PROBLEMS AND RESULTS ON RANDOM WALKS

P.Erdös
Mathematical Institute
Reáltanoda u 13-15
1053 Budapest
Hungary

P.Révész
Mathematical Institute Budapest
and Technical University
Wiedner Hauptstraße 8-10/107
A-1040 Vienna, Austria

ABSTRACT: In this papes we present a number of unsolved problems of the simple, symmetric random walk together with the relevant known results.

1. INTRODUCTION

We consider the simple, symmetric random walk on the r-dimensional integer lattice. It is perhaps surprising how many unsolved problems remain in this old subject. In this paper we present a number of unsolved problems together with the relevant known results. We do not give any proofs but we give as many references as possible. Together with the presented unsolved problems we try to indicate whether we believe that they can be solved by the methods standing at our disposal or we feel that some new ideas of methods are necessary to settle them.

2. RANDOM WALK ON THE LINE

Let $X_1, X_2, \ldots$ be a sequence of i.i.d.r.v.'s with

$$P(X_1=+1) = P(X_1=-1) = 1/2$$

and

$$S_0 = 0, \quad S_n = \sum_{i=1}^{n} X_i \qquad (n=1,2,\ldots).$$

S_n is considered as the location of the particle (involved in the random walk) after n steps.

2.1. The favourite values of a random walk

Let

$$\xi(x,n) = \# \{k: \ 0 \le k \le n, \ S_k = x\}$$

be the local time of the random walk, i.e. $\xi(x,n)$ is the number of visits in x up to n. A point x_n is called a favourite value at the moment n if the particle visits x_n most often during the first n steps i.e.

P. Bauer et al. (eds.), Mathematical Statistics and Probability Theory, Vol. B, 59–65.

$$\xi(x_n,n) = \max_x \xi(x,n).$$

The investigation of the properties of the favourite values started simultaneously by Bass and Griffin (1985) and ourselves (1984). One can easily observe that for infinitely many n there are two favourite values and also for infinitely many n there is only one favourite value with probability one. More formally speaking let F_n be the set of favourite values i.e.

$$F_n = \{x: \xi(x,n) = \max \xi(x,n)\}$$

and let $|F_n|$ be the cardinality of F_n. Then

$$P\{|F_n| = 2 \text{ i.o.}\} = P\{|F_n| = 1 \text{ i.o.}\} = 1.$$

1.) We do not know wether 3 or more favourite values can occur infinitely often i.e. we ask:

$$P\{|F_n| = r \text{ i.o.}\} = ? \qquad (r=3,4,5,\ldots).$$

We thought that 0 is a favourite value i.o. that is $P\{0 \in F_n \text{ i.o.}\} = 1$. To our great surprise Bass and Griffin showed that it is not so and they proved that the favourite values are going to infinity faster than $n^{1/2}(\log n)^{-11}$. In fact they have

$$P\{\lim_{n\to\infty} \frac{(\log n)^{\alpha}}{n^{1/2}} \inf\{|x|, x \in F_n\} = \infty\} = 1$$

if $\alpha>11$. We showed that the favourite value i.o. larger than $(1-\varepsilon)(2n \log\log n)^{1/2}$ i.e.

$$P\{\limsup_{n\to\infty} (1+\varepsilon) (2n \log\log n)^{-1/2} \inf\{|x|, x\in F_n\}=1 \text{ i.o.}\}=1.$$

2.) We do not know whether the ε can be replaced by 0 in the above statement.

Let $\alpha(n)$ be the number of different favourite values up to n, i.e.

$\alpha(n) = |\sum_{k=1}^{n} F_k|$. We guess that $\alpha(n)$ is very small i.e. $\alpha(n) <$ $(\log n)^c$ for some $c > 0$ but we cannot prove it. Hence we ask

3.) How can one describe the limit behaviour of $\alpha(n)$?

4.) We also ask how long can a point stay as a favourite value i.e. let $1 \le i=i(n) < j=j(n) \le n$ be two integers for which

$$|\prod_{k=i}^{j} F_k| \ge 1$$

and $j-i=\beta(n)$ is as big as possible. The question is to describe the limit behaviour of $\beta(n)$.

5.) Further if x was a favourite value once, can it happen that the favourite value moves away from x but later it returns to x again, i.e. do sequences $a_n<b_n<c_n$ of positive random integers exist such that

$$F_{a_n} F_{b_n} = \emptyset \quad \text{and} \quad F_{a_n} F_{c_n} \neq \emptyset \qquad (n=1,2,\ldots)?$$

6.) To investigate the jumps of the favourite values looks also interesting. Let $n=n(\omega)$ be a positive integer for which $F_n F_{n+1} = \emptyset$. Then the jump j_n is defined as

$$j_n = \rho(F_n,F_{n+1}) = \min\{|x-y|; x \in F_n, y \in F_{n+1}\}.$$

The theorem of Bass and Griffin implies that $j_n \ge n^{1/2}(\log n)^{-11}$ i.o.a.s.

It looks very likely that $\lim_{n\to\infty} j_n = \infty$ a.s. We do not see how one can describe the limit behaviour of j_n.

2.2 Long head-runs

We (1976) studied the length of the longest head-run Z_n, i.e. Z_n is the largest integer for which

$$I(n,Z_n)=Z_n$$

where

$$I(n,k) = \max_{0\le j\le n-k} (S_{j+k}-S_j) \qquad (0 \le k \le n).$$

Our 1976 paper contained a complete enough characterization of Z_n but the result was extended by Guibas-Odlyzko (1980), Samarova (1981), Révész (1982), Ortega-Wschebor (1984), Deheuvels (1985), Deheuvels, Devroye-Lynch (1986), Deheuvels-Steinebach (1986), Erdős-Révész (1986) among others. Now we propose some further problems.

Let Z_n^* be the length of the longest tail run, i.e. Z_n^* is the largest integer for which

$$I^*(n,Z_n^*) = -Z_n^*$$

where

$$I^*(n,k) = \min_{0\le j\le n-k} (S_{j+k}-S_j).$$

7.) How can we characterize the limit properties of $|Z_n-Z_n^*|$?

A trivial argument shows $P\{Z_n=Z_n^* \text{ i.o.}\}=1$ but it is not clear at all how big $|Z_n-Z_n^*|$ can be.

Let $Z_n^{(1)} = Z_n$ and let $Z_n^{(2)}, Z_n^{(3)}, \ldots$ be the length of the second, third,... longest head-run up to n.

8.) We ask about the properties of $Z_n^{(1)}-Z_n^{(2)}$. It is clear again that $P(Z_n^{(1)}=Z_n^{(2)} \text{ i.o.}) = 1$. The lim sup properties of $Z_n^{(1)}-Z_n^{(2)}$ look harder.

9.) Let k_n be the largest integer for which

$$P\{Z_n^{(1)}=Z_n^{(2)}=\ldots=Z_n^{(k_n)} \text{ i.o.}\} = 1.$$

Characterize the limit properties of k_n.

2.3. On logarithmic and other densities

In many problems on random walk one gets density results only if we replace ordinary density by some more general concept of density. As an example we consider the sequence of time points when the particle returns to the origin. Let

$$Y_k = \begin{cases} 1 & \text{if} \quad S_k=0, \\ 0 & \text{if} \quad S_k\neq 0. \end{cases}$$

Then the sequence $\xi(0,n) = \sum_{k=1}^{n} Y_k$ does not obey law of large numbers but Chung and Erdős (1951) proved

$$\lim_{n\to\infty} (\log n)^{-1} \sum_{k=1}^{n} k^{-1/2} Y_k = \pi^{1/2} \quad \text{a.s.}$$

A similar result is due to P.Lévy who investigated the logarithmic density of those n's for which $S_n>0$. In fact let

$$V_n = \begin{cases} 1 & \text{if} \quad S_n > 0, \\ 0 & \text{if} \quad S_n \le 0. \end{cases}$$

Then P.Lévy proved that

$$(\log n)^{-1} \sum_{k=1}^{n} k^{-1} V_k = 1/2 \quad \text{a.s.}$$

In connection with result we ask:

10.) Does the sequence

$$(\log n)^{-1/2} \left(\sum_{k=1}^{n} k^{-1} V_k - \frac{1}{2} \log n \right)$$

satisfy the central limit theorem. This question does not seem to be very hard.

Our next problem is connected to the problem of long head-runs. Using the same notation as above let

$$U_n = \begin{cases} 0 & \text{if} \quad Z_n < Z_n^*, \\ 1 & \text{if} \quad Z_n > Z_n^*. \end{cases}$$

i.e. $U_n = 1$ if the longest head run up to n is longer than the longest tail run.

11.) Does the logarithmic density

$$\lim_{n\to\infty} (\log n)^{-1} \sum_{k=1}^{n} k^{-1} U_k$$

exist with probability one. This problem seems to be not very hard.

2.4. Rarely visited points

It is easy to see that for infinitely many n almost all paths assume every value at least twice which they assume at all, i.e. let

$$f_r(n) = \# \{k: \xi(k,n) = r\}$$

be the number of points visited exactly r-times up to n. Then

$$P\{f_1(n) = 0 \text{ i.o.}\} = 1.$$

12.) We do not know if for inifinitely many n almost all paths assume every value at least (r+1)-times (r=2,3,...) which they assume at all, i.e. let

$$\sum_{j=1}^{r} f_j(n) = g_r(n)$$

and we ask

$$P\{g_r(n) = 0 \text{ i.o.}\} = ?$$

We would guess that this probability is 0 if r > 2 but perhaps it is 1 if r=2.

13.) For every r (r may depend on n) investigate

$$\liminf_{n\to\infty} f_r(n) \text{ and } \limsup_{n\to\infty} f_r(n).$$

As already stated $\liminf_{n\to\infty} f_1(n) = 0$. P.Major (1986) proved

$$\limsup_{n\to\infty} \frac{f_1(n)}{\log^2 n} = C \qquad \text{a.s.}$$

where $0 < C < \infty$ but its exact value is unknown.
Let e_i be the i-th unit-vector on R^d i.e. $e_i=(0,0,\dots 0,i,0,\dots 0)$ and let $X_1, X_2,\dots$ be a sequence of i.i.d.r.v.'s with

$$P\{X_1=e_i\} = P\{X_1=-e_i\} = \frac{1}{2d} \qquad (i=1,2,\dots,d).$$

Further let

$$S_0=0, \quad S_n=X_1+X_2+\dots+X_n \qquad (n=1,2,\dots).$$

3. RANDOM WALK IN THE SPACE

Most of the problems formulated in Section 2 can be reformulated in d-dimension and a number of new problems can be found. At first we give a few remarks to the already stated problems. Later we present some new problems.

3.1. Multivariate versions of the one dimensional problems.

In connection with the favourite values it is natural to ask

14.) Does the favourite value of the random walk converge to inifinity in case d=2?

The answer is clearly positive when $d \geq 3$ and very likely it is also so in case d=2 but the proof is not clear.

In connection with the rarely visited points a result of Dvoretzky and Erdős (1950) implies that in case $d \geq 2$ a.s. there will be many points visited exactly once if n is big enough. In fact we have $\lim_{n\to\infty} f_r(n) = \infty$ a.s. if $d \geq 2$. Dvoretzky and Erdős as well as Erdős and Taylor (1960 and 1960) have some results to describe the limit properties of $f_r(n)$ but a complete description is missing.

3.2. Special problems in case $d \geq 2$.

Let us consider the largest square around the origin completely covered by the path during the first n steps. Clearly we say that a square $[-A,A] \times [-A,A]$ is completely covered during the first n steps if for any $x \in [-A,A] \times [-A,A]$ there exists a $1 \leq k = k_n(x) \leq n$ such that $S_k=x$. Let A_n be the largest integer for which the square $[-A_n,A_n] \times [-A_n,A_n]$ is completely covered. Clearly $\lim_{n\to\infty} A_n = \infty$ a.s. We ask

15.) How rapidly converges A_n to infinity?

Clearly in case $d \geq 3$ the volume of the largest completely covered cube around the origin does not go to infinity. However there will be a completely covered large cube somewhere.

16.) What is the volume of the largest completely covered cube in case $d \geq 2$?

17.) Where is the largest completely covered cube located?

a) in case $d=2$ we ask whether the center of this cube converges to infinity

b) in case $d \geq 3$ it is clear that the center is going to infinity but the speed is not clear.

Instead of the largest completely covered cube we can consider the largest "essentially" covered one. For example one can consider the largest integer $B_n=B_n(p)$ $(0<p<1)$ for which 100p% of the cube $[-B_n, B_n]^2$ is covered during the first n steps.

18.) Question 15-17 should be reformulated for essentially covered cubs.

Question 15 is already formulated in Erdös-Taylor (1960) where an intutive solution is also given.

The following questions look connected to the above ones

19.) Between n and $n+t_n$ $(t_n \uparrow \infty)$ how many new points will be covered? $S_j (n \leq j \leq t_n)$ can be considered as a newly covered point if

(i) $S_j \neq S_k$ $(k=0,1,2,\ldots,j-1)$

or

(ii) $S_j \neq S_k$ $(k=n,n+1,\ldots,j-1)$.

20.) How long time do we have to wait after n steps to obtain a new point? In fact let Z_n be the smallest integer for which

$$S_{n+Z_n} \neq S_i \qquad (i=1,2,\ldots,n).$$

How can we characterize Z_n?

References:

Bass, R.F.-Griffin, P.S.(1985) 'The most visited site of Brownian motion and simple random walk'. Z. Wahrscheinlichkeitstheorie verw. Gebiete 70, 417-436.

Chung, K.L.-Erdös, P.(1951) 'Probability limit theorems assuming only the first moment I'. Four papers on probability. Mem. Amer. Math. Soc. No. 6.

Deheuvels, P. (1985) 'On the Erdös-Rényi theorem for random fields and sequences and its relationships with the theory of runs and spacings'. Z. Wahrscheinlichkeitstheorie verw. Gebiete 70, 91-115.

Deheuvels, P.-Devroye, L.-Lynch, J. (1986) 'Exact convergence rate in the limit theorems of Erdös-Rényi and Shepp'. Ann. Probability 14, 209-223.

Deheuvels, P.-Steinebach, J. (1986) 'Exact convergence rates in strong approximation laws for large increments of partial sums'. Preprint

Dvoretzky, A.-Erdös, P. (1950) 'Some problems on random walk in space'. Proc. Second Berkeley Symposium 353-368.

Erdös, P.-Révész, P. (1976) 'On the length of the longest head run'. Coll. Math. Soc. J.Bolyai 16, Topics in Information Theory, North Holland.

Erdös, P.-Révész, P. (1984) 'On the favourite points of a random walk'. Mathematical structures-Computational mathematics-Mathematical Modelling, 2. Sofia.

Erdös, P.-Révész, P.(1986) 'Many heads in a short block'. This volume.

Erdős, P.-Taylor, S.J. (1960) 'Some problems concerning the structure of random walk paths'. Acta Math. Acad. Sci. Hung. 11, 137-162.

Erdős, P.-Taylor, S.J. (1960) 'Some intersection properties of random walk paths'. Acta Math. Sci. Hung. 11, 231-248.

Guibas, L.J.-Odlyzko, A.M. (1980) 'Long repetitive patterns in random sequence'. Z. Wahrscheinlichkeitstheorie verw. Gebiete 53, 241-262.

Major, P. (1986) 'On the set visited once by a random walk'. Preprint.

Ortega, J.-Wschebor, M. (1984) 'On the increments of the Wiener process' Z. Wahrscheinlichkeitstheorie verw. Gebiete 65, 329-339.

Révész, P. (1982) 'On the increments of Wiener and related processes'. Ann. Probability 10, 613-622.

Samarova, S.S. (1981) 'On the length of the longest head-run for a Markov chain with two states'. Theory Probab. Appl. 26, 489-509.