THE LIMIT OF A RANDOM WALK AND THE NUMBER THEORETIC DENSITY

by

David C. Heath[1] and William D. Sudderth[2]

Technical Report No. 215

University of Minnesota
Minneapolis, Minnesota

August, 1973

[1]Research supported by the National Science Foundation under NSF Grant GP-28683.

[2]Research supported by the National Science Foundation under NSF Grant GP-39316.

Abstract

It is shown that the limiting behavior of many random walks on the line is closely related to the number theoretic density.

1. Introduction.

Let $X_1, X_2, \ldots$ be independent integer-valued random variables with the same distribution and let $S_n = X_1 + \ldots + X_n$ for all n. Suppose that the random walk $\{S_n\}$ is aperiodic and let E be the collection of even integers. Then, by considering the random walk modulo 2, one can easily see that

$$(*) \qquad \lim_{n\to\infty} P[S_n \in E] = d(E),$$

where

$$d(E) = \lim (2n)^{-1} |E \cap \{-n,\ldots,n\}| .$$

Here $|\cdot|$ denotes cardinality. (The set function d is essentially the "density" of number theory. For an interesting discussion about d, see [2], p.53.) The obvious question is for what other sets does $(*)$ hold. A partial answer is given here. In particular, it follows from Corollary 2 that, if $EX_1 = 0$ and $E(X_1^2) < \infty$, then $(*)$ holds for all sets for which either side is well-defined. Corollary 1 is the corresponding result for non-lattice random walks. The proofs are quite easy but use a Tauberian theorem and the local limit theorem for random walks. A related ratio limit theorem (Theorem 4) is proved without moment conditions under the assumption that X_1 has a symmetric density which decreases on $[0,+\infty)$.

2. Results.

For $x > 0$ and t real, let

$$\varphi_x(t) = (\sqrt{2\pi x})^{-1} e^{-t^2/2x} .$$

Theorem 1. Let f be a bounded, real-valued, Lebesgue measurable function of a real variable. If either

$$\lim_{x\to\infty} (2x)^{-1} \int_{-x}^{x} f(t)dt$$

or

$$\lim_{x\to\infty} \int_{-\infty}^{+\infty} f(t)\varphi_x(t)dt$$

is well-defined, then both limits exist and are equal.

Proof: Assume $f \geq 0$. (If not, just consider $f + c$ where $c \geq - \inf f$.) Set

$$g(x) = f(x) + f(-x) ,$$

$$V(x) = \int_0^x g(t)dt ,$$

and

$$U(x) = \int_0^x (\sqrt{2t})^{-1} g(\sqrt{2t})dt$$

for $x \geq 0$. Notice that $V(\sqrt{2x}) = U(x)$.

Let $\lambda = x^{-1}$ and let ω be the Laplace transform of U. Then, for $x > 0$,

$$\int_{-\infty}^{\infty} f(t)\varphi_x(t)dt = \int_0^{\infty} \varphi_x(t)V(dt)$$

$$= \sqrt{\lambda/2\pi}\int_0^{\infty} e^{-\lambda y}\, U(dy)$$

$$= (\sqrt{\lambda/2\pi})\omega(\lambda) .$$

By a Tauberian theorem (Theorem 2, section XIII.5 of [1]) , for any constant c ,

$$\omega(\lambda) \sim \sqrt{2\pi}\,\lambda^{-1/2} c \qquad \text{as} \quad \lambda \to 0$$

if and only if

$$U(x) \sim 2\sqrt{2x}\, c \qquad \text{as} \quad x \to \infty$$

if and only if

$$V(x) \sim 2x\, c \qquad \text{as} \quad x \to \infty .$$

Since $V(x) = \int_{-x}^{x} f(t)dt$, the proof is complete. □

In the sequel, $X_1, X_2, \ldots$ is a sequence of independent, identically distributed random variables and $S_n = X_1 + \ldots + X_n$ for all n . Also, assume $P[X_1 = 0] < 1$.

Theorem 2. Suppose $EX_1 = 0$, $E(X_1^2) = 1$, and, for some n , the characteristic function of S_n is integrable. Let f be a bounded, real-valued, Lebesgue measurable function on the reals. Then

$$\lim_{k\to\infty} \left|Ef(S_k) - \int_{-\infty}^{\infty} f(x)\varphi_k(x)dx\right| = 0 .$$

Proof: By the local limit theorem (Theorem 2, section XV.5 in [1]), S_k has a density f_k for k sufficiently large and

$$\sqrt{k}\,\|f_k - \varphi_k\| \to 0 \quad \text{as} \quad k \to \infty .$$

(Here, $\|\cdot\|$ denotes the sup norm.) Hence, for k large and $\alpha > 0$,

$$\left|Ef(S_k) - \int_{-\infty}^{\infty} f(x)\varphi_k(x)dx\right|$$

$$\leq \|f\| \int_{-\infty}^{\infty} |f_k(x) - \varphi_k(x)|dx$$

$$\leq \|f\| \left\{ \int_{|x|<\sqrt{k}\,\alpha} |f_k(x) - \varphi_k(x)|dx + \int_{|x|\geq\sqrt{k}\,\alpha} f_k(x)dx + \int_{|x|\geq\sqrt{k}\,\alpha} \varphi_k(x)dx \right\}$$

$$\leq \|f\| \{\alpha\sqrt{k}\,\|f_k - \varphi_k\| + 2\alpha^{-1}\} .$$

Since α is arbitrary, the result follows. □

It is easy to check that the second limit in Theorem 1 could as well have been taken over the positive integers. Thus the following corollary is immediate from Theorems 1 and 2.

<u>Corollary 1</u>: Under the hypotheses of Theorem 2,

$$\lim_{x\to\infty} (2x)^{-1} \int_{-x}^{x} f(t)dt = c$$

if and only if

$$\lim_{n\to\infty} E\, f(S_n) = c .$$

The next theorem corresponds to Theorem 2 in the case when X_1 has integer values. The proof is similar except that a version of the local limit theorem (Theorem 3, section XV.5 of [1]) for lattice random walks must be quoted.

Theorem 3. Suppose X_1 is integer-valued, $EX_1 = 0$, and $E(X_1^2) = 1$. Assume $\{S_n\}$ is aperiodic and let f be a bounded, real-valued function defined on the integers. Then

$$\lim_{n\to\infty} \left| Ef(S_n) - \sum_{k=-\infty}^{\infty} f(k)\varphi_n(k) \right| = 0 .$$

If f is defined on the integers, let $g(x) = f(k)$ for $k \leq x < k+1$. Then, to get the next result, just apply Theorem 1 to g and use Theorem 3.

Corollary 2: Under the hypotheses of Theorem 3.

$$\lim_{n\to\infty} (2n)^{-1} \sum_{k=-n}^{n} f(k) = c$$

if and only if

$$\lim_{n\to\infty} Ef(S_n) = c .$$

Theorem 4. Let S_n be a symmetric random walk with densities f_n which are non-increasing on $[0,\infty)$. Let A and B be Borel subsets of $\mathbb{R}$ and suppose that for each $L > 0$,

$$(*) \qquad \lim_{n\to\infty} \frac{P\{|S_n| \leq L\}}{P\{S_n \in B\}} = 0 .$$

Then if

$$\lim_{x\to\infty} \frac{m(A \cap [-x,x])}{m(B \cap [-x,x])} = c$$

where m is Lebesgue measure, we have

$$\lim_{n\to\infty} \frac{P\{S_n \in A\}}{P\{S_n \in B\}} = c \ .$$

Proof. Fix $\epsilon > 0$. Choose L so that $x \geq L \Rightarrow$

$$\left| \frac{m(A \cap [-x,x])}{m(B \cap [-x,x])} - c \right| \leq \epsilon \ .$$

Using (*) , choose M so large that $m \geq M \Rightarrow \int_{-L}^{L} f_m(\xi)d\xi \leq \epsilon \int_B f_m(\xi)d\xi$.

Set $E = \mathbb{R} \setminus [-L,L]$. Clearly

$$\int_E f_m \cdot 1_A(\xi)d\xi = \lim_{\delta\searrow 0} \delta \sum_{n=1}^{\infty} m\{\xi: f_m \cdot 1_A \cdot 1_E(\xi) \geq n\delta\}$$

and

$$\int_E f_m \cdot 1_B(\xi)d\xi = \lim_{\delta\searrow 0} \delta \sum_{n=1}^{\infty} m\{\xi: f_m \cdot 1_B \cdot 1_E(\xi) \geq n\delta\} \ .$$

Set

$$n(\delta) = \sup\{n: m\{\xi: f_m \cdot 1_A \cdot 1_E(\xi) \geq n\delta\} > 0$$

$$\text{or} \quad m\{\xi: f_m \cdot 1_B \cdot 1_E(\xi) \geq n\delta\} > 0\} \ .$$

The upper limit in both of the previous sums can be taken as $n(\delta)$ (which may be ∞) . We can choose δ small enough that both

$$\int_A f_m(\xi)d\xi - 2\epsilon \int_B f_m(\xi)d\xi \le \int_E f_m \cdot 1_A(\xi)d\xi - \epsilon \int_B f_m(\xi)d\xi$$

$$\le \delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_A \cdot 1_E(\xi) \ge n\delta\} \le \delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_A(\xi) \ge n\delta\}$$

$$\le \int_A f_m(\xi)d\xi$$

and

$$(1-2\epsilon) \int_B f_m(\xi)d\xi \le \int_E f_m \cdot 1_B(\xi)d\xi - \epsilon \int_B f_m(\xi)d\xi$$

$$\le \delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_B \cdot 1_E(\xi) \ge n\delta\} \le \delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_B(\xi) \ge n\delta\}$$

$$\le \int_B f_m(\xi)d\xi$$

are satisfied.

These inequalities imply

$$(1-2\epsilon) \frac{\delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_A(\xi) \ge n\delta\}}{\delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_B(\xi) \ge n\delta\}} \le \frac{\int_A f_m(\xi)d\xi}{\int_B f_m(\xi)d\xi} \le \frac{\delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_A(\xi) \ge n\delta\}}{\delta \sum_{n=1}^{n(\delta)} m\{\xi: f_m \cdot 1_B(\xi) \ge n\delta\}} + \frac{2\epsilon}{1-2\epsilon} .$$

Since $1 \le n \le n(\delta) \Rightarrow m\{\xi: f_m \cdot 1_E(\xi) \ge n\delta\} > 0$, it is easily checked that $\{\xi: f_m(\xi) \ge n\delta\}$ is a symmetric interval about the origin of length at least $2L$. We therefore have $1 \le n \le n(\delta) \Rightarrow$

$$\left| \frac{m\{\xi: f_m \cdot 1_A(\xi) \ge n\delta\}}{m\{\xi: f_m \cdot 1_B(\xi) \ge n\delta\}} - c \right| \le \epsilon .$$

Multiplying through by the denominator and summing over all n with $1 \leq n \leq n(\delta)$, we obtain

$$\left| \frac{\delta \sum_{n=1}^{n(\delta)} m\{\xi : f_m \cdot 1_A(\xi) \geq n\delta\}}{\delta \sum_{n=1}^{n(\delta)} m\{\xi : f_m \cdot 1_B(\xi) \geq n\delta\}} - c \right| \leq \epsilon \text{ , which}$$

completes the proof of Theorem 4. □

The proof of this theorem also works for random walks on the integers; in this case m is counting measure, and f_n is the density of S_n with respect to m .

In many cases the condition labeled (*) is satisfied. In particular, if $B = \mathbb{R}$, one can obtain (*) by considering the random walk modulo some large constant and applying Theorem 3 of section VIII.7 of [1]. The remarks following that theorem allow one to obtain (*) for random walks on the integers $\mathbb{Z}$ with $B = \mathbb{Z}$.

Finally, in the case of the integers, (*) is implied by the usual ratio limit theorem (see p. 4,section 5 of [3]) if B is an infinite set and S_n is recurrent.

References

1. Feller, William (1966). An Introduction to Probability Theory and Its Applications. v.2 , John Wiley, New York.

2. Kac, Mark (1959). Statistical Independence in Probability, Analysis and Number Theory. John Wiley, New York.

3. Spitzer, Frank (1964). Principles of Random Walk. Van Nostrand, Princeton, New Jersey.