应用概率统计　第 37 卷
第 5 期　2021 年 10 月

Chinese Journal of Applied Probability and Statistics
Oct., 2021, Vol. 37, No. 5, pp. 515-522
doi: 10.3969/j.issn.1001-4268.2021.05.007

Deviations for Weak Record Numbers in Simple Random Walks *

LI Yuqiang　　YAO Qiang⋆

(*Key Laboratory of Advanced Theory and Application in Statistics and Data Science-MOE, School of Statistics, East China Normal University, Shanghai, 200062, China*)

Abstract: Record numbers are important statistics in random walk models. Their deviation principles are unknown as far as we know. In this article, we provide the asymptotic probabilities of large and moderate deviations for the number of weak records in one-dimensional symmetric simple random walks.

Keywords: random walks; weak record numbers; large deviations principle; moderate deviations principle

2020 Mathematics Subject Classification: 60F10; 60G50

Citation: LI Y Q, YAO Q. Deviations for weak record numbers in simple random walks [J]. Chinese J Appl Probab Statist, 2021, 37(5): 515–522.

§1. Introduction

"Record", according to the Oxford dictionary, can be referred as an extreme attainment, that is, the best (or worst) performance ever attested in a particular activity. The study of record statistics has become an integral part of diverse fields such as meteorology, hydrology, economics, sports, etc. In mathematics, record statistics in the setting of i.i.d. random variables are well understood in many situations. For example, suppose a family of i.i.d. random variables $\{X_n, n \geqslant 0\}$ is a stochastic model for achievements in a sequence of activities. Let $M_n = \max\limits_{0\leqslant i\leqslant n} X_i$, then M_n is the record at time n and the statistic

$$K_n = \sum_{i=1}^{n} \mathbf{1}_{\{X_i = M_n\}}$$

*The project was supported by the National Natural Science Foundation of China (Grant No. 11671145), the Natural Science Foundation of Shanghai (Grant No. 16ZR1409700) and the Program of China Scholarships Council (Grant No. 201806145024). The project was supported in part by the Science and Technology Commission of Shanghai Municipality (Grant No. 18dz2271000).

⋆Corresponding author, E-mail: qyao@sfs.ecnu.edu.cn.

Received October 15, 2020. Revised November 19, 2020.

counts the number of the current record. Brands et al. [1] and the references therein studied the asymptotic behaviors of K_n as n tends to infinity, and Khmaladze et al. [2] discussed the number of the so-called ε-repetitions of the current record value.

In real applications, it is more reasonable that the series of X_n are correlated. In this case, we say that a record event happens at time k, if X_k is larger than all previous values in the series. For example, Majumdar and Ziff [3] used random walks to model the time series of achievements in some particular activities and discussed the growth of record numbers and surviving ages. Now let us recall some important definitions. Let $\{X_k,\, k \geqslant 1\}$ be a sequence of i.i.d. random variables. $S = \{S_n,\, n \geqslant 0\}$ denotes the general one-dimensional random walk on $\mathbb{R}$, that is, $S_0 = 0$, and $S_n = \sum\limits_{k=1}^{n} X_k$ for $n \geqslant 1$, where X_k's are i.i.d. Define

$$M_n = \max_{0 \leqslant m \leqslant n} S_m, \qquad n \geqslant 1. \tag{1}$$

Let $T_0 = 0$,

$$T_n = \inf\{m > T_{n-1},\, S_m \geqslant M_{m-1}\} \ \text{ for } \ n \geqslant 1, \tag{2}$$

and define

$$A_n = \sup\{k \geqslant 1,\, T_k \leqslant n\} \tag{3}$$

for each $n \geqslant 1$, where $\inf \emptyset := +\infty$ and $\sup \emptyset := 0$ by convention. Obviously, S_{T_k} is the maximum value among $S_0, S_1, \cdots, S_{T_k}$ and $\{A_n,\, n \geqslant 1\}$ is a counting process which records the number that S arrives at or exceeds its previous maximum value. In this paper, we call A_n the weak record numbers up to time n. Here, we use "weak" to emphasize that we not only consider the time when a new record appears, but also keep eyes on the time when the current record is repeated. We remark that our weak record numbers up to time n is also different from the "record numbers" studied in [4–6], and the references therein, where they discussed the number of the events $\{S_k = M_n\}$ (rather than $\{S_k = M_k\}$) that occur up to time n.

In the field of random walks, A_n is also called the number of "weak ladder points" which is a footstone in the fluctuation theory of random walks. The fluctuation theory was set forth by Spitzer [7] and Feller [8], and has drawn much attention since then because of its wide applications and elaborated but fascinating theory. For more details, one can refer to [9; Chapter 17]. Greenwood et al. [10] proved that a normed version of the bivariate ladder process (T_n, S_{T_n}) converges in law to the bivariate ladder process of a Lévy process X whenever the normed (S_n) converges in law to X. As an immediate corollary, one can derive that a normed version of A_n (number of ladder points) of S converges in distribution to the local time at the supremum of X. Chaumont and Doney[11] extended this result to a more general case, where they proved that when a normed sequence of random walks (S_n) converges almost surely on the Skorokhod space to a Lévy process X, then a normed version of the counting processes of ladder points of (S_n) converges uniformly on compact sets in probability to the local time at the supremum of X. Based on these results, one may further ask how about the deviations between the normed version of A_n and its limit. As far as we know, there is little literature to investigate such problems.

In this short paper, we aim to discuss the deviations between the normed version of A_n and its limit. To simplify calculations, we suppose that $\{S_n\}$ is a one-dimensional simple symmetric random walk on $\mathbb{Z}$, that is, we assume that

$$\mathsf{P}(X_k = 1) = \mathsf{P}(X_k = -1) = 1/2.$$

In this case, we can show that as n tends to infinity,

$$A_n/\sqrt{n} \to 2 \max_{0\leqslant t\leqslant 1} B(t) =: 2B^*(1)$$

in distribution, where $\{B(t)\}$ is a standard Brownian motion. We study the asymptotic probabilities of $\mathsf{P}(A_n \geqslant \sqrt{n}c_n)$, where c_n tends to ∞ besides other constraints. We will get the large deviations principle (LDP), moderate deviations principle (MDP) for A_n, respectively. The main results and their proofs are provided in the next section.

§2. Main Results

Let $S = \{S_n\}_{n\geqslant 0}$ be a simple symmetric random walk on $\mathbb{Z}$, and A_n defined by (3) is the corresponding weak record number up to time n. For any $\lambda \leqslant 0$, define

$$M(\lambda) := \frac{1 + \mathrm{e}^{\lambda} - \sqrt{1 - \mathrm{e}^{2\lambda}}}{2},$$

and

$$\Lambda(\lambda) := \ln M(\lambda) = \ln(1 + \mathrm{e}^{\lambda} - \sqrt{1 - \mathrm{e}^{2\lambda}}\,) - \ln 2.$$

Then

$$G(\lambda) := \Lambda'(\lambda) = 1 - \frac{1}{2(1 + \mathrm{e}^{\lambda})} + \frac{1}{2\sqrt{1 - \mathrm{e}^{2\lambda}}}$$

is a continuous monotone function with $G(0) = +\infty$ and $G(-\infty) = 1$. Therefore, for any $x \in (1, +\infty)$, there exists a unique $\lambda_0 \in [-\infty, 0)$ such that $G(\lambda_0) = x$. Denote this unique

λ_0 by $G^{-1}(x)$. By direct computation, we get that for any $x \geqslant 0$,

$$\Lambda^*(x) = \sup_{\lambda \leqslant 0}\{x\lambda - \Lambda(\lambda)\} = \begin{cases} xG^{-1}(x) - \Lambda(G^{-1}(x)), & \text{if } x > 1; \\ \ln 2, & \text{if } x = 1; \\ +\infty, & \text{if } x < 1. \end{cases}$$

Let $\overline{S}_0 = 0$ and $\overline{S}_n := M_n - S_n$ for $n \geqslant 1$. Then $\{\overline{S}_n, n \geqslant 0\}$ is a nonnegative homogeneous Markov chain with the transition probability matrix $(p_{i,j})_{i,j\geqslant 0}$, where

$$p_{0,1} = p_{0,0} = 1/2 \quad \text{and} \quad p_{k,k+1} = p_{k,k-1} = 1/2, \ \ k \geqslant 1.$$

Let $L_n^0(\overline{S})$ be the occupation time of $\overline{S}$ at site 0 from time 1 and up to time n, that is,

$$L_0^0(\overline{S}) = 0, \quad \text{and} \quad L_n^0(\overline{S}) := \sum_{k=1}^{n} \mathbf{1}_{\{\overline{S}_k = 0\}} \ \text{for} \ n \geqslant 1.$$

It is easy to see that

$$A_n = L_n^0(\overline{S}). \tag{4}$$

Let $\overline{\tau}_1 := \inf\{n > 0, \overline{S}_n = 0\}$ and $\overline{\tau}_{k+1} := \inf\{n > \overline{\tau}_k, \overline{S}_n = 0\}$ for $k \geqslant 1$. (4) suggests that

$$A_n = \sup\{k \geqslant 1, \overline{\tau}_k \leqslant n\}. \tag{5}$$

The Markov property indicates that $\overline{\tau}_1$ and $\overline{\tau}_{k+1} - \overline{\tau}_k$, $k \geqslant 1$ are i.i.d.

We have the following large deviations principle (LDP) for A_n.

Theorem 1 For any $x > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathsf{P}(A_n \geqslant xn) = -x\Lambda^*(1/x). \tag{6}$$

Proof Let $\{Y_i, i \geqslant 1\}$ be a sequence of i.i.d. random variables with the same distribution as $(\overline{\tau}_1 \,|\, \overline{S}_0 = 0)$. Then we have

$$L_n^0(\overline{S}) = \sup\Big\{k \geqslant 0, \sum_{i=1}^{k} Y_i \leqslant n\Big\}.$$

Therefore, for any $0 < x \leqslant 1$,

$$\mathsf{P}(L_n^0(\overline{S}) \geqslant \lceil xn \rceil) \leqslant \mathsf{P}(A_n \geqslant xn) = \mathsf{P}(L_n^0(\overline{S}) \geqslant xn) \leqslant \mathsf{P}(L_n^0(\overline{S}) \geqslant \lfloor xn \rfloor),$$

which implies that

$$\mathsf{P}\Big(\sum_{i=1}^{\lceil xn \rceil} Y_i \leqslant n\Big) \leqslant \mathsf{P}(A_n \geqslant xn) \leqslant \mathsf{P}\Big(\sum_{i=1}^{\lfloor xn \rfloor} Y_i \leqslant n\Big),$$

where $\lceil a\rceil$ and $\lfloor a\rfloor$ denote the minimal integer no smaller than a and the maximal integer no larger than a, respectively.

Next, let

$$f_0(\lambda)=\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_0=0)\quad\text{and}\quad f_1(\lambda)=\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_0=1)$$

for $\lambda\leqslant 0$. From the basic "first-step decomposition" of Markov chains, we know that

$$f_0(\lambda)=\frac{1}{2}\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=0)+\frac{1}{2}\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=1)$$

and

$$f_1(\lambda)=\frac{1}{2}\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=0)+\frac{1}{2}\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=2).$$

It is obvious that $\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=0)=\mathrm{e}^{\lambda}$. From the time homogeneity, we have

$$\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=1)=\mathrm{e}^{\lambda}\mathsf{E}(\mathrm{e}^{\lambda(\overline{\tau}_1-1)}\,|\,\overline{S}_1=1)=\mathrm{e}^{\lambda}f_1(\lambda).$$

Furthermore, the strong Markov property together with the time and space homogeneity imply

$$\mathsf{E}(\mathrm{e}^{\lambda\overline{\tau}_1}\,|\,\overline{S}_1=2)=\mathrm{e}^{\lambda}\mathsf{E}(\mathrm{e}^{\lambda(\overline{\tau}_1-1)}\,|\,\overline{S}_1=2)=\mathrm{e}^{\lambda}f_1^2(\lambda).$$

Therefore, we have that $f_0(\lambda)$ and $f_1(\lambda)$ are the minimal nonnegative solutions of the following equations:

$$f_0(\lambda)=\frac{1}{2}\mathrm{e}^{\lambda}+\frac{1}{2}\mathrm{e}^{\lambda}f_1(\lambda),\qquad f_1(\lambda)=\frac{1}{2}\mathrm{e}^{\lambda}+\frac{1}{2}\mathrm{e}^{\lambda}f_1^2(\lambda),$$

from which we get that

$$f_0(\lambda)=M(\lambda),$$

which implies that $\mathsf{E}(\overline{\tau}\,|\,\overline{S}_0=0)=+\infty$.

From the assumptions that $\overline{S}_0=0$ and $Y\overset{\mathrm{d}}{=}\overline{\tau}_1$, we have that $\mathsf{E}(\mathrm{e}^{\lambda Y_i})=M(\lambda)$ for $\lambda<0$, and that $\mathsf{E}(Y_k)=+\infty$. Applying the Cramér's theorem (see [12; P. 27]), we obtain that

$$\lim_{n\to\infty}\frac{1}{n}\mathsf{P}\Big(\sum_{i=1}^{n}Y_i\leqslant xn\Big)=-\Lambda^*(x).$$

The remainder is same as the proof of Theorem 2 in [13]. So we omit the details. □

To get the moderate deviations principle (MDP) for A_n, we need an auxiliary result which is a special presentation of [14; Theorem 2] in our case.

Lemma 2 Suppose that there is a non-decreasing positive function $a(t)$ on $[1,+\infty)$ such that $a(t)\uparrow+\infty$, then

$$\lim_{n\to\infty}\frac{1}{a(n)}\sum_{k=1}^{n}\mathsf{P}(\overline{S}_n=0\,|\,\overline{S}_0=0)=1,$$

and there exists $p\in[0,1)$ such that

$$\lim_{\lambda\to\infty}a(\lambda t)/a(t)=t^p,\qquad\forall\,t>0.$$

Let $\{b_n\}$ be a positive sequence satisfying $b_n\to\infty$ and $b_n/n\to 0$ as $n\to\infty$. Then

$$\lim_{n\to\infty}\frac{1}{b_n}\ln\mathsf{P}\Big[\sum_{k=1}^{n}\mathbf{1}_{\{\overline{S}_k=0\}}>\lambda a\Big(\frac{n}{b_n}\Big)b_n\Big]=-(1-p)\Big[\frac{p^p\lambda}{\Gamma(p+1)}\Big]^{(1-p)^{-1}}.$$

The MDP for A_n reads as follows.

Theorem 3 If $\{c_n\}$ is a sequence of positive numbers such that $c_n\to\infty$ and $c_n=o(n^{1/2})$ as $n\to\infty$, then for any $x>0$,

$$\lim_{n\to\infty}\frac{1}{c_n^2}\ln\mathsf{P}(A_n\geqslant x\sqrt{n}c_n)=-x^2/8. \tag{7}$$

Proof By the first entrance decomposition of Markov chains,

$$\sum_{n=0}^{\infty}\mathsf{P}(\overline{S}_n=0\,|\,\overline{S}_0=0)s^n=\frac{1}{1-f_0(\ln s)}=\frac{2}{1-s+\sqrt{1-s^2}}$$

for any $s\in[0,1)$. Therefore,

$$\sum_{n=0}^{\infty}\mathsf{P}(\overline{S}_n=0\,|\,\overline{S}_0=0)s^n\sim\sqrt{2}(1-s)^{-1/2}$$

as s tends to $1-$. By Tauberian's theorem [8], we know that

$$\lim_{n\to\infty}\frac{\sum\limits_{k=0}^{n}\mathsf{P}(\overline{S}_k=0\,|\,\overline{S}_0=0)}{n^{1/2}}=\frac{\sqrt{2}}{\Gamma(3/2)}. \tag{8}$$

Note that $A_n=\sum\limits_{k=1}^{n}\mathbf{1}_{\{\overline{S}_k=0\}}$. From (8), we know that Lemma 2 is fulfilled for $p=1/2$, $a(t)=\sqrt{2t}/\Gamma(3/2)$ and $b_n=c_n^2$. Consequently,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{c_n^2}\ln\mathsf{P}(A_n\geqslant x\sqrt{n}c_n)&=\lim_{n\to\infty}\frac{1}{c_n^2}\ln\mathsf{P}(A_n\geqslant x\sqrt{n/c_n^2}c_n^2)\\&=\lim_{n\to\infty}\frac{1}{c_n^2}\ln\mathsf{P}\Big[A_n\geqslant\frac{x\Gamma(3/2)}{\sqrt{2}}a(n/c_n^2)c_n^2\Big]\\&=-(1-p)\Big[\frac{p^px\Gamma(3/2)}{\Gamma(p+1)\sqrt{2}}\Big]^{(1-p)^{-1}}\end{aligned}$$

$$= -\frac{x^2}{8},$$

where we use the fact $p = 1/2$ in the last equality. □

From the MDP, we can get the following law of the iterated logarithm (LIL).

Theorem 4

$$\limsup_{n\to\infty} \frac{A_n}{\sqrt{n \ln\ln n}} = \frac{\pi}{2\sqrt{2}}.$$

Since Theorem 4 is a straightforward application of Theorem 3 in [14] to our case, we omit the details of its proof.

Remark 5 Let $Y_k = T_k - T_{k-1}$ for $k \geqslant 1$. The strong Markov property of random walks implies that Y_k's are i.i.d., and

$$A_n = \sup\Big\{k, \sum_{i=1}^{k} Y_i \leqslant n\Big\}.$$

Namely, $\{A_n\}_{n\geqslant 1}$ is a discrete time renewal process with the inter-occurrence time sequence $\{Y_n\}$. There are many results on the theory of deviations for renewal processes or renewal reward processes. See, for example, [15–21] and the references therein. However, most of the above results need constraints on moments or moment generating functions for inter-occurrence times, which are not fulfilled by A_n in the symmetric simple random walk case.

References

[1] BRANDS J J A M, STEUTEL F W, WILMS R J G. On the number of maxima in a discrete sample [J]. *Statist Probab Lett*, 1994, **20(3)**: 209–217.

[2] KHMALADZE E, NADAREISHVILI M, NIKABADZE A. Asymptotic behaviour of a number of repeated records [J]. *Statist Probab Lett*, 1997, **35(1)**: 49–58.

[3] MAJUMDAR S N, ZIFF R M. Universal record statistics of random walks and Lévy flights [J]. *Phys Rev Lett*, 2008, **101(5)**: 050601 (4 pages).

[4] KATZENBEISSER W, PANNY W. On the number of times where a simple random walk reaches its maximum [J]. *J Appl Probab*, 1992, **29(2)**: 305–312.

[5] KIRSCHENHOFER P, PRODINGER H. Return statistics of simple random walks [J]. *J Statist Plann Inference*, 1996, **54(1)**: 67–74.

[6] PĂLTĂNEA E. On the asymptotic behaviour of the number of maximum points of a simple random walk [J]. *Carpathian J Math*, 2007, **23(1-2)**: 156–164.

[7] SPITZER F. *Principles of Random Walk* [M]. New York: Springer, 1964.

[8] FELLER W. *An Introduction to Probability Theory and Its Applications, Vol. II* [M]. 2nd ed. New York: Wiley, 1971.

[9] KARLIN S, TAYLOR H M. *A Second Course in Stochastic Processes* [M]. New York: Academic Press, 1981.

[10] GREENWOOD P, OMEY E, TEUGELS J L. Harmonic renewal measures and bivariate domains of attraction in fluctuation theory [J]. *Z Wahrsch Verw Gebiete* 1982, **61(4)**: 527–539.

[11] CHAUMONT L, DONEY R A. Invariance principles for local times at the maximum of random walks and Lévy processes [J]. *Ann Probab*, 2010, **38(4)**: 1368–1389.

[12] DEMBO A, ZEITOUNI O. *Large Deviations Techniques and Applications* [M]. 2nd ed. New York: Springer-Verlag, 1998.

[13] GANTERT N, ZEITOUNI O. Large and moderate deviations for the local time of a recurrent Markov chain on $\mathbb{Z}^2$ [J]. *Annales de l'I H P (B)*, 1998, **34(5)**: 687–704.

[14] CHEN X. Moderate deviations for Markovian occupation times [J]. *Stochastic Process Appl*, 2001, **94(1)**: 51–70.

[15] JIANG T F. Large deviations for renewal processes [J]. *Stochastic Process Appl*, 1994, **50(1)**: 57–71.

[16] CHI Z Y. Uniform convergence of exact large deviations for renewal reward processes [J]. *Ann Appl Probab*, 2007, **17(3)**: 1019–1048.

[17] LEFEVERE R, MARIANI M, ZAMBOTTI L. Large deviations for renewal processes [J]. *Stochastic Process Appl*, 2011, **121(10)**: 2243–2271.

[18] BOROVKOV A A, MOGULSKII A A. Large deviation principles for trajectories of compound renewal processes. I [J]. *Theory Probab Appl*, 2016, **60(2)**: 207–224.

[19] BOROVKOV A A, MOGULSKII A A. Large deviation principles for trajectories of compound renewal processes. II [J]. *Theory Probab Appl*, 2016, **60(3)**: 349–366.

[20] TSIRELSON B. From uniform renewal theorem to uniform large and moderate deviations for renewal-reward processes [J]. *Electron Commun Probab*, 2013, **18(52)**: 1–13.

[21] LOGACHOV A V, MOGULSKII A A. Anscombe-type theorem and moderate deviations for trajectories of a compound renewal process [J]. *J Math Sci* (*NY*), 2018, **229(1)**: 36–50.

简单随机游动弱记录数的偏差

李育强　　姚　强

(华东师范大学统计学院, 统计与数据科学前沿理论及应用教育部重点实验室, 上海, 200062)

摘　要: 记录数是随机游动模型中的一个基础统计量. 本文给出了直线上简单对称随机游动弱记录数的大偏差和中偏差.

关键词: 随机游动; 记录数; 大偏差; 中偏差

中图分类号: O211.4