduction. Let $X_1, \cdots, X_n$ be i.i.d. random variables with common distribution function $F(x - \theta)$. The problem is to estimate θ when F is unknown. In (1.6) below, we define a sequence of estimators $\{\tilde{\theta}_n\}$ which is then proved to be asymptotically efficient in the sense that its asymptotic variance coincides with the Cramér–Rao lower bound for any F in the class of distributions defined by Assumption A below.

Such theorems have been the goal of many investigations in recent years. We mention, only in passing, the work of Huber [6], [7] and many others who deal with related problems but with a class of F's more narrow than the class defined by Assumption A below. More directly related to this paper is the work of Stein [9], Bhattacharya [1], van Eeden [11], Weiss and Wolfowitz [12], Takeuchi [10], and Fabian [3].

Stein [9] was the first to indicate the possibility of theorems of the kind given below. Bhattacharya [1] discusses (among other things) the asymptotic efficiency of the maximum likelihood estimator when the observations are grouped and thereby obtains an almost asymptotically efficient sequence of estimators. Van Eeden [11] converts asymptotically efficient tests (see Hájek [5]) into an asymptotically efficient estimator under assumptions stronger than those of Assumption A (in particular, van Eeden requires h' to be monotone on $(0, \frac{1}{2})$). Weiss and Wolfowitz [12] give a sequence of estimators which is asymptotically efficient among all estimators based on trimming away p-percent of the largest observations and q-percent of the smallest observations. It is not hard to see that $\tilde{\theta}_n$ could be easily modified to work in their formulation and without assuming (as Weiss and Wolfowitz do) that h'' exists (see Assumption A for the definition of h). $\tilde{\theta}_n$ is close in spirit to the estimator proposed by Weiss and Wolfowitz. Takeuchi [10] proposes a complicated sequence of estimators but gives no proof of asymptotic efficiency and some of the data in [10] suggest that Takeuchi's estimators may behave poorly when, for example, F is Cauchy. In [3], Fabian gives an asymptotically efficient sequence of estimators by use of stochastic

Received June 1972; revised August 1973.

[1] Research supported by NSF Grant No. GP-28576.

AMS 1970 *subject classifications.* Primary 62G20; Secondary 62G35.

Key words and phrases. Adaptive estimates, location, asymptotic efficiency, order statistics.

approximation methods. The conditions under which Fabian's results hold are overlapping with ours. In particular, he requires a first moment for F, h' continuous on $(0, 1)$, and h increasing on $[0, \frac{1}{2}]$. Our conditions are different in a neighborhood of 0 and are minimal outside a neighborhood of 0. The estimators $\tilde{\theta}_n$ seem to be easier to calculate than Fabian's.

We turn now to the Assumptions and definition of $\tilde{\theta}_n$.

ASSUMPTION A. 1. F has a density f with respect to Lebesgue measure, f is symmetric around 0, the support of f is an interval which may be infinite and $f > 0$ on the interior of the interval.

2. Let $h(t) = f(F^{-1}(t))$ with $h(0) = h(1) = 0$. Then h is absolutely continuous and

$$0 < \int_0^1 h'^2 = E_f(f'/f)^2 < \infty . \tag{1.1}$$

3. There is an $\varepsilon > 0$ (ε may depend on h) such that $h' > 0$ on $(0, \varepsilon)$.

4. Either

4.1. $\liminf_{t\to 0} h'(t) > 0$ and there is $\varepsilon > 0$ and a constant M_ε such that for all $0 < a < b < \varepsilon$,

$$\int_a^b h' \leqq (b - a)M_\varepsilon h(a)/a \tag{1.2}$$

or

4.2. $\liminf_{t\to 0} h'(t) = 0$, h' is bounded on $(0, \varepsilon)$, and

$$h'(b) \leqq M_\varepsilon h(b)/b \qquad \text{for all} \quad 0 < b < \varepsilon . \tag{1.3}$$

A.1 needs little comment except to note that one could relax the condition that $f > 0$; the estimator given in (1.6) would then have to be altered. A.2 is essential for the theorem. A.3 is a mild regularity condition. (1.2) (also (1.3)) is an oscillatory condition which we need in our arguments. Stated as a condition on f, (1.2) becomes, for some $y_0 < 0$,

$$\frac{f(y) - f(x)}{F(y) - F(x)} \leqq M_\varepsilon \frac{f(x)}{F(x)} \tag{1.4}$$

for all $-\infty < x < y < y_0$. Note that (1.2) is satisfied if h' is bounded away from 0 and ∞ on $(0, \varepsilon)$ so that (1.2) has force when $\limsup_{t\to 0} h'(t) = \infty$. When h' is decreasing on $(0, \varepsilon)$, (1.2) is satisfied. Most of the "usual" densities can be handled by A.4.1, but A.4.2 is needed to handle such densities as the Cauchy and other members of the t-family. A.4.2 is satisfied if $h'(t) \sim ct^\alpha$ $(\alpha > 0)$ as $t \to 0$. A.4.2 can be replaced by other conditions, e.g., $h' \in \mathrm{Lip}(\frac{1}{2})$ and $\lim_{t\to 0} h'(t) = 0$. We do not give the proof when A.4.2 holds, but it is easy to do so along the lines of the given proof under A.4.1 with usually less complication.

Everything that follows depends on n, and to alleviate the cumbersome notation we shall suppress this dependence. All limits are as $n \to \infty$ unless stated explicitly otherwise. We shall use the usual o and O notation and the o_p and O_p notation. Statements of the form $\varphi_i = O(1)$ will mean $\sup_i \varphi_i = O(1)$. Similarly with O replaced by o, o_p, or O_p.

Let $\{\gamma\}$, $\{\lambda\}$, $\{d\}$ be sequences of positive numbers satisfying

ASSUMPTION B. 1. $\gamma, \lambda, d \to 0$,
2. $d^{\frac{1}{2}}\gamma^{-1} = o(1)$,
3. $n^{-\frac{1}{2}}d^{-1} = O(1)$,
4. $\gamma|\log \gamma|\lambda^{-2} = O(1)$.

Examples of λ, γ, d which satisfy Assumption B are easy to give. $d = n^{-\frac{1}{2}}$ is an obvious choice for d from B.3 as we shall see below. Then, for example, $\gamma = n^{-\frac{1}{4}+\delta}$, $\lambda = n^{-\frac{1}{8}+\delta}$ for $0 < \delta < \frac{1}{8}$ works as does $\gamma = n^{-\frac{1}{4}} \log n$, $\lambda = n^{-\frac{1}{8}}(\log n)$. If $\{\gamma, \lambda, d\}$ satisfy B then so do $\{\gamma, \lambda', d\}$ where $\lambda' \geqq \lambda$ and $\lambda' \to 0$.

Let $t_0, t_1, \cdots, t_{k+1}$ (k depends on n) satisfy $0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = 1$, $t_i + t_{k-i+1} = 1$, $t_1 > \gamma$, $t_{i+1} - t_i = d$. For simplicity we will assume that $t_i(n+1)$ and $\gamma(n+1)/2$ are integers. When this is not satisfied some minor adjustments have to be made in the arguments in Section 3. When $t(n+1)$ is an integer let $X(t)$ be the $t(n+1)$ ordered observation from the sequence $X_1, \cdots, X_n$. Define

$$(1.5) \qquad \begin{aligned} \tau_i &= t_i + \gamma/2 , & \sigma_i &= t_i - \gamma/2 , \\ \hat{h}_i &= \gamma/(X(\tau_i) - X(\sigma_i)) , & \tilde{h}_i &= (\hat{h}_i + \hat{h}_{k-i+1})/2 . \end{aligned}$$

Thus $\hat{h}_i$ is an estimate of $f(F^{-1}(t_i))$ and $\tilde{h}$ is a symmetrization of $\hat{h}$.

Let I_1 be the smallest i such that, for all $i \leqq j \leqq k - i + 1$, $\tilde{h}_j \geqq \lambda$. Let $I = I_1 + 1$, $J = k - I$ and define

$$(1.6) \qquad \tilde{\theta}_n = \textstyle\sum_I^J (\tilde{h}_{i+1} - \tilde{h}_i)(\tilde{h}_{i+1}X(t_{i+1}) - \tilde{h}_i X(t_i))/\sum_I^J (\tilde{h}_{i+1} - \tilde{h}_i)^2 .$$

A word is in order to explain the motivation behind the definition of $\tilde{\theta}_n$ and the significance of Assumption B. When F is known, then, as is well known, there is an asymptotically efficient linear function of order statistics which, under enough regularity conditions is given by

$$n^{-1} \textstyle\sum_1^n - h''\left(\frac{i}{n+1}\right) h\left(\frac{i}{n+1}\right) X\left(\frac{i}{n+1}\right)\Big/\int_0^1 h'^2 .$$

This is, of course, based on the so-termed score function $J(t) = -h''h/\int_0^1 h'^2$. If we use only the k order statistics $X(t_1), \cdots, X(t_k)$ then the corresponding estimator is

$$(1.7) \qquad d^{-1} \textstyle\sum_1^k - h''(t_i)h(t_i)X(t_i)/\int_0^1 h'^2 .$$

Approximate h'' by a second difference and $\int_0^1 h'^2$ by $d^{-1} \sum_1^k (h(t_{i+1}) - h(t_i))^2$. Then sum the numerator of (1.7) by parts and ignore the end terms. This brings us to (1.6) with h instead of $\tilde{h}$ and the sum from 1 to k instead of from I to J. When F is unknown it is natural to try to estimate $h(t_i)$ which we do by $\tilde{h}(t_i)$ and proceed from there (this idea is common to all previous investigations). Since $X(t)$ is large when t is near 0 and $h(t_i)$ may be hard to estimate accurately when t_i is near 0 it seems natural to "trim" or "censor" a suitable portion of the observations. An obstacle to this program lies in the determination of the appropriate trimming percentage. We determine this by letting the percentage depend on the observations and $\{\lambda\}$.

The choice of γ clearly affects the error in estimating h by $\hat{h}$ and the build-up of errors in $\tilde{\theta}_n$ will be affected by the number ($\sim d^{-1}$) of terms in the sums appearing in (1.6) and we are able to control it only when d, γ satisfy the requirements in Assumption B. The choice of λ is affected by the observation that if λ were too small the effect of trimming would tend to be lost. Some of the requirements in Assumption B are probably technical but $d = O(n^{-\frac{1}{4}})$ is probably as good as one could expect.

In Section 3 we prove

THEOREM. *If Assumptions* A *and* B *hold then* $n^{\frac{1}{2}}(\tilde{\theta}_n - \theta)$ *is asymptotically normal with mean* 0 *and variance* $1/\int_0^1 h'^2$.

When F is known, the estimator of (1.6) with $\hat{h}$ replaced by h can be shown, by arguing as in Section 3, to be asymptotically efficient under Assumptions A and B. This result is related to results of Shorack [8a] (see particularly his Theorem 2) but does not appear to be derivable from Shorack's results.

The proof of the theorem is given in Section 3 and is essentially direct. Lemma 1 is needed to deal with the randomness of I and J. Lemma 2 gives the asymptotic behavior of the denominator of (1.6) and the behavior of the numerator is then studied at (3.16) *et seq*. In Section 2, we gather some needed facts.

We wish to acknowledge some helpful discussions with Donald Ylvisaker. We also wish to thank Jack Kiefer for helpful comments about the result listed below as (2.6).

2. Preliminaries. We collect some facts here that are used frequently in the arguments in Section 3. Most are well known or readily obtained. Assume that $\theta = 0$ with no loss of generality.

Let $U(0) = 0$, $U(1) = 1$, for $0 < t < 1$ with $t(n+1)$ an integer, let $U(t) = F(X(t))$, and for other $t \in [0, 1]$ define $U(t)$ by linear interpolation. The definition of U implies that

$$X(t) = F^{-1}(t) + \int_t^{U(t)} \frac{1}{h}\,. \tag{2.1}$$

When F is the underlying distribution, $U(1/(n+1)), \cdots, U(n/(n+1))$ have the same joint distribution as the n order statistics from a sample of n i.i.d. uniform (on $[0, 1]$) random variables. Let $W(t) = n^{\frac{1}{2}}(U(t) - t)$. Then

$$\sup_{0 \leq t \leq 1} W(t) = O_p(1)\,. \tag{2.2}$$

$$\lim_{\varepsilon \to 0} \sup_{|t-s| \leq \varepsilon} |W(t) - W(s)| = 0\,. \tag{2.3}$$

$$EW(s)W(t) = s(1-t)(1 + O(n^{-1}))\,, \qquad s \leqq t\,. \tag{2.4}$$

If (r_i, s_i), $i = 1, \cdots, 4$ are disjoint intervals then for $2 \leqq m \leqq 4$

$$\begin{aligned} &\text{(a)} \quad E \textstyle\prod_1^m (W(s_i) - W(r_i)) = O(1) \prod_1^m (s_i - r_i)\,, \\ &\text{(b)} \quad E(W(s_1) - W(r_1))^2 = O(1)(s_1 - r_1)\,, \\ &\text{(c)} \quad E(W(s_2) - W(r_2))^2(W(s_1) - W(r_1))^2 = O(1)(s_1 - r_1)(s_2 - r_2)\,. \end{aligned} \tag{2.5}$$

We will use

$$\sup_{0<t<1} |W(t+\varepsilon_n) - W(t)| = O_p(1)\varepsilon_n^{\frac{1}{2}}|\log \varepsilon_n|^{\frac{1}{2}} \quad \text{as } \varepsilon_n \to 0 \tag{2.6}$$

provided $|\log \varepsilon_n|^{\frac{1}{2}}\varepsilon_n^{-\frac{1}{2}}n^{-1} \to 0$. The verification of (2.6) is carried out by estimating $\bar{w}_n = \sup_{0\leqq j \leqq 1/\varepsilon_n} |W((j+1)\varepsilon_n) - W(j\varepsilon_n)|$ by use of estimates of binomial probabilities obtained from Feller [4], pages 168–179 as done by Kiefer [8] (Lemma 2) and then, again following Kiefer, by showing that $P[\bar{w}_n > c/6 \mid w_n > c] \geqq a > 0$ where w_n denotes the left side of (2.6). We will apply (2.6) when $\varepsilon_n = \gamma_n$ and Assumption B guarantees that $|\log \gamma|^{\frac{1}{2}}\gamma^{\frac{1}{2}}n^{-1} = o(1)$.

Let $x_{i_0} < x_{i_0+1} < \cdots < x_{i_1}$. Let y_i, $i_0 \leqq i \leqq i_1$ be numbers in (0, 1). Let z_0 be the smallest of $\{x_{i_0}, y_i, i_0 \leqq i \leqq i_1\}$ and let z_1 be the largest element in the same set. Let $x_{i_0-1} = z_0$, $x_{i_1+1} = z_1$. Then $y_i \in [x_{i+N_i-1}, x_{i+N_i}]$ for some integer N_i which may be positive or negative. Thus, if $y_i > x_i$,

$$|h(y_i) - h(x_i)| \leqq \int_{x_i}^{y_i} |h'| \leqq \sum_{j=i}^{i+N_i-1} \int_{x_j}^{x_{j+1}} |h'|$$

and there is a similar estimate if $y_i < x_i$. Hence

$$\sum_{i_0}^{i_1} |h(y_i) - h(x_i)| \leqq \max_{i_0 \leqq i \leqq i_1} (|N_i| + 1) \int_{z_0}^{z_1} |h'| \tag{2.7}$$

which is a crude estimate but enough for our purposes. (2.7) will be applied when $x_i = \sigma_i$, $y_i = \tau_i$ to yield

$$\sum_{i_0}^{i_1} |h(\tau_i) - h(\sigma_i)| = O(1)\gamma d^{-1} \int_{\sigma_{i_0}}^{\tau_{i_1}} |h'| \,. \tag{2.8}$$

(2.7) will also be applied when $x_i = t_i$ and $y_i \in (t_i, U(t_i))$ so that $|y_i - x_i| = O_p(1)n^{-\frac{1}{2}} = O_p(1)d$ and then $\max_i (|N_i| + 1) = O_p(1)$ and

$$\sum_{i_0}^{i_1} |h(t_i) - h(y_i)| = O_p(1) \int_{z_0}^{z_1} |h'| \,. \tag{2.9}$$

An estimate related to (2.7) is

$$\sum_{i_0}^{i_1} |h(y_i) - h(x_i)|^2 \leqq \max_{i_0 \leqq i \leqq i_1} |y_i - x_i| \max_{i_0 \leqq i \leqq i_1} (|N_i| + 1) \int_{z_0}^{z_1} h'^2 \tag{2.10}$$

which is obtained by noting that $|h(y_i) - h(x_i)|^2 \leqq |y_i - x_i| \int_{x_i}^{y_i} h'^2$, and then applying an argument like that used to get (2.7). When $x_i = \sigma_i$, $y_i = \tau_i$ we get

$$\sum_{i_0}^{i_1} (h(\tau_i) - h(\sigma_i))^2 = O(1)\gamma^2 d^{-1} \int_{\sigma_{i0}}^{\tau_{i1}} h'^2 \tag{2.11}$$

and when $|x_i - y_i| = d\, O_p(1)$ we get

$$\sum_{i_0}^{i_1} (h(y_i) - h(x_i))^2 = O_p(1)d \int_{z_0}^{z_1} h'^2 \,. \tag{2.12}$$

Note also that

$$\lim_{t\to 0} h(t)t^{-\frac{1}{2}} = 0 \tag{2.13}$$

because $h' \in L_2[0, 1]$ and $h(t) = \int_0^t h' \leqq t^{\frac{1}{2}}(\int_0^t h'^2)^{\frac{1}{2}} = o(1)t^{\frac{1}{2}}$ as $t \to 0$. Because $h' \in L_2[0, 1]$ it is also true that, when $t_{I_1} \to 0$, $t_{I_2} \to 1$,

$$\sum_{I_1}^{I_2} (\textstyle\int_{t_i}^{t_{i+1}} h')^2/d \to \int_0^1 h'^2 \tag{2.14}$$

$$\sum_{I_1}^{I_2} (\textstyle\int_{\sigma_i}^{\tau_i} h')^2/\gamma \sim \gamma d^{-1} \int_0^1 h'^2 \,. \tag{2.15}$$

3. Proof of Theorem. As noted in the introduction we will not give the proof when A.4.2 holds instead of A.4.1. Otherwise Assumptions A and B will be assumed throughout. We set $\theta = 0$ with no loss of generality. Whenever we write $\sum$ we mean $\sum_I^J$.

From (2.1), the symmetry of F, and the symmetry in the definition of $\tilde{h}$ (see (1.5)) we obtain

$$\tilde{\theta}_n = \frac{\sum (\tilde{h}_{i+1}-\tilde{h}_i)(\tilde{h}_{i+1}(F^{-1}(U(t_{i+1}))-F^{-1}(t_{i+1}))-\tilde{h}_i(F^{-1}(U(t_i))-F^{-1}(t_i)))}{\sum (\tilde{h}_{i+1}-\tilde{h}_i)^2}. \tag{3.0}$$

Let $V_i = F^{-1}(U(\tau_i)) - F^{-1}(U(\sigma_i))$ and then $\hat{h}_i = \gamma / V_i$. Let $s = \sup\{t \mid 0 \leqq t \leqq \frac{1}{2}, h(t) = \lambda\}$. From A.1, A.2, A.3 we conclude that $h > \lambda$ on $(s, 1-s)$, $h(s) = \lambda$, and $h < \lambda$ on $[0, s)$ if λ is small enough.

LEMMA 1. *$P[|t_I - s| > 2\gamma] \to 0$ as $n \to \infty$.*

PROOF. Let $\delta > 0$ and let $\mathscr{E}$ be the event $\{|\sup_j W(t_j)| \leqq K_1$, all $n \geqq n_0\}$ where n_0, K_1 are such that $P(\mathscr{E}) > 1 - \delta$. Then to prove Lemma 1 it is sufficient to show $P[t_I - s > 2\gamma, \mathscr{E}] \to 0$ and $P[t_I - s < -2\gamma, \mathscr{E}] \to 0$. Let $J_1 = \{j \mid t_j > 2\gamma - 2d + s, t_j \leqq \frac{1}{2}\}$. Then

$$\begin{aligned} P[t_I - s > 2\gamma, \mathscr{E}] &\leqq \textstyle\sum_{J_1} P[\tilde{h}_j \leqq \lambda, \mathscr{E}] \\ &\leqq 2 \textstyle\sum_{J_1} P[\hat{h}_j \leqq \lambda, \mathscr{E}] \\ &\leqq 2 \textstyle\sum_{J_1} P[\gamma\lambda^{-1} \leqq V_j, \mathscr{E}] \\ &= 2 \textstyle\sum_{J_1} P\left[\frac{\gamma n^{\frac{1}{2}}}{\lambda}(h(x_j) - \lambda) < W(\tau_j) - W(\sigma_j), \mathscr{E}\right] \end{aligned} \tag{3.1}$$

where x_j is a (random) point in $(U(\sigma_j), U(\tau_j))$ and because of the definition of $\mathscr{E}$, $\sigma_j - K_1 n^{-\frac{1}{2}} < x_j < \tau_j + K_1 n^{-\frac{1}{2}}$. Let $t_{I_\varepsilon} = \varepsilon$ define I_ε for small ε. Then, if $j \leqq I_\varepsilon, j \in J_1$, we have from A.4.1, $h(x_j) - \lambda > A_\varepsilon(x_j - s) > A_\varepsilon(\sigma_j - s - K_1 n^{-\frac{1}{2}})$. If $j > I_\varepsilon$, $j \in J_1$, and n large enough we get $h(x_j) - \lambda > \rho(\varepsilon)$ from the continuity and positivity of h on $[\varepsilon, \frac{1}{2}]$. Then, using the Chebyshev inequality, (2.5), and Assumption B, the right side of (3.1) is dominated by

$$\begin{aligned} &\textstyle\sum_{j \in J_1; j \leqq I_\varepsilon} P[W(\tau_j) - W(\sigma_j) > \gamma n^{\frac{1}{2}} \lambda^{-1} A_\varepsilon(\sigma_j - s - K_1 n^{-\frac{1}{2}})] \\ &\quad + \textstyle\sum_{j \in J_1; j > I_\varepsilon} P[W(\tau_j) - W(\sigma_j) > \gamma n^{\frac{1}{2}} \lambda^{-1} \rho(\varepsilon)] \\ &\quad = O(1) \textstyle\sum_{j \in J_1; j \leqq I_\varepsilon} \gamma^{-1} n^{-1} \lambda^2 \dfrac{1}{(\sigma_j - s - K_1 n^{-\frac{1}{2}})^2} + O(1)\gamma^{-1} n^{-1} \lambda^2 d^{-1} \\ &\quad = O(1)\gamma^{-1} n^{-1} \lambda^2 \textstyle\sum_{m=1}^{\varepsilon/d} (\gamma + md)^{-2} + o(1) \\ &\quad = O(1)\gamma^{-1} n^{-1} \lambda^2 (\gamma^{-2} \textstyle\sum_1^{\gamma/d} 1 + d^{-2} \sum_{\gamma/d}^{\varepsilon/d} m^{-2}) + o(1) \\ &\quad = O(1)\gamma^{-2} d^{-1} n^{-1} \lambda^2 + O(1)\gamma^{-1} n^{-1} \lambda^2 d^{-2} \textstyle\sum_{\gamma/d}^{\infty} m^{-2} + o(1) \\ &\quad = O(1)\gamma^{-2} d^{-1} n^{-1} \lambda^2 + o(1) = o(1). \end{aligned}$$

Thus $P[t_I - s > 2\gamma] \to 0$.

Let $J_2 = \{j \mid t_j - s < -2\gamma\}$. Then, as above,

$$P[t_I - s < -2\gamma, \mathscr{E}] \leqq 2 \sum_{J_2} P[\lambda^{-1}\gamma n^{\frac{1}{2}}(h(x_j) - \lambda) > W(\tau_j) - W(\sigma_j), \mathscr{E}]$$
$$= O(1)\lambda^2\gamma^{-1}n^{-1} \sum_1^{(s-2\gamma)/d} (\gamma + md)^{-2} = o(1) .$$

The proof of Lemma 1 is now done.

LEMMA 2. *$d^{-1} \sum (\tilde{h}_{i+1} - \tilde{h}_i)^2 \rightarrow \int_0^1 h'^2$ in probability.*

PROOF. To prove Lemma 2, it is enough to show that it holds with $\tilde{h}_i$ replaced by $\hat{h}_i$. Recall the definition of V_i made before Lemma 1. It is easy to obtain

$$V_i = \int_{U(\sigma_i)}^{U(\tau_i)} \left(\frac{1}{h} - \frac{1}{h(\sigma_i)}\right) + \frac{\gamma + n^{-\frac{1}{2}}o_p(1)}{h(\sigma_i)} \tag{3.2}$$
$$= \frac{\gamma}{h(\sigma_i)}\left[1 + \frac{h(\sigma_i) - h(q_i)}{h(q_i)} + o_p(1)\gamma^{-1}n^{-\frac{1}{2}}\right]$$

where $q_i \in (\sigma_i, \tau_i)$. Let

$$a(\tau_i) = \int_{U(\tau_i)}^{U(\tau_{i+1})} \left(\frac{1}{h} - \frac{1}{h(\tau_i)}\right) \tag{3.3}$$
$$= (d + n^{-\frac{1}{2}}\Delta_i^{**}) \left(\frac{1}{h(a_i)} - \frac{1}{h(\tau_i)}\right)$$

for some $a_i \in (U(\tau_i), U(\tau_{i+1}))$ where $\Delta_i^{**} = W(\tau_{i+1}) - W(\tau_i)$. Let $\alpha(\sigma_i)$ be defined similarly with $\Delta_i^* = W(\sigma_{i+1}) - W(\sigma_i)$ instead of Δ_i^{**} and then let

$$\alpha_i = \alpha(\tau_i) - \alpha(\sigma_i) . \tag{3.4}$$

$$\beta_i = \frac{1}{h(\tau_i)} - \frac{1}{h(\sigma_i)} \tag{3.5}$$

$$\delta_i = n^{-\frac{1}{2}}\Delta_i^{**}\beta_i \tag{3.6}$$

$$\xi_i = \frac{n^{-\frac{1}{2}}}{h(\sigma_i)} (\Delta_i^{**} - \Delta_i^*) . \tag{3.7}$$

Then

$$V_{i+1} - V_i = \alpha_i + d\beta_i + \delta_i + \xi_i \tag{3.8}$$

and

$$\hat{h}_{i+1} - \hat{h}_i = \gamma(V_i - V_{i+1})/V_i V_{i+1} \tag{3.9}$$
$$= -h(\sigma_i)h(\sigma_{i+1})\gamma^{-1}(\alpha_i + d\beta_i + \delta_i + \xi_i)(1 + \rho_i)$$

where

$$\rho_i = o_p(1)\gamma^{-1}n^{-\frac{1}{2}} + O(1)\left(\frac{h(\sigma_i) - h(q_i)}{h(q_i)} + \frac{h(\sigma_{i+1}) - h(q_{i+1})}{h(q_{i+1})}\right) . \tag{3.10}$$

From (3.3) we get, with the use of (2.12),

$$\gamma^{-2}d^{-1} \sum h^2(\sigma_i)h^2(\sigma_{i+1})\alpha_i^2(1 + \rho_i)^2$$
$$= (1 + o_p(1))\gamma^{-2}d^{-1}O_p(1)d^2[\sum (\int_{a_i}^{\tau_i} h')^2 + (\int_{a_i'}^{\sigma_i} h')^2] \tag{3.11}$$
$$= O_p(1)\gamma^{-2}d^2 .$$

From (3.5) and (2.15) we obtain

$$\text{(3.12)}\quad \gamma^{-2}d \sum h^2(\sigma_i)h^2(\sigma_{i+1}) \left(\frac{1}{h(\tau_i)} - \frac{1}{h(\sigma_i)}\right)^2 (1+\rho_i)^2 \\ = \gamma^{-2}d(1+o_p(1)) \sum \left(\textstyle\int_{\sigma_i}^{\tau_i} h'\right)^2 = (1+o_p(1)) \textstyle\int_0^1 h'^2 .$$

We also obtain from (3.6)

$$\text{(3.13)}\quad \gamma^{-2}d^{-1} \sum h^2(\sigma_i)h^2(\sigma_{i+1})n^{-1}\Delta_i^{**2}\beta_i^2(1+\rho_i)^2 \\ = (1+o_p(1))n^{-1}d^{-1}\gamma^{-2} \sum \Delta_i^{**2}\left(\textstyle\int_{\sigma_i}^{\tau_i} h'\right)^2 \\ = O_p(1)n^{-1}d^{-1}$$

where we have used

$$\text{(3.14)}\quad E \sum \Delta_i^{**2}\left(\textstyle\int_{\sigma_i}^{\tau_i} h'\right)^2/\gamma \leqq \sum_1^k E\Delta_i^{**2}\left(\textstyle\int_{\sigma_i}^{\tau_i} h'\right)^2/\gamma \\ = O(1)d\gamma d^{-1} = O(1)\gamma .$$

The same kind of argument (using the boundedness of h) and Assumption B produces

$$\text{(3.15)}\quad \gamma^{-2}d^{-1} \sum h^2(\sigma_i)h^2(\sigma_{i+1})\xi_i^2(1+\rho_i)^2 \\ = (1+o_p(1))n^{-1}\gamma^{-2}d^{-1} \sum h^2(\sigma_{i+1})O_p(1)d \\ = O_p(1)n^{-1}\gamma^{-2}d^{-1} = o_p(1) .$$

Combining (3.11), (3.12), (3.13), and (3.15) proves Lemma 2.

We turn now to the numerator in (3.0) and replace $\tilde{h}$ by $\hat{h}$ (this will not affect the conclusions), and put $G_i = n^{\frac{1}{2}}(F^{-1}(U(t_i)) - F^{-1}(t_i))$. Then, using (3.9),

$$\text{(3.16)}\quad n^{\frac{1}{2}}d^{-1} \cdot \text{numerator of (3.0)} \\ = -d^{-1} \sum h(\sigma_i)h(\sigma_{i+1})(\alpha_i + d\beta_i + \delta_i + \xi_i)(1+\rho_i)\left(\frac{G_{i+1}}{V_{i+1}} - \frac{G_i}{V_i}\right) .$$

We will show that the right side of (3.16) is $\sim N(0, \int_0^1 h'^2)$, and this used with Lemma 2 gives the theorem. To this end we prove Lemma 3 below, which, together with (3.29)—(3.34), reduces the problem to the statement at (3.36). A further reduction (described at (3.36)) to (3.42) then follows, after which we prove (3.42), thus concluding the proof of the theorem.

LEMMA 3.

$$d^{-1} \sum h(\sigma_i)h(\sigma_{i+1})(\alpha_i + \delta_i + \xi_i)(1+\rho_i)\left(\frac{G_{i+1}}{V_{i+1}} - \frac{G_i}{V_i}\right) = o_p(1) .$$

(α_i, δ_i, ξ_i are defined in (3.4)—(3.7).)

PROOF. Let $c_i = h(\sigma_i)h(\sigma_{i+1})(\alpha_i + \delta_i + \xi_i)(1+\rho_i)$. Then, from (3.11), (3.13), and (3.15), we get

$$\text{(3.17)}\quad \gamma^{-2}d^{-1} \sum c_i^2/h^2(t_i) = o_p(1) .$$

Hence, using $G_i = O_p(1)/h(t_i)$, Cauchy–Schwarz, (3.17), and Lemma 2,

$$\text{(3.18)}\quad d^{-1} \sum c_i G_i \left(\frac{1}{V_{i+1}} - \frac{1}{V_i}\right) = O_p(1)\gamma^{-1}d^{-1} \sum |c_i| \frac{1}{h(t_i)} |\hat{h}_{i+1} - \hat{h}_i| \\ = o_p(1) .$$

Next, we let $\Delta_i = W(t_{i+1}) - W(t_i)$, note that $V_i = \gamma/h(y_i)$ for some $y_i \in (U(\sigma_i), U(\tau_i))$, and obtain

$$\frac{G_{i+1} - G_i}{V_i} = \Delta_i(\gamma + O_p(1)n^{-\frac{1}{2}})^{-1}h(y_i)/h(x_i) + W(t_i)\frac{(h(x_i) - h(x_{i+1}))h(y_i)}{h(x_i)h(x_{i+1})(\gamma + O_p(1)n^{-\frac{1}{2}})} \tag{3.19}$$

for some $x_i \in (t_i, U(t_i))$. Since $(h(y_i)/h(x_i))W(t_i) = O_p(1)$ we can use Cauchy-Schwarz, (2.12) and (3.17) to conclude that the contribution of the second term on the right side of (3.19) to $d^{-1}\sum c_i(G_{i+1} - G_i)/V_i$ is $o_p(1)$. Use of (3.3), (2.12), and Assumption B yields

$$\begin{aligned} O_p(1)\gamma^{-1}d^{-1}\textstyle\sum h(\sigma_i)h(\sigma_{i+1})|\alpha_i|\,|\Delta_i| \\ = O_p(1)\gamma^{-1}(\textstyle\sum_1^k \Delta_i^2)^{\frac{1}{2}}(\sum (h(\tau_i) - h(a_i))^2)^{\frac{1}{2}} \\ = O_p(1)d^{\frac{1}{2}}\gamma^{-1} = o_p(1)\,. \end{aligned} \tag{3.20}$$

Referring to (3.5) and (3.6) we obtain

$$\begin{aligned} d^{-1}\textstyle\sum h(\sigma_i)h(\sigma_{i+1})\delta_i(1 + \rho_i)\Delta_i\dfrac{h(y_i)}{h(x_i)}\dfrac{1}{\gamma + O_p(n^{-\frac{1}{2}})} \\ = \gamma^{-1}d^{-1}n^{-\frac{1}{2}}\textstyle\sum (1 + \theta_i)\Delta_i^{**}\Delta_i(h(\tau_i) - h(\sigma_i)) \end{aligned} \tag{3.21}$$

where $\theta_i = o_p(1)$. Since $E|\Delta_i^{**}\Delta_i| = O(d)$ by (2.5), we get, with the use of (2.11),

$$\begin{aligned} \gamma^{-1}d^{-1}n^{-\frac{1}{2}}\textstyle\sum \theta_i\Delta_i^{**}\Delta_i(h(\tau_i) - h(\sigma_i)) \\ = o_p(1)\gamma^{-1}n^{-\frac{1}{2}}\textstyle\sum |h(\tau_i) - h(\sigma_i)| \\ = o_p(1)n^{-\frac{1}{2}}d^{-1} \\ = o_p(1)\,. \end{aligned} \tag{3.22}$$

Similarly,

$$\begin{aligned} \gamma^{-1}d^{-1}n^{-\frac{1}{2}}\textstyle\sum_I^{I_\varepsilon} \Delta_i^{**}\Delta_i(h(\tau_i) - h(\sigma_i)) \\ = O_p(1)n^{-\frac{1}{2}}\gamma^{-1}\textstyle\sum_I^{I_\varepsilon} |h(\tau_i) - h(\sigma_i)| \\ = o_p(1)n^{-\frac{1}{2}}d^{-1} = o_p(1)\,. \end{aligned} \tag{3.23}$$

From (2.5), (2.11) we get

$$\begin{aligned} E(\textstyle\sum_{I_\varepsilon}^{J_\varepsilon} \Delta_i^{**}\Delta_i(h(\tau_i) - h(\sigma_i)))^2 = O(1)\sum_{I_\varepsilon}^{J_\varepsilon} d^2(h(\tau_i) - h(\sigma_i))^2 + O(d^2) \\ = O(1)d\gamma^2\,, \end{aligned}$$

and then

$$\begin{aligned} \gamma^{-1}d^{-1}n^{-\frac{1}{2}}\textstyle\sum_{I_\varepsilon}^{J_\varepsilon} \Delta_i^{**}\Delta_i(h(\tau_i) - h(\sigma_i)) = O_p(1)n^{-\frac{1}{2}}d^{-\frac{1}{2}} \\ = o_p(1)\,. \end{aligned} \tag{3.24}$$

Combining (3.22), (3.23), (3.24), shows that the left side of (3.21) is $o_p(1)$.

Use (3.10) and argue as in (3.22), (3.23), (3.24) to obtain

$$\begin{aligned} d^{-1}\textstyle\sum h(\sigma_i)h(\sigma_{i+1})\xi_i(1 + \rho_i)\Delta_i\dfrac{h(y_i)}{h(x_i)}\dfrac{1}{\gamma + O_p(n^{-\frac{1}{2}})} \\ = o_p(1) + \gamma^{-1}n^{-\frac{1}{2}}d^{-1}\textstyle\sum \Delta_i(\Delta_i^{**} - \Delta_i^*)h(y_i)\,\cdot \end{aligned} \tag{3.25}$$

Let $t_{I_s} = s$. From $E|\Delta_i \Delta_i^{**}| = O(d)$, we get

$$\begin{aligned}\gamma^{-1}n^{-\frac{1}{2}}d^{-1} \textstyle\sum_{\max(I_s-2\gamma d-1,1)}^{I_s} |\Delta_i \Delta_i^{**}|\,|h(y_i)| &= \gamma^{-1}n^{-\frac{1}{2}}d^{-1}O_p(1)\gamma \max_{i\leq I_s}|h(y_i)| \\ &= o_p(1)n^{-\frac{1}{2}}d^{-1} = o_p(1)\end{aligned}$$

and use of Lemma 1 then shows that

$$\gamma^{-1}n^{-\frac{1}{2}}d^{-1} \textstyle\sum_{I}^{I_s} \Delta_i \Delta_i^{**} h(y_i) = o_p(1) \,.$$

$E(\sum_{I_s}^{J_s} \Delta_i \Delta_i^{**} h(\sigma_i))^2 = O(d)$ so that

$$\gamma^{-1}n^{-\frac{1}{2}}d^{-1} \textstyle\sum_{I_s}^{J_s} \Delta_i \Delta_i^{**} h(\sigma_i) = O_p(1)\gamma^{-1}n^{-\frac{1}{2}}d^{-\frac{1}{2}} = o_p(1) \,. \tag{3.26}$$

Now let ψ be a continuous function with $\int_0^1 (h' - \psi)^2 < \delta$. Then

$$h(y_i) - h(\sigma_i) = \textstyle\int_{\sigma_i}^{y_i} (h' - \psi) + \gamma\psi(\sigma_i) + o_p(1)\gamma \,.$$

The argument which yields (3.26) and some by-now obvious calculations produces

$$\begin{aligned}|\gamma^{-1}n^{-\frac{1}{2}}d^{-1} &\textstyle\sum_{I_s}^{J_s} \Delta_i \Delta_i^{**}(h(y_i) - h(\sigma_i))| \\ &\leq \gamma^{-1}n^{-\frac{1}{2}}d^{-1}(\textstyle\sum_{I_s}^{J_s} (\Delta_i \Delta_i^{**})^2)^{\frac{1}{2}}(\sum_{I_s}^{J_s} (\int_{\sigma_i}^{y_i} (h' - \psi))^2)^{\frac{1}{2}} + o_p(1) \\ &= O_p(1)\gamma^{-1}n^{-\frac{1}{2}}d^{-1}d^{\frac{1}{2}}(\textstyle\int_0^1 (h' - \psi)^2)^{\frac{1}{2}}\gamma d^{-\frac{1}{2}} = O_p(1)\delta \,.\end{aligned}$$

These calculations imply that the left side of (3.25) is $o_p(1)$, and this together with (3.20), (3.21), and the remark preceding (3.20) yields

$$d^{-1} \sum c_i \frac{G_{i+1} - G_i}{V_i} = o_p(1)$$

which with (3.18) gives the conclusion of Lemma 3.

To enable the further reduction of (3.16) we note

$$\sum \gamma^{-1}|h(t_{i+1}) - h(t_i)|\,|W(t_i)|(h(\tau_i) - h(\sigma_i))^2/h^2(t_i) = o_p(1) \,, \tag{3.27}$$

which follows with little difficulty from Lemma 1 and the observation that $W(t_i)(h(\tau_i) - h(\sigma_i))/h^2(t_i) = o_p(1)$, which itself is trivial when $I_\varepsilon \leqq i \leqq J_\varepsilon$, while for $I_s \leqq i \leqq I_\varepsilon$,

$$\begin{aligned}W(t_i)(h(\tau_i) - h(\sigma_i))/h^2(t_i) &= O_p(1)\gamma/t_i^{\frac{1}{2}}h(t_i) \\ &= o_p(1)\gamma/h^2(t_i) \\ &= o_p(1)\gamma\lambda^{-2} = o_p(1)\end{aligned}$$

(here we have used A.4.1, (2.4), (2.13) and Assumption B). We also use

$$\begin{aligned}\gamma^{-1} \sum |h(\tau_{i+1}) - h(\tau_i) - (h(\sigma_{i+1}) - h(\sigma_i))|\,|W(t_i)| & \\ \times (|h(\tau_i) - h(\sigma_i)|/h(t_i) &= o_p(1) \,.\end{aligned} \tag{3.28}$$

To get (3.28) first restrict the summation to $I_\varepsilon \leqq i \leqq J_\varepsilon$ so that $h(t_i)$ is bounded away from 0 for such i, then let ψ be a continuous function with $\int_0^1 (h' - \psi)^2 < \varepsilon$

and proceed much as in the argument following (3.26) to obtain that the (restricted) sum is $o_p(1)$. For $I_s \leqq i \leqq I_\varepsilon$ take the second moment which is

$$O(1)\gamma^{-2} \sum_{I_s \leqq i \leqq j \leqq I_\varepsilon} |h(\tau_{i+1}) - h(\tau_i)|\,|h(\tau_{j+1}) - h(\tau_j)|\,|h(\tau_i) - h(\sigma_i)| \times \frac{|h(\tau_j) - h(\sigma_j)|t_i}{h(t_i)h(t_j)}$$

and use (A.4.1) and $\sum_{I_s}^{j} |h(\tau_{i+1}) - h(\tau_i)| = \sum_{I_s}^{j} (h(\tau_{i+1}) - h(\tau_i)) = O(1)h(t_j)$ to gain the estimate for the second moment

$$O(1)\gamma^{-1} \sum_{I_s}^{I_\varepsilon} |h(\tau_{j+1}) - h(\tau_j)|\,|h(\tau_j) - h(\sigma_j)| = o(1)\omega(\varepsilon)$$

where $\omega(\varepsilon) \to 0$ as $\varepsilon \to 0$. For $I_{s-2\gamma} \leqq i \leqq I_{s+2\gamma}$ we use A.4.1 and (2.13) to get to

$$O(1)d \sum_{I_{s-2\gamma}}^{I_{s+2\gamma}} |W(t_i)|h(t_i)/t_i^2 = o_p(1)s^{-1}d\gamma d^{-1} = o_p(1)\,.$$

These estimates imply (3.28).

We obtain

$$\gamma^{-1} \sum \Delta_i(h(\tau_i) - h(\sigma_i))^2/h(t_i) = o_p(1) \tag{3.28a}$$

(Δ_i is defined above (3.19)) by first summing (3.28a) by parts and then using (3.27) and (3.28). It is not hard to obtain from results like (3.27), (3.28), (3.28a) and arguments like those in Lemma 3 that

$$\sum h(\sigma_i)h(\sigma_{i+1})\beta_i \rho_i \left(\frac{G_{i+1}}{V_{i+1}} - \frac{G_i}{V_i}\right) = o_p(1) \tag{3.29}$$

(recall ρ_i is given in (3.10)).

Let

$$A_i = h^2(\sigma_i)h(\sigma_{i+1})\beta_i \frac{W(t_i)}{h(x_i)} \tag{3.30}$$

$$B_i = h(\sigma_i)h^2(\sigma_{i+1})\beta_i \frac{W(t_{i+1})}{h(x_{i+1})}$$

where x_i is some point in $(t_i, U(t_i))$. Then we can sum by parts and use (3.2) to obtain

$$\begin{aligned} &\sum h(\sigma_i)h(\sigma_{i+1})\beta_i \left(\frac{G_{i+1}}{V_{i+1}} - \frac{G_i}{V_i}\right) \\ &\quad = \sum (B_i - A_{i+1})\gamma^{-1}\left(1 + \frac{h(\sigma_i) - h(q_i)}{h(q_i)} + o_p(1)\gamma^{-1}n^{-\frac{1}{2}}\right) \\ &\qquad + (A_{J+1} - A_I)\gamma^{-1}(1 + o_p(1))\,. \end{aligned} \tag{3.31}$$

It is straightforward to see from (3.5) and Assumption B that

$$\begin{aligned} &\sum (B_i - A_{i+1})o_p(1)\gamma^{-2}n^{-\frac{1}{2}} \\ &\quad = o_p(1)\lambda^{-1}\gamma^{-2}n^{-\frac{1}{2}} \sum |h(\sigma_i) - h(\sigma_{i+2})|\,|h(\tau_i) - h(\sigma_i)| \\ &\qquad + o_p(1)\gamma^{-2}n^{-\frac{1}{2}} \sum |h(\tau_{i+1}) - h(\tau_i) - (h(\sigma_{i+1}) - h(\sigma_i))| \\ &\quad = o_p(1) \end{aligned} \tag{3.32a}$$

and from (3.27) and (3.28) we get

$$\sum (B_i - A_{i+1})(h(\sigma_i) - h(q_i))/h(q_i) = o_p(1) . \tag{3.32b}$$

Similar reasoning permits us to conclude that

$$\gamma^{-1} \sum (B_i' - A'_{i+1}) = o_p(1) + \gamma^{-1} \sum (B_i - A_{i+1}) \tag{3.33}$$

where A_i', B_i' are obtained from A_i, B_i be replacing $h(x_i)$ by $h(t_i)$.

We next show

$$\text{(a)}\quad \gamma^{-1}A_I = o_p(1) , \qquad \text{(b)}\quad \gamma^{-1}A_{J+1} = o_p(1) . \tag{3.34}$$

From A.4.1, we have

$$\begin{aligned} h(\sigma_I)h(\sigma_{I+1})\beta_I &= O_p(1)(h(\tau_I) - h(\sigma_I)) \\ &= O_p(1)h(\sigma_I)\gamma/\sigma_I \end{aligned}$$

and then, with the use of (2.13),

$$\gamma^{-1}A_I = O_p(1)W(t_I)h(t_I)/t_I = o_p(1)W(t_I)/t_I^{\frac{1}{2}} . \tag{3.35}$$

From $\gamma \leqq \lambda^2 = (\int_0^s h')^2 = o(1)s$, Lemma 1, (2.5), (2.6), and Assumption B, we get

$$\begin{aligned} W(t_I)/t_I^{\frac{1}{2}} &= O_p(1)\left(\frac{W(t_I) - W(s)}{s^{\frac{1}{2}}} + W(s)/s^{\frac{1}{2}}\right) \\ &= O_p(1)\gamma^{\frac{1}{2}}|\log \gamma|^{\frac{1}{2}}s^{-\frac{1}{2}} + O_p(1) \\ &= o_p(1)\gamma^{\frac{1}{2}}|\log \gamma|^{\frac{1}{2}}\lambda^{-1} + O_p(1) \\ &= O_p(1) \end{aligned}$$

which we use in (3.35) to conclude (3.34) (a). (3.34) (b) follows similarly. The effect of (3.29), (3.31), (3.32) (a) and (b), (3.33), (3.34), and Lemma 3 is to reduce the study of the asymptotic distribution of the right side of (3.16) to the study of $\gamma^{-1} \sum (B_i' - A'_{i+1})$. To prove

$$\gamma^{-1} \sum (B_i' - A'_{i+1}) \sim N(0, \textstyle\int_0^1 h'^2) \tag{3.36}$$

we first use Lemma 1 to obtain (3.37) and (3.38) below, which permits us to replace I by I_s and J by J_s in the summation in (3.36). (3.39) and (3.40) then enable us to reduce (3.36) to (3.42). The verification of (3.42) is then the final step.

Because of Lemma 1 we can restrict attention to the set where $|t_I - s| \leqq 2\gamma$, and as in the proof of (3.27) we get

$$\begin{aligned} &\gamma^{-1} \textstyle\sum_I^{I_s} (h(\sigma_i) - h(\sigma_{i+2}))h^2(\sigma_{i+1})\beta_i W(t_{i+1})/h(t_{i+1}) \\ &\qquad = O_p(1)d \textstyle\sum_{I_s-2\gamma}^{I_s+2\gamma} |W(t_{i+1})|h(t_{i+1})/t_{i+1}^2 \\ &\qquad = O_p(1)d \textstyle\sum_{I_s-2\gamma}^{I_s+2\gamma} h(t_{i+1})/t_{i+1}^{\frac{3}{2}} \\ &\qquad = o_p(1)d(4\gamma d^{-1})s^{-1} = o_p(1) . \end{aligned} \tag{3.37}$$

For similar reasons

$$
\begin{aligned}
&\gamma^{-1} \textstyle\sum_{I_s-2\gamma}^{I_s+2\gamma} h^2(\sigma_{i+1})h(\sigma_{i+2})(\beta_i - \beta_{i+1})W(t_{i+1})/h(t_{i+1}) \\
(3.38) \qquad &= O(1)\gamma^{-1} \textstyle\sum_{I_s-2\gamma}^{I_s+2\gamma} (|h(\sigma_{i+1}) - h(\sigma_i)| + |h(\tau_{i+1}) - h(\tau_i)|) \cdot |W(t_{i+1})| \\
&= O(1)d\gamma^{-1} \textstyle\sum_{I_s-2\gamma}^{I_s+2\gamma} h(\sigma_i)|W(t_{i+1})|/\sigma_i = o_p(1) .
\end{aligned}
$$

The argument producing (3.27) also produces

$$
(3.39) \qquad \gamma^{-1} \textstyle\sum_{I_s}^{J_s} |h(\sigma_{i+2}) - h(\sigma_i)|\,|\beta_i|h(\sigma_{i+1})|W(t_{i+1})| \left|\dfrac{h(\sigma_{i+1})}{h(t_{i+1})} - 1\right| = o_p(1)
$$

and (3.28) produces

$$
(3.40) \qquad \gamma^{-1} \textstyle\sum_{I_s}^{J_s} |\int_{\tau_i}^{\tau_{i+1}} h' - \int_{\sigma_i}^{\sigma_{i+1}} h'|\,|W(t_{i+1})| \left|\dfrac{h(\sigma_{i+1}) - h(t_{i+1})}{h(t_{i+1})}\right| = o_p(1) .
$$

From (3.37)–(3.40) and a summation by parts

$$
\begin{aligned}
&\gamma^{-1} \textstyle\sum (B_i' - A_{i+1}') \\
(3.41) \qquad &= o_p(1) + \textstyle\sum_{I_s}^{J_s} \gamma^{-1}(h(\sigma_i)h(\sigma_{i+1})\beta_i - h(\sigma_{i+1})h(\sigma_{i+2})\beta_{i+1})W(t_{i+1}) \\
&= o_p(1) + \textstyle\sum_{I_s}^{J_s} \gamma^{-1}h(\sigma_i)h(\sigma_{i+1})\beta_i \Delta_i \\
&= o_p(1) + \textstyle\sum_{I_s}^{J_s} \gamma^{-1}(h(\tau_i) - h(\sigma_i))\Delta_i
\end{aligned}
$$

where $\Delta_i = W(t_{i+1}) - W(t_i)$ as in the proof of Lemma 3.

The last step is to show

$$
(3.42) \qquad \textstyle\sum_{I_s}^{J_s} \gamma^{-1}(h(\tau_i) - h(\sigma_i))\Delta_i \sim N(0, \int_0^1 h'^2) .
$$

This is done quite easily if we use the representation of order statistics by exponential random variables as in Breiman [2], let S_m be the mth partial sum of independent and identically distributed exponential random variables, put $m_i = t_i(n+1)$, write

$$
\eta_{in} = \gamma^{-1}(h(\tau_i) - h(\sigma_i))(S_{m_{i+1}} - S_{m_i} - (m_{i+1} - m_i))n^{-\frac{1}{2}} ,
$$

and note that the left side of (3.42) is asymptotically the same as $\sum_{I_s}^{J_s} \eta_{in}$. Since $\{\eta_{in}\}$ are independent,

$$
E\eta_{in} = 0 , \qquad \textstyle\sum_{I_s}^{J_s} E\eta_{in}^2 = (1 + o(1))\int_0^1 h'^2 ,
$$

and the estimate $(E\eta_{in}^4)^{\frac{1}{2}} = O(1)d\gamma^{-2}(h(\tau_i) - h(\sigma_i))^2$ permits us to use the Lindeberg condition for normal convergence, we conclude (3.42) and thereby the theorem.

REFERENCES

[1] Bhattacharya, P. K. (1967). Efficient estimation of a shift parameter from grouped data. *Ann. Math. Statist.* **38** 1770–1787.

[2] Breiman, L. (1968). *Probability*. Addison-Wesley, Reading.

[3] Fabian, V. (1973). Asymptotically efficient stochastic approximation: the RM case. *Ann. Statist.* **1** 486–495.

[4] Feller, W. (1957). *An Introduction to Probability Theory and its Applications*, **1** (2nd ed.). Wiley, New York.

[5] Hájek, J. (1962). Asymptotically most powerful rank order tests. *Ann. Math. Statist.* **33** 1124–1147.

[6] HUBER, P. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.
[7] HUBER, P. (1969). *Theorie de l'inference statistique robuste.* Les Presses de l'Universite de Montreal.
[8] KIEFER, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.
[8a] SHORACK, G. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412–427.
[9] STEIN, C. (1956). Efficient non-parametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 187–196. Univ. of California Press.
[10] TAKEUCHI, K. (1971). A uniformly asymptotically efficient estimator of a location parameter. *J. Amer. Statist. Assoc.* **66** 292–301.
[11] VAN EEDEN, C. (1970). Efficiency-robust estimation of location. *Ann. Math. Statist.* **41** 172–181.
[12] WEISS, L. and WOLFOWITZ, J. (1970). Asymptotically efficient non-parametric estimators of location and scale parameters. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **16** 134–150.

DEPARTMENT OF MATHEMATICS
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60201

Added in proof. Since this paper was revised R. Beran has published "Asymptotically efficient adaptive rank estimates in location models," [*Ann. Statist.* **2** (1974) 63–74]. Beran's estimators are based on ranks and are more complicated, but are valid under minimal conditions. A recent manuscript by C. Stone "Empirical approximate maximum likelihood estimators of a location parameter" also gives an asymptotically efficient sequence of estimators which works under minimal conditions.