The Annals of Probabilty
1982, Vol. 10, No. 1, 132–136

INTERSECTIONS OF TRACES OF RANDOM WALKS WITH FIXED SETS

I. Z. RUZSA AND G. J. SZÉKELY

Mathematical Institute of the Hungarian Academy of Sciences and Eötvös Loránd University

To the memory of Alfréd Rényi

The probability of the event $|S \cap T| = \infty$ is investigated, where S is the trace of a random walk on the set of positive integers and T is a fixed set of natural numbers.

1. Introduction. Let $X, X_1, X_2, \cdots$ be a sequence of positive integer-valued i.i.d. rv's, $S_n = \sum_{k=1}^{n} X_k$ and $S = \{S_1, S_2, \cdots\}$, i.e., S is the trace of the random walk $\{S_n\}$. For sake of simplicity suppose that X is aperiodic, i.e., the g.c.d. of the numbers n for which $P(X = n) > 0$ is 1. In the sequel the term "trace" will be preserved for sequences S of this kind.

We are going to investigate the event

$$|S \cap T| = \infty, \tag{1.1}$$

where T is a fixed set of natural numbers. The value $|S \cap T|$ can be considered as the number of visits of the random walk $\{S_n\}$ into a fixed set T of positions. By the 0 or 1 law of Hewitt and Savage (1955) the event (1.1) happens with probability 0 or 1. Our object is to study when it is 0 and when 1.

2. Universally intersecting traces and sets. Let $X, X_1, X_2, \cdots$ be as above. A theorem of Erdös, Feller and Pollard (1949) asserts that

$$\lim_{k\to\infty} P(k \in S) = E(X)^{-1} \tag{2.1}$$

where S is a trace and E denotes the expectation (the right-hand side is to be interpreted as 0 in case $E(X) = \infty$). (2.1) easily yields:

THEOREM 1. a. *If $E(X) < \infty$, then*

$$P(|S \cap T| = \infty) = 1 \tag{2.2}$$

for any infinite set T of natural numbers.

b. *If $E(X) = \infty$, then one can always find an infinite set T of natural numbers such that*

$$P(|S \cap T| = \infty) = 0. \tag{2.3}$$

That is we know which traces intersect every infinite set; now we determine which sets intersect every trace.

THEOREM 2. a. *If $T = \{T_1, T_2, \cdots\}$ is an infinite increasing sequence of natural numbers and the differences $T_{n+1} - T_n (n = 1, 2, \cdots)$ are bounded, then* (2.2) *holds for every trace S.*

b. *If $T_{n+1} - T_n (n = 1, 2, \cdots)$ is unbounded, then there always exists a trace S such that* (2.3) *holds.*

Received February 26, 1980; revised December 8, 1980.
AMS 1980 *classification numbers.* 60G50, 60J10.
Key words and phrases. Regenerative phenomena, Borel-Cantelli lemma, Markov chains.

3. Borel-Cantelli does not decide. If $E(X) = \infty$ and $T_{n+1} - T_n$ is unbounded, then Theorems 1 and 2 do not provide any information. The Borel-Cantelli lemma yields

$$P(|S \cap T| = \infty) = 0 \tag{3.1}$$

if

$$\textstyle\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} P(S_n = T_m) < \infty. \tag{3.2}$$

The condition (3.2) is, however, far from being necessary. Indeed, we have the following result.

THEOREM 3. *Given an arbitrary trace S such that* $E(X) = \infty$, *there always exists a sequence T such that* (3.1) *holds, though the series* (3.2) *is divergent.*

REMARK. Under some further restrictions (3.2) is, however, necessary for (3.1). Put

$$\beta_k = |S \cap \{T_1, T_2, \cdots, T_k\}|;$$

note that $E(\beta_k) > 0$ for large k. (3.2) means that $E(\beta_k) = O(1)$, and (3.1) that β_k is a.s. bounded; thus if (3.1) holds but (3.2) does not, then the normalized variables $\beta_k/E(\beta_k)$ tend to 0 a.s. and hence also in distribution. This cannot happen if $\beta_k/E(\beta_k)$ has a uniformly integrable subsequence, in other words if one can find a positive real function f such that $f(x)/x \to \infty$ for $x \to \infty$ but

$$\liminf_{k\to\infty} E(f(\beta_k/E(\beta_k))) < \infty. \tag{3.3}$$

This condition with $f(x) = x^2$ will be used in Section 4.

The problem is much simpler if T is not a fixed set but itself a trace, independent from S.

THEOREM 4. *If S and T are independent traces, then* (3.1) *and* (3.2) *are equivalent.*

4. Random Dirichlet problem. The dual of Theorem 3 does not seem to be true, i.e., we do not expect that if $T_{n+1} - T_n (n = 1, 2, \cdots)$ is unbounded (e.g., $T_n = n^2$) then one can always find a trace S such that (3.1) holds while (3.2) is divergent. The characterization of the sequences T for which (3.1) implies (3.2) is unsolved and probably difficult.

PROBLEM. Does the sequence of primes have this property?

We believe it has but we can prove only some special cases. The answer is affirmative if e.g. the sequence $u_n = P(n \in S) (n = 1, 2, \cdots)$ is asymptotically nonincreasing, by which we mean that for $m > n > n_0$ the inequality $u_m < cu_n$ holds with a constant $c = c(S)$ independent of m and n, for in this case (3.3) holds with $f(x) = x^2$. (If u_k is asymptotically nonincreasing and T is an arbitrary set such that

$$\tau(x+y) - \tau(y) \le K\,\tau(x) \tag{4.1}$$

for some finite K where $\tau(x)$ denotes the number of elements belonging to T and not larger than x, then (3.3) always holds with $f(x) = x^2$. In case when T is the set of primes (4.1) is known to hold.) If the restriction "asymptotically nonicreasing" could be omitted, we would get a random analogue of Dirichlet's celebrated theorem on the infinitude of primes in arithmetic progressions (c.f. also Erdös and Chung (1952) concerning a related problem).

5. Proof of Theorems 1 and 4. Given the sequence $\{X_i\}$ and $S_n = \sum_{i=1}^{n} X_i$ we define

$$M_n^S = \min\{d : d \ge 0, \quad n - d \in S\}.$$

It is easy to check that M_n^S is a Markov chain. Hence by the Erdös-Feller-Pollard theorem (see e.g. Kingman (1972), page 12)

$$P(n \in S) = P(M_n^S = 0) \to E(X)^{-1}$$

which was stated in (2.1).

PROOF OF THEOREM 1. If $E(X) < \infty$, then

$$P(T_n \in S \quad \text{for infinitely many} \quad n) \geq \lim P(T_n \in S) = E(X)^{-1} > 0,$$

thus this probability must be 1 by the cited 0–1 law of Hewitt and Savage. On the other hand, if $E(X) = \infty$, then we can choose T_n so that

$$\textstyle\sum_n P(T_n \in S) < \infty,$$

and hence the Borel-Cantelli lemma implies $P(T \cap S| = \infty) = 0$.

PROOF OF THEOREM 4. We consider the two Markov chains M_n^S, M_n^T and define their composition $M_n = (M_n^S, M_n^T)$ which will be also a Markov chain. The existence of infinitely many $n \in S \cap T$ is equivalent to M_n returning to the state (0, 0) infinitely many times. This happens with probability 0 or 1 according to the convergence or divergence of the series $\sum_n P(M_n = (0, 0)) = \sum_n P(n \in S \cap T) = \sum_n P(n \in S)P(n \in T)$. (Cf. Chung (1960), Theorems 4.3 and 5.4).

6. Proof of Theorem 2. To prove part a. we need a lemma.

(6.1) LEMMA. *For a trace S, a fixed set T and integer n*

$$P(S \cap (T + n) = \infty)$$

is either 0 *for all n or* 1 *for all n.*

PROOF. Write

$$A_j = \{n : P(S \cap (T + n) = \infty) = j\}, \qquad j = 0, 1.$$

A_0 and A_1 are disjoint and $A_0 \cup A_1 = Z$ by the much quoted Hewitt-Savage theorem. Now we show that if $n \in A_j$ and $P(S_k = a) > 0$ for some k, then $n + a \in A_j$. Namely if $j = 1$, then

$$P(|S \cap (T + a + n)| = \infty) \geq P(S_k = a)P(|S \cap (T + n)| = \infty) > 0,$$

thus $a + n \notin A_0$; for $j = 0$ similarly we obtain $a + n \notin A_1$.

As X is aperiodic, for every large a there is a k such that $P(S_k = a) > 0$. Hence if A_j is not empty, it contains all large numbers. As this cannot hold for both $j = 0, 1$, one of them must be empty.

PROOF OF THEOREM 2a. Suppose the contrary. The above lemma yields

$$P(|S \cap (T + n)| < \infty) = 1$$

for all n. Hence with

$$Z_k = T \cup (T + 1) \cup \cdots \cup (T + k)$$

we would also have

$$P(|S \cap Z_k| < \infty) = 1. \tag{6.2}$$

But for $k > \limsup_{n\to\infty} (T_{n+1} - T_n)$,

Z_k contains all but finitely many natural numbers, and hence (6.2) is impossible.
To prove the second part of the Theorem we need the following lemma.

(6.3) LEMMA. *Let $0 < r_1 < r_2 < \cdots$ be an infinite sequence of integers and let $X, X_1, X_2, \cdots$ be* i.i.d. rv's with the distribution

$$P(X = r_m) = m^{-\alpha} - (m + 1)^{-\alpha}$$

for some fixed $\alpha \in (0, 1)$. Let M_n be the maximum of $X_1, X_2, \cdots, X_n$ and M'_n the maximum of the remaining $n - 1$ terms. Then we have

$$P(M_n < r_n \text{ i.o.}) = 0,$$
$$P(M_n = M'_n \text{ i.o.}) = 0.$$

PROOF. We have $P(M_n < r_n) = (1 - (n + 1)^{-\alpha})^n \le c_1 e^{-n^{1-\alpha}}$ and

$$P(M_n = M'_n = r_k) = \binom{n}{2} (k^{-\alpha} - (k + 1)^{-\alpha}]^2(1 - k^{-\alpha})^{n-2}$$
$$\le c_2 \binom{n}{2} k^{-2(1+\alpha)} e^{-nk^{-\alpha}}.$$

Hence an easy computation yields

$$\textstyle\sum_n P(M_n < r_n) < \infty$$

and

$$\textstyle\sum_n P(M_n = M'_n) = \sum_k \sum_n P(M_n = M'_n = r_k) \le \sum_k c_2 k^{\alpha-2} < \infty,$$

and our assertions follow from the Borel-Cantelli lemma.

PROOF OF THEOREM 2b. The underlying idea in the proof is the observation that one can choose the distribution of X so that the partial sums and the partial maxima will be generally close.

If $T_{n+1} - T_n$ is unbounded then we can find intervals $[r_i, r_i + s_i]$ containing no T_j and such that $s_i > ir_{i-1}$, $r_i > r_{i-1} + s_{i-1}$. Now choose $X_1, X_2, \cdots$ as in Lemma (6.3). By this lemma with probability 1 for all large n there is a maximal r_j among $X_1, X_2, \cdots, X_n$, $j > n$, and it occurs only once. Hence

$$S_n = \textstyle\sum_{k=1}^n X_k \in [r_j, r_j + (n - 1)r_{j-1}] \subset [r_j, r_j + s_j]$$

as required. (X is aperiodic assuming $r_1 = 1$).

7. Proof of Theorem 3. Let $u_n = P(n \in S)$; we know by (2.1) that $u_n \to 0$ and evidently $\sum_{n=1}^\infty u_n = \infty$. Given an r, let $s = s(r)$ denote the smallest number such that

$$\textstyle\sum_{n=r}^{n+s} u_n > 1;$$

obviously we have

$$E(|S \cap [r, r + s]|) = \textstyle\sum_{n=r}^{r+s} u_n \in [1, 2]. \tag{7.1}$$

Let K be large but fixed. Divide the event $S \cap [r, r + s] \ne \emptyset$ into the subevents

$$Z_1 : S \cap [r, r + s - K] \ne \emptyset, \qquad Z_2 : S \cap [r + s - K, r + s] \ne \emptyset.$$

Obviously for fixed K we have $P(Z_2) \to 0$ as $r \to \infty$. On the other hand

$$E(|S \cap [r, r + s]|) \ge P(Z_1)E(|S \cap [1, K]|),$$

so that we obtain

$$P(S \cap [r, r + s] \ne \emptyset) \le \frac{2}{E(|S \cap [1, K]|)} + o(1).$$

Since this holds for arbitrary K and $E(|S \cap [1, K]|) \to \infty$ as $K \to \infty$, we have

$$P(S \cap [r, r+s] \neq \varnothing) \to 0 \qquad (r \to \infty). \tag{7.2}$$

Now choose a sequence r_n such that

$$r_{n+1} > r_n + s(r_n)$$

and

$$\textstyle\sum_{n=1}^{\infty} P(S \cap [r_n, r_n + s(r_n)] \neq \varnothing) < \infty.$$

This means that with probability one S intersects only finitely many intervals $[r_n, r_n + s(r_n)]$. Then its intersection with

$$T = \cup_n [r_n, r_n + s(r_n)]$$

is finite with probability one; on the other hand (7.1) ensures that the series (3.2) is divergent.

Acknowledgements. We are grateful to A. A. Balkema and L. de Haan for their valuable remarks.

REFERENCES

CHUNG, K. L. (1960). *Markov Chains with Stationary Transition Probabilities.* Springer, Berlin.

ERDÖS, P. and CHUNG, K. L. (1952). On the application of the Borel-Cantelli lemma. *Trans. Amer. Math. Soc.* **72** 179–186.

ERDÖS, P., FELLER, W. and POLLARD, U. (1949). A theorem on power series. *Bull. Amer. Math. Soc.* **55** 201–204.

HEWITT, E. and SAVAGE, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501.

KINGMAN, J. F. C. (1972). *Regenerative Phenomena.* Wiley, New York.

MATHEMATICAL INSTITUTE OF
THE HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, REALTANODA U. 13–15
H-1053 HUNGARY

DEPARTMENT OF PROBABILITY THEORY
EÖTVÖS LORÁND UNIVERSITY
MUZEUM KÖRUT 6–8,
BUDAPEST,
H-1088 HUNGARY