PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 5, Fasc. 1 (1985), p. 83-89

RANDOM LIMIT THEOREMS FOR RANDOM WALKS CONDITIONED TO STAY POSITIVE

BY

A. SZUBARGA AND D. SZYNAL (LUBLIN)

Abstract. Let $\{X_k, k \geqslant 1\}$ be a sequence of independent, identically distributed random variables with $EX_1 = 0$, $EX_1^2 = \sigma^2 < \infty$, and let $\{N_n, n \geqslant 1\}$, $N_0 = 0$ a.s., be a sequence of positive integer-valued random variables. Form the random walk $\{S_{N_n}, n \geqslant 0\}$ by setting $S_0 = 0$ and $S_{N_n} = X_1 + \ldots + X_{N_n}$, $n \geqslant 1$. This paper investigates the limit behaviour of $P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, S_2 > 0, \ldots, S_{N_n} > 0]$.

1. Introduction. Let $\{X_k, k \geqslant 1\}$ be a sequence of independent, identically distributed random variables with $EX_1 = 0$, $0 < EX_1^2 = \sigma^2 < \infty$. Define

$$S_0 = 0, \quad S_n = \sum_{k=1}^{n} X_k.$$

The conditional central limit theorem is given in [2], i.e., it is shown that

$$(1.1) \quad \lim_{n\to\infty} P[S_n < x\sigma\sqrt{n} | S_1 > 0, \ldots, S_n > 0] = 1 - \exp(-x^2/2), \quad x \geqslant 0,$$

provided $E|X_1|^3 < \infty$ and X_1 is nonlattice or integer-valued with span 1. However these extra assumptions are superfluous (see [1]).

2. The conditional random central limit theorem. First we note that common conditions for the random central limit theorem are not, in general, sufficient for the conditioned random central limit theorem.

Let $\{N_n, n \geqslant 1\}$ be a sequence of positive integer-valued random variables with

$$P[N_n = 1] = 1/\sqrt{n}, \quad P[N_n = n] = 1 - 1/\sqrt{n}, \quad n = 1, 2, \ldots,$$

and suppose that $\{X_k, k \geqslant 1\}$ is a sequence of independent identically distributed random variables satisfying the conditions of Section 1 and independent of $\{N_n, n \geqslant 1\}$.

One can see that $N_n/\alpha_n \xrightarrow{P} 1$, $n \to \infty$ (P. — in probability), where $\alpha_n = EN_n$, which implies $N_n \xrightarrow{P} \infty$, $n \to \infty$, and the random limit theorem ([3], p. 472). But, by (1.1) and $P[S_1 > 0, \ldots, S_n > 0] \sim C/\sqrt{n}$ (see [4]), where C is a positive constant, we have

$$P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, \ldots, S_{N_n} > 0]$$
$$= \frac{\sqrt{n}\, P[X_1 < x\sigma,\, X_1 > 0] + (1-1/\sqrt{n})\, P[S_n < x\sigma\sqrt{n},\, S_1 > 0, \ldots, S_n > 0]}{\sqrt{n}\, P[X_1 > 0](1/\sqrt{n}) + (1-1/\sqrt{n})\, P[S_1 > 0, \ldots, S_n > 0]}$$
$$\to \frac{P[X_1 < \sigma,\, X_1 > 0] + C(1-\exp(-x^2/2))}{P[X_1 > 0] + C} \neq 1-\exp(-x^2/2), \quad x \geqslant 0.$$

We prove the following

THEOREM 1. *Let $\{X_k, k \geqslant 1\}$ be a sequence of independent, identically distributed random variables with* $\mathrm{E}X_1 = 0$, $0 < \mathrm{E}X_1^2 = \sigma^2 < \infty$.

Suppose that $\{N_n, n \geqslant 1\}$, $N_0 = 0$ a.s., is a sequence of positive integer-valued random variables independent of $\{X_k, k \geqslant 1\}$ and $\{\alpha_n, n \geqslant 1\}$ is a sequence of positive numbers with $\lim_{n\to\infty} \alpha_n = \infty$. If

(1.2) $N_n/\alpha_n \xrightarrow{P} a$, $n \to \infty$, *and* $\sqrt{\alpha_n}\, P[N_n = k] \to 0$, $n \to \infty$, $k \geqslant 1$, *where a is a positive constant,*

then

(1.3) $$\lim_{n\to\infty} P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, \ldots, S_{N_n} > 0] = 1-\exp(-x^2/2),\ x \geqslant 0.$$

Proof. Note that

$$P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, \ldots, S_{N_n} > 0]$$
$$= \sum_{k=1}^{\infty} (r_k/t_n)\, P[N_n = k]\, P[S_k < x\sigma\sqrt{k} | S_1 > 0, \ldots, S_k > 0],$$

where $r_k = P[S_1 > 0, \ldots, S_k > 0]$ while $t_n = \mathrm{E}R_n = P[S_1 > 0, \ldots, S_{N_n} > 0]$ and R_n is a random variable taking values r_k with probability $p_k = P[N_n = k]$, $k \geqslant 1$. To complete the proof, it is enough to show that $C_{nk} = (r_k/t_n)\, P[N_n = k]$ is a permanent Toeplitz matrix ([3], p. 472) and to use (1.1).

By the independence of $\{X_k, k \geqslant 1\}$ and $\{N_n, n \geqslant 1\}$ we have

$$t_n = \sum_{k=1}^{\infty} r_k P[N_n = k] = \sum_{A_n^c} r_k P[N_n = k] + \sum_{A_n} r_k P[N_n = k],$$

where, for any given $\varepsilon > 0$, $A_n = \{k\colon (a-\varepsilon)\alpha_n < k < (a+\varepsilon)\alpha_n\}$ and A_n^c is the complement of A_n.

Taking into account that $r_n \sim C/\sqrt{n}$ (see [4]), we get, for sufficiently large n,

(1.4) $$\frac{C}{\sqrt{[(a+\varepsilon)\alpha_n]}} \sum_{A_n} P[N_n = k] \leqslant \sum_{A_n} r_k P[N_n = k]$$

as

$$\sum_{A_n} r_k P[N_n = k] \geqslant \sum_{A_n} P[S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0] P[N_n = k]$$

$$\geqslant \frac{C}{\sqrt{[(a+\varepsilon)\alpha_n]}} \sum_{A_n} P[N_n = k].$$

Therefore, by (1.2) and (1.4), we obtain

$$(1.5) \quad 0 \leqslant \frac{r_k}{t_n} P[N_n = k] = \frac{r_k P[N_n = k]}{\sum_{A_n} r_m P[N_n = m] + \sum_{A_n^c} r_m P[N_n = m]}$$

$$\leqslant \frac{r_k \sqrt{[(a+\varepsilon)\alpha_n]}\, P[N_n = k]}{C \sum_{A_n} P[N_n = m]} \to 0, \ n \to \infty, \qquad k \geqslant 1.$$

It is obvious that $C_{nk} \geqslant 0, n \geqslant 1$, and that $\sum_{k=1}^{\infty} C_{nk} = 1, n \geqslant 1$. These facts, (1.1) and (1.5) imply (1.3).

Remark. From the proof it follows that, under the assumptions of independence of $\{X_k, k \geqslant 1\}$ and $\{N_n, n \geqslant 1\}$, the conditioned random central limit theorem holds when $N_n/\alpha_n \xrightarrow{P} a > 0, \ n \to \infty$, and

$$P[N_n = k]/ER_n \to 0, \ n \to \infty, \qquad \text{for all } k \geqslant 1.$$

The following theorem considers the case when $\{N_n, n \geqslant 1\}$ and $\{X_k, k \geqslant 1\}$ can be dependent.

THEOREM 2. *Let $\{X_k, k \geqslant 1\}$ be a sequence of independent, identically distributed random variables with* $EX_1 = 0$, $EX_1^2 = \sigma^2 < \infty$.

If $\{N, n \geqslant 1\}$, $N_0 = 0$ a.s., is a sequence of positive integer-valued random variables and $\{\alpha_n, n \geqslant 1\}$ is a sequence of positive numbers such that for any given $\varepsilon > 0$

$$(1.6) \qquad P[|N_n/\alpha_n - a| \geqslant \varepsilon] = o(1/\sqrt{\alpha_n}), \ n \to \infty,$$

then (1.3) *holds.*

Proof. With the notation of Theorem 1, we have

$$(1.7) \qquad \lim_{n\to\infty} \sqrt{\alpha_n}\, P[N_n \in A_n^c] = 0,$$

which implies, for sufficiently large n,

$$(1.8) \qquad \frac{C}{\sqrt{[(a+\varepsilon)\alpha_n]}} - o(1/\sqrt{\alpha_n}) \leqslant t_n \leqslant \frac{C}{\sqrt{[(a-\varepsilon)\alpha_n]}} + o(1/\sqrt{\alpha_n}).$$

Hence, by (1.7) and (1.8),

$$(1.9)\qquad P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n^c]/t_n \to 0, n \to \infty.$$

Therefore, to prove (1.3), it is enough to consider

$$P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n]/t_n$$

as

$$(1.10)\qquad P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, \ldots, S_{N_n} > 0] \\ = P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n]/t_n + \\ + P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n^c]/t_n.$$

Put now

$$Z_n = \max_{k \in A_n} \frac{S_k - S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}}, \qquad Z_n^* = \max_{k \in A_n} \frac{S_k - S_{[(a-\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a-\varepsilon)\alpha_n]}},$$

$$u_n = \sqrt{\frac{[(a-\varepsilon)\alpha_n]}{[(a+\varepsilon)\alpha_n]}}, \qquad b_n = [(a+\varepsilon)\alpha_n] - [(a-\varepsilon)\alpha_n].$$

Then we have

$$(1.11)\qquad P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n]$$

$$= \sum_{k \in A_n} P[S_k < x\sigma\sqrt{k}, S_1 > 0, \ldots, S_k > 0, N_n = k]$$

$$\geqslant \sum_{k \in A_n} P[S_k < x\sigma\sqrt{[(a-\varepsilon)\alpha_n]}, S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0, N_n = k]$$

$$\geqslant \sum_{k \in A_n} P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}} + Z_n < xu_n, S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0, N_n = k\right]$$

$$\geqslant P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}} + Z_n < xu_n, S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right] - P[N_n \in A_n^c]$$

$$\geqslant P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}} + Z_n < xu_n, S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0, Z_n < \sqrt[4]{\varepsilon}\right] - \\ - P[N_n \in A_n^c]$$

$$\geqslant P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}} + \sqrt[4]{\varepsilon} < xu_n, S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right] - \\ - P[S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0, Z_n \geqslant \sqrt[4]{\varepsilon}] - P[N_n \in A_n^c]$$

$$\geqslant P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}}+\sqrt[4]{\varepsilon} < xu_n | S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right] r_{[(a+\varepsilon)\alpha_n]} -$$

$$-P[Z_n \geqslant \sqrt[4]{\varepsilon}]\, P[S_1 > 0, \ldots, S_{[(a-\varepsilon)\alpha_n]} > 0] - P[N_n \in A_n^c]$$

as Z_n does not depend on $S_1, \ldots, S_{[(a-\varepsilon)\alpha_n]}$.

The similar evaluations lead us to

$$\text{(1.12)} \quad P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n]$$

$$\leqslant P\left[\frac{S_{[(a-\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a-\varepsilon)\alpha_n]}}-\sqrt[4]{\varepsilon} < x/u_n | S_1 > 0, \ldots, S_{[(a-\varepsilon)\alpha_n]} > 0\right] r_{[(a-\varepsilon)\alpha_n]} +$$

$$+P[Z_n^* \geqslant \sqrt[4]{\varepsilon}]\, P[S_1 > 0, \ldots, S_{[(a-\varepsilon)\alpha_n]} > 0].$$

Note now that, by the Kolmogorov's inequality,

$$\text{(1.13)} \quad P[Z_n \geqslant \sqrt[4]{\varepsilon}] \leqslant P\left[\max_{0\leqslant k\leqslant b_n}\left|\frac{S_k}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}}\right| \geqslant \sqrt[4]{\varepsilon}\right]$$

$$\leqslant \frac{b_n}{\sqrt{\varepsilon}\,[(a+\varepsilon)\alpha_n]} \to \frac{2\sqrt{\varepsilon}}{a+\varepsilon} \quad \text{as } n\to\infty,$$

and

$$\text{(1.14)} \quad P[Z_n^* \geqslant \sqrt[4]{\varepsilon}] \leqslant \frac{b_n}{\sqrt{\varepsilon}\,[(a-\varepsilon)\alpha_n]} \to \frac{2\sqrt{\varepsilon}}{a-\varepsilon} \quad \text{as } n\to\infty.$$

Therefore, by (1.11)-(1.14), we obtain

$$\text{(1.15)} \quad P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}}+\sqrt[4]{\varepsilon} < xu_n | S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right] \frac{r_{[(a+\varepsilon)\alpha_n]}}{t_n} -$$

$$-\frac{b_n}{\sqrt{\varepsilon}\,[(a+\varepsilon)\alpha_n]}(r_{[(a-\varepsilon)\alpha_n]}/t_n) - P[N_n \in A_n^c]/t_n$$

$$\leqslant P[S_{N_n} < x\sigma\sqrt{N_n}, S_1 > 0, \ldots, S_{N_n} > 0, N_n \in A_n]/t_n$$

$$\leqslant P\left[\frac{S_{[(a-\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a-\varepsilon)\alpha_n]}}-\sqrt[4]{\varepsilon} < x/u_n | S_1 > 0, \ldots, S_{[(a-\varepsilon)\alpha_n]} > 0\right] \frac{r_{[(a-\varepsilon)\alpha_n]}}{t_n} +$$

$$+\frac{b_n}{\sqrt{\varepsilon}\,[(a-\varepsilon)\alpha_n]}(r_{[(a-\varepsilon)\alpha_n]}/t_n).$$

Hence, by (1.10) and (1.15), we get

$$(1.16)\qquad -P[N_n \in A_n^c]/t_n - \frac{b_n}{\sqrt{\varepsilon}\,[(a+\varepsilon)\alpha_n]}(r_{[(a-\varepsilon)\alpha_n]}/t_n) +$$
$$+P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}}+\sqrt[4]{\varepsilon} < xu_n \mid S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right]\frac{r_{[(a+\varepsilon)\alpha_n]}}{t_n}$$
$$\leqslant P[S_{N_n} < x\sigma\sqrt{N_n} \mid S_1 > 0, \ldots, S_{N_n} > 0]$$
$$\leqslant P\left[\frac{S_{[(a-\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a-\varepsilon)\alpha_n]}}-\sqrt[4]{\varepsilon} < x/u_n \mid S_1 > 0, \ldots, S_{[(a-\varepsilon)\alpha_n]} > 0\right]\frac{r_{[(a-\varepsilon)\alpha_n]}}{t_n}+$$
$$+\frac{b_n}{\sqrt{\varepsilon}\,[(a-\varepsilon)\alpha_n]}(r_{[(a-\varepsilon)\alpha_n]}/t_n)+P[N_n \in A_n^c]/t_n.$$

But, by (1.1), we have

$$(1.17)\qquad P\left[\frac{S_{[(a+\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a+\varepsilon)\alpha_n]}}+\sqrt[4]{\varepsilon} < xu_n \mid S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right]$$
$$\to 1-\exp\left(-\frac{\left(x\sqrt{\frac{a-\varepsilon}{a+\varepsilon}}-\sqrt[4]{\varepsilon}\right)^2}{2}\right) \quad \text{as } n\to\infty,$$

and

$$(1.18)\qquad P\left[\frac{S_{[(a-\varepsilon)\alpha_n]}}{\sigma\sqrt{[(a-\varepsilon)\alpha_n]}}-\sqrt[4]{\varepsilon} < x/u_n \mid S_1 > 0, \ldots, S_{[(a+\varepsilon)\alpha_n]} > 0\right]$$
$$\to 1-\exp\left(-\frac{\left(x\sqrt{\frac{a+\varepsilon}{a-\varepsilon}}+\sqrt[4]{\varepsilon}\right)^2}{2}\right) \quad \text{as } n\to\infty.$$

Moreover, taking into account that $r_n \sim c/\sqrt{n}$ [4], and (1.8), we get

$$(1.19)\qquad \liminf_{n\to\infty}(r_{[(a+\varepsilon)\alpha_n]}/t_n) \geqslant \sqrt{\frac{a-\varepsilon}{a+\varepsilon}}$$

and

$$(1.20)\qquad \limsup_{n\to\infty}(r_{[(a-\varepsilon)\alpha_n]}/t_n) \leqslant \sqrt{\frac{a+\varepsilon}{a-\varepsilon}}.$$

Therefore, for any given $\varepsilon > 0$, by (1.9), and (1.16)-(1.20), we have

$$-\frac{2\sqrt{\varepsilon}}{a+\varepsilon}\sqrt{\frac{a+\varepsilon}{a-\varepsilon}}+\sqrt{\frac{a-\varepsilon}{a+\varepsilon}}\left(1-\exp\left(-\frac{\left(x\sqrt{\frac{a-\varepsilon}{a+\varepsilon}}-\sqrt[4]{\varepsilon}\right)^2}{2}\right)\right)$$

$$\leqslant \liminf_{n\to\infty} P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, \ldots, S_{N_n} > 0]$$

$$\leqslant \limsup_{n\to\infty} P[S_{N_n} < x\sigma\sqrt{N_n} | S_1 > 0, \ldots, S_{N_n} > 0]$$

$$\leqslant \sqrt{\frac{a+\varepsilon}{a-\varepsilon}}\left(1-\exp\left(-\frac{\left(x\sqrt{\frac{a+\varepsilon}{a-\varepsilon}}+\sqrt[4]{\varepsilon}\right)^2}{2}\right)\right)+\frac{2\sqrt{\varepsilon}}{a-\varepsilon}\sqrt{\frac{a+\varepsilon}{a-\varepsilon}}.$$

Letting now $\varepsilon \to 0$, we obtain (1.3), which completes the proof of Theorem 2.

COROLLARY 2. *If* $\alpha_n = \mathrm{E}N_n$ *and* $\sigma^2(N_n)/\alpha_n^{3/2} \to 0$, $n \to \infty$, *then* (1.3) *holds.*

Remark. One can note that the sequence $\{N_n, n \geqslant 0\}$ of random indices of [3] satisfies (1.6).

REFERENCES

[1] F. Bolthausen, *On a functional central limit theorem for random walks conditioned to stay positive*, Ann. Probability 4 (1976), p. 480-485

[2] D. L. Iglehart, *Functional central limit theorem for random walks conditioned to stay positive*, Ann. Probability 2 (1974), p. 608-619.

[3] A. Renyi, *Probability theory*, North Holland, Amsterdam, London 1970.

[4] F. Spitzer, *A Tauberian theorem and its probability interpretation*, Trans. Amer. Soc. 94 (1960), p. 150-169.

Mathematical Institute
Maria Curie-Skłodowska University
ul. Nowotki 10
20-031 Lublin, Poland

Received on 20. 11. 1981;
revised version on 29. 6. 1982