Topics in Applied Stochastic Processes

Siva Athreya

Spring 2023

Contents

1 Finite length random walks on $\mathbb{Z}$ **2**

1.1 Definitions . . . 2

1.2 Stopping times . . . 6

1.3 Exercises . . . 8

2 More on random walks **9**

2.1 Reflection Principle . . . 9

2.2 Arc-Sine Law . . . 11

2.3 SRW of length N in $\mathbb{Z}^d$ . . . 12

2.3.1 Notations and notions in higher dimension . . . 12

2.3.2 Infinite length random walk . . . 12

2.3.3 Speed of the walk . . . 14

2.3.4 Typical position of the walk . . . 14

2.3.5 Large deviation principle . . . 14

2.4 Exercises . . . 15

3 Random Walks on Graphs **16**

3.1 Introduction . . . 16

3.2 Random Walks on Weighted Graphs . . . 19

3.3 Exercises . . . 22

4 Discrete Time Martingales **23**

Week 1 **January 13, 2023**

Finite length random walks on $\mathbb{Z}$

LECTURER: SIVA ATHREYA SCRIBE: SRIVATSA B, VENKAT TRIVIKRAM†

1.1 Definitions

Random walks serve as very useful models in many applications. They are simple to state and understand, yet they lead to lots of intractable questions.

Notation. $\mathbb{N} = \{k \in \mathbb{Z} : k \geq 1\}$ *and* $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$

We now proceed to construct what is called a "simple random walk" on $\mathbb{Z}$ of finite length $N \in \mathbb{N}$. The sample space Ω_N and the event space $\mathcal{F}_N$ are described below.

$$\Omega_N := \{(\omega_1, \dots, \omega_N) : \omega_i \in \{-1, 1\} \ \forall \ 1 \leq i \leq N\}$$

$$\mathcal{F}_N := \{A : A \subseteq \Omega_N\}$$

The probability function $\mathbf{P}_N : \Omega_N \to [0, 1]$ is defined as

$$\mathbf{P}_N(A) := |A| \ 2^{-N}$$

We also define random variables X_k and S_k on Ω_N for $1 \leq k \leq N$ as

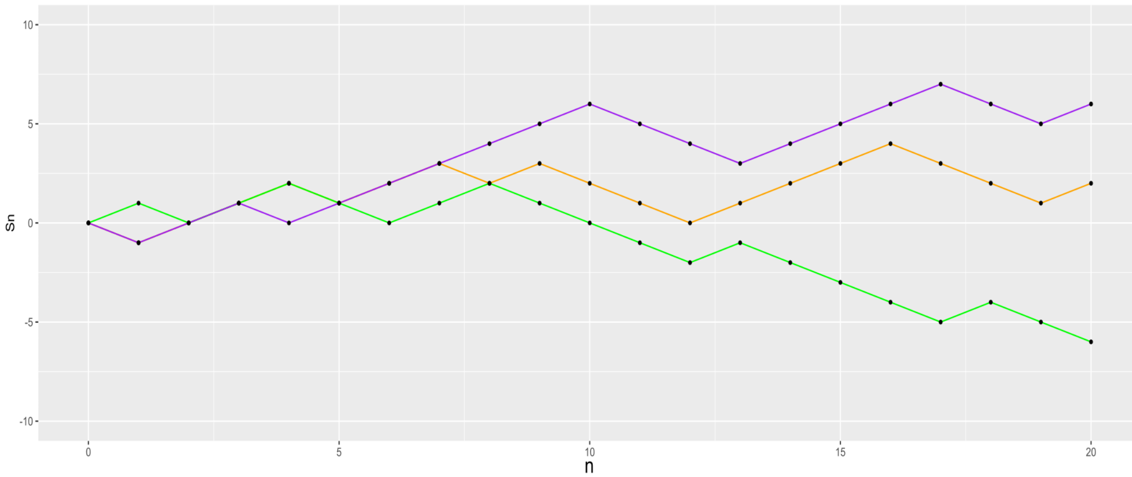
$$X_k : \Omega_N \to \{-1, 1\} \ ; \ X_k(\omega) := \omega_k$$

$$S_k : \Omega_N \to \mathbb{Z} \ ; \ S_k(\omega) := \sum_{i=1}^{k} X_k(\omega) \ ; \ S_0(\omega) := 0 \text{ for all } \omega \in \Omega_N$$

Definition 1.1.1. *Fix $N \in \mathbb{N}$. The sequence of random variables $\{S_k\}_{k=1}^{N}$ on $(\Omega_N, \mathcal{F}_N, \mathbf{P}_N)$ is called a (symmetric) simple random walk on $\mathbb{Z}$, of finite length N, starting at 0.*

† added illustrations

Figure 1.1: Three possible trajectories for $(S_n)_{n=0}^N$



In what follows, we suppress the subscript N while referring to the probability space $(\Omega_N, \mathcal{F}_N, \mathbf{P}_N)$, and we assume that $N \in \mathbb{N}$ is fixed.

Observations.

(a) $\{X_k\}_{k=1}^N$ are iid, i.e. independent and identically distributed.

Proof.

$$\begin{aligned} \mathbf{P}(X_k = 1) = \mathbf{P}(\{\omega \in \Omega : \omega_k = 1\}) &= 2^{-N} |\{\omega \in \Omega : \omega_k = 1\}| \\ &= 2^{-N} 2^{N-1} \\ &= \frac{1}{2} \\ &= \mathbf{P}(X_k = -1) \end{aligned}$$

So $\{X_k\}_{k=1}^N$ are identically distributed. Independence is left as an exercise. □

(b) (Independent increments) For $1 \leq k_1 \leq k_2 \leq \ldots \leq N$, $\{S_{k_i} - S_{k_{i-1}} : 1 \leq i \leq N\}$ are independent random variables.

Proof. Observe that, for $1 \leq k < l \leq N$, we have $S_l - S_k = \sum_{i=k+1}^{l} X_i$. Therefore, if $1 \leq a < b \leq c < d \leq N$, we see that $S_b - S_a$ and $S_d - S_c$ are functions of disjoint sets of independent random variables, and hence the claim is true. □

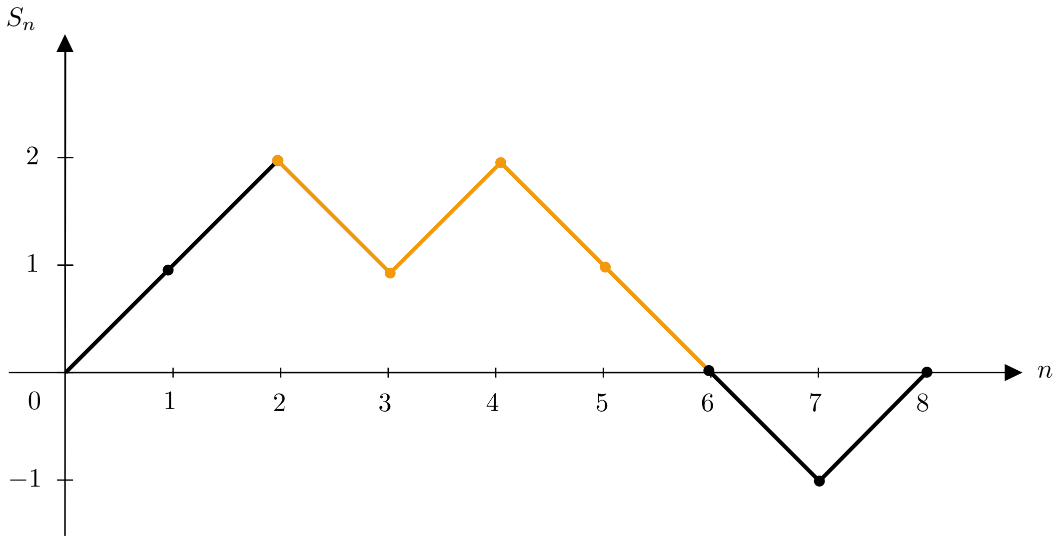


Figure 1.2: Independent (colored) increments in a simple random walk

(c) (Stationary in increments) For $1 \leq k < m \leq N$, $\mathbf{P}(S_m - S_k = \alpha) = \mathbf{P}(S_{m-k} = \alpha)$ for every $\alpha \in \mathbb{Z}$.

Proof. We use the fact that $\{X_i\}_{i=1}^N$ are identically distributed in the following argument.

$$\mathbf{P}(S_m - S_k = \alpha) = \mathbf{P}\left(\sum_{i=k+1}^{m} X_i = \alpha\right) = \mathbf{P}\left(\sum_{i=1}^{m-k} X_i = \alpha\right) = \mathbf{P}(S_{m-k} = \alpha)$$

□

(d) (Markov Property) For $\alpha_i \in \mathbb{Z},\ 1 \leq i \leq N$ and $0 \leq n \leq N$,

$$\mathbf{P}(S_n = \alpha_n \mid S_{n-1} = \alpha_{n-1}, \ldots, S_1 = \alpha_1) = \mathbf{P}(S_n = \alpha_n \mid S_{n-1} = \alpha_{n-1}),$$

assuming (of course) that the conditional probabilities are well defined.

Proof. Left as an exercise. □

(e) (Conditional Law) For $1 \leq k < m \leq N$, $\mathbf{P}(S_m = b \mid S_k = a) = \mathbf{P}(S_{m-k} = b - a)$.

Proof. Left as an exercise. □

(f) (Moments) For $1 \leq k \leq N$, we have $\mathbf{E}[X_k] = \mathbf{E}[S_k] = 0$ and $\mathrm{Var}[S_k] = k$.

Proof. By definition of expected value, $\mathbf{E}[X_k] = 1(1/2) - 1(1/2) = 0$. By linearity of expected values, $\mathbf{E}[S_k] = \sum_{i=1}^{k} \mathbf{E}[X_i] = 0$.

Since $\mathbf{E}[S_k] = 0$, $\mathrm{Var}[S_k] = \mathbf{E}[(\sum_{i=1}^{k} X_i)^2] = \sum_{i=1}^{k} \mathbf{E}[X_k^2] = k$. As an exercise, show that $\mathbf{E}[(\sum_{i=1}^{k} X_i)^2] = \sum_{i=1}^{k} \mathbf{E}[X_k^2]$. □

(g) (Distribution of S_n) For $x \in \{-n, -n+2, \ldots, n-2, n\}$, we have

$$\mathbf{P}(S_n = x) = \mathbf{P}(S_n = -x) = \binom{n}{\frac{n+x}{2}} 2^{-n}$$

Proof. We only provide a sketch of the proof, which is left as an exercise. For $0 \leq j \leq N$, $\{S_n = 2j - n\} = \{S_n = j - (n-j)\}$. So there must be a total of j steps to the right and $n - j$ steps to the left. Therefore

$$\mathbf{P}(S_n = 2j - n) = 2^{-N} \ |\{\omega \in \Omega : \cdots\}| = 2^{-n} \binom{n}{j}$$

□

(h) (Mode) The mode of the above distribution is achieved in the middle, i.e. at $x = 0$ and at $x = 1, -1$ for S_{2n} and S_{2n-1} respectively.

Proof.

$$\mathbf{P}(S_{2n} = 0) = \mathbf{P}(S_{2n-1} = 1) = \binom{2n}{n} 2^{-2n}$$

□

(i) (Stirling's formula) Using Stirling's approximation, for large n, we have

$$\binom{2n}{n} = \frac{2n!}{n!n!} \sim \frac{(2n)^{2n} e^{-2n} \sqrt{4\pi n}}{n^{2n} e^{-2n} \sqrt{2\pi n} \sqrt{2\pi n}} \sim \frac{2^{2n}}{\sqrt{\pi n}} \qquad (*)$$

Therefore,

$$\mathbf{P}(S_{2n} = 0) = \binom{2n}{n} \frac{1}{2^{2n}} \sim \frac{1}{\sqrt{\pi n}} \quad \text{as} \quad n \to \infty$$

This approximation, although correct, has a caveat - we chose to keep N fixed, but as $n \to \infty$, we must also let $N \to \infty$, and this requires subtler arguments. A few consequences of this approximation are mentioned in the exercises.

1.2 Stopping times

Motivation for this section comes from the classic Gambler's ruin problem. We can interpret a simple random walk as a fair game between two players, where in round k, a player wins the amount X_k. Then S_n denotes the capital of one player over the other after n rounds.

We would like to answer the following question - "Is it possible to stop the game in a favorite moment, i.e., can clever stopping lead to a positive expected gain?". In other words, can we design a $T(\omega)$ for every $\omega \in \Omega$ such that $\mathbf{E}[S_T] > 0$? Of course, the decision to stop may only depend on the trajectory until that time: no "insider knowledge" about the future of the trajectory is permitted.

To formalize this setting, we make the following definition.

Definition 1.2.1. *An event $A \subseteq \Omega$ is said to be observable by time n if it is a (possibly empty) union of basic / elementary events of the form*

$$\{\omega \in \Omega : \omega_1 = o_1, \ldots, \omega_n = o_n\}$$

where $o_i \in \{-1, 1\}$ for $1 \leq i \leq n$.

We also define $\mathcal{A}_0 = \{\phi, \Omega\}$ and set

$$\mathcal{A}_n := \{A \in \mathcal{F} : A \text{ is observable by time } n\}.$$

Immediately, we observe that

$$\{\phi, \Omega\} = \mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \ldots \subseteq \mathcal{A}_{N-1} \subseteq \mathcal{A}_N = \mathcal{F}$$

As an easy exercise, verify that each $\mathcal{A}_n$ is closed with respect to taking complement, union and intersection. Such a sequence $\{\mathcal{A}_i\}_{i=0}^N$ is called a *filtration.*

Definition 1.2.2. *A function $T : \Omega \to \{0, 1, \ldots, N\} \cup \{\infty\}$ is called a stopping time if for each $0 \leq n \leq N$,*

$$\{T = n\} = \{\omega \in \Omega : T(\omega) = n\} \in \mathcal{A}_n$$

Example. For $a \in \mathbb{Z}$, let $\sigma_a = \inf\{n : S_n = a, 0 \leq n \leq N\}$ denote the *first* hitting time of a. As an exercise, show that σ_a is a stopping time.

Example. For $a \in \mathbb{Z}$, let $L_a = \max\{n : S_n = a, 0 \leq n \leq N\}$ denote the *last* hitting time of a. As an exercise, show that L_a is NOT a stopping time.

Theorem 1.2.1. *Let $T : \Omega \to \{0, 1, \ldots, N\}$ be a stopping time. Then*

$$\mathbf{E}[S_T] = 0$$

where $S_T : \Omega \to \mathbb{Z}$ maps $\omega \mapsto S_{T(\omega)}(\omega)$.

Proof.

$$
\begin{aligned}
S_T = \sum_{k=1}^{N} S_k \mathbb{1}\{T = k\} &= \sum_{k=1}^{N} S_k(\mathbb{1}\{T \geq k\} - \mathbb{1}\{T \geq k+1\}) \\
&= \sum_{k=1}^{N} (S_k - S_{k-1})\mathbb{1}\{T \geq k\} \\
&= \sum_{k=1}^{N} X_k \mathbb{1}\{T \geq k\}
\end{aligned}
$$

where we take $\mathbb{1}\{T \geq N+1\} = 0$. Now, we can write $\mathbf{E}[S_T]$ as

$$
\mathbf{E}[S_T] = \sum_{k=1}^{N} \mathbf{E}[X_k \mathbb{1}\{T \geq k\}] \tag{$\dagger$}
$$

Observe that for $1 \leq k \leq N$, we have

$$
X_k \mathbb{1}\{T \geq k\} = \begin{cases} 1, & \text{for } X_k = 1,\ T \geq k \\ -1, & \text{for } X_k = -1,\ T \geq k \\ 0, & \text{otherwise.} \end{cases}
$$

$$
\mathbf{E}[X_k \mathbb{1}\{T \geq k\}] = \mathbf{P}(X_k = 1, T \geq k) - \mathbf{P}(X_k = -1, T \geq k) \tag{$\dagger\dagger$}
$$

Now,

$$
\{T \geq k\} = \{T < k\}^c = \left(\bigcup_{l=0}^{k-1} \{T = l\} \right)^c \in \mathcal{A}_{k-1}
$$

Using the fact that $\{T \geq k\} \in \mathcal{A}_{k-1}$, one can show that (details left as an exercise)

$$
\mathbf{P}(X_k = 1, T \geq k) = \mathbf{P}(X_k = -1, T \geq k) = \frac{1}{2}\mathbf{P}(T \geq k)
$$

Substituting the above values in ($\dagger$) and ($\dagger\dagger$), we finally have

$$
\mathbf{E}[S_T] = 0
$$

□

As an exercise, compute $\mathrm{Var}[S_T]$.

Definition 1.2.3. *A bet sequence / game system is a sequence of random variables $V_k : \Omega \to \mathbb{R}$ such that*

$$
\{V_k = c\} \in \mathcal{A}_{k-1} \textit{ for every } c \in \mathbb{R} \textit{ and } 1 \leq k \leq N
$$

Theorem 1.2.2. *Let $\{V_k\}_{k=1}^{N}$ be a bet sequence. Then*

$$
\mathbf{E}[S_N^V] = 0 \quad \textit{where} \quad S_N^V = \sum_{k=1}^{N} V_k X_k
$$

In this setting, S_N^V is interpreted as the "total gain".

Proof. Since Ω is finite, we may write

$$\text{Range}(V_k) = \{c_i^k : 1 \le i \le m_k\} \text{ where } c_i^k \in \mathbb{R}$$

$$V_k = \sum_{i=1}^{m_k} c_i^k \mathbb{1}\{V_k = c_i^k\}$$

Now, since $\mathbf{E}[X_k] = 0$, and since $X_k \perp \mathbb{1}\{V_k = c_i^k\}$, we get

$$\begin{aligned}
\mathbf{E}[S_N^V] = \sum_{k=1}^{N} \mathbf{E}[V_k X_k] &= \sum_{k=1}^{N} \mathbf{E}\left[X_k \sum_{i=1}^{m_k} c_i^k \mathbb{1}\{V_k = c_i^k\}\right] \\
&= \sum_{k=1}^{N} \sum_{i=1}^{m_k} c_i^k \mathbf{E}[X_k \mathbb{1}\{V_k = c_i^k\}] \\
&= \sum_{k=1}^{N} \sum_{i=1}^{m_k} c_i^k \mathbf{E}[X_k] \mathbf{P}(V_k = c_i^k) \\
&= 0
\end{aligned}$$

□

1.3 Exercises

1. Show that $\{X_k\}_{k=1}^N$ are independent.
2. Show that $\{S_n\}_{n=0}^N$ satisfies the Markov property.
3. For $1 \le k < m \le N$, show that $\mathbf{P}(S_m = b \mid S_k = a) = \mathbf{P}(S_{m-k} = b - a)$.
4. Show that $\mathbf{E}[S_n^2] = \sum_{i=1}^n \mathbf{E}[X_i^2]$.
5. (a) Show that for any $a, b \in \mathbb{R}$,

   $$\mathbf{P}(a \le S_n < b) \le (b - a)\ \mathbf{P}(S_n \in \{-1, 0, 1\}).$$

   (b) Using (a), conclude that

   $$\mathbf{P}(a \le S_n < b) \to 0 \quad \text{as} \quad n \to \infty.$$

   Thus, we observe that the walk exits any finite interval as $n \to \infty$.
6. Verify that each $\mathcal{A}_n$, $0 \le n \le N$, is closed with respect to taking complement, union and intersection.
7. For $a \in \mathbb{Z}$, let $\sigma_a = \inf\{n : S_n = a, 0 \le n \le N\}$. Show that σ_a is a stopping time.
8. For $a \in \mathbb{Z}$, let $L_a = \max\{n : S_n = a, 0 \le n \le N\}$. Show that L_a is not a stopping time.
9. Let $T : \Omega \to \{0, 1, \ldots, N\}$ be a stopping time. Compute $\text{Var}[S_T]$.
10. Show that X_k and $\mathbb{1}\{T \ge k\}$ are independent.

Week 2 **January 20, 2023**

More on random walks

LECTURER: SIVA ATHREYA SCRIBE: SANCHAYAN BHOWAL, VENKAT TRIVIKRAM

Theorem 2.0.1. *Let $T : \Omega \to 0, 1, \ldots, N$ be a stopping time. Then,*

$$\mathbf{E}[S_T^2] = E[T].$$

Proof.

$$\begin{aligned} S_T^2 &= \sum_{k=1}^{N} S_k^2 \mathbb{1}\{T = k\} \\ &= \sum_{k=1}^{N} (S_k^2 - S_{k-1}^2) \mathbb{1}\{T \geq k\} \\ &= \sum_{k=1}^{N} (X_k + S_{k-1})^2 - S_{k-1}^2 \mathbb{1}\{T \geq k\} \\ &= \sum_{k=1}^{N} (1 + 2X_k S_{k-1}) \mathbb{1}\{T \geq k\}. \end{aligned}$$

Now, consider $V_k = S_{k-1}\mathbb{1}\{T \geq k\}$. Note that this is a bet sequence. Hence,

$$\begin{aligned} \mathbf{E}[S_T^2] &= \mathbf{E}\left[\sum_{k=1}^{N} \mathbb{1}\{T \geq k\}\right] + 2\sum_{k=1}^{N} \mathbf{E}[X_k V_k] \\ &= \sum_{k=1}^{N} \mathbf{P}(T \geq k) + 0 \\ &= E[T]. \end{aligned}$$

□

2.1 Reflection Principle

Assume that $a \in \mathbb{Z}$ and $c > 0$. There is a bijection between the paths that cross $a + c$ and those that do not. This bijection is obtained by reflecting the part of the path crossing $a + c$ as shown in the Figure 2.1. So,

$$|S_n = a + c| = |\sigma_a \leq n \;\&\; S_n = a + c| = |\sigma_a \leq n \& S_n = a - c|$$

Now, we know that all the paths have equal probability. Hence, we get the following lemma.

Lemma 2.1.1. $\mathbf{P}(S_n = a + c) = \mathbf{P}(\sigma_a \leq n \;\&\; S_n = a - c)$ *where* $a \in \mathbb{Z}$ *and* $c > 0$*. This is also known as the* reflection principle*.*

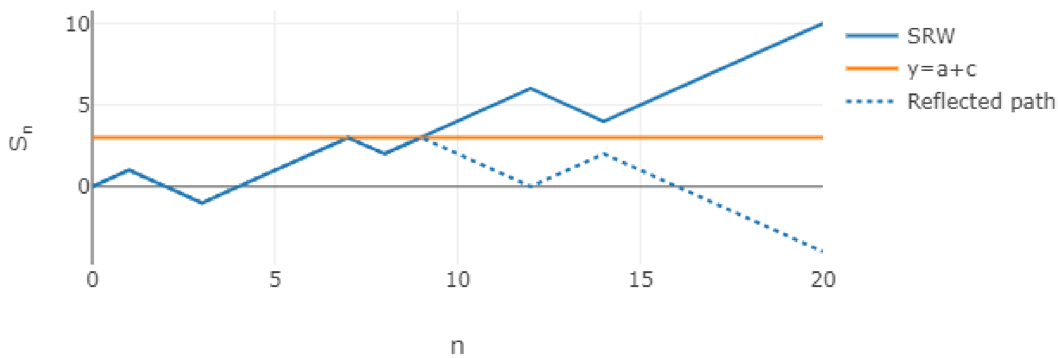


Figure 2.3: The figure shows that the bijection between the paths that cross a+c=3 and those that do not.

Theorem 2.1.1. $\mathbf{P}(\sigma_a \leq n) = \mathbf{P}(S_n \notin [-a, a))$ *where* $a \in \mathbb{Z}$
$\{0\}$*.*

Proof.

$$\begin{aligned}
\mathbf{P}(\sigma_a \leq n) &= \mathbf{P}(\sigma_a \leq n, \bigcup_{b \in \mathbb{Z}} S_n = b) \\
&= \sum_{b \in \mathbb{Z}} \mathbf{P}(\sigma_a \leq n, S_n = b) \\
&= \sum_{b \in \mathbb{Z}, b \geq a} \mathbf{P}(\sigma_a \leq n, S_n = b) + \sum_{b \in \mathbb{Z}, b < a} \mathbf{P}(\sigma_a \leq n, S_n = b) \\
&= \sum_{b \in \mathbb{Z}, b \geq a} \mathbf{P}(S_n = b) + \sum_{b \in \mathbb{Z}, b < a} \mathbf{P}(S_n = 2a - b) \\
&= \mathbf{P}(S_n \geq a) + \mathbf{P}(S_n > a) \\
&= \mathbf{P}(S_n \geq a) + \mathbf{P}(S_n < -a) \\
&= \mathbf{P}(S_n \notin [-a, a))
\end{aligned}$$

□

Corollary 2.1.1. $\mathbf{P}(\sigma_a = n) = \frac{1}{2}\left[\mathbf{P}(S_{n-1} = a - 1) - \mathbf{P}(S_{n-1} = a + 1)\right]$ *where* $a \in \mathbb{Z}$*.*

Proof. □

2.2 Arc-Sine Law

Let L denote the last time the random walk hits 0, i.e., $L = \max_{0 \leq n \leq 2N} S_n = 0$, where N denotes the length of the walk.

Theorem 2.2.1.

$$\mathbf{P}(L = 2n) = \frac{1}{2^{2N}} \binom{2n}{n} \binom{2N-2n}{N-n}.$$

Remark. By Stirling's approximation,

$$\mathbf{P}(L = 2n) \sim \frac{1}{\pi N} \frac{1}{\sqrt{\left(\frac{n}{N}\right)\left(1 - \frac{n}{N}\right)}}.$$

$$\begin{aligned}
\mathbf{P}\left(\frac{L}{2N} \leq x\right) &= \mathbf{P}(L \leq 2Nx) \\
&= \sum_{n=0}^{[2Nx]} \mathbf{P}(L = 2n) \\
&\sim \sum_{n=0}^{[2Nx]} \frac{1}{\pi N} \frac{1}{\sqrt{(x)(1-x)}} \\
&\sim \int_0^x \frac{dy}{pi\sqrt{y(1-y)}} \\
&= \frac{2}{\pi} \sin^{-1}(\sqrt{x}).
\end{aligned}$$

Proof of Theorem 2.2.1. Define $\tilde{\sigma_0} \inf\{n : S_n = 0, 0 < n \leq N\}$. Consider a path of length $2N$ with $L = 2n$. This path can be formed by a path which takes $S_2 n = 0$ and followed by a path of length $2N - 2n$ with $\sigma_0 > 2N - 2n$. Hence, number of paths of length $2N$ with $L = 2n$ is the product of the number of paths of length $2n$ with $S_{2n} = 0$ and the number of paths of length $2N - 2n$ with $\sigma_0 > 2N - 2n$. Hence,

$$\mathbf{P}(L = 2n) = \mathbf{P}(S_{2n} = 0)\mathbf{P}(\tilde{\sigma_0} > 2N - 2n), \tag{2.1}$$

Now let us compute the distribution of $\tilde{\sigma_0}$.

$$\begin{aligned}
\mathbf{P}(\tilde{\sigma_0} > 2k) &= \mathbf{P}(S_1 \neq 0, \ldots, S_{2k} \neq 0) \\
&= 2\mathbf{P}(S_1 > 0, \ldots, S_{2k} > 0) \\
&= \frac{2}{2^{2k}}\{\text{No. of paths start at 0 and stay above -1 for } 2k-1 \text{ steps}\} \\
&= \frac{2}{2^{2k}}\{\text{No. of paths start at 0 and stay below 1 for } 2k-1 \text{ steps}\} \qquad (2.2) \\
&= \mathbf{P}(\sigma_1 > 2k-1) \\
&= 1 - \mathbf{P}(\sigma_1 \geq 2k-1) \\
&= \mathbf{P}(S_{2k-1} = -1) + \mathbf{P}(S_{2k-1} = 0) \\
&= \mathbf{P}(S_{2k-1} = -1)
\end{aligned}$$

Using (2.1) and (2.2),

$$\begin{aligned}\mathbf{P}(L=2n) &= \mathbf{P}(S_{2n}=0)\mathbf{P}(S_{2N-2n-1}=-1)\\ &= \mathbf{P}(S_{2n}=0)\mathbf{P}(S_{2N-2n}=0)\\ &= \frac{1}{2^{2N}}\binom{2n}{n}\binom{2N-2n}{N-n}.\end{aligned}$$

The first step analysis of S_{2n} shows that, $\mathbf{P}(S_{2N-2n}=0) = \frac{1}{2}\mathbf{P}(S_{2N-2n-1}=1) + \frac{1}{2}\mathbf{P}(S_{2N-2n-1}=-1)$. Using the symmetry of the walk we know that $\mathbf{P}(S_{2N-2n-1}=1) = \mathbf{P}(S_{2N-2n-1}=-1)$. This gives the second inequality. $\square$

2.3 SRW of length N in $\mathbb{Z}^d$

2.3.1 Notations and notions in higher dimension

- $e_i \in \mathbb{Z}^d$, $\forall\, i \in \{1, 2, \cdots, d\}$, defined as the vector of length d with all entries zeroes except i^{th} being 1.

$$e_i = (0, 0, \cdots, \underbrace{1}_{i^{th}}, 0, \cdots, 0)$$

- For $x \in \mathbb{Z}^d$,

$$x = \sum_{i=1}^{d} x_i e_i,\ x_i \in \mathbb{Z} \qquad\qquad \|x\| = \left(\sum_{i=1}^{d} x_i^2\right)^{\frac{1}{2}}$$

- $\Omega_N = \{(\omega_1, \omega_2, \cdots, \omega_N) \mid \omega_i \in \mathbb{Z}^d, \|\omega_i\| = 1\, \forall\, 1 \leq i \leq N\}$
- We have, for $1 \leq k, n \leq N$

$$X_k : \Omega_N \to \mathbb{Z}^d,\ X_k(\omega) = \omega_k \qquad\qquad S_n : \Omega_N \to \mathbb{Z}^d,\ S_n(\omega) = \sum_{k=1}^{n} X_k(\omega)$$

  with $S_0(\omega) = 0$. We can consider S_n as a d-dimensional vector given by $S_n = \left(S_n^{(1)}, S_n^{(2)}, \cdots S_n^{(d)}\right)$, where each $S_n^{(i)}$ is a random walk on $\mathbb{Z}$.

- The probability function $\mathbf{P}^N$, given by,

$$\mathbf{P}^N : \mathcal{P}(\Omega_N) \to [0,1], \quad \mathbf{P}(A) = \frac{|A|}{(2d)^N}\, \forall\, A \subseteq \Omega_N$$

2.3.2 Infinite length random walk

On extending $N \to \infty$, we preserve something called as “consistency”. First, let us define, for $0 < N < M$,

$$\pi_N : \Omega_M \to \Omega_N,\ \ \pi_N(\omega_1, \omega_2, \cdots, \omega_M) = (\omega_1, \omega_2, \cdots, \omega_N)$$

Under $(\Omega_N, \mathcal{P}(\Omega_N), \mathbf{P}^N)$ and $(\Omega_M, \mathcal{P}(\Omega_M), \mathbf{P}^M)$, if we observe the walk till time $n < N$ the probability of evenets concerning the walk should be same under $\mathbf{P}^N$ or $\mathbf{P}^M$. For any event $\{\tilde{\omega} \in \Omega_N\}$, there exists a corresponding same event namely $\{\omega \in \Omega_M : \pi_N(\omega) = \tilde{\omega}\}$. We have,

$$\mathbf{P}^N(\{\tilde{\omega}\}) = \frac{1}{(2d)^N} \qquad \mathbf{P}^M(\{\omega \in \Omega_M : \pi_N(\omega) = \tilde{\omega}\}) = \frac{(2d)^{M-N}}{(2d)^M} = \frac{1}{(2d)^N}$$

So, we say the sequence of probability spaces $(\Omega_1, \mathbf{P}^1), (\Omega_2, \mathbf{P}^2), \cdots, (\Omega_N, \mathbf{P}^N)$ satisfies the consistency condition

$$\mathbf{P}^N(\{\tilde{\omega}\}) = \frac{1}{(2d)^N} = \frac{(2d)^{M-N}}{(2d)^M} = \mathbf{P}^M(\{\omega \in \Omega_M : \pi_N(\omega) = \tilde{\omega}\}),\ 0 < N < M,\ \tilde{\omega} \in \Omega_N$$

We define the space of infinite sequences,

$$\Omega_\infty = \{\omega = (\omega_k)k \geq 1 \mid \omega_k \in \mathbb{Z}^d,\ \|\omega_k\| = 1\}$$

$$\mathcal{A}_\infty\ (\equiv \mathcal{P}(\Omega_\infty))\ \text{denotes the class of events observable “for ever”}$$

For $N \in \mathbb{N}$,

$$\pi_N : \Omega_\infty \to \Omega_N,\ \ \pi_N(\omega) = (\omega_1, \omega_2, \cdots, \omega_N)$$

Theorem 2.3.1 (Kolmogorov Consistency Theorem). There exists a unique probability measure on $(\Omega_\infty, \mathcal{A}_\infty)$ such that $\forall\, N \geq 1$, $\forall\, \tilde{\omega} \in \Omega_N$,

$$\mathbf{P}^N(\{\tilde{\omega}\}) = \mathbf{P}^M(\{\omega \in \Omega_M : \pi_N(\omega) = \tilde{\omega}\}) = \frac{1}{(2d)^N}$$

Now, we can define,

$$X_k : \Omega_\infty \to \mathbb{Z}^d,\ \ X_k(\omega) = \omega_k \qquad S_n = \sum_{k=1}^{n} X_k\ \forall\, n \geq 1$$

under $\mathbf{P}$, $\{S_n\}_{n\geq 1}$ is a simple random walk starting at $S_0 = 0$.

Definition 2.3.1. $A \subseteq \Omega_\infty$ is said to be **observable** by time n if A is a union of the events of the form

$$\{\omega \in \Omega_\infty : \omega_i = o_i,\ 1 \leq i \leq N\} \text{ with } o_i \in \mathbb{Z}^d,\ \|o_i\| = 1$$

For, $k \in \mathbb{N}_0$, $\mathcal{A}_k$ denotes the set of all events in Ω_∞ observable by time k.

Definition 2.3.2. $T : \Omega_\infty \to \mathbb{N} \cup \{\infty\} \cup \{0\}$ is a **stopping time** if

$$\text{for any }\ k \in \mathbb{N}_0,\ \{T = k\} \in \mathcal{A}_k$$

For example, $\sigma_a = \min\{n \geq 0 \mid S_n = a\}$ is a stopping time.

2.3.3 Speed of the walk

Definition 2.3.3. For, $S_n = \sum_{k=1}^{n} X_k$, we define **speed of the walk** as

$$\text{Speed} \ = \ \frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

We have, $X_k = \left(X_k^{(1)}, X_k^{(2)}, \cdots, X_k^{(d)}\right)$, $\{X_k\}_{k\geq 1}$ which is an i.i.d sequence of random variables with

$$\mathbf{P}(X_k = e_i) = \frac{1}{2d} = \mathbf{P}(X_k = -e_i)$$

$\Rightarrow$ $\mathbf{E}[X_k] = 0$ and $\mathbf{E}[\|X_k\|] = 1$ $(\leq \infty)$

Theorem 2.3.2 (**Strong law of large numbers**)**.** For simple random walk on $\mathbb{Z}^d$,

$$\frac{S_n}{n} \to 0 \text{ with probability 1 under } (\Omega_\infty, \mathcal{A}_\infty, \mathbf{P})$$

2.3.4 Typical position of the walk

For $d = 1$,

$$\frac{S_n - (n)(0)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1)$$
$$\Rightarrow \ \sqrt{n}\left(\frac{S_n}{n}\right) \xrightarrow{d} \mathcal{N}(0,1)$$

For $d > 1$, $\mu \in \mathbb{R}^d$ and a positive definite matrix $\Sigma \in \mathbb{R}^{d\times d}$, we have d-dimensional normal distribution as,

$$\Phi_{d,\mu,\Sigma}(y) = \frac{1}{(2\pi)^{d/2}}\frac{1}{\det(\Sigma)^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^t\Sigma^{-1}(x-\mu)\right)$$

$$\mathbf{P}\left(\frac{S_n}{\sqrt{n}} \in \prod_{i=1}^{d}[a_i, b_i]\right) \xrightarrow[n\to\infty]{} \int\limits_{\prod_{i=1}^{d}[a_i,b_i]} \Phi_{d,0,\Sigma^d}(y)\,dy$$

where, $\mu = 0$, $\Sigma^d = \text{diag}\,(\frac{1}{d}, \cdots, \frac{1}{d})$

2.3.5 Large deviation principle

From the CLT, we have that

$$\mathbf{P}(\|S_n\| > a\sqrt{n}) \xrightarrow[n\to\infty]{} \int\limits_{\|x\|>a} \Phi_{d,0,\Sigma^d}(y)\,dy$$

We consider the events of the form $\{\|S_n\| > an\}$, $a \in [0, \infty)$, which are "rare" in the sense that their probabiility tends to 0 as $n \to \infty$. On formal application of CLT shows that probability of these rare events are exponentially small.

Theorem 2.3.3 (Cramer's theorem). For, $a > 0$,

$$\lim_{n\to\infty} \frac{\log(\mathbf{P}(\|S_n\| > an))}{n} = -I(a)$$

where,

$$I(a) = \begin{cases} \log 2 + \frac{1+a}{2}\log\frac{1+a}{2} + \frac{1-a}{2}\log\frac{1-a}{2}, & \text{for } a \in [-1,1] \\ \infty, & \text{otherwise} \end{cases}$$

It can be vaguely interpreted as, $\mathbf{P}(\|S_n\| > na) \sim e^{-nI(a)}$

2.4 Exercises

1. Complete the proof of Reflection Principle (Lemma 2.1.1).
2. Find the distribution of $M_k = \max_{1\leq k\leq n} S_k$.
3. Show that $\mathbf{E}[\|X_k\|] = 1$.

Week 3 **January 27, 2023**

Random Walks on Graphs

LECTURER: SIVA ATHREYA SCRIBE: ABHITI MISHRA, DEVESH BAJAJ

3.1 Introduction

- A random walk on a graph is basically a reversible Markov chain on the graph.
- many results of random walks will hold true for general markov chains but we will not go into it
- we will study some of the geometric properties of the Graph which translate to different properties of the Random walks

$\Gamma = (V, E)$
$V \equiv$ Vertex set = finite or countably infinite set.
$E \equiv$ Edge set $= E \subset \mathcal{P}(V) = \{\{x, y\} : |x, y \in V, x \neq y\}$.
(No self loops, No multiple edges)

1. $x \in V; y \in V$ is a neighbour of x in $\{x, y\} \in E$ $(x \sim y)$
2. A path $\gamma \in \Gamma$ is any sequence $\{x_i\}_{i=0}^{n}$ such that $x_{i-1} \sim x_i$ in Γ for some $n \geq 1, x_i \in V, 1 \leq i \leq n$
   •γ is a loop if $x_0 = x_n$
   •γ is self avoiding if $x_i \neq x_j$ $\forall$ $i \neq j$.
3. "chemical metric" $d : V \times V \longrightarrow [0, \infty) \bigcup \{\infty\}$
   $d(x, x) = 0$,
   $$d(x, y) = \begin{cases} \text{length of smallest path from x to y} \\ \infty \text{ if no path exixts} \end{cases}$$
4. Γ is connected if $d(x, y) < \infty, \forall x, y \in V$ **(H1 property)**
5. Γ is locally finite if $\forall x \in V$,
   $N(x) = \{y \in V | y \sim x\} \Rightarrow |N(x)| < \infty$ **(H2 property)**
6. we say Γ has a bounded geometry if $\sup_{x \in V} |N(x)| < \infty$ **(H3 property)**

Definition 3.1.1. $\forall$ $x, y \in V$, *we assume that thre is a weight* μ_{xy} *such that:*

1. $\mu_{xy} = \mu_{yx}$
2. $\mu_{xy} \geq 0$
3. *if* $x \neq y$ *then,* $\mu_{xy} > 0 \Leftrightarrow x \sim y$

we will call (Γ, μ) *a weighted graph.*
Using property 3 above, $E = \{\{x, y\} | x, y \in V, \mu_{xy} > 0, x \neq y\}$

Definition 3.1.2. (Γ, μ) *has bounded weights if* $\exists\ C_1, C_2 > 0$ *such that* $C_1 < \mu_{xy} \leq C_2\ \forall\ x, y \in V, x \neq y$. *This is called the* ***(H4 Property)***.

Definition 3.1.3. (Γ, μ) *has controlled weights if* $\exists\ c > 0$ *such that* $\frac{\mu_{xy}}{\mu_x} \geq c^{-1}\ \forall\ x, y \in V, x \neq y$. *This is called the* ***(H5 Property)***.

Define for $x \in V$: $\mu_x = \sum_{y \sim x} \mu_{xy}$

Definition 3.1.4. *Natural weights:*

$$\mu_{xy} = \begin{cases} 1 \text{ if } x \sim y \\ 0 \text{ otherwise} \end{cases}$$

Lemma 3.1.1. *Suppose* (Γ, μ) *is a weighted graph then,*

1. *(H3), (H5) holds.*

2. $\forall x \in V, n > 0\ , B(x, n) = \{y \in V | d(x, y) \leq n\}$ *(balls are not exponentially large)*

3. $\forall x \in V, n \geq 0, \mu(B(x, n)) = \sum_{y \in B(x,n)} \mu_y \leq 2\mu_x(c_2)^n$ *(Balls have bounded weights)*

Proof. 1. Take $x \in V$.

$$\begin{aligned} N(x) &= c \sum_{y \in V} \frac{1}{c} 1_{\{x \sim y\}} \\ &\leq c \sum_{y \in V} \frac{\mu_{xy}}{\mu_x} 1_{\{x \sim y\}} \\ &= c \frac{1}{\mu_x} \sum_{y \in V} \mu_{xy} = c \end{aligned}$$

2. $S(x, n) = \{y \in V | d(x, y) = n\}$

$$|S(x, n)| \leq c |S(x, n-1)|\ \ \forall\ n \geq 1$$

Arguing inductively,

$$\begin{aligned}|B(x,n)| &= \sum_{k=0}^{n} |S(x,k)| \\ &\leq \sum_{k=0}^{n} c^k \\ &= \frac{c^{n+1}-1}{c-1} \leq 2c^n\end{aligned}$$

3. $n = 1$.

$$\begin{aligned}\mu(B(x,1)) &= \mu_x + \sum_{y\sim x} \mu_y \\ &\leq c\sum_{y\sim x} \mu_{xy} + \mu_x \\ &= c\mu_x + \mu_x\end{aligned}$$

Second step follows from the H5 assumption.
We also note

$$\mu(B(x,2)) = \sum_{y\in B(x,2)} \mu_y = \mu(B(x,1)) + \sum_{y\sim x}\sum_{z\sim y} \mu_z$$

Therefore

$$\begin{aligned}\mu(B(x,2)) &\leq \mu_x + c\mu_x + \sum_{y\sim x} c \sum_{z\sim y} \mu_{zy} \\ &= \mu_x + c\mu_x c \sum_{y\sim x} \mu_y \\ &\leq \mu_x + c\mu_x + c^2\mu_x\end{aligned}$$

$\square$

Example. $V = \mathbb{Z}^d$. Take $x, y \in V, |x-y| = \sum_{i=1}^{d} |x_i - y_i|$
$E = \{(x,y) | |x-y| = 1\}$. $\mu_{xy} = 1$ whenever $(x,y) \in E$. $N(x) = 2d \ \ \forall x \in V$
$|B(x,n)| \sim n^d \leq 2c^n \ \ \forall c \geq 2$.
Example. Rooted Binary Tree- Let the root be $B_0 = \{\rho\}$.
$\forall\ n \geq 1, B_n = \{0,1\}^n$

$$V = \cup_{n=1}^{\infty} B_n \cup \{\rho\}$$

For $x \in B_n, n \geq 2, x = (x_1, \ldots, x_n), x_i \in \{0,1\}$.
Let the parent of x be- $\alpha(x) = (x_1, \ldots, x_{n-1})$
For $n = 1, x \in B_1, \alpha(x) = \rho$

$$E = \{(x, \alpha(x)) |\ x \in V, x \notin B_0\}$$

$$|N(\rho)| = 2, |N(x)| = 3 \ \ \forall\ x \notin B_0$$

Canopy Tree

$$\bar{V} = \{x \in V | x = (x_1, \ldots, x_n) \text{ and } x_i = 0 \ \ \forall\, 1 \leq i \leq n \text{ forsome } n \geq 1\} \cup \{\rho\}$$

$f(x)$ is the element in $\bar{V}$ closest to x.
V_{canopy} is a subset of V such that-

$$V_{canopy} = \{x \in V |\ d(x, f(x)) \leq d(\rho, f(x))\}$$

Observe that in the canopy tree, there is only one self-avoiding path to infinity, but the size of the balls $B(\rho, n)$ still grows exponentially. It shows that one does not need too many paths to infinity for the size of your graph to grow exponentially. Denoted by $\mathbb{T}^2_{canopy}$

3.2 Random Walks on Weighted Graphs

(This section will be done as a discrete time reversible Markov Chain)
Formally, X_n jumps from $x \sim y_i$ with probability proportional to μ_{xy_i}. It stays at x with probability proportional to μ_{xx}.

Our graph is denoted by $\Gamma = (V, E)$. We assume there are no isolated edges that is $\{\mu_x \neq 0 \ \ \forall x \in V\}$. Also assume H(1) and H(2).

$$\Omega = \{f : \mathbb{N} \cup \{0\} \to V\} \equiv V^{\mathbb{N} \cup \{0\}}$$

$\forall\ n \geq 0$, $X_n : \Omega \to V$ where $X_n(\omega) = \omega(n)$
Let $\mathcal{A}_n \equiv$ observable events upto time n (all events that can be derived from $X_1, \ldots, X_n$). This will be a filtration.

$$\mathcal{F} \equiv \cup_{n \geq 1}\ \mathcal{A}_n$$

Set $\mathcal{P}(x, y) = \frac{\mu_{xy}}{\mu_x}\ \ \forall x, y \in V$.
$\forall x\ \in V$, there exists a unique $\mathcal{P}^x(.)$ on $(\Omega, \mathcal{F})$.
(Existence can be shown using Kolmogorov consistency theorem).
$\forall\ n \geq 1$

$$\mathbb{P}^x(X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = 1_{\{x\}}(x_0) \prod_{i=1}^{n} P(x_n, x_{n-1})$$

$$\begin{aligned}
\mathbb{P}^x(X_1 = y) &= \mathbb{P}^x(X_1 = y, \cup_{z \in V} X_0 = z) \\
&= \sum_{z \in V} \mathbb{P}^x(X_1 = y, X_0 = z) \\
&= \sum_{z \in V} \mathcal{P}(y, z) 1_{\{x\}}(z) \\
&= \mathcal{P}(y, x)
\end{aligned}$$

One-step transition probability-

$$\mathbb{P}(X_n = y | X_{n-1} = z) = \frac{\mathbb{P}(X_n = y, X_{n-1} = z)}{\mathbb{P}(X_{n-1} = z)} = \mathcal{P}(y, z)$$

The last equality is left as an exercise.
Reversibility-

$$\mu_x \mathcal{P}(x, y) = \mu_x \frac{\mu_{xy}}{\mu_x} = \mu yx = \mu_y \mathcal{P}(y, x)$$

$(X_n, \mathcal{P})$ markov chain is symmetric with repsect to $\{\mu_x\}_{x \in V}$

Lemma 3.2.1. *Let $x_0, \ldots, x_n \in V$*

$$\mu_{x_0} \mathbb{P}^{x_0}(X_n = x_n, \ldots, X_0 = x_0) = \mu_{x_n} \mathbb{P}^{x_n}(X_n = x_0, \ldots, X_0 = x_n)$$

The above shows the reversibility of the markov chain wrt μ.

Proof.

$$\begin{aligned}
\mu_{x_0} \mathbb{P}^{x_0}(X_n = x_n, \ldots, X_0 = x_0) &= \mu_{x_0} \prod_{i=1}^{n} \mathcal{P}(x_i, x_{i-1}) \\
&= \mu_{x_0} \prod_{i=1}^{n} \frac{\mu_{x_i, x_{i-1}}}{\mu_{x_{i-1}}} \\
&= \mu_{x_n} \prod_{i=1}^{n} \frac{\mu_{x_{n-i}, x_{n-i+1}}}{\mu_{x_{n-i+1}}} \\
&= \mu_{x_n} \mathbb{P}^{x_n}(X_n = x_0, \ldots, X_0 = x_n)
\end{aligned}$$

□

Remark. If $\mu(V) = \sum_{x \in V} \mu_x = 1$ and $\mu(A) = \sum_{x \in A}$, then μ is the reversible distribution for $\{X_n\}_{n \geq 0}$ that is

$$\mu_x \mathcal{P}(x, y) = \mu_y \mathcal{P}(y, x)$$

Hence $\{\mu_x\}_{x \in V}$ is the stationary distribution.

Definition 3.2.1. *$A \subseteq V$. The hitting time of A be given by*

$$T_A = \min\{n \geq 0 | X_n \in A\}$$

By convention, $T_A = \infty$ iff X_n does not visit A.

Definition 3.2.2. *The return time of A is defined as -*

$$T_A^+ = \min\{n \geq 1 | X_n \in A\}$$

Note that $X_0 \notin A \implies T_A^+ = T_A$

Definition 3.2.3. *The exit time of A is-*

$$\tau_A = T_{A^c}$$

Theorem 3.2.1. *Let Γ be H(1) and H(2) and $|V| = \infty$. Then TFAE-*

1. $\exists\ x \in V$ *such that* $\mathbb{P}^x(\tau_x^+ < \infty) < 1$

2. $\forall\ x \in V$, $\mathbb{P}^x(\tau_x^+ < \infty) < 1$

3. $\forall\ x \in V$, $\sum_{n=0}^{\infty} \mathbb{P}^x(X_n = x) < \infty$

4. $\forall\ x, y \in V$, $\mathbb{P}^x(\tau_y < \infty) < 1$

5. $\mathbb{P}^x(\sum_{n\geq 0} 1_{\{X_n = x\}} < \infty) = 1 \quad \forall x, y \in V$

If the above is satisfied, the Markov Chain is transient.

Theorem 3.2.2. *Let Γ be H(1) and H(2) and $|V| = \infty$. Then TFAE-*

1. $\exists\ x \in V$ *such that* $\mathbb{P}^x(\tau_x^+ < \infty) = 1$

2. $\forall\ x \in V$, $\mathbb{P}^x(\tau_x^+ < \infty) = 1$

3. $\forall\ x \in V$, $\sum_{n=0}^{\infty} \mathbb{P}^x(X_n = x) = \infty$

4. $\forall\ x, y \in V$, $\mathbb{P}^x(\tau_y < \infty) = 1$

5. $\mathbb{P}^x(\sum_{n\geq 0} 1_{\{X_n = x\}} = \infty) = 1 \quad \forall x, y \in V$

If the above is satisfied, the Markov Chain is recurrent.

Definition 3.2.4. *If $\{X_n\}_{n\geq 0}$ random walk on (Γ, μ) satisfies*

1. *any statement of theorem 1.6, the graph (Γ, μ) is transient.*

2. *any statement of theorem 1.7, the graph (Γ, μ) is recurrent.*

3.3 Exercises

1. Show that $H_3, H_4 \Rightarrow H_5$

2. When is (Γ, μ) transient or recurrent?
   Partial answer- When $|V| < \infty$, (Γ, μ) is recurrent.

3. **Kesten Problem-** G is a finitely generated group with generating set A. Look at the Cayley graph of G. Which groups provide transient graphs?

Week 4 **January 27, 2023**

Discrete Time Martingales

LECTURER: SIVA ATHREYA SCRIBE: ABHITI MISHRA, DEVESH BAJAJ

Origin is from horse-racing (betting system). The dictionary meaning of the word 'martingale' is the harness of a horse.
Let $\{Z_n\}_{n\geq 1}$ is a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 4.0.1. *A sequence of random variables $\{Z_n\}_{n\geq 1}$ is said to be a Martingale if*

$$\mathbb{E}(Z_n | Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1) = z_{n-1} \quad \forall\ n \geq 2 \tag{4.3}$$

Things to understand- conditional expectation for discrete and conditional r.v. Reference- Ch6 of Siva's book.
Things we will explore-

1. Examples of $\{Z_n\}_{n\geq 1}$ that are martingales.

2. How different are martingales from iid sequences and markov chains?

3. How to interpret 4.3?

Example. $\{S_n\}_{n\geq 1}$ and $S_0 \equiv 0$.

$$X_i = \begin{cases} 1, & w.p\ \ 1/2 \\ -1, & w.p\ \ 1/2 \end{cases}$$

$$S_n = \sum_{i=1}^{n} X_i$$

Let $s_{n-1}, s_{n-2}, \ldots, s_1 \in \mathbb{Z}$ such that $\mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1) > 0$

$$\begin{aligned}
\mathbb{E}(S_n | S_{n-1} = s_{n-1}, \ldots, S_1 = s_1) &= \sum_{k \in \mathbb{Z}} k \mathbb{P}(S_n = k | S_{n-1} = s_{n-1}, \ldots, S_1 = s_1) \\
&= \sum_{k \in \mathbb{Z}} k \frac{\mathbb{P}(S_n = k, S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)}{\mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)} \\
&= \sum_{k \in \mathbb{Z}} k \frac{\mathbb{P}(S_{n-1} + X_n = k, S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)}{\mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)} \\
&= \sum_{k \in \mathbb{Z}} k \frac{\mathbb{P}(X_n = k - s_{n-1}, S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)}{\mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)} \\
&= \sum_{k \in \mathbb{Z}} k \frac{\mathbb{P}(X_n = k - s_{n-1}) \mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)}{\mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)} \\
&= (s_{n-1} + 1) \mathbb{P}(X_n = -1) + (s_{n-1} - 1) \mathbb{P}(X_n = 1) \\
&= (s_{n-1} + 1) \frac{1}{2} + (s_{n-1} - 1) \frac{1}{2} = s_{n-1}
\end{aligned}$$

Note that the summations here are "finite" sums.
As $s_{n-1}, \ldots, s_1 \in \mathbb{Z}$ were arbitrary, $\{S_n\}_{n \geq 1}$ is a martingale.

Example. $\{X_i\}_{i \geq 1}$ be an iid sequence on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $Z_n = \prod_{i=1}^n X_i$ and $\text{Range}(Z_n) \subset \mathbb{R} \ \forall \ n \geq 1$.
Let $z_{n-1}, \ldots, z_1 \in \mathbb{R}$ such that $\mathbb{P}(Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1) > 0$. Then

$$\begin{aligned}
\mathbb{E}(Z_n | Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1) &= \sum_{k \in Range(Z_n)} k \mathbb{P}(Z_n = k | Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1) \\
&= \sum_{k \in Range(Z_n)} k \frac{\mathbb{P}(Z_n = k, Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)}{\mathbb{P}(Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)} \\
&= \sum_{k \in Range(Z_n)} k \frac{\mathbb{P}(Z_{n-1} X_n = k, Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)}{\mathbb{P}(Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)} \\
&= \sum_{k \in Range(Z_n)} k \frac{\mathbb{P}(z_{n-1} X_n = k, Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)}{\mathbb{P}(Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)} \\
&= \sum_{k \in Range(Z_n)} k \mathbb{P}(Z_{n-1} X_n = k) \frac{\mathbb{P}(Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)}{\mathbb{P}(Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1)} \\
&= \sum_{u \in S^1, Range(X_n) = S^1} u z_{n-1} \mathbb{P}(X_n = u) \\
&= z_{n-1} \mathbb{E}[X_n] = z_{n-1}
\end{aligned}$$

Note that the sums here might be infinite. In the last step we assume $\mathbb{E}[X_i] = 1$. Now since $\{z_i\}_{i=1}^{n-1}$ were arbitrary, $\{Z_n\}_{n\geq 1}$ is a martingale.

Example.

$$X_i = \begin{cases} 2, & w.p \;\; 1/2 \\ 0, & w.p \;\; 1/2 \end{cases}$$

Then $\mathbb{E}(X_i) = 1$. Therefore, $Z_n = \prod_{i=1}^{n} X_i$ is a martingale. Range (Z_n)= $\{2^n, 0\}$. Note that the mean stays constant and

$$\mathbb{P}(Z_n = 0) = 1 - \frac{1}{2^n}$$

$$\mathbb{P}(Z_n = 2^n) = \frac{1}{2^n}$$

Intuition- The first equation shows that the martingale takes a very low value with very high probability and the second one shows that it takes a very large value with very low probability
Idea behind Markov Chains -

$$\text{“}X_n | X_{n-1}, \ldots, X_1\text{”} \stackrel{d}{=} X_n | X_{n-1}$$

Idea behind Martingales - Expected value of Z_n conditioned on the past depends only on Z_{n-1}. $\{Z_n\}_{n\geq 1}$ in law could depend on the entire past!