Journal of the Korean Statistical Society 45 (2016) 285–294

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

A moderate deviation for associated random variables

Tonguç Çağın, Paulo Eduardo Oliveira *, Nuria Torrado

CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal

ARTICLE INFO

Article history:
Received 15 July 2014
Accepted 5 November 2015
Available online 27 November 2015

AMS 2000 subject classifications:
primary 60E15
secondary 60F10

Keywords:
Moderate deviation
Association
Coupling
Approximation

ABSTRACT

Moderate deviations are an important topic in many theoretical or applied statistical areas. We prove two versions of a moderate deviation for associated and strictly stationary random variables with finite moments of order $q > 2$. The first one uses an assumption depending on the rate of a Gaussian approximation, while the second one discusses more natural assumptions to obtain the approximation rate. The control of the dependence structure relies on the decay rate of the covariances, for which we assume a relatively mild polynomial decay rate. The proof combines a coupling argument together with a suitable use of Berry–Esséen bounds.

1. Introduction

Sums of random variables have always been a central subject in the probabilistic literature, with a special interest on their asymptotics. Among results on this topic the important Central Limit Theorems (CLT) describe the limiting distributional behavior of such sums, providing useful approximate descriptions of the tail probabilities. These, besides their natural theoretical interest, are extremely relevant in statistical applications. There is, however, a limitation inherent to the properties of convergence in distribution, requiring that the tails considered through the limiting process should be computed at points that behave like the variance. More specifically, if the random variables $X_n,\ n \geq 1$, are assumed centered and we define $S_n = X_1 + \cdots + X_n$, $s_n^2 = \mathrm{E}S_n^2$, the CLT provide the approximation of $\mathrm{P}(S_n > xs_n)$ by $N(x) = 1 - \Phi(x)$, for $x > 0$ fixed, where Φ is the distribution function of a standard Gaussian variable. If we allow x to depend on n, converging to infinity, then the approximation for $\mathrm{P}(S_n > xs_n)$ is known as a moderate or large deviation, depending on how fast x grows to infinity, moderate deviations corresponding to the case where $x = O(s_n)$. In such cases where x is growing to infinity, the approximating function N is no longer necessarily the tail of a standard Gaussian, depending on the growth rate of x to infinity.

First large deviations were proved by Ibragimov and Linnik (1971), Linnik (1961), Nagaev (1965, 1969) or Rozovski (1982) for independent and identically distributed variables. We refer the reader to the survey paper by Nagaev (1979) for a nice account of these early results. The techniques of proof were based on suitable exponential bounds, the so called Fuk–Nagaev inequalities, on the tail probabilities. A typical result, given in Theorem 1.9 in Nagaev (1979), states that

$$\mathrm{P}(S_n > xs_n) = (1 - \Phi(x) + n\mathrm{P}(X_1 > xs_n))\,(1 + o(1)), \tag{1}$$

provided that $x \geq 1$, $s_n = n^{1/2}$ and the right tail of the X_n's is a regularly varying function. Extensions of such results have been recently proved by Peligrad, Sang, Zhong, and Wu (2012) considering weighted sums $\tilde{S}_n = \sum c_{n,i}X_i$ instead of

* Corresponding author.
E-mail address: paulo@mat.uc.pt (P.E. Oliveira).

http://dx.doi.org/10.1016/j.jkss.2015.11.004

S_n. These authors prove a result similar to (1) under essentially the same assumptions on the random variables (i.i.d. and regularly varying tails) and a regularity condition on the weights: $\max_i c_{n,i}/\mathrm{E}\tilde{S}_n^2 \longrightarrow 0$. The proof of this extension relies on moderate or large deviations for triangular arrays of random variables and convenient strong approximations between the tails of $\tilde{S}_n$ and the sums of tails of the X_n's, much in the same spirit of the proof technique used in Theorem 1.9 in Nagaev (1979). Going back to early results, moderate or large deviations for triangular arrays of row-wise independent variables were considered by Amosova (1972), Rubin and Sethuraman (1965), Slastnikov (1979) or, more recently, by Frolov (2005). All the results mentioned so far characterize the tail probabilities directly. Concerning large deviations, that is, x growing fast to infinity, a lot of attention was given to the logarithms of the tail probabilities instead, thus providing exponential bounds for the tail probabilities themselves. The bound for these logarithms appears then as the Fenchel–Légendre transform of the normalized logarithm of the Laplace transform of S_n (notice that we are now back to non-weighted sums). A good account of results in this direction can be found in the book by Dembo and Zeitouni (1998). The interest on logarithmic tails meant that there are much fewer results available in the non-logarithmic scale in recent literature, particularly for weighted sums. Another recent direction of development is concerned with dependent variables. Here, available results seem even more scarce. Looking at large deviations, some results were proved by Bryc (1992), Bryc and Dembo (1996) or Nummelin (1990) considering mixing variables or, Henriques and Oliveira (2008) for associated random variables. These authors were interest on logarithmic scale results and their proof techniques relied on suitable exponential bounds and required a rather fast decay rate on the coefficients characterizing the dependence structure, meaning they should decrease faster than geometrically. More recently, for mixing variables Merlevède, Peligrad, and Rio (2009) relaxed the assumption on the mixing coefficients, requiring only a geometric decay rate to prove the large deviation. Their proof technique, called by the authors a "Cantor set construction", adapts the block decomposition of sums, popular for proving CLT, to large deviations. These authors have more recently extended their results to other forms of weak dependent variables (see Merlevède, Peligrad, & Rio, 2011). Efforts in the non-logarithmic scale for dependent variables were made by Grama (1997), Grama and Haeusler (2006) for martingales, Wu and Zhao (2008) for stationary processes, Tang (2006) for negatively dependent variables or Lui (2009) for negatively dependent heavy tailed variables.

In this paper we present a moderate deviation in the non-logarithmic scale for sums of associated random variables. In Section 2 we give some definitions and recall some auxiliary results, in Section 3 we prove a first moderate deviation based on an assumption depending on a Gaussian approximation. In Section 4 we discuss this approximation issue, giving a general moderate deviation based on more natural assumptions. Finally, in Section 5 we give an application to moving averages of our results.

2. Framework and auxiliary results

To define appropriately our framework let $X_n,\ n \geq 1$, be strictly stationary, centered and associated random variables with finite variances. Denote $S_n = X_1 + \cdots + X_n$ and $s_n^2 = \mathrm{E}S_n^2$. Recall that association means that for any $m \geq 1$ and any two real-valued coordinatewise nondecreasing functions f and g,

$$\mathrm{Cov}\Big(f\,(X_1,\ldots,X_m)\,,\, g\,(X_1,\ldots,X_m)\Big) \geq 0,$$

whenever this covariance exists. For basic results on associated random variables, please refer to the monographs Bulinski and Shashkin (2007), Oliveira (2012) or Prakasa Rao (2012). It is well known that the covariance structure of associated random variables characterizes their asymptotics, so it is natural to seek assumptions on the covariances.

A common assumption when proving Central Limit Theorems is $\frac{1}{n}s_n^2 \longrightarrow \sigma^2 > 0$ (see, for example, Newman & Wright, 1981, 1982 or Oliveira & Suquet, 1995, 1996), so we will be assuming that this is fulfilled in the sequel. Notice this assumption implies that $s_n^2 \sim n\sigma^2$. Finally, define the Cox–Grimmett coefficients, commonly used to control dependence for associated variables:

$$u(n) = \sum_{k=n}^{\infty} \mathrm{Cov}(X_1, X_k). \tag{2}$$

Our proof will rely on a suitable approximation to independent variables that will be chosen so they satisfy the moderate deviation we want to extend. We quote next a result by Frolov (2005), providing a moderate deviation for triangular arrays of row-wise independent random variables. This will be the tool to prove the moderate deviation for the approximating variables.

Theorem 2.1 (*Theorem 1.1 in Frolov, 2005*)**.** *Let $X_{n,k},\ k = 1,\ldots,k_n,\ n \geq 1$, be an array of row-wise independent variables with $F_{n,k}(y) = \mathrm{P}(X_{n,k} \leq y)$, $\mathrm{E}X_{n,k} = 0$ and $\mathrm{E}X_{n,k}^2 = \sigma_{n,k}^2 < \infty$. Denote $T_n = \sum_{k=1}^{k_n} X_{n,k}$ and $B_n = \sum_{k=1}^{k_n} \sigma_{n,k}^2$. For $q > 2$, let $\beta_{n,k} = \int_0^{\infty} y^q\, F_{n,k}(dy) < +\infty$, and define*

$$M_n = \sum_{k=1}^{k_n} \beta_{n,k} \quad \textit{and} \quad L_n = B_n^{-q/2} M_n.$$