The Annals of Statistics
1990, Vol. 18, No. 3, 1188–1202

A BAHADUR-TYPE REPRESENTATION FOR EMPIRICAL QUANTILES OF A LARGE CLASS OF STATIONARY, POSSIBLY INFINITE-VARIANCE, LINEAR PROCESSES

BY C. H. HESSE

University of California, Berkeley

Bahadur has obtained an asymptotic almost sure representation for empirical quantiles of independent and identically distributed random variables. In this paper we present an analogous result for a large class of stationary linear processes.

0. Introduction. Bahadur (1966) has initiated the asymptotic representation theory of sample quantiles via the empirical distribution function. In particular, he demonstrated that under certain fairly mild regularity conditions on the distribution F and the density f of the iid sequence $X(1), X(2), \ldots$ the following is true with probability 1:

$$X_{p,n} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + R_n$$

and

$$R_n = \mathbf{O}\big(n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}\big).$$

Here, for $0 < p < 1$, ξ_p is the unique p quantile of F, i.e., $F(\xi_p) = p$, $X_{p,n}$ is the pth sample quantile based on $X(1), X(2), \ldots, X(n)$ and F_n is the empirical distribution function based on the same sample.

Bahadur's theorem and proof give great insight into the relation between empirical quantiles and the empirical distribution function. It has triggered a number of refined studies in the iid case and subsequent extension to nonindependent sequences: Analysis by Eicker (1966) has revealed that the remainder R_n is $o_p(n^{-3/4}g(n))$ if and only if $g(n) \to \infty$ as $n \to \infty$ and Kiefer (1967) gave the definite answer

$$\text{w.p. 1} \quad \limsup_{n\to\infty} \pm \frac{n^{3/4}R_n}{(\log\log n)^{3/4}} = \frac{2^{5/4}(p(1-p))^{1/4}}{3^{3/4}}$$

for either choice of sign.

Other references in the iid case include Duttweiler (1973) and Ghosh (1971), who obtained a simpler proof of Bahadur's representation but for a weaker result.

Received June 1988; revised June 1989.
AMS 1980 *subject classifications*. Primary 62G30; secondary 60F15, 60G10.
Key words and phrases. Stationary linear processes, quantiles, Bahadur representation, almost sure convergence.

There are some extensions to sequences of random variables with certain dependency structures, e.g., m-dependence, ϕ-mixingness and strong mixingness; compare Sen (1968, 1972) and Babu and Singh (1978). In this paper we obtain an analogous strong representation for a very broad class of stationary linear processes with parameters decreasing at a polynomial rate. In particular, the sequences considered are

$$X(n) = \sum_{i=0}^{\infty} \delta(i)\varepsilon(n-i), \tag{0.1}$$

where $\varepsilon(n)$ are iid innovations with $E(|\varepsilon(n)|^{\alpha}) < \infty$ for some $\alpha > 0$ and $|\delta(i)| \leq c \cdot i^{-q}$ for some $c, q > 0$ and $i \geq 1$.

The class of linear processes in (0.1) is very broad. It includes both finite and infinite variance linear processes and also incorporates processes based on both continuous and certain (due to restrictions on the stationary distribution function that will be imposed later) discrete innovation series $\varepsilon(n)$. It covers m-dependent sequences, all autoregressive-moving average processes and certain sequences which are neither ϕ mixing nor strong mixing. Examples of sequences within the class (0.1) which are not strong mixing are easily obtained: The first order autoregressive process

$$X(n) - \tfrac{1}{2}X(n-1) = \varepsilon(n) \tag{0.2}$$

is strongly mixing iff $\varepsilon(n)$ has a distribution with absolutely integrable characteristic function, such as the normal distribution; see Chanda (1974). If for example the $\varepsilon(n)$ are iid symmetric Bernoulli, $X(n)$ is not strongly mixing; compare also Andrews (1984). However, (0.2) is in the class (0.1) (even has an absolutely continuous stationary distribution function), which is easily demonstrated by obtaining the stationary solution of the difference equation, namely,

$$X(n) = \sum_{i=0}^{\infty} 2^{-i}\varepsilon(n-i).$$

1. Statement of results. Theorem 1 in this section is the main contribution of the paper. It gives a Bahadur-type result for empirical quantiles of the broad class of stationary processes introduced in (0.1) with parameters $\delta(i)$ decreasing to zero in absolute value at a polynomial rate. For $0 < p < 1$, what is here and below meant by pth sample quantile $X_{p,n}$ of a sequence of random variables $X(1), \dots, X(n)$ is the $\lceil n \cdot p \rceil$th order statistic, where $\lceil x \rceil$ denotes the smallest integer larger or equal to x.

As before, F_n and F denote the empirical distribution and the stationary distribution of X, respectively, ξ_p is such that $F(\xi_p) = p$ and c denotes a generic positive constant, not always the same one. Other notation will be introduced as needed.

THEOREM 1. *Let*

$$X(n) = \sum_{i=0}^{\infty} \delta(i)\varepsilon(n-i), \tag{1.1}$$

where the innovations $\varepsilon(n)$ *are* iid *with* $E(|\varepsilon(n)|^{\alpha}) < \infty$ *for some* $\alpha > 0$. *Assume also that for* $i \geq 1$, $|\delta(i)| \leq c \cdot i^{-q}$ *with* $q > 1 + 2/\alpha$ *and that the density* f *of the stationary distribution of* X *is bounded away from* 0 *and* ∞ *in a neighborhood* B_p *of* ξ_p. *Then*

$$X_{p,n} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + R_n \quad a.s.,$$

where $R_n = \mathbf{O}(n^{-3/4+\gamma})$ *for all* $\gamma > (\alpha^2(8q-5) + 2\alpha(10q-9) - 13)/(4(2\alpha q - \alpha - 1)^2)$.

REMARK 1. The lower bound for γ in Theorem 1 is decreasing both for increasing α and for increasing q.

2. Proof of main result. The proof is based on extensions of Lemmas 1–3 in Bahadur (1966) to the present context. Two of these extensions are straightforward while one (our Theorem 2) is somewhat involved.

THEOREM 2. *Under the conditions of Theorem* 1 *let, for given* n, $0 < \beta < 1$ *and* $i = 1, \dots, n$,

$$X_\beta(i) = \sum_{j=0}^{\lceil n^\beta \rceil - 1} \delta(j)\varepsilon(i-j).$$

Then it holds true for the empirical pth *quantile* $(X_\beta)_{p,n}$ *of* $X_\beta(i)$, $i = 1, \dots, n$, *that for all* n *sufficiently large*,

$$(X_\beta)_{p,n} \in I_n = (\xi_p - a_n, \xi_n + a_n) \quad a.s.$$

with

$$a_n = \mathbf{O}\Big(\max\big\{n^{1/2(\beta-1)}(\log n)^{1/2}, n^{\alpha/(\alpha+1)(\beta(1+1/\alpha-q)+1/\alpha)}\big\}\Big).$$

REMARK 2. The parameter β determines the order of truncation of the infinite linear combination of innovations in (1.1). To obtain the strongest result in Theorem 1, an optimal choice for β will have to be made later.

PROOF OF THEOREM 2. For later use we first evaluate the difference between $X(i)$ and $X_\beta(i)$. Clearly,

$$\big|X(i) - X_\beta(i)\big| \leq c \sum_{j=\lceil n^\beta \rceil}^{\infty} j^{-q}|\varepsilon(i-j)| \leq c\varepsilon^* n^{1/\alpha} \sum_{j=\lceil n^\beta \rceil}^{\infty} j^{-q+1/\alpha},$$

where $\varepsilon^* = \sup\{|\varepsilon(0)|, \sup_{|k|\geq 1}(|\varepsilon(k)|/|k|^{1/\alpha})\}$ is almost sure finite due to the conditions on $\varepsilon(k)$ and moreover $P(\varepsilon^* \geq \xi) \leq c\xi^{-\alpha}$ for all $\xi > 0$, as is easily proved. Hence, uniformly in i (up to n),

$$|X(i) - X_\beta(i)| \le c\varepsilon^* n^{\beta(1+1/\alpha-q)+1/\alpha}. \tag{2.1}$$

We write $\eta = -\beta(1 + 1/\alpha - q) - 1/\alpha$. Equation (2.1) implies that the difference between the pth sample quantiles based on $X_\beta(i)$ and $X(i)$, $i = 1,\dots,n$, respectively, is

$$|(X_\beta)_{p,n} - X_{p,n}| = \mathbf{O}(n^{-\eta}). \tag{2.2}$$

Here and below order relations are to be interpreted to hold almost surely.

Then, using Lemma 1 from the Appendix, we get

$$\sup_{y\in B_p} |F(y) - F_\beta(y)| = \mathbf{O}(n^{-\alpha/(\alpha+1)\eta}), \tag{2.3a}$$

and

$$\sup_{y\in B_p} |F(y) - F_\beta(y\,-)| = \mathbf{O}(n^{-\alpha/(\alpha+1)\eta}), \tag{2.3b}$$

where F_β is the distribution of X_β and $F_\beta(y\,-)$ denotes the limit from the left, i.e., $F_\beta(y\,-) = \lim_{y_0 \uparrow y} F_\beta(y_0)$.

We now exploit the independence of the truncated series after lag $\lceil n^\beta \rceil$ by defining

$$\mathbf{S}^\beta_{n,k} = \{X_\beta(k), X_\beta(k + \lceil n^\beta \rceil), \dots, X_\beta(k + (n_k - 1)\lceil n^\beta \rceil)\},$$

$$k = 1, 2, \dots, \lceil n^\beta \rceil,$$

where n_k is either $\lceil n^{1-\beta} \rceil$ or $\lceil n^{1-\beta} \rceil - 1$, its dependence on k being of no concern. If for given n, X^k_{p,n_k} denotes the pth sample quantile of the kth set $S^\beta_{n,k}$ containing n_k random variables, then by Lemma 2,

$$\min_{1\le k\le\lceil n^\beta\rceil} X^k_{p,n_k} \le (X_\beta)_{p,n} \le \max_{1\le k\le\lceil n^\beta\rceil} X^k_{p,n_k}. \tag{2.4}$$

At this point it is necessary to point out one of the defects of the distribution F_β: its possible discontinuity. We remedy this by introducing the slightly perturbed but continuous random variables

$$G_\beta(X_\beta(i)) = U(i) F_\beta(X_\beta(i)\,-) + (1 - U(i)) F_\beta(X_\beta(i)),$$

where $U(i)$ has the uniform distribution over $(0, 1)$ and is independent of $X_\beta(i)$. If we also define $G_\beta(y)$ as

$$G_\beta(y) = U(i) F_\beta(y\,-) + (1 - U(i)) F_\beta(y),$$

then, using (2.3a) and (2.3b), we get

$$\sup_{y\in B_p} |F_\beta(y) - F_\beta(y\,-)| \le \sup_{y\in B_p} |F_\beta(y) - F(y)| + \sup_{y\in B_p} |F(y) - F_\beta(y\,-)|$$

$$= \mathbf{O}(n^{-\alpha/(\alpha+1)\eta})$$

and hence

$$\sup_{y \in B_p} \left| F_\beta(y) - G_\beta(y) \right| = \mathbf{O}\left(n^{-\alpha/(\alpha+1)\eta}\right). \tag{2.5}$$

Since, in particular, for all n sufficiently large, $(X_\beta)_{p,n} \in B_p$ a.s., by Lemma 5, (2.5) therefore implies that

$$\left| F_\beta\left((X_\beta)_{p,n}\right) - G_\beta\left((X_\beta)_{p,n}\right) \right| = \mathbf{O}\left(n^{-\alpha/(\alpha+1)\eta}\right),$$

from which we deduce

$$\left(F_\beta(X_\beta)\right)_{p,n} = \left(G_\beta(X_\beta)\right)_{p,n} + \mathbf{O}\left(n^{-\alpha/(\alpha+1)\eta}\right), \tag{2.6}$$

by monotonicity of F_β and G_β. In (2.6), $(F_\beta(X_\beta))_{p,n}$ is the $\lceil n \cdot p \rceil$th order statistic of $F_\beta(X_\beta(i))$, $i = 1, \dots, n$, and $(G_\beta(X_\beta))_{p,n}$ is defined similarly. Equation (2.6) demonstrates that the effect introduced by the perturbation with respect to the corresponding pth quantiles may be ignored. Keeping in mind (2.6) we are in the sequel concerned with $G_\beta(X_\beta(i))$, $i = 1, \dots, n$ only.

We will first determine how close $(-\log(G_\beta(X_\beta)))_{p,n}$, the pth quantile of $-\log G_\beta(X_\beta(i))$, $i = 1, \dots, n$, is to $\log p^{-1}$. Since $X_\beta(i)$ has distribution F_β, $-\log G_\beta(X_\beta(i))$ has an exponential distribution with mean 1. On each set $S^\beta_{n,k}$, we may therefore apply the Renyi representation [see, e.g. Shorack and Wellner (1986), page 723] to these transformed random variables. In particular, for the $\lceil n_k \cdot p \rceil$th order statistic $(-\log G_\beta(X_\beta))^k_{p,n_k}$ of the transformed random variables in $S^\beta_{n,k}$, we obtain

$$\left(-\log G_\beta(X_\beta)\right)^k_{p,n_k} = \sum_{v=\lceil p \cdot n_k \rceil}^{n_k} \frac{E_{n,k,v}}{v}, \tag{2.7}$$

where for each n, k, the $E_{n,k,v}$ for different v are independent random variables with exponential distribution (with mean 1). This representation is used to establish that

$$\limsup_{n\to\infty} \max_{1 \le k \le \lceil n^\beta \rceil} \left(\frac{n^{1-\beta}}{\log n} \right)^{1/2} \left| \left(-\log G_\beta(X_\beta)\right)^k_{p,n_k} - \log p^{-1} \right| < \infty. \tag{2.8}$$

In view of (2.7) and the rectangular rule of quadrature it suffices to show, in place of (2.8), that

$$\max_{1 \le k \le \lceil n^\beta \rceil} \left| \sum_{v=\lceil p \cdot n_k \rceil}^{n_k} \frac{E_{n,k,v} - 1}{v} \right| = \mathbf{O}\left(n^{1/2(\beta-1)} (\log n)^{1/2}\right). \tag{2.9}$$

Lemma 3 in the Appendix proves the statement in (2.9).

Equations (2.9) and (2.7) together with (2.4) imply that

$$\left(-\log G_\beta(X_\beta)\right)_{p,n} = \log p^{-1} + \mathbf{O}\left(n^{1/2(\beta-1)} (\log n)^{1/2}\right)$$

and hence

$$\left(G_\beta(X_\beta)\right)_{p,n} = p + \mathbf{O}\left(n^{1/2(\beta-1)} (\log n)^{1/2}\right). \tag{2.10}$$

Since

$$\left|G_\beta\big((X_\beta)_{p,n}\big) - F\big((X_\beta)_{p,n}\big)\right| = \Big|U(i)\left[F_\beta\big((X_\beta)_{p,n}-\big) - F\big((X_\beta)_{p,n}\big)\right]$$
$$+(1-U(i))\left[F_\beta\big((X_\beta)_{p,n}\big) - F\big((X_\beta)_{p,n}\big)\right]\Big|.$$

and because of (2.3a), (2.3b) and Lemma 5 we get

$$(G_\beta(X_\beta))_{p,n} = (F(X_\beta))_{p,n} + \mathbf{O}(n^{-\alpha/(\alpha+1)\eta}), \tag{2.11}$$

for n large enough. Then, combining (2.11) and (2.10) establishes

$$(F(X_\beta))_{p,n} = p + \mathbf{O}\Big(\max\big\{n^{1/2(\beta-1)}(\log n)^{1/2}, n^{-\alpha/(\alpha+1)\eta}\big\}\Big).$$

Since, over the neighborhood B_p, the derivative of F is bounded away from 0, which implies that F^{-1} has bounded derivatives, we may transform from $(F(X_\beta))_{p,n}$ to $(X_\beta)_{p,n}$ and obtain

$$(X_\beta)_{p,n} \in I_n \quad \text{with probability 1 for all sufficiently large } n.$$

This completes the proof of Theorem 2. □

REMARK 3. Because of (2.2) and since $\alpha/(\alpha+1) < 1$ for $\alpha > 0$, the statement of Theorem 2 holds with $(X_\beta)_{p,n}$ replaced by $X_{p,n}$.

PROPOSITION 1. *If $F_n^{\beta*}$ is the empirical distribution function of $X_{\beta*}(i)$, $i = 1,\dots,n$, where β^* is not necessarily equal to β above and*

$$V_n(y) = \left|F_n^{\beta^*}(y) - F_n^{\beta^*}(\xi_p) - \big(F_{\beta^*}(y) - F_{\beta^*}(\xi_p)\big)\right|,$$

then

$$\limsup_{n\to\infty} (\gamma(n))^{-1} \sup_{y\in I_n} V_n(y) < \infty \quad a.s.$$

for any $\gamma(n)$ with

$$\gamma(n) = \mathbf{O}\big(\max\{n^{-(1/2)\alpha/(\alpha+1)\eta - 1/2(1-\beta^*)+\varepsilon}, n^{-3/4+\beta/4+\beta^*/2+\varepsilon}\}\big), \qquad \varepsilon > 0,$$

and

$$I_n = (\xi_p - a_n, \xi_p + a_n),$$

where

$$a_n = \mathbf{O}\Big(\max\big\{n^{1/2(\beta-1)}(\log n)^{1/2}, n^{-\alpha/(\alpha+1)\eta}\big\}\Big). \tag{2.12}$$

PROOF. Without loss of generality we assume that F_{β^*} is continuous. If not we use the method introduced after (2.4) and consider G_{β^*} instead of F_{β^*}. Write $F_{n_k}^{\beta^*}$ for the empirical distribution function of $X_{\beta^*}(i)$ based on the subset $S_{n,k}^{\beta^*}$ and $V_{n_k}^k(y)$ accordingly. Then

$$\sup_{y\in I_n} V_n(y) \le \sup_{y\in I_n} \max_{1\le k\le \lceil n^{\beta^*}\rceil} V_{n_k}^k(y).$$

Let

$$b_n = \lceil cn^s \rceil \tag{2.13}$$

be an integer sequence with an optimal exponent s to be selected later. Also, write $\mathbf{w}_{n,v} = \xi_p + a_n b_n^{-1} v$, I_n^v for the interval $[\mathbf{w}_{n,v}, \mathbf{w}_{n,v+1}]$ and $u_{n,v} = F_{\beta^*}(\mathbf{w}_{n,v+1}) - F_{\beta^*}(\mathbf{w}_{n,v})$ for all n and integers v with $-b_n \le v \le b_n - 1$. Then for all $y \in I_n^v$,

$$V_{n_k}^k(y) \le V_{n_k}^k(\mathbf{w}_{n,v+1}) + u_{n,v},$$

$$V_{n_k}^k(y) \ge V_{n_k}^k(\mathbf{w}_{n,v+1}) - u_{n,v}$$

and hence

$$\begin{aligned} \sup_{y \in I_n} V_n(y) &\le \max_{1 \le k \le \lceil n^{\beta^*} \rceil} \max_{-b_n \le v \le b_n} V_{n_k}^k(\mathbf{w}_{n,v}) + \max_{-b_n \le v \le b_n - 1} u_{n,v} \\ &\le T_1(n) + T_2(n), \quad \text{say}. \end{aligned} \tag{2.14}$$

Since $\mathbf{w}_{n,v+1} - \mathbf{w}_{n,v} \le a_n b_n^{-1}$ for each v and since F_{β^*} (or better G_{β^*}) is sufficiently well behaved in a fixed neighborhood of ξ_p, it follows that $T_2(n) = \mathbf{O}(a_n b_n^{-1})$.

As far as $T_1(n)$ is concerned, it suffices, in view of the Borell–Cantelli lemma and Bonferroni's inequality, to show that

$$\sum_{n=N_0}^{\infty} \sum_{k=1}^{\lceil n^{\beta^*} \rceil} \sum_{v} P\left(V_{n_k}^k(\mathbf{w}_{n,v}) \ge \gamma(n)\right) < \infty, \tag{2.15}$$

for N_0 sufficiently large so that here and below degeneracies are avoided. To demonstrate this, we exploit the fact that the distribution of $V_{n_k}^k(\mathbf{w}_{n,v})$ is the same as that of $n_k^{-1}|B(n_k, \delta(n,v)) - n_k\delta(n,v)|$, where $B(n_k, \delta(n,v))$ is a binomial random variable with parameters n_k and $\delta(n,v) = |F_{\beta^*}(\mathbf{w}_{n,v}) - F_{\beta^*}(\xi_p)|$. Using Bernstein's inequality,

$$P\left(|B(n_k, \delta(n,v)) - n_k\delta(n,v)| \ge \gamma(n)\right) \le 2\exp(-h), \tag{2.16}$$

with $h = \gamma(n)^2/\{2[n_k\delta(n,v)(1-\delta(n,v)) + (\gamma(n)/3)\max\{\delta(n,v), 1-\delta(n,v)\}]\}$. N_0 in (2.15) has to be chosen so large that $F_{\beta^*}(\xi_p + a_n) - F_{\beta^*}(\xi_p) < c_1 a_n$ and $F_{\beta^*}(\xi_p) - F_{\beta^*}(\xi_p - a_n) < c_1 a_n$ for all $n > N_0$ and some constant c_1. Using $h = h(n_k, \delta(n,v), \gamma(n)) \ge \gamma(n)^2/(2[n_k\delta(n,v) + \gamma(n)])$ and since $|v| \le b_n$ implies $\delta(n,v) \le c_1 a_n$ for $n > N_0$, it follows

$$P\left(V_{n_k}^k(\mathbf{w}_{n,v}) \ge \gamma(n)\right) \le 2\exp(-h_1),$$

where $h_1 = h_1(n_k, \gamma(n)) = n_k^2\gamma(n)^2/(2[c_1 n_k a_n + n_k\gamma(n)])$ which depends on k only through n_k and is independent of v. Hence

$$\sum_{k=1}^{\lceil n^{\beta^*} \rceil} \sum_{v} P\left(V_{n_k}^k(\mathbf{w}_{n,v}) \ge \gamma(n)\right) \le 4b_n \lceil n^{\beta^*} \rceil \exp\left(-h_1(\lceil n^{1-\beta^*} \rceil - 1, \gamma(n))\right),$$

where $\gamma(n)$ and the exponent of n in $b_n = \lceil cn^s \rceil$ have to be chosen so that for all $n \geq N_0$, the expression $\beta^* + s - h_1(\lceil n^{1-\beta^*} \rceil - 1, \gamma(n))/\log n$ is less than -1. Hence, given β^*, we choose s so that the exponent of $a_n b_n^{-1} n^{1-\beta^*}$ is larger than s and since

$$\mathbf{O}(a_n b_n^{-1}) = \mathbf{O}\left(\max\{n^{-\alpha/(\alpha+1)\eta - s}, n^{1/2(\beta-1)-s}(\log n)^{1/2}\}\right),$$

this requires

$$-\frac{\alpha}{\alpha+1}\eta - s + 1 - \beta^* > s \quad \text{or} \quad 1 - \beta^* - \frac{1}{2}(1-\beta) - s > s,$$

which leads to

$$s < \frac{1}{2}\left(1 - \frac{\alpha}{\alpha+1}\eta - \beta^*\right) \quad \text{or} \quad s < \frac{1}{4} + \frac{\beta}{4} - \frac{\beta^*}{2}$$

and $\gamma(n) = \mathbf{O}(\max\{n^{-(1/2)\alpha/(\alpha+1)\eta-(1/2)(1-\beta^*)+\varepsilon}, n^{-3/4+\beta/4+\beta^*/2+\varepsilon}\})$, $\varepsilon > 0$. This completes the proof of Proposition 1. □

REMARK 4. The rate $\gamma(n)$ essentially determines the rate of convergence in Theorem 1. An optimal choice for β in Theorem 2 is $\beta_0 = (\alpha+3)/(2\alpha q - \alpha - 1)$ so that the optimal a_n is $a_n^0 = \mathbf{O}(n^{-1/2+\lambda})$ for all $\lambda > (\alpha+3)/(2\alpha(2q-1) - 2)$. Since $\sup_{y \in J_n} |F_{\beta^*}(y) - F(y)| = \mathbf{O}(n^{-\alpha/(\alpha+1)\eta^*})$ with $J_n = (\xi_p - a_n^0, \xi_p + a_n^0)$ and $\eta^* = -\beta^*(1 + 1/\alpha - q) - 1/\alpha$, the optimal β^* is (in view of Lemma 4)

$$\beta_0^* = \beta_0 + \frac{\beta_0 - 1}{2 - 4\alpha q/(\alpha+1)}.$$

(Note also that $\beta_0^* \geq \beta_0$.) This implies the optimal

$$R_n = \mathbf{O}(n^{-3/4+\gamma}) \tag{2.17}$$

for all

$$\gamma > \frac{\alpha^2(8q-5) + 2\alpha(10q-9) - 13}{4(2\alpha q - \alpha - 1)^2}. \tag{2.18}$$

PROOF OF THEOREM 1. Theorem 2 and Proposition 1 provide us with the technology to establish the main result. Due to Theorem 2 and Remark 3, we may select N_0 such that for all $n > N_0$, $X_{p,n} \in I_n^0 = J_n$. Then, also, $F_n(X_{p,n}) = \lceil n \cdot p \rceil / n$. Since $\sup_{y \in J_n} |F_{\beta_0^*}(y) - F(y)| = \mathbf{O}(n^{-\alpha/(\alpha+1)\eta_0^*})$ and $\sup_{y \in J_n} |F_n(y) - F_n^{\beta_0^*}(y)| = \mathbf{O}(n^{-\alpha/(\alpha+1)\eta_0^*})$, by Lemma 4, using Proposition 1 (with $\beta^* = \beta_0^*$) applied to $y = X_{p,n}$ gives

$$\frac{\lceil n \cdot p \rceil}{n} = F_n(\xi_p) + F(X_{p,n}) - F(\xi_p) + \mathbf{O}(n^{-3/4+\gamma}) \tag{2.19}$$

for all γ satisfying (2.18). Since, by assumption, F is sufficiently smooth within the neighborhood B_p of ξ_p, we may use Taylor's theorem (in Young's form) to assert that

$$F(X_{p,n}) = F(\xi_p) + (X_{p,n} - \xi_p) f(\xi_p) + \mathbf{O}\big((a_n^0)^2\big). \tag{2.20}$$

Consequently, combining (2.19) and (2.20),

$$X_{p,n} = \xi_p + \frac{\lceil n \cdot p \rceil / n - F_n(\xi_p)}{f(\xi_p)} + \mathbf{O}\Big(\max\big\{(a_n^0)^2, n^{-3/4+\gamma}\big\}\Big).$$

Comparing the rates $(a_n^0)^2$ and $n^{-3/4+\gamma}$ and observing that $\lceil n \cdot p \rceil / n = p + \mathbf{O}(n^{-1})$ gives the desired result. This completes the proof of Theorem 1. □

APPENDIX

The appendix contains the lemmas used in the proof of Theorem 1.

LEMMA 1. *Let $X(i)$ and $X_\beta(i)$, $i = 1, \dots, n$, be two sequences of random variables with stationary distribution functions F and F_β, respectively. F_β may depend on n. Assume that F has bounded derivative in some neighborhood B_p of ξ_p with $F(\xi_p) = p$. Assume also that*

$$\max_{1 \le i \le n} \big| X(i) - X_\beta(i) \big| \le c\varepsilon^* n^{-\rho}, \quad a.s., \tag{A.1}$$

where ρ is a positive constant and ε^* is a random variable independent of n and such that

$$P(\varepsilon^* \ge \xi) \le c\xi^{-\alpha} \tag{A.2}$$

for some $\alpha > 0$ and any $\xi > 0$. Then

$$\limsup_{n \to \infty} n^{\alpha\rho/(1+\alpha)} \sup_{y \in B_p} \big| F(y) - F_\beta(y) \big| < \infty.$$

PROOF. We may conclude from (A.1) that for any λ with $0 < \lambda < \rho$ and all $y \in B_p$,

$$\begin{aligned} &P(X(i) \le y - n^{-\lambda}) - P\left(\varepsilon^* \ge \frac{1}{c} n^{\rho-\lambda}\right) \\ &\quad \le P(X_\beta(i) \le y) \le P(X(i) \le y + n^{-\lambda}) + P\left(\varepsilon^* \ge \frac{1}{c} n^{\rho-\lambda}\right). \end{aligned}$$

Using (A.2), it is clear that

$$F(y - n^{-\lambda}) - \mathbf{O}(n^{-\alpha(\rho-\lambda)}) \le F_\beta(y) \le F(y + n^{-\lambda}) + \mathbf{O}(n^{-\alpha(\rho-\lambda)})$$

and consequently, since F has bounded derivative over B_p,

$$\big| F(y) - F_\beta(y) \big| = \mathbf{O}(n^{-\lambda} + n^{-\alpha(\rho-\lambda)}).$$

Selecting $\lambda = \alpha\rho/(1+\alpha)$, we obtain the best possible rate:

$$\sup_{y \in B_p} |F(y) - F_\beta(y)| = \mathbf{O}(n^{-\alpha\rho/(1+\alpha)}).$$

This completes the proof. □

Theorem 2 utilizes this lemma with $\rho = -\beta(1 + 1/\alpha - q) - 1/\alpha$.

LEMMA 2. *Let* $\mathbf{J} = \{X(i): i \in \{1, 2, \dots, n\}\}$ *be a set of random variables and* $S_{n,k}$, $k = 1, \dots, r$, *be* r *nonempty disjoint subsets of cardinality* n_k *of the set* $\mathbf{J}$ *with* $\bigcup_{k=1}^r S_{n,k} = \mathbf{J}$. *Then, for any* $0 < p < 1$, *the* pth *sample quantile* $X_{p,n}$ *of* $\mathbf{J}$ *and the* pth *sample quantiles* X_{p,n_k}^k *of* $S_{n,k}$ *satisfy the inequalities*

$$\min_{1 \le k \le r} X_{p,n_k}^k \le X_{p,n} \le \max_{1 \le k \le r} X_{p,n_k}^k.$$

PROOF. Since $\lceil n_k \cdot p \rceil \ge n_k \cdot p > \lceil n_k \cdot p \rceil - 1$ for all $k = 1, \dots, r$, it is true that

$$\sum_{k=1}^r \lceil n_k \cdot p \rceil \ge \lceil n \cdot p \rceil > \sum_{k=1}^r \lceil n_k \cdot p \rceil - r$$

and hence

$$\#\Big\{X(i): X(i) < \min_{1 \le k \le r} X_{p,n_k}^k\Big\} \le \sum_{k=1}^r (\lceil n_k \cdot p \rceil - 1) < \lceil n \cdot p \rceil \tag{A.3}$$

and

$$\#\Big\{X(i): X(i) \le \max_{1 \le k \le r} X_{p,n_k}^k\Big\} \ge \sum_{k=1}^r \lceil n_k \cdot p \rceil \ge \lceil n \cdot p \rceil. \tag{A.4}$$

(A.3) implies that $\min_{1 \le k \le r} X_{p,n_k}^k \le X_{p,n}$ and (A.4) implies that $\max_{1 \le k \le r} X_{p,n_k}^k \ge X_{p,n}$. □

REMARK 5. Both r and n_k may be functions of n.

LEMMA 3. *For integers* n, k, v, *let* $E_{n,k,v}$ *be random variables having an exponential distribution with mean* 1. *Also, for all* n, k, *let* E_{n,k,v_1} *and* E_{n,k,v_2} *be independent whenever* $v_1 \ne v_2$. *Then, for any* $0 < \beta < 1$,

$$\limsup_{n \to \infty} \left(\frac{n^{1-\beta}}{\log n}\right)^{1/2} \max_{1 \le k \le \lceil n^\beta \rceil} \left| \sum_{v = \lceil p \cdot n_k \rceil}^{n_k} \frac{E_{n,k,v} - 1}{v} \right| < \infty \quad a.s.,$$

where the n_k *are defined after* (2.3b) *in Section* 2.

PROOF. We start by deriving sharp bounds for

$$P\left(\left(\frac{n^{1-\beta}}{\log n}\right)^{1/2} \sum_{v = \lceil p \cdot n_k \rceil}^{n_k} \frac{E_{n,k,v} - 1}{v} \ge M \right). \tag{A.5}$$

Using Chernoff's bound [Chernoff (1952)] we obtain that this probability is bounded by

$$\left(\prod_{v=\lceil p\cdot n_k\rceil}^{n_k} \frac{1}{1-t/v}\exp\left(\frac{-t}{v}\right)\right)\exp\left(-tMn^{1/2(\beta-1)}(\log n)^{1/2}\right)$$

for all $0 \le t \le \lceil p \cdot n_k \rceil$. We choose $t = c_3 n^{1/2(1-\beta)}(\log n)^{1/2}$ with some positive constant c_3 to be determined later. Then

$$\begin{aligned}\log\left(\prod_{v=\lceil p\cdot n_k\rceil}^{n_k} \frac{1}{1-t/v}\exp\left(\frac{-t}{v}\right)\right) &= \sum_{v=\lceil p\cdot n_k\rceil}^{n_k} -\log\left(1-\frac{t}{v}\right)-\frac{t}{v}\\ &= \sum_{v=\lceil p\cdot n_k\rceil}^{n_k} \frac{t^2}{2v^2}+o(1)\\ &\le \frac{c_3^2 n^{1-\beta}\log n}{pn_k} \quad \text{(for } n \text{ large enough)}\\ &\le \frac{c_3^2}{p}\log n \quad \text{(for } n \text{ large enough)}.\end{aligned}$$

Hence

$$P\left(\left(\frac{n^{1-\beta}}{\log n}\right)^{1/2}\sum_{v=\lceil p\cdot n_k\rceil}^{n_k}\frac{E_{n,k,v}-1}{v} \ge M\right) \le n^{c_3^2/p-c_3M}. \tag{A.6}$$

Taking $c_3 = pM/2$, we may make, for given p, the exponent of n on the right-hand side of (A.6) as small as we want by increasing M. Hence for n and M large enough, the probability in (A.6) is bounded by $n^{-pM^2/4} = \psi(n, p, M)$, say.

The same argument can be applied to

$$(-1)\left(\frac{n^{1-\beta}}{\log n}\right)^{1/2}\sum_{v=\lceil p\cdot n_k\rceil}^{n_k}\frac{E_{n,k,v}-1}{v}.$$

Combining both, it is obvious that

$$P\left(\left(\frac{n^{1-\beta}}{\log n}\right)^{1/2}\left|\sum_{v=\lceil p\cdot n_k\rceil}^{n_k}\frac{E_{n,k,v}-1}{v}\right| \ge M\right) \le 2\psi(n,p,M)$$

and the bound is independent of k. Exploiting this uniformity and choosing M large enough we get

$$\sum_{n=N_0}^{\infty} P\left(\max_{1\le k\le\lceil n^{\beta}\rceil}\left(\frac{n^{1-\beta}}{\log n}\right)^{1/2}\left|\sum_{v=\lceil p\cdot n_k\rceil}^{n_k}\frac{E_{n,k,v}-1}{v}\right| \ge M\right) < \infty \tag{A.7}$$

and the Borel–Cantelli lemma produces the desired result. □

LEMMA 4. *With the notation of Section* 2,

$$\sup_{y\in J_n}\left|F_n(y)-F_n^{\beta_0^*}(y)\right|=\mathbf{O}\left(n^{-\alpha/(\alpha+1)\eta_0^*}\right),$$

where $J_n=(\xi_p-a_n^0,\xi_p+a_n^0)$ *and* a_n^0,β_0,β_0^* *are as in Remark* 4.

PROOF. Choose r such that $\alpha/(\alpha+1)(\eta_0^*-\eta_0)<r<\eta_0^*-\alpha/(\alpha+1)\eta_0$, $d_n=\lceil cn^r\rceil$ and define $y_{n,v}=\xi_p+a_n^0 d_n^{-1}v$.

For n sufficiently large and all v with $|v|\le d_n$, the empirical distribution function evaluated at $y_{n,v}$,

$$F_n(y_{n,v})=\frac{1}{n}\sum_{i=1}^n\chi\left(X(i)\le y_{n,v}\right)$$

is upperbounded a.s. by

$$\frac{1}{n}\sum_{i=1}^n\chi\left(X_{\beta_0^*}(i)\le y_{n,v}+a_n^0 d_n^{-1}\right)+\frac{1}{n}\sum_{i=1}^n\chi\left(X_{\beta_0^*}(i)-X(i)>a_n^0 d_n^{-1}\right).$$

Similarly,

$$\begin{aligned}F_n(y_{n,v})\ge\ &\frac{1}{n}\sum_{i=1}^n\chi\left(X_{\beta_0^*}(i)\le y_{n,v}-a_n^0 d_n^{-1}\right)\\&-\frac{1}{n}\sum_{i=1}^n\chi\left(X(i)-X_{\beta_0^*}(i)>a_n^0 d_n^{-1}\right),\end{aligned}$$

so that

$$\begin{aligned}\left|F_n(y_{n,v})-F_n^{\beta_0^*}(y_{n,v})\right|\le\ &\max\left\{\left|F_n^{\beta_0^*}(y_{n,v}\pm a_n^0 d_n^{-1})-F_n^{\beta_0^*}(y_{n,v})\right|\right\}\\&+\frac{1}{n}\sum_{i=1}^n\chi\left(\left|X(i)-X_{\beta_0^*}(i)\right|>a_n^0 d_n^{-1}\right),\end{aligned}\tag{A.8}$$

where here the max is to be taken over the choice of signs in the argument of $F_n^{\beta_0^*}(y_{n,v}\pm a_n^0 d_n^{-1})$.

It is also easy to show that

$$\begin{aligned}\sup_{y\in J_n}\left|F_n(y)-F_n^{\beta_0^*}(y)\right|\le\ &\max_{|v|\le d_n}\left|F_n(y_{n,v})-F_n^{\beta_0^*}(y_{n,v})\right|\\&+\max_{-d_n\le v<d_n}\left|F_n^{\beta_0^*}(y_{n,v+1})-F_n^{\beta_0^*}(y_{n,v})\right|.\end{aligned}$$

Combining (A.8) and the previous inequality we see that

$$\begin{aligned}\sup_{y\in J_n}\left|F_n(y)-F_n^{\beta_0^*}(y)\right|\le\ &2\max_{|v|\le d_n}\left|F_n^{\beta_0^*}(y_{n,v+1})-F_n^{\beta_0^*}(y_{n,v})\right|\\&+\frac{1}{n}\sum_{i=1}^n\chi\left(\left|X(i)-X_{\beta_0^*}(i)\right|>a_n^0 d_n^{-1}\right)\\=\ &2S(1)+S(2),\quad\text{say.}\end{aligned}$$

To show the required rate for $S(1)$, one makes use of the fact that $n(F_n^{\beta\delta^*}(y_{n,v+1}) - F_n^{\beta\delta^*}(y_{n,v}))$ has the same distribution as

$$\sum_{i=1}^{n} \chi\left(y_{n,v} < X_{\beta_0^*}(i) \le y_{n,v+1}\right)$$

and $\chi(y_{n,v} < X_{\beta_0}^*(i) \le y_{n,v+1})$, $i = 1, \dots, n$, is a sequence of $\lceil n^{\beta\delta^*} \rceil$-dependent Bernoulli random variables with parameter equal to

$$F_{\beta_0^*}(y_{n,v+1}) - F_{\beta_0^*}(y_{n,v}) = \mathbf{O}\left(a_n^0 d_n^{-1}\right).$$

The inequalities in Hoeffding (1963), Section 5d, admit a straightforward extension to $\lceil n^{\beta\delta^*} \rceil$-dependent random variables and using these inequalities together with the Borel–Cantelli lemma proves the required rate for $S(1)$.

To show that

$$\frac{1}{n} \sum_{i=1}^{n} \chi\left(\left|X(i) - X_{\beta_0^*}(i)\right| > a_n^0 d_n^{-1}\right) = \mathbf{O}\left(n^{-\alpha/(\alpha+1)\eta\delta^*}\right),$$

it is sufficient to realize that

$$\max_{1 \le i \le n} \left|X(i) - X_{\beta_0^*}(i)\right| \le c\varepsilon^* n^{-\eta\delta^*} \quad \text{a.s.}$$

by (2.1), that $a_n^0 d_n^{-1} > c\varepsilon^* n^{-\eta\delta^*}$ for all sufficiently large n and that ε^* is a.s. finite. □

LEMMA 5. *With the notation of Section 2 it is true for all β with $0 < \beta < 1$ that*

$$(X_\beta)_{p,n} \in B_p \quad a.s.\ \text{for all } n \text{ large enough}.$$

Here B_p is the fixed neighborhood of ξ_p over which the density of $X(i)$ is bounded away from 0 and infinity.

PROOF. We will show that $(X_\beta)_{p,n} \to \xi_p$ a.s. from which the statement follows. For all $\delta > 0$ it is clear that

$$F(\xi_p - \delta) < p < F(\xi_p + \delta).$$

Now, if we can also show that

$$\text{(A.9)} \qquad \begin{aligned} F_n^\beta(\xi_p - \delta) &\to F(\xi_p - \delta) \quad \text{a.s.,} \\ F_n^\beta(\xi_p + \delta) &\to F(\xi_p + \delta) \quad \text{a.s.,} \end{aligned}$$

then it follows that

$$F_n^\beta(\xi_p - \delta) < p < F_n^\beta(\xi_p + \delta) \quad \text{a.s. for all } n \text{ large enough}$$

and therefore

$$\xi_p - \delta < (X_\beta)_{p,n} \le \xi_p + \delta \quad \text{a.s. for all } n \text{ large enough},$$

because clearly $F_n^\beta(\xi_p + \delta) \ge p$ iff $\xi_p + \delta \ge (F_n^\beta)^{-1}(p) = (X_\beta)_{p,n}$.

So we only need to show (A.9). We prove only that

$$F_n^\beta(\xi_p - \delta) \to F(\xi_p - \delta) \quad \text{a.s.}$$

That is, for all $\varepsilon > 0$,

$$\text{(A.10)} \qquad \sum_n P\left(\left| \frac{1}{n} \sum_{i=1}^{n} \chi(X_\beta(i) \le \xi_p - \delta) - F(\xi_p - \delta) \right| > \varepsilon \right) < \infty.$$

The left-hand side of (A.10) is upperbounded by

$$\text{(A.11)} \qquad \sum_n P\left(\left| \frac{1}{n} \sum_{i=1}^{n} \chi(X_\beta(i) \le \xi_p - \delta) - F_\beta(\xi_p - \delta) \right| > \varepsilon - cn^{-\alpha/(\alpha+1)\eta} \right)$$

by (2.3a). By construction, $\chi(X_\beta(i) \le \xi_p - \delta)$, $i = 1, \dots, n$, is an $\lceil n^\beta \rceil$-dependent sequence of Bernoulli random variables and again we may use the results in Hoeffding (1963), Section 5d, to prove the finiteness of (A.11). □

Acknowledgments. The author would like to thank Professor E. J. Hannan for many stimulating discussions.

REFERENCES

ANDREWS, D. W. K. (1984). Nonstrong mixing autoregressive processes. *J. Appl. Probab.* **21** 930–934.

BABU, G. J. and SINGH, K. (1978). On deviations between empirical and quantile processes for mixing random variables. *J. Multivariate Anal.* **8** 532–549.

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

CHANDA, K. C. (1974). Strong mixing properties of linear stochastic processes. *J. Appl. Probab.* **11** 401–408.

CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

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

EICKER, F. (1966). On the asymptotic representation of sample quantiles (Abstract). *Ann. Math. Statist.* **37** 1424.

GHOSH, J. K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Ann. Math. Statist.* **42** 1957–1961.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

KIEFER, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.

KIEFER, J. (1970a). Deviations between the sample quantile and the sample df. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 299–319. Cambridge Univ. Press, New York.

KIEFER, J. (1970b). Old and new methods for studying order statistics and sample quantiles. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 349–357. Cambridge Univ. Press, New York.

SEN, P. K. (1968). Asymptotic normality of sample quantiles for m-dependent processes. *Ann. Math. Statist.* **39** 1724–1730.

SEN, P. K. (1972). On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables. *J. Multivariate Anal.* **2** 77–95.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

DEPARTMENT OF STATISTICS
STATISTICAL LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720