CHAPTER 4: RANDOM WALKS

September 7, 2021

Contents

1 Random Walks

Definition. Let $X_1, X_2, \ldots$ be i.i.d. random variables in $\mathbb{R}^d$. Let $S_n = X_1 + \cdots + X_n$. Then $\{S_n : n \geq 1\}$ is called a **random walk**. We usually take $S_0 = 0$.

Definition. In the above setting, if $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$, then $\{S_n : n \geq 1\}$ is called a **simple random walk (SRW)** in $\mathbb{R}^1$. If $P(X_i = (1,1)) = P(X_i = (1,-1)) = P(X_i = (-1,1)) = P(X_i = (-1,-1)) = \frac{1}{4}$, then $\{S_n : n \geq 1\}$ is called a SRW in $\mathbb{R}^2$. Each step of SRW in $\mathbb{R}^d$ choose a uniform random direction and goes forward or backward by amount 1 with equal probability in that random direction.

As opposed to previous chapters, in this chapter, we focus on the whole trajectory of $\{S_n : n \geq 1\}$.

We first give a generalization of Kolmogorov's 0-1 law to permutable events defined below.

Definition. An event is **permutable (or exchangeable)** if it does not change under finite permutation of $\{X_1, X_2, \ldots\}$.

Note that all tail events are permutable. However, there are permutable events which are not tail events. For example, $\{\limsup_{n\to\infty} S_n \geq c\}$ for a given constant c. In this sense, the following Hewitt-Savage 0-1 law is more general. We do not give the proof, which is slightly more complicated.

Theorem 4.1.1. If A is permutable, then $P(A) = 0$ or 1.

With the help of the 0-1 law, we can give a classification of the behavior of all the random walks in $\mathbb{R}^1$.

Theorem 4.1.2. For a random walk on $\mathbb{R}^1$, one of the following has probability 1:
(i) $S_n = 0$ for all n.
(ii) $S_n \to \infty$, as $n \to \infty$.
(ii) $S_n \to -\infty$, as $n \to \infty$.
(iv) $-\infty = \liminf_{n\to\infty} S_n < \limsup_{n\to\infty} S_n = \infty$.

Proof. By the 0-1 law, $\{\limsup_{n\to\infty} S_n \geq c\}$ has probability 0 or 1. This mean with probability 1, $\limsup_{n\to\infty} S_n$ equals some value in $\mathbb{R} \cup \{-\infty, \infty\}$. Similarly, $\liminf_{n\to\infty} S_n$ equals some value in $\mathbb{R} \cup \{-\infty, \infty\}$. If $\limsup_{n\to\infty} S_n = c$ for some $c \in \mathbb{R}$, then from

$$\limsup_{n\to\infty} S_n \stackrel{d}{=} \limsup_{n\to\infty} (S_{n+1} - X_1) = \limsup_{n\to\infty} S_{n+1} - X_1 = \limsup_{n\to\infty} S_n - X_1,$$

we must have $X_1 = 0$ and it belongs to case (i). Other cases correspond to other combinations of possibilities of $\limsup_{n\to\infty} S_n$ and $\liminf_{n\to\infty} S_n$. $\square$

Next, we introduce filtration and stopping times.

Definition. Let $X_1, X_2, \dots$ be a sequence of random variables,

$$\mathcal{F}_n := \sigma(X_1, \dots, X_n),$$

as an increasing sequence of σ-fields, is called a **filtration**. We usually take $\mathcal{F}_0$ to be the trivial σ-field $\{\emptyset, \Omega\}$.

Definition. $\tau \in \{1, 2 \dots, \} \cup \{\infty\}$ is a **stopping time** (or optional random variable, optimal time, Markov time) with respect to the filtration $\{\mathcal{F}_n\}$ if

$$\{\tau = n\} \in \mathcal{F}_n, \ \forall\ n = 1, 2, \dots$$

Remarks. We can equivalently define stopping times by requiring

$$\{\tau \le n\} \in \mathcal{F}_n, \ \forall\ n = 1, 2, \dots$$

or

$$\{\tau \ge n+1\} \in \mathcal{F}_n, \ \forall\ n = 1, 2, \dots$$

All of them result in the same definition.

Fact. If τ_1, τ_2 are both stopping times with respect to (w.r.t.) $\{\mathcal{F}_n\}$, then $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$, $\tau_1 + \tau_2$ are all stopping times w.r.t to $\{\mathcal{F}_n\}$.

Example. Let $X_1, X_2, \dots$ be a sequence of random vectors in $\mathbb{R}^d$ and

$$\mathcal{F}_n := \sigma(X_1, \dots, X_n).$$

Let $S_n = X_1 + \dots + X_n$. Let A be a measurable subset of $\mathbb{R}^d$. Then,

$$\tau := \inf\{n \ge 1 : S_n \in A\}$$

is a stopping time.

The next two results are useful in computations involving stopping times. In the following, for a sequence of random variables $X_1, X_2, \dots$, stopping times are by default w.r.t. the filtration $\{\mathcal{F}_n\}$ where $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.

Theorem 4.1.5 (Wald's Equation). Let $X_1, X_2, \dots$ be i.i.d. τ is a stopping time. If

$$E|X_1| < \infty, \quad E\tau < \infty,$$

then

$$E(S_\tau) = E(X_1)E(\tau). \tag{1.1}$$

Proof. First consider the case $X_i \geq 0$. We have

$$\begin{aligned}
E(S_\tau) =& E(\sum_{i=1}^{\tau} X_i) \\
=& E(\sum_{i=1}^{\infty} X_i 1_{\{\tau \geq i\}}) \\
=& \sum_{i=1}^{\infty} E(X_i 1_{\{\tau \geq i\}}) && \text{(Fubini)} \\
=& \sum_{i=1}^{\infty} E(X_i) E(1_{\{\tau \geq i\}}) && (\{\tau \geq i\} \in \mathcal{F}_{i-1}) \\
=& E(X_1) E(\tau).
\end{aligned}$$

For the general case, (1.1) again follows by interchanging integrals and the conditions $E|X_1| < \infty$ and $E\tau < \infty$ ensure that Fubini still applies. □

Theorem 4.1.6 (Wald's Second Equation). Let $X_1, X_2, \ldots$ be i.i.d. τ is a stopping time. If

$$EX_1 = 0, \quad EX_1^2 = \sigma^2 < \infty, \quad E\tau < \infty,$$

then

$$E(S_\tau^2) = \sigma^2 E(\tau).$$

Proof. Write

$$S_{\tau \wedge n}^2 = S_{\tau \wedge (n-1)}^2 + (2X_n S_{n-1} + X_n^2) 1_{\{\tau \geq n\}}.$$

We have, note that all the following expectations exist,

$$\begin{aligned}
ES_{\tau \wedge n}^2 =& ES_{\tau \wedge (n-1)}^2 + E\left(2X_n S_{n-1} 1_{\{\tau \geq n\}}\right) + E[X_n^2 1_{\{\tau \geq n\}}] \\
=& ES_{\tau \wedge (n-1)}^2 + \sigma^2 P(\tau \geq n) && \text{(again using stopping time and independence)} \\
=& \cdots && (\text{reduce to } n-2, n-3, \ldots) \\
=& \sigma^2 \sum_{i=1}^{n} P(\tau \geq i).
\end{aligned}$$

Letting $n \to \infty$ gives the result. □

Example. Let $X_1, X_2, \ldots$ be i.i.d. $\sim$ Uniform$(0,1)$. $S_n = X_1 + \cdots + X_n$. Let $\tau := \inf\{n : S_n > 1\}$. Then

$$E(\tau) = e, \quad E(S_\tau) = \frac{e}{2}.$$

Proof. We have

$$\begin{aligned}
P(\tau > n) =& P(S_n \le 1) \\
=& \int_0^1 \cdots \int_0^1 1_{\{x_1+\cdots+x_n \le 1\}} dx_1 \cdots dx_n && \text{(from the uniform distribution)} \\
=& \int_0^1 \int_0^{y_n} \cdots \int_0^{y_2} dy_1 dy_2 \cdots dy_n && \text{(by a change of variable)} \\
=& \frac{1}{n!}.
\end{aligned}$$

This implies

$$E(\tau) = \sum_{n=0}^{\infty} P(\tau > n) = e.$$

By Wald's first equation,

$$E(S_\tau) = E(X_1)E(\tau) = \frac{e}{2}.$$

□

Example. Let $X_1, X_2, \ldots$ be i.i.d. with $P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}$. Let $S_n = X_1 + \cdots + X_n$ (SRW). Let a, b be two integers with $a < 0 < b$. Let

$$N := \inf\{n : S_n \notin (a, b)\}.$$

Then

1. $E(N) < \infty$,
2. $S_N = a$ or b,
3. $P(S_N = a) = \frac{b}{b-a}, \quad P(S_N = b) = \frac{-a}{b-a}$,
4. $E(N) = E(S_N^2) = (-a)b$.

Proof. For any positive integer k, by dividing the interval $(0, k(b-a))$ into k subintervals of equal length and considering an extreme case behavior (keep going upwards) of the random walk within each subinterval, we obtain

$$P(N > k(b-a)) \le (1 - \frac{1}{2^{b-a}})^k.$$

This implies 1.

2. is obvious.
3. follows from Wald's first equation.
4. follows from Wald's second equation. □

2 Recurrence vs. Transience

A common topic in the study of random processes is whether the process will ever return to its initial position. We consider SRW on $\mathbb{R}^d$ and define its first, second, ..., nth returning time to the origin to be

$$\tau_1 = \inf\{m \geq 1 : S_m = 0\},$$

$$\tau_n = \inf\{m > \tau_{n-1} : S_m = 0\}.$$

Theorem. The following are equivalent:

(i) $P(\tau_1 < \infty) = 1$.
(ii) $P(\tau_n < \infty) = 1, \ \forall\, n = 1, 2, 3, \ldots$.
(iii) $P(S_m = 0 \ i.o.) = 1$.
(iv) $\sum_{m=1}^{\infty} P(S_m = 0) = \infty$.

Proof. We have

$$\begin{aligned}
P(\tau_2 < \infty) =& P(\tau_1 < \infty, \tau_2 - \tau_1 < \infty) \\
=& \sum_{m,n=1}^{\infty} P(\tau_1 = m, \tau_2 - \tau_1 = n) \\
=& \sum_{m,n=1}^{\infty} P(X_1 + \cdots + X_m = 0, X_1 + \cdots + X_u \neq 0, \forall\, 1 \leq u < m; \\
& \qquad X_{m+1} + \cdots + X_{m+n} = 0, X_{m+1} + \cdots + X_{m+v} \neq 0, \forall\, 1 \leq v < n) \\
=& \sum_{m,n=1}^{\infty} P(\tau_1 = m) P(\tau_1 = n) \qquad \text{(by i.i.d. assumption)} \\
=& (P(\tau_1 < \infty))^2.
\end{aligned}$$

Similarly, we can prove

$$P(\tau_n < \infty) = (P(\tau_1 < \infty))^n. \tag{2.1}$$

Therefore, (i) and (ii) are equivalent. They are equivalent to (iii) by examining their meanings. Finally, we have

$$
\begin{aligned}
\sum_{m=0}^{\infty} P(S_m = 0) = \sum_{m=0}^{\infty} E 1_{\{S_m=0\}} = E \sum_{m=0}^{\infty} 1_{\{S_m=0\}} \\
&= 1 + E \sum_{n=0}^{\infty} 1_{\{\tau_n < \infty\}} \quad \text{(both count the number of times the SRW is at 0)} \\
&= 1 + \sum_{n=0}^{\infty} P(\tau_n < \infty) \\
&= \sum_{n=0}^{\infty} (P(\tau_1 < \infty))^n. \qquad \text{(by (2.1))}
\end{aligned}
$$

This shows (iv) is equivalent to the others by considering the convergence of the geometric series. □

Definition. If $P(\tau_1 < \infty) = 1$, then the random walk is called **recurrent**.

If $P(\tau_1 < \infty) < 1$, then the random walk is called **transient**.

Theorem 4.2.3. SRW is recurrent in $\mathbb{R}^1$ and $\mathbb{R}^2$ and is transient in $\mathbb{R}^d, d \geq 3$.

Proof. In $\mathbb{R}^1$,

$$
\begin{aligned}
\sum_{m=1}^{\infty} P(S_m = 0) &= \sum_{n=1}^{\infty} P(S_{2n} = 0) && \text{(can only return to 0 at even steps)} \\
&= \sum_{n=1}^{\infty} \binom{2n}{n} (\frac{1}{2})^{2n} && \text{(combinatorics)} \\
&\sim \sum_{n=1}^{\infty} \frac{\sqrt{2\pi 2n}(\frac{2n}{e})^{2n}}{(\sqrt{2\pi n}(\frac{n}{e})^n)^2} \frac{1}{2^{2n}} && \text{(Stirling's formula)} \\
&= \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}} = \infty;
\end{aligned}
$$

hence recurrent by the previous theorem. By $\mathbb{R}^2$, similar calculation yields

$$P(S_{2n} = 0) \asymp \frac{1}{n},$$

still sum to infinity; hence still recurrent.

In $\mathbb{R}^3$, more complicated combinatorics give $P(S_{2n} = 0) \asymp \frac{1}{n^{3/2}}$ summing up to a finite number; hence transient. In even higher dimensions, the probabilities become even smaller; hence all transient. □

For $a, b \in \mathbb{Z}^d$, we define the first time a random walk starting from a reaches b

$$\tau_{a\to b} := \inf\{m \geq 1 : a + S_m = b\}.$$

It can be proved that

$$P(\tau_1 < \infty) = 1 \text{ if and only if } P(\tau_{a\to b} < \infty) = 1, \ \forall \ a, b.$$

Therefore, a drunk man will eventually find his way home, but a drunk bird may get lost forever.

3 Reflection Principle and Arcsine Distribution

We consider the SRW in $\mathbb{R}^1$.

Theorem (Reflection Principle). Let $X_1, X_2, \ldots$ be i.i.d. with $P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}$, $S_n = X_1 + \cdots + X_n$. For any positive integer b, we have

$$P(\max_{1\leq k\leq n} S_k \geq b) = 2P(S_n > b) + P(S_n = b).$$

Proof. We have

$$\begin{aligned} &P(\max_{1\leq k\leq n} S_k \geq b) \\ =&P(\max_{1\leq k\leq n} S_k \geq b, S_n > b) + P(\max_{1\leq k\leq n} S_k \geq b, S_n < b) + P(\max_{1\leq k\leq n} S_k \geq b, S_n = b). \end{aligned}$$

The first two terms are equal by reflecting the random walk trajectory along the horizontal line $y = b$ after it first hits level b. They both equal $P(S_n > b)$. The third term is simply $P(S_n = b)$. This gives the result. $\square$

Arcsine distribution is a continuous distribution on $(0, 1)$ with density $\frac{1}{\pi\sqrt{x(1-x)}}$, $x \in (0, 1)$. For the SRW in $\mathbb{R}^1$, define

$$L_{2n} := \sup\{m \leq 2n : S_m = 0\}, \quad \text{(last time at 0)}$$

$$F_n := \inf\{0 \leq m \leq n : S_m = \max_{0\leq k\leq n} S_k\}, \quad \text{(first time at maximum)}$$

π_{2n} := number of $k : 1 \leq k \leq 2n$ such that the line $(k-1, S_{k-1}) \to (k, S_k)$ is above the x-axis.

Theorem. $\frac{L_{2n}}{2n}, \frac{F_n}{n}, \frac{\pi_{2n}}{2n}$ all converge in distribution to the arcsine distribution.

Proof, not discussed in this course, can be found in An Introduction to Probability Theory and its Applications, Volume II, by William Feller.