New asymptotics of first-passage times for random walks in the triangular array setting

Alexander Sakhanenko

Joint works with

Vitali Wachtel and Denis Denisov

Applied Probability Workshop 2020

Mathematical Center in Academgorodok,

Novosibirsk State University

August 27, 2020, Online

Denis Denisov

School of Mathematics, University of Manchester,

Oxford Road, Manchester M13 9PL, UK

denis.denisov@manchester.ac.uk

Alexander Sakhanenko[a]

Novosibirsk State University,

Pirogova street, 2, Novosibirsk, 630090, Russia

and Sobolev Institute of Mathematics,

aisakh@mail.ru

Vitali Wachtel[b]

Institut für Mathematik, Universität Augsburg,

86135 Augsburg, Germany

vitali.wachtel@math.uni-augsburg.de

[a)] The work of A.S. is supported by Mathematical Center in Akademgorodok under agreement No. 075-15-2019-1675 with the Ministry of Science and Higher Education of the Russian Federation and it is supported by RFBR (grant No 20-51-12007).

Statement of the Problem — I

Consider a random walk $S_1, S_2, \ldots$ with $S_0 = 0$.

For arbitrary non-random real numbers $g_1, g_2, \ldots$ let

(1) $$T := \min\{n \geq 1 : S_n \leq g_n\}$$

be the time of the first crossing of the moving boundary $\{g_n\}$ by the random walk $\{S_n\}$.

We are going to study the asymptotic behaviour of the distributions of first-passage times over moving (non-constant) boundaries

(2) $$\mathbf{P}(T > n) = \mathbf{P}(\min_{1 \leq k \leq n} (S_k - g_k) > 0) \downarrow$$

for non-classical random walks $\{S_n\}$.

(In what follows, all unspecified limits are taken with respect to $n \to \infty$.)

Natural assumptions — I

Assumptions (A).

(A1) increments $X_k = S_k - S_{k-1}$, $k = 1, 2, \ldots$, are independent random variables;

(A2) they have zero means;

(A3) they satisfy the *classical Lindeberg condition*.

In particular, **(A4)** $B_n^2 := \sum_{i=1}^n \mathbf{E} X_{i,n}^2 \to \infty$.

Assumption (G) $g_n = o(B_n)$, i.e. the boundary $\{g_n\}$ is of a small magnitude;

in particular, **(G+)** $G_n := \max_{k \le n} |g_k| = o(B_n)$.

Main notation:

(3)
$$E_n := \mathbf{E}[S_n - g_n : T_n > n] \ge 0.$$

Note that for all $n \ge 1$

(4)
$$E_n > 0 \quad \text{iff} \quad \mathbf{P}(T > n) > 0.$$

Theorem 1+. *Under Assumptions (A) and (G)*

$$B_n \mathbf{P}(T > n) = \sqrt{2/\pi} E_n (1 + \alpha_n), \qquad (AM)$$

where $\alpha_n \to 0$.

In particular, we have the universal formula **(M)** for asymptotic of $\mathbf{P}(T > n)$:

Theorem 1. (See **[1]**) *Suppose that*

$$\mathbf{P}(T > n) > 0 \quad \textit{for all} \quad n \geq 1. \qquad (T+)$$

Then under Assumptions (A) and (G)

$$\mathbf{P}(T > n) \sim \sqrt{\frac{2}{\pi}} \frac{E_n}{B_n} = \sqrt{\frac{2}{\pi}} \frac{\mathbf{E}[S_n - g_n; T > n]}{B_n}. \qquad (M)$$

[1] Denisov, D., Sakhanenko, A. and Wachtel, V.
First-passage times for random walks with non-identically distributed increments.
Ann. Probab. **46**(6): 3313-3350, 2018.

Theorem 1. (See [1]) *Suppose that*

$$\mathbf{P}(T > n) > 0 \quad \textit{for all} \quad n \geq 1. \tag{$T+$}$$

Then under Assumptions (A) and (G)

$$\mathbf{P}(T > n) \sim \sqrt{\frac{2}{\pi}} \frac{E_n}{B_n} \to 0. \tag{M}$$

Remark 1. *When condition (T+) is not true, then there exists an* $N_0 < \infty$ *such that*

(5) $$E_n = 0 = \mathbf{P}(T > n) \qquad \forall n > N_0.$$

But Theorem 1+ remains valid in all cases for all $n \geq 1$.

Theorem 1+. *Under Assumptions (A) and (G)*

$$B_n \mathbf{P}(T > n) = \sqrt{2/\pi} E_n (1 + \alpha_n) \to 0, \quad \textit{where} \quad \alpha_n \to 0. \tag{AM}$$

Statement of the Problem — II

a) to generalize Theorem 1 for the case of random walks in the triangular array setting;

b) to obtain a rate of convergence for $\alpha_n \to 0$;

c) to obtain an estimate for α_n up to some absolute constants.

Assumptions — II

Assumptions (B).

(A1-) increments $X_1, \ldots, X_n$, are independent for a fixed $n \geq 1$;

(A2-) they have zero means;

(B3) they are *bounded* by a constant $r_n < \infty$, i.e.

$$\mathbf{P}\big(|X_i| \leq r_n\big) = 1 \quad \text{for all} \quad i = 1, \ldots, n; \qquad (B3)$$

(A4-) $B_n^2 := \sum_{i=1}^n \mathbf{E}X_{i,n}^2 > 0.$

Main result

Theorem 2+. *Under Assumptions (B)*

$$\alpha_n \le C\rho_n^{2/3} \qquad \text{if} \qquad \rho_n := \frac{r_n + G_n}{B_n} \le \frac{1}{24}, \qquad (BM)$$

where $C < \infty$ *is an absolute constant and* $G_n := \max_{k\le n} |g_k|$.

Under additional Assumptions (G+) and (T+) a weaker variant of Theorem 2+ have been proved in our joint paper:

[2] Denisov, D., Sakhanenko, A. and Wachtel, V.

First-passage times for random walks in the triangular array setting.

(invited to the volume in Honor of R.A. Doney in the series Progress in Probability)

https://arxiv.org/pdf/2005.00240.pdf

A comment on the main result

Theorem 2+. *Under Assumptions (B)*

$$\mathbf{P}(T > n) = \sqrt{\frac{2}{\pi}}\frac{E_n}{B_n}\big(1 + C\theta_n\rho_n^{2/3}\big) \qquad \textit{if} \quad \rho_n := \frac{r_n + G_n}{B_n} \le \frac{1}{24}, \qquad (\mathbf{BM})$$

where $|\theta_n| \le 1$ *and* $C < \infty$ *is an absolute constant.*

Assumptions (B).

(A1-) increments $X_1, \dots, X_n$, are independent for a fixed $n \ge 1$;

(A2-) they have zero means;

(B3) they are *bounded* by a constant $r_n < \infty$, i.e.

$$\mathbf{P}\big(|X_i| \le r_n\big) = 1 \quad \text{for all} \quad i = 1, \dots, n; \qquad (B3)$$

(A4-) $B_n^2 := \sum_{i=1}^n \mathbf{E}X_{i,n}^2 > 0.$

Natural assumptions in the triangular array setting

Assumptions (A+).

(A1+) For each $n \geq 1$ we are given independent random variables $X_{1,n}, \ldots, X_{n,n}$ such that

(A2+) they have zero means;

(A3+) they satisfy the *classical Lindeberg condition*.

In particular, for some $n_0 < \infty$ **(A4)** $B_n^2 := \sum_{i=1}^n \mathbf{E} X_{i,n}^2 > 0$ for all $n > n_0$.

Assumption (G+) Let $g_{1,n}, \ldots, g_{n,n}$ be deterministic real numbers such that $G_n := \max_{k \leq n} |g_k| = o(B_n)$.

For each $n \geq 1$ we consider a random walk

(6)
$$S_{k,n} := X_{1,n} + \cdots + X_{k,n}, \quad k = 1, 2, \ldots, n;$$

and let

(7)
$$T_n := \inf\{k \geq 1 : S_{k,n} \leq g_{k,n}\},$$

be the first crossing of the moving boundary $\{g_{k,n}\}$ by the random walk $\{S_{k,n}\}$.

Here the main purpose is to study the asymptotic behavior of the probability

$$\mathbf{P}(T_n > n) = \mathbf{P}\left(\min_{1\le k\le n}(S_{k,n} - g_{k,n}) > 0\right).$$

Main notation:

(8) $$E_n := \mathbf{E}[S_{n,n} - g_{n,n} : T_n > n] \ge 0.$$

Note that for all $n \ge 1$

(9) $$E_n > 0 \quad \text{iff} \quad \mathbf{P}(T_n > n) > 0.$$

Assumptions (T+). $\mathbf{P}(T_n > n) > 0$ for all $n > n_1$ with some $n_1 < \infty$.

Natural Hypothesis.

Based on the validity of Theorem 1 from [1] one can expect that the Lindeberg condition will again be sufficient for the following natural generalization of (M):

$$\mathbf{P}(T_n > n) \sim \sqrt{\frac{2}{\pi}} \frac{E_n}{B_n} = \sqrt{\frac{2}{\pi}} \frac{\mathbf{E}[S_{n,n} - g_{n,n}; T_n > n]}{B_n}. \qquad (M+)$$

However the following result shows that this is not the case and the situation is more complicated.

Proposition 1. *There exists a triangular array of independent random variables* $\{X_{k,n}\}$ *such that conditions* (A+), (G+), (T+) *hold with* $B_n^2 = n$, *but*

(10)
$$\mathbf{P}(T_n > n) = o(E_n/B_n).$$

Main result from [2] in the triangular array setting

Assumptions (B+). Conditions (A1+), (A2+), (A4+) hold; in addition

(B3+) for each n, i random variables $X_{n,i}$ are *bounded* by a constant $r_n = o(B_n)$, i.e.

$$r_n = o(B_n) \quad \text{and} \quad \mathbf{P}\big(\max_{1 \le i \le n} |X_{i,n}| \le r_n \big) = 1 \quad \text{for all} \quad n \ge 1. \qquad (B3+)$$

Corollary 1. *Under conditions* (G+), (T+) *and* (B+) *(with* $r_n = o(B_n)$ *) relation* (M+) *takes place.*

Thus, we have obtained the desired asimptotic (M+) but under assumption (B3+) with $r_n = o(B_n)$, which is stronger than the Lindeberg condition (A3+).

A generalization of Corollary 1.

For an arbitrary random variable Y_n let

(11) $$\tau_n := \inf\left\{k \ge 1 : Y_n + S_{k,n} \le g_{k,n}\right\},$$

be the first crossing of the moving boundary $\{g_{k,n}\}$ by the random walk $\{Y_n + S_{k,n}\}$. From Theorem 2 with variable Y_n truncated on the level $h_n = o(B_n)$ we obtain that

(12) $$\mathbf{P}\left(\tau_n > n : |Y_n| \le h_n\right) \sim \sqrt{\frac{2}{\pi}} \frac{E_n(h_n)}{B_n}, \qquad \text{where}$$

$$E_n(h) := \mathbf{E}\left[Y_n + S_{n,n} - g_{n,n} : \tau_n > n,\ |Y_n| \le h\right], \qquad h > 0.$$

On the other hand, from known estimates in functional CLT it is not difficult to find that

(13) $$\mathbf{P}\left(\tau_n > n : Y_n > h_n\right) \sim \mathbf{E}\left[\Psi\left(\frac{Y_n}{B_n}\right) : Y_n > h_n\right] \qquad \text{with}$$

$$\Psi(x) := 2\int_0^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = \mathbf{P}\left(\min_{0 \le t \le 1} W(t) > -x\right), \quad (x > 0),$$

when $r_n + G_n = o(h_n)$, where $W(t)$ is a standard Wiener process.

Summing up (12) and (13) we obtain:

Corollary 2. *Suppose that for each* $n = 1, 2, \ldots$ *random variable* Y_n *is independent of* $X_{1,n}, \ldots, X_{n,n}$. *Then under assumptions of Corollary 1*

$$\mathbf{P}\left(\tau_n > n\right) \sim \sqrt{\frac{2}{\pi}} \frac{E_n(h_n)}{B_n} + \mathbf{E}\left[\Psi\left(\frac{Y_n}{B_n}\right) : Y_n > h_n\right] \tag{14}$$

for all non-random numbers $h_n > 0$ *such that*

$$h_n = o(B_n) \qquad \text{and} \qquad r_n + G_n = o(h_n). \tag{15}$$

In particular, we may everywhere take

$$h_n = \sqrt{(r_n + G_n)B_n}. \tag{16}$$

Simple proof of Corollary 2 (given above) shows the power of Theorem 2.

Proof of Proposition 1.

For a sequence of integers $r_n \geq 1$ let random variables $X_{1,n}$ be defined as follows

(17) $$X_{1,n} := \begin{cases} r_n, & \text{with probability } p_n := \frac{1}{2r_n^2} \leq \frac{1}{2}, \\ 0, & \text{with probability } 1 - 2p_n, \\ -r_n, & \text{with probability } p_n. \end{cases}$$

Let each $X_{1,n}$ is independent of the sequence $X_2, X_3, \ldots$ of i.i.d. random variables with the Rademacher distribution: $\mathbf{P}(X_k = \pm 1) = 1/2$.

So that $\mathbf{E}X_{k,n} = 0$ and $\mathbf{E}X_{k,n}^2 = 1$ for all $k \geq 1$; and the triangular array

(18) $$X_{1,n},\ X_{k,n} := X_k,\ k = 2, 3, \ldots, n;\ n \geq 1$$

satisfies the Lindeberg condition with $B_n^2 = n$.

Put $g_{k,n} \equiv 0$; then conditions (G+) and (T+) also hold.

Lemma 1. $E_n = p_n r_n$.

Proof. For integers $m \geq 2$ and $r_n \geq 1$ define

$$U_m := X_2 + X_3 + \ldots + X_m, \tag{19}$$

$$\nu := \inf\{m \geq 2 : r_n + U_m \leq 0\} = \inf\{m \geq 2 : r_n + U_m = 0\}. \tag{20}$$

Since U_m is a martingale and ν is a stopping time, we obtain from (20) that

$$\begin{aligned} E_{2,n} &:= \mathbf{E}\left[r_n + U_n : \nu > n\right] \\ &= \mathbf{E}\left[r_n + U_n\right] - \mathbf{E}\left[r_n + U_n : \nu \leq n\right] \\ &= \left[r_n + 0\right] - \mathbf{E}\left[r_n + U_\nu : \nu \leq n\right] = r_n - \mathbf{E}[0] = r_n. \end{aligned}$$

Now, with $g_{k,n} \equiv 0$ one has:

$$E_n = \mathbf{E}[S_{n,n}; T_n > n] = \mathbf{P}(X_{1,n} = r_n) E_{2,n}(r_n) = p_n r_n. \tag{21}$$

Proof of Proposition 1. In particular, from (17) and (21) we conclude that

$$\mathbf{P}(T_n > n) < \mathbf{P}(X_{1,n} = r_n) = p_n = \frac{E_n}{r_n} = \frac{E_n}{B_n} \times \frac{B_n}{r_n}. \tag{22}$$

Thus, Proposition 1 is proved since (10) follows from (22) when $r_n/B_n \to \infty$.

Proposition 2. *More precise calculations in [2] show that, for the triangular array from* (17) *and* (18) *with* $g_{k,n} \equiv 0$*, the desired asymptotic (M+) takes place iff* $r_n = o(B_n)$*.*

Theorem 2 and Propositions 1 and 2 have been proved in our joint paper:

[2] Denisov, D., Sakhanenko, A. and Wachtel, V.

First-passage times for random walks in the triangular array setting.

(invited to the volume in Honor of R.A. Doney in the series Progress in Probability)

https://arxiv.org/pdf/2005.00240.pdf

======================================

THANK YOU !