Statistics & Probability Letters 49 (2000) 113–118

STATISTICS & PROBABILITY LETTERS

www.elsevier.nl/locate/stapro

A note on logarithmic tail asymptotics and mixing

Nina Gantert

Fachbereich Mathematik, Technische Universität Berlin, MA 7-4, Straße des 17. Juni 136, D-10623 Berlin, Germany

Received November 1998; received in revised form October 1999

Abstract

Let $Y_1, Y_2, \ldots$ be a stationary, ergodic sequence of non-negative random variables with heavy tails. Under mixing conditions, we derive logarithmic tail asymptotics for the distributions of the arithmetic mean. If not all moments of Y_1 are finite, these logarithmic asymptotics amount to a weaker form of the Baum–Katz law. Roughly, the sum of i.i.d. heavy-tailed non-negative random variables has the same behaviour as the largest term in the sum, and this phenomenon persists for weakly dependent random variables. Under mixing conditions, the rate of convergence in the law of large numbers is, as in the i.i.d. case, determined by the tail of the distribution of Y_1. There are many results which make these statements more precise. The paper describes a particularly simple way to carry over logarithmic tail asymptotics from the i.i.d. to the mixing case. © 2000 Elsevier Science B.V. All rights reserved

MSC: 60F10; 60G10; 60G50

Keywords: Baum–Katz law; Rate of convergence in the strong law; Mixing; Absolutely regular sequence; Regularly varying tail; Semiexponential distributions; β-mixing

1. Introduction

Throughout this note, $Y_1, Y_2, \ldots$ will be non-negative random variables, $S_n := \sum_{i=1}^n Y_i$, and we will denote $m_0 := E[Y_1] \leqslant \infty$. We recall the Baum–Katz law, see Baum and Katz (1965) or Peligrad (1985).

Theorem 1. *Assume $Y_1, Y_2, \ldots$ are i.i.d. Let $p > 1$. Then, the following two statements are equivalent*:

(i) $E[Y_1^p] < \infty$.

(ii) *For each $\psi \geqslant 1$ and for each $m > m_0$, we have*

$$\sum_{n=1}^{\infty} n^{\psi p-2} P[S_n \geqslant n^{\psi} m] < \infty. \tag{0}$$

Theorem 1 (in a more general form) was extended under conditions on the ρ-mixing coefficients or the ϕ-mixing coefficients in Peligrad (1985). We will extend a weaker form of Theorem 1 under conditions on

E-mail address: gantert@math.tu-berlin.de (N. Gantert)

PII: S0167-7152(00)00037-7

the β-mixing coefficients. Since the β-mixing coefficients are always smaller than the ϕ-mixing coefficients, we have to assume a decay which is faster than logarithmic as in Peligrad (1985). In Section 3, we give a counterexample (where $\beta(n)=n^{-c}$ for a constant c) to stress that a logarithmic decay of β-mixing coefficients is in general not sufficient. Our assumption neither implies nor does it follow from the assumptions in Peligrad (1985), Daley et al. (1996) or Kiesel (1997).

Let us first give a weaker form of Theorem 1 in terms of logarithmic asymptotics.

Theorem 2. *Assume $Y_1, Y_2, \dots$ are i.i.d. Let $\alpha > 1$. Then the following holds.*

(a) *If $E[Y_1^\alpha] = \infty$, and $P[Y_1 > t] \geqslant L(t)/t^\alpha$ for some slowly varying function L, then for $m > 0$ and $\psi \geqslant 1$*

$$\liminf_n \frac{1}{\log n} \log P[S_n \geqslant n^\psi m] \geqslant -(\alpha\psi - 1). \tag{1}$$

(b) *If $E[Y^p] < \infty$ for each $p < \alpha$, then for $m > m_0$ and $\psi \geqslant 1$*

$$\limsup_n \frac{1}{\log n} \log P[S_n \geqslant n^\psi m] \leqslant -(\alpha\psi - 1). \tag{2}$$

It is not difficult to prove Theorem 2 directly, i.e. without making use of Theorem 1. We conjecture that (1) holds true without the additional assumption on Y_1. But, even then, Theorem 2 is weaker than Theorem 1 because from Theorem 2, one does not know if the sum in (0) converges or diverges at the critical point $p_c := \sup\{p: E[Y_1^p] < \infty\}$. More precisely, Theorem 2 is equivalent to the following.

Theorem 3. *Assume $Y_1, Y_2, \dots$ are i.i.d. Let $p_c := \sup\{p: E[Y_1^p] < \infty\}$ and assume $p_c > 1$. Then the following holds.*

(a) *If $P[Y_1 > t] \geqslant L(t)/t^{p_c}$ for some slowly varying function L and if $p > p_c$, then for $m > 0$ and $\psi \geqslant 1$*

$$\sum_{n=1}^{\infty} n^{\psi p - 2} P[S_n \geqslant n^\psi m] = \infty. \tag{3}$$

(b) *If $p < p_c$, then for $m > m_0$ and $\psi \geqslant 1$*

$$\sum_{n=1}^{\infty} n^{\psi p - 2} P[S_n \geqslant n^\psi m] < \infty. \tag{4}$$

The aim of this note is to show how logarithmic asymptotics, as given in Theorem 2, can be extended to a stationary, mixing sequence $Y_1, Y_2, \dots$ under assumptions on the decay of the β-mixing coefficients. We use only β-mixing coefficients; for the definition of other mixing coefficients and for the relations between them, we refer to Bradley (1986). We recall the definition of the β-mixing coefficient. For two random variables X and Z, denote the distribution of (X,Z) by $\mu_{(X,Z)}$ and the distributions of X and Z by μ_X and μ_Z. The β-mixing coefficient of X and Z is defined by

$$\beta(X,Z) := \tfrac{1}{2} \| \mu_{(X,Z)} - \mu_X \otimes \mu_Z \|, \tag{5}$$

where $\|\mu - v\|$ denotes the (total) variation norm of the signed measure $\mu - v$. Now, for a sequence of random variables $(Y_1, Y_2, \dots)$, define

$$\beta(n) := \sup_{k \in \mathbb{N}} \beta((Y_1, \dots, Y_k), (Y_{k+n}, Y_{k+n+1}, \dots)). \tag{6}$$

The sequence is called absolutely regular if $\beta(n) \to 0$ for $n \to \infty$, cf. Bradley (1986). Our main results are the following two theorems.

Theorem 4. *Assume $Y_1, Y_2, \dots$ is a stationary sequence such that $\log \beta(n)/\log n \to -\infty$. Let $\alpha > 1$. Then* (a) *and* (b) *in Theorem 2 hold.*

Theorem 5. *Let a_1, a_2 and L be slowly varying functions, with $L(t)/\log t \to \infty$ for $t \to \infty$ and*

$$\liminf_{t\to\infty} \frac{L(t^\gamma)}{L(t)} = \limsup_{t\to\infty} \frac{L(t^\gamma)}{L(t)} =: h(\gamma) > 0 \tag{7}$$

for $\gamma \in]0,1[$. Assume that $\log \beta(n)/L(n) \to -\infty$. Then the following hold:

(a) *If $P[Y_1 \geqslant t] \geqslant a_1(t)\exp(-L(t))$, then for $m > 0$ and $\psi \geqslant 1$*

$$\liminf_{n} \frac{1}{L(n^\psi)} \log P[S_n \geqslant n^\psi m] \geqslant -1. \tag{8}$$

(b) *If $P[Y_1 \geqslant t] \leqslant a_2(t)\exp(-L(t))$, then $m_0 = E[Y_1] < \infty$ and for $m > m_0$ and $\psi \geqslant 1$*

$$\limsup_{n} \frac{1}{L(n^\psi)} \log P[S_n \geqslant n^\psi m] \leqslant -1. \tag{9}$$

Our mixing conditions are satisfied in many circumstances. Often, one has even an exponential decay of the β-mixing coefficients; for instance, every Doeblin–Markov chain is ϕ-mixing with $\phi(n) \to 0$ exponentially fast, hence $\beta(n) \to 0$ exponentially fast, cf. Bradley (1986).

In the i.i.d. case, one can of course give very precise expansions for $P[S_n \geqslant nm]$ instead of just logarithmic asymptotics. For instance, if the distribution of Y_1 has a regularly varying tail and $m_0 < \infty$, it is known that for $m > m_0$,

$$P[S_n \geqslant nm] \sim P\left[\max_{1\leqslant i\leqslant n} Y_i \geqslant n(m-m_0)\right] \sim nP[Y_1 \geqslant n(m-m_0)] \tag{10}$$

(see for instance Vinogradov, 1994).

If the distribution of Y_1 is semiexponential as in the setting of Theorem 5, precise expansions for $P[S_n \geqslant nm]$ can be found in Rozovskii (1994).

We show here only that, under mixing conditions, the *logarithmic* asymptotics are the same as in the i.i.d. case. One might ask about the conditions to carry over (10), possibly with some constants, to the mixing case. Davis and Hsing (1995) have some results in this direction. Example 1 shows, however, that if there is not enough mixing, *not even* the logarithmic asymptotics are in general the same as in the i.i.d. case.

For simplicity, we consider only non-negative random variables. However, it is obvious from the proofs that the results can be extended under suitable assumptions on the lower tail of the distribution of Y_1.

2. Proofs

Proof of Theorem 4. We will use the following decoupling lemma due to H. Berbee.

Lemma 1. *Let $X_1, X_2, \ldots, X_n$ be random variables on a probability space $(\Omega, \mathscr{A}, P)$ and let $\beta_k := \beta((X_1,\ldots,X_k), (X_{k+1},\ldots,X_n))$. Then there exist independent random variables $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ on the same probability space such that $\tilde{X}_i$ and X_i have the same distribution $(1 \leqslant i \leqslant n)$ and*

$$\|\mu_{(X_1,X_2,\ldots,X_n)} - \mu_{(\tilde{X}_1,\tilde{X}_2,\ldots,\tilde{X}_n)}\| \leqslant \beta_1 + \cdots + \beta_n. \tag{11}$$

See Berbee (1987) and Schwarz (1980) for the proof. We will apply the lemma to decouple Y_i and Y_j when $|i-j|$ is big enough.

Take any $\gamma \in]0,1[$. Decompose the set $\{1,\ldots,n\}$ into $\ell(n)$ blocks of length $k(n)$ and a block of length less than $k(n)$, where $k(n)$, $\ell(n)$ are integers with $k(n)/n^\gamma \to 1$, $\ell(n)/n^{1-\gamma} \to 1$ for $n \to \infty$.

Proof of (a). Since $Y_1,\dots,Y_n$ are non-negative, we have

$$P[S_n \geqslant n^{\psi} m] = P\left[\sum_{i=1}^{n} Y_i \geqslant n^{\psi} m\right] \geqslant P\left[\sum_{i=1}^{\ell(n)} Y_{(i-1)k(n)+1} \geqslant n^{\psi} m\right]. \tag{12}$$

We apply Lemma 1 to $\{Y_{(i-1)k(n)+1},\ 1 \leqslant i \leqslant \ell(n)\}$ and get

$$P[S_n \geqslant n^{\psi} m] \geqslant P\left[\sum_{i=1}^{\ell(n)} \tilde{Y}_i \geqslant n^{\psi} m\right] - \ell(n)\beta(k(n)), \tag{13}$$

where $\tilde{Y}_1,\dots,\tilde{Y}_{\ell(n)}$ are i.i.d. with the same distribution as Y_1. The second term on the r.h.s. of (13) is negligible: In fact, we have $\log\beta(k(n))/\log n \to -\infty$, since $\log\beta(k(n))/\log k(n) \to -\infty$ and $\log k(n)/\log n \to \gamma$. Therefore, $\log(\ell(n)\beta(k(n)))/\log n \to -\infty$. For the first term on the r.h.s. of (13), we have, using Theorem 2,

$$\liminf_n \frac{1}{\log n} \log P\left[\sum_{i=1}^{\ell(n)} \tilde{Y}_i \geqslant n^{\psi} m\right]$$

$$= \liminf_n \frac{1-\gamma}{\log \ell(n)} \log P\left[\sum_{i=1}^{\ell(n)} \tilde{Y}_i \geqslant \ell(n)^{\psi/(1-\gamma)} m\right] \geqslant (1-\gamma)\left(-\frac{\alpha\psi}{1-\gamma} + 1\right). \tag{14}$$

Since $\gamma \in]0,1[$ is arbitrary, (1) holds.

Proof of (b). Note that

$$P[S_n \geqslant n^{\psi} m] \leqslant P\left[\frac{1}{\ell(n)^{\psi} k(n)^{\psi}} \sum_{i=1}^{n} Y_i \geqslant m\right]$$

$$\leqslant P\left[\frac{1}{k(n)^{\psi}} \sum_{i=1}^{k(n)} \frac{1}{\ell(n)^{\psi}} \sum_{j=1}^{\ell(n)+1} Y_{(j-1)k(n)+i} \geqslant m\right]$$

$$\leqslant k(n)^{\psi} P\left[\frac{1}{\ell(n)^{\psi}} \sum_{j=1}^{\ell(n)+1} Y_{(j-1)k(n)+1} \geqslant m\right]. \tag{15}$$

We apply Lemma 1 to $\{Y_{(j-1)k(n)+1},\ 1 \leqslant j \leqslant \ell(n)+1\}$ here, yielding

$$P[S_n \geqslant n^{\psi} m] \leqslant k(n)^{\psi} P\left[\frac{1}{\ell(n)^{\psi}} \sum_{i=1}^{\ell(n)+1} \tilde{Y}_i \geqslant m\right] + k(n)^{\psi}(\ell(n)+1)\beta(k(n)), \tag{16}$$

where $\tilde{Y}_1,\dots,\tilde{Y}_{\ell(n)+1}$ are i.i.d. with the same distribution as Y_1. The second term in (16) can be shown to be negligible, exactly as the second term in (13). To estimate the first term, we use Theorem 2 and get

$$\limsup_n \frac{1}{\log \ell(n)} \log P\left[\frac{1}{\ell(n)^{\psi}} \sum_{i=1}^{\ell(n)+1} \tilde{Y}_i \geqslant m\right] \leqslant -(\alpha\psi - 1). \tag{17}$$

Since $\log \ell(n)/\log n \to 1-\gamma$ and $\log k(n)/\log n \to \gamma$, (16) and (17) yield

$$\limsup_n \frac{1}{\log n} \log P[S_n \geqslant n^{\psi} m] \leqslant \gamma\psi - (1-\gamma)(\alpha\psi - 1). \tag{18}$$

Since $\gamma \in]0,1[$ is arbitrary, (2) holds. □

Proof of Theorem 5. First one has to prove (a) and (b) for the i.i.d. case, see Gantert (1996) or Rozovskii (1994). Then, to show (a), one uses $P[S_n \geqslant n^{\psi} m] \geqslant P[Y_1 \geqslant n^{\psi} m]$, and one proceeds directly without making use of the assumption on $\beta(n)$. (The mixing condition is needed, however, if we drop the assumption that Y_1 is non-negative.)

Proof of (b). Let $\gamma \in]0,1[$. Decompose the set $\{1,\dots,n\}$ into $\ell(n)$ blocks of length $k(n)$ and a block of length less than $k(n)$, where $k(n)$, $\ell(n)$ are integers with $k(n)/n^{\gamma} \to 1$, $\ell(n)/n^{1-\gamma} \to 1$ for $n \to \infty$. As in (15), we have

$$P[S_n \geqslant n^{\psi} m] \leqslant k(n)^{\psi} P\left[\frac{1}{\ell(n)^{\psi}} \sum_{i=1}^{\ell(n)+1} \tilde{Y}_i \geqslant m\right] + k(n)^{\psi}(\ell(n)+1)\beta(k(n)), \tag{19}$$

where $\tilde{Y}_1,\dots,\tilde{Y}_{\ell(n)+1}$ are i.i.d. with the same distribution as Y. We first show that the second term on the r.h.s. of (19) is negligible: we have

$$\frac{1}{L(n^{\psi})}\log \beta(k(n)) = \frac{\log \beta(k(n))}{L(k(n))}\frac{L(k(n))}{L(n^{\psi})} \to -\infty \tag{20}$$

due to our assumptions on β and L. Further, $\log \ell(n)/L(n^{\psi}) \to 0$ and $\log k(n)/L(n^{\psi}) \to 0$, hence

$$\frac{1}{L(n^{\psi})}\log\left(k(n)^{\psi}(\ell(n)+1)\beta(k(n))\right) \to -\infty. \tag{21}$$

We turn to the first term on the r.h.s. of (19). The tail asymptotics for the i.i.d. case imply that

$$\limsup_n \frac{1}{L(\ell(n)^{\psi})}\log P\left[\frac{1}{\ell(n)^{\psi}} \sum_{i=1}^{\ell(n)+1} \tilde{Y}_i \geqslant m\right] \leqslant -1. \tag{22}$$

Our assumptions on L imply $\lim_n L(\ell(n)^{\psi})/L(n^{\psi}) = h(1-\gamma) > 0$. Therefore, (19) and (22) yield

$$\limsup_n \frac{1}{L(n^{\psi})}\log P[S_n \geqslant n^{\psi} m] \leqslant -h(1-\gamma). \tag{23}$$

It is easy to see that due to our assumptions on L, $h(1-\gamma) \to 1$ for $\gamma \to 0$. The claim follows by letting $\gamma \to 0$ in (23). □

3. A counterexample

We give an example of a Markov chain where $\beta(n) \leqslant n^{-c}$ for some constant c, but the upper bound (b) in Theorem 2 does not hold. As a consequence, the equivalence of (i) and (ii) in Theorem 1 breaks down. This shows that if one uses β-mixing coefficients, a logarithmic mixing rate as in Peligrad (1985) is not sufficient.

Example 1. Let $\alpha > 1$ and $a(x) := \exp(-(1/x^{\delta}))$ where $\delta > 1/(\alpha-1)$. For two probability distributions ν and μ specified below, let $Y_1, Y_2, \dots$ be the Markov chain with kernel $\pi(x,\cdot) = a(x)\delta_x(\cdot) + (1-a(x))\nu(\cdot)$ and starting distribution μ. We assume that $P[Y_1 \geqslant t] = 1/t^{\alpha}$ for $t \geqslant 1$, i.e. μ has density $f(t) = \alpha/t^{\alpha+1}$ for $t \geqslant 1$, and ν is chosen such that μ is the invariant distribution for the Markov chain. More precisely, ν is given by $\mathrm{d}\nu/\mathrm{d}\mu(x) = (1-a(x))(\int(1-a(x))\mu(\mathrm{d}x))^{-1}$.

Claim. $\beta(n) \leqslant n^{-\alpha/\delta}$const. and, for $m > m_0 = E[Y_1]$,

$$\liminf_n \frac{1}{\log n}\log P[S_n \geqslant nm] > -(\alpha-1). \tag{24}$$

Proof. We see that either $(Y_1,\dots Y_k)$ and $(Y_{k+n},Y_{k+n+1},\dots)$ are independent or we have $Y_k=Y_{k+1}=\dots=Y_{k+n}$, and the second case occurs, given Y_k, with probability $a(Y_k)^n$. This implies

$$\beta(n)\leqslant\int a(x)^n\mu(\mathrm{d}x)=\int_1^\infty \exp\left(-\frac{n}{x^\delta}\right)\frac{\alpha}{x^{\alpha+1}}\,\mathrm{d}x\leqslant n^{-\alpha/\delta}\mathrm{const.} \tag{25}$$

Let $k(n)$, $n=1,2,\dots$, be a sequence of integers with $k(n)/n^\gamma\to 1$, where $\gamma=\delta/(1+\delta)$. We then have

$$P[S_n\geqslant nm]\geqslant P\left[Y_1\geqslant\frac{nm}{k(n)},Y_{k(n)}=Y_{k(n)-1}=\cdots=Y_1\right]=P\left[Y_1\geqslant\frac{nm}{k(n)}\right]a\left(\frac{nm}{k(n)}\right)^{k(n)}.$$

Taking logarithms gives

$$\log P[S_n\geqslant nm]\geqslant\log\left(\frac{k(n)^\alpha}{n^\alpha m^\alpha}\right)+k(n)\left(\frac{-k(n)^\delta}{n^\delta m^\delta}\right). \tag{26}$$

Since $k(n)^{1+\delta}/n^\delta\to 1$, dividing by $\log n$ in (26) and letting $n\to\infty$ yields

$$\liminf_n\frac{1}{\log n}\log P[S_n\geqslant nm]\geqslant-\alpha+\frac{\delta}{1+\delta}\alpha>-\alpha+1, \tag{27}$$

where the last inequality holds because $\delta>1/(\alpha-1)$. □

Note that this Markov chain does not satisfy Doeblin's condition.

Acknowledgement

I would like to thank Dirk Tasche for several discussions.

References

Baum, L.E., Katz, M., 1965. Convergence rates in the law of large numbers. Trans. Amer. Math. Soc. 120, 108–123.
Berbee, H., 1987. Convergence rates in the strong law for a bounded mixing sequence. Probab. Theory Related Fields 74, 253–270.
Bradley, R.C., 1986. Basic properties of strong mixing conditions. In: Eberlein, E., Taqqu, M.S. (Eds.), Birkhäuser, Dependence in Probability and Statistics, Basel, Progress in Probability and Statistics.
Daley, D.J., Foley, R.D., Rolski, T., 1996. A note on convergence rates in the strong law for strong mixing sequences. Probab. Math. Statist. 16 (1), 19–28.
Davis, R.A., Hsing, T., 1995. Point processes and partial sum convergence for weakly dependent random variables with infinite variance. Ann. Probab. 23 (2), 879–917.
Gantert, N., 1996. Large deviations for a heavy-tailed mixing sequence. Preprint.
Kiesel, R., 1997. Strong laws and summability for sequences of φ-mixing random variables in Banach spaces. Electron. commun. Probab. 2, 27–41.
Peligrad, M., 1985. Convergence rates for the strong law for stationary mixing sequences. Z. Wahrsch. Verw. Gebiete 70, 307–314.
Schwarz, G., 1980. Finitely determined processes – an indiscrete approach. J. Math. Anal. Appl. 76, 146–158.
Rozovskii, L.V., 1994. Probabilities of large deviations on the whole axis. Theory Probab. Appl. 38, 53–79.
Vinogradov, V., 1994. Refined Large Deviation Limit Theorems. Pitman Research Notes in Mathematics Series 315.