文章编号:0253-2778(2019)08-0606-08

Berry-Eséen type bound of sample quantiles for positively associated sequence

HU Xueping, ZHANG Tong, GUO Zhidong, ZHANG Heng

(*School of Mathematics and Computation Science, Anqing Normal University, Anqing* 246133, *China*)

Abstract: By utilizing some inequalities for positively associated (PA) random variables, a Berry-Esséen type bound of sample quantiles for PA samples under mild conditions was studied. The rate of uniform asymptotic normality was presented and the rate of convergence is near $O(n^{-1/6})$ when the third moment is finite.

Key words: Berry-Esséen bound; sample quantiles; PA random variables

CLC number: O211.4 **Document code**: A doi:10.3969/j.issn.0253-2778.2019.08.002

2010 Mathematics Subject Classification: 62F12; 62E20; 60F05

Citation: HU Xueping, ZHANG Tong, GUO Zhidong, et al. Berry-Esséen type bound of sample quantiles for positively associated sequence[J]. Journal of University of Science and Technology of China, 2019, 49(8): 606-613.

胡学平,张彤,郭志东,等. PA 序列样本分位数估计的 Berry-Esséen 型界[J]. 中国科学技术大学学报,2019,49(8):606-613.

PA 序列样本分位数估计的 Berry-Esséen 型界

胡学平,张 彤,郭志东,张 恒

(安庆师范大学数学与计算科学学院,安徽安庆 246133)

摘要:主要利用 PA 随机序列的有关不等式,在合适条件下探讨了 PA 样本分位数估计的 Berry-Esséen 型界,获得了其一致渐近正态性的收敛速度且在三阶矩有限时,其收敛速度近似为 $O(n^{-1/6})$.

关键词:Berry-Esséen 界;样本分位数;PA 随机序列

0 Introduction

Let $\{X_n\}_{n\geqslant 1}$ be a sequence of random variables defined on a fixed probability space $(\Omega, \mathscr{F}, P)$ with a common marginal distribution function $F(x)=P(X_1\leqslant x)$. F is a distribution function (continuous from the right, as usual). For $p\in(0,1)$, let

$$\xi_p=\inf\{x: F(x)\geqslant p\},$$

denote the pth quantile of F, and be alternately denoted by $F^{-1}(p)$. $F^{-1}(u), 0<u<1$, is called the inverse function of F. An estimator of the population quantile $F^{-1}(p)$ is given by the sample pth quantile

$$F_n^{-1}(p)=\inf\{x: F_n(x)\geqslant p\},$$

Received: 2018-05-08; **Revised**: 2018-06-11

Foundation item: Supported by the NSFC (11626031), the NSF of Anhui Province Ministry of Education (KJ2016A428).

Biography: HU Xueping (corresponding author), male, born in 1972, master/Prof. Research field: Probability limit theory. E-mail: hxprob@163.com

where $F_n(x)=\frac{1}{n}\sum_{i=1}^{n}I(X_i\leqslant x), x\in\mathbb{R}$, denotes the empirical distribution function based on the sample $X_1, X_2, \cdots, X_n, n\geqslant 1$, $I(A)$ denotes the indicator function of a set A and $\mathbb{R}$ is the real line.

The concept of positively associated sequence was proposed by Joag-Dev and Proschan[1]. A finite of random variables $\{X_i\}_{1\leqslant i\leqslant n}$ is said to be positively associated, if for any disjoint subsets A, $B\subset\{1,2,\cdots,n\}$

$$\mathrm{Cov}(f(X_i, i\in A), g(X_i, i\in B))\geqslant 0,$$

where f and g are real coordinate-wise nondecreasing functions such that this covariance exists. A sequence $\{X_n\}_{n\geqslant 1}$ of random variables is said to be PA if for every $n\geqslant 2, X_1, X_2, \cdots, X_n$ are PA.

For a fixed $p\in(0,1)$, denote $\xi_p=F^{-1}(p)$, $\xi_{p,n}=F_n^{-1}(p)$ and $\Phi(u)$ is the distributing function of $N(0,1)$. The Berry-Esséen bound of the sample quantiles for i. i. d. random variables is given in Ref. [2] as follows:

Theorem 0. 1 Let $p\in(0,1)$ and $\{X_n\}_{n\geqslant 1}$ be a sequence of i. i. d. random variables. Suppose that F possesses a positive continue density f and a bounded second derivative F'' in a neighborhood of ξ_p. Then

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{n^{1/2}(\xi_{p,n}-\xi_p)}{[p(1-p)]^{1/2}/f(\xi_p)}\leqslant x\right)-\Phi(x)\right|= O(n^{-1/2}), n\to\infty.$$

Berry-Esséen theorem, which is known as the rate of convergence in the central limit theorem, can be available in many monographs such as Refs. [3-4]. Under the i. i. d. random variables, the optimal rate is $O(n^{-1/2})$, and for the case of martingales, the rate is $O(n^{-1/4}\lg n)$[5, Chapter 3]. Recently, Ref. [6] obtained the Berry-Esséen bound of the sample quantiles for α-mixing sequence. Their result has an optimal rate of $O(n^{-1/2})$ under the strong condition of mixing coefficients satisfying $\alpha(n)=O(n^{-\alpha_0}), \alpha_0>12$. Yang et al.[7-9] investigated the Berry-Esséen bound of the sample quantiles for NA random sequence and ϕ-mixing sequence, respectively, and obtained the same convergence rate: $O(n^{-1/6}\lg n\lg\lg n)$. In other papers about Berry-Esséen bound, Ref. [10] studied the Berry-Esséen bound for the smooth estimator of a function under association samples. Refs. [11-12] obtained the Berry-Esséen bound in kernel density estimator for associated samples. Refs. [13-15] investigated uniformly asymptotic normality of the regression weighted estimator for NA, PA and strong mixing samples, respectively. Ref. [16] obtained the Berry-Esséen bound in kernel density estimation for α-mixing censored samples. Under associated samples, Ref. [17] studied the consistency and uniformly asymptotic normality of wavelet estimator in the regression model.

There are very few literature works on Berry-Esséen bound of sample quantiles for a sequence of PA random variables. Inspired by Refs. [2, 6-10, 16], we investigate the Berry-Esséen bound of the sample quantiles for PA random variables under some mild conditions and obtain two preliminary lemmas and a theorem. The proof of the theorem is provided in Section 1. The proofs of two preliminary lemmas are given in Section 2. The appendix contains some known results (Lemmas A. 1～A. 5).

Throughout the paper, $C, C_0, C_1, \cdots$ denote some positive constants not depending on n, which may be different in various places. $\lfloor x\rfloor$ denotes the largest integer not exceeding x and second-order stationary means that $(X_1, X_{1+k})\overset{\mathrm{d}}{=}(X_i, X_{i+k})$, $i\geqslant 1, k\geqslant 1$.

1 Assumptions and main results

In order to formulate our main results, we now list some assumptions as follows:

Assumption 1. 1 Let $\{X_n\}_{n\geqslant 1}$ be a second-order stationary PA sequence with zero means and common marginal distribution function F, and F possesses a positive continue density f and a bounded second derivative F'' in a neighborhood of ξ_p, for $p\in(0,1)$.

Assumption 1. 2 There exist some $r>2$ and $\delta>0$ such that

$$\left.\begin{aligned}&\sup_{j\geqslant 1} E\mid X_j\mid^{r+\delta}<\infty,\\&u(n):=\sum_{j=n}^{\infty}\mathrm{Cov}(X_1,X_{j+1})=O(n^{-(r-2)(r+\delta)/(2\delta)})\end{aligned}\right\}\tag{1}$$

Assumption 1.3 There exists an $\varepsilon_0>0$ such that for $x\in[\xi_p-\varepsilon_0,\xi_p+\varepsilon_0]$

$$\sum_{j=n}^{\infty} j\,\mathrm{Cov}[I(X_1\leqslant x),I(X_{j+1}\leqslant x)]=O(n^{-\frac{(r-2)(r+\delta)}{2\delta}+1})\tag{2}$$

where $0<\delta<\frac{r(r-2)}{4-r}$, if $2<r<4$; $\delta>0$, if $r\geqslant 4$, respectively.

Assumption 1.4 There exit positive integers $p:=p_n$ and $q:=q_n$ such that for sufficiently large n

$$p+q\leqslant n,\ qp^{-1}\leqslant C<\infty\tag{3}$$

and let $k:=k_n=\lfloor n/(p+q)\rfloor$, as $n\to\infty$

$$\gamma_{1n}=qp^{-1}\to 0,\ \gamma_{2n}=pn^{-1}\to 0,\ kp/n\to 1\tag{4}$$

Assumption 1.5 There exist some $r>2$ and $\delta>0$ such that

$$\sum_{j=n}^{\infty} j\,\mathrm{Cov}(X_1,X_{j+1})=O(n^{-\frac{(r-2)(r+\delta)}{2\delta}+1})\tag{5}$$

where $0<\delta<\frac{r(r-2)}{4-r}$, if $2<r<4$; $\delta>0$, if $r\geqslant 4$, respectively.

Remark 1.1 Assumptions 1.1, 1.2 and 1.4 are used commonly in the literatures. For example, Refs. [12, 14, 16-17] used Assumption 1.4. Assumptions 1.1 and 1.2 were used by Refs. [14,17] and Assumption 1.1 was assumed in Refs. [7-9]. Assumption 1.4 is easily satisfied, for example, when $p=\lfloor n^{2/3}\rfloor$, $q=\lfloor n^{1/3}\rfloor$, $k=\lfloor\frac{n}{p+q}\rfloor=\lfloor n^{1/3}\rfloor$. It is seen that $pk/n\to 1$ implies $qk/n\to 0$, as $n\to\infty$.

Our main results are as follows.

Theorem 1.1 Assume that Assumptions 1.1, 1.3 and 1.4 are satisfied, and let

$$\mathrm{Var}[I(X_1\leqslant\xi_p)]+2\sum_{j=2}^{\infty}\mathrm{Cov}[I(X_1\leqslant\xi_p),I(X_j\leqslant\xi_p)]:=\sigma^2(\xi_p)>0.$$

Suppose that $p_n\leqslant\frac{n\lg\lg n}{\beta\lg n}$, where $\beta=\frac{144\alpha^2}{\sigma^2(\xi_p)}$, for some $\alpha\geqslant 1$, and

$$C(p_n)=\frac{\lg n}{n^{\alpha/2}p_n\lg\lg n}\cdot\exp\{(\beta n\lg n/(p_n\lg\lg n))^{1/2}\}v(p_n)\leqslant C_0<\infty,$$

where

$$v(n)=\sum_{j=n}^{\infty}\mathrm{Cov}[I(X_1\leqslant x),I(X_{j+1}\leqslant x)],$$

then

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{n^{1/2}(\xi_{p,n}-\xi_p)}{\sigma(\xi_p)/f(\xi_p)}\leqslant x\right)-\Phi(x)\right|=O(a_n)\tag{6}$$

where $a_n=\gamma_{1n}^{1/3}+\gamma_{2n}^{1/3}+\gamma_{2n}^{(r-2)/2}+u^{1/3}(q)\to 0$, $n\to\infty$.

Remark 1.2 The above condition $C(p_n)<\infty$ is similar to (2.3) in Ref. [18]. When $p_n=\frac{n\lg\lg n}{\beta\lg n}$, $v(n)=n^{-\rho}$, for some $\rho>0$, we can obtain $C(p_n)\leqslant C_0<\infty$, while for some $\rho(n)$, if $v(n)=O(\mathrm{e}^{-\rho(n)})$, it follows that $C(p_n)\leqslant C_0<\infty$.

Corollary 1.1 Suppose all the assumptions of Theorem 1.1 are satisfied, and $r=3$, then

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{n^{1/2}(\xi_{p,n}-\xi_p)}{\sigma(\xi_p)/f(\xi_p)}\leqslant x\right)-\Phi(x)\right|=O(n^{-(3+\delta)/(18+2\delta)}).$$

Remark 1.3 The rate of convergence is near $O(n^{-1/6})$ as $\delta\to 0$ by Corollary 1.1. First, we give some preliminaries, which will be used to prove Theorem 1.1.

Lemma 1.1 Let $\{X_n\}_{n\geqslant 1}$ be a stationary random variable sequence with zero mean and $|X_n|\leqslant d<\infty$ for $n=1,2,\cdots$. Suppose that Assumption 1.2 is satisfied. If

$$\liminf_{n\to\infty} n^{-1}\mathrm{Var}(\sum_{i=1}^{n}X_i)=\sigma_1^2>0\tag{7}$$

then

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{\sum_{i=1}^{n}X_i}{\sqrt{\mathrm{Var}(\sum_{i=1}^{n}X_i)}}\leqslant x\right)-\Phi(x)\right|=O(a_n)\tag{8}$$

Corollary 1.2 Suppose all the assumptions of Lemma 1.1 are fulfilled, and $r=3$, $u(n)=$

$O(n^{-(3+\delta)/(2\delta)})$, then

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{\sum_{i=1}^{n}X_i}{\sqrt{\mathrm{Var}(\sum_{i=1}^{n}X_i)}}\leqslant x\right)-\Phi(x)\right|= O(n^{-(3+\delta)/(18+2\delta)}).$$

Remark 1.4 The rate of convergence is near $O(n^{-1/6})$ as $\delta\to 0$ by Corollary 1.2.

Lemma 1.2 Let $\{X_n\}_{n\geqslant 1}$ be a second-order stationary PA sequence with common marginal distribution function F and $EX_n=0$, $|X_n|\leqslant d<\infty$, $n\geqslant 1$. Assumption 1.5 is satisfied and let

$$\mathrm{Var}(X_1)+2\sum_{j=2}^{\infty}\mathrm{Cov}(X_1,X_j):=\sigma_0^2>0,$$

then

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{\sum_{i=1}^{n}X_i}{\sqrt{n}\sigma_0}\leqslant x\right)-\Phi(x)\right|=O(a_n) \quad (9)$$

Similar to Remark 1.3, it follows that the rate of convergence is near $O(n^{-1/6})$ as $\delta\to 0$, if $r=3$ in Eq. (9).

Proof of Theorem 1.1 By taking the same notation as that in the proof of Ref. [9, Theorem 1.1], denote $A=\sigma(\xi_p)/f(\xi_p)$ and

$$G_n(t)=P(n^{1/2}(\xi_{p,n}-\xi_p)/A\leqslant t).$$

Similar to the proof of Ref. [9, Eq. (3.3)]. Let $L_n=(p_n\lg n\lg\lg n)^{1/2}$, we have

$$\sup_{|t|>L_n}|G_n(t)-\Phi(t)|\leqslant P(|\xi_{p,n}-\xi_p|\geqslant AL_nn^{-1/2})+1-\Phi(L_n) \quad (10)$$

Let $\varepsilon_n=\frac{A}{2}L_nn^{-1/2}$, it follows that

$$P(|\xi_{p,n}-\xi_p|\geqslant AL_nn^{-1/2})\leqslant P(|\xi_{p,n}-\xi_p|\geqslant\varepsilon_n) \quad (11)$$

by Lemma A.5(iii), we have

$$P(\xi_{p,n}>\xi_p+\varepsilon_n)= P\Big(\sum_{i=1}^{n}I(X_i>\xi_p+\varepsilon_n)>n(1-p)\Big)= P\Big(\sum_{i=1}^{n}(V_i-EV_i)>n\delta_{n1}\Big),$$

where $V_i=I(X_i>\xi_p+\varepsilon_n)$ and $\delta_{n1}=F(\xi_p+\varepsilon_n)-p$. Likewise,

$$P(\xi_{p,n}<\xi_p-\varepsilon_n)=P(p\leqslant F_n(\xi_p-\varepsilon_n))= P\Big(\sum_{i=1}^{n}(W_i-EW_i)>n\delta_{n2}\Big),$$

where $W_i=I(X_i>\xi_p-\varepsilon_n)$ and $\delta_{n2}=p-F(\xi_p-\varepsilon_n)$. It is easy to see that $\{V_i-EV_i, 1\leqslant i\leqslant n\}$ and $\{W_i-EW_i, 1\leqslant i\leqslant n\}$ are still PA sequences, and $|V_i-EV_i|\leqslant 1$, $|W_i-EW_i|\leqslant 1$. According to Assumption 1.3, we have

$$v(n)\leqslant n^{-1}\sum_{j=n}^{\infty}j\,\mathrm{Cov}[I(X_1\leqslant x),I(X_{j+1}\leqslant x)]= O(n^{-\frac{(r-2)(r+\delta)}{2\delta}}).$$

Combining Ref. [18, Remark 2.1] with Assumption 1.4, Assumptions (A1)～(A3) in Ref. [18] were satisfied for n large enough. According to Lemma A.1, for some $\theta>0$, $\theta p_n\leqslant 1$, we obtain

$$P\Big(\sum_{i=1}^{n}(V_i-EV_i)>n\delta_{n1}\Big)\leqslant 2\{\theta^2nv(p_n)e^{n\theta}+e^{C_1n\theta^2}\}e^{-\frac{n\theta\delta_{n1}}{2}},$$

and

$$P\Big(\sum_{i=1}^{n}(W_i-EW_i)>n\delta_{n2}\Big)\leqslant 2\{\theta^2nv(p_n)e^{n\theta}+e^{C_1n\theta^2}\}e^{-\frac{n\theta\delta_{n2}}{2}}.$$

Since $F(x)$ is continuous at ξ_p with $f(\xi_p)>0$, by the assumption on $f(x)$ and Taylor's expansion

$$\delta_{n1}=F(\xi_p+\varepsilon_n)-p=f(\xi_p)\varepsilon_n+o(\varepsilon_n);$$
$$\delta_{n2}=p-F(\xi_p-\varepsilon_n)=f(\xi_p)\varepsilon_n+o(\varepsilon_n).$$

Therefore, we obtain that for n large enough

$$\frac{f(\xi_p)\varepsilon_n}{2}=\frac{\sigma(\xi_p)L_n}{4n^{1/2}}\leqslant F(\xi_p+\varepsilon_n)-p=\delta_{n1},$$
$$\frac{f(\xi_p)\varepsilon_n}{2}=\frac{\sigma(\xi_p)L_n}{4n^{1/2}}\leqslant p-F(\xi_p-\varepsilon_n)=\delta_{n2}.$$

Taking $\theta=\left(\frac{\beta\lg n}{np_n\lg\lg n}\right)^{1/2}$, it is clear that from $\theta p_n\leqslant 1$

$$e^{-n\theta\delta_{n1}/2}\leqslant e^{-n\theta f(\xi_p)\varepsilon_n/4}=e^{-\frac{3\alpha}{2}\lg n} \quad (12)$$

Note that $p_n\to\infty$, we have for n large enough

$$e^{C_1n\theta^2}=\exp\{C_1\beta\lg n/(p_n\lg\lg n)\}\leqslant e^{\frac{\alpha}{2}\lg n} \quad (13)$$

by the assumption $C(p_n)<\infty$ in Theorem 1.1

$$\theta^2nv(p_n)e^{n\theta}= \frac{\beta\lg n}{p_n\lg\lg n}\exp\left\{\left(\frac{\beta n\lg n}{p_n\lg\lg n}\right)^{1/2}\right\}v(p_n)\leqslant e^{\frac{\alpha}{2}\lg n} \quad (14)$$

From (11)～(14) we obtain

$P(|\xi_{p,n}-\xi_p|>\varepsilon_n)\leqslant C\exp\{-\alpha\lg n\}\leqslant O(n^{-1})$.

Since $1-\Phi(L_n)\leqslant\frac{(2\pi)^{-1/2}}{L_n}\exp\{-L_n^2/2\}=o(n^{-1})$, we have $\sup_{|t|>L_n}|G_n(t)-\Phi(t)|\leqslant O(n^{-1})$.

According to the proof of convergence rate of $|\sigma^2(n,t)-\sigma^2(\xi_p)|$ in Ref.[9]. Taking

$$p_n=(n/(\lg n\lg\lg n))^{1/3},$$

we obtain that for $|t|\leqslant L_n$,

$$|\sigma^2(n,t)-\sigma^2(\xi_p)|=O(n^{-1/3}(p_n\lg n\lg\lg n)^{1/2})+O(n^{-\frac{(r-2)(r+\delta)}{12\delta}}).$$

By taking $r=3,\delta=3$, we obtain $|\sigma^2(n,t)-\sigma^2(\xi_p)|=O(n^{-1/6})$. On the other hand, seeing the proof of Ref.[9, Eq.(3.9)], by Lemma 1.2 it follows that,

$$\sup_{|t|\leqslant L_n}|G_n(t)-\Phi(t)|\leqslant\sup_{|t|\leqslant L_n}\left|P\left[\frac{\sum_{i=1}^n Z_i}{\sqrt{n}\sigma(n,t)}<-c_{nt}\right]-\Phi(-c_{nt})\right|+\sup_{|t|\leqslant L_n}|\Phi(t)-\Phi(c_{nt})|\leqslant$$
$$C\{\gamma_{1n}^{1/3}+\gamma_{2n}^{1/3}+\gamma_{2n}^{(r-2)/2}+u^{1/3}(q)\}+\sup_{|t|\leqslant L_n}|\Phi(t)-\Phi(c_{nt})|\leqslant C\{\gamma_{1n}^{1/3}+\gamma_{2n}^{1/3}+\gamma_{2n}^{(r-2)/2}+u^{1/3}(q)\}.$$

Therefore, Eq.(6) follows the same steps as those in the proof of Ref.[9, Theorem 1.1].

Proof of Corollary 1.1 We obtain it by taking $p=\lfloor n^{3(\delta+1)/(6+4\delta)}\rfloor, q=\lfloor n^{\delta/(3+2\delta)}\rfloor$.

2 Proof of preliminary lemmas

Proof of Lemma 1.1 We employ Bernstein's big-block and small-block procedure and partition the set $\{1,2,\cdots,n\}$ into $2k_n+1$ subsets with a large block of size $p=p_n$ and small blocks of size $q=q_n$, and let $k=k_n:=\lfloor\frac{n}{p_n+q_n}\rfloor$. Define $Z_{n,i}=X_i/\sqrt{\mathrm{Var}(\sum_{i=1}^n X_i)}$, then S_n may be split as

$$S_n:=\frac{\sum_{i=1}^n X_i}{\sqrt{\mathrm{Var}(\sum_{i=1}^n X_i)}}=\sum_{i=1}^n Z_{n,i}=S_{n1}+S_{n2}+S_{n3},$$

where $S_{n1}=\sum_{j=1}^k\eta_j$, $S_{n2}=\sum_{j=1}^k\xi_j$, $S_{n3}=\zeta_k$, and $\eta_j=\sum_{i=k_j}^{k_j+p-1}Z_{n,i}$, $\xi_j=\sum_{i=l_j}^{l_j+q-1}Z_{n,i}$, $\zeta_k=\sum_{i=k(p+q)+1}^{n}Z_{n,i}$, $k_j=(j-1)(p+q)+1, l_j=(j-1)(p+q)+p+1, j=1,2,\cdots,k$.

According to Lemma A.2 with $a=\gamma_{1n}^{1/3}+\gamma_{2n}^{1/3}$, we have

$$\sup_{-\infty<t<\infty}|P(S_n\leqslant t)-\Phi(t)|=$$
$$\sup_{-\infty<t<\infty}|P(S_{n1}+S_{n2}+S_{n3}\leqslant t)-\Phi(t)|\leqslant$$
$$\sup_{-\infty<t<\infty}|P(S_{n1}\leqslant t)-\Phi(t)|+\frac{a}{\sqrt{2\pi}}+$$
$$P(|S_{n2}|\geqslant\gamma_{1n}^{1/3})+P(|S_{n3}|\geqslant\gamma_{2n}^{1/3})\quad(15)$$

Step 1 We estimate $E(S_{n2})^2$ and $E(S_{n3})^2$, which will be used to estimate $P(|S_{n2}|\geqslant\gamma_{1n}^{1/3})$ and $P(|S_{n3}|\geqslant\gamma_{2n}^{1/3})$ in (15). By the conditions $|X_i|\leqslant d$ and (7), it is easy to find that $|Z_{n,i}|\leqslant\frac{C}{\sqrt{n}}$. Combining the definition of PA with the definition ξ_j, $j=1,2,\cdots,k$, we can easily prove that $\{\xi_i\}_{1\leqslant i\leqslant k}$ is PA. According to the stationary and $EX_n=0$, $n\geqslant1$, we have

$$E(S_{n2})^2=\sum_{j=1}^k E\xi_j^{\,2}+2\sum_{1\leqslant i<j\leqslant k}\mathrm{Cov}(\xi_i,\xi_j)=$$
$$\sum_{j=1}^k\sum_{i=l_j}^{l_j+q-1}E(Z_{n,i})^2+$$
$$2\sum_{j=1}^k\sum_{l_j\leqslant i_1<i_2\leqslant l_j+q-1}\mathrm{Cov}(Z_{n,i_1},Z_{n,i_2})+$$
$$2\sum_{1\leqslant i<j\leqslant k}\sum_{i_1'=l_i}^{l_i+q-1}\sum_{i_2'=l_j}^{l_j+q-1}\mathrm{Cov}(Z_{n,i_1'},Z_{n,i_2'})\leqslant$$
$$Cn^{-1}kq+\sum_{j=1}^k\sum_{i=1}^{q-1}(q-i)\mathrm{Cov}(Z_{n,1},Z_{n,i+1})+$$
$$\sum_{i=1}^{k-1}\sum_{i_1'=l_i}^{l_i+q-1}\sum_{j=i+1}^k\sum_{i_2'=l_j}^{l_j+q-1}\mathrm{Cov}(Z_{n,i_1'},Z_{n,i_2'})\leqslant$$
$$C[kq+kqu(1)+kqu(p)]/n\leqslant$$
$$Ckq/n=Cqp^{-1}=C\gamma_{1n}\quad(16)$$

$$E(S_{n3})^2=\sum_{i=k(p+q)+1}^n E(Z_{n,i})^2+$$
$$2\sum_{k(p+q)+1\leqslant i_1<i_2\leqslant n}\mathrm{Cov}(Z_{n,i_1},Z_{n,i_2})\leqslant$$
$$C\{n^{-1}[n-k(p+q)]+$$

$$p \sum_{i=1}^{n-k(p+q)-1} \mathrm{Cov}(Z_{n,1}, Z_{n,i+1})\} \leqslant C\{ n^{-1}[n-k(p+q)] + pn^{-1}u(1)\} = C(p+q)/n = C\gamma_{2n} \quad (17)$$

Hence, by Markov's inequality, (16) and (17), we have

$$P(|S_{n2}| > \gamma_{1n}^{1/3}) \leqslant C\gamma_{1n}^{-2/3}E(S_{n2})^2 \leqslant C\gamma_{1n}^{1/3} \quad (18)$$

$$P(|S_{n3}| > \gamma_{2n}^{1/3}) \leqslant C\gamma_{2n}^{-2/3}E(S_{n3})^2 \leqslant C\gamma_{2n}^{1/3} \quad (19)$$

Step 2 We estimate $\sup_{-\infty<t<\infty} |P(S_{n1} \leqslant t) - \Phi(t)|$. Define

$$s_n^2 := \sum_{j=1}^{k} \mathrm{Var}(\eta_j),\ \Gamma_n := \sum_{1\leqslant i<j\leqslant k} \mathrm{Cov}(\eta_i, \eta_j).$$

Clearly $s_n^2 = E(S_{n1})^2 - 2\Gamma_n$, and since $ES_n^2 = 1$, by (16) and (17) we get that

$$|E(S_{n1})^2 - 1| = |E(S_{n2}+S_{n3})^2 - 2E[S_n(S_{n2}+S_{n3})]| \leqslant E(S_{n2})^2 + E(S_{n3})^2 + 2E[(S_{n2})^2]^{1/2}E[(S_{n3})^2]^{1/2} + 2E[(S_n)^2]^{1/2}E[(S_{n2})^2]^{1/2} + 2E[(S_n)^2]^{1/2}E[(S_{n3})^2]^{1/2} \leqslant C(\gamma_{1n}^{1/2} + \gamma_{2n}^{1/2}) \quad (20)$$

On the other hand, similarly to the process of (16),

$$\Gamma_n = \sum_{1\leqslant i<j\leqslant k} \sum_{s=k_i}^{k_i+p-1} \sum_{t=k_j}^{k_j+p-1} \mathrm{Cov}(Z_{n,s}, Z_{n,t}) = \sum_{i=1}^{k-1} \sum_{s=k_i}^{k_i+p-1} \sum_{j=i+1}^{k} \sum_{t=k_j}^{k_j+p-1} \mathrm{Cov}(Z_{n,s}, Z_{n,t}) \leqslant C[kpu(q)]/n \leqslant Cu(q) \quad (21)$$

From (20) and (21), it follows that

$$|s_n^2 - 1| \leqslant C[\gamma_{1n}^{1/2} + \gamma_{2n}^{1/2} + u(q)] \quad (22)$$

We assume that η_j' are the independent random variables and η_j' have the same distribution as η_j, $j=1, 2, \cdots, k$. Let $H_n := \sum_{j=1}^{k} \eta_j'$. It is easily seen that

$$\sup_{-\infty<t<\infty} |P(S_{n1} \leqslant t) - \Phi(t)| \leqslant \sup_{-\infty<t<\infty} |P(S_{n1} \leqslant t) - P(H_n \leqslant t)| + \sup_{-\infty<t<\infty} |P(H_n \leqslant t) - \Phi(t/s_n)| + \sup_{-\infty<t<\infty} |\Phi(t/s_n) - \Phi(t)| := D_1 + D_2 + D_3.$$

Let $\phi(t)$ and $\psi(t)$ be the characteristic function of S_{n1} and H_n, respectively. Thus applying Esséen inequality (see Ref. [3, Theorem 5.3]), for any $T>0$,

$$D_1 \leqslant \int_{-T}^{T} \left| \frac{\phi(t) - \varphi(t)}{t} \right| \mathrm{d}t + T \sup_{-\infty<t<\infty} \int_{|u|\leqslant C/T} |P(H_n \leqslant u+t) - P(H_n \leqslant t)| \mathrm{d}u := D_{1n} + D_{2n}.$$

By Lemma A.3, we have that

$$|\phi(t) - \varphi(t)| = \left| E\exp(\mathrm{i}t\sum_{j=1}^{k} \eta_j) - \prod_{j=1}^{k} E\exp(\mathrm{i}t\eta_j) \right| \leqslant 4t^2 \sum_{1\leqslant i<j\leqslant k} \sum_{s=k_i}^{k_i+p-1} \sum_{t=k_j}^{k_j+p-1} \mathrm{Cov}(Z_{n,s}, Z_{n,t}) \leqslant Ct^2u(q).$$

Therefore

$$D_{1n} = \int_{-T}^{T} \left| \frac{\phi(t) - \varphi(t)}{t} \right| \mathrm{d}t \leqslant Cu(q)T^2 \quad (23)$$

It follows from Berry-Esséen inequality[3, Theorem 5.7] and Lemma A.4, that

$$\sup_{-\infty<t<\infty} |P(H_n/s_n \leqslant t) - \Phi(t)| \leqslant \frac{C}{s_n^r} \sum_{j=1}^{k} E|\eta_j'|^r = \frac{C}{s_n^r} \sum_{j=1}^{k} E|\eta_j|^r \leqslant \frac{Ck[(p/n)]^{r/2}}{s_n^r} \leqslant C\frac{\gamma_{2n}^{(r-2)/2}}{s_n^r} \quad (24)$$

Note that $s_n \to 1$, as $n \to \infty$ by (22). From (24), we get that

$$\sup_{-\infty<t<\infty} |P(H_n/s_n \leqslant t) - \Phi(t)| \leqslant C\gamma_{2n}^{(r-2)/2} \quad (25)$$

which implies that

$$\sup_{-\infty<t<\infty} |P(H_n \leqslant t+u) - P(H_n \leqslant t)| \leqslant \sup_{-\infty<t<\infty} \left| P\left(\frac{H_n}{s_n} \leqslant \frac{t+u}{s_n}\right) - \Phi\left(\frac{t+u}{s_n}\right) \right| + \sup_{-\infty<t<\infty} \left| P\left(\frac{H_n}{s_n} \leqslant \frac{t}{s_n}\right) - \Phi\left(\frac{t}{s_n}\right) \right| + \sup_{-\infty<t<\infty} \left| \Phi\left(\frac{t+u}{s_n}\right) - \Phi\left(\frac{t}{s_n}\right) \right| \leqslant 2\sup_{-\infty<t<\infty} \left| P\left(\frac{H_n}{s_n} \leqslant t\right) - \Phi(t) \right| + \sup_{-\infty<t<\infty} \left| \Phi\left(\frac{t+u}{s_n}\right) - \Phi\left(\frac{t}{s_n}\right) \right| \leqslant C\left(\gamma_{2n}^{(r-2)/2} + \left|\frac{u}{s_n}\right|\right) \quad (26)$$

By (26), we obtain

$$D_{2n}=T\sup_{-\infty<t<\infty}\int_{|u|\leqslant C/T}|P(H_n\leqslant t+u)-P(H_n\leqslant t)|\,\mathrm{d}u\leqslant C(\gamma_{2n}^{(r-2)/2}+1/T)\quad(27)$$

Combining (23) with (27), and choosing $T=u^{-1/3}(q)$, it is easily see that

$$D_1\leqslant C(u^{1/3}(q))+\gamma_{2n}^{(r-2)/2})\quad(28)$$

and by (25),

$$D_2=\sup_{-\infty<t<\infty}\left|P\left(\frac{H_n}{s_n}\leqslant\frac{t}{s_n}\right)-\Phi(\frac{t}{s_n})\right|\leqslant C\gamma_{2n}^{(r-2)/2}\quad(29)$$

On the other hand, from (22) and Ref. [3, Lemma 5.2], it follows that

$$D_3\leqslant(2\pi\mathrm{e})^{-1/2}(s_n-1)I(s_n\geqslant 1)+(2\pi\mathrm{e})^{-1/2}(s_n^{-1}-1)I(0<s_n<1)\leqslant C\,|\,s_n^2-1\,|\leqslant C[\gamma_{1n}^{1/2}+\gamma_{2n}^{1/2}+u(q)]\quad(30)$$

Consequently, combining (28),(29) with (30), we can get

$$\sup_{-\infty<t<\infty}|P(S_{n1}\leqslant t)-\Phi(t)|\leqslant C[\gamma_{1n}^{1/2}+\gamma_{2n}^{1/2}+\gamma_{2n}^{(r-2)/2}+u^{1/3}(q)]\quad(31)$$

Finally, by (15),(18),(19) and (31), (8) is verified.

Proof of Corollary 1.2 We obtain it by taking $p=\lfloor n^{3(\delta+1)/(6+4\delta)}\rfloor, q=\lfloor n^{\delta/(3+2\delta)}\rfloor$.

Proof of Lemma 1.2 Define $\sigma_n^2:=\mathrm{Var}(\sum_{i=1}^n X_i)$ and $\gamma(k)=\mathrm{Cov}(X_i,X_{i+k})$ for $i=1,2,\cdots$, according to (5), it is checked that

$$\sum_{j=n}^{\infty}\mathrm{Cov}(X_1,X_{j+1})\leqslant n^{-1}\sum_{j=n}^{\infty}j\,\mathrm{Cov}(X_1,X_{j+1})=O(n^{-\frac{(r-2)(r+\delta)}{2\delta}})\quad(32)$$

therefore Assumption 1.2 holds true. For the second-order stationary process $\{X_n\}_{n\geqslant 1}$ with common marginal distribution function, by Eq. (5) it follows that

$$|\sigma_n^2-n\sigma_0^2|=\left|n\gamma(0)+2n\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)\gamma(j)-n\gamma(0)-2n\sum_{j=1}^{\infty}\gamma(j)\right|=\left|2n\sum_{j=1}^{n-1}\frac{j}{n}\gamma(j)+2n\sum_{j=n}^{\infty}\gamma(j)\right|\leqslant 2\sum_{j=1}^{\infty}j\gamma(j)+2\sum_{j=n}^{\infty}j\gamma(j)\leqslant 4\sum_{j=1}^{\infty}j\gamma(j)=O(1)\quad(33)$$

On the other hand,

$$\sup_{-\infty<t<\infty}\left|P\left(\frac{\sum_{i=1}^n X_i}{\sqrt{n}\sigma_0}\leqslant t\right)-\Phi(t)\right|\leqslant\sup_{-\infty<t<\infty}\left|P\left(\frac{\sum_{i=1}^n X_i}{\sigma_n}\leqslant\frac{\sqrt{n}\sigma_0}{\sigma_n}t\right)-\Phi\left(\frac{\sqrt{n}\sigma_0}{\sigma_n}t\right)\right|+\sup_{-\infty<t<\infty}\left|\Phi\left(\frac{\sqrt{n}\sigma_0}{\sigma_n}t\right)-\Phi(t)\right|:=I_1+I_2\quad(34)$$

By (33), it is easy to see that $\lim_{n\to\infty}\frac{\sigma_n^2}{n\sigma_0^2}=1$. Thus, applying Lemma 1.1, one has

$$I_1\leqslant C\{\gamma_{1n}^{1/3}+\gamma_{2n}^{1/3}+\gamma_{2n}^{(r-2)/2}+u^{1/3}(q)\}\quad(35)$$

and according to (33) again, similarly to the proof of (30), we obtain that

$$I_2\leqslant C\left|\frac{\sigma_n^2}{n\sigma_0^2}-1\right|=\frac{C}{n\sigma_0^2}\,|\,\sigma_n^2-n\sigma_0^2\,|=O(n^{-1})\quad(36)$$

Combining (34),(35) with (36), (9) holds true.

Appendix

Lemma A.1[18] Let $\{X_n\}_{n\geqslant 1}$ be PA random variables with zero means and $\max_{1\leqslant i\leqslant n}|X_n|\leqslant c_n<\infty$, a.s. for $n=1,2,\cdots$. Denote

$$u(n)=\sup_{i\geqslant 1}\sum_{j:|i-j|\geqslant n}\mathrm{Cov}(X_i,X_j),$$

and satisfies $\sum_{i=1}^{\infty}u^{1/2}(2^i)<\infty$. Assume that $\theta p_n c_n\leqslant 1$ for some $\theta>0$. Then there exits a positive constant C_1, which does not depend on n, such that for every $\varepsilon>0$

$$P(\left|\sum_{i=1}^n X_i\right|>n\varepsilon)\leqslant 4\{\theta^2 nu(p_n)\mathrm{e}^{n\theta c_n}+\mathrm{e}^{C_1 n\theta^2 c_n^2}\}\,\mathrm{e}^{-n\theta\varepsilon/2}.$$

Lemma A.2 Let X and Y be random variables, then for any $a>0$,

$$\sup_t|P(X+Y\leqslant t)-\Phi(t)|\leqslant\sup_t|P(X\leqslant t)-\Phi(t)|+\frac{a}{\sqrt{2\pi}}+P(|Y|>a).$$

Lemma A.3[14] Let $\{X_n\}_{n\geqslant 1}$ be a PA sequence, and let $\{a_n, n\geqslant 1\}$ be a real constant sequence, $1=m_0<m_1<\cdots<m_k=n$. Denote by

$\eta_l := \sum_{j=m_{l-1}+1}^{m_l} a_j X_j$ for $1 \leqslant l \leqslant k$. Then

$$\left| E\exp(\mathrm{i}t\sum_{l=1}^{k}\eta_l) - \prod_{l=1}^{k} E\exp(\mathrm{i}t\eta_l) \right| \leqslant 4t^2 \sum_{1\leqslant s<t\leqslant n} |a_s a_t| \operatorname{Cov}(X_s, X_t).$$

Lemma A. 4[14]　Let $\{X_j\}_{j\geqslant 1}$ be a stationary PA sequence with $EX_j = 0$ for $j = 1, 2, \cdots$, and there exist some $r > 2$ and $\delta > 0$ such that

$$\sup_{j\geqslant 1} E|X_j|^{r+\delta} < \infty,$$

$$u(n) = \sum_{j=n}^{\infty} \operatorname{Cov}(X_1, X_{j+1}) = O(n^{-(r-2)(r+\delta)/2\delta}).$$

Let $\{a_j\}_{j\geqslant 1}$ be a real constant sequence, $a := \sup_j |a_j| < \infty$. Then there is a constant C not depending on n such that

$$E\left|\sum_{j=1}^{n} a_j X_j\right|^r \leqslant C a^r n^{r/2}.$$

Especially, if $\{X_n\}_{n\geqslant 1}$ is a stationary PA sequence with $EX_n = 0$, $|X_n| \leqslant d < \infty$, for $n = 1, 2, \cdots$, and assume $u(n) = O(n^{-(r-2)/2})$ for some $r > 2$, then $E\left|\sum_{j=1}^{n} X_j\right|^r \leqslant C n^{r/2}$.

Lemma A. 5[2]　Let $F(x)$ be a right-continuous distribution function. The inverse function $F^{-1}(t)$, $0 < t < 1$, is nondecreasing and left-continuous, and satisfies

(ⅰ) $F^{-1}(F(x)) \leqslant x$, $-\infty < x < \infty$;

(ⅱ) $F(F^{-1}(t)) \geqslant t$, $0 < t < 1$;

(ⅲ) $F(x) \geqslant t$ if and only if $x \geqslant F^{-1}(t)$.

References

[1] JOAG-DEV K, PROSCHAN F. Negative association of random variables with applications[J]. Ann Stat, 1983, 11(1): 286-295.

[2] SERFILING R J. Approximation Theorems of Mathematical Statistics[M]. New York: John Willey & Sons, 1980.

[3] PETROV V V. Limit Theorem of Probability Theory: Sequences of Independent Random Variables[M]. New York: Oxford Univ Press Inc, 1995.

[4] SHIRYAEV A N. Probability[M]. 2nd ed. New York: Springer-Verlag, 1989.

[5] HALL P. Martingale Limit Theory and Its Application [M]. New York: Academic Press Inc, 1980.

[6] LAHIRI S N, SUN S. A Berry-Esséen theorem for samples quantiles under weak dependence[J]. Ann Appl Probab, 2009, 19(1): 108-126.

[7] YANG W Z, WANG X J, LI X Q, et al. Berry-Esséen bound of sample quantiles for φ-mixing random variables[J]. J Math Anal, 2012, 388(1): 451-462.

[8] YANG W Z, HU S H, WANG X J, et al. The Berry-Esséen type bound of sample quantiles for strong mixing sequence[J]. J Statist Plann Infer, 2012, 142(3): 660-672.

[9] YANG W Z, HU S H, WANG X J, et al. Berry-Esséen bound of sample quantiles for negatively associated sequence[J]. J Inequal Appl, 2011, 2011: 83.

[10] CAI Z W, ROUSSAS G G. Berry-Esséen bounds for smooth estimator of a function under association[J]. J Nonparameter Stat, 1999, 10: 79-106.

[11] ROUSSAS G G. Asymptotic normality of the kernel estimate of a probability density function under association[J]. Statis Probab Lett, 2000, 50: 1-12.

[12] ROUSSAS G G. An Esséen-type inequality for probability density functions with an application[J]. Stat Probab Lett, 2001, 51: 397-408.

[13] YANG S C. Uniformly asymptotic normality of the regression weighted estimator for negatively associated samples[J]. Stat Probab Lett, 2003, 62(2): 101-110.

[14] YANG Shanchao, LI Yufang. Uniform asymptotic normality of the regression weighted estimator for positively associated samples[J]. Chinese J Appl Probab Statist, 2005, 21(2): 150-160.

[15] YANG Shanchao, LI Yongming. Uniform asymptotic normality of the regression weighted estimator for strong mixing samples[J]. Acta Math Sinica (Chin Ser), 2006, 49(5): 1163-1170.

[16] LIANG H Y, DE UÑA-ÁLVAREZ J. A Berry-Esséen type bound in kernel density estimation for strong mixing censored samples[J]. J Multivariate Anal, 2009, 100: 1219-1231.

[17] LI Y M, YANG S C, ZHOU Y. Consistency and uniformly asymptotic normality of wavelet estimator in regression model with associated samples[J]. Statist Probab Lett, 2008, 78: 2947-2956.

[18] YANG S C, CHEN M. Exponential inequality for associated random variables and strong laws of large number[J]. Science in China Series A: Mathematics, 2007, 50(5): 705-714.