A note on the Bahadur representation of sample quantiles for mixing random variables*

Shanchao Yang, Keming Yu†

Address of the authors
Department of Mathematical Sciences, Brunel University, UK
e-mail: scyang@mailbox.gxnu.edu.cn keming.yu@brunel.ac.uk

Abstract: Bahadur representation and its applications have attracted a large number of publications and presentations on a wide variety of problems. Mixing dependency is weak enough to describe the dependent structure of random variables, including observations in time series and longitudinal studies. This note proves the Bahadur representation of sample quantiles for strongly mixing random variables (including ρ-mixing and ϕ-mixing) under very weak mixing coefficients. As application, the asymptotic normality is derived. These results greatly improves those recently reported in literature.

AMS 2000 subject classifications: Primary 60F05, 60F17; secondary 60G35.
Keywords and phrases: Bahadur's representation, Central limit theorem, mixing random variable, Sample quantile.

1. Introduction

The sample quantiles, often used as the estimator of distribution quantiles, are very important statistics. The applications of sample quantiles are beyond the area of statistics. For example, Dowd (2001) used sample quantile to derive a simple and accurate estimates of parametric Value-at-Risk (VaR) for financial risk analysis.

Let X be a random variable with a continuous distribution function $F(x)$ and a density function $f(x)$. For $0 < p < 1$, we call q_p the p-quantile of $F(x)$ if $F(q_p) = p$. Let $\{X_t\}_{t=1}^n$ be a sample drawn from the population X. We define the sample p-quantile as

$$Z_{n,p} = X_{[np]+1}, \tag{1.1}$$

where $[s]$ denotes the largest integer m such that $m \leq s$. The corresponding empirical distribution function is

$$F_n(x) = n^{-1} \sum_{i=1}^{n} I(X_i \leq x) \tag{1.2}$$

*This research is supported by the National Science Foundation of China(10161004) and the National Science Foundation of Guanxi(04047033).
†Correspondence author: E-mail:keming.yu@brunel.ac.uk

where $I(\cdot)$ is an indicator function.

Bahadur (1966) first established an almost sure representation of sample quantile for independent random variables and its convergent rate is $O(n^{-3/4}\log^{1/2} n \log_2^{1/4} n)$, where $\log_2 n = \log\log n$. Later, Kiefer (1967) proved the exact rate is $O(n^{-3/4}\log_2^{3/4} n)$. The Bahadur representation allows one to express asymptotically a sample quantile as a sample mean of certain (bounded) random variables, from which many important properties of the sample quantile, e.g., the central limit theorem, the law of iterated logarithm, may be easily proved.

Bahadur representation and its applications have attracted a large number of publications and presentations on a wide variety of problems. Hall (1991) developed it under bootstrap resampling. Xiang (1995), among others, derived it for kernel quantile estimator. He and Shao (1996) discussed Bahadur representation of M-estimation. Several researchers (Jureckova and Sen, 1984; Koenker and Portnoy,1987) derived it for parametric quantile regression. Further, Chaudhuri (1991), and references therein, extended it for conditional kernel quantile regression. Wu (2005) dealt with it for stochastic processes. Chen and Qin (1993) explored it for empirical likelihood sample quantiles. Besides its well-known applications in asymptotic theory, Bahadur representation has been applied in parameter estimation of complex models (Humphreys and Titterington, 2000), longitudinal analysis (Fitzmaurice, Laird and Rotnitzky, 1993, among others) and spatial data analysis (Koltchinskii, 1994).

Mixing dependent structure of random variables has been proved to be suitable to describe most of time series models, in particular, financial time series models. For example, Chanda (1974) first presented that linear stochastic process is strong mixing. Gorodeskii (1977) showed that linear process is strong mixing under certain conditions and also provided the convergent rate for the strong mixing coefficient. Withers (1981) further gave an alternative set of conditions for linear processes to be strong mixing, and proved the strong mixing coefficient is polynomial decay under some conditions. Pham and Tran (1985) studied some sufficient conditions for strong mixing linear process. In recent years, Genon-Catalot, Jeantheau and Laredo (2000) proved that continuous time diffusion models and stochastic volatility models are strong mixing under certain conditions. These models are the most popular models in pricing theory of financial assets, such as Black-Scholes pricing theory of option. The geometrically ergodic properties of various time series processes (including ARCH processes and Markov processes), which imply β-mixing and hence strong mixing, have been widely investigated by Lu (1996), Carrasco and Chen (2002), Lee and Shin (2004), Hwang and Kim (2004), Francq and Zakoian (2006), and references therein. Therefore, the mixing dependent structures of stochastic process have widely been concerned.

In view of Bahadur representation for mixing processes, the early work in the area includes the study of ϕ-mixing stationary processes by Sen (1972) and Bahu and Singh (1978). Later work included Yoshihara (1995) which derived the Bahadur representation by assuming that the random variables are uniformly bounded and the strong mixing coefficient $\alpha(n) = O(n^{-\beta})$ where $\beta > 5/2$, that is

$$Z_{n,p} - q_p = (p - F_n(q_p))/f(q_p) + O\left(n^{-3/4}\log n\right), a.s. \tag{1.3}$$

Unfortunately, we think this proof contains an error, i.e. the crucial inequality (20) of Yoshihara (1995) is untrue. For the sake of convenience, the inequality can be stated follows: let $\theta_n = n^{-1}$ and $m_n = [n^{1/2}\log n] + 1$, then for any ℓ $(\ell\theta_n < 1)$,

$$E\left|\sum_{i=1}^{n} W_{i,\ell}\right|^4 \leq C(n\ell\theta_n)^{1+\gamma} \tag{1.4}$$

for some γ $(0 < \gamma < 1)$, where

$$W_{i,\ell} = I(q_p < X_i \leq q_p + \ell\theta_n) - P(q_p < X_i \leq q_p + \ell\theta_n). \tag{1.5}$$

As a counter example of (1.4), we suppose that $\{X_i : i \geq 1\}$ are strictly stationary and independent random variables. Since $EW_{i,\ell} = 0$, we have

$$E\left|\sum_{i=1}^{n} W_{i,\ell}\right|^4 = nEW_{1,\ell}^4 + n(n-1)(EW_{1,\ell}^2)^2.$$

It is easy to obtain that $EW_{1,\ell}^2 = p_\ell(1-p_\ell)$ and $EW_{1,\ell}^4 = p_\ell(1-4p_\ell+6p_\ell^2-3p_\ell^3)$, where $p_\ell = Pr(q_p < X_j \leq q_p + \ell\theta_n)$. Hence

$$\begin{aligned} E\left|\sum_{i=1}^{n} W_{i,\ell}\right|^4 &= np_\ell(1-4p_\ell+6p_\ell^2-3p_\ell^3) + n(n-1)[p_\ell(1-p_\ell)]^2 \\ &= (np_\ell)^2(1-p_\ell)^2 + np_\ell(1-5p_\ell+8p_\ell^2-4p_\ell^3) \\ &= (np_\ell)^2(1-p_\ell)^2 - np_\ell + np_\ell(2-5p_\ell+8p_\ell^2-4p_\ell^3) \\ &\geq (np_\ell)^2(1-p_\ell)^2 - np_\ell \end{aligned}$$

by the fact: $2-5p_\ell+8p_\ell^2-4p_\ell^3 \geq 0$ for any p_ℓ $(0 \leq p_\ell \leq 1)$.

Under the conditions of Theorem 1 of Yoshihara (1995), there exist positive constants c_1 and c_2 such that $c_1\ell\theta_n \leq p_\ell \leq c_2\ell\theta_n$. Note that $\ell\theta_n < 1$ for $1 \leq \ell \leq m_n$. Let us consider the case $\ell = m_n$. Then $p_\ell \leq c_2\ell\theta_n = c_2 m_n\theta_n \to 0$ and $np_\ell \geq c_1 n\ell\theta_n \geq c_1 n\ell\theta_n = c_1 m_n \to \infty$. Thus, for n sufficiently large, we have that

$$E\left|\sum_{i=1}^{n} W_{i,\ell}\right|^4 \geq (np_\ell)^2(1-p_\ell)^2 - np_\ell \geq (np_\ell)^2/2 > (np_\ell)^{1+\gamma}, \tag{1.6}$$

which implies that (1.4) doesn't hold. Therefore, the inequality (20) of Yoshihara (1995) is not true. In fact, for generally independent random variables, by Rosenthal type moment inequality one can only prove that

$$E\left|\sum_{i=1}^{n} W_{i,\ell}\right|^r \leq C(n\ell\theta_n)^{r/2}. \tag{1.7}$$

Of course, Appendix B of this note show that (1.7) holds also for ρ-mixing and ϕ-mixing random variables, see (4.4) in Appendix B for details. Up to now, however, it is still unclear if (1.7) is true for α-mixing random variables. Therefore, the proof of Theorem 1 in Yoshihara (1995) can't be corrected for α-mixing random variables although it can for ϕ-mixing random variables.

Without noticing the error of Yoshihara's proof Sun (2006) tried to improve the result (Yoshihara, 1995) by removing the bound restriction. Sun (2006) stated his result as, for any $\delta \in (\frac{11}{4(\beta+1)}, \frac{1}{4})$ and $\beta > 10$

$$Z_{n,p} - q_p = (p - F_n(q_p))/f(q_p) + O\left(n^{-3/4+\delta} \log n\right), a.s. \tag{1.8}$$

However, this result requires the order of strong mixing coefficient to be $\beta > 10$ and reduces the rate of convergence by $\delta > \frac{11}{4(\beta+1)}$.

This note proves the Bahadur representation under weaker condition and provides a better convergent rate than (1.8). Our results don't need to assume that the random variables are uniformly bounded either.

2. Main results

We first give the definitions of mixing sequences. Let

$$\alpha(n) = \sup_{k\geq 1} \sup_{A\in\mathcal{F}_1^k, B\in\mathcal{F}_{k+n}^\infty} |P(AB) - P(A)P(B)|,$$

$$\phi(n) = \sup_{k\geq 1} \sup_{A\in\mathcal{F}_1^k, B\in\mathcal{F}_{k+n}^\infty, P(A)>0} |P(B) - P(B|A)|,$$

$$\rho(n) = \sup_{k\geq 1} \sup_{X\in L^2(\mathcal{F}_1^k), Y\in L^2(\mathcal{F}_{k+n}^\infty)} |\text{corr}(X,Y)|,$$

where $\mathcal{F}_1^k = \sigma(X_j, 1 \leq j \leq k)$ and $\mathcal{F}_{k+n}^\infty = \sigma(X_j, j > k+n)$.

The sequence $\{X_i\}_{i\geq 1}$ is called α-mixing (or strong mixing), ϕ-mixing and ρ-mixing if $\lim_{n\to\infty} \alpha(n) = 0$, $\lim_{n\to\infty} \phi(n) = 0$ and $\lim_{n\to\infty} \rho(n) = 0$ respectively.

Assumption (A) *Let $\{X_i : i \geq 1\}$ be a strictly stationary sequence of random variables with a common distribution function $F(x)$, where $F(x)$ is absolutely continuous in some neighborhood of its p-quantile q_p and has a continuous density function $f(x)$ such that $0 < f(q_p) < \infty$.*

This paper will derive the following results.

Theorem 2.1 *Suppose that Assumption (A) holds, $0 < \tau \leq 1$ and one of the following conditions holds:*

(i) $\{X_i : i \geq 1\}$ is α-mixing and $\alpha(n) = O(n^{-\beta})$ for some $\beta > 1/\tau$;

(ii) $\{X_i : i \geq 1\}$ is ρ-mixing and $\sum_{j=1}^\infty \rho^{2/r}(2^j) < \infty$ where $r > 2/\tau$;

(iii) $\{X_i : i \geq 1\}$ is ϕ-mixing and $\sum_{j=1}^\infty \phi^{1/r}(2^j) < \infty$ where $r > 2/\tau$.

Then, as $n \to \infty$,

$$Z_{n,p} - q_p = o\left(n^{-1/2} \log^\tau n\right) \ a.s., \tag{2.1}$$

Remark 2.1 When $\tau = 1$, we only need $\beta > 1$ for strong mixing sequence. On the other hand, $\alpha(n) = O(n^{-\beta})$ for some $\beta > 1$ is equivalent to $\sum_{j=1}^{\infty} \alpha^{\delta/(2+\delta)}(j) < \infty$ for some $\delta > 0$, which, as far as we know, is the weakest condition for asymptotic properties of strong mixing sequences. In addition, the condition $\sum_{j=1}^{\infty} \phi^{1/r}(2^j) < \infty$, where $r > 2/\tau$, is much weaker than $\phi(n) = O(n^{-2})$ in Yoshihara (1995).

Theorem 2.2 *Suppose the conditions in Theorem 2.1 are satisfied. For α-mixing random variables, assume further that (1.7) holds for $r > 2/\tau$, $1 \le \ell \le \lambda_n$, $\theta_n = n^{-3/4}\log^{\tau} n$ and $\lambda_n = [n^{1/4}]$. Then, as $n \to \infty$,*

$$\sup_{x \in J_n} |(F_n(x) - F(x)) - (F_n(q_p) - p)| = O\left(n^{-3/4}\log^{\tau} n\right) \ a.s., \tag{2.2}$$

where $J_n = \{x : |x - q_p| \le n^{-1/2}\log^{\tau} n\}$.

Theorem 2.3 *Suppose the conditions in Theorem 2.2 are satisfied and $f'(x)$ is bounded in some neighborhood of q_p. Then, as $n \to \infty$,*

$$Z_{n,p} - q_p = (p - F_n(q_p))/f(q_p) + O\left(n^{-3/4}\log^{\tau} n\right) \ a.s.. \tag{2.3}$$

Although it is unclear if (1.7) is till true for α-mixing, we are able to derive some further results for α-mixing.

Theorem 2.4 *Suppose that Assumption (A) holds and $\{X_i : i \ge 1\}$ is α-mixing with geometric mixing coefficients. Then as $n \to \infty$,*

$$\sup_{x \in J_n} |(F_n(x) - F(x)) - (F_n(q_p) - p)| = O\left(n^{-3/4}\log^{1+\tau} n\right) \ a.s., \tag{2.4}$$

for any $\tau > 0$. Further, assume that $f'(x)$ is bounded in some neighborhood of q_p, then as $n \to \infty$,

$$Z_{n,p} - q_p = (p - F_n(q_p))/f(q_p) + O\left(n^{-3/4}\log^{1+\tau} n\right) \ a.s., \tag{2.5}$$

for any $\tau > 0$.

Theorem 2.5 *Suppose that Assumption (A) holds and $\{X_i : i \ge 1\}$ is α-mixing with mixing coefficient $\alpha(n) = O(n^{-\beta})$ for $\beta > 3$. Then as $n \to \infty$,*

$$\sup_{x \in J_n} |(F_n(x) - F(x)) - (F_n(q_p) - p)| = O\left(n^{-\frac{3}{4} + \frac{7}{4(2\beta+1)}}\log n\right) \ a.s., \tag{2.6}$$

for any τ $(1/\beta < \tau < 1)$. Further, assume that $f'(x)$ is bounded in some neighborhood of q_p, then as $n \to \infty$,

$$Z_{n,p} - q_p = (p - F_n(q_p))/f(q_p) + O\left(n^{-\frac{3}{4} + \frac{7}{4(2\beta+1)}}\log n\right) \ a.s.. \tag{2.7}$$

Remark 2.2 Note that $-\frac{3}{4}+\frac{7}{4(2\beta+1)}<0$ for $\beta>3$. The result (2.7) improves greatly the (1.8) of Sun (2006).

Now, we discuss the asymptotic normality of sample quantiles for strong mixing sequence by the Bahadur representations. From strictly stationary,

$$\begin{aligned}\sigma_{p,n}^2 &=: E\{\sqrt{n}(p-F_n(q_p))\}^2\\ &= \mathrm{Var}(I(X_1\le q_p))+2\sum_{j=1}^{n-1}(1-j/n)\mathrm{Cov}(I(X_1\le q_p),I(X_{j+1}\le q_p))\\ &= p(1-p)+2\sum_{j=1}^{n-1}(1-j/n)\mathrm{Cov}(I(X_1\le q_p),I(X_{j+1}\le q_p)). \qquad (2.8)\end{aligned}$$

and by Lemma 3.7 in Appendix A,

$$\begin{aligned}&\left|\sum_{j=1}^{n-1}(1-j/n)\mathrm{Cov}(I(X_1\le q_p),I(X_{j+1}\le q_p))\right|\\ &\le \sum_{j=1}^{n-1}(1-j/n)\left|\mathrm{Cov}(I(X_1\le q_p),I(X_{j+1}\le q_p))\right|\\ &\le C\sum_{j=1}^{n-1}\alpha(j)\le C\sum_{j=1}^{n-1}j^{-\beta}. \qquad (2.9)\end{aligned}$$

It implies that $\sigma_{p,n}^2$ converges absolutely as $n\to\infty$ under the conditions of Theorem 2.1 and say $\sigma_p^2=\lim_{n\to\infty}\sigma_{p,n}^2$. Let

$$U_n=\sqrt{n}f(q_p)(Z_{n,p}-q_p)/\sigma_p. \qquad (2.10)$$

From (2.7),

$$U_n=\sqrt{n}(p-F_n(q_p))/\sigma_p+O\left(n^{-1/4+\frac{7}{4(2\beta+1)}}\log n\right)\ a.s., \qquad (2.11)$$

then we can establish the uniformly asymptotic normality of sample quantiles.

Theorem 2.6 *Suppose the conditions in Theorem 2.5 are satisfied and* $\sigma_p^2>0$. *Assume further that* $0<b<1$ *and*

$$\beta\ge\max\left\{1+\frac{1-b}{6b},\ \frac{7(1-b)}{6b}\right\}. \qquad (2.12)$$

Then for any $\epsilon>0$,

$$\sup_u|F_{U_n}(u)-\Phi(u)|=O\left(n^{-(1-b)/6}+n^{-1/4+\frac{7}{4(2\beta+1)}+\epsilon}\right) \qquad (2.13)$$

where Φ *is the standardize normal distribution function and* $F_X(x)$ *denotes the distribution function of any random variable* X.

Remark 2.3 We know that $1+\frac{1-b}{6b} \leq \frac{7(1-b)}{6b}$ for $0 < b \leq 1/2$, while $1+\frac{1-b}{6b} > \frac{7(1-b)}{6b}$ for $1/2 < b < 1$. Therefore, the low bound in (2.12) tends to 1 as $b \to 1^-$. Thus, we have

$$\sup_u |F_{U_n}(u) - \Phi(u)| = o(1) \tag{2.14}$$

provided with $\beta > 3$. Also, under geometric coefficient of strong mixing, the convergence rate of uniformly asymptotic normality is near to $n^{-1/6}$ by choosing $b \to 0$ and $\beta \to \infty$ in (2.13), that is

$$\sup_u |F_{U_n}(u) - \Phi(u)| = O(n^{-1/6+\epsilon}), \tag{2.15}$$

for any $\epsilon > 0$.

Appendix A

Auxiliary lemmas

There are some known lemmas that will be used in the next section.

Lemma 3.1 (Roussas and Ioannides, 1987) *Let $\{X_j : j \geq 1\}$ be a sequence of α-mixing random variables. Suppose that ξ and η are $\mathcal{F}_1^k$ - measurable and $\mathcal{F}_{k+n}^\infty$ - measurable random variables, respectively. (i) If $|\xi| \leq C_1$ a.s. and $|\eta| \leq C_2$ a.s., then*

$$|E(\xi\eta) - (E\xi)(E\eta)| \leq 4C_1C_2\alpha(n) \ ;$$

(ii) If $E|\xi|^p < \infty$ a.s. and $|\eta| \leq C$ a.s. with $1/p + 1/q = 1$, then

$$|E(\xi\eta) - (E\xi)(E\eta)| \leq 6C\alpha^{1/q}(n)(E|\xi|^p)^{1/p} \ ;$$

(iii) If $E|\xi|^p < \infty$ a.s. and $E|\eta|^q < \infty$ a.s. with $1/p + 1/q + 1/t = 1$, then

$$|E(\xi\eta) - (E\xi)(E\eta)| \leq 10\alpha^{1/t}(n)(E|\xi|^p)^{1/p}(E|\eta|^q)^{1/q} \ .$$

Lemma 3.2 *Let $\{X_i : i \geq 1\}$ be a sequence of α-mixing random variables with zero mean. If $E|X_i|^{2+\delta} < \infty$ for some $\delta > 0$ and $\sum_{j=1}^\infty \alpha^{\delta/(2+\delta)}(j) < \infty$, then there exists a positive constant C, which does't depend on n, such that*

$$E\left(\sum_{i=1}^n X_i\right)^2 \leq C\sum_{i=1}^n ||X_i||_{2+\delta}^2 \ .$$

It is easy to prove Lemma 3.2 by Lemma 3.1 (iii). In fact, it also be found in Yang (2006, Lemma 2.1; 2000, Theorem 2.1).

Lemma 3.3 (Yang, 2006, Theorem 2.2) *Let $\{X_i : i \geq 1\}$ be a sequence of α-mixing random variables with zero mean and $E|X_i|^{r+\delta} < \infty$ for some $r > 2$ and $\delta > 0$. If*

$$\beta > r(r+\delta)/(2\delta) \tag{3.1}$$

and $\alpha(n) \leq Cn^{-\beta}$ for some $C > 0$, then for any $\varepsilon > 0$, there exists a positive constant $K = K(\varepsilon, r, \delta, \beta, C) < \infty$ such that

$$E\max_{1\leq j\leq n}\left|\sum_{i=1}^j X_i\right|^r \leq K\left\{n^\varepsilon\sum_{i=1}^n E|X_i|^r + \left(\sum_{i=1}^n \left(E|X_i|^{r+\delta}\right)^{2/(r+\delta)}\right)^{r/2}\right\}.$$

Lemma 3.4(Shao, 1995, Corollary 1.1) *Let $\{X_i : i \geq 1\}$ be a sequence of ρ-mixing random variables with zero mean, $E|X_i|^r < \infty$ for some $r \geq 2$ and $\sum_{j=1}^{\infty} \rho^{2/r}(2^j) < \infty$. Then there exists a positive constant $K = K(r, \rho(\cdot)) < \infty$ such that*

$$E \max_{1\leq j\leq n} \left| \sum_{i=1}^{j} X_i \right|^r \leq K \left\{ n \max_{1\leq i\leq n} E|X_i|^r + \left(n \max_{1\leq i\leq n} E|X_i|^2 \right)^{r/2} \right\}.$$

Remark 3.1 Since $\rho(n) \leq 2\varphi^{1/2}(n)$, Lemma 3.2 also holds for the sequence of φ-mixing random variables with zero mean, $E|X_i|^r < \infty$ for some $r \geq 2$ and $\sum_{j=1}^{\infty} \varphi^{1/r}(2^j) < \infty$. Its direct proof can be found in Shao (1988).

Lemma 3.5 Let $\{X_i : i \geq 1\}$ be an α-mixing sequence of mean-zero real-valued random variables with $|X_i| \leq b < \infty$ *a.s.*. Suppose furthermore that k_n are positive integers such that $1 \leq k_n \leq n/2$. Then for any $\epsilon > 0$,

$$P\left(\left| \sum_{i=1}^{n} X_i \right| > n\epsilon \right) \leq 4 \exp\left(-\frac{n\epsilon^2}{12(3\sigma_n^2 + bk_n\epsilon)} \right) + 36b\epsilon^{-1}\alpha(k_n)$$

where $\sigma_n^2 = n^{-1} \sum_{j=1}^{2m_n+1} E|U_j|^2,\ m_n = [\frac{n}{2k_n}], U_j = \sum_{(j-1)k_n<i\leq jk_n} X_i\ (j = 1, 2, \cdots, 2m_n)$ and $U_{2m_n+1} = \sum_{i=2m_n+1}^{n} X_i$.

Proof Rio (1995) used the coupling method to prove a Bennett type exponential inequality (Theorem 5) for strong mixing random variables. Here we will use his method to show our Bernstein type exponential inequality. Obviously, $2k_n m_n \leq n$ and $n - 2k_n m_n \leq n - 2k_n(\frac{n}{2k_n} - 1) = 2k_n$. Also,

$$\sum_{i=1}^{n} X_i = \sum_{j=1}^{2m_n+1} U_j = \sum_{j=0}^{m_n} U_{2j+1} + \sum_{j=1}^{m_n} U_{2j}. \tag{3.2}$$

Note that $|U_j| \leq 2k_n b, a.s.$ for any j. From the proof of Theorem 5 in Rio (1995), there exist the random variables $\{U_j^*\}_{1\leq j\leq 2m_n+1}$ which have the following properties:

1. For any positive j, the random variable U_j^* has the same distribution as U_j.
2. The random variables $\{U_{2j}^*\}_{1\leq j\leq m_n}$ are independent and the random variables $\{U_{2j+1}^*\}_{0\leq j\leq m_n}$ are also independent.
3. Moreover,

$$\sum_{j=1}^{2m_n+1} E|U_j - U_j^*| \leq 12bn\alpha(k_n). \tag{3.3}$$

Now, from (3.2), we get that

$$\begin{aligned} P\left(\left| \sum_{i=1}^{n} X_i \right| > n\epsilon \right) &\leq P\left(\left| \sum_{j=1}^{m_n} U_{2j}^* \right| > n\epsilon/3 \right) + P\left(\left| \sum_{j=0}^{m_n} U_{2j+1}^* \right| > n\epsilon/3 \right) \\ &\quad + P\left(\left| \sum_{j=1}^{2m_n} (U_j - U_j^*) \right| > n\epsilon/3 \right) \\ &=: I_1 + I_2 + I_3. \end{aligned} \tag{3.4}$$

By (3.3) and Markov's inequality,

$$I_3 \le 36b\epsilon^{-1}\alpha(k_n). \tag{3.5}$$

By Bernstein exponential inequality for independent random variables, we have

$$I_1 \le 2\exp\left(-\frac{(n\epsilon/3)^2}{2\left(2\sum_{j=1}^{m_n} EU_{2j}^2 + 2k_n bn\epsilon/3\right)}\right) \le 2\exp\left(-\frac{n\epsilon^2}{12\,(3\sigma_n^2 + bk_n\epsilon)}\right). \tag{3.6}$$

In the same way,

$$I_2 \le 2\exp\left(-\frac{n\epsilon^2}{12\,(3\sigma_n^2 + bk_n\epsilon)}\right). \tag{3.7}$$

Combining (3.4)-(3.7) yields the desired result. ♯

Lemma 3.6 *Let θ be a positive real number and λ be a positive integer. If $f(x)$ is continuous and bounded (say b) in $[s-\lambda\theta, s+\lambda\theta]$, then*

$$\sup_{s\le t\le s+\lambda\theta} |Y_n(t) - Y_n(s)| \le 2\max_{1\le j\le\lambda} |Y_n(s+j\theta) - Y_n(s)| + 2b\sqrt{n}\theta \tag{3.8}$$

and

$$\sup_{s-\lambda\theta\le t\le s} |Y_n(t) - Y_n(s)| \le 2\max_{1\le j\le\lambda} |Y_n(s-j\theta) - Y_n(s)| + 2b\sqrt{n}\theta \tag{3.9}$$

where $Y_n(t) = \sqrt{n}(F_n(t) - F(t))$.

Proof It is similar to Lemma 3 of Yoshihara (1995), but it is not exactly the same. Denote $U(t) =: F(t) - F(s)$ and $U_n(t) =: F_n(t) - F_n(s)$. Note that $U(t)$ and $U_n(t)$ are non-decreasing for $t \in [s, \infty)$, and $Y_n(t) - Y_n(s) = \sqrt{n}[U_n(t) - U(t)]$, we have

$$\begin{aligned}
&\sup_{s\le t\le s+\lambda\theta} |Y_n(t) - Y_n(s)| = \max_{1\le j\le\lambda} \sup_{s+(j-1)\theta<t\le s+j\theta} \sqrt{n}|U_n(t) - U(t)| \\
&\le \max_{1\le j\le\lambda} \sup_{s+(j-1)\theta<t\le s+j\theta} \sqrt{n}|U_n(s+j\theta) - U(s+(j-1)\theta)| \\
&+ \max_{1\le j\le\lambda} \sup_{s+(j-1)\theta<t\le s+j\theta} \sqrt{n}|U_n(s+(j-1)\theta) - U(s+j\theta)| \\
&\le \max_{1\le j\le\lambda} \sqrt{n}|U_n(s+j\theta) - U(s+j\theta) + a_{n,j}| \\
&+ \max_{1\le j\le\lambda} \sqrt{n}|U_n(s+(j-1)\theta) - U(s+(j-1)\theta) - a_{n,j}| \\
&\le 2\max_{1\le j\le\lambda} \sqrt{n}|U_n(s+j\theta) - U(s+j\theta)| + 2\sqrt{n}a_{n,j} \\
&\le 2\max_{1\le j\le\lambda} |Y_n(s+j\theta) - Y_n(s)| + 2\sqrt{n}a_{n,j}
\end{aligned}$$

where $a_{n,j} = F(s+j\theta) - F(s+(j-1)\theta)$. Furthermore, $a_{n,j} \le b\theta$. So we've got (3.8). Similarly, (3.9) follows. ♯

Lemma 3.7(Moricz, 1976, Theorem 1) *Let $\{\xi_i\}$ be a sequence of random variables (It is not assumed that they are independent or stationary) and $g(F_{b,n})$ denote a non-negative function depending on the joint distribution function of $(\xi_{b+1}, \cdots, \xi_{b+n})$ and satisfying*

$$g(F_{b,k}) + g(F_{b+k,l}) \le g(F_{b,k+l}), \ \textit{for all } b \ge 0 \textit{ and } 1 \le k \le k+l.$$

If

$$E\left|\sum_{i=b+1}^{b+n}\xi_i\right|^r \le g^{\upsilon}(F_{b,n}),\ \textit{for all } b\ge 0, n\ge 1,$$

where $r>0$ and $\upsilon>1$. Then exists a positive constant $C_{r,\upsilon}$, which depends only on r and υ, such that

$$E\max_{1\le k\le n}\left|\sum_{i=b+1}^{b+k}\xi_i\right|^r \le C_{r,\upsilon}g^{\upsilon}(F_{b,n}),\ \textit{for all } b\ge 0, n\ge 1.$$

Lemma 3.8(Yang, 2003, Lemma 3.7) *Suppose that $\{\zeta_n : n\ge 1\}$ and $\{\eta_n : n\ge 1\}$ are two random variable sequences, $\{\gamma_n : n\ge 1\}$ is a positive constant sequence, and $\gamma_n\to 0$. If*

$$\sup_u |F_{\zeta_n}(u)-\Phi(u)|\le C\gamma_n,$$

then for any $\varepsilon>0$,

$$\sup_u |F_{\zeta_n+\eta_n}(u)-\Phi(u)|\le C\{\gamma_n+\varepsilon+P(|\eta_n|\ge\varepsilon)\}.$$

Lemma 3.9(Yang and Li, 2006, Lemma 3.2) *Let $\{X_j : j\ge 1\}$ be a sequence of α-mixing random variables, p,q be two positive integers. Denote $\eta_l := \sum_{j=(l-1)(p+q)+1}^{(l-1)(p+q)+p} X_j$ for $1\le l\le k$. If $r>0, s>0$ and $\frac{1}{r}+\frac{1}{s}=1$, then*

$$\left|E\exp\left(it\sum_{l=1}^{k}\eta_l\right)-\prod_{l=1}^{k}E\exp\left(it\eta_l\right)\right|\le C|t|\alpha^{1/s}(q)\sum_{l=1}^{k}||\eta_l||_r.$$

Proof It is Lemma 3.2 of Yang and Li (2006). For the sake of convenience, we give the proof. Obviously

$$\begin{aligned}
&\left|E\exp\left(it\sum_{l=1}^{k}\eta_l\right)-\prod_{l=1}^{k}E\exp\left(it\eta_l\right)\right|\\
&\le \left|E\exp\left(it\sum_{l=1}^{k}\eta_l\right)-E\exp\left(it\sum_{l=1}^{k-1}\eta_l\right)E\exp\left(it\eta_k\right)\right|\\
&+\left|E\exp\left(it\sum_{l=1}^{k-1}\eta_l\right)-\prod_{l=1}^{k-1}E\exp\left(it\eta_l\right)\right|\\
&=: I_1+I_2.
\end{aligned}$$

Note that $e^{ix}=\cos(x)+i\sin(x)$, $\sin(x+y)=\sin(x)\cos(y)+\cos(x)\sin(y)$, $\cos(x+y)=\cos(x)\cos(y)-\sin(x)\sin(y)$. We have

$$\begin{aligned}
I_1 \quad\le\quad & \left|\mathrm{cov}\left(\cos\left(\mathrm{t}\sum_{\mathrm{l}=1}^{\mathrm{k}-1}\eta_{\mathrm{l}}\right),\cos(\mathrm{t}\eta_{\mathrm{k}})\right)\right|+\left|\mathrm{cov}\left(\sin\left(\mathrm{t}\sum_{\mathrm{l}=1}^{\mathrm{k}-1}\eta_{\mathrm{l}}\right),\sin(\mathrm{t}\eta_{\mathrm{k}})\right)\right|\\
+\quad & \left|\mathrm{cov}\left(\sin\left(\mathrm{t}\sum_{\mathrm{l}=1}^{\mathrm{k}-1}\eta_{\mathrm{l}}\right),\cos(\mathrm{t}\eta_{\mathrm{k}})\right)\right|+\left|\mathrm{cov}\left(\cos\left(\mathrm{t}\sum_{\mathrm{l}=1}^{\mathrm{k}-1}\eta_{\mathrm{l}}\right),\sin(\mathrm{t}\eta_{\mathrm{k}})\right)\right|\\
=:\quad & I_{11}+I_{12}+I_{13}+I_{14}.
\end{aligned}$$

Using Lemma 3.1 and note that $|\sin(x)| \le |x|$, have

$$I_{12} \le C\alpha^{1/s}(q)||\sin(t\eta_k)||_r \le C|t|\alpha^{1/s}(q)||\eta_k||_r,\ I_{14} \le C|t|\alpha^{1/s}(q)||\eta_k||_r$$

Using $\cos(2x) = 1 - 2\sin^2(x)$, we obtain

$$\begin{aligned} I_{11} &= \left|\mathrm{cov}\left(\cos\left(\mathrm{t}\sum_{\mathrm{l}=1}^{\mathrm{k}-1}\eta_{\mathrm{l}}\right), 1-2\sin^2(\mathrm{t}\eta_{\mathrm{k}}/2)\right)\right| = 2\left|\mathrm{cov}\left(\cos\left(\mathrm{t}\sum_{\mathrm{l}=1}^{\mathrm{k}-1}\eta_{\mathrm{l}}\right), \sin^2(\mathrm{t}\eta_{\mathrm{k}}/2)\right)\right| \\ &\le C\alpha^{1/s}(q)E^{1/r}|\sin(t\eta_k/2)|^{2r} \le C\alpha^{1/s}(q)E^{1/r}|\sin(t\eta_k/2)|^r \le C|t|\alpha^{1/s}(q)||\eta_k||_r. \end{aligned}$$

Similarly,

$$I_{13} \le C|t|\alpha^{1/s}(q)||\eta_k||_r.$$

Combining the equations above yields

$$\left|E\exp\left(it\sum_{l=1}^{k}\eta_l\right) - \prod_{l=1}^{k}E\exp\left(it\eta_l\right)\right| \le C|t|\alpha^{1/s}(q)||\eta_k||_r + I_2.$$

From that, and repeating the procedure above, we obtain the result.

Appendix B

Proofs of Theorems

We will first prove the Bahadur representation (i.e. Theorem 2.1-2.5) and then show the uniformly asymptotic normality (i.e. Theorem 2.6).

4.1 Proof of Bahadur representation

Proof of Theorem 2.1 Clearly,

$$\begin{aligned} A_n &=: \{\omega : |Z_{n,p} - q_p| \ge \epsilon n^{-1/2}\log^\tau n\} \\ &= \{\omega : Z_{n,p} \le q_p - \epsilon n^{-1/2}\log^\tau n\} + \{\omega : Z_{n,p} \ge q_p + \epsilon n^{-1/2}\log^\tau n\} \\ &=: A_{1n} + A_{2n}. \end{aligned}$$

Also,

$$\begin{aligned} \bigcup_{2^k \le n < 2^{k+1}} A_{1n} &= \bigcup_{2^k \le n < 2^{k+1}} \left\{\sum_{i=1}^{n} I(X_i \le q_p - \epsilon n^{-1/2}\log^\tau n) \ge [np]+1\right\} \\ &\subseteq \bigcup_{2^k \le n < 2^{k+1}} \left\{\sum_{i=1}^{n} I(X_i \le q_p - c_1 2^{-k/2}k^\tau) \ge [np]+1\right\} \\ &\subseteq \bigcup_{2^k \le n < 2^{k+1}} \left\{\sum_{i=1}^{n} \xi_{k,i} \ge [np]+1 - nF(q_p - c_1 2^{-k/2}k^\tau)\right\} \end{aligned}$$

where $\xi_{k,i} = I(X_i \le q_p - c_1 2^{-k/2}k^\tau) - F(q_p - c_1 2^{-k/2}k^\tau)$. Note that $F(q_p - c_1 2^{-k/2}k^\tau) = p - c_1 f(q_p)2^{-k/2}k^\tau + O(2^{-k}k^{2\tau})$, we have

$$[np] + 1 - nF(q_p - c_1 2^{-k/2}k^\tau) = c_2 2^{-k/2}k^\tau n + O(2^{-k}k^{2\tau}n) \ge c_3 2^{k/2}k^\tau$$

for $2^k \le n < 2^{k+1}$. Hence

$$\bigcup_{2^k \le n < 2^{k+1}} A_{1n} \subseteq \bigcup_{2^k \le n < 2^{k+1}} \left\{ \sum_{i=1}^{n} \xi_{k,i} \ge c_3 2^{k/2} k^{\tau} \right\} \subseteq \left\{ \max_{1 \le n < 2^{k+1}} \left| \sum_{i=1}^{n} \xi_{k,i} \right| \ge c_3 2^{k/2} k^{\tau} \right\}.$$

For α-mixing, we can first choose a $r > 2$ such that $\beta > r/2 > 1/\tau$ since $\beta > 1/\tau \ge 1$, and then choose a $\delta > 0$ sufficiently large such that $\beta > r^2/(2\delta) + r/2$. These imply that $\beta > r(r+\delta)/(2\delta)$ and $r\tau/2 > 1$. By Lemma 3.3 and choosing an $\epsilon > 0$ sufficiently small, we have

$$\begin{aligned} P\left(\bigcup_{2^k \le n < 2^{k+1}} A_{1n} \right) &\le C 2^{-rk/2} k^{-r\tau} E \left(\max_{1 \le n < 2^{k+1}} \left| \sum_{i=1}^{n} \xi_{k,i} \right|^r \right) \\ &\le C 2^{-rk/2} k^{-r\tau} \left\{ 2^{\epsilon(k+1)} \sum_{i=1}^{2^{k+1}} E|\xi_{k,i}|^r + \left(\sum_{i=1}^{2^{k+1}} (E|\xi_{k,i}|^{r+\delta})^{2/(r+\delta)} \right)^{r/2} \right\} \\ &\le C 2^{-rk/2} k^{-r\tau} 2^{r(k+1)/2} \le C k^{-r\tau}. \end{aligned} \tag{4.1}$$

Similarly, $P\left(\bigcup_{2^k \le n < 2^{k+1}} A_{2n}\right) \le C k^{-r\tau}$. So $\sum_{k=1}^{\infty} P\left(\bigcup_{2^k \le n < 2^{k+1}} A_n\right) < \infty$. From the Borel-Cantelli lemma, we have that $P\left(\bigcup_{2^k \le n < 2^{k+1}} A_n, \text{i.o.}\right) = 0$, hence $P(A_n, \text{i.o.}) = 0$, i.e.

$$|Z_{n,p} - q_p| < \epsilon n^{-1/2} \log^{\tau} n, a.s.$$

for any $\epsilon > 0$. It implies the desired result. For ρ-mixing and φ-mixing, (4.1) is still true and followed by Lemma 3.4 and Remark 3.1 instead of Lemma 3.3. Thus we finish the proof. ♯

Proof of Theorem 2.2 Let $\lambda_n = [n^{1/4}]$ which is the integer part of $n^{1/4}$ and $\theta_n = n^{-3/4} \log^{\tau} n$. By Lemma 3.3, we have

$$\begin{aligned} &\sup_{x \in J_n} |(F_n(x) - F(x)) - (F_n(q_p) - p)| \\ &= n^{-1/2} \sup_{x \in J_n} |Y_n(x) - Y_n(q_p)| \\ &\le n^{-1/2} \sup_{q_p - \lambda_n\theta_n < x < q_p + \lambda_n\theta_n} |Y_n(x) - Y_n(q_p)| \\ &\le 2n^{-1/2} \max_{1 \le \ell \le \lambda_n} |Y_n(q_p + \ell\theta_n) - Y_n(q_p)| \\ &\quad + 2n^{-1/2} \max_{1 \le \ell \le \lambda_n} |Y_n(q_p - \ell\theta_n) - Y_n(q_p)| + C\theta_n \\ &\le 2n^{-1/2} \max_{1 \le \ell \le \lambda_n} |Y_n(q_p + \ell\theta_n) - Y_n(q_p)| \\ &\quad + 2n^{-1/2} \max_{1 \le \ell \le \lambda_n} |Y_n(q_p - \ell\theta_n) - Y_n(q_p)| + O\left(n^{-3/4} \log^{\tau} n\right). \end{aligned}$$

Hence, it suffices to show that

$$\max_{1 \le \ell \le \lambda_n} |Y_n(q_p + \eta\ell\theta_n) - Y_n(q_p)| = O\left(n^{-1/4} \log^{\tau} n\right). \tag{4.2}$$

where $\eta = 1$ or -1. We will only discuss the case of $\eta = 1$ because of its similarity for the case of $\eta = -1$. Clearly,

$$\begin{aligned}
&Y_n(q_p + \ell\theta_n) - Y_n(q_p)\\
&= \sqrt{n}(F_n(q_p + \ell\theta_n) - F(q_p + \ell\theta_n)) - \sqrt{n}(F_n(q_p) - F(q_p))\\
&= \sqrt{n}\left\{(F_n(q_p + \ell\theta_n) - F_n(q_p)) - (F(q_p + \ell\theta_n) - F(q_p))\right\}\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{I(q_p < X_i \le q_p + \ell\theta_n) - P(q_p < X_i \le q_p + \ell\theta_n)\right\}\\
&= n^{-1/2}\sum_{i=1}^{n} W_{i,\ell}.
\end{aligned}$$

where $W_{i,\ell}$ are defined in (1.5). Therefore, it remains to prove that

$$\max_{1\le\ell\le\lambda_n}\left|\sum_{i=1}^{n} W_{i,\ell}\right| = O\left(n^{1/4}\log^{\tau} n\right). \tag{4.3}$$

Denote $Z_{i,j} = I\left(q_p + (j-1)\theta_n \le X_i < q_p + j\theta_n\right) - P\left(q_p + (j-1)\theta_n \le X_i < q_p + j\theta_n\right)$. Clearly, for any given ℓ, $\{W_{i,\ell} : i \ge 1\}$ is still a strong mixing (ρ-mixing or ϕ-mixing) sequence with zero-mean and bounded, and $W_{i,\ell} = \sum_{j=1}^{\ell} Z_{i,j}$.

Now for the cases of ρ-mixing and φ-mixing, by Lemma 3.4 and Remark 3.1 it is easy to get

$$E\left|\sum_{i=1}^{n}\sum_{j=1}^{\ell} Z_{i,j}\right|^{r} = E\left|\sum_{i=1}^{n} W_{i,\ell}\right|^{r} \le C\left(nEW_{1,\ell}^{2}\right)^{r/2} \le C(n\ell\theta_n)^{r/2}. \tag{4.4}$$

For α-mixing, (4.4) just is (1.7). Furthermore, using Lemma 3.7 (If necessary, we can change the double sum into single sum by re-arrange the random variables $Z_{i,j}$), we have

$$E\max_{1\le k\le n}\max_{1\le\ell\le\lambda_n}\left|\sum_{i=1}^{k}\sum_{j=1}^{\ell} Z_{i,j}\right|^{r} \le C(n\lambda_n\theta_n)^{r/2}.$$

Therefore,

$$\begin{aligned}
&P\left(\bigcup_{2^k\le n<2^{k+1}}\left\{\max_{1\le\ell\le\lambda_n}\left|\sum_{i=1}^{n} W_{i,\ell}\right| \ge n^{1/4}\log^{\tau} n\right\}\right)\\
&\le P\left(\max_{2^k\le n<2^{k+1}}\max_{1\le\ell\le\lambda_n}\left|\sum_{i=1}^{n}\sum_{j=1}^{\ell} Z_{i,j}\right| \ge c2^{k/4}k^{\tau}\right)\\
&\le C(2^{k/4}k^{\tau})^{-r}(2^k\lambda_{2^k}\theta_{2^k})^{r/2} \le Ck^{-r\tau/2}
\end{aligned}$$

which, together with $r\tau/2 > 1$, yields (4.3). The proof is completed. ♯

Proof of Theorem 2.3 Note that

$$F_n(Z_{n,p}) = n^{-1}\sum_{i=1}^{n} I(X_i \le Z_{n,p}) = \frac{[np]+1}{n} = p + O(n^{-1}).$$

Expand $F(Z_{n,p})$ at q_p, we get that

$$F(Z_{n,p}) = p + f(q_p)(Z_{n,p} - q_p) + \frac{1}{2} f'(q_p + \theta(Z_{n,p} - q_p))(Z_{n,p} - q_p)^2.$$

Thus

$$\begin{aligned}&(F_n(Z_{n,p}) - F(Z_{n,p})) + f(q_p)(Z_{n,p} - q_p)\\ &= O(n^{-1}) - \frac{1}{2} f'(q_p + \theta(Z_{n,p} - q_p))(Z_{n,p} - q_p)^2.\end{aligned}$$

Since $f'(x)$ is bounded in some neighborhood of q_p, by Theorem 2.1

$$|(F_n(Z_{n,p}) - F(Z_{n,p})) + f(q_p)(Z_{n,p} - q_p)| = o(n^{-1} \log^{2\tau} n), a.s.,$$

which, together with Theorem 2.2, leads to

$$\begin{aligned}&|f(q_p)(Z_{n,p} - q_p) + (F_n(q_p) - p)|\\ &\le |(F_n(Z_{n,p}) - F(Z_{n,p})) + f(q_p)(Z_{n,p} - q_p)| + |(F_n(Z_{n,p}) - F(Z_{n,p})) - (F_n(q_p) - p)|\\ &\le |(F_n(Z_{n,p}) - F(Z_{n,p})) + f(q_p)(Z_{n,p} - q_p)| + \sup_{x \in J_n} |(F_n(x) - F(x)) - (F_n(q_p) - p)|\\ &= O\left(n^{-3/4} \log^{\tau} n\right),\ a.s..\end{aligned}$$

So the proof is completed. ♯

Proof of Theorem 2.4 Because of geometric mixing coefficient, (2.1) holds for any $\tau > 0$. To get (2.4), it is sufficient to prove that

$$\max_{1 \le \ell \le \lambda_n} \left| \sum_{i=1}^n W_{i,\ell} \right| = O\left(n^{1/4} \log^{1+\tau} n\right). \tag{4.5}$$

instead of (4.3). We will show it by using the similar method to that in the proof of Theorem 1 of Sun (2006). Now recall that the definition of σ_n^2 in Lemma 3.5, we have that

$$\sigma_n^2 = n^{-1} \sum_{j=1}^{m_n} E|U_j|^2 \le C k_n E W_{1,\ell}^2 \le k_n \ell \theta_n \le C k_n n^{-1/2} \log^{\tau} n. \tag{4.6}$$

Hence, applying Lemma 3.5 and choosing $k_n = [B \log n]$ for some positive constant B sufficiently large,

$$\begin{aligned}&P\left(\left|\sum_{i=1}^n W_{i,\ell}\right| > n^{1/4} \log^{1+\tau} n\right) = P\left(\left|\sum_{i=1}^n W_{i,\ell}\right| > n n^{-3/4} \log^{1+\tau} n\right)\\ &\le 4 \exp\left(-\frac{n^{-1/2} \log^{2+2\tau} n}{C k_n n^{-1/2} \log^{\tau} n + 12 k_n n^{-3/4} \log^{1+\tau} n}\right) + 36\alpha(k_n) n^{3/4} \log^{-1-\tau} n\\ &\le 4 \exp\left(-\frac{\log^{2+\tau} n}{C k_n + 12 k_n n^{-1/4} \log n}\right) + 36 n^{3/4} \rho^{k_n} \log^{-\tau} n \le C n^{-2},\end{aligned}$$

where $0 < \rho < 1$. Thus,

$$\left|\sum_{i=1}^n W_{i,\ell}\right| \le C n^{1/4} \log^{1+\tau} n, a.s.$$

for any $\ell(1 \le \ell \le \lambda_n)$. It follows (4.5). Finally, we get (2.5) by the same way in the proof of Theorem 2.3. So the proof is completed. ♯

Proof of Theorem 2.5 Since $\beta > 1/\tau$ and $0 < \tau < 1$, (2.1) holds. To get (2.6), hence, we only need to show that

$$\max_{1\le \ell\le \lambda_n}\left|\sum_{i=1}^n W_{i,\ell}\right| = O\left(n^{1/4+\epsilon}\log n\right),\ a.s. \tag{4.7}$$

instead of (4.3), where $\epsilon = \frac{7}{4(2\beta+1)}$. Recall (4.6) and choose $k_n = [n^{2\epsilon}\log^{(1-\tau)/2} n]$, by Lemma 3.5 we have

$$\begin{aligned}
&P\left(\left|\sum_{i=1}^n W_{i,\ell}\right| > n^{1/4+\epsilon}\log n\right) = P\left(\left|\sum_{i=1}^n W_{i,\ell}\right| > n\cdot n^{-3/4+\epsilon}\log n\right)\\
&\le 4\exp\left(-\frac{n^{-1/2+2\epsilon}\log^2 n}{Ck_n n^{-1/2}\log^\tau n + 12k_n n^{-3/4+\epsilon}\log n}\right) + 36\alpha(k_n)n^{3/4-\epsilon}\log^{-1} n\\
&\le 4\exp\left(-C\log^{1+(1-\tau)/2} n\right) + 36n^{3/4-\epsilon-2\epsilon\beta}\log^{-1-\beta(1-\tau)/2} n\\
&\le 4\exp\left(-C\log^{1+(1-\tau)/2} n\right) + 36n^{-1}\log^{-1-\beta(1-\tau)/2} n.
\end{aligned}$$

Yields (4.7). On the other hand, it follows (2.7) by the same way in the proof of Theorem 2.3. So the proof is completed. ♯

4.2 Proof of uniformly asymptotic normality (Theorem 2.6)

Let $S_n = \sqrt{n}(p - F_n(q_p))/\sigma_p$. By (2.11),

$$U_n = S_n + O\left(n^{-1/4+\frac{7}{4(2\beta+1)}}\log n\right), a.s. \tag{4.8}$$

Hence, by the equation (4.8) and Lemma 3.8, in order to prove Theorem 2.4 it is sufficient to show that

$$\sup_u |F_{S_n}(u) - \Phi(u)| \le Cn^{-(1-b)/6}. \tag{4.9}$$

To prove (4.9), we need some lemmas.

Let $Y_{ni} = (P(X_i \le q_p) - I(X_i \le q_p))/(\sqrt{n}\sigma_p)$, then $S_n = \sum_{i=1}^n Y_{ni}$. Let $q_n = [n^b], p_n = [n^{(1+b)/2}] + 1$ and $k_n = [n^{(1-b)/2}] + 1$ be integers, where $[x]$ denotes the integer part of x. Then $k_n(p_n + q_n) \ge n$. Therefore, S_n may be split as

$$S_n = S_n' + S_n'', \tag{4.10}$$

where $S_n' = \sum_{m=1}^{k_n} y_{nm},\ S_n'' = \sum_{m=1}^{k_n} y'_{nm}$, and

$$y_{nm} = \sum_{i=(m-1)(p_n+q_n)+1}^{(m-1)(p_n+q_n)+p_n} Y_{ni},\quad y'_{nm} = \sum_{i=(m-1)(p_n+q_n)+p_n+1}^{m(p_n+q_n)} Y_{ni}$$

for $m = 1, 2, \ldots, k_n$, where $Y_{ni} = 0$ if $i > n$.

Lemma 4.1 *Under condition of Theorem 2.6,*

$$E(S_n'')^2 \le Cn^{-(1-b)/2}, \tag{4.11}$$

and

$$P(|S''_n| \geq n^{-(1-b)/6}) \leq Cn^{-(1-b)/6}. \tag{4.12}$$

Proof Since $\beta > 1$, we can choose a $\delta > 0$ such that $\beta\delta/(2+\delta) > 1$. It implies that $\sum_{j=1}^{\infty} \alpha^{\delta/(2+\delta)}(j) < \infty$. By Lemma 3.2, we have

$$E(S''_n)^2 \leq C \sum_{m=1}^{k_n} \sum_{i=(m-1)(p_n+q_n)+p_n+1}^{m(p_n+q_n)} ||Y_{ni}||^2_{2+\delta} \leq Ck_n q_n n^{-1} \leq Cn^{-(1-b)/2},$$

yields (4.11). Moreover, we get immediately (4.12) by the Markov's inequality and (4.11). ♯

Set $s_n^2 := \sum_{m=1}^{k} \mathrm{Var}(y_{nm})$. We have the following lemma.

Lemma 4.2 *Under condition of Theorem 2.6,*

$$|s_n^2 - 1| = O(n^{-(1-b)/6}). \tag{4.13}$$

Proof Let $\Gamma_n = \sum_{1\leq i<j\leq k} \mathrm{Cov}(y_{ni}, y_{nj})$. We know that $E(S_n)^2 = \sigma^2_{p,n}/\sigma^2_p$ from (2.8) and $\sigma^2_{p,n} = \sigma^2_p + O(n^{-\beta+1})$ from (2.9). Hence, $E(S_n)^2 = 1 + O(n^{-\beta+1}) = 1 + O(n^{-(1-b)/6})$ due to $\beta \geq 1 + (1-b)/(6b)$ from (2.12). Clearly

$$s_n^2 = E(S'_n)^2 - 2\Gamma_n. \tag{4.14}$$

Moreover, $E(S'_n)^2 = E[S_n - S''_n]^2 = 1 + E(S''_n)^2 - 2E(S_n S''_n) + O(n^{-(1-b)/6})$. Hence by Lemma 4.1,

$$|E(S'_n)^2 - 1| = |E(S''_n)^2 - 2E(S_n S''_n) + O(n^{-(1-b)/6})| \leq Cn^{-(1-b)/6}. \tag{4.15}$$

On the other hand, by Lemma 3.1,

$$\begin{aligned}
|\Gamma_n| &\leq \sum_{1\leq i<j\leq k_n} \sum_{s=(i-1)(p_n+q_n)+1}^{(i-1)(p_n+q_n)+p_n} \sum_{t=(j-1)(p_n+q_n)+1}^{(j-1)(p_n+q_n)+p_n} |\mathrm{Cov}(Y_{ns}, Y_{nt})| \\
&\leq Cn^{-1} \sum_{1\leq i<j\leq k_n} \sum_{s=(i-1)(p_n+q_n)+1}^{(i-1)(p_n+q_n)+p_n} \sum_{t=(j-1)(p_n+q_n)+1}^{(j-1)(p_n+q_n)+p_n} \alpha(t-s) \\
&\leq Cn^{-1} \sum_{1\leq i<j\leq k_n} \sum_{s=(i-1)(p_n+q_n)+1}^{(i-1)(p_n+q_n)+p_n} \sum_{t=(j-1)(p_n+q_n)+1}^{(j-1)(p_n+q_n)+p_n} (t-s)^{-\beta} \\
&\leq Cn^{-1} \sum_{i=1}^{k_n-1} \sum_{s=(i-1)(p_n+q_n)+1}^{(i-1)(p_n+q_n)+p_n} \sum_{t=q_n}^{k_n(p_n+q_n)} t^{-\beta} \\
&\leq C \sum_{t=q_n}^{\infty} t^{-\beta} \leq Cq_n^{-\beta+1} \leq Cn^{-(\beta-1)b} \leq Cn^{-(1-b)/6}.
\end{aligned} \tag{4.16}$$

Combining (4.14)–(4.16) implies the desired result. ♯

Assume that $\{\eta_{nm} : m = 1, \ldots, k_n\}$ are independent random variables, and the distribution of η_{nm} is the same as that of y_{nm} for $m = 1, \ldots, k_n$. Let $T_n = \sum_{m=1}^{k_n} \eta_{nm}$ and $B_n = \sum_{m=1}^{k_n} Var(\eta_{nm})$. Clearly

$$B_n = s_n^2, \quad F_{T_n}(u) = F_{T_n/\sqrt{B_n}}(u/s_n). \tag{4.17}$$

Lemma 4.3 *Under conditions of Theorem 2.6,*

$$\sup_u |F_{T_n/\sqrt{B_n}}(u) - \Phi(u)| = O\left(n^{-(1-b)/6}\right) \tag{4.18}$$

and

$$\sup_u |F_{S'_n}(u) - F_{T_n}(u)| = O\left(n^{-(1-b)/6}\right). \tag{4.19}$$

Proof Let $r = 2 + 2/3$. Since $\beta > 1 + 1/3 = r/2$, we can choose a sufficiently large $\delta > 0$ such that $\beta > r^2/\delta + r/2 = r(r+\delta)/(2\delta)$. Using Lemma 3.3, we have that for $\epsilon > 0$ sufficiently small,

$$\begin{aligned} E|y_{nm}|^r &\le C\left\{ n^\epsilon \sum_{i=(m-1)(p_n+q_n)+1}^{(m-1)(p_n+q_n)+p_n} E|Y_{ni}|^r + \left(\sum_{i=(m-1)(p_n+q_n)+1}^{(m-1)(p_n+q_n)+p_n} ||Y_{ni}||_{r+\delta}^2 \right)^{r/2} \right\} \\ &\le C\left\{ n^\epsilon p_n n^{-r/2} + \left(p_n n^{-1}\right)^{r/2} \right\} \le C\left(p_n n^{-1}\right)^{r/2} \\ &\le C n^{-(1-b)r/4} = C n^{-2(1-b)/3} \end{aligned}$$

Thus

$$\sum_{m=1}^{k_n} E|y_{nm}|^3 \le C k_n n^{-2(1-b)/3} = C n^{-(1-b)/6},$$

and Lemma 4.2 implies $B_n = s_n^2 \to 1$, so $B_n^{-3/2} \sum_{m=1}^{k_n} E|\eta_{nm}|^3 \le C n^{-(1-b)/6}$. Applying Berry-Esseen theorem, we get (4.18).

Assume that $\varphi(t)$ and $\psi(t)$ are the characteristic functions of S'_n and T_n respectively. Note that $\psi(t) = E(\exp\{itT_n\}) = \prod_{m=1}^{k_n} E\exp\{it\eta_{nm}\} = \prod_{m=1}^{k_n} E\exp\{ity_{nm}\}$. Using Lemma 3.9 and Lemma 3.2,

$$\begin{aligned} |\varphi(t) - \psi(t)| &= |E\exp(it\sum_{m=1}^{k_n} y_{nm}) - \prod_{m=1}^{k_n} E\exp(ity_{nm})| \\ &\le C|t|\alpha^{1/2}(q_n) \sum_{m=1}^{k_n} ||y_{nm}||_2 \le C|t|\alpha^{1/2}(q_n) k_n (p_n n^{-1})^{1/2} \\ &\le C|t| q_n^{-\beta/2} k_n (p_n n^{-1})^{1/2} \le C|t| n^{-b\beta/2 + (1-b)/4} \le C|t| n^{-(1-b)/3} \end{aligned}$$

by $\beta \ge 7(1-b)/(6b)$ from (2.12). Therefore

$$I_1 =: \int_{-T}^{T} \left| \frac{\varphi(t) - \psi(t)}{t} \right| dt \le CTn^{-(1-b)/3}. \tag{4.20}$$

On the other hand, note that $F_{T_n}(u) = F_{T_n/\sqrt{B_n}}(u/s_n)$ and (4.20), we have

$$\sup_u |F_{T_n}(u+y) - F_{T_n}(u)| \le \sup_u \left|F_{T_n/\sqrt{B_n}}((u+y)/s_n) - F_{T_n/\sqrt{B_n}}(u/s_n)\right|$$
$$\le \sup_u \left|F_{T_n/\sqrt{B_n}}((u+y)/s_n) - \Phi((u+y)/s_n)\right| + \sup_u |\Phi((u+y)/s_n) - \Phi(u/s_n)|$$
$$+ \sup_u \left|F_{T_n/\sqrt{B_n}}(u/s_n) - \Phi(u/s_n)\right|$$
$$\le 2\sup_u \left|F_{T_n/\sqrt{B_n}}(u) - \Phi(u)\right| + \sup_u |\Phi((u+y)/s_n) - \Phi(u/s_n)|$$
$$\le C\left(n^{-(1-b)/6} + |y|/s_n\right) \le C\left(n^{-(1-b)/6} + |y|\right).$$

Hence,

$$I_2 =: T\sup_u \int_{|y|\le c/T} |F_{T_n}(u+y) - F_{T_n}(u)|\, dy$$
$$\le CT\int_{|y|\le c/T} \{n^{-(1-b)/6} + |y|\}dy \le C\{n^{-(1-b)/6} + 1/T\}. \tag{4.21}$$

By the Esseen inequality (Pollard, 1984) and taking $T = n^{(1-b)/6}$, one has that

$$\sup_u |F_{S_n'}(u) - F_{T_n}(u)| \le I_1 + I_2$$
$$\le C\left\{Tn^{-(1-b)/3} + n^{-(1-b)/6} + 1/T\right\} \le Cn^{-(1-b)/6}.$$

That is (4.19). We have finished the proof. ♯

Proof of (4.9) By Lemma 4.2 and 4.3,

$$\sup_u |F_{S_n'}(u) - \Phi(u)|$$
$$\le \sup_u |F_{S_n'}(u) - F_{T_n}(u)| + \sup_u |F_{T_n}(u) - \Phi(u/\sqrt{B_n})| + \sup_u |\Phi(u/\sqrt{B_n}) - \Phi(u)|$$
$$\le \sup_u |F_{S_n'}(u) - F_{T_n}(u)| + \sup_u |F_{T_n/\sqrt{B_n}}(u/\sqrt{B_n}) - \Phi(u/\sqrt{B_n})| + C|s_n^2 - 1|$$
$$= \sup_u |F_{S_n'}(u) - F_{T_n}(u)| + \sup_u |F_{T_n/\sqrt{B_n}}(u) - \Phi(u)| + C|s_n^2 - 1|$$
$$\le Cn^{-(1-b)/6}$$

Note that (4.10) and (4.12) in Lemma 4.1, and using Lemma 3.8, we obtain the desired result of (4.9). ♯

Acknowledgement

This research is supported by the National Science Foundation of China(10161004) and the National Science Foundation of Guanxi(04047033)

References

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

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

Carrasco, M. and Chen, X. (2002). Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models, *Econometrics Theory* **18**, 17–39.

Chanda, K. C. (1974). strong mixing properties of linear stochastic processes. *Journal of Applied Probability* **11,** 401–408.

Chaudhuri, P. (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. *Ann. Statist.* **19**, 760–777.

Chen, J. and Qin, J. (1993). Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. *Biometrika* **80**, 107–116.

Dowd, K. (2001). Estimating VaR with order statistics. *J. Derivatives* **8**, 23–30.

Fitzmaurice, G.M., Laird, N. M. and Rotnitzky, A.G. (1993). Regression models for discrete longitudinal response, *Statistical Sciences* **8,** 284–309.

Francq, C. and Zakoian, J. M. (2006). Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* **22,** 815–834.

Genon-Catalot, V., Jeantheau, T. and Laredo, C. (2000). Stochastic Volatility models as hidden Markov models and statistical applications. *Bernoulli* **6**, 1051–1079.

Gorodeskii, V. V. (1977). On the strong mixing property for linear sequences. *Theory of Probability And Its Applications* **22**, 411–413.

Hall, P. (1991). Bahadur representation for uniform resampling and importance sampling, with applications to asymptotic relative efficiency, *Ann. Statist.*, **19,** 1062–1072.

He, X. and Shao, Q.M. (1996). A general Bahadur representation of M-estimators and its application to linear regression with nonstochatic designs. *Ann. Statist.*, **24**, 2608–2630.

Humphreys K and Titterington DM. (2000). Improving the mean field approximation in belief networks using Bahadur's reparameterisation of the multivariate binary distribution. *Neural Processing Letters* **12**, 183-197.

Hwang, S. Y. and Kim, T. Y.(2004). Power transformation and threshold modeling for ARCH innovations with applications to tests for ARCH structure. *Stochastic Processes and Their Applications* **110**, 295–314.

Jureckova, J. and Sen, P. K. (1984). On adaptive scale-equivariant M-estimators in linear. models. Statist. Decisions, *Suppl.* **1**, 3146.

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

Koenker, R. and Portnoy, S. (1987). L-estimation for linear models. *J. Amer. Statist. Assoc.* **82**, 851-857.

Koltchinskii, V. (1994). Bahadur-Kiefer approximation for spatial quantiles. In *Probability in Banach Spaces* (J. Hoffman-Jrgensen, J. Kuelbs, and M. Marcus, eds.), pp. 394-408. Birkhauser.

Lee, O. and Shin, D. W. (2004). Strict stationarity and mixing properties of asymmetric power GARCH models allowing a signed volatility. *Economics Letters* **84**, 167-173.

Lu, Z. (1996). A note on geometric ergodicity of autoregressive conditional heteroscedasticity ARCH model. *Statistics and Probability Letters* **30**, 305–311.

Moricz, F. (1976). Moment inequalities and the strong laws of large numbers. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **35**, 299-314.

Pham, T. D. and Tran, L. T. (1985). Some mixing properties of time series models. *Stoch. Process. Appl.*, **19**, 297-303.

Pollard, D., (1984). Convergence of Stochastic Processes. *Spring-Verlag, Berlin.*

Rio, E. (1995). The functional law of the iterated logarithm for stationary stronly mixing sequences. *Ann. Probab.*, **23**, 1188-1203.

Roussas, G. G. and Ioannides, D. A. (1987). Moment inequalities for mixing sequences of random variables. *Stochastic Anal. Appl.*, **5**, 61-120.

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

Shao, Q. (1988). A moment inequality and its applications. *Acta Mathematica Sinica (in chinese)*, **31**, 736-747

Shao, Q. (1995). Maximal inequalities for partial sums of ρ-mixing sequences. *Ann. Probab.* **23,** 948–965.

Sun, S. (2006). The Bahadur representation for sample quantiles under weak dependence. *Statistics and Probability Letters*, **76**, 1238–1244.

Withers, C. S.(1981). Conditions for linear processes to be strong-mixing. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **57**,477–480.

Wu, W.B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *Ann. Statist.*, **33**, 1934-1963.

Xiang, X. (1995). Bahadur representation of the kernel quantile estimator under random censorship, *Journal of Multivariate Analysis*, **54** , 193–209.

Yang, S. (2000). Moment Bounds for strong mixing sequences and their application. *J. Math. Research and Exposition*, **20**, 349-359

Yang, S. (2003). Uniformly asymptotic normality of the regression weighted estimator for negatively associated samples. *Statist. and Probab. Lett.*, **62**, 101-110.

Yang, S. (2006). Maximal Moment Inequality for Partial Sums of Strong Mixing Sequences and Application, *Acta Mathematica Sinica, English series*, Available online 30 September 2006

Yang, S. and Li, Y. (2006). Uniformly asymptotic normality of the regression weighted estimator for strong mixing samples (in Chinese), *Acta Mathematica Sinica, Chinese series*, **49(5)**, 1163-1170.

Yoshihara, K. (1995). The Bahadur representation of sample quantiles for sequences of strongly mixing random variables. *Statist. Prob. Letters* **24**, 299–304.