Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH **Open Access**

The Berry-Esséen bound of sample quantiles for NA sequence

Tingting Liu[1], Zhimian Zhang[2], Shuhe Hu[1] and Wenzhi Yang[1*]

*Correspondence: wzyang827@163.com
[1]School of Mathematical Science, Anhui University, Hefei, 230039, P.R. China
Full list of author information is available at the end of the article

Abstract
By using the exponential inequality, we investigate the Berry-Esséen bound of sample quantiles for negatively associated (NA) random variables and obtain the rate $O(n^{-1/6}\log n)$. Our result extends the corresponding one obtaining $O(n^{-1/9})$.
MSC: 62F12; 62E20; 60F05

Keywords: Berry-Esséen bound; sample quantile; NA random variables

1 Introduction

First, we will recall the definition of negatively associated (NA) random variables.

Definition 1.1 A finite family $\{X_1,\dots,X_n\}$ is said to be negatively associated (NA) if for any disjoint subsets $A, B \subset \{1,2,\dots,n\}$, and any real coordinatewise nondecreasing functions f on R^A, g on R^B,

$$\mathrm{Cov}\big(f(X_k, k\in A), g(X_k, k\in B)\big) \le 0.$$

A sequence of random variables $\{X_n\}_{n\ge1}$ is said to be negatively associated (NA) if for every $n\ge 2$, $X_1, X_2, \dots, X_n$ are NA.

The concept of an NA sequence was introduced by Joag-Dev and Proschan [1]. There are many good results of NA random variables. For example, Matula [2] obtained the three series theorem, Su *et al.* [3] gave the moment inequality, Shao [4] investigated the maximal inequality, Yuan *et al.* [5] studied the central limit theorem, Yang [6] and Sung [7] investigated the exponential inequality, *etc.*

In this article, we investigate the Berry-Esséen bound of sample quantiles for NA random variables and obtain the rate $O(n^{-1/6}\log n)$. Our result extends the corresponding one of Yang *et al.* [8] obtaining $O(n^{-1/9})$. Let us give some details of the pth quantile.

Let $\{X_n\}_{n\ge1}$ be a sequence of random variables defined on a fixed probability space $(\Omega,\mathcal{F},P)$ with a common marginal distribution function $F(x) = P(X_1 \le x)$, where F is a distribution function (continuous from the right, as usual). For $0 < p < 1$, the pth quantile of F is defined as

$$\xi_p = \inf\big\{x : F(x) \ge p\big\}$$

and is alternately denoted by $F^{-1}(p)$. The function $F^{-1}(t)$, $0 < t < 1$, is called the inverse function of F. With a sample $X_1, X_2, \dots, X_n$, $n \ge 1$, let F_n represent the empirical distribution function based on $X_1, X_2, \ldots, X_n$, which is defined as $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$, $x \in \mathbb{R}$, where $I(A)$ denotes the indicator function of a set A and $\mathbb{R}$ is the real line. Let $0 < p < 1$, we define

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

as the pth quantile of sample.

Throughout the paper, $C, C_1, C_2, \ldots, d$ 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

$$(X_1, X_{1+k}) \stackrel{d}{=} (X_i, X_{i+k}), \quad i \ge 1, k \ge 1.$$

For $0 < p < 1$, denote $\xi_p = F^{-1}(p)$, $\xi_{p,n} = F_n^{-1}(p)$ and $\Phi(t)$ is the distribution function of a standard normal variable. Yang *et al.* [8, Theorem 1.1] presented the Berry-Esséen bound of sample quantiles for an NA sequence as follows.

Theorem 1.1 *Let* $0 < p < 1$ *and* $\{X_n\}_{n\ge 1}$ *be a second-order stationary NA sequence with common marginal distribution function* F *and* $EX_n = 0$ *for* $n = 1, 2, \ldots$ *. Assume that in a neighborhood of* ξ_p, F *possesses a positive continuous density* f *and a bounded second derivative* F''. *Suppose that there exists an* $\varepsilon_0 > 0$ *such that for* $x \in [\xi_p - \varepsilon_0, \xi_p + \varepsilon_0]$,

$$\sum_{j=2}^{\infty} j \bigl|\mathrm{Cov}\bigl[I(X_1 \le x), I(X_j \le x)\bigr]\bigr| < \infty \tag{1.1}$$

and

$$\mathrm{Var}\bigl[I(X_1 \le \xi_p)\bigr] + 2\sum_{j=2}^{\infty} \mathrm{Cov}\bigl[I(X_1 \le \xi_p), I(X_j \le \xi_p)\bigr] = \sigma^2(\xi_p) > 0. \tag{1.2}$$

Then

$$\sup_{-\infty<t<\infty} \left| P\left(\frac{n^{1/2}(\xi_{p,n} - \xi_p)}{\sigma(\xi_p)/f(\xi_p)} \le t\right) - \Phi(t)\right| = O\bigl(n^{-1/9}\bigr), \quad n \to \infty. \tag{1.3}$$

For the work on Berry-Esséen bounds of sample quantiles, one can refer to Reiss [9] or Chapter 2 of Serfling [10]. Cai and Roussas [11] studied the smooth estimate of quantiles under an association sample, Rio [12] obtained the Berry-Esséen bounds of sample quantiles under a φ-mixing sequence, Lahiri and Sun [13] and Yang *et al.* [14] investigated the Berry-Esséen bounds of sample quantiles under an α-mixing sequence, *etc.* For more work on Berry-Esséen bounds, we can refer to Chapter 3 of Hall and Heyde [15], Chapter 5 of Petrov [16], Gao *et al.* [17], Chapter 5 of Härdle *et al.* [18], and to the references therein too.

Moreover, value-at-risk (VaR) is a popular measure of the market risk associated with an asset or a portfolio of assets. It has been chosen by the Basel Committee on Banking Supervision as a benchmark risk measure and has been used by financial institutions for asset management and minimization of risk. Let $\{X_t\}_{t=1}^{n}$ be the market value of an asset over n periods of a time unit, and let $Y_t = \log(X_t/X_{t-1})$ be the log-returns. Suppose $\{Y_t\}_{t=1}^n$ is a strictly stationary dependent process with marginal distribution function F. Given a positive value p close to zero, the $1-p$ level VaR is

$$\nu_p = \inf\{x : F(x) \geq p\},$$

which specifies the smallest amount of loss such that the probability of the loss in market value being large than ν_p is less than p. So, the study of VaR is a specific application of the pth quantile. For more details, one can refer to Chen and Tang [19] and the references therein.

In this paper, by the exponential inequality and properties of NA random variables, we go on studying the Berry-Esséen bound of sample quantiles for an NA sequence and get a better rate of normal approximation. For the details, see Theorem 2.1 in Section 2. Some preliminaries and the proof of Theorem 2.1 are presented in Section 3.

2 Main result

Theorem 2.1 *Let $0 < p < 1$ and $\{X_n\}_{n\geq 1}$ be a second-order stationary NA sequence with common marginal distribution function F. Assume that in a neighborhood of ξ_p, F possesses a positive continuous density f and a bounded second derivative F''. Let n_0 be some positive integer. Suppose that there exists an $\varepsilon_0 > 0$ such that for $x \in [\xi_p - \varepsilon_0, \xi_p + \varepsilon_0]$*

$$\bigl|\mathrm{Cov}\bigl[I(X_1 \leq x), I(X_j \leq x)\bigr]\bigr| \leq Cj^{-5/2}, \quad j \geq n_0 \tag{2.1}$$

and condition (1.2) *holds. Then*

$$\sup_{-\infty<t<\infty}\left|P\left(\frac{n^{1/2}(\xi_{p,n} - \xi_p)}{\sigma(\xi_p)/f(\xi_p)} \leq t\right) - \Phi(t)\right| = O\bigl(n^{-1/6}\log n\bigr), \quad n \to \infty. \tag{2.2}$$

Remark 2.1 Obviously, the condition (2.1) of Theorem 2.1 is relatively stronger than (1.1) of Theorem 1.1, but the normal approximation rate $O(n^{-1/6}\log n)$ in (2.2) is better than $O(n^{-1/9})$ in (1.3). So our result Theorem 2.1 extends Theorem 1.1 of Yang *et al.* [8]. It is pointed out that the condition of mean zero in Theorem 1.1 should be removed. In fact, the process of estimating (3.9) on page 12 of Yang *et al.* [8], was used the Lemma 2.2 of Yang *et al.* [8], which requires the condition of mean zero, but Z_i in (3.9) of Yang *et al.* [8], defined by $Z_i = I[X_i \leq \xi_p + tAn^{-1/2}] - EI[X_i \leq \xi_p + tAn^{-1/2}]$, satisfies the condition of mean zero. Thus, the mean zero condition of Theorem 1.1 of Yang *et al.* [8] is not needed. It coincides with the independent case, which does not need the mean zero condition. For the details, one can see Serfling [10, Theorem C, p.81] or Theorem A of Yang *et al.* [8].

3 Some preliminaries and the proof of Theorem 2.1

First, we give some preliminaries, which will be used to prove our Theorem 2.1.

Lemma 3.1 [6, Lemma 3.5] *Let $\{X_n\}_{n\geq 1}$ be a NA sequence with $EX_n = 0$, $|X_n| \leq b$, a.s. $n = 1, 2, \ldots$. Denote $\Delta_n = \sum_{i=1}^n EX_i^2$. Then for $\forall\varepsilon > 0$,*

$$P\left(\left|\sum_{i=1}^n X_i\right| > \varepsilon\right) \leq 2\exp\left\{-\frac{\varepsilon^2}{2(2\Delta_n + b\varepsilon)}\right\}.$$

Lemma 3.2 *Let $\{X_n\}_{n\geq 1}$ be a stationary NA sequence with $EX_n = 0$ and $|X_n| \leq d < \infty$, $n = 1, 2, \ldots$. Assume that there exists a $\beta \geq 3/2$ such that*

$$\sum_{j=b_n}^{\infty} \bigl|\mathrm{Cov}(X_1, X_j)\bigr| = O\bigl(b_n^{-\beta}\bigr) \tag{3.1}$$

for all $0 < b_n \to \infty$ as $n \to \infty$ and

$$\liminf_{n\to\infty} n^{-1} \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sigma_1^2 > 0. \tag{3.2}$$

Then

$$\sup_{-\infty<t<\infty} \left| P\left(\frac{\sum_{i=1}^n X_i}{\sqrt{\operatorname{Var}(\sum_{i=1}^n X_i)}} \leq t \right) - \Phi(t) \right| = O\bigl(n^{-1/6} \log n\bigr), \quad n \to \infty. \tag{3.3}$$

Proof By taking the same notation as that in the proof of Lemma 2.1 of Yang *et al.* [8], we partition the set $\{1, 2, \ldots, n\}$ into $2k_n + 1$ subsets with large block of size $\mu = \mu_n$ and small block of size $\nu = \nu_n$. Let

$$\mu_n = \bigl\lfloor n^{2/3} \bigr\rfloor, \qquad \nu_n = \bigl\lfloor n^{1/3} \bigr\rfloor, \qquad k = k_n = \left\lfloor \frac{n}{\mu_n + \nu_n} \right\rfloor = \bigl\lfloor n^{1/3} \bigr\rfloor$$

and $Z_{n,i} = X_i / \sqrt{\operatorname{Var}(\sum_{i=1}^n X_i)}$. Define η_j, ξ_j, ζ_k as follows:

$$\eta_j = \sum_{i=j(\mu+\nu)+1}^{j(\mu+\nu)+\mu} Z_{n,i}, \quad 0 \leq j \leq k-1, \qquad \xi_j = \sum_{i=j(\mu+\nu)+\mu+1}^{(j+1)(\mu+\nu)} Z_{n,i}, \quad 0 \leq j \leq k-1,$$

$$\zeta_k = \sum_{i=k(\mu+\nu)+1}^{n} Z_{n,i}.$$

Denote

$$S_n := \frac{\sum_{i=1}^n X_i}{\sqrt{\operatorname{Var}(\sum_{i=1}^n X_i)}} = \sum_{j=0}^{k-1} \eta_j + \sum_{j=0}^{k-1} \xi_j + \zeta_k := S_n' + S_n'' + S_n'''.$$

By Lemma A.3 in Yang *et al.* [8] with $a = 2\varepsilon_n = 2Mn^{-1/6} \log n$ we have

$$\begin{aligned}
\sup_{-\infty<t<\infty} \bigl|P(S_n \leq t) - \Phi(t)\bigr| &= \sup_{-\infty<t<\infty} \bigl|P\bigl(S_n' + S_n'' + S_n''' \leq t\bigr) - \Phi(t)\bigr| \\
&\leq \sup_{-\infty<t<\infty} \bigl|P\bigl(S_n' \leq t\bigr) - \Phi(t)\bigr| + \frac{2\varepsilon_n}{\sqrt{2\pi}} \\
&\quad + P\bigl(\bigl|S_n''\bigr| > \varepsilon_n\bigr) + P\bigl(\bigl|S_n'''\bigr| > \varepsilon_n\bigr),
\end{aligned} \tag{3.4}$$

where M is a positive constant.

Combining the definition of NA with the definition of ξ_j, $j = 0, 1, \ldots, k-1$, we can easily prove that $\{\xi_0, \xi_1, \ldots, \xi_{k-1}\}$ is NA. Together the condition (3.2) with (2.8) of Yang *et al.* [8], it has $E(S_n'')^2 \leq C_1 n^{-1/3}$. On the other hand, it can be seen that $|\xi_j| \leq C_2 n^{-1/6}$, $j = 0, 1, \ldots, k-1$. Thus, we take M large enough and apply Lemma 3.1, and we obtain for n large enough

$$\begin{aligned} P\bigl(\bigl|S_n''\bigr| > \varepsilon_n\bigr) &\leq 2\exp\biggl\{-\frac{M^2 n^{-1/3} \cdot \log^2 n}{2(2C_1 n^{-1/3} + C_2 M n^{-1/6} n^{-1/6} \cdot \log n)}\biggr\} \\ &= 2\exp\biggl\{-\frac{M^2 \cdot \log^2 n}{2(2C_1 + C_2 M \log n)}\biggr\} \\ &\leq C_3 n^{-1}. \end{aligned} \tag{3.5}$$

Meanwhile, by (2.9) of Yang *et al.* [8], it follows $E(S_n''')^2 \leq C_4 n^{-1/3}$. Since $|Z_{n,i}| \leq C_5 n^{-1/2}$, by Lemma 3.1 again, one has for n large enough

$$\begin{aligned} P\bigl(\bigl|S_n'''\bigr| > \varepsilon_n\bigr) &\leq 2\exp\biggl\{-\frac{M^2 n^{-1/3} \cdot \log^2 n}{2(2C_4 n^{-1/3} + C_5 M n^{-1/2} n^{-1/6} \cdot \log n)}\biggr\} \\ &\leq 2\exp\biggl\{-\frac{M^2 \log^2 n}{2(2C_4 + C_5 M)}\biggr\} \\ &\leq C_6 n^{-1}. \end{aligned} \tag{3.6}$$

Similar to the proof of (2.18) in Yang *et al.* [8], by (3.1), it can be seen that

$$\begin{aligned} \bigl|\phi(t) - \psi(t)\bigr| &= \left|E\exp\left(it\sum_{j=0}^{k-1}\eta_j\right) - \prod_{j=0}^{k-1} E\exp(it\eta_j)\right| \\ &\leq 4t^2 \sum_{0\leq i<j\leq k-1} \sum_{l_1=1}^{\mu_n}\sum_{l_2=1}^{\mu_n}\bigl|\operatorname{Cov}(Z_{n,\lambda_i+l_1}, Z_{n,\lambda_j+l_2})\bigr| \\ &\leq \frac{C_1 t^2}{n} \sum_{\substack{1\leq i<j\leq n \\ j-i\geq \nu_n}} \bigl|\operatorname{Cov}(X_i, X_j)\bigr| \\ &\leq C_2 t^2 \sum_{j\geq \nu_n}\bigl|\operatorname{Cov}(X_1, X_j)\bigr| \\ &\leq C_3 t^2 \nu_n^{-\beta} \leq C_4 t^2 n^{-\beta/3}. \end{aligned}$$

Combining the above inequality with $T = n^{\frac{2\beta-1}{12}}$, $\beta \geq 3/2$, we obtain

$$D_{1n} = \int_{-T}^{T}\left|\frac{\phi(t) - \psi(t)}{t}\right| dt \leq C n^{-\beta/3} \cdot T^2 = O\bigl(n^{-1/6}\bigr).$$

On the other hand, we take $T = n^{\frac{2\beta-1}{12}}$, $\beta \geq 3/2$, in (2.23) of Yang *et al.* [8] and have $D_{2n} = O(n^{-1/6})$.

Consequently, by the proof of (2.26) of Yang *et al.* [8], it is easy to check that

$$\sup_{-\infty<t<\infty}\bigl|P\bigl(S_n' \leq t\bigr) - \Phi(t)\bigr| = O\bigl(n^{-1/6}\bigr). \tag{3.7}$$

Finally, by (3.4)-(3.7), (3.3) holds. □

Lemma 3.3 *Let $\{X_n\}_{n\geq 1}$ be a stationary NA sequence with $EX_n = 0$ and $|X_n| \leq d < \infty$, $n = 1, 2, \ldots$. Assume that there exists an n_0 such that*

$$\left|\operatorname{Cov}(X_1, X_j)\right| \leq Cj^{-5/2}, \quad j \geq n_0 \tag{3.8}$$

and

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

Then

$$\sup_{-\infty<t<\infty}\left|P\left(\frac{\sum_{i=1}^n X_i}{\sqrt{n}\sigma_0} \leq t\right) - \Phi(t)\right| = O\left(n^{-1/6}\log n\right). \tag{3.9}$$

Proof By the condition (3.8), it is checked that

$$\sum_{j=b_n}^{\infty}\left|\operatorname{Cov}(X_1, X_j)\right| \leq C\sum_{j=b_n}^{\infty} j^{-5/2} = O\left(b_n^{-3/2}\right),$$

providing $b_n \to \infty$ as $n \to \infty$. So by (3.8), the condition (3.1) of Lemma 3.2 holds. Combining Lemma 3.2 with the proof of Lemma 2.2 of Yang *et al.* [8], we have (3.9) finally. □

Proof of Theorem 2.1 By taking the same notation as that in the proof of Theorem 1.1 in Yang *et al.* [8], one checks the proof of (3.9) in Yang *et al.* [8] and obtains by Lemma 3.3

$$\begin{aligned}\sup_{|t|\leq L_n}\left|G_n(t) - \Phi(t)\right| &\leq \sup_{|t|\leq L_n}\left|P\left[\frac{\sum_{i=1}^n Z_i}{\sqrt{n}\sigma(n,t)} < -c_{nt}\right] - \Phi(-c_{nt})\right| + \sup_{|t|\leq L_n}\left|\Phi(t) - \Phi(c_{nt})\right| \\ &\leq C_1\left(\sigma^2(\xi_p)\right)n^{-\frac{1}{6}}\log n + \sup_{|t|\leq L_n}\left|\Phi(t) - \Phi(c_{nt})\right|,\end{aligned}$$

where $C_1(\sigma^2(\xi_p))$ is a positive constant depending only on $\sigma^2(\xi_p)$. Therefore, (2.2) follows by the same steps as those in the proof of Theorem 1.1 of Yang *et al.* [8]. □

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors read and approved the final manuscript.

Author details
[1]School of Mathematical Science, Anhui University, Hefei, 230039, P.R. China. [2]School of Electronics and Information Engineering, Anhui University, Hefei, 230039, P.R. China.

Acknowledgements
The authors are deeply grateful to the editor and two anonymous referees whose insightful comments and suggestions have contributed substantially to the improvement of this paper. This work was supported by the NNSF of China (11171001, 11201001, 11301004, 11326172), Natural Science Foundation of Anhui Province (1208085QA03, 1308085QA03), Talents Youth Fund of Anhui Province Universities (2012SQRL204), Higher Education Talent Revitalization Project of Anhui Province (2013SQRL005ZD) and the Academic and Technology Leaders to Introduction Projects of Anhui University.

Received: 28 July 2013 Accepted: 5 February 2014 Published: 17 Feb 2014

References

1. Joag-Dev, K, Proschan, F: Negative association of random variables with applications. Ann. Stat. **11**(1), 286-295 (1983)
2. Matula, P: A note on the almost sure convergence of sums of negatively dependent random variables. Stat. Probab. Lett. **15**(3), 209-213 (1992)
3. Su, C, Zhao, LC, Wang, YB: Moment inequalities and weak convergence for negatively associated sequences. Sci. China Ser. A **40**(2), 172-182 (1997)
4. Shao, QM: A comparison theorem on moment inequalities between negatively associated and independent random variables. J. Theor. Probab. **13**(2), 343-356 (2000)
5. Yuan, M, Su, C, Hu, TZ: A central limit theorem for random fields of negatively associated processes. J. Theor. Probab. **16**(2), 309-323 (2003)
6. Yang, SC: Uniformly asymptotic normality of the regression weighted estimator for negatively associated samples. Stat. Probab. Lett. **62**(2), 101-110 (2003)
7. Sung, SH: On the exponential inequalities for negatively dependent random variables. J. Math. Anal. Appl. **381**(2), 538-545 (2011)
8. Yang, WZ, Hu, SH, Wang, XJ, Zhang, QC: Berry-Esséen bound of sample quantiles for negatively associated sequence. J. Inequal. Appl. **2011**, 83 (2011)
9. Reiss, RD: On the accuracy of the normal approximation for quantiles. Ann. Probab. **2**(4), 741-744 (1974)
10. Serfling, RJ: Approximation Theorems of Mathematical Statistics. Wiley, New York (1980)
11. Cai, ZW, Roussas, GG: Smooth estimate of quantiles under association. Stat. Probab. Lett. **36**(3), 275-287 (1997)
12. Rio, E: Sur le théorème de Berry-Esseen pour les suites faiblement dépendantes. Probab. Theory Relat. Fields **104**(2), 255-282 (1996)
13. Lahiri, SN, Sun, S: A Berry-Esseen theorem for samples quantiles under weak dependence. Ann. Appl. Probab. **19**(1), 108-126 (2009)
14. Yang, WZ, Hu, SH, Wang, XJ, Ling, NX: The Berry-Esséen type bound of sample quantiles for strong mixing sequence. J. Stat. Plan. Inference **142**(3), 660-672 (2012)
15. Hall, P, Heyde, CC: Martingale Limit Theory and Its Application. Academic Press, New York (1980)
16. Petrov, VV: Limit Theorems of Probability Theory: Sequences of Independent Random Variables. Oxford University Press, New York (1995)
17. Gao, JT, Hong, SY, Liang, H: Berry-Esséen bounds of error variance estimation in partly linear models. Chin. Ann. Math., Ser. B **17**(4), 477-490 (1996)
18. Härdle, W, Liang, H, Gao, J: Partially Linear Models. Springer Series in Economics and Statistics. Physica-Verlag, New York (2000)
19. Chen, SX, Tang, CY: Nonparametric inference of value-at-risk for dependent financial returns. J. Financ. Econom. **3**(2), 227-255 (2005)

10.1186/1029-242X-2014-79
Cite this article as: Liu et al.: The Berry-Esséen bound of sample quantiles for NA sequence. *Journal of Inequalities and Applications* 2014, 2014:79