Statistics and Probability Letters 109 (2016) 22–29

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Asymptotic results for random sums of dependent random variables

Ümit Işlak

Georgia Institute of Technology, Mathematics Department, School of Mathematics Atlanta, GA 30332, United States

ARTICLE INFO

Article history:
Received 30 August 2015
Received in revised form 15 October 2015
Accepted 17 October 2015
Available online 10 November 2015

Keywords:
Stein's method
Random sums
Central limit theorem
Concentration inequality
Local dependence

ABSTRACT

Our main result is a central limit theorem for random sums of the form $\sum_{i=1}^{N_n} X_i$, where $\{X_i\}_{i\geq 1}$ is a stationary m-dependent process and N_n is a random index independent of $\{X_i\}_{i\geq 1}$. This extends the work of Chen and Shao on the i.i.d. case to a dependent setting and provides a variation of a recent result of Shang on m-dependent sequences. Further, a weak law of large numbers is proven for $\sum_{i=1}^{N_n} X_i$, and the results are exemplified with applications on moving average and descent processes.

1. Introduction

In the following, we analyze the asymptotic behavior of random sums of the form $\sum_{i=1}^{N_n} X_i$ as $n \to \infty$, where X_i's are non-negative random variables that are stationary and m-dependent, and N_n is a non-negative integer valued random variable independent of X_i's (below, sums such as $\sum_{i=1}^{0} X_i$ are considered as empty sums and their values are set to be zero). Limiting distribution of random sums of independent and identically distributed (i.i.d.) random variables is well studied, see Chen and Shao (2007a), Kläver and Schmitz (2006), Robbins (1948) and the references therein. There is also considerable amount of work on the asymptotic normality of deterministic sums of m-dependent random variables. See, for example, Bergström (1970), Hoeffding and Robbins (1948) and Orey (1958). To the best of author's knowledge, previous work on the case of random sums of the form $\sum_{i=1}^{N_n} X_i$ where X_i's are dependent are limited to Shang (2012) where he works on sums of m-dependent random variables and to Barbour and Xia (2006) where they investigate random variables that appear as a result of integrating a random field with respect to point processes. Our results here will be in the lines of Chen and Shao (2007a) generalizing their work to an m-dependent setting. Throughout the way, we will also obtain a variation of Shang's distributional approximation result in Shang (2012) and will prove a weak law of large numbers for random sums with dependent summands.

Let us now recall stationary and m-dependent processes. Let $\{X_i\}_{i\geq 1}$ be a stochastic process and let $F_X(X_{i_1+m}, \ldots, X_{i_k+m})$ be the cumulative distribution function of the joint distribution of $\{X_i\}_{i\geq 1}$ at times $i_1 + m, \ldots, i_k + m$. Then $\{X_i\}_{i\geq 1}$ is said to be *stationary* if the identity

$$F_X(X_{i_1+m}, \ldots, X_{i_k+m}) = F_X(X_{i_1}, \ldots, X_{i_k})$$

holds for all k, m and for all $i_1, \ldots, i_k$. For more on stationary processes, we refer to Shiryaev (1996). If we define the distance between two subsets of A and B of $\mathbb{N}$ by

$$\rho(A, B) := \inf\{|i - j| : i \in A, j \in B\},$$

E-mail addresses: islak002@umn.edu, uislak6@gatech.edu.

http://dx.doi.org/10.1016/j.spl.2015.10.015
0167-7152/

then the sequence $\{X_i\}_{i\geq 1}$ is said to be *m-dependent* if $\{X_i, i \in A\}$ and $\{X_j, j \in B\}$ are independent whenever $\rho(A, B) > m$ for $A, B \subset \mathbb{N}$.

An example of a stationary m-dependent process can be given by the moving average process. Assume that $\{T_i\}_{i\geq 1}$ is a sequence of i.i.d. random variables with finite mean μ and finite variance σ^2. Letting $X_i = (T_i + T_{i+1})/2$, $\{X_i\}_{i\geq 1}$ is a stationary 1-dependent process with $\mathbb{E}[X_i] = \mu$, $\mathrm{Var}(X_i) = \sigma^2/2$ and $\mathrm{Cov}(X_1, X_2) = \sigma^2/4$.

Let us now introduce some notation that will be used throughout this paper. We will be using $\longrightarrow_d$ and $\longrightarrow_{\mathbb{P}}$ for convergence in distribution and convergence in probability, respectively. Also $N(0, 1)$ will denote a standard normal random variable and Φ will be its cumulative distribution function.

The rest of the paper is organized as follows: In the next section, we state our main results and discuss two examples on moving average and descent processes. Proofs of the results are given in Section 3.

2. Results

We have the following notation and assumptions in this section: $\{X_i\}_{i\geq 1}$ is a stationary m-dependent process with $\mu := \mathbb{E}[X_1] \neq 0$, $\sigma^2 := \mathrm{Var}(X_1) \in (0, \infty)$, $a_j := \mathrm{Cov}(X_1, X_{1+j})$. $\{Y_i\}_{i\geq 1}$ is a sequence of i.i.d. non-negative integer valued random variables with $\nu := \mathbb{E}[Y_1] \in (0, \infty)$, $\tau^2 := \mathrm{Var}(Y_1) \in (0, \infty)$. We assume that $\{X_i\}_{i\geq 1}$ and $\{Y_i\}_{i\geq 1}$ are also independent. For given $n \geq 1$, we define $N_n = \sum_{i=1}^n Y_i$.

We begin the main discussion with two propositions.

Proposition 2.1. *For any fixed $N \geq 1$, we have*

$$\mathrm{Var}\left(\sum_{i=1}^N X_i\right) = N\sigma^2 + 2\sum_{j=1}^m (N-j)a_j \mathbb{1}(N \geq j+1).$$

When $N \geq m+1$, this reduces to $\mathrm{Var}\left(\sum_{i=1}^N X_i\right) = N\sigma^2 + 2\sum_{j=1}^m (N-j)a_j$.

Proof of Proposition 2.1 follows immediately from decomposing the left-hand side as $\mathrm{Var}\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^N \mathrm{Var}(X_i) + 2\sum_{1\leq i<j\leq N} \mathrm{Cov}(X_i, X_j)$. Details are omitted.

Proposition 2.2. *We have* $\mathbb{E}\left[\sum_{i=1}^{N_n} X_i\right] = n\nu\mu$ *and*

$$\mathrm{Var}\left(\sum_{i=1}^{N_n} X_i\right) = n\left(\nu\sigma^2 + 2\nu\sum_{j=1}^m a_j + \mu^2\tau^2\right) + \alpha(m), \tag{2.1}$$

where $\alpha(m) = -2\sum_{j=1}^m ja_j - 2\sum_{j=1}^m \sum_{k=0}^j (k-j)a_j \mathbb{P}(N_n = k)$. In particular, $\alpha(m) \longrightarrow -2\sum_{j=1}^m ja_j$ as $n \to \infty$. When X_i's are also independent (i.e., $m = 0$), (2.1) simplifies to $\mathrm{Var}\left(\sum_{i=1}^{N_n} X_i\right) = n(\nu\sigma^2 + \mu^2\tau^2)$.

Proof of Proposition 2.2 is standard and is given at the end of Section 3.

Now we are ready to present our first result which is a central limit theorem for random sums of m-dependent random variables. The proof is a direct generalization of Chen and Shao's result for the i.i.d. case, and in return we recover their result from Chen and Shao (2007a) (see (2.2)).

Theorem 2.3. *Assume that $\sigma^2 + 2\sum_{j=1}^m a_j > 0$, $\mathbb{E}|X_1|^3 < \infty$ and $\mathbb{E}|Y_1|^3 < \infty$. Then we have*

$$\frac{\sum_{i=1}^{N_n} X_i - n\mu\nu}{\sqrt{n(\nu\sigma^2 + 2\nu\sum_{j=1}^m a_j + \mu^2\tau^2)}} \longrightarrow_d N(0, 1)$$

as $n \to \infty$. When $m = 0$, we have the stronger conclusion

$$\sup_{z\in\mathbb{R}} \left| \mathbb{P}\left(\frac{\sum_{i=1}^{N_n} X_i - n\mu\nu}{\sqrt{n(\nu\sigma^2 + \mu^2\tau^2)}} \leq z \right) - \Phi(z) \right| \leq \frac{C}{\sqrt{n}} \left(\frac{\tau^2}{\nu^2\sqrt{n}} + \frac{\mathbb{E}[Y_1^3]}{\tau^3} + \frac{\mathbb{E}|X_1|^3}{\nu^{1/2}\sigma^3} + \frac{\sigma}{|\mu|\sqrt{\nu}} \right) \tag{2.2}$$

for any $n \geq 1$, where C is a constant independent of n.