Sequential Analysis

ISSN: 0747-4946 (Print) 1532-4176 (Online) Journal homepage: http://www.tandfonline.com/loi/lsqa20

Recursive U-quantiles

Ola Hössjer

To cite this article: Ola Hössjer (1997) Recursive U-quantiles, Sequential Analysis, 16:1, 119-129, DOI: 10.1080/07474949708836376

To link to this article: http://dx.doi.org/10.1080/07474949708836376

Published online: 29 Mar 2007.

SEQUENTIAL ANALYSIS, 16(1), 119–129 (1997)

Recursive U-quantiles

OLA HÖSSJER
Lund University
Department of Mathematical Statistics

Key words and phrases. Asymptotic normality, incomplete designs, on-line estimator, recursive designs, U-statistics.

AMS 1991 subject classifications. Primary 62L20; secondary 62E20.

Abstract

Suppose we have a function h with m arguments and i.i.d. random variables $\{X_i\}_{i=1}^{\infty}$ with marginal distribution F. Let H_F be the distribution of $h(X_1, \ldots, X_m)$, $m \geq 2$. We consider on-line schemes for estimating quantiles of H_F. Such an estimator is based on a design D_n, which is a small subset of all $n!/(n-m)!$ possible index vectors $I = (i_1, \ldots, i_m)$ having distinct entries not exceeding n. When a new observation X_n arrives, $\gamma = |D_n \setminus D_{n-1}|$ new vectors $(X_{i_1}, \ldots, X_{i_m})$ with $I \in D_n \setminus D_{n-1}$ are used to modify the current estimate. When $\gamma \to \infty$, the asymptotic relative efficiency of the recursive estimator compared to the off-line estimator (U-quantile) tends to one. The on-line estimator is closely related to incomplete U-quantiles (Hössjer, 1996), and it generalizes a recursive quantile estimator considered by Holst (1987) for $m = 1$.

1 Introduction

Assume we have a sequence $\{X_i\}_{i=1}^{\infty}$ of $(\mathcal{X}, \mathcal{F})$-measurable random variables that are independent and identically distributed (i.i.d.) with common distribution F. Let $h : \mathcal{X}^m \to \mathbb{R}$ be a measurable function, and define another distribution function $H_F(t) = P(h(X_1, \ldots, X_m) \leq t)$, which depends on F and h. We consider estimating the quantile

$$\theta = H_F^{-1}(p) = \inf\{t;\ H_F(t) > p\},$$

given some fixed $0 < p < 1$. For each $I = (i_1, \ldots, i_m)$, introduce the short-hand notation $h(X_I) = h(X_{i_1}, \ldots, X_{i_m})$. Let also $S_n(m) = \{I = (i_1, \ldots, i_m);\ 1 \leq i_j \leq n, i_j \neq i_{j'} \text{ if } j \neq j'\}$ be the collection of all $n!/(n-m)!$ possible multi-indices I with entries not exceeding n. For any *design* $D_n \subset S_n(m)$ of multi-indices, we may define the distribution function

$$H_n(t) = \frac{1}{N(n)} \sum_{I \in D_n} 1_{h(X_I) \leq t},$$

which is an empirical analogue of H_F. Here $N(n)$ is the number of elements contained in D_n. A natural estimator of θ is

$$\tilde{\theta}_n = H_n^{-1}(p).$$

If $D_n = S_n(m)$, $\tilde{\theta}_n$ is a U-quantile (UQ). The most well known UQ is the Hodges-Lehmann estimator, which is the median of all $(X_i + X_j)/2$ in the location model (Hodges and Lehmann, 1963) The UQ based on the kernel $h(x_1, x_2) = c|x_1 - x_2|$ results in a measure of spread, with $c = c(p)$ a constant that ensures consistency if we want to estimate the standard deviation, interquartile range or some other scale functional (cf. Bickel and Lehmann, 1979, Choudhury and Serfling, 1988 and Rousseeuw and Croux, 1993). Another UQ is the Theil-Sen estimator of slope in simple linear regression (Theil, 1950 and Sen, 1968)

If $D_n \neq S_n(m)$, $\tilde{\theta}_n$ is an incomplete U-quantile (IUQ). This notion was introduced in Hössjer (1996), but an IUQ estimator was already considered by Brown and Kildea (1978) for the Hodges-Lehmann kernel. By generalizing quantiles to arbitrary L-functionals we obtain so called generalized L-statistics (Serfling, 1984) when $D_n = S_n(m)$ and incomplete generalized L-statistics (Hössjer, 1996) for general D_n.

There are several advantages of using an incomplete design D_n. Since $|S_n(m)| = O(n^m)$, the computation of $\tilde{\theta}_n$ may be intractable for large n and $m \geq 2$. On the other hand, it is possible to choose designs with $N(n) = O(n)$ and asymptotic relative efficiency (ARE) arbitrarily close to one w.r.t. the corresponding UQ. This phenomenon was first noted by Blom (1976) for incomplete U-statistics (defined as $\int x dH_n(x)$). Certain IUQ can be used for estimating the scale parameter in nonparametric regression with homoscedastic errors, and they can also be used in time series applications (Hössjer, 1996).

In this paper, we will focus on another application of incomplete designs: On-line estimation of θ. Following Hössjer (1996), we refer to a design as *recursive and on-line* (RO) if

$$D_{n-1} \subset D_n \text{ for all } n \geq 2$$

$$|D_n \setminus D_{n-1}| = O(1).$$

This means that D_n is generated from D_{n-1} by simply adding a number of multi-indices, and this number doesn't increase with n. The two designs considered here are (cf. Hössjer, 1996, Section 2)

(D1) <u>RO design based on cyclic permutations:</u> Given a positive integer $\gamma \in \mathbb{Z}^+$, define vectors $\boldsymbol{i} = (i, \ldots, i)$, $\boldsymbol{d}_1 = (d_{11}, \ldots, d_{1m}), \ldots, \boldsymbol{d}_\gamma = (d_{\gamma 1}, \ldots, d_{\gamma m})$ of length m, so that all $d_{jk} - d_{jk'}, k \neq k'$ are different, $0 \leq d_{j1} < \ldots < d_{jm}$ and $d_{1m} < \ldots < d_{\gamma m}$. Then put $D_n = \{\boldsymbol{i} + \boldsymbol{d}_j;\ 1 \leq j \leq \gamma, 1 \leq i \leq n - d_{jm}\}\}$. Examples are:

$m = 2$, $\boldsymbol{d}_k = (0, k), k = 1, \ldots, \gamma$.

$m = 3$, $\gamma = 1$ and $\boldsymbol{d}_1 = (0, 1, 3)$.

$m = 4$, $\gamma = 1$ and $\boldsymbol{d}_1 = (0, 1, 4, 6)$.

$m = 3$, $\gamma = 2$, $\boldsymbol{d}_1 = (0, 1, 3)$ and $\boldsymbol{d}_2 = (0, 4, 9)$

(D2) <u>RO design, $m = 2$:</u> $D_n = \{(i, j);\ 1 \leq i < j \leq n, j - i \leq \gamma\}$ for some $\gamma \in \mathbb{Z}^+$

In fact, both (D1) and (D2) satisfy

$$|D_n \setminus D_{n-1}| = \gamma \text{ for } n \geq \bar{m}, \tag{1.1}$$

with $\bar{m} - 1 + d_{s,m}$ for (D1) and $\bar{m} = 1 + \gamma$ for (D2). Hence, the number of added I:s remains fixed for large n. We imposed that all $d_{jk} - d_{jk'}$ are different for (D1) to ensure that estimators based on this design have a tractable asymptotic variance. A detailed account of various designs that have been used in the incomplete U-statistics literature may be found in Lee (1990, Chapter 4).

Before introducing our recursive estimator, notice that $\tilde{\theta}_n$ may be written as an M-estimator

$$\sum_{I \in D_n} \psi\left(h(X_I) - \tilde{\theta}_n\right) = 0,$$

with score function

$$\psi(x) = \begin{cases} p, & x > 0, \\ p - 1, & x \leq 0. \end{cases}$$

To define a recursive estimator of θ, let $\hat{\theta}_{\bar{m}-1}$ and $h_{\bar{m}-1}$ be fixed numbers, and put

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \frac{1}{n\gamma b_{n-1}} \sum_{I \in D_n \setminus D_{n-1}} \psi(h(X_I) - \hat{\theta}_{n-1}) \tag{1.2}$$

$$h_n = h_{n-1} + \frac{1}{n\gamma} \sum_{I \in D_n \setminus D_{n-1}} \left(n^r K\left(n^r(h(X_I) - \hat{\theta}_n)\right) - h_{n-1}\right), \tag{1.3}$$

for $n \geq \bar{m}$, with

$$b_n = [h_n]_{\rho/\log n}^{\nu \log n}.$$

Here $\nu, \rho > 0$ are fixed numbers, $[x]_a^b = \max(a, \min(x, b))$ and h_n is a recursive density estimator of $h_F(\theta)$. Finally, K is a non-negative function that integrates to one and r a fixed positive number. If $m = 1$ and $D_n = S_n(1)$, $\hat{\theta}_n$ is essentially the recursive estimator of θ considered by Holst (1987).

In Section 2, we first review some asymptotic theory for (incomplete) U-quantiles and then, in Section 3, we consider the asymptotic behaviour of $\hat{\theta}_n$. Our main result (Theorem 1) is that $\hat{\theta}_n$ is asymptotically equivalent to an IUQ based on the same design ((D1) and (D2) respectively). The (ARE) of $\hat{\theta}_n$ w.r.t. the corresponding U-quantile approaches 1 as $\gamma \to \infty$. Hence, we have found an on-line estimator of θ with negligible loss in asymptotic efficiency. Finally, the proof of Theorem 1 is given Section 4.

2 Asymptotics results for incomplete U-quantiles

Serfling (1984) considered generalized L-statistics (and in particular U-quantiles) as statistical functionals, operating on the U-process H_n. This approach was also adopted by Hössjer (1996) for incomplete generalized L-statistics. The linear, first order von Mises expansion of $\tilde{\theta}_n$ is

$$\tilde{\theta}_n = \theta + \frac{1}{N(n)} \sum_{I \in D_n} A(X_I) + R_n, \tag{2.1}$$

with $A(x_I) = \psi(h(x_I) - \theta)/h_F(\theta)$. Here R_n is a remainder term of Bahadur type. It has been analyzed by Choudhury and Serfling (1988) and Arcones (1995) for U-quantiles. The linear main term in (2.1) is an incomplete U-statistic. Asymptotic normality of $\tilde{\theta}_n$ is established using asymptotic theory of incomplete U-statistics and proving that R_n is negligible. To this end we need some notation:

Let

$$\sigma_{1ij}^2 = E\left(A(X_{I_i})A(X_{J_j})\right),$$

with I_i is a cyclic rearrangement of $(1,\ldots,m)$ with 1 in position i, and J_j is a cyclic rearrangement of $(1, m+1, \ldots, 2m-1)$ with 1 in position j. Let also

$$\sigma^2 = \sum_{i,j=1}^{m} \sigma_{1ij}^2$$

and

$$\sigma_m^2 = EA(X_I)^2.$$

The following result is a special case of Theorem 4.1 in Hössjer (1996):

Theorem 1 *Suppose $\sigma^2 > 0$ and that H_F has a positive derivative $h_F(\theta)$ at θ. Then, an IUQ based on design (D1) or (D2) has an asymptotically normal distribution,*

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} N\left(0, \sigma^2(\gamma)\right),$$

with asymptotic variance given by

$$\sigma^2(\gamma) = \sigma^2 + \frac{\sigma_m^2 - \sum_{i=1}^{m} \sigma_{1ii}^2}{\gamma}.$$

Notice that $\sigma^2(\gamma) \to \sigma^2$ as $\gamma \to \infty$, which is the asymptotic variance for U-quantiles (Serfling, 1984). By choosing γ sufficiently large, we obtain an asymptotic relative efficiency arbitrarily close to one.

If h is symmetric w.r.t. permutation of indices, the asymptotic variance simplifies to

$$\sigma^2(\gamma) = \sigma^2 + \frac{\sigma_m^2 - \sigma^2/m}{\gamma},$$

with $\sigma^2 = m^2 E\left(A(X_1, X_2, \ldots, X_m)A(X_1, X_{m+1}, \ldots, X_{2m-1})\right)$.

3 On-line estimator

Consider now the recursive estimator $\hat{\theta}_n$ defined in Section 1. We will prove below that $\hat{\theta}_n \xrightarrow{P} \theta$ and $h_n \xrightarrow{P} h_F(\theta)$. In fact, h_n is a recursive kernel density estimator of $h_F(\theta)$. Heuristically, this means

$$\hat{\theta}_n \approx \hat{\theta}_{n-1} + \frac{1}{n\gamma} \sum_{I \in D_n \setminus D_{n-1}} A(X_I).$$

In view of (2.1), this motivates why $\hat{\theta}_n$ is asymptotically equivalent to an IUQ based on the same recursive design.

We will impose the following regularity conditions:

(A) $\hat{\theta}_1, \ldots, \hat{\theta}_{m-1}, h_1, \ldots, h_{m-1}$ are arbitrary finite numbers.

(B) In some neighbourhood U of θ and for some $0 < \varepsilon_0 < 1$, $H_F' = h_F$ exists and is Hölder continuous of order ε_0, i.e. for some $L < \infty$, $|h_F(y) - h_F(x)| \leq L|x-y|^{\varepsilon_0}$ whenever $x, y \in U$.

(C) H_F is Hölder continuous of order η, $1/2 < \eta < 1$, i.e. $|H_F(y) - H_F(x)| \leq L|x-y|^{\eta}$ for all $x, y \in \Re$, with $L < \infty$.

(D) For some $\varepsilon_1 > 0$, $0 < \varepsilon_1 < r < 1/2$.

(E) The kernel function K satisfies $\int K(t)dt = 1$, has compact support, is non-negative, bounded and Lipschitz continuous, i.e. for some $L < \infty$ we have $|K(x) - K(y)| \leq L|x - y|$.

Theorem 2 *Assume a design of type (D1) or (D2), that H_F has a positive derivative at θ, and that (A)-(E) hold. Then $\hat{\theta}_n$ has an asymptotically normal distribution,*

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} N\left(0, \sigma^2(\gamma)\right),$$

with $\sigma^2(\gamma)$ as defined in Theorem 1.

4 Proof of Theorem 2

Throughout this section, C will refer to a constant whose value may change from line to line. Unless otherwise stated all convergence $\rightarrow$ means $\xrightarrow{a.s.}$, i.e. convergence almost surely. To simplify the notation, introduce $Y_n = (X_{n-\bar{m}+1}, \ldots, X_n)$, $y_n = (x_{n-\bar{m}+1}, \ldots, x_n)$ and

$$M(\theta, y_n) = \frac{1}{\gamma} \sum_{I \in D_n \setminus D_{n-1}} \psi(h(x_I) - \theta),$$

so that

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \frac{1}{nb_{n-1}} M(\hat{\theta}_{n-1}, Y_n) \tag{4.1}$$

for $n \geq \bar{m}$. Notice that $\{Y_n\}_{n \geq \bar{m}}$ is an $\bar{m}$-dependent sequence. Let also $\mathcal{F}_n$ be the σ-algebra generated by $X_1, \ldots, X_n$. With

$$\tilde{C}_k(\theta_1, \theta_2) = \mathrm{Cov}\left(M(\theta_1, Y_n), M(\theta_1, Y_{n+k})\right), \tag{4.2}$$

it may be shown that

$$\sigma^2(\gamma) = \frac{1}{h_F(\theta)^2} \sum_{k=-\bar{m}+1}^{\bar{m}-1} \tilde{C}_k(\theta, \theta) \tag{4.3}$$

This relation will be useful later on in the proof. We will start by proving a series of Lemmas. The proof of the first lemma is simple and therefore omitted. The proofs of Lemmas 2 and 3 are similar to the proofs of Theorem 3.1 and Theorem 3.2 in Holst (1987).

Lemma 1 *Assume $n \geq 2\bar{m}$ and $0 < k \leq \bar{m}$. Then*

$$|\hat{\theta}_n - \hat{\theta}_{n-k}| \leq Cn^{-1} \log n \tag{4.4}$$

and

$$\left|\frac{1}{b_n} - \frac{1}{b_{n-k}}\right| \leq Cn^{r-1} (\log n)^2 \tag{4.5}$$

Lemma 2

$$\hat{\theta}_n \rightarrow \theta \text{ as } n \rightarrow \infty.$$

Proof. Assume $n \geq 2\bar{m}$. After some manipulations, using (4.1) and $E\left(M(\hat{\theta}_{n-\bar{m}}, Y_n)|\mathcal{F}_{n-\bar{m}}\right) = p - H_F(\hat{\theta}_{n-\bar{m}})$, we get

$$\begin{aligned}\hat\theta_n - \theta &= \hat\theta_{n-1} - \theta + \frac{1}{nb_{n-1}} M(\hat\theta_{n-1}, Y_n) \\ &= \hat\theta_{n-1} - \theta + \frac{p - H_F(\hat\theta_{n-1})}{nb_{n-1}} + \frac{V_n}{nb_{n-\bar m}} + w_n + R_n,\end{aligned} \tag{4.6}$$

with

$$V_n = M(\hat\theta_{n-\bar m}, Y_n) - E\left(M(\hat\theta_{n-\bar m}, Y_n)|\mathcal{F}_{n-m}\right), \tag{4.7}$$

$$w_n = \frac{1}{nb_{n-1}}\left(M(\hat\theta_{n-1}, Y_n) - M(\hat\theta_{n-\bar m}, Y_n)\right) \tag{4.8}$$

and

$$\begin{aligned}R_n &= \frac{H_F(\hat\theta_{n-1}) - H_F(\hat\theta_{n-\bar m})}{nb_{n-1}} \\ &+ \left(\frac{1}{b_{n-1}} - \frac{1}{b_{n-\bar m}}\right)\frac{1}{n}\left(M(\hat\theta_{n-\bar m}, Y_n) - E\left(M(\hat\theta_{n-\bar m}, Y_n)|\mathcal{F}_{n-\bar m}\right)\right).\end{aligned} \tag{4.9}$$

By Lemma 1 and (C), $|R_n| \leq Cn^{-\zeta}$ for some $\zeta > 3/2$, so

$$\sum_{k=2\bar m}^{\infty} |R_k| < \infty. \tag{4.10}$$

Since V_n is adapted to $\mathcal{F}_n$ and $E(V_n|\mathcal{F}_{n-\bar m}) = 0$, $\{V_n\}_{2\bar m}^{\infty}$ is a uniformly bounded sequence of mixingale differences. By McLeish (1975, Corollary (1.8)),

$$\sum_{k=2\bar m}^{n} \frac{V_k}{kb_{k-\bar m}} \text{converges,} \tag{4.11}$$

since $\sum_k k^{-2}(\log k)^2 < \infty$. By Lemma 1,

$$\begin{cases} |w_n| \leq Cn^{-1}\log n \\ P(w_n \neq 0) \leq Cn^{-\eta}(\log n)^{\eta} \\ P(w_n \neq 0, w_l \neq 0) \leq Cn^{-\eta}(\log n)^{\eta} l^{-\eta}(\log l)^{\eta} \text{ for } |n-l| \geq \bar m. \end{cases} \tag{4.12}$$

Actually, (4.12) implies that

$$S_n = \sum_{k=2\bar m}^{n} w_k \text{ converges.} \tag{4.13}$$

Put $S_n = S_n^+ - S_n^- = \sum_{2\bar m}^n w_k^+ - \sum_{2\bar m}^n w_k^-$, where $w_k^+ = 0 \vee w_k$ and $w_k^- = 0 \vee (-w_k)$. We will show that

$$S_n^+ = \sum_{k=2\bar m}^{n} w_k^+ \text{ converges.} \tag{4.14}$$

The convergence of S_n^- is handled in the same way. First note that $E(w_k^+) \leq Ck^{-1-\eta}(\log k)^2$ because of (4.12), so

$$\sum_{k=2\bar m}^{\infty} E(w_k^+) < \infty. \tag{4.15}$$

After some calculation, it also follows from (4.12), if $l > n$, that

$$\mathrm{Var}(S_l^+ - S_n^+) \leq Cn^{-2\eta+\epsilon}$$

for any $\varepsilon > 0$. In connection with Chebyshev's inequality this gives

$$\sum_{k:2^k \geq 2\bar m} P\left(|S_{2^{k+1}}^+ - S_{2^k}^+| \geq \varepsilon k^{-2}\right) < \infty \tag{4.16}$$

for any $\varepsilon > 0$, so $S^+_{2^k}$ converges. Since S^+_n is a non-decreasing sequence, (4.14) follows. Put now $\delta_n = R_n + V_n/(nb_{n-\bar{m}}) + w_n$. Then, by (4.10), (4.11) and (4.13),

$$\sum_{k=2\bar{m}}^{n} \delta_k \text{ converges.} \tag{4.17}$$

Choose now $\alpha_n \to 0$ s.t.

$$\sum_{n=2\bar{m}}^{\infty} \alpha_n/(nb_{n-1}) = \infty. \tag{4.18}$$

This is possible since $b_{n-1} \leq \nu \log(n-1)$. Define $\beta_n = C\alpha_n$, with C so large that $|x-\theta| \geq \beta_n$ implies $|p - H_F(x)| \geq \alpha_n$ for all but finitely many n. Then

$$\begin{cases} \hat{\theta}_n - \theta \geq \beta_n \Longrightarrow \hat{\theta}_n - \theta \leq \hat{\theta}_{n-1} - \theta - \frac{\alpha_n}{nb_{n-1}} + \delta_n \\ \hat{\theta}_n - \theta \leq -\beta_n \Longrightarrow \hat{\theta}_n - \theta \geq \hat{\theta}_{n-1} - \theta + \frac{\alpha_n}{nb_{n-1}} + \delta_n. \end{cases}$$

Also, find $\gamma_n \to 0$ s.t. $|x-\theta| \leq \beta_n$ implies $|x - \theta + (p - H_F(x))/(nb_{n-1}) + \delta_n| \leq \gamma_n$. Then, for large enough n,

$$\begin{cases} (\hat{\theta}_n - \theta)_+ \leq \max\left(\gamma_n, (\hat{\theta}_{n-1} - \theta)_+ - \frac{\alpha_n}{nb_{n-1}} + \delta_n\right) \\ (\hat{\theta}_n - \theta)_- \leq \max\left(\gamma_n, (\hat{\theta}_{n-1} - \theta)_- - \frac{\alpha_n}{nb_{n-1}} - \delta_n\right). \end{cases}$$

The lemma now follows from (4.17), (4.18) and Lemma 1 in Derman and Sacks (1959). □

Lemma 3

$$h_n \to h_F(\theta) \text{ as } n \to \infty.$$

Proof. Let, for $x \in U$ (cf. (B)),

$$\bar{h}_n(x) = n^r \int K(n^r(y-x))\, dH_F(y)$$

and

$$v_n = \frac{2\bar{m}-1}{n} h_{2\bar{m}-1} + \frac{1}{n} \sum_{k=2\bar{m}}^{n} \bar{h}_n(\hat{\theta}_{k-m}).$$

Conditions (B), (D) and (E) imply $\lim \bar{h}_n(x) = h(\theta)$ as $n \to \infty$ and $x \to \theta$. Hence, by Lemma 2,

$$v_n \to h_F(\theta) \tag{4.19}$$

Now

$$h_n - v_n = \frac{1}{n} \sum_{k=2\bar{m}}^{n} U_k + \frac{1}{n} \sum_{k=2\bar{m}}^{n} R_k,$$

with

$$U_k = \frac{1}{\gamma} \sum_{I \in D_k \setminus D_{k-1}} \left(k^r K\left(k^r(h_F(X_I) - \hat{\theta}_{k-\bar{m}})\right) - E\left(k^r K\left(k^r(h_F(X_I) - \hat{\theta}_{k-\bar{m}})\right) | \mathcal{F}_{k-\bar{m}}\right)\right)$$

and

$$R_k = \frac{1}{\gamma} \sum_{I \in D_k \setminus D_{k-1}} k^r \left(K\left(k^r(h_F(x_I) - \hat{\theta}_{k-1})\right) - K\left(k^r(h_F(x_I) - \hat{\theta}_{k-\bar{m}})\right)\right).$$

Observe that $E(U_k|\mathcal{F}_{k-\bar{m}}) = 0$, so $\{U_k\}$ are mixingale differences. McLeish (1975, Corollary (1.8)) and Kronecker's Lemma gives

$$\frac{1}{n} \sum^{n} U_k \to 0,$$

using the fact that $r < 1/2$ Finally, (E) and (4.4) imply $|R_k| \leq Ck^{2r-1}\log k$, which results in

$$\frac{1}{n}\sum_{k=2\bar{m}}^{n} R_k \to 0.$$

□

Lemma 4 *For any $\delta < 1/2$,*

$$n^{\delta}(\hat{\theta}_n - \theta) \to 0.$$

Proof. In view of (B) and (4.6),

$$\begin{aligned} n^{\delta}(\hat{\theta}_n - \theta) &= (n-1)^{\delta}\left((\hat{\theta}_{n-1} - \theta)(1 - \frac{h_F(\theta)}{nb_{n-1}} + \frac{\delta}{n} + o(n^{-1})\right) \\ &+ n^{\delta}(R_n + \frac{V_n}{nb_{n-m}} + w_n), \end{aligned}$$

with R_n, V_n and w_n as in Lemma 2. It follows as in Lemma 2 that the three sums $\sum_{2\bar{m}}^{n} k^{\delta}R_k$, $\sum_{2\bar{m}}^{n} k^{\delta-1}V_k/b_{k-\bar{m}}$ and $\sum_{2\bar{m}}^{n} k^{\delta}w_k$ converge as $n \to \infty$. (For the second sum, use McLeish (1975, Corollary (1.8)) since $\sum_{2\bar{m}}^{\infty} k^{2\delta-2}(\log k)^2 < \infty$.) The lemma now follows from Lemma 1 in Venter (1967) and the fact that

$$\liminf_{n\to\infty}(\frac{h_F(\theta)}{b_{n-1}} - \delta) > 0$$

by Lemma 3. □

Lemma 5 *For some $\varepsilon_2 > 0$,*

$$n^{\varepsilon_2}(h_n - h_F(\theta)) \to 0.$$

Proof. We will see below that the choice

$$\varepsilon_2 < \min\left(\varepsilon_0\varepsilon_1, \varepsilon_0\delta, \frac{1}{2} - r\right)$$

will do, where δ is any admissible number in Lemma 4. Choose now $K_0 > 0$ so that $\mathrm{supp}(K) \subset [-K_0, K_0]$. Then, if $[x \pm K_0 n^{-r}] \subset U$, if follows from Assumption (B) that

$$|\bar{h}_n(x) - h_F(\theta)| \leq C\left(|x-\theta|^{\varepsilon_0} + n^{-r\varepsilon_0}\right).$$

Hence, by (D) and Lemma 4,

$$n^{\varepsilon_2}(v_n - h_F(\theta)) \to 0.$$

Next, by Kronecker's Lemma,

$$n^{\varepsilon_2}(h_n - v_n) \to 0,$$

provided $(\sum_{k=2\bar{m}}^{n} U_k)/n^{\varepsilon_2 - 1}$ and $(\sum_{k=2\bar{m}}^{n} R_k)/n^{\varepsilon_2-1}$ converge. This follows as in the proof of Lemma 3, since $\sum_{2\bar{m}}^{\infty} k^{2r+2\varepsilon_2-2} < \infty$ and

$$\frac{1}{n^{1-\varepsilon_2}}\sum_{k=2\bar{m}}^{n} |R_k| \leq Cn^{2r-1+\varepsilon_2}\log n \to 0.$$

□

Proof of Theorem 2 Define the sequence $\{\tilde{\theta}_n\}_{n=2\bar{m}-1}^{\infty}$ through $\tilde{\theta}_{2\bar{m}-1} = \hat{\theta}_{2\bar{m}-1}$ and

$$\tilde{\theta}_n - \theta = (\tilde{\theta}_{n-1} - \theta)(1 - \frac{1}{n}) + \frac{V_n}{nb'_{n-\bar{m}}}, \quad n \geq 2\bar{m}$$

with $b'_n = \max(\rho, b_n)$. We will first show that $\tilde{\theta}_n$ is asymptotically equivalent to $\hat{\theta}_n$, that is

$$\delta_n \to 0, \tag{4.20}$$

with $\delta_n = \sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n)$. Observe that $\{\delta_n\}_{2\bar{m}}^{\infty}$ satisfy the recursion

$$\delta_n = \sqrt{1 - \frac{1}{n}}\delta_{n-1} + \sqrt{n}R_n + \sqrt{n}w_n + \left(\frac{1}{b_{n-\bar{m}}} - \frac{1}{b'_{n-\bar{m}}}\right)\frac{V_n}{n}, + \frac{1 - A_n}{\sqrt{n}}(\hat{\theta}_{n-1} - \theta)$$

with

$$A_n = \frac{H(\hat{\theta}_{n-1}) - p}{b_{n-1}(\hat{\theta}_{n-1} - \theta)}.$$

As in the proof of Lemma 2, one shows that

$$\sum_{k=2\bar{m}}^{n} \sqrt{k}w_k \text{ converges} \tag{4.21}$$

and

$$\sum_{k=2\bar{m}}^{\infty} \sqrt{k}|R_k| < \infty. \tag{4.22}$$

By Lemma 3, $b_n = b'_n$ for all but finitely many n. Hence,

$$\sum_{k=2\bar{m}}^{n} \left(\frac{1}{b_{k-\bar{m}}} - \frac{1}{b'_{k-\bar{m}}}\right)\frac{V_k}{k} \text{ converges.} \tag{4.23}$$

Lemma 2-3 and Condition (B) imply

$$|A_n - 1| \leq C\left(|b_{n-1} - h_F(\theta)| + |\hat{\theta}_{n-1} - \theta|^{\epsilon_0}\right).$$

According to Lemma 4-5 this yields $n^{\zeta}(1 - A_n)(\hat{\theta}_n - \theta) \to 0$ for some $\zeta > 1/2$. Therefore,

$$\sum_{k=2\bar{m}}^{\infty} \frac{\left|(1 - A_k)(\hat{\theta}_k - \theta)\right|}{\sqrt{k}} < \infty. \tag{4.24}$$

Now (4.21)-(4.24) and Lemma 1 of Venter (1967) imply (4.20). By Slutsky's Lemma, it remains to prove asymptotic normality of $\tilde{\theta}_n$. Observe that

$$\begin{aligned}\tilde{\theta}_n - \theta &= \frac{2\bar{m} - 1}{n}(\hat{\theta}_{2\bar{m}-1} - \theta) + \frac{1}{n}\sum_{k=2\bar{m}}^{n}\frac{V_k}{b'_{k-\bar{m}}} \\ &= O(n^{-1}) + \frac{1}{n}\sum_{k=2\bar{m}}^{n}\frac{V_k}{b'_{k-\bar{m}}}.\end{aligned}$$

Hence, it suffices to prove that

$$\frac{1}{\sqrt{n}}\sum_{k=2\bar{m}}^{n}\frac{V_k}{b'_{k-\bar{m}}} \xrightarrow{\mathcal{L}} N(0, \sigma^2(\gamma)). \tag{4.25}$$

Actually, (4.25) follows from a Central Limit Theorem for mixingales in McLeish (1977), with $X_{ni} = V_i/(\sqrt{n}b'_{i-m}\sigma(\gamma))$ if $2\bar{m} \leq i \leq n$, $X_{ni} = 0$ if $1 \leq i \leq 2\bar{m} - 1$, $k_n(t) = [nt]$, $\sigma^2_{n,i} = 1/n$, $\mathcal{F}_{n,i} = \mathcal{F}_n$, $\psi_k = 4\|M\|_\infty/(\rho\sigma(\gamma))$ for $0 \leq k \leq \bar{m}$ and $\psi_k = 0$ for $k \geq \bar{m}$ (cf (4.2), $\|M\|_\infty = \sup_{x,y}|M(x,y)|$). Notice that $\{X_{n,i}/\sigma_{n,i}\}$ are uniformly bounded in n and i, because of the choice of $\{b'_n\}_{2\bar{m}}^\infty$. Conditions (2.2)-(2.5) in McLeish (1977) are easily checked. It remains to check (2.6), which requires that for any $s < t < u \leq 1$

$$\left\| E\left((\sum_{i=k_n(t)}^{k_n(u)} X_{ni})^2 | \mathcal{F}_{k_n(s)} \right) - (u-t) \right\|_1 \to 0 \quad \text{as } n \to \infty. \tag{4.26}$$

where $\|\cdot\|_1$ denotes the L_1-norm. Assume that n is so large that $k_n(t) \geq 3\bar{m}$ and $k_n(t) - k_n(s) \geq 2\bar{m}$. In view of (4.3),

$$\begin{aligned} &\left\| E\left((\sum_{i=k_n(t)}^{k_n(u)} X_{ni})^2 | \mathcal{F}_{k_n(s)} \right) - (u-t) \right\|_1 \\ &= \left\| \sum_{i,j=k_n(t)}^{k_n(u)} \left(\mathrm{Cov}(X_{ni}|\mathcal{F}_{k_n(s)}, X_{nj}|\mathcal{F}_{k_n(s)}) - \tfrac{\tilde{C}_{i-j}(\theta,\theta)}{nh_F(\theta)^2\sigma^2(\lambda)} \right) \right\|_1 + O(\tfrac{1}{n}) \\ &\leq \sum_{\substack{i,j=k_n(t) \\ |i-j|\leq\bar{m}}}^{k_n(u)} \left\| \mathrm{Cov}(X_{ni}|\mathcal{F}_{k_n(s)}, X_{nj}|\mathcal{F}_{k_n(s)}) - \tfrac{\tilde{C}_{i-j}(\theta,\theta)}{nh_F(\theta)^2\sigma^2(\lambda)} \right\|_1 + O(\tfrac{1}{n}), \end{aligned} \tag{4.27}$$

since $\tilde{C}_k(\theta,\theta) = 0$ for $|k| \geq \bar{m}$ and $\mathrm{Cov}(X_{ni}|\mathcal{F}_{k_n(s)}, X_{nj}|\mathcal{F}_{k_n(s)}) = 0$ for $|i-j| \geq \bar{m}$. To proceed further from (4.27), we will show below that

$$\mathrm{Cov}(X_{ni}|\mathcal{F}_{k_n(s)}, X_{nj}|\mathcal{F}_{k_n(s)}) = E\left(\frac{\tilde{C}_{i-j}\left(\hat{\theta}_{i-2\bar{m}}, \hat{\theta}_{j-2\bar{m}}\right)}{n\sigma^2(\gamma)b'_{i-2\bar{m}}b'_{j-2\bar{m}}} | \mathcal{F}_{k_n(s)} \right) + n^{-1}O(i^{-\varepsilon} + j^{-\varepsilon}), \tag{4.28}$$

for some $\varepsilon > 0$, and with the O-term holding uniformly for $k_n(t) \leq i, j \leq k_n(u), |i-j| < \bar{m}$. Now $(x_1, x_2, x_3, x_4) \to \tilde{C}_k(x_3, x_4)/(\max(x_1, \rho)\max(x_2, \rho))$ is a continuous and bounded function for $|k| < \bar{m}$. This is because $\rho > 0$ and since Condition (C) implies continuity of the $\tilde{C}_k$-factor. Putting $(x_1, x_2, x_3, x_4) = (b_{i-2\bar{m}}, b_{j-2\bar{m}}, \hat{\theta}_{i-2\bar{m}}, \hat{\theta}_{j-2\bar{m}})$ implies, via Lemma 2 3 and (4.28), that the RHS of (4.27) tends to zero as $n \to \infty$. Finally, (4.28) is deduced by introducing

$$X'_{ni} = \frac{V'_i}{\sqrt{n}b'_{n-2\bar{m}}\sigma(\gamma)}$$

if $2\bar{m} \leq i \leq n$, with

$$V'_i = M(\hat{\theta}_{i-2\bar{m}}, Y_n) - E\left(M(\hat{\theta}_{i-2\bar{m}}|\mathcal{F}_{i-2\bar{m}})\right).$$

Then, because the $\bar{m}$-depencence of the sequence $\{Y_n\}$, and since $k_n(s) \leq \min(i,j) - 2\bar{m}$,

$$\mathrm{Cov}(X'_{ni}|\mathcal{F}_{k_n(s)}, X'_{nj}|\mathcal{F}_{k_n(s)}) = E\left(\frac{\tilde{C}_{i-j}\left(\hat{\theta}_{i-2\bar{m}}, \hat{\theta}_{j-2\bar{m}}\right)}{n\sigma^2(\gamma)b'_{i-2\bar{m}}b'_{j-2\bar{m}}} | \mathcal{F}_{k_n(s)} \right)$$

when $|i-j| < \bar{m}$. Finally, the proof is completed by noting that

$$|\mathrm{Cov}(X_{ni}|\mathcal{F}_{k_n(s)}, X_{nj}|\mathcal{F}_{k_n(s)}) - \mathrm{Cov}(X'_{ni}|\mathcal{F}_{k_n(s)}, X'_{nj}|\mathcal{F}_{k_n(s)})| \leq Cn^{-1}(i^{-\varepsilon} + j^{-\varepsilon})$$

for any $\varepsilon < \min(1-r, \eta)$, which follows from Lemma 1 and Condition (C), using estimates similar to (4.12). □

Acknowledgement

References

Arcones, M.A. (1995). The Bahadur Kiefer representation for U-quantiles. Manuscript

Bickel, P.J. and Lehmann, E.L. (1979). Decriptive statistics for nonparametric models. IV Spread In *Contributions to Statistics. Hajek Memorial Volume*, ed. J Jurečková, Academia, Prague, 33-40.

Blom, G. (1976). Some properties of incomplete U-statistics. *Biometrika* **63**, 573-580.

Brown, B.M. and Kildea, D.G. (1978). Reduced U-statistics and the Hodges-Lehmann estimator. *Ann. Statist.* **6**, 828-835.

Choudhury, J. and Serfling, R.J. (1988). Generalized order statistics, Bahadur representations, and sequential nonparametric fixed-width confidence intervals, *J. Statist. Planning and Inference* **19**, 269-282.

Derman, C. and Sacks, J. (1959). On Dvoretzky's stochastic approximation theorem, *Ann. Math. Statist.* **30**, 601-605.

Hodges, J.L., Jr. and Lehmann, E. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598-611.

Holst, U. (1987). Recursive estimation of quantiles using recursive kernel density estimators, *Seq. Anal.* **6**, 219-237.

Hössjer, O. (1996). Incomplete generalized L-statistics. To appear in *Annals of Statistics.*

Lee, A.J. (1990). *U-statistics, Theory and Practice.* Statistics, Textbooks and Monographs, Vol. 110: Marcel Dekker, Inc. New York and Basel.

McLeish, D.L. (1975). A maximal inequality and dependent strong laws, *Ann. Probab.* **3**, 829-839.

McLeish, D.L. (1977). On the invariance principle for nonstationary martingales, *Ann. Probab.* **5**, 616-621.

Rousseeuw, P.J. and Croux, C. (1993). Alternatives to the median absolute deviation, *J. Amer. Statist. Assoc.* **88**, 1273-1283.

Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau, *J. Amer. Statist. Assoc.* **63**, 1379-1389.

Serfling, R.J. (1984). Generalized L-, M- and R-statistics. *Ann. Statist.* **12**, 76-86.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis, I, II and III, *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings* **53**, 386-392, 521-525, 1397-1412.

Venter, J.H. (1967). An extension of the Robbins-Monro procedure, *Ann. Math. Statist.* **38**, 181-190.

Received January 1996; Revised November 1996