A Bahadur - Type Representation for Empirical Quantiles of a Large Class of Stationary, Possibly Infinite-Variance, Linear Processes

By

C.H. Hesse
Department of Statistics
University of California, Berkeley

Technical Report No. 132
December 1987
(revised June 1989)

Department of Statistics
University of California
Berkeley, California 94720
A Bahadur - Type Representation for Empirical Quantiles of a Large Class of Stationary, Possibly Infinite-Variance, Linear Processes

By

C.H. Hesse
Department of Statistics
University of California, Berkeley

Abstract
Bahadur (1966) 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.

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

Key Words: Stationary linear processes, Quantiles, Bahadur representation, Almost sure convergence.

AMS subject classifications: Primary 62G30;
Secondary 60F15, 60G10
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), ... the following is true with probability one:

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

and

\[ R_n = O(n^{-3/4}(\log n)^{1/2}(\log \log n)^{1/4}). \]

Here, for 0 < p < 1, \( \xi_p \) is the unique p-quantile of F, i.e. \( F(\xi_p) = p \), \( X_{p,n} \) is the p-th 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 non-independent 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 following definite answer,

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

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.

There are some extensions to sequences of random variables with certain dependency
structures, e.g. m-dependence, \( \phi \)-mixingness, 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 following sequences are considered:

\begin{equation}
X(n) = \sum_{i=0}^{\infty} \delta(i) \varepsilon(n - i)
\end{equation}

where \( \varepsilon(n) \) are iid innovations with \( \mathbb{E}(|\varepsilon(n)|^p) < \infty \) 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 \( \phi \)-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

\begin{equation}
X(n) - \frac{1}{2} X(n - 1) = \varepsilon(n)
\end{equation}

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 (1983). 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 RESULT

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 $p$-th sample quantile $X_{p,n}$ of a sequence of random variables $X(1), \ldots, X(n)$ 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 $F_n$, $F$ denote the empirical distribution and the stationary distribution of $X$, respectively, $\xi_p$ 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

\begin{equation}
X(n) = \sum_{i=0}^{\infty} \delta(i) \varepsilon(n-i)
\end{equation}

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 $\infty$ in a neighborhood $B_p$ of $\xi_p$. Then

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

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

**Remark 1:** The lower bound for $\gamma$ in Theorem 1 is decreasing both for increasing $\alpha$ 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 \( i = 1, \ldots, n \)

\[
X_\beta(i) = \sum_{j=0}^{[n^{\beta}] - 1} \delta(j)e(i - j).
\]

Then it holds true for the empirical p-th quantile \((X_\beta)_{p,n}\) of \( X_\beta(i) \) that for all \( n \) sufficiently large

\[
(X_\beta)_{p,n} \in I_n = (\xi_{\beta p} - a_n, \xi_{\beta p} + a_n) \text{ a.s.}
\]

with \( a_n = O\left(\max \left\{ n^{1/2(\beta - 1)}(\log n)^{1/2}, n^{\frac{-\alpha}{\alpha + 1}}(\beta(1 + 1/\alpha + \alpha) + 1/\alpha) \right\} \right) \).

Remark 2: The parameter \( \beta \) 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 \( \beta \) will have to be made later.

Proof of Theorem 2: For latter use we first evaluate the difference between \( X(i) \) and \( X_\beta(i) \). Clearly,

\[
|X(i) - X_\beta(i)| \leq c \cdot \sum_{j=0}^{[n^{\beta}] - 1} j^{-q}|e(i - j)|
\]

\[
\leq c \cdot e^* n^{1/\alpha} \sum_{j=0}^{[n^{\beta}] - 1} j^{-q+1/\alpha}
\]

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

\begin{equation}
|X(i) - X_\beta(i)| \leq c \cdot \varepsilon^* n^\beta (1 + 1/\alpha - q) + 1/\alpha.
\end{equation}

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

\begin{equation}
|\{X_\beta\}_{p,n} - X_{p,n}| = O(n^{-\eta})
\end{equation}

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

Then, using Lemma 1 from the Appendix, we get

\begin{equation}
\sup_{y \in B_p} |F(y) - F_\beta(y)| = O(n^{\alpha / (\alpha + 1) \eta})
\end{equation}

and

\begin{equation}
\sup_{y \in B_p} |F(y) - F_\beta(y^-)| = O(n^{\alpha / (\alpha + 1) \eta})
\end{equation}

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

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

$$S_{n,k}^\beta = \{X_\beta(k), X_\beta(k + [n^\beta]), \ldots, X_\beta(k + (n_k - 1) \cdot [n^\beta])\}, \ k = 1, 2, \ldots, [n^\beta].$$

where $n_k$ is either $[n^{1-\beta}]$ or $[n^{1-\beta}] - 1$, its dependence on $k$ being of no concern.

If for given $n$, $X_{p,n_k}^k$ denotes the p-th sample quantile of the k-th set $S_{n,k}^\beta$ containing $n_k$ random variables, then by Lemma 2

\begin{equation}
\min_{1 \leq k \leq [n^\beta]} X_{p,n_k}^k \leq (X_\beta)_{p,n} \leq \max_{1 \leq k \leq [n^\beta]} X_{p,n_k}^k.
\end{equation}

At this point it is necessary to point out one of the defects of the distribution $F_\beta$: its possible discontinuity. We remedy this by introducing the slightly perturbed but continuous random variables
\[ G_\beta(X_\beta(i)) = U(i) \cdot F_\beta(X_\beta(i)-) + (1 - U(i)) F_\beta(X_\beta(i)) \]

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

\[ G_\beta(y) = U(i) \cdot F_\beta(y-) + (1 - U(i)) F_\beta(y) \]

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

\[
\sup_{y \in B_p} |F_\beta(y) - F_\beta(y-)| \leq \sup_{y \in B_p} |F_\beta(y) - F(y)| + \sup_{y \in B_p} |F(y) - F_\beta(y-)| = O(n^{-\frac{\alpha}{\alpha+1} \eta})
\]

and hence

\[
(2.5) \quad \sup_{y \in B_p} |F_\beta(y) - G_\beta(y)| = O(n^{-\frac{\alpha}{\alpha+1} \eta})
\]

Since, in particular, for all \( n \) sufficiently large \((X_\beta)p_n \in B_p\) a.s., by Lemma 5, Equation (2.5) therefore implies that

\[
|F_\beta((X_\beta)p_n) - G_\beta((X_\beta)p_n)| = O(n^{-\frac{\alpha}{\alpha+1} \eta})
\]

from which we deduce

\[
(2.6) \quad (F_\beta(X_\beta))p_n = (G_\beta(X_\beta))p_n + O(n^{-\frac{\alpha}{\alpha+1} \eta})
\]

by monotonicity of \( F_\beta \) and \( G_\beta \). In Equation (2.6) \((F_\beta(X_\beta))p_n\) is the \( \lceil n \cdot p \rceil \)-th order statistic of \( F_\beta(X_\beta(i)), i = 1, \ldots, n \) 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 \( p \)-th quantiles may be ignored. Keeping in mind (2.6) we are in the sequel concerned with \( G_\beta(X_\beta(i)), i = 1, \ldots, n \) only.

We will first determine how close \((-\log(G_\beta(X_\beta)))p_n\), the \( p \)-th quantile of \(-\log(G_\beta(X_\beta))), i = 1, \ldots, n \) is to \( \log p^{-1} \). Since \( X_\beta(i) \) has distribution \( F_\beta \), \(-\log(G_\beta(X_\beta))) \) 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) p.723)
to these transformed random variables. In particular, for the \([n_k \cdot p]-th\) order statistic 
\((-\log G_\beta (X_\beta))_{p,n_k}^k\) of the transformed random variables in \(S_{n,k}^\beta\) we obtain

\[
(-\log G_\beta (X_\beta))_{p,n_k}^k = \sum_{v=[p \cdot n_k]}^{n_k} \frac{E_{n,k,v}}{v}
\]

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

\[
\limsup_{n \to \infty} \max_{1 \leq k \leq \lfloor \frac{n}{p} \rfloor} \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \left( (-\log G_\beta (X_\beta))_{p,n_k}^k - \log p^{-1} \right) < \infty.
\]

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

\[
\max_{1 \leq k \leq \lfloor \frac{n}{p} \rfloor} \left| \sum_{v=[p \cdot n_k]}^{n_k} \frac{E_{n,k,v} - 1}{v} \right| = O(n^{1/2(\beta-1)}(\log n)^{1/2}).
\]

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

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

\[
(-\log G_\beta (X_\beta))_{p,n} = \log p^{-1} + O(n^{1/2(\beta-1)}(\log n)^{1/2})
\]

and hence

\[
(G_\beta (X_\beta))_{p,n} = p + O(n^{1/2(\beta-1)}(\log n)^{1/2}).
\]

Since

\[
|G_\beta ((X_\beta)_{p,n}) - F(X_\beta)_{p,n}| = |U(i)[F_\beta (X_\beta)_{p,n} - F(X_\beta)_{p,n}] + (1 - U(i))[F_\beta (X_\beta)_{p,n} - F(X_\beta)_{p,n}]|
\]

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

\[
(G_\beta (X_\beta))_{p,n} = (F(X_\beta))_{p,n} + O\left(n^{-\frac{\alpha}{\alpha+1}}\right)
\]

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

\[
(F(X_\beta))_{p,n} = p + O\left(\max \left\{ n^{1/2(\beta-1)}(\log n)^{1/2}, n^{-\frac{\alpha}{\alpha+1}} \right\} \right).
\]

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

\[(X_{\beta})_{p,n} \in I_n \text{ with probability one for all sufficiently large } n.\]

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 \(X_{\beta*}(i)\), \(i = 1, \ldots, n\) where \(\beta^*\) is not necessarily equal to \(\beta\) above, and

\[V_n(y) = |F_n^{\beta*}(y) - F_n^{\beta*}(\xi_p) - (F_{\beta^*}(y) - F_{\beta^*}(\xi_p))|,\]

Then

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

for any \(\gamma(n)\) with

\[\gamma(n) = O\left(\max\left\{\frac{n^{-1/2} \alpha}{\eta} - \frac{1}{2} (1-\beta^*) + \epsilon, \frac{3}{4} + \frac{\beta + \beta^*}{2} + \epsilon\right\}\right), \epsilon > 0\]

and

\[I_n = (\xi_p - a_n, \xi_p + a_n)\]

where

\[(2.12) \quad a_n = O\left(\max\left\{n^{1/2} (\beta - 1) (\log n)^{1/2}, n^{-\alpha/(\alpha+1)\eta}\right\}\right)\]

Proof of Proposition 1: Without loss of generality we assume that \(F_{\beta^*}\) is continuous.

If not we use the method introduced after (2.4) and consider \(G_{\beta^*}\) instead of \(F_{\beta^*}\). Write \(F_{\beta^*}\) for the empirical distribution function of \(X_{\beta^*}(i)\) based on the subset \(S_{\beta^*}\), and \(V_{\beta^*}^{k}(y)\) accordingly, then

\[
\sup_{y \in I_n} V_n(y) \leq \sup_{y \in I_n} \max_{1 \leq k \leq \lfloor n^{\beta^*}\rfloor} V_{\beta^*}^{k}(y).
\]
Let

\[(2.13) \quad b_n = \lceil cn^s \rceil\]

be an integer sequence with an optimal exponent \(s\) to be selected later. Also, write

\[w_{n,v} = \xi_p + a_n \cdot b_n^{-1} \cdot v, \quad I_n^v\]

for the interval \([w_{n,v}, w_{n,v+1}]\) and \(u_{n,v} = F_{\beta^*}(w_{n,v+1}) - F_{\beta^*}(w_{n,v})\) for all \(n\) and integers \(v\) with \(-b_n \leq v \leq b_n - 1\). Then for all \(y \in I_n^v\)

\[
V_{n_k}^k(y) \leq V_{n_k}^k(w_{n,v+1}) + u_{n,v}
\]

\[
V_{n_k}^k(y) \geq V_{n_k}^k(w_{n,v+1}) - u_{n,v}
\]

and hence

\[(2.14) \quad \sup_{y \in I_n^v} V_n(y) \leq \max_{l \leq k} \max_{v \in \mathbb{Z}} V_{n_k}^k(w_{n,v}) + \max_{v \in \mathbb{Z}} u_{n,v}
\]

\[
\leq T_1(n) + T_2(n), \quad \text{say.}
\]

Since \(w_{n,v+1} - w_{n,v} \leq a_n b_n^{-1}\) for each \(v\), and since \(F_{\beta^*}\) (or better \(G_{\beta^*}\)) is sufficiently well-behaved in a fixed neighborhood of \(\xi_p\), it follows that \(T_2(n) = O(a_n b_n^{-1})\).

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

\[(2.15) \quad \sum_{n=N_0}^{\infty} \sum_{k=1}^{[n^b]} \sum_{v} P(V_{n_k}^k(w_{n,v}) \geq \gamma(n)) < \infty
\]

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

\[(2.16) \quad P(|B(n_k, \delta(n,v)) - n_k \cdot \delta(n,v)| \geq \gamma(n)) \leq 2 \exp(-h)
\]

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 \text{ for all } n > N_0 \text{ and some constant } c_1. \text{ Using } h = h(n_k, \delta(n, \nu), \gamma(n)) \geq \gamma(n)^{2/2} [n_k \cdot \delta(n, \nu) + \gamma(n)] \text{ and since } |\nu| \leq b_n \text{ implies } \\
\delta(n, \nu) \leq c_1 \cdot a_n \text{ for } n > N_0 \text{ it follows } \\
P(V_{n_k}(w_{n, \nu}) \geq \gamma(n)) \leq 2 \exp(-h_1) \\
where h_1 = h_1(n_k, \gamma(n)) = n_k^2 \gamma(n)^2 / 2 \cdot n_k \cdot c_1 a_n + n_k \cdot \gamma(n) \text{ which depends on } k \text{ only through } n_k \text{ and is independent of } \nu. \text{ Hence } \\
\sum_{k=1}^{[n^\beta^*]} \sum_{\nu} P(V_{n_k}(w_{n, \nu}) \geq \gamma(n)) \leq 4 b_n \cdot [n^\beta^*] \cdot \exp(-h_1([n^{1-\beta^*}] - 1, \gamma(n))) \\
where \gamma(n) \text{ and the exponent of } n \text{ in } b_n = [cn^\alpha] \text{ have to be chosen so that for all } n \geq N_0 \text{ the expression } \beta^* + s - h_1([n^{1-\beta^*}] - 1, \gamma(n))/\log n \text{ is less than } -1. \text{ Hence, given } \beta^*, \text{ we choose } s \text{ so that the exponent of } a_n b_n^{-1} \cdot n^{1-\beta^*} \text{ is larger than } s \text{ and since } \\
O(a_n b_n^{-1}) = O\left(\max\left\{ n^{\frac{-\alpha}{\alpha+1} - \frac{s}{\eta} - \frac{s}{1 - \beta^*}}, n^{\frac{1}{\eta} \cdot (1 - \beta^*) - \frac{s}{\eta}}, \frac{\alpha}{\alpha + 1} \cdot \frac{\beta^*}{4} + \frac{1}{4} \cdot \frac{\beta^*}{2}, n^{\frac{1}{\eta} \cdot (1 - \beta^*) + \varepsilon} \right\} \right), \varepsilon > 0. \text{ 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 $\beta$ in Theorem 2 is $\beta_0 = \frac{\alpha + 3}{2\alpha q - \alpha - 1}$ so that the optimal $a_n$ is \\
a_n^0 = O\left(n^{-1/2 + \lambda}\right) \text{ for all } \lambda > \frac{\alpha + 3}{2\alpha (2q - 1) - 1}.\]
Since \( \sup_{y \in I_n} \left| F'_{\beta^*}(y) - F(y) \right| = O\left( n^{-\frac{\alpha}{\alpha+1} \eta^*} \right) \) with \( J_n = (\xi_p - a_n^0, \xi_p + a_n^0) \), and
\[
\eta^* = -\beta^* (1 + 1/\alpha - q) - 1/\alpha,
\]
the optimal \( \beta^* \) is (in view of Lemma 4)
\[
\beta_0^* = \beta_0 + \frac{\beta_0 - 1}{2 \alpha + 4\alpha q}.
\]
(Note also that \( \beta_0^* \geq \beta_0 \)). This implies the optimal
\[
R_n = O\left( n^{-3/4 + \gamma} \right)
\]
for all
\[
\gamma > \frac{\alpha^2 (8q - 5) + 2\alpha (10q - 9) - 13}{4(2\alpha q - \alpha - 1)^2}.
\]
End of Remark.

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 \( n \geq N_0 \), \( X_{p,n} \in I_n^0 = J_n \). Then, also, \( F_n(X_{p,n}) = \left\lfloor n \cdot p \right\rfloor / n \). Since
\[
\sup_{y \in I_n} \left| F'_{\beta^*}(y) - F(y) \right| = O\left( n^{-\frac{\alpha}{\alpha+1} \eta_0^*} \right)
\]
and
\[
\sup_{y \in I_n} \left| F_n(y) - F'_{\beta_0^*}(y) \right| = O\left( n^{-\frac{\alpha}{\alpha+1} \eta_0^*} \right)
\]
by Lemma 4, using Proposition 1 (with \( \beta^* = \beta_0^* \)) applied to \( y = X_{p,n} \) gives
\[
\left\lfloor n \cdot p \right\rfloor / n = F_n(\xi_p) + F(X_{p,n}) - F(\xi_p) + O\left( n^{-3/4 + \gamma} \right)
\]
for all \( \gamma \) satisfying (2.18). Since, by assumption, \( F \) is sufficiently smooth within the neighborhood \( B_p \) of \( \xi_p \) we may use Taylor’s theorem (in Young’s form) to assert that
\[
F(X_{p,n}) = F(\xi_p) + (X_{p,n} - \xi_p) f(\xi_p) + O\left( (a_n^0)^2 \right).
\]
Consequently, combining (2.19) and (2.20)
\[
X_{p,n} = \xi_p + \frac{\left\lfloor n \cdot p \right\rfloor / n - F_n(\xi_p)}{f(\xi_p)} + O\left( (a_n^0)^2, n^{-3/4 + \gamma} \right).
\]
Comparison of the rates \( (a_n^0)^2 \) and \( n^{-3/4 + \gamma} \), and observing that \( \left\lfloor n \cdot p \right\rfloor / n = p + O\left( n^{-1} \right) \) 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 $X(i), X_\beta(i), i = 1, \ldots, n$ be two sequences of random variables with stationary distribution functions $F$ and $F_\beta$, respectively. $F_\beta$ may depend on $n$. Assume that $F$ has bounded derivative in some neighborhood $B_p$ of $\xi_p$ with $F(\xi_p) = p$. Assume also that

\begin{equation}
\max_{1 \leq i \leq n} |X(i) - X_\beta(i)| \leq c \cdot \epsilon^* \cdot n^{-\rho} \text{ a.s.}
\end{equation}

where $\rho$ is a positive constant and $\epsilon^*$ is a random variable independent of $n$ and such that

\begin{equation}
P(\epsilon^* \geq \xi) \leq c \cdot \xi^{-\alpha}
\end{equation}

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

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

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

\[
P(X(i) \leq y - n^{-\lambda}) - P(\epsilon^* \geq \frac{1}{c} n^{\rho - \lambda}) \leq P(X_\beta(i) \leq y) \leq P(X(i) \leq y + n^{-\lambda}) + P(\epsilon^* \geq \frac{1}{c} n^{\rho - \lambda})
\]

Using (A.2), it is clear that

\[
F(y - n^{-\lambda}) - O(n^{-\alpha(\rho - \lambda)}) \leq F_\beta(y) \leq F(y + n^{-\lambda}) + O(n^{-\alpha(\rho - \lambda)})
\]

and consequently, since $F$ has bounded derivative over $B_p$,

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

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

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

This completes the proof.

Theorem 2 utilizes this Lemma with $\rho = -\beta(1 + 1 / \alpha - q) - 1 / \alpha$. 
Lemma 2: Let $J = \{X(i): i \in \{1, 2, \ldots, n\}\}$ be a set of random variables and $S_{n,k}$ $k = 1, \ldots, r$ be $r$ nonempty disjoint subsets of cardinality $n_k$ of the set $J$ with $\bigcup_{k=1}^{r} S_{n,k} = J$.

Then, for any $0 < p < 1$, the $p$-th sample quantile $X_{p,n}$ of $J$ and the $p$-th sample quantiles $X_{p, n_k}$ of $S_{n,k}$ satisfy the inequalities

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

Proof: Since $\lceil n_k \cdot p \rceil \geq n_k \cdot p > \lceil n_k \cdot p \rceil - 1$ for all $k = 1, \ldots, r$ it is true that

$$\sum_{k=1}^{r} (\lceil n_k \cdot p \rceil - 1) < [n \cdot p]$$

and hence

(A.3) \[\# \{X(i): X(i) < \min_{1 \leq k \leq r} X_{p, n_k}\} \geq \sum_{k=1}^{r} (\lceil n_k \cdot p \rceil - 1) < [n \cdot p]\]

and

(A.4) \[\# \{X(i): X(i) \leq \max_{1 \leq k \leq r} X_{p, n_k}\} \geq \sum_{k=1}^{r} \lceil n_k \cdot p \rceil \geq [n \cdot p].\]

(A.3) implies that $\min_{1 \leq k \leq r} X_{p, n_k} \leq X_{p,n}$ and (A.4) implies that $\max_{1 \leq k \leq r} X_{p, n_k} \geq X_{p,n}$.

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 $v_1 \neq v_2$. Then, for any $0 < \beta < 1$,

$$\limsup_{n \to \infty} \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \max_{1 \leq k \leq r} \left| \sum_{v=1}^{n} \frac{E_{n,k,v} - 1}{v} \right| < \infty \text{ a.s.,}$$

where the $n_k$ are defined after (2.3a) in Section 2.

Proof: We start by deriving sharp bounds for

(A.5) \[P \left( \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \sum_{v=1}^{n} \frac{E_{n,k,v} - 1}{v} \geq M \right) \]
Using Chernoff's bound (Chernoff (1952)) we obtain that this probability is bounded by

\[
\left[ \prod_{v \in \mathbb{P} \cdot n_k} \frac{1}{1-t/v} \right] \cdot \exp \left( -tMn^{1/2} \cdot (\log n)^{1/2} \right)
\]

for all \( 0 \leq t \leq \left[ \mathbb{P} \cdot n_k \right] \). We choose \( t = c_3 n^{1/2} \cdot (\log n)^{1/2} \) with some positive constant \( c_3 \) to be determined later. Then

\[
\log \left[ \prod_{v \in \mathbb{P} \cdot n_k} \frac{1}{1-t/v} \right] = \sum_{v \in \mathbb{P} \cdot n_k} -\log (1 - \frac{t}{v}) - \frac{t}{v} = \sum_{v \in \mathbb{P} \cdot n_k} \frac{t^2}{2v} + o(n^{-1/2+\varepsilon}), \quad \varepsilon > 0
\]

\[
\leq \frac{c_3^2 n^{1-\beta} \log n}{2p \cdot n_k} + o(n^{-1/2+\varepsilon})
\]

\[
\leq \frac{c_3^2}{2p} \log n \quad \text{for } n \text{ large enough.}
\]

Hence

\[
(A.6) \quad P \left[ \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \sum_{v \in \mathbb{P} \cdot n_k} \frac{E_{n,k,v} - 1}{v} \geq M \right] \leq n^{c_3^2/2p - c_3 M}
\]

Taking \( c_3 = p \cdot M \) we may make, for given \( p \), the exponent of \( n \) on the RHS 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 \( n^{-p \cdot M/2} = \psi(n,p,M) \) say.

The same argument can be applied to

\[
(-1) \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \sum_{v \in \mathbb{P} \cdot n_k} \frac{E_{n,k,v} - 1}{v}.
\]

Combining both, it is obvious that

\[
P \left[ \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \sum_{v \in \mathbb{P} \cdot n_k} \frac{E_{n,k,v} - 1}{v} \geq M \right] \leq 2\psi(n,p,M)
\]
and the bound is independent of k. Exploiting this uniformity and choosing $M$ large enough we get

\[
(A.7) \quad \sum_{n=N_0}^{\infty} P \left( \max_{1 \leq k \leq [n^{\beta}] \left( \frac{n^{1-\beta}}{\log n} \right)^{1/2} \left| \sum_{v=p \cdot n_0}^{n_k} \frac{E_{n,k,v} - 1}{v} \right| \geq M \right) < \infty
\]

and the Borel-Cantelli lemma produces the desired result.

**Lemma 4:** With the notation of Section 2

\[
\sup_{y \in J_n} |F_n(y) - F_n^{\beta_0^*}(y)| = O(n^{-\alpha \eta_0^*})
\]

where $J_n = (\xi_p - a_n^0, \xi_p + a_n^0)$, and $a_n^0$, $\beta_0$, $\beta_0^*$ as in Remark 4.

**Proof:** Choose $r$ such that

\[
\frac{\alpha}{\alpha + 1} (\eta_0^* - \eta_0) < r < \frac{\alpha}{\alpha + 1} \eta_0, \quad d_n = \lceil c n^r \rceil
\]

define $y_{n,v} = \xi_p + a_n^0 d_n^{-1} \cdot v$.

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

\[
F_n(y_{n,v}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X(i) \leq y_{n,v})
\]

is upperbounded a.s. by

\[
\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_{\beta_0^*}(i) \leq y_{n,v} + a_n^0 d_n^{-1}) + \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_{\beta_0^*}(i) - X(i) > a_n^0 d_n^{-1})
\]

Similarly,

\[
F_n(y_{n,v}) \geq \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_{\beta_0^*}(i) \leq y_{n,v} - a_n^0 d_n^{-1}) - \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X(i) - X_{\beta_0^*}(i) > a_n^0 d_n^{-1})
\]

so that

\[
(A.8) \quad |F_n(y_{n,v}) - F_n^{\beta_0^*}(y_{n,v})| \leq \max \{ |F_n^{\beta_0^*}(y_{n,v} + a_n^0 d_n^{-1}) - F_n^{\beta_0^*}(y_{n,v})| \}

+ \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(|X(i) - X_{\beta_0^*}(i)| > a_n^0 d_n^{-1})
\]

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

It is also easy to show that

$$\sup_{y \in \mathcal{F}_n} |F_n(y) - F_n^{\beta_0^*}(y)| \leq \max_{|v| \leq d_n} |F_n(y_{n,v}) - F_n^{\beta_0^*}(y_{n,v})|$$

$$+ \max_{-d_n \leq v < d_n} |F_n^{\beta_0^*}(y_{n,v+1}) - F_n^{\beta_0^*}(y_{u,v})|$$

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

$$\sup_{y \in \mathcal{F}_n} |F_n(y) - F_n^{\beta_0^*}(y)| \leq 2 \cdot \max_{|v| \leq d_n} |F_n^{\beta_0^*}(y_{n,v+1}) - F_n^{\beta_0^*}(y_{n,v})|$$

$$+ \frac{1}{n} \sum_{i=1}^{n} \chi (|X(i) - X_{\beta_0^*}(i)| > d_n^{-1})$$

$$= 2 S(1) + S(2), \text{ say.}$$

To show the required rate for $S(1)$ one makes use of the fact that

$n (F_n^{\beta_0^*}(y_{n,v+1}) - F_n^{\beta_0^*}(y_{n,v}))$ has the same distribution as

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

and $\chi (y_{n,v} < X_{\beta_0^*}(i) \leq y_{n,v+1}), i = 1, \ldots, n$ is a sequence of $[n^{\beta_0^*}]$-dependent Bernoulli random variables with parameter equal to

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

The inequalities in Hoeffding (1963), Section 5d admit a straightforward extension to $[n^{\beta_0^*}]$-dependent random variables and, using these inequalities together with the Borel Cantelli lemma proves the required rate for $S(1)$.

To show that

$$\frac{1}{n} \sum_{i=1}^{n} \chi (|X(i) - X_{\beta_0^*}(i)| > a_n^0 d_n^{-1}) = O(n^{-\alpha \alpha' + \eta_0^*})$$

it is sufficient to realize that
\[ \max_{1 \leq i \leq n} |X(i) - X_{\alpha_n}(i)| \leq c \cdot \varepsilon^* n^{-\eta_0^*} \text{ a.s.} \]

by (2.1), that \( a_n^0 d_n^{-1} > c \varepsilon^* n^{-\eta_0^*} \) for all sufficiently large \( n \), and that \( \varepsilon^* \) is a.s. finite.

**Lemma 5:** With the notation of Section 2 it is true for all \( \beta \) with \( 0 < \beta < 1 \) that

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

Here \( B_p \) is the fixed neighborhood of \( \xi_p \) 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 \( \delta > 0 \) it is clear that

\[ F(\xi_p - \delta) < p < F(\xi_p + \delta) \]

Now, if we can also show that

\[ (A.9) \quad F_n^\beta(\xi_p - \delta) \to F(\xi_p - \delta) \text{ a.s.} \]
\[ F_n^\beta(\xi_p + \delta) \to F(\xi_p + \delta) \text{ a.s.} \]

then it follows that

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

and therefore

\[ \xi_p - \delta \leq (X_{\beta})_{p,n} \leq \xi_p + \delta \text{ a.s. for all } n \text{ large enough,} \]

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

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

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

i.e. for all \( \varepsilon > 0 \)

\[ (A.10) \quad \sum_{n} P(\{|n^{-1} \sum_{i=1}^{n} \chi(X_{\beta}(i) \leq \xi_p - \delta) - F(\xi_p - \delta)| > \varepsilon\} < \infty \]

The LHS of (A.10) is upperbounded by

\[ (A.11) \quad \sum_{n} P(\{|n^{-1} \sum_{i=1}^{n} \chi(X_{\beta}(i) \leq \xi_p - \delta) - F(\xi_p - \delta)| > \varepsilon - c n^{-\eta} \}) \]
by (2.3). By construction $\chi(X_p(i) \leq \xi_p - \delta)$, $i = 1, \ldots, n$ is an $[n^\delta]$-dependent sequence of Bernoulli random variables and again we may use the results in (Hoeffding 1963 Section 5d) to prove the finiteness of (A11).

References:


Department of Statistics

University of California

Berkeley, CA 94720


18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.


32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.


43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?


56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.


64. O’SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.


70. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.


76. O’SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.


80. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.


90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.
95. CANCELLED


114. RITOY, Y. (September 1987). Estimation in a linear regression model with censored data.


118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.


127. Same as No. 133.

128. DONHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.


131. Same as No. 140


166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.
172. ADLER, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:
Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.