Moment inequalities and weak convergence for negatively associated sequences*

SU Chun (苏　淳), ZHAO Lincheng (赵林城)

(Department of Statistics and Finance, University of Science and Technology of China, Hefei 230026, China)

and WANG Yuebao (王岳宝)

(Department of Mathematics, Zhanjiang Normal Institute, Zhanjiang 524048, China)

Received July 25, 1996

Abstract A probability inequality for S_n and some pth moment ($p \geqslant 2$) inequalities for $|S_n|$ and $\max_{1 \leqslant k \leqslant n} |S_k|$ are established. Here S_n is the partial sum of a negatively associated sequence. Based on these inequalities, a weak invariance principle for strictly stationary negatively associated sequences is proved under some general conditions.

Keywords: NA (negatively associated) sequences, partial sums, maximum partial sums, moment inequalities, weak invariance principle.

Following introducing concepts of association and some other types of dependences of random variables in the 1960s (see refs. [1, 2]), more concepts of dependence, including negative association (NA) of random variables, were introduced in the 1980s (see refs. [3, 4]). Since these concepts have a lot of applications, e.g. in reliability theory, percolation theory and multivariate statistical analysis, their limit properties have aroused wide interest.

A finite family of random variables $X_1, \cdots, X_n$ ($n \geqslant 2$) is said to be negatively associated (NA)[3], if for any two disjoint nonempty subsets A_1 and A_2 of $\{1, 2, \cdots, n\}$, the following inequality holds:

$$\operatorname{cov}(f_1(X_i, i \in A_1), f_2(X_j, j \in A_2)) \leqslant 0,$$

where f_1 and f_2 are any two coordinatewise nondecreasing or nonincreasing functions, whenever the covariance is finite. A finite family $X_1, \cdots, X_n$ ($n \geqslant 2$) is said to be associated (or positively associated, PA for short) if for any two coordinatewise nondecreasing or nonincreasing functions of f_1 and f_2,

$$\operatorname{cov}(f_1(X_1, \cdots, X_n), f_2(X_1, \cdots, X_n)) \geqslant 0,$$

whenever the covariance is finite. A sequence of random variables is NA or PA, if every finite subsequence is NA or PA respectively.

In 1984, Newman reported in his survey paper[5] some research results on asymptotic inde-

* Project supported by the National Natural Science Foundation of China, the Doctoral Program Foundation of the State Education Commission of China and the High Eductional Natural Science Foundation of Guangdong Province.

pendence and limit theorems for various types of positively and negatively dependent random variables, including some CLTs for strictly stationary PA or NA sequences, weak invariance principles for strictly stationary PA sequences and Berry-Esseen theorem for several other dependent types of sequences. Since 1984, many new results on limit properties of PA and other positively dependent sequences have been obtained. However, researches on NA and other negatively dependent sequences are left behind. Until 1992 in Matula's work[6] a Kolmogorov type of upper bound inequality and a three-series theorem for NA sequences were established. Matula's results and a generalized Borel-Cantelli lemma proved by Petrov in ref. [7] opened a path to researches on almost sure and complete convergence for NA sequences. Using these results, it is discovered that identically distributed NA random variables have the same Kolmogorov and Marcinkiewicz SLLN as that of iid random variables (see ref. [6] and footnote 1)). It is also found that NA sequences also have complete convergence properties similar to that of iid random variables (see ref. [8]). All these results are very useful in the applications of NA sequences and they also show that these limit properties of NA sequences are furthermore quite different from that of PA and other dependent sequences.

At the same time, it is obvious that there are difficulties in researching limit properties for NA sequences. For example, it is difficult to establish a Levy-type inequality for NA sequences. It is well known that if $\{X_j, j\in\mathbb{N}\}$ is a sequence of PA random variables, and $\mathrm{E}X_j=0$, $\mathrm{E}X_j^2<\infty$, for every $j\in\mathbb{N}$, then[9]

$$\mathrm{P}\left(\max_{1\leqslant k\leqslant n}|S_k|\geqslant\lambda s_n\right)\leqslant 2\mathrm{P}\left(|S_n|\geqslant(\lambda-\sqrt{2})s_n\right),$$

where $\lambda>\sqrt{2}$, $S_k=\sum_{j=1}^{k}X_j$, and $s_n^2=\operatorname{var}S_n$. This inequality plays an important role in studying weak invariance principle and some other limit theorems for PA sequences. Since it is difficult to establish a similar inequality for NA sequences, we have to find another way to prove some corresponding results for NA sequences, which motivates the study of this paper.

1 Some inequalities for NA sequences

Firstly, we introduce two preliminary lemmas.

Lemma 1[4]. *Let $X_1,\cdots,X_n$ be NA variables, $A_1,\cdots,A_m$ be some pairwise disjoint nonempty subsets of $\{1,\cdots,n\}$, $\alpha_i=\#(A_i)$, where $\#(A)$ denotes the number of elements in set A. If $f_i:\mathbb{R}^{\alpha_i}\to\mathbb{R}$, $i=1,\cdots,m$, are m coordinatewise nondecreasing functions (or m nonincreasing), then $f_1(X_j, j\in A_1),\cdots,f_m(X_j, j\in A_m)$ are also NA variables. Besides, if $f_i\geqslant 0$ for $i=1,\cdots,m$, then*

$$\mathrm{E}\left(\prod_{i=1}^{m}f_i(X_j,j\in A_i)\right)\leqslant\prod_{i=1}^{m}\mathrm{E}f_i(X_j,j\in A_i).\tag{1.1}$$

Lemma 2[7] **(Matula's lemma).** *Let $\{X_j, j\in\mathbb{N}\}$ be a sequence of NA random variables, $\mathrm{E}X_j^2<\infty$, $\mathrm{E}X_j=0$ for every $j\in\mathbb{N}$. Denote $S_k=\sum_{j=1}^{k}X_j$, then*

1) Su Chun, Wang Yuebao, On the strong convergence of NA sequence with a common marginal distribution, (to appear).

$$\mathrm{E}\left(\max_{1\leqslant k\leqslant n} S_k\right)^2 \leqslant \sum_{j=1}^{n} \mathrm{E}X_j^2. \tag{1.2}$$

Proof. See the proof of lemma 4 in reference [6].

In this section, we will prove the following two theorems.

Theorem 1. *Let $X_1, \cdots, X_n$ be NA random variables, $\mathrm{E}X_j = 0$ and $\mathrm{E}|X_j|^p < \infty$, where $j = 1, \cdots, n$ and $p \geqslant 2$. Then for $S_n = \sum_{j=1}^{n} X_j$ and any $t > \frac{p}{2}$, $x > 0$, we have*

$$\mathrm{P}(|S_n| \geqslant x) \leqslant \sum_{j=1}^{n} \mathrm{P}(|X_j| \geqslant x/t) + 2\mathrm{e}^t \left\{1 + x^2 \Big/ \left(t \sum_{j=1}^{n} \mathrm{E}X_j^2\right)\right\}^{-t}; \tag{1.3}$$

$$\mathrm{E}|S_n|^p \leqslant c_p \left(\sum_{j=1}^{n} \mathrm{E}|X_j|^p + \left(\sum_{j=1}^{n} \mathrm{E}X_j^2 \right)^{p/2} \right), \tag{1.4}$$

$$\mathrm{E}|S_n|^p \leqslant c_p n^{p/2-1} \sum_{j=1}^{n} \mathrm{E}|X_j|^p, \tag{1.5}$$

where $c_p > 0$ depends only on p.

Theorem 2. *Let $\{X_j, j \in \mathbb{N}\}$ be an NA sequence, $\mathrm{E}X_j = 0$ for every $j \in \mathbb{N}$ and $\beta_p =: \sup_j \mathrm{E}|X_j|^p < \infty$ for some $p \geqslant 2$. If we define*

$$S_{a,k} = \sum_{j=0}^{k-1} X_{a+j}, \quad S_{1,k} = S_k, \quad \beta_2 = \sup_j \mathrm{E}X_j^2,$$

then there exists a constant K_p which depends only on p such that

$$\mathrm{E}\left(\max_{1\leqslant k\leqslant n} |S_{a,k}|\right)^p \leqslant K_p(n\beta_p + (n\beta_2)^{p/2}), \tag{1.6}$$

$$\mathrm{E}\left(\max_{1\leqslant k\leqslant n} |S_{a,k}|\right)^p \leqslant K_p \beta_p n^{p/2}, \tag{1.7}$$

for all $a \in \mathbb{N}$ and $n \in \mathbb{N}$.

Proof of Theorem 1. Following ref. [10], let $y > 0$ and

$$Y_j = \min(X_j, y), \quad T_n = \sum_{j=1}^{n} Y_j,$$

we know from Lemma 1 that

$$\mathrm{E}\mathrm{e}^{hT_n} = \mathrm{E}\prod_{j=1}^{n} \mathrm{e}^{hY_j} \leqslant \prod_{j=1}^{n} \mathrm{E}\mathrm{e}^{hY_j}. \tag{1.8}$$

Denoting $F_j(x) = \mathrm{P}(X_j < x)$, we have

$$\mathrm{E}\mathrm{e}^{hY_j} = \int_{-\infty}^{y} \mathrm{e}^{hx} \mathrm{d}F_j(x) + \mathrm{e}^{hy}\mathrm{P}(X_j \geqslant y)$$

$$= 1 + h\mathrm{E}Y_j + \int_{-\infty}^{y}(\mathrm{e}^{hx} - 1 - hx)\mathrm{d}F_j(x) + (\mathrm{e}^{hy} - 1 - hy)\mathrm{P}(X_j \geqslant y)$$

$$\leqslant 1 + \int_{-\infty}^{y}(\mathrm{e}^{hx} - 1 - hx)\mathrm{d}F_j(x) + (\mathrm{e}^{hy} - 1 - hy)\mathrm{P}(X_j \geqslant y)$$

since $\mathrm{E}Y_j \leqslant 0$ and $h > 0$. We observe that the function

$$g(x) =: x^{-2}(\mathrm{e}^{hx} - 1 - hx)$$

is nondecreasing for any constant $h > 0$. It follows from the above inequality that

$$\mathrm{E}\mathrm{e}^{hY_j} \leqslant 1 + \frac{\mathrm{e}^{hy} - 1 - hy}{y^2}\left(\int_{-\infty}^{y} x^2\mathrm{d}F_j(x) + y^2\mathrm{P}(X_j \geqslant y)\right)$$
$$\leqslant 1 + \frac{\mathrm{e}^{hy} - 1 - hy}{y^2}\mathrm{E}X_j^2 \leqslant \exp\left\{\frac{\mathrm{e}^{hy} - 1 - hy}{y^2}\mathrm{E}X_j^2\right\}.$$

Here the last inequality follows from $1 + x \leqslant \mathrm{e}^x$, $\forall x \in \mathbb{R}$. Denoting $B_n = \sum_{j=1}^{n}\mathrm{E}X_j^2$, we conclude from the above inequality and (1.8) that for any $x > 0$ and $h > 0$, we have

$$\mathrm{e}^{-hx}\mathrm{E}\mathrm{e}^{hT_n} \leqslant \exp\{-hx + (\mathrm{e}^{hy} - 1 - hy)y^{-2} \cdot B_n\}. \tag{1.9}$$

Letting $h = \ln\left(\frac{xy}{B_n} + 1\right)\Big/ y > 0$, we get

$$(\mathrm{e}^{hy} - 1 - hy)y^{-2} \cdot B_n = \frac{x}{y} - \frac{B_n}{y^2}\ln\left(\frac{xy}{B_n} + 1\right) \leqslant \frac{x}{y}.$$

Putting this one into (1.9), we get furthermore

$$\mathrm{e}^{-hx}\mathrm{E}\mathrm{e}^{hT_n} \leqslant \exp\left\{\frac{x}{y} - \frac{x}{y}\ln\left(\frac{xy}{B_n} + 1\right)\right\}. \tag{1.10}$$

If now we let

$$X_j = \min(-X_j, y), \quad T_n' = \sum_{j=1}^{n} Z_j,$$

then Z_j is a nonincreasing function for X_j. $Z_1, \cdots, Z_n$ are also NA variables by Lemma 1. Similarly, (1.10) is also true when T_n is replaced by T_n'. Therefore, for any $x > 0$ and $y > 0$, we have

$$\mathrm{P}(|S_n| \geqslant x) \leqslant \mathrm{P}(S_n \geqslant x) + \mathrm{P}(-S_n \geqslant x)$$

$$\leqslant \sum_{j=1}^{n}\mathrm{P}(|X_j| \geqslant y) + \mathrm{e}^{-hx}\mathrm{E}\mathrm{e}^{hT_n} + \mathrm{e}^{-hx}\mathrm{E}\mathrm{e}^{hT_n'}$$

$$\leqslant \sum_{j=1}^{n}\mathrm{P}(|X_j| \geqslant y) + 2\exp\left\{\frac{x}{y} - \frac{x}{y}\ln\left(\frac{xy}{B_n} + 1\right)\right\}.$$

Putting $\frac{x}{y} = t > \frac{p}{2}$ into the above inequality, we get

$$P(|S_n| \geqslant x) \leqslant \sum_{j=1}^{n} P\left(|X_j| \geqslant \frac{x}{t}\right) + 2\exp\left\{t - t\ln\left(\frac{x^2}{tB_n} + 1\right)\right\},$$

which is just (1.3).

Multiplying (1.3) by px^{p-1}, and integrating over $0 < x < \infty$, according to

$$E|X|^p = p\int_0^{+\infty} x^{p-1} P(|X| \geqslant x)\mathrm{d}x,$$

we obtain

$$E|S_n|^p \leqslant t^p \sum_{j=1}^{n} E|X_j|^p + 2pe^t \int_0^{-\infty} x^{p-1}\left(\frac{x^2}{tB_n} + 1\right)^{-t} \mathrm{d}x$$

$$= t^p \sum_{j=1}^{n} E|X_j|^p + pe^t t^{p/2} B\left(\frac{p}{2}, t - \frac{p}{2}\right) B_n^{p/2}, \tag{1.11}$$

where $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x = \int_0^{+\infty} x^{\alpha-1}(1+x)^{-(\alpha+\beta)}\mathrm{d}x$. Letting $t = p$ and $c_p = \max\left(p^p, p^{1+p/2}e^p B\left(\frac{p}{2}, \frac{p}{2}\right)\right)$ we can deduce (1.4) from (1.11). It is not difficult to get (1.5) from (1.4).

Proof of Theorem 2. Obviously, (1.7) follows from (1.6), hence it suffices to prove (1.6). Following the proof of Theorem 3.7.5 of ref. [11], we prove (1.6) by induction on p.

From Matula's Lemma (i.e. Lemma 2) we observe that (1.6) is also true for $p = 2$. Suppose that (1.6) is true for integer $p = \nu \geqslant 2$, we will prove that (1.6) is also true for $p = \nu + \delta$ with $0 < \delta \leqslant 1$.

For natural numbers a and k, define

$$f_{a,k} = \max(0, S_{a,1}, S_{a,2}, \cdots, S_{a,k}), \quad f_{1,k} = f_k,$$

$$\tilde{f}_{a,k} = \max(0, -S_{a,1}, -S_{a,2}, \cdots, -S_{a,k}), \quad \tilde{f}_{1,k} = \tilde{f}_k,$$

$$g_{a,k} = \max(S_{a,1}, S_{a,2}, \cdots, S_{a,k}), \quad g_{1,k} = g_k,$$

$$\tilde{g}_{a,k} = \max(-S_{a,1}, -S_{a,2}, \cdots, -S_{a,k}), \quad \tilde{g}_{1,k} = \tilde{g}_k.$$

Then it is obvious that

$$\max_{1 \leqslant k \leqslant n} |S_{a,k}| \leqslant \max(0, S_{a,1}, \cdots, S_{a,n}) + \max(0, -S_{a,1}, \cdots, -S_{a,n}) = f_{a,n} + \tilde{f}_{a,n}, \tag{1.12}$$

and the above functions have the following properties:

(i) $f_{a,k}$ is a nondecreasing function of $X_a, \cdots, X_{a+k-1}$; $\tilde{f}_{a,k}$ is a nondecreasing function of $-X_a, \cdots, -X_{a+k-1}$;

(ii) $f_{a,k}$ is a nondecreasing function of k, furthermore it has the following subadditivity property:

$$f_{a,j+k} \leqslant f_{a,j} + f_{a+j,k}; \tag{1.13}$$

(iii)

$$0 \leqslant f_{a,k} \leqslant | g_{a,k} | \leqslant \max_{1 \leqslant j < k} | S_{a,j} |; \tag{1.14}$$

and $\tilde{f}_{a,k}$ and $\tilde{g}_{a,k}$ have properties similar to (1.13) and (1.14) respectively.

Write $\beta_2 = \sigma^2$ and

$$L_{p,n} = n\beta_p + n^{p/2}\sigma^p. \tag{1.15}$$

Then property (iii), Lemma 2 and the induction hypotheses imply that

$$\mathrm{E}f_{a,k}^2 \leqslant k\sigma^2, \quad \mathrm{E}f_{a,k}^{\nu} \leqslant K_\nu L_{\nu,k}, \quad \forall\, a, k \in \mathbb{N}. \tag{1.16}$$

We will firstly prove that there exists a constant $M > 0$, depending only on ν and σ, such that

$$\mathrm{E}f_{a,n}^{\nu+\delta} \leqslant M(L_{\nu+\delta,n} + n^{1+\delta/2}\beta_\nu\sigma^\delta) = M(n\beta_{\nu+\delta} + n^{(\nu+\delta)/2}\delta^{\nu+\delta} + n^{1+\delta/2}\beta_\nu\sigma^\delta) \tag{1.17}$$

for any $a, n \in \mathbb{N}$.

For $n = 2m$ or $2m-1$, let $\nu = n - m$, i.e. $\nu = m$ or $m-1$. Without loss of generality, we may take $a = 1$. From (1.13) we get

$$\mathrm{E}f_n^{\nu+\delta} \leqslant \mathrm{E}(f_m + f_{m+1,\mu})^\nu (f_m^\delta + f_{m+1,\mu}^\delta) = \mathrm{E}f_m^{\nu+\delta} + \mathrm{E}f_{m+1,\mu}^{\nu+\delta} + I_n + J_n, \tag{1.18}$$

where

$$I_n = \sum_{j=1}^{\nu-1} \binom{\nu}{j} \mathrm{E}f_m^{j+\delta} f_{m+1,\mu}^{\nu-j} + \mathrm{E}f_m^\delta f_{m+1,\mu}^{\nu},$$

$$J_n = \sum_{j=1}^{\nu-1} \binom{\nu}{j} \mathrm{E}f_m^{j} f_{m+1,\mu}^{\nu+\delta-j} + \mathrm{E}f_m^\nu f_{m+1,\mu}^{\delta}.$$

In the following, $M_1, M_2, \cdots$ are some constants depending only on ν and δ. We will prove that

$$I_n + J_n \leqslant M_1(m^{1+\delta/2}\beta_\nu\sigma^\delta + m^{(\nu+\delta)/2}\sigma^{\nu+\delta}). \tag{1.19}$$

From properties of NA variables, properties (ii) and (iii) of $f_{a,k}$ and $\nu \leqslant m$, we observe that

$$I_n \leqslant \sum_{j=1}^{\nu-1}\binom{\nu}{j}\mathrm{E}f_m^{j+\delta}\cdot \mathrm{E}f_{m+1,\mu}^{\nu-j} + \mathrm{E}f_m^{\delta}\cdot \mathrm{E}f_{m+1,\mu}^{\nu}$$

$$\leqslant \sum_{j=1}^{\nu-1}\binom{\nu}{j}\mathrm{E}f_m^{j+\delta}\mathrm{E}f_{m+1,m}^{\nu-j} + \mathrm{E}f_m^{\delta}\mathrm{E}f_{m+1,m}^{\nu}, \tag{1.20}$$

and from (1.16), we know that

$$\mathrm{E}f_m^{\delta}\cdot \mathrm{E}f_{m+1,m}^{\nu} \leqslant (\mathrm{E}f_m^2)^{\delta/2}\cdot \mathrm{E}f_{m+1,m}^{\nu} \leqslant K_\nu m^{\delta/2}\sigma^{\delta}L_{\nu,m}$$

$$= K_\nu(m^{1+\delta/2}\beta_\nu\sigma^{\delta} + m^{(\nu+\delta)/2}\sigma^{\nu+\sigma}), \tag{1.21}$$

where the last equality is obtained from (1.15). Since $\nu \geqslant 4$, it follows from (1.16) and the Hölder inequality that for $2 \leqslant j \leqslant \nu-2$,

$$\mathrm{E}f_m^{j+\delta} \leqslant (\mathrm{E}f_m^{\nu})^{(j+\delta-2)/(\nu-2)}(\mathrm{E}f_m^2)^{(\nu-j-\delta)/(\nu-2)}$$

$$\leqslant (K_\nu L_{\nu,m})^{(j+\delta-2)/(\nu-2)}(m\sigma^2)^{(\nu-j-\delta)/(\nu-2)}, \tag{1.22}$$

$$\mathrm{E}f_{m+1,m}^{\nu-j} \leqslant (\mathrm{E}f_{m+1,m}^{\nu})^{(\nu-j-2)/(\nu-2)}(\mathrm{E}f_{m+1,m}^2)^{j/(\nu-2)}$$

$$\leqslant (K_\nu L_{\nu,m})^{(\nu-j-2)/(\nu-2)}(m\sigma^2)^{j/(\nu-2)}, \tag{1.23}$$

and

$$\mathrm{E}f_m^{j+\delta}\cdot \mathrm{E}f_{m+1,m}^{\nu-j} \leqslant K_\nu(m\sigma^2)^{(\nu-\delta)/(\nu-2)}L_{\nu,m}^{(\nu+\delta-4)/(\nu-2)}$$

$$\leqslant K_\nu\sigma^{\nu+\delta}m^{(\nu-\delta)/(\nu-2)}(m\beta_\nu\sigma^{-\nu} + m^{\nu/2})^{(\nu+\delta-4)/(\nu-2)}$$

$$\leqslant K_\nu\sigma^{\nu+\delta}[m^2(\beta_\nu\sigma^{-\nu})^{(\nu+\delta-4)/(\nu-2)} + m^{(\nu+\delta)/2}], \tag{1.24}$$

where the first inequality in (1.24) is obtained by (1.22), (1.23) and $K_\nu \geqslant 1$, and the last inequality by the Cr-inequality. By Jensen's inequality

$$m^2(\beta_\nu\sigma^{-\nu})^{(\nu+\delta-4)/(\nu-2)} = (m^{1+\delta/2}\beta_\nu\sigma^{-\nu})^{(\nu+\delta-4)/(\nu-2)}(m^{(\nu+\delta)/2})^{\frac{2-\delta}{\nu-2}}$$

$$\leqslant \frac{\nu+\delta-4}{\nu-2}m^{1+\delta/2}\beta_\nu\sigma^{-\nu} + \frac{2-\delta}{\nu-2}m^{(\nu+\delta)/2}.$$

Substituting it into (1.24) we get for every $j = 2, \cdots, \nu-2$,

$$\mathrm{E}f_m^{j+\delta}\cdot \mathrm{E}f_{m+1,m}^{\nu-j} \leqslant 2K_\nu(m^{1+\delta/2}\beta_\nu\sigma^{\delta} + m^{(\nu+\delta)/2}\sigma^{\nu+\delta}). \tag{1.25}$$

Inequality (1.25) for $j=1$ or $\nu-1$ can be derived in a manner similar to that of $\mathrm{E}f_m^{\nu-1+\delta}$ and $\mathrm{E}f_{m+1,m}^{\nu-1}$ via $\mathrm{E}f_m \leqslant (\mathrm{E}f_m^2)^{1/2}$ and $\mathrm{E}f_{m+1,m}^{1+\delta} \leqslant (\mathrm{E}f_{m+1,m}^2)^{(1+\delta)/2}$. Substituting (1.21) and (1.25) into (1.20) we get

$$I_n \leqslant 2^{\nu+1}K_\nu(m^{1+\delta/2}\beta_\nu\sigma^{\delta} + m^{(\nu+\delta)/2}\sigma^{\nu+\delta}). \tag{1.26}$$

A similar inequality for J_n can also be obtained, and (1.19) has been proved with $M_1 = 2^{\nu+2}K_\nu$.

To prove (1.17), fix $\delta' \in (0, \delta)$, and choose a natural number m_0 such that

$$(2m/(2m-1))^{1+\delta/2} \leqslant 2^{(\delta-\delta')/2} \text{ for all } m \geqslant m_0. \tag{1.27}$$

From (1.4) there exists a constant $C_{\nu+\delta}>0$, depending only on ν and δ, such that for all integers a and k,

$$\mathrm{E}\,|S_{a,k}|^{\nu+\delta} \leqslant C_{\nu+\delta}(k\beta_{\nu+\delta} + (k\sigma^2)^{(\nu+\delta)/2}) = C_{\nu+\delta}L_{\nu+\delta,k}.$$

Hence for any a, and $n \leqslant 2m_0$, we have

$$\mathrm{E}(\max_{1\leqslant k\leqslant n}|S_{a,k}|)^{\nu+\delta} \leqslant \mathrm{E}(|S_{a,1}|^{\nu+\delta} + |S_{a,2}|^{\nu+\delta} + \cdots + |S_{a,n}|^{\nu+\delta}) \leqslant C_{\nu+\delta}nL_{\nu+\delta,n}.$$

From this and (1.14),

$$M_2 =: \max_{1\leqslant n\leqslant 2m_0}\sup_{a\in\mathbb{N}}[\mathrm{E}f_{a,n}^{\nu+\delta}/(L_{\nu+\delta,n} + n^{1+\delta/2}\beta_\nu\sigma^\delta)] \leqslant 2m_0C_{\nu+\delta} < \infty.$$

Choose $M > M_2$ such that

$$2 + M_1/M \leqslant 2^{1+\delta'/2}, \tag{1.28}$$

where M_1 is the same as in (1.19). Obviously, $M = M(\nu, \delta)$, and (1.17) holds for all $a \in \mathbb{N}$ and $n \leqslant 2m_0$.

Fix $k > 2m_0$, and assume (1.17) is true for all $a \in \mathbb{N}$ and every $n \leqslant k-1$. Then we proceed to prove that (1.17) is also true for $n = k$.

Put $k = 2m$ or $m-1$. From (1.18), (1.19), the induction hypotheses and $\mu \leqslant m$, we know that

$$\begin{aligned}
\mathrm{E}f_k^{\nu+\delta} &\leqslant \mathrm{E}f_m^{\nu+\delta} + \mathrm{E}f_{m+1,\mu}^{\nu+\delta} + M_1(m^{1+\delta/2}\beta_\nu\sigma^\delta + m^{(\nu+\delta)/2}\sigma^{\nu+\delta}) \\
&\leqslant M(m\beta_{\nu+\delta} + \mu\beta_{\nu+\delta} + 2m^{1+\delta/2}\beta_\nu\sigma^\delta + 2m^{(\nu+\delta)/2}\sigma^{\nu+\delta}) \\
&\quad + M_1(m^{1+\delta/2}\beta_\nu\sigma^\delta + m^{(\nu+\delta)/2}\sigma^{\nu+\delta}) \\
&=: M(k\beta_{\nu+\delta} + k^{1+\delta/2}\beta_\nu\sigma^\delta H(k) + k^{(\nu+\delta)/2}\sigma^{\nu+\delta}\widetilde{H}(k)),
\end{aligned} \tag{1.29}$$

where $H(k) = \left(2 + \dfrac{M_1}{M}\right)\left(\dfrac{m}{k}\right)^{1+\delta/2}$, $\widetilde{H}(k) = \left(2 + \dfrac{M_1}{M}\right)\left(\dfrac{m}{k}\right)^{(\nu+\delta)/2}$

From (1.27), (1.28) and $k = 2m$ or $2m-1$, we have

$$\widetilde{H}(k) \leqslant H(k) \leqslant 2^{1+\delta'/2}\left(\frac{m}{2m-1}\right)^{1+\delta/2} = 2^{(\delta'-\delta)/2}\left(\frac{2m}{2m-1}\right)^{1+\delta/2} \leqslant 1.$$

By this and (1.29) we get (1.17) for $n=k$, and (1.17) is true for all $a\in\mathbb{N}$ and all $n\in\mathbb{N}$.

For $\nu=2$, we get from (1.17) that

$$\mathrm{E}f_{a,n}^{2+\delta}\leqslant M(n\beta_{2+\delta}+2n^{1+\delta/2}\sigma^{2+\delta})\leqslant M'L_{2+\delta,n}, \tag{1.30}$$

where $M'=2M$.

For $\nu>2$, by Hölder's inequality and Jensen's inequality,

$$\begin{aligned}n^{1+\delta/2}\beta_\nu\sigma^\delta &\leqslant n^{1+\delta/2}\beta_{\nu+\delta}^{(\nu+2)/(\nu+\delta-2)}\sigma^{2\delta/(\nu+\delta-2)+\delta}\\ &=\sigma^{\nu+\delta}(n\beta_{\nu+\delta}\sigma^{-(\nu+\delta)})^{\frac{\nu-2}{\nu+\delta-2}}(n^{(\nu+\delta)/2})^{\frac{\delta}{\nu+\delta-2}}\\ &=\sigma^{\nu+\delta}\left(\frac{\nu-2}{\nu+\delta-2}n\beta_{\nu+\delta}\sigma^{-(\nu+\delta)}+\frac{\delta}{\nu+\delta-2}n^{\frac{\nu+\delta}{2}}\right)\leqslant L_{\nu+\delta,n}.\end{aligned}$$

By this and (1.17) we obtain

$$\mathrm{E}f_{a,n}^{\nu+\delta}\leqslant 2ML_{\nu+\delta,n}=M'L_{\nu+\delta,n}. \tag{1.31}$$

Since $-X_1,\cdots,-X_n$ are also NA variables, (1.30) and (1.31) are also true for $\tilde{f}_{a,n}$.

At last, it follows from (1.12) that

$$\mathrm{E}(\max_{1\leqslant k\leqslant n}|S_{a,k}|)^{\nu+\delta}\leqslant 2^{\nu+\delta-1}(\mathrm{E}f_{a,n}^{\nu+\delta}+\mathrm{E}\tilde{f}_{a,n}^{\nu+\delta})\leqslant 2^{\nu+\delta}M'L_{\nu+\delta,n}.$$

Putting $K_{\nu+\delta}=2^{\nu+\delta}M'=2^{\nu+\delta+1}M(\nu,\delta)$, we know that inequality (1.6) is also true for all $a\in\mathbb{N}$ and $n\in\mathbb{N}$ when $p=\nu+\delta$.

2 Weak invariance principle for strictly stationary NA sequences

At first we give an example to show that there exist non-degenerate non-independent strictly stationary NA sequences.

Example. Let $\{r_j, j\in\mathbb{N}\}$ be a sequence of positive numbers and

$$D_n=\begin{pmatrix}1 & -r_1 & -r_2 & \cdots & -r_{n-2} & -r_{n-1}\\ -r_1 & 1 & -r_1 & \cdots & -r_{n-3} & -r_{n-2}\\ -r_2 & -r_1 & 1 & \cdots & -r_{n-4} & -r_{n-3}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ -r_{n-2} & -r_{n-3} & -r_{n-4} & \cdots & 1 & -r_1\\ -r_{n-1} & -r_{n-2} & -r_{n-3} & \cdots & -r_1 & 1\end{pmatrix},\quad n\in\mathbb{N}.$$

It is well known that if $\sum_{j=1}^{\infty}r_j$ is small enough then D_n is positive definite for all $n\in\mathbb{N}$. Suppose that $\{X_j, j\in\mathbb{N}\}$ is a sequence of random variables, where $n\in\mathbb{N}$, $(X_1,\cdots,X_n)$ is an n-dimensional normal random vector with mean vector $(0,\cdots,0)$ and covariance matrix D_n for any $n\in\mathbb{N}$, then from 3.4 of ref.[4], one knows that $X_1,\cdots,X_n$ are negatively associated for each

$n \geqslant 2$. Hence, $\{X_j,\ j \in \mathbb{N}\}$ is just a sequence of non-degenerate non-independent strictly stationary NA random variables. Furthermore, suppose that $F(x)$ is any non-degenerate distribution function, and let

$$F^{-1}(y) = \inf\{x : F(x) \geqslant y\}, \quad 0 < y \leqslant 1, \quad F^{-1}(0) = F^{-1}(0_+),$$

and

$$Y_j = F^{-1}(\Phi(X_j)), j \in \mathbb{N},$$

where Φ is the distribution function of the standard normal random variable. Since Y_j is a non-decreasing function of X_j for every $j \in \mathbb{N}$, from Lemma 1, $\{Y_j,\ j \in \mathbb{N}\}$ is a sequence of non-independent non-normal strictly stationary random variables.

The main result of this section is as follows:

Theorem 3 (Weak invariance principle). *Let $\{X_j,\ j \in \mathbb{N}\}$ be a sequence of strictly stationary NA random variables such that*

$$\mathrm{E}X_1 = b,\ 0 < \mathrm{var}X_1 = \sigma^2 < \infty,\ A^2 = \mathrm{var}X_1 + 2\sum_{j=2}^{\infty}\mathrm{cov}(X_1, X_j) > 0. \tag{2.1}$$

Put $S_m = \sum_{j=1}^{m} X_j$ and

$$W_n(t) = \frac{1}{A\sqrt{n}}[S_m + (nt - m)X_{m+1} - ntb],$$

where $m \leqslant nt < m + 1$, $0 \leqslant t \leqslant T$. Then the sequence $\{W_n,\ n \in \mathbb{N}\}$ of stochastic process converges weakly in $\mathrm{C}[0, T]$ to a standard Wiener process.

Proof. Following the proof in ref. [9] for PA case, we can conclude that any finite-dimensional distribution of W_n weakly converges to the corresponding distribution of a standard Wiener process. The details are omitted.

Now we only need to prove the tightness of $\{W_n\}$. Note that the proof for the NA case is quite different from that for the PA case. Without loss of generality, we can take $b = 0$. From the remark to Theorem 8.4 in ref. [12], we know that for the proof of tightness of strictly stationary sequences with zero mean, it suffices to prove that for any $\varepsilon > 0$, there exist $\lambda > 1$ and $n_1 \in \mathbb{N}$ such that when $n > n_1$, we have

$$\mathrm{P}\Big(\max_{1\leqslant k\leqslant n} |S_k| \geqslant \lambda A\sqrt{n}\Big) \leqslant \varepsilon\lambda^{-2}. \tag{2.2}$$

Put

$$Z_j(n) = Z_j = -\sqrt{n}I(X_j < -\sqrt{n}) + X_j I(|X_j| \leqslant \sqrt{n}) + \sqrt{n}I(X_j > \sqrt{n}),$$

$$T_k = \sum_{j=1}^{k} Z_j; \qquad k = 1, \cdots, n.$$

Since Z_j is a non-decreasing function of X_j, it is obvious that $Z_1, \cdots, Z_n$ are also NA variables.

From $\mathrm{E}X_1^2 < \infty$,

$$\sqrt{n}\mathrm{E}Z_1(n) \to 0, \quad n \to \infty.$$

So there exists $n_1 \in \mathbb{N}$ such that for $n \geqslant n_1$, we have $|\sqrt{n}\mathrm{E}Z_1(n)| < \frac{A}{2}$ and

$$\mathrm{P}\left(\max_{1 \leqslant k \leqslant n} |S_k| \geqslant \lambda A \sqrt{n}\right) \leqslant \mathrm{P}\left(\bigcup_{j=1}^{n} (|X_j| > \sqrt{n})\right) + \mathrm{P}\left(\max_{1 \leqslant k \leqslant n} |T_k| \geqslant \lambda A \sqrt{n}\right)$$

$$\leqslant n\mathrm{P}(|X_1| > \sqrt{n}) + \mathrm{P}\left(\max_{1 \leqslant k \leqslant n} |T_k - \mathrm{E}T_k| \geqslant \frac{\lambda}{2} A \sqrt{n}\right) =: I_1 + I_2.$$

By inequality (1.6) with $p = 3$, we can get

$$I_2 \leqslant 8K_3 A^{-3} \lambda^{-3} (2\sigma^2 + \sigma^3).$$

Hence for any given $\varepsilon > 0$,

$$I_2 < \frac{\varepsilon}{2\lambda^2},$$

whenever $\lambda > \max(1, 16K_3(2\tau^2 + \tau^3)A^{-3}\varepsilon^{-1})$, $n \geqslant n_1$.

Furthermore, since $\mathrm{E}X_1^2 < \infty$, we have

$$I_1 \leqslant \mathrm{E}(X_1^2 I(|X_1| > \sqrt{n})) < \frac{\varepsilon}{2\lambda^2}$$

for any given $\varepsilon > 0$ and the above $\lambda > 1$, when n is large enough.

This completes the proof.

References

1 Esary, J., Proschan, F., Walkup, D., Association of random variables with applications, *Ann. Math. Statist.*, 1967, 38: 1466.

2 Lehmann, E. L., Some concepts of dependence, *Ann. Math. Statist.*, 1966, 37: 1137.

3 Joag-Dev, K., Proschan, F., Negative association of random variables with applications, *Ann. Statist.*, 1983, 11: 286.

4 Block, H. W., Savits, T. H., Shaked, M., Some concepts of negative dependence, *Ann. Probab.*, 1982, 10: 765.

5 Newman, C. L., Asymptotic independence and limit theorems for positively and negatively dependent random variables, in *Inequalities in Statistics and Probability* (ed. Tong, Y. L.), *IMS Lecture Notes-Monograph Series*, Vol. 5, 1984, 127—140.

6 Matula, P., A note on the almost sure convergence of sums of negatively dependent random variables, *Statist. Probab. Lett.*, 1992, 15: 209.

7 Petrov, V. V., *Limit Theorems of Sums of Independent Random Variables* (in Russian), Moscow: Nauka, 1987.

8 Su Chun, A theorem of Hsu-Robbins type for negatively associated sequence, *Chinese Science Bulletin*, 1996, 41(6): 441.

9 Newman, C. M., Wright, A. L., An invariance principle for certain dependent sequences, *Ann. Probab.*, 1981, 9: 671.

10 Nagaev, S. V., Large deviation of sums of independent random variables, *Ann. Probab.*, 1979, 7: 745.

11 Stout, W., *Almost Sure Convergence*, New York: Academic Press, 1974, 204—209.

12 Billingsley, P., *Convergence of Probability Measures*, New York: John Wiley, 1968.