Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 78 (2008) 84–89
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Localized large sums of random variables

Kevin Ford[a,*,1], Gérald Tenenbaum[b]

[a]Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green St., Urbana, IL 61801, USA
[b]Institut Élie Cartan, Université Henri– Poincaré Nancy 1, B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex, France

Received 13 September 2006; received in revised form 14 March 2007; accepted 23 May 2007
Available online 8 June 2007

Dedicated to the memory of Walter Philipp

Abstract

We study large partial sums, localized with respect to the sums of variances, of a sequence of centered random variables. An application is given to the distribution of prime factors of typical integers.

MSC: 60G50; 11N25

Keywords: Sums of random variables; Prime divisors

1. Introduction

Consider random variables $X_1, X_2, \ldots$ with $\mathbb{E}X_j = 0$ and $\mathbb{E}X_j^2 = \sigma_j^2$. Let

$$S_n = X_1 + \cdots + X_n, \quad s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2,$$

and assume that (a) $s_n \to \infty$ as $n \to \infty$.

Given a positive function $f_N \geqslant 1 + 1/N$, we are interested in the behavior of

$$I = \liminf_{N\to\infty} \max_{N < s_n^2 \leqslant Nf_N} |S_n|/s_n.$$

If we replace lim inf by lim sup, it immediately follows from the law of the iterated logarithm that $I = \infty$ almost surely when f_N is bounded. Our results answer a question originally raised, in oral form, by A. Sárközy and for which a partial answer had previously been given by the second author, see Chap. 3 of Oon (2005, Chapter 3).

*Corresponding author.
E-mail addresses: ford@math.uiuc.edu (K. Ford), gerald.tenenbaum@iecn.u-nancy.fr (G. Tenenbaum).
[1]Research supported by National Science Foundation Grants DMS-0301083 and DMS-0555367.

0167-7152/$ - see front matter
doi:10.1016/j.spl.2007.05.006

2. Independent random variables

Assume that the X_j are independent. Then $\mathbb{E}S_n^2 = s_n^2$. In addition to condition (a), we will work with two other mild assumptions, (b) $s_{j+1}/s_j \ll 1$ when $s_j > 0$ and (c) for every $\lambda > 0$, there is a constant $c_\lambda > 0$ such that if n is large enough and $s_m^2 > 2s_n^2$, then

$$\mathbb{P}(|S_m - S_n| \geqslant \lambda s_m) \geqslant c_\lambda.$$

Condition (b) says that no term in S_n dominates the others. Condition (c) follows if the central limit theorem (CLT) holds for the sequence of S_n, since CLT for S_n implies CLT for $S_m - S_n$ as $(m-n) \to \infty$. For example, (c) holds for i.i.d. random variables, under the Lindeberg condition

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \sum_{1 \leqslant j \leqslant n} \mathbb{E}(X_j^2/s_n^2 \colon |X_j| > \varepsilon s_n) = 0$$

and the stronger Lyapunov condition

$$\exists \delta > 0 \colon \sum_{1 \leqslant j \leqslant n} \mathbb{E}|X_j|^{2+\delta} = \mathrm{o}(s_n^{2+\delta}).$$

Condition (c) is weaker, however, than CLT.

Theorem 1. (i) *Suppose* (*a*), (*b*), *and* $f_N = (\log N)^M$ *for some constant* $M > 0$. *Then* $I < \infty$ *almost surely.*
(ii) *Suppose* (*a*), (*b*), (*c*) *and* $f_N = (\log N)^{\xi(N)}$ *with* $\xi(N)$ *tending monotonically to* ∞. *Then* $I = \infty$ *almost surely.*

Remark. In the first statement of the theorem we show in fact that almost surely $I \leqslant 15\sqrt{M+1}(\max_{s_j>0} s_{j+1}/s_j)^2$.

Lemma 2 (*Kolmogorov's inequality, 1929*). *We have*

$$\mathbb{P}(\max_{1\leqslant j\leqslant k} |S_j| \geqslant \lambda s_k) \leqslant 1/\lambda^2 \quad (k \geqslant 1).$$

Proof of Theorem 1. By (a) and (b), there is a constant D so that $s_{j+1}/s_j \leqslant D$ for all large j. Define

$$h(n) := \max\{k \colon s_k^2 \leqslant n\} \quad (n \in \mathbb{N}^*)$$

so that the conditions $N < s_n^2 \leqslant Nf_N$ and $h(N) < n \leqslant h(Nf_N)$ are equivalent.

We first consider the case when $f_N := (\log N)^M$. Let

$$N_j := j^{(M+3)j}, \quad t(j) := \lfloor (M+1)(\log j)/\log 2 \rfloor, \quad H_j := 2^{t(j)},$$

and

$$U_j := h(N_j), \quad U_{j,t} := h(2^t N_j) \quad (0 \leqslant t \leqslant t(j)), \quad V_j := h(H_j N_j) = U_{j,t(j)}.$$

It is possible that $U_{j,t+1} = U_{j,t}$ for some t. Note that for large j, $H_j N_j \geqslant N_j f_{N_j}$.

Let k be a constant depending only on M and D. For $j \geqslant 1$ define the events

$$A_j := \{|S_{V_j}| \leqslant s_{U_{j+1}}\},$$

$$B_j := \bigcap_{0 \leqslant t \leqslant t(j)-1} B_{j,t} \quad \text{where } B_{j,t} := \left\{ \max_{U_{j+1,t} \leqslant n \leqslant U_{j+1,t+1}} |S_{U_{j+1,t+1}} - S_n| \leqslant k s_{U_{j+1,t}} \right\},$$

$$C_j := \{|S_{U_{j+1}} - S_{V_j}| \leqslant 2 s_{U_{j+1}}\}.$$

By (b) and the definition of $h(N)$, we have

$$D^{-1}\sqrt{2^t N_j} \leqslant s_{U_{j,t}} \leqslant \sqrt{2^t N_j} \tag{1}$$