Topics in Applied Stochastic Processes
Topics in Discrete Probability

Siva Athreya

Spring 2023

Preface

This compilation of lecture notes is an introduction to the course 'Topics in Applied Stochastic Processes and Topics in Discrete Probability' offered at ISI Bangalore for BMath and MMath students. It is based on the assumption that the reader has a basic pre-requisite of Probability Theory, and an exposure to Measure Theoretic Probability is desired but not mandatory.

The primary aim of this course is to explore and discuss a number of topics from modern probability theory that are centred around random walks. The lecture notes will provide students with the essential knowledge and understanding of the topics, as well as a platform to apply the same practically.

We hope that this compilation of lecture notes will be a useful resource to the students, and serve as a valuable reference material in the future. The main references of this course is [1, 2].

Contents

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.

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

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

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

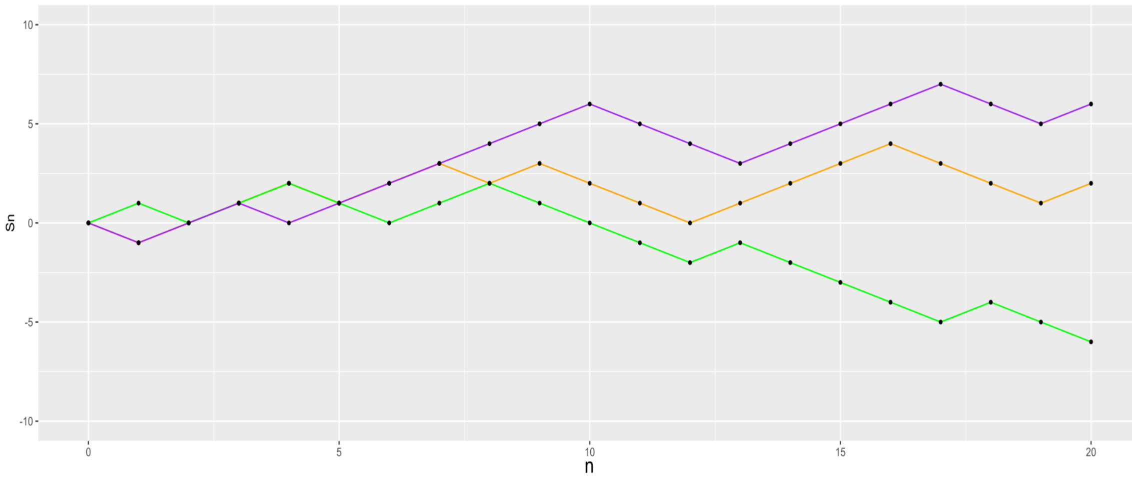
$$X_k : \Omega_N \to \{-1, 1\} \ ; \ X_k(\omega) := \omega_k$$

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

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. □

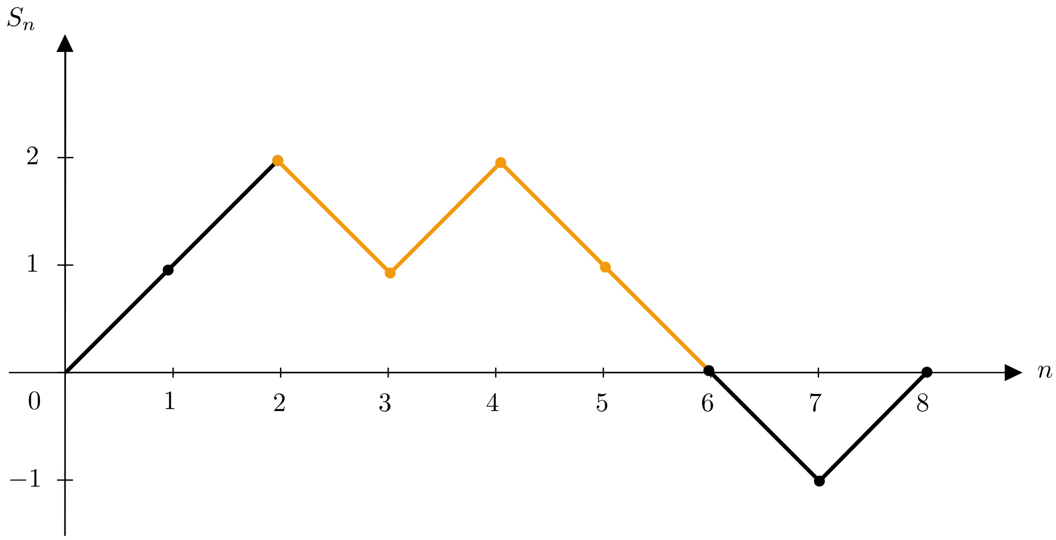


Figure 1.2: Independent (colored) increments in a simple random walk

(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 $x \in \{-n, -n+2, \dots, n-2, n\}$, 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 $0 \le j \le N$, $\{S_n = 2j - n\} = \{S_n = j - (n-j)\}$. 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

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

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

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

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

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

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

Now,

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

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

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

Substituting the above values in (†) and (††), 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*

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

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

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

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

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 $\mathrm{Var}[S_T]$.
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*.*

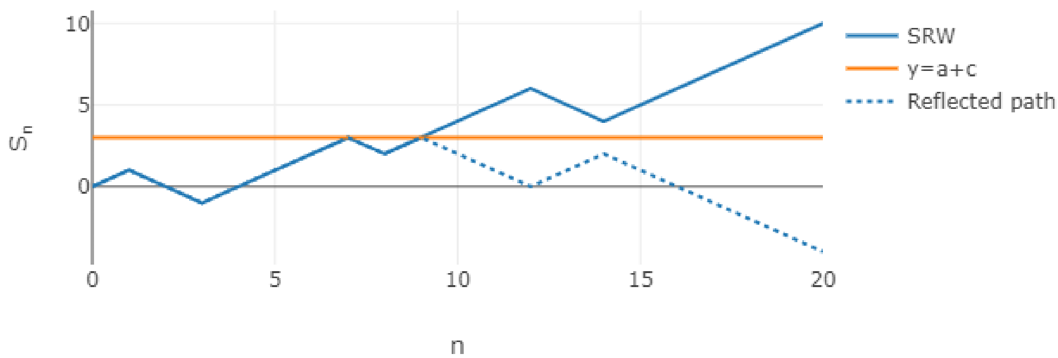


Figure 2.3: The figure shows that the bijection between the paths that cross a+c=3 and those that do not.

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., $L = \max_{0 \le n \le 2N} S_n = 0$, 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,

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

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

Proof of Theorem 2.2.1. Define $\tilde{\sigma_0} \inf\{n : S_n = 0, 0 < n \le N\}$. Consider a path of length $2N$ with $L = 2n$. This path can be formed by a path which takes $S_2n = 0$ 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 $\tilde{\sigma_0}$.

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

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

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

- $\Omega_N = \{(\omega_1,\omega_2,\cdots,\omega_N) \mid \omega_i \in \mathbb{Z}^d, \|\omega_i\| = 1\,\forall\, 1\le i\le N\}$

- We have, for $1 \le k,n \le N$

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

  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, $\mu = 0$, $\Sigma^d = \operatorname{diag}\left(\frac{1}{d}, \cdots, \frac{1}{d}\right)$

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

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

□

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 $(x,y) \in E$. $N(x) = 2d \quad \forall x \in V$
$|B(x,n)| \sim n^d \leq 2c^n \quad \forall c \geq 2$.
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 $x \in B_n, n \geq 2, x = (x_1, \dots, x_n), x_i \in \{0,1\}$.
Let the parent of x be- $\alpha(x) = (x_1, \dots, x_{n-1})$
For $n = 1, x \in B_1, \alpha(x) = \rho$

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

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

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 $x_0, \dots, x_n \in V$*

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

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

Proof.

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

□

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 **February 3, 2023**

Energy and Variational Methods

LECTURER: SIVA ATHREYA SCRIBE: ATREYA CHOUDHURY

4.1 Transition densities and Laplacian

We can recall that $\mathbf{P}(X_1 = y) = \mathcal{P}(x,y) = \frac{\mu_{xy}}{\mu_x}$.

Definition 4.1.1. *The transition density w.r.t weights μ of a random walk $\{X_n\}$ is given by:*

$$p_n(x,y) = \frac{\mathbf{P}^x(X_n = y)}{\mu_y} := \frac{\mathcal{P}_n(x,y)}{\mu_y} \qquad \mu \geq 1$$

$p_0(x,y) = \frac{1_{\{x\}}(y)}{\mu_y} = \frac{1_{\{x\}}(y)}{\mu_x}$.
$p(x,y) \equiv p_1(x,y) = \frac{\mu_{xy}}{\mu_x \mu_y}$

Lemma 4.1.1. *Let p_n be the transition densities of $\{X_n\}_{n\geq 0}$*

1. $p_{n+m}(x,y) = \sum_{z\in V} p_n(x,z)p_m(z,y)\mu_z$
2. $\forall\, x,y \in V$, $p_n(x,y) = p_n(y,x)$
3. $\forall\, x,y \in V$, $\sum_{z\in V} p_n(x,z)\mu_z = 1 = \sum_{z\in V} p_n(z,y)\mu_z$

Proof. 1.

$$\begin{aligned}
p_{n+m}(x,y) &= \frac{\mathbf{P}^x(X_{n+m} = y)}{\mu_y} \\
&= \sum_{z\in V} \frac{\mathbf{P}^x(X_{n+m} = y, X_n = z)}{\mu_y} \\
&= \frac{1}{\mu_y} \sum_{z\in V} \sum_{o\leq i<n+m, x_i\in V} 1_{\{x\}}(x_0) \prod_{i=0}^{n-1} \mathcal{P}(x_i, x_{i+1}) 1_{\{z\}}(x_n) \prod_{i=n}^{n+m} \mathcal{P}(x_i, x_{i+1}) 1_{\{y\}}(x_{n+m}) \\
&\overset{\text{H1}}{=} \frac{1}{\mu_y} \sum_{z\in V} \mathbf{P}^x(X_n = z)\mathbf{P}^z(X_m = y) \\
&= \frac{1}{\mu_y} \sum_{z\in V} p_n(x,z)\mu_z \; p_m(z,y)\mu_y \\
&= \sum_{z\in V} p_n(x,z)p_m(z,y)\mu_z
\end{aligned}$$

2.

$$p_n(x,y) = \frac{\mathbf{P}^x(X_n = y)}{\mu_y} = \frac{\mathbf{P}^y(X_n = x)}{\mu_x} = p_n(y,x)$$

The second equality is obtained by applying the Detailed Balance equations.

3.

$$\sum_{z\in V} p_n(x,z)\mu_z = \sum_{z\in V} \mathbf{P}^x(X_n = z) = 1$$
$$\sum_{z\in V} p_n(z,y)\mu_z = \sum_{z\in V} p_n(y,z)\mu_z = \sum_{z\in V} \mathbf{P}^y(X_n = z) = 1$$

□

4.2 Function Spaces

Definition 4.2.1.

$$C(V) = \{f : V \to \mathbb{R}\} = \mathbb{R}^V$$
$$\mathrm{Co}(V) = \{f : V \to \mathbb{R},\ f \neq 0 \textit{ on finitely many points}\}$$
$$C_+(V) = \{f : f \in C(V),\ f \geq 0\}$$
$$\mathrm{Supp}(f) = \{x : x \in V,\ f(x) \neq 0\}$$

Definition 4.2.2. *We define the* ***norm*** *of a function as the following*

$$\forall\, p \in [1,\ \infty),\ \|f\|_p = (\sum_{x\in V} |f(x)|^p\, \mu_x)^{\frac{1}{p}}$$
$$\|f\|_\infty = \sup\{|f(x)| : x \in V\}$$

f is said to be L^p on the graph with vertex set V and weights μ if and only if f is a function defined on the vertex set, V and its p-norm is finite everywhere.

$$f \in L^p(V,\mu) \iff f \in C(V) \text{ and } \|f\|_p < \infty$$

Definition 4.2.3. *We define an inner product on the $L^2(V,\ \mu)$ space in the following way*

$$\langle f,\ g\rangle = \sum_{x\in V} f(x)g(x)\mu_x$$

$$\mathbf{E}[f(X_n)] = \sum_{x\in V} f(z)\mathbf{P}^x(X_n = z)$$
$$= \sum_{x\in V} f(z)p_n(x,z)\mu_z$$
$$= \langle f,\ p_n(x,\ .)\rangle$$

which brings us to define a new function

Definition 4.2.4. *$\mathcal{P}_n : C(V) \to C(V)$ given by*

$$\mathcal{P}_n f(x) = \sum_{x \in V} f(z) p_n(x, z)\mu_z = \langle f,\ p_n(x,\ .)\rangle$$

where $\Delta : C(V) \to C(V)$ as an "operation" on C(V) is

$$\Delta = P - I$$

We write $\mathcal{P}_1 f(x)$ as $\mathcal{P} f(x)$ and proceed to look at computations and lemmas involving $\mathcal{P} f$.

Lemma 4.2.1.

$$\forall x \in V,\ \mathcal{P} f(x) - f(x) = \Delta f(x)$$

Proof.

$$\begin{aligned}
\mathcal{P} f(x) - f(x) &= \sum_{x \in V} f(z) p(x, z)\mu_z - f(x) \\
&\overset{*}{=} \sum_{x \in V} p(x, z)\mu_z (f(z) - f(x)) \\
&= \sum_{x \in V} \frac{\mu_{xz}}{\mu_x \mu_z} \mu_z (f(z) - f(x)) \\
&= \frac{1}{\mu_x} \sum_{x \in V} \mu_{xz} (f(z) - f(x)) \\
&= \Delta f(x)
\end{aligned}$$

* is left as an exercise and can be proved using property 2 from (4.1.1) □

Corollary 4.2.1.

$$\Delta f = 0 \iff f(x) = \mathcal{P} f(x) = \mathbf{E}^x[f(X_1)]$$

Definition 4.2.5. *We define a function $A : C(V) \to C(V)$ as*

$$\|A\|_{p \to p} = \sup\{\|Af\|_p : \|f\|_p \leq 1\}$$

Proposition 4.2.1.

1. $\mathcal{P}1 = 1$
   where $1(x) = 1\ \forall x \in V$

2. $|\mathcal{P} f| \leq \mathcal{P}|f|$
   where $f \in C(V)$

3. $\|\mathcal{P}\|_{p \to p} \leq 1$
   $\|\Delta\|_{p \to p} \leq 2$
   where $p \in [1, \infty) \cup \{\infty\}$

Proof. 1.

$$\mathcal{P}1(x) = \sum_{x \in V} p(x,z)\mu_z = 1 = 1(x)$$

2.

$$\begin{aligned}
|\mathcal{P}f(x)| &= \left| \sum_{x \in V} f(z)p(x,z)\mu_z \right| \\
&\leq \sum_{x \in V} |f(z)|\, p(x,z)\mu_z \\
&= \mathcal{P}|f|(x)
\end{aligned}$$

3.

$$\begin{aligned}
\|\mathcal{P}f\|_p^p &= \sum_{x \in V} |\mathcal{P}f(x)|^p \mu_x \\
&= \sum_{x \in V} \left| \sum_{z \in V} f(z)p(x,z)\mu_z \right|^p \mu_x \\
&\overset{*}{\leq} \sum_{x \in V} \left(\sum_{z \in V} |f(z)|^p\, p(x,z)\mu_z \right) \left(\sum_{z \in V} 1^q p(x,z)\mu_z \right) \mu_x \\
&= \sum_{x \in V} \left(\sum_{z \in V} |f(z)|^p\, p(x,z)\mu_z \right) \mu_x \\
&\overset{**}{=} \sum_{z \in V} |f(z)|^p \mu_z \\
&= \|f\|_p \\
\implies \|\mathcal{P}\|_{p \to p} &\leq 1
\end{aligned} \tag{4.3}$$

where $\frac{1}{p} + \frac{1}{q} = 1$

We leave the proofs of the following as exercises
*, which can be proved using Holder's inequality, ** and the $p = \infty$ case

$$\begin{aligned}
\|\Delta f\|_p^p &= \|\mathcal{P}f - f\|_p^p \\
&\leq (\|\mathcal{P}f\|_p + \|f)\|_p)^p \\
&\leq 2^{p-1}(\|\mathcal{P}f\|_p^p + \|f\|_p^p) \\
&\leq (2\,\|f\|_p)^p \qquad [\because \|\mathcal{P}f\|_p \leq \|f\|_p] \\
\implies \|\Delta\|_{p \to p} &\leq 1
\end{aligned}$$

The final inequality is obtained from (4.3).

$\square$

Proposition 4.2.2. *$\mathcal{P}$ is self-adjoint on $L^2(V, \mu)$*

$$\forall\ f, g \in L^2(V, \mu),\ \langle \mathcal{P}f,\ g\rangle = \langle f,\ \mathcal{P}g\rangle$$

Proof.

$$\begin{aligned}
\langle \mathcal{P}f,\ g\rangle &= \sum_{x\in V} \mathcal{P}f(x)g(x)\mu_x \\
&= \sum_{x\in V}(\sum_{z\in V} f(z)p(x,z)\mu_z)g(x)\mu_x \\
&\overset{Ex}{=} \sum_{z\in V} f(z)\mu_z \sum_{x\in V} p(z,x)g(x)\mu_x \\
&= \sum_{z\in V} f(z)\mathcal{P}g(z)\mu_z \\
&= \langle f,\ \mathcal{P}g\rangle
\end{aligned}$$

□

4.3 Dirichlet forms

Definition 4.3.1. *We define the quadratic form on $L^2(V, \mu)$, $\mathcal{E}$ as*

$$\mathcal{E}(f,g) = \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x) - f(y))(g(x) - g(y))\mu_{xy}$$

whenever the series converges absolutely.

Theorem 4.3.1 (Discrete Green's Theorem)**.** $\forall\ f, g \in C(V)$,

$$\begin{aligned}
&\sum_{x\in V}\sum_{y\in V} |f(x) - f(y)|\,|g(x)|\,\mu_x < \infty \\
&\implies \mathcal{E}(f,g) = -\langle \Delta f,\ g\rangle
\end{aligned}$$

We present an application of (4.3.1)

Lemma 4.3.1. *Let $(\Gamma,\ \mu)$ be a weighted graph such that $\mu(V) < \infty$. Then, $(\Gamma,\ \mu)$ is **recurrent**.*

Proof. Fix $Z \in V$ Define $\Phi : V \to \mathbb{R}$ where $\Phi(x) \coloneqq \mathbf{P}^x(\mathcal{T}_z = \infty)$

1. Firstly observe that $\Phi(z) = \mathbf{P}^z(\mathcal{T}_z = \infty) = 0$

2. $\forall\ n \geq 1,\ x \neq z$
   $\Phi_n(x) \coloneqq \mathbf{P}^x(\mathcal{T}_z = n) = \sum_{u\in V} \mathcal{P}(x,u)\Phi_{n-1}(x)$

   This holds true from a simple logical argument. Starting from x, hitting z in n steps is equivalent to jumping from x to some vertex u and hitting z in $n-1$ steps.

3. $1 - \Phi(x) = \mathbf{P}^x(\mathcal{T}_z < \infty) = \sum_{n=0}^{\infty} \mathbf{P}^x(\mathcal{T}_z = n) = \sum_{n=1}^{\infty} \Phi_n(x)$
4. $\Phi \equiv 0$

$$
\begin{aligned}
&\sum_{n=1}^{k} \Phi_n(x) = \sum_{n=1}^{k} \sum_{u \in V} \mathcal{P}(x,u)\Phi_{n-1}(u) \\
\implies &\sum_{n=1}^{k} \Phi_n(x) = \sum_{u \in V} \mathcal{P}(x,u) \sum_{n=1}^{k} \Phi_{n-1}(u) \\
\implies &\sum_{n=1}^{\infty} \Phi_n(x) = \sum_{u \in V} \mathcal{P}(x,u) \sum_{n=1}^{\infty} \Phi_{n-1}(u) \\
\implies &1 - \Phi(x) = \sum_{u \in V} \mathcal{P}(x,u)(1 - \Phi(u)) \\
\implies &1 - \Phi(x) = \sum_{u \in V} p(x,u)(1 - \Phi(u))\mu_u \\
\implies &1 - \Phi = \mathcal{P}(1 - \Phi) \\
\implies &\Delta(1 - \Phi) = 0
\end{aligned}
$$

Then, by theorem (4.3.1),

$$
\begin{aligned}
&\mathcal{E}(1-\Phi, 1-\Phi) = \langle \Delta(1-\Phi),\ 1-\Phi \rangle = 0 \\
\implies &\frac{1}{2} \sum_{x \in V} \sum_{y \in V} (\Phi(x) - \Phi(y))^2 \mu_{xy} = 0 \\
\implies &\Phi(x) = \Phi(y) \qquad \forall\, x, y \in V \\
\implies &\Phi \text{ is constant} \\
\implies &\Phi \equiv 0
\end{aligned}
$$

The last equality holds as $\Phi(z) = 0$.

Since, $\Phi \equiv 0$ for arbitrary z, $(\Gamma,\ \mu)$ is recurrent.

$\square$

To proof theorem (4.3.1), we start with some prerequisites.

Definition 4.3.2.

$$
\begin{aligned}
&\mathcal{H}^2(V) = \{f : f \in C(v),\ \mathcal{E}(f,f) < \infty\} \\
&\|f\|_{\mathcal{H}^2} = \sqrt{\mathcal{E}(f,f) + f^2(\rho)} \qquad \textit{for some fixed } \rho \in V
\end{aligned}
$$

Proposition 4.3.1. *Let $(\Gamma,\ \mu)$ be a graph satisfying properties, H1 and H2.*

1. $|f(x) - f(y)| \leq \frac{1}{\sqrt{\mu_{xy}}} \sqrt{\mathcal{E}(f,f)} \qquad \forall\, x \sim y$
2. $\mathcal{E}(f,f) = 0 \iff f$ *is constant*

3. $f \in L^2 \implies \mathcal{E}(f,f) \leq 2\|f\|_2^2$

Proof. 1. If $\mathcal{E}(f,f) = \infty$, then we are done

Let $\mathcal{E}(f,f)$ be finite

$$\begin{aligned}\mathcal{E}(f,f) &= \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x)-f(y))^2\mu_{xy} \\ &\geq (f(x)-f(y))^2\mu_{xy} \qquad \forall\, x,y \in V\end{aligned}$$

2. The forward direction is left as an exercise. The reverse direction follows from the definition.

3.

$$\begin{aligned}\mathcal{E}(f,f) &\leq \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x)-f(y))^2\mu_{xy} \\ &\leq \sum_{x\in V}\sum_{y\in V}(|f(x)|^2+|f(y)|^2)\mu_{xy} \\ &\overset{Ex}{=} \sum_{x\in V}|f(x)|^2\mu_x + \sum_{y\in V}|f(y)|^2\mu_y \\ &= 2\|f\|_2^2\end{aligned}$$

The second last equality is left as an exercise.

□

Proposition 4.3.2. *Let* $f \in \mathcal{H}^2(V)$*. Then,*

$$\|\Delta f\|_2^2 \leq 2\mathcal{E}(f,f)$$

Proof.

$$\begin{aligned}\|\Delta f\|_2^2 &= \sum_{x\in V}(\Delta f(x))^2\mu_x \\ &= \sum_{x\in V}\left[\frac{1}{\mu_x}\sum_{y\in V}(f(y)-f(x))^2\mu_{xy}\right]^2\mu_x \\ &= \sum_{x\in V}\frac{1}{\mu_x}\left[\sum_{y\in V}(f(x)-f(y))^2\mu_{xy}\right]^2 \\ &\overset{Ex}{\leq} \sum_{x\in V}\frac{1}{\mu_x}\left[\sum_{y\in V}(f(x)-f(y))^2\mu_{xy}\right]\left[\sum_{y\in V}\mu_{xy}\right] \\ &= 2\mathcal{E}(f,f)\end{aligned}$$

The second last inequality is an exercise and can be shown using Cauchy-Schwarz inequality. □

Proof of Discrete Green's Theorem (4.3.1).

$$\begin{aligned}
\langle \Delta f,\ g\rangle &= \sum_{x\in V} \Delta f(x) g(x)\mu_x \\
&= \sum_{x\in V} \frac{1}{\mu_x} \sum_{y\in V} (f(y)-f(x))\mu_{xy} g(x)\mu_x \\
&= -\sum_{x\in V}\sum_{y\in V} (f(x)-f(y))\mu_{xy} g(x)
\end{aligned}$$

$$\begin{aligned}
\mathcal{E}(f,g) &= \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x)-f(y))(g(x)-g(y))\mu_{xy} \\
&= \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x)-f(y))g(x)\mu_{xy} - \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x)-f(y))g(y)\mu_{xy} \\
&= -\frac{1}{2}\langle \Delta f,\ g\rangle - \frac{1}{2}\sum_{y\in V}\sum_{x\in V}(f(x)-f(y))g(y)\mu_{xy} \\
&\overset{*}{=} -\frac{1}{2}\langle \Delta f,\ g\rangle - \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(y)-f(x))g(x)\mu_{yx} \\
&= -\frac{1}{2}\langle \Delta f,\ g\rangle + \frac{1}{2}\sum_{x\in V}\sum_{y\in V}(f(x)-f(y))g(x)\mu_{xy} \\
&= -\frac{1}{2}\langle \Delta f,\ g\rangle - \frac{1}{2}\langle \Delta f,\ g\rangle \\
&= -\langle \Delta f,\ g\rangle
\end{aligned}$$

where * is obtained by flipping the labels of x and y. ☐

Example.
Let $V = \mathbb{N}$ and μ be the usual weights.
Define $f, g : \mathbb{N} \to \mathbb{R}$ such that

$$\begin{aligned}
f(n) &:= \sum_{i=1}^{n} \frac{(-1)^i}{i} \\
g(n) &:= 1
\end{aligned}$$

Then,

$$\begin{aligned}
\mathcal{E}(f,f) &= \frac{1}{2}\left[\sum_{k\geq 1}(f(k+1)-f(k))^2 + \sum_{k\geq 1}(f(k-1)-f(k))^2\right] \\
&\leq \sum_{k\geq 2}\frac{1}{k^2} < \infty
\end{aligned}$$

$$
\begin{aligned}
\mathcal{E}(g, g) &= 0 \\
\mathcal{E}(f, g) &= 0
\end{aligned}
$$

$$
\begin{aligned}
\Delta f(n) &= \frac{1}{2}[f(n+1) + f(n-1) - 2f(n)] \\
&= \frac{(-1)^{n+1}}{2} \frac{2n+1}{n(n+1)} \\
\implies \langle \Delta f,\ g \rangle &= \frac{3}{4} + \sum_{n=2}^{\infty} \frac{(-1)^{n+1}}{2} \frac{2n+1}{n(n+1)} 2 \\
&= \frac{3}{4} - \frac{1}{2} \neq 0
\end{aligned}
$$

which contradicts the Discrete Green's Theorem (4.3.1)

Week 5 **February 17, 2023**

Killed process and Green's function

Lecturer: Siva Athreya Scribe: Aniket Sen, Arun Sharma

5.1 Introduction

(Γ, μ) is a weighted graph which is H1(Locally finite) and H2(Connected).$\{X_n\}$ is a simple random walk on it.

Transition density: $p_n^x(y) = \frac{\mathbf{P}^x(X_n)=y}{\mu_y}$
$p_0(x,y) = \frac{\mathbf{1}_x(y)}{\mu_y}$
The transition density satisfies the following:

- $p_{n+m}(x,y) = \sum_{z\in\mathbb{V}} p_n(x,z)p_m(z,y)\mu_z$ [Chapman-Kolmogorov Equation]
- $p_n(x,y) = p_n(y,x)$ [Symmetry]
- $P(p_n^x(y)) = \sum_{z\in\mathbb{V}} p(y,z)p_n^x(z)\mu_z = \sum_{z\in\mathbb{V}} p(y,z)p_n(x,z)\mu_z = p_{n+1}^x(y)$ [*details left as Exercise*]
- $p_t^x(y) = P(t;x,y) = \frac{e^{-(x-y)^2}/2t}{\sqrt{2\pi t}}$
  $\Leftrightarrow \frac{\delta}{\delta t}p_t^x = \Delta p_t^x = \frac{\delta^2}{\delta y^2}p_t^x$
- $\Delta p_n^x(y) = (P-I)p_n^x y = p_{n+1}^x(y) - p_n^x(y)$
- $\|p_n^x\|_2^2 = \langle p_n^x, p_n^x\rangle = p_{2n}(x,x) = \frac{\mathbf{P}^x(X_2n=x)}{\mu_x} \leq \frac{1}{\mu_x}$

Dirichlet form/Energy form
$\varepsilon(f,g) = \frac{1}{2}\sum_{x\in\mathbb{V}}\sum_{y\in\mathbb{V}}$
Domain of $\varepsilon : D(\varepsilon) = \{f : \mathbb{V} \to \mathbb{R} | \varepsilon(f,f) < \infty\}$

$$\begin{aligned}\varepsilon(f,g) &= -\langle \Delta f, g\rangle \\ &= -\langle (P-I)f, g\rangle \\ &= -\langle Pf, g\rangle + \langle f, g\rangle\end{aligned}$$

where the first equality comes from Discrete Gauss-Green theorem.

$$\varepsilon \leftrightarrow \Delta \leftrightarrow P \leftrightarrow \{X_n\}_{n\geq 1}$$

on $\mathbb{R}^n$

$$\varepsilon(f,g) = \int_{\mathbb{R}^n} \nabla f(x) \nabla g(x) dx$$

it can be shown that if $f \in D(\varepsilon)$, $-\langle \Delta f, g \rangle_n$

$$\varepsilon \leftrightarrow \Delta \leftrightarrow \{P_t\}_{t \geq 0} \leftrightarrow \{X_t\}_{t \geq 0}$$

$$\begin{aligned} \varepsilon(p_n^x, p_m^y) &= -\langle \Delta p_n^x, p_m^y \rangle \\ &= -\langle p_{n+1}^x - p_n^x, p_m^y \rangle \\ &= -\langle p_{n+1}, p_m^y \rangle + \langle p_n^x, p_m^y \rangle \\ &= -p_{n+m+1}(x,y) + p_{n+m}(x,y) \end{aligned}$$

where the first equality comes from Discrete Gauss-Green theorem. *As an Exercise* check that $p_n^x(.)$ and $p_m^y(.)$ satisfies the hypothesis of Discrete Gauss- Green Theorem.

$x \in \mathbb{V}$, $I_x(z) = \begin{cases} 1, z = x \\ 0, otherwise \end{cases}$

$$\begin{aligned} \varepsilon(I_x, I_y) &= -\langle \Delta I_x, I_y \rangle \\ &= -\sum_{z \in \mathbb{V}} I_y(x) \Delta I_x(z) \mu_z \\ &= -\Delta I_x(y) \mu_y \\ &= \mu_y \frac{\sum_{z \in \mathbb{V}} (I_x(z) - I_x(y) \mu_{zy}}{\mu_y} \\ &= \begin{cases} -\mu_{xy}, if y \neq x \\ \mu_x - \mu_{xx}, if y = x \end{cases} \end{aligned}$$

5.2 Killed Process

<u>Gambler's ruin</u>

N: Total capital of 2 players

X_k : Capital of Player 1 in k^{th} step

$$\mathbf{P}^x(X_{T_{\{0,N\}}} = 0) = h(X) \leftrightarrow h(x) = \begin{cases} \frac{1}{2}h(x-1) + \frac{1}{2}h(x+1), 0 < x < N \\ 1, x = 0 \\ 1, x = N \end{cases}$$

$$h = Ph \Leftrightarrow \Delta h = 0$$

Let the graph $\Gamma = (\mathbb{V}, E)$ be H1 and H2 with weights μ. $A \subset \mathbb{V}$.

$\tau_A = \tau_{A^c} = inf\{n \geq 1 | X_n \in A^c\}$

We define the kill density, i.e. the transition density of the random walk until it exits A by:

$$p_n^A(x,y) = \frac{\mathbf{P}^x(X_n = y, n < \tau_A}{\mu_y}$$

- if $y \notin A$, then $p_n^A(x,y) = 0 \ \forall n \geq 1$
- $I_A f(x) = I_A(x) f(x)$
- $n \geq 1$, $P_n^A f(x) = \sum_{z \in \mathbb{V}} p_n^A(x,z) f(z) \mu_z = F^x[f(X_n); n < \tau_A]$
- $\Delta^A := P^A - I^A$

Lemma 5.2.1. *(a)* $p_n^A(x,y) = 0 \ \forall x, y \notin A, n \geq 1$

(b) $p_{n+1}^A(x,y) = \sum_{z \in \mathbb{V}} p_n^A(x,z) p^A(z,y) \mu_z$

(c) $\Delta p_n^{A,x} = p_{n+1}^{A,x} - p_n^{A,x}$
[$p_n^{A,x} = p_n^A(x,y)$*]*

(d) $p_n^A(x,y) = p_n^A(y,x) \ \forall x, y \in \mathbb{V}$

(e) $P_n^A f(x) = (P^A)^n f(x) \ \forall n \geq 1$

(f) $P^A f(x) = I_A P I_A f(x)$

Proof. Left as an Exercise. □

5.3 Green's function

Let $A \subset \mathbb{V}$. We define Green's function of $\{X_n\}_{n \geq 0}$ as:

$$g_A(x,y) = \sum_{n=0}^{\infty} p_n^A(x,y)$$

$x, y \in \mathbb{V}$.

Notation. • *if* $A = \mathbb{V}$ *then* $g_A = g$

- $x \in \mathbb{V}$ *fixed, then* $g_A^x(y) = g_A(x, y \ \forall y \in \mathbb{V}$

Observations. • $g_A(x,y) = g_A(y,x) \ \forall \ x, y \in \mathbb{V}$.

- *Define Local time at* y *before exiting* A *i.e. time spent by the walk at* y *before exiting* A *by* $L^y_{\tau_A} = \sum_{n=0}^{\infty} \mathbf{1}_{X_n=y}$.

$$\begin{aligned} g_A(x,y) &= \sum_{n=0}^{\infty} p_n^A(x,y) \\ &= \frac{\sum_{n=0}^{\infty} E^x[\mathbf{1}_{X_n=y}; n < \tau_A]}{\mu_y} \\ &= \frac{E^x[\sum_{n=0}^{\infty}(\mathbf{1}_{X_n=y}\mathbf{1}_{n<\tau_A})]}{\mu_y} \\ &= \frac{E^x[\sum_{n=0}^{\tau_A-1}(\mathbf{1}_{X_n=y})]}{\mu_y} \\ &= \frac{E^x[L^y_{\tau_A}]}{\mu_y}. \end{aligned}$$

- *if* $A = \mathbb{V}$ *and* $\mathbb{V}$ *is recurrent then* $g(x,.) = \infty$

Theorem 5.3.1. $A \subset \mathbb{V}$. *Suppose either* (Γ, μ) *is transient or* $A \neq V$. *Then*

1. $g_A(x,y) = \mathbb{P}(\tau_y < \tau_A) g_A(y,y)$
2. $g_A(y,y) = \frac{1}{\mu_y \mathbb{P}(\tau_a \leq \tau_y^+)}$

Lemma 5.3.1. *Let* $x, y \in A$. *Then,*

1. $\mathbf{P} g_A^x(y) = g_A(x,y) - \frac{\mathbf{1}_x(y)}{\mu_x}$
2. $\Delta g_A^x(y) = \begin{cases} -\frac{1}{\mu_x} \text{ if } y{=}x \\ 0 \text{ otherwise} \end{cases}$

Proof. 1.

$$\begin{aligned}
Pg_A^x &= \sum_{z\in\mathbb{V}} p(y,z)g_A^x(z)\mu_z \\
&= \sum_{z\in\mathbb{V}} p(y,z)\mu_z(\sum_{n=0}^{\infty} p_n^A(xz) \\
&= \sum_{n=0}^{\infty}\sum_{z\in\mathbb{V}} p(y,z)\mu_z p_n^A(x,z) \\
&= \sum_{n=0}^{\infty}\sum_{z\in A} p(y,z)\mu_z p_n^A(x,z) \\
&= \sum_{n=0}^{\infty}\sum_{z\in A} p_1^A(y,z)p_n^A(x,z)\mu_z \\
&= \sum_{n=0}^{\infty} p_{n+1}^A(x,y) \\
&= g_A(x,y) - p_0^A(x,y) \\
\Rightarrow Pg_A^x(y) &= g_A(x,y) - \frac{\mathbf{1}_x(y)}{\mu_x}
\end{aligned}$$

2. follows from definition of $D = P - I$ $\square$

Proof of Theorem.
Notations: Given $f : \mathbb{V} \to \mathbb{R}$, $E^X f(X_n) = \sum_{y\in\mathbb{V}} \mathbf{P}^x(X_n = y)f(y)$.
let ξ be a random variable. $h_n(\xi) = E^\xi f(X_n)$
1.

$$\begin{aligned}
g_A(x,y)\mu_y &= E^x(L_{\tau_A}^y \\
&= E^x(\mathbf{1}_{\tau_y<\tau_A} \times L_{\tau_A}^y \\
&= E^x(\mathbf{1}_{\tau_y<\tau_A}\mathbf{E}^y(L_{\tau_A}^y) \\
\Rightarrow g_A(x,y) &= g_A(y,y)\mathbf{P}^x(\tau_y < \tau_A)\square
\end{aligned}$$

2. $p = \mathbf{P}(\tau_y^+ < \tau_A)$
if (Γ,μ) is transient then $p < 1$ and if recurrent and $A \neq \mathbb{V}$ then $p < 1$.[$\exists z \in A^c$ such that $\mathbf{P}^y(\tau_A < \tau_y^+) \geq \mathbf{P}^y(\tau_z < \tau_y^+) > 0$]
$\therefore p < 1$

$$
\begin{aligned}
\mathbf{P}^y(L^y_{\tau_A} = k) &= p^k(1-p) \\
\Rightarrow \mu_y g_A(y,y) &= E^y(L^y_{\tau_A}) \\
&= \sum_{k=0}^{\infty} p^k(1-p) \\
&= \frac{1}{1-p} \\
&= \frac{1}{\mathbf{P}(\tau_A \le \tau_y^+)} \\
\Rightarrow g_A(y,y) &= \frac{1}{\mu_y \mathbf{P}(\tau_A \le \tau_y^+)} \square
\end{aligned}
$$

Combining 1 and 2, we get

$$
g_A(x,y) = \frac{\mathbf{P}^x(\tau_y < \tau_A)}{\mu_y \mathbf{P}(\tau_A \le \tau_y^+)}.
$$

Week 6 **March 3, 2023**

Harmonic Functions

Lecturer: Siva Athreya Scribe: Rahil Miraj, Jainam Khakra

Harmonic Functions:

Let $\Gamma = (V, E, \mu)$ be weighted graph and let $A \subseteq V$, $\bar{A} = A \bigcup \partial A$. Then $f : \bar{A} \to \mathbb{R}$ is said to be:
Harmonic if $\Delta f = 0,\ i.e.\ (Pf = f)$,
Super Harmonic if $\Delta f \leq 0,\ i.e.\ (Pf \leq f)$,
Sub Harmonic if $\Delta f \geq 0,\ i.e.\ (Pf \geq f)$.

Examples:
1. For $x, y \in V$ and $A \subseteq V$,

$$g_A(x, y) = \sum_{n=0}^{\infty} p_n^A(x, y) = \frac{\mathbb{E}[L^y_{\tau_A}]}{\mu_y}.$$

$$\Delta g_A(x, y) = -\frac{1}{\mu_x}\mathbf{1}_{\{\mathbf{x}\}}(\mathbf{y}) = \begin{cases} -\frac{1}{\mu_x}, & \text{if } y = x \\ 0, & \text{if } y \neq x. \end{cases}$$

Therefore, $g_A(x, \cdot)$ is Harmonic in $A \backslash x$ and Super Harmonic in A.
2. For $z \in V$ and $x \neq z$,$\phi(x) = \mathbb{P}^x(T_z = \infty)$ is Harmonic in $V \backslash z$.

Theorem 6.0.1. ***(Foster's Criteria/Lyapunov Function)****: Let $A \subseteq V$ be a finite set. Then (Γ, μ) is recurrent iff there exists a function h, which is:
non negative, Super Harmonic on $V \backslash A$ and $|\{x : h(x) < M\}| < \infty\ \forall\ M > 0$.*

Proof:
($\Rightarrow$)
WLOG, we take $A = \{\rho\}$. Suppose $\exists h : V \to [0, \infty)$ such that h is super harmonic on $V \backslash \{\rho\}$ and $|\{x : h(x) < M\}| < \infty\ \forall\ M > 0$.
$T_\rho = min\{n \geq 0 | X_n = \rho\}$, $\{X_n\}_{n \geq 0}$ is a random walk on (Γ, μ). Let $Y_n = h(X_{n \bigcap T_\rho})$ and let $\mathcal{A}_n$ be the observable events upto time n. So for $X_0 = x$,

$$\begin{aligned} \mathbb{E}^x[Y_n | \mathcal{A}_{n-1}] &= \mathbb{E}^x[h(X_{n \bigcap T_\rho}) | \mathcal{A}_{n-1}] \\ &= \mathbb{E}^{X_{n-1}}[h(X_{n \bigcap T_\rho})],\ \text{(SMP)} \\ &= Ph(X_{n-1 \bigcap T_\rho}), \\ &\leq h(X_{n-1 \bigcap T_\rho}),\ \text{(super harmonic)}, \\ &= Y_{n-1}. \end{aligned}$$

Super Martingle: Let $\{Z_n\}_{n\geq 1}$ be a sequence of random variables such that $\mathbb{E}[Z_n] < \infty$. Then $\{Z_n\}_{n\geq 1}$ is a super Martingle if $\mathbb{E}[Z_n|Z_{n-1}, Z_{n-2}, ...Z_1] \leq z_{n-1}$.

Theorem 6.0.2. ***(Martingle Convergence Theorem):*** *Let $Y_n \geq 0$ be Super Martingle, $\exists Y \equiv Y_\infty$ such that $Y_n \to Y_\infty$ wp 1 and $\mathbb{E}[Y_\infty] \leq \mathbb{E}[Y_0]$ as $n \to \infty$.*

So, in our case, $Y_0 = h(X_{0 \bigcap T_\rho}) = h(X_0) < \infty$.
Therefore, from the above theorem, $Y_\infty < \infty$ wp 1.
Now, suppose (Γ, μ) is Transient. Then $\exists x \in V \backslash \{\rho\}$ such that $\mathbb{P}^x(T_{\rho)} < 1$.
Let $C_n = \{y \in V\backslash\{\rho\} | h(y) \geq n\}$ $\forall\ n \geq 1$. Then $|C_n^c| < \infty$.
$N = \{T_\rho = \infty\} \bigcap \{\exists n_k \geq 1 : X_{n_k} \in C_k\}$. Let $w \in N$ and n_k be as given by N.

$$Y_{n_k} = h(X_{n_k \bigcap T_\rho}),$$
$$\Rightarrow Y_{n_k} \geq k,$$
$$\Rightarrow N \subseteq \{Y_\infty = \infty\},$$

$\Rightarrow \mathbb{P}(Y_\infty = \infty) > 0$, which contradicts Martingle Convergence Theorem.
$\Rightarrow \{X_n\}_{n\geq 0}$ can not be Transient.
Hence, $\{X_n\}_{n\geq 0}$ is Recurrent.

($\Leftarrow$)

Suppose (Γ, μ) is Recurrent.
Let $B(\rho, n) = \{x \in V | d(x, \rho) \leq n\}$ and let $h_n : V \to [0, 1]$ such that

$$h_n(x) = \mathbb{P}^x(\tau_{B(\rho,n)} < T_\rho);\ \tau_{B(\rho,n)} = T_{B(\rho,n)^c}.$$

$$(SMP) \Rightarrow Ph_n = h_n,\ x \neq \rho,$$
$$\Rightarrow \Delta h_n = 0,\ \forall\ x \neq \rho.$$

In particular, $h_n(\cdot)$ is super harmonic in $V\backslash\{\rho\}$.

$$\lim_{n\to\infty} h_n(x) = 0,$$
$$h_n(x) = 1\ \forall\ x \in B(\rho, n)^c.$$

$\exists\{n_k\}_{k\geq 1}$ such that $h_{n_k}(x) \leq \frac{1}{2^k}$ $\forall\ x \in B(\rho, n)$.
Let $x \in V$, $\sum\limits_{k=1}^{\infty} h_{n_k}(x)$ (Ex: $h(\cdot) < \infty$).
(I) $h \geq 0$.

(II) $x \in V \backslash \{\rho\}$,

$$\begin{aligned} Ph(x) &= \sum_{x \sim y} \frac{\mu_{xy}}{\mu_x} h(y), \\ &= \sum_{x \sim y} \frac{\mu_{xy}}{\mu_x} \sum_{k=1}^{\infty} h_{n_k}(x), \\ &= \sum_{k=1}^{\infty} \sum_{x \sim y} \frac{\mu_{xy}}{\mu_x} h_{n_k}(x), \\ &= \sum_{k=1}^{\infty} Ph_{n_k}(x), \\ &= \sum_{k=1}^{\infty} h_{n_k}(x) = h(x). \end{aligned}$$

(III) Let $M > 0$ and $U = \{x \in V | h(x) < M\}$.

$$\forall j \geq h_{n_j}(x) = 1 \; \forall \; x \in B(\rho, n_j)^c.$$

Claim: $U \subseteq B(\rho, n_m)^c$.
Proof: Given $M > 0$, $\exists n_m$ such that
$\forall j \geq h_{n_j}(x) = 1 \; \forall \; x \in B(\rho, n_m)^c$, for $1 \leq j \leq M$.

$$\sum_{j=1}^{m} h_{n_j}(x) = M \; \forall \; x \in B(\rho, n_m)^c.$$

$$\Rightarrow h(x) \geq M \; \forall \; x \in B(\rho, n_m)^c.$$

Theorem (Maximum Principle):
$A \subset V$, connected, $h : V \to \mathbb{R}$ such that $\Delta h \geq 0$ on A

a) If $\exists x \in A$ such that $h(x) = \max\limits_{z \in A \cup \partial A} h(z)$ then h is constant on $\overline{A}$

b) $|A| < \infty$, $h(z) = \max\limits_{z \in \partial A} h(z)$

Proof
a) $B = \{y \in \overline{A} \mid h(y) = h(x)\}$
$B \neq \emptyset$ as $x \in B$
$y \in B \cap A$, $z \sim y \Rightarrow z_0 \in \overline{A} \Rightarrow z_0 \in B$
$h(z_0) \leq \max\limits_{u \in \overline{A}} h(u)$

$y \in B \cap A$ and $z \sim y$ then
$z_0 \in \overline{A} \Rightarrow h(z_0) \leq \max\limits_{u \in \overline{A}} h(u) = h(x) = h(y)$

But $\Delta h(y) \geq 0$
$\frac{1}{\mu_y} \Sigma_{z \sim y} \; \mu_{zy}(h(y) - h(z)) \geq 0$

Hence if $z \sim y \Rightarrow h(z) = h(y)$
Inductively, as A is connected, we have that $B = \overline{A}$

b) $|A| < \infty \Rightarrow \exists x_0 \in \overline{A}$, $h(x) = \max\limits_{z \in A} \ h(z)$

If $x_0 \in \partial A \Rightarrow \max\limits_{u \in \partial A} \ h(u) = \max\limits_{z \in \overline{A}} \ h(z)$

If $x_0 \in A \Rightarrow$ (a) h is constant on $\overline{A}$ and $\max\limits_{u \in \partial A} \ h(u) = \max\limits_{x \in \overline{A}} \ h(x)$

Liouville Property: (Γ, μ) is said to have the Liouville Property if all bounded harmonic functions are constant.

Strong Liouville Property: A graph is said to have the strong Liouville Property if all positive harmonic functions are constant.

Theorem: Let (Γ, μ) be recurrent. Any positive superharmonic function is constant (in particular, (Γ, μ) has the strong Liouville Property).

<u>Notes for continuation of Martingales</u>

Let $\{Z_n\}$ be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.
Let $E[Z_n] < \infty$

If $E[Z_n | Z_{n-1}, Z_{n-2}, ..., Z_1] = Z_{n-1}$, then $\{Z_n\}_{n \geq 1}$ is a martingale. $\{h(.)$ is bounded, harmonic, X_n r.v on (Γ, μ), $\{h(X_n)\}_{n \geq 1}$ $\}$

If $E[Z_n | Z_{n-1}, Z_{n-2}, ..., Z_1] \geq Z_{n-1}$, then $\{Z_n\}_{n \geq 1}$ is a submartingale. $\{h(.)$ is bounded, subharmonic, X_n r.v on (Γ, μ), $\{h(X_n)\}_{n \geq 1}$ $\}$

If $E[Z_n | Z_{n-1}, Z_{n-2}, ..., Z_1] \leq Z_{n-1}$, then $\{Z_n\}_{n \geq 1}$ is a supermartingale. $\{h(.)$ is bounded, superbharmonic, X_n r.v on (Γ, μ), $\{h(X_n)\}_{n \geq 1}$ $\}$

Jensen's Inequality

Let $f : \mathbb{R} \to \mathbb{R}$ be such that $\forall a \in \mathbb{R}, \exists c \in \mathbb{R}$ such that $f(x) \geq f(a) + c(x - a)$
Then f is said to be a convex function.

Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Then Jensen's Inequality states that $E[f(X)] \geq f(E[X])$ wherever both expectations are well defined.

Theorem (Kolmogorov's Maximal Inequality): Let $\{Z_n\}_{n \geq 1}$ be a non-negative submartingale. Then $\mathbb{P}(\max\limits_{1 \leq i \leq m} \ Z_i \geq a) \leq \frac{E[Z_m]}{a} \ \ \forall a > 0, \forall m \geq 1$

Proof:
Let $m \in \mathbb{N}$ and $a > 0$ be given.
$J = min.\{min.\{n \geq 1 | Z_n > a\}, m\}$
$\tilde{J} = min.\{n \geq 1 | Z_n > a\}$
It is left as an exercise to show that J is a bounded stopping time. $Z_J \geq a \Leftrightarrow Z_n \geq a$ for some $n \leq m$

$\mathbb{P}(\max\limits_{1 \leq i \leq m} Z_i \geq a) = \mathbb{P}(Z_J \geq a) \leq \frac{E[Z_J]}{a}$

This follows from the Markov Inequality which can be applied here since $Z_i \geq 0$

$Z_J = \sum_{k=1}^{m} Z_k \mathbf{1}_{J=k} + Z_m \mathbf{1}_{\tilde{J}>m}$

$E[Z_J] = \sum_{k=1}^{m} E[Z_k \mathbf{1}_{J=k}] + E[Z_m \mathbf{1}_{\tilde{J}>m}]$
Since $\{Z_k\}_{k \geq 1}$ is a submartingale, we have that the above is less than or equal to
$\sum_{k=1}^{m} E[E[Z_m | \mathcal{A}_k] \mathbf{1}_{J=k}] + E[Z_m \mathbf{1}_{\tilde{J}>m}]$
The above is equal to
$\sum_{k=1}^{m} E[E[Z_m \mathbf{1}_{J=k} | \mathcal{A}_k]] + E[Z_m \mathbf{1}_{\tilde{J}>m}]$ (Since $E[XY|\mathcal{A}_Y] = YE[X|\mathcal{A}_Y]$)
$= \sum_{k=1}^{m} E[Z_m \mathbf{1}_{J=k}] + E[Z_m \mathbf{1}_{\tilde{J}>m}]$
$= E[Z_m(\sum_{k=1}^{m} \mathbf{1}_{J=k} + \mathbf{1}_{\tilde{J}>m}]$
$= E[Z_m] = 0$

Corollary: Let $\{Z_m\}_{m \geq 1}$ be a martingale.

1) $E[Z_m^2] < \infty \quad \forall m \geq 1, \mathbb{P}(\max\limits_{1 \leq i \leq m} |Z_i| \geq a) \leq \frac{E[|Z_m|^2]}{a^2}$

2) $Z_m \geq 0, \mathbb{P}(\sup\limits_{n \geq 1} Z_n > a) \leq \frac{E[Z_1]}{a}$

Proof
For 1) Let $Y_n = Z_n^2$. Then we apply Kolmogorov's Maximal Inequality. The proof of 2) will be done later.

Theorem (Martingale Convergence):
Let $\{Z_n\}_{n \geq 1}$ be a martingale and $\sup\limits_{n \geq 1} E[Z_n^2] < \infty$. Then $\exists\ Z$ such that $Z_n \to Z$ w.p. 1 as $n \to \infty$

Proof
$f(x) = x^2$ is a convex function. Hence, Jensen's inequality applies to conditional expectation.
$E[Z_n^2 | Z_{n-1}^2, ..., Z_1^2] \geq Z_{n-1}^2 \quad \forall n \geq 2$

From the Tower Property, it follows that
$E[Z_n^2 | Z_i^2, ..., Z_1^2] \geq Z_i^2$ for $1 \leq i \leq n$

Thus, $E[Z_n^2] \geq E[Z_i^2] \quad \forall 1 \leq i \leq n$

In particular, $E[Z_n^2] \geq E[Z_{n-1}^2]$ and $\sup\limits_{n\geq 1} \; E[Z_n^2] < \infty$

Thus $\exists \;\; \alpha > 0$ such that $E[Z_n^2] \to \alpha$ as $n \to \infty$

Let $k \geq 1$, $Y_m = Z_{k+m} - Z_k \;\; \forall m \geq 1$
Exercise: $\{Y_m\}_{m\geq 1}$ is also a martingale.

From the Tower Property, we have that $E[Z_{k+m}Z_k] = E[E[Z_{k+m}Z_k|\mathcal{A}_k]]$ which is equal to $E[Z_k E[Z_{k+m}|\mathcal{A}_k]]$ which is equal to $E[Z_k^2]$ since $\{Z_n\}_{n\geq 1}$ is a martingale and Z_k is observable in A_k.

Hence $E[Y_m^2] = E[Z_{k+m}^2] - E[Z_k^2]$

By Corollary 1,
$\mathbb{P}(\max\limits_{1\leq i\leq m} \;\; |Y_i| > a) \leq \frac{E[Z_m^2]}{a^2} = \frac{E[Z_{k+m}^2]-E[Z_k^2]}{a^2}$

$\mathbb{P}(\max\limits_{1\leq i\leq m} \;\; |Y_i| > a) = \mathbb{P}(\max\limits_{1\leq i\leq m} \;\; |Z_{i+k} - Z_k| > b)$

Letting m go to 0 on both sides, we get
$\mathbb{P}(\cup_{i\geq 1} \;\; |Z_{i+k} - Z_k| > a) \leq \frac{\alpha - E[Z_k^2]}{a^2}$

(Exercise:
i) $0 \leq \mathbb{P}(\sup\limits_{i\geq 1} \;\; |Z_{i+k} - Z_k| > a) \leq \frac{\alpha - E[Z_k^2]}{a^2}$

ii) $\lim\limits_{k\to\infty} \;\; \mathbb{P}(\cup_{i\geq 1} \;\; |Z_{i+k} - Z_k| > a) = 0 \;\; \forall a)$

From ii) above and the Borel Cantelli Lemma, we have that if $E = \{\{Z_k\}_{k\geq 1}$ is a Cauchy sequence $\}$ then $\mathbb{P}(E) = 1$.
$\Rightarrow \exists \;\; Z$ such that $Z_n \to Z$ w.p. 1 as $n \to \infty$.

Week 8 **March 10, 2023**

Isoperimetric Inequalities and Applications

Lecturer: Siva Athreya Scribe: Atreya Choudhury, Srivatsa Balaji

The focus of this chapter is to look at how the geometry of weighted graph affects the properties of the corresponding random walk on it.

Definition 8.0.1 (Isoperimetric Inequality). *Let $A, B \subseteq V$, $\mu_E(A,B) = \sum_{x\in A}\sum_{y\in B}\mu_{xy}$ and $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function.*

$(\Gamma,\ \mu)$ is said to satisfy the $\psi-$isoperimetric inequality if $\exists\, c_0 > 0$ such that

$$\frac{\mu_E(A, V \setminus A)}{\psi(\mu(A))} \geq \frac{1}{c_0} \qquad \forall A \subseteq V \text{ and } |A| < \infty$$

If a weighted graph satisfies the $\psi-$isoperimetric inequality, we say it has the I_ψ property.

A graph is said to have the property I_α for $\alpha \in [0,\infty)$ when $\psi(t) = t^{1-\frac{1}{\alpha}}$ and said to have the property I_∞ when $\psi(t) = t$

Example. $\mathbb{R}^d$. We look at $A = B(0,r)$

$$S_B \equiv \text{surface area of A} = c_d r^{d-1}$$
$$V_B \equiv \text{volume of A} = \widetilde{c_d} r^d$$
$$\therefore \frac{S_B}{V_B^{\frac{d-1}{d}}} \geq \frac{1}{c_0}$$

We can take $\psi(t) = t^{1-\frac{1}{d}}$

Show that Z^d has the I_d property for all such A such that $|A| < \infty$

Example. $\prod_2$, the binary tree has the I_∞ property with $c_0 = 3$

Observations. *If $(\Gamma,\ \mu)$ satisfies $I_{\alpha+\delta}$, then it satisfies I_α*

Definition 8.0.2 (Nash Inequality). *$\alpha \in [1,\infty)$, $(\Gamma,\ \mu)$ is said to have the property N_α if $\forall f \in \mathbb{L}^1(V) \cap \mathbb{L}^2(V)$,*

$$\mathcal{E}(f,f) \geq C_N \|f\|_1^{-\frac{4}{\alpha}} \|f\|_2^{2+\frac{4}{\alpha}}$$

Remark.

1. $(\Gamma,\ \mu)$ satisfies I_α for $\alpha \in [1, \infty) \implies (\Gamma,\ \mu)$ satisfies N_α
2. Z^d satisfies N_α

Theorem 8.0.1. *Let $\alpha \geq 1$. Then the following are equivalent*

1. $(\Gamma,\ \mu)$ *satisfies* N_α
2. $\exists\, C_H > 0$ *such that*
$$p_n(x, x) \leq \frac{C_H}{(n \vee 1)^{\frac{\alpha}{2}}} \qquad \forall\, n \geq 0 \text{ and } x \in V$$
3. $\exists\, C'_H > 0$ *such that*
$$p_n(x, y) \leq \frac{C'_H}{(n \vee 1)^{\frac{\alpha}{2}}} \qquad \forall\, n \geq 0 \text{ and } x, y \in V$$

Corollary 8.0.1.

1. *Suppose $(\Gamma,\ \mu)$ satisfies I_α. Then, $\exists\, C > 0$ such that*
$$p_n(x, y) \leq \frac{C}{(n \vee 1)^{\frac{\alpha}{2}}} \qquad \forall\, n \geq 0 \text{ and } x, y \in V$$
2. *Let V be infinite and $\mu_{xy} \geq c_0 > 0\ \ \forall\, x \sim y$. Then, $\exists\, C_1 > 0$ such that*
$$p_n(x, y) \leq \frac{C_1}{(n \vee 1)^{\frac{1}{2}}} \qquad \forall\, n \geq 0 \text{ and } x, y \in V$$

Remark.

1. $p_n(x, x) \equiv$ on-diagonal bounds
2. Theorem provides global upper bounds
3. part b of corollary 8.0.1 applied to $V = \mathbb{Z}$
$\implies$ the shortest possible on-diagonal upper bounds with natural weights
4. Let $\Gamma = \mathbb{Z}^d$ have natural weights $\mu^{(0)}_{xy}$ and $\Gamma' = \mathbb{Z}^d$ have natural weights $\mu^{(1)}_{xy}$ such that $\mu^{(1)}_{xy} \geq c_0 \mu^{(0)}_{xy}$ Let $(\Gamma,\ \mu^0)$ satisfy N_d
$\implies (\Gamma',\ \mu^1)$ satisfies N_d
$\implies$ the upper bound of the theorem holds
5. $\Gamma = \mathbb{Z}^d \cup_{(0,\ldots,0)} \mathbb{Z}^d$
$\implies \Gamma$ also satisfies N^d
6. 8.0.1 does not give us any information on upper bounds when we fix $n \geq 0$ and let $d(x, y)$ get large.

Theorem 8.0.2. *Let $(\Gamma,\ \mu)$ be a weighted graph. Then,*
$$p_n(x, y) \leq \frac{2}{\sqrt{\mu_x \mu_y}} e^{-\frac{d(x,y)^2}{2n}} \qquad \forall\, x, y \in V \text{ and } n \geq 1$$

Example. Consequences for $\mathbb{Z}^d$
We expect

$$p_n(x,y) \leq \frac{c_1}{n^{\frac{d}{2}}} e^{-c_2 \frac{d(x,y)^2}{n}}$$

$\mathbb{Z}^d$ satisfies $I_d \implies \mathbb{Z}^d$ satisfies $N_d \overset{8.0.1}{\implies} p_n(x,y) \leq \frac{c}{n^{\frac{d}{2}}} \qquad \forall\, x,y \in V$ and $n \geq 1$

$$\therefore\; p_n(x,y) \leq \frac{c}{n^{\frac{d}{2}}} \leq \frac{c}{n^{\frac{d}{2}}} e^{-\frac{d(x,y)^2}{n}} \qquad \text{when } d(x,y) \leq \sqrt{n}$$

When, $d(x,y) \geq \sqrt{2dn \log n}$,

$$p_n(x,y) \leq c_1 e^{-\frac{d(x,y)^2}{n}} = c_1 e^{-\frac{2c_2}{4}\frac{d(x,y)^2}{n}} e^{-\frac{2c_2}{4}\frac{d(x,y)^2}{n}} \leq \frac{\widetilde{c_1}}{n^{\frac{d}{2}}} e^{-\frac{c_2^2 d(x,y)^2}{n}}$$

Definition 8.0.3. $(\Gamma,\ \mu)$ *is said to have* ***polynomial volume growth*** *if* $\exists\, C_V$ *and* θ *such that*

$$\max\{|B(x,r)|\,,\ \mu(B(x,r))\} \leq C_V r^\theta \qquad \forall\, x \in V \text{ and } r \geq 1$$

Lemma 8.0.2. $(\Gamma,\ \mu)$ *has polynomial volume growth with index* θ*. Then,*

$$\mathbf{P}^x(d(x,X_n) > r) \leq c r^\theta e^{-\frac{r^2}{4n}}$$

This implies $\exists\, c_2 > 0$ *such that*

$$d(x,X_n) \leq c_2\sqrt{n \log n} \qquad \forall \text{ large } n \text{ w.p. } 1$$

Proof. We define $\mathcal{D}_k = B(x,2^k r) \setminus B(x,2^{k-1}r)$

$$\begin{aligned}
\mathbf{P}^x(d(x,X_n) > r) &\overset{Ex}{=} \sum_{k=1}^{\infty} \sum_{y\in\mathcal{D}_k} p_n(x,y)\mu_x \\
&\overset{Ex}{\leq} \sum_{k=1}^{\infty} \sum_{y\in\mathcal{D}_k} \frac{2}{\sqrt{\mu_x}}\sqrt{\mu_y} e^{-\frac{(2^{k-1}r)^2}{2n}} \\
&= \sum_{k=1}^{\infty} \frac{2}{\sqrt{\mu_x}} e^{-\frac{(2^{k-1}r)^2}{2n}} \sum_{y\in\mathcal{D}_k} \sqrt{\mu_y} \\
&\overset{Ex}{\leq} \sum_{k=1}^{\infty} \frac{2}{\sqrt{\mu_x}} e^{-\frac{(2^{k-1}r)^2}{2n}} \sqrt{|\mathcal{D}_k|}\sqrt{\mu(\mathcal{D}_k)} \\
&\overset{Ex}{\leq} \sum_{k=1}^{\infty} \frac{2}{\sqrt{\mu_x}} e^{-\frac{(2^{k-1}r)^2}{2n}} c(2^k r)^\theta
\end{aligned}$$

□

Week 8 **March 10, 2023**

Large Deviations for Random Walks

Lecturer: Siva Athreya Scribe: Venkat Trivikram, Srivatsa B

Nash Inequality (continued)

Theorem 8.0.1. *Let (Γ, μ) be a weighted graph, and let $\alpha \geq 1$. TFAE.*

(a) (Nash Inequality) (Γ, μ) satisfies (N_α)

(b) (On Diagonal Bounds) There exists $C_H > 0$ such that for every $x \in V$ and $n \geq 0$

$$p_n(x,x) \leq \frac{C_H}{(n \vee 1)^{\alpha/2}}$$

(c) (Off Diagonal Bounds) There exists $C'_H > 0$ such that for every $x, y \in V$ and $n \geq 0$

$$p_n(x,y) \leq \frac{C'_H}{(n \vee 1)^{\alpha/2}}$$

Proof. We provide only a sketch of the proof. From Worksheet 2, $(a) \implies (b)$ holds, and $(c) \implies (b)$ is trivial. First, we show $(b) \implies (c)$. So assume (b). Let $m \geq 0$. If n is even, with $n = 2m$, then for any $x, y \in V$, we have

$$p_{2m}(x,x) \leq \frac{C_H}{(2m \vee 1)^{\alpha/2}} \quad \text{and} \quad p_{2m}(y,y) \leq \frac{C_H}{(2m \vee 1)^{\alpha/2}}$$

As an exercise, show that $p_{2m}(x,y) \leq \sqrt{p_{2m}(x,x)p_{2m}(y,y)}$, and using this, we get $(b) \implies (c)$ with $C'_H = C_H$. If $n = 2m+1$, then, since $p_{2m+1}(x,y) \leq \sqrt{p_{2m}(x,x)p_{2m+2}(y,y)}$ (by a similar exercise), we get

$$p_{2m+1}(x,y) \leq \sqrt{\frac{C_H^2}{(2m \vee 1)^{\alpha/2}(2m+2 \vee 1)^{\alpha/2}}} \leq \frac{C'}{(2m+1 \vee 1)^{\alpha/2}}$$

for some $C' > 0$. To show the last inequality above, use the fact that there exists $C_\alpha > 0$ such that $(2m)^{\alpha/2}(2m+2)^{\alpha/2} \leq C_\alpha (2m+1)^{\alpha/2}$ (details left as exercises). Thus $(b) \implies (c)$.
Now, we show $(c) \implies (a)$. Assuming (c), observe that (by taking supremum over $x \in V$)

$$|P_n f(x)| \leq \sum_{y \in V} p_n(x,y)\,|f(y)|\,\mu_y \implies \|P_n f\|_\infty \leq \frac{C_H}{(n \vee 1)^{\alpha/2}}\,\|f\|_1$$

$$\text{and} \quad \|P_n f\|_2^2 = \langle P_n f,\ P_n f\rangle = \langle P_{2n} f,\ f\rangle \le \|P_{2n} f\|_\infty \|f\|_1 \le \frac{C_H}{(2n \vee 1)^{\alpha/2}} \|f\|_1^2 \tag{8.4}$$

Now, we make use of the following inequality - (verify!)

$$\mathcal{E}(f,f) \ge \frac{1}{2n}[\|f\|_2^2 - \|P_n f\|_2^2]$$

Using this, and (8.1), we get

$$\mathcal{E}(f,f) \ge \frac{1}{2n}\left[\|f\|_2^2 - \frac{C_H}{(2n \vee 1^{\alpha/2})}\|f\|_1^2\right]$$

WLOG, assume $\|f\|_1 = 1$, and choose smallest possible k such that

$$\frac{C_H}{(2n \vee 1)^{\alpha/2}} \le \frac{\|f\|_2^2}{2} \quad \text{so that} \quad \mathcal{E}(f,f) \ge \frac{1}{4k}\|f\|_2^2$$

Since $k \ge 1$, we have $k^{-\alpha/2} \le C^2 \|f\|_2^2$ for some $C > 0$, and hence $k^{-\alpha/2} \le C\|f\|_2$. Therefore,

$$\mathcal{E}(f,f) \ge \frac{C_2\|f\|_2^2}{\|f\|_2^{\frac{4}{\alpha}}} = C_2\|f\|_2^{2-4/\alpha} \implies (N_\alpha)$$

□

Carne-Varopoulos Bound

We begin with a few lemmas and some results involving Chebyshev polynomials.

Lemma 8.0.3. *Let $\{S_n\}_{n\ge 0}$ denote the simple symmetric random walk on $\mathbb{Z}$ with $S_0 = 0$. Then*

(a)

$$\mathbf{P}(S_n \ge D) \le \exp\left(-\frac{D^2}{2n}\right)$$

(b)

$$\mathbf{E}[\lambda^{S_n}] = \sum_{r\in\mathbb{Z}} \lambda^r \mathbf{P}(S_n = r) = 2^{-n}\sum_{r=0}^{n}\binom{n}{r}\left(\frac{1}{\lambda}\right)^{2n-r}$$

Proof. (a) was given in Worksheet 2, and (b) is trivial using results from Week 1. □

Definition 8.0.4. *(Chebyshev Polynomials) For $-1 \le t \le 1$, define*

$$H_k(t) := \frac{1}{2}(t + i\sqrt{1-t^2})^k + \frac{1}{2}(t - i\sqrt{1-t^2})^k$$

Lemma 8.0.4. *For each $k \ge 0$, we have*

(a) H_k is a real polynomial of degree k.

(b) $t^n = \sum_{k\in\mathbb{Z}} \mathbf{P}(S_n = k) H_{|k|}(t)$

Proof. To show (a), fix $t \in [-1, 1]$ and set $s = \sqrt{1-t^2}$. Observe that

$$H_k(t) = \frac{1}{2}\sum_{r=0}^{k}\binom{k}{r}t^{k-r}[(is)^r + (-is)^r] = \frac{1}{2}\sum_{r=0}^{k/2}\binom{k}{2r}t^{k-2r}\psi(s)$$

where ψ is some real function of s.
To show (b) set $z_1 = t + is$ and $z_2 = t - is$ so that $|z_1| = |z_2| = 1$ and $z1z2 = 1$. Then,

$$H_k(t) = \frac{1}{2}(z_1^k + z_2^k) = H_{-k}(t) \implies |H_k(t) \le 1|$$

Now, observe that $t = (z_1 + z_2)/2$, so that

$$t^n = \sum_{k=0}^{n}\frac{1}{2^n}\binom{n}{k}z_1^k z_2^{n-k} = \sum_{k=0}^{n}\frac{1}{2^n}\binom{n}{k}z_1^{2k-n} = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)z_1^r$$

Repeating the same arguments above, we get

$$t^n = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)z_1^r = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)z_2^r$$

$$\implies t^n = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)\left(\frac{z_1^r + z_2^r}{2}\right) = \sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)H_{|r|}(t)$$

□

Theorem 8.0.2. *(Carne-Varopoulos bound) Let (Γ, μ) be a weighted graph. Then, for every $x, y \in V$ and $n \ge 1$*

$$p_n(x, y) \le \frac{2}{\sqrt{\mu_x\mu_y}}\exp\left(-\frac{d(x,y)^2}{2n}\right)$$

Proof. Proved in Worksheet 2. □

Large Deviations for Random Walks

Let $\{\xi_i\}_{i\ge1}$ be IID $\mathbb{Z}$ valued random variables such that $\mathbf{E}[\xi_1] = \mu$ and $\mathrm{Var}[\xi_1] < \infty$. Define $S_0 = 0$ and $S_n = \sum_{i=1}^n \xi_i$. Then, the strong law of large numbers (SLLN) and the central limit theorem (CLT) respectively state that

$$\mathbf{P}\left(\lim_{n\to\infty}\frac{S_n}{n} = \mu\right) = 1 \text{ and } \frac{S_n - n\mu}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1)$$

Thus, the CLT loosely states that $S_n \approx n\mu + \sqrt{n}Z$, where $Z \sim \mathcal{N}(0,1)$.

As an exercise, show that for every $\epsilon > 0$, $\mathbf{P}(A_n^\epsilon) \to 0$ as $n \to \infty$, where $A_n^\epsilon = \{S_n \ge n(\mu+\epsilon)\}$. What is the rate of decay of $\mathbf{P}(A_n^\epsilon)$ (as $n \to \infty$)?
(Hint: $\mathbf{P}(S_n \ge n(\mu+\epsilon)) \approx \mathbf{P}(\xi_i > \mu+\epsilon \ \forall\ 1 \le i \le n) = [\mathbf{P}(\xi_1 > \mu+\epsilon)]^n \approx e^{-Cn}$ for some $C > 0$)

Theorem 8.0.3. *Let $\{\xi_i\}_{i\geq 0}$ be IID random variables with $\mathbf{P}(\xi_1 = 0) = \mathbf{P}(\xi_1 = 1) = 1/2$. Then, for every $a > 1/2$,*

$$\lim_{n\to\infty} \frac{1}{n} \log[\mathbf{P}(S_n \geq an)] = -I(a)$$

where

$$I(z) = \begin{cases} \log 2 + a\log a + (1-a)\log a & \text{if } 0 \leq z \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

Observations:

(1) Minima of $I(z)$ is achieved at $z = 1/2$, and the graph increases from $[1/2, 1]$. This implies rate of exponential decay increases as $1/2 \to a \to 1$.

(2) Symmetry of the function $I(\cdot)$ around $1/2$ suggests that for $a < 1/2$, (Requires a proof)

$$\frac{1}{n}\log[\mathbf{P}(S_n \geq an)] \to -I(a)$$

(3) The theorem implies SLLN. The idea of the proof makes use of the following inequality

$$\mathbf{P}(S_n > (1/2+\delta)n) \leq \exp\{-I(n(1/2+\delta))\}$$

Proof.
If, $a > 1$ then, since S_n can be atmost n, $\mathbf{P}(S_n > an) = 0$ so the result follows. Now, consider $\frac{1}{2} < a \leq 1$, then

$$\mathbf{P}(S_n > an) = \sum_{an<k\leq n} \mathbf{P}(S_n = k) = \sum_{an<k\leq n} \binom{n}{k}\frac{1}{2^n} = \frac{1}{2^n}\sum_{an<k\leq n}\binom{n}{k}$$

Let, $Q_n(a) = \max\limits_{an<k\leq n}\binom{n}{k}$. So, we have,

$$2^{-n}\, Q_n(a) \leq \mathbf{P}(S_n > an) \leq 2^{-n}\, Q_n(a)\,(n+1) \tag{8.5}$$

First equality follows from the fact that one summand in the $\sum\limits_{an<k\leq n}\binom{n}{k}$ attains maximum and the second equality follows since, each summand of $\sum\limits_{0\leq k\leq n}\binom{n}{k}$ is $\leq Q_n(a)$.

Claim:

For, $\frac{1}{2} < a < 1$,

$$\frac{1}{n}\log Q_n(a) \xrightarrow[n\to\infty]{} -a\log a - (1-a)\log(1-a)$$

Now, from (8.2),

$$-\log 2 + \frac{1}{n}\log Q_n(a) \leq \frac{1}{n}\log \mathbf{P}(S_n > an) \leq -\log 2 + \frac{1}{n}\log Q_n(a) + \frac{1}{n}\log(n+1) \tag{8.6}$$

assuming the claim as LHS and RHS of (8.3) goes to $-I(a)$, the result follows. We now prove the claim.

Proof of claim:

Since, $a > \frac{1}{2}$, $\max_{an<k\leq n} \binom{n}{k} = \binom{n}{\lceil an \rceil}$. Now, from stirling's approximation

$$\binom{n}{\lceil an \rceil} = \frac{n!}{\lceil an \rceil!(n - \lceil an \rceil)!} \sim \frac{n^n e^{-n}\sqrt{2\pi n}}{\lceil an \rceil^{\lceil an \rceil} e^{-\lceil an \rceil}\sqrt{2\pi \lceil an \rceil}} \cdot \frac{1}{(n-\lceil an \rceil)^{n-\lceil an \rceil} e^{n-\lceil an \rceil}\sqrt{2\pi(n-\lceil an \rceil)}}$$

For, $a > \frac{1}{2}, a < 1$; $\lceil an \rceil \to \infty$ and $n - \lceil an \rceil \to \infty$ as $n \to \infty$ (Check!) and

$$\begin{aligned}\frac{1}{n}\log Q_n(a) &\sim \frac{1}{n}\left[(n+\frac{1}{2})\log n - (\lceil an \rceil + \frac{1}{2})\log\lceil an \rceil - (n - \lceil an \rceil + \frac{1}{2})\log(n - \lceil an \rceil) - \log(\sqrt{2\pi})\right] \\ &= \log n + \frac{1}{2n}\log n - \frac{\lceil an \rceil}{n}\log\lceil an \rceil - \frac{1}{2n}\log\lceil an \rceil - \frac{1}{n}\log\sqrt{2\pi} - \frac{n-\lceil an \rceil}{n}\log(n-\lceil an \rceil) - \frac{1}{2}\log(n-\lceil an \rceil)\end{aligned}$$

the second, fourth, fifth and seventh summand of the above equation tends to 0 as n tends to ∞ and from the exercise (?) we have that

$$\frac{\lceil an \rceil}{n}\log\frac{\lceil an \rceil}{n} \xrightarrow[n\to\infty]{} a\log a \quad \text{and} \quad \frac{n-\lceil an \rceil}{n}\log\frac{n-\lceil an \rceil}{n} \xrightarrow[n\to\infty]{} (1-a)\log(1-a)$$

which proves the claim. $\square$

Cramer, 1930's

$\{\xi_i\}_{i\geq 1}$ i.i.d random variables with $\mathbf{E}[\xi_i] = \mu < \infty$, $\mathbf{E}[e^{r\xi_i}] < \infty, \forall r \in \mathbb{R}$. For any $a > \mu$,

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbf{P}(S_n > an) = -I(a)$$

where, $I(a) = \sup_{z\in\mathbb{R}}[za - \mathbf{E}[e^{z\xi}]]$

Sanov, 1961 (`Level 2 of LDP`)

$\mathbf{P}(S_n > an) = \mathbf{P}\circ S_n^{-1}((an,\infty)) \coloneqq \mu_n((an,\infty))$

$$-\frac{1}{n}\log\mu_n((an,\infty)) \xrightarrow[n\to\infty]{} \infty$$

Varadhan's LDP setup, 1960's

Let, $X_n : \Omega \to \mathbb{R}$ be a random variable of $(\Omega, \mathcal{F}, \mathbf{P})$. A be an event, $\mathbf{P}_n(A) \coloneqq \mathbf{P}(S_n \in A)$, then $\mathbf{P}(\cdot)$ is a probability on $\mathbb{R}$.

A sequence $\{\mathcal{P}_n\}_{n\geq 1}$ of probaility measures on $\mathbb{R}$ (can be any metric space (X,d)) is said to satisfy large deviation principle with rate n and rate function $I : \mathbb{R} \to [0,\infty)\cup\{\infty\}$, if

1. $I \not\equiv \infty$, I is lower-semi continuous and has compact level sets.
2. $\overline{\lim\limits_{n\to\infty}} \frac{1}{n} \log \mathbf{P}_n(\mathcal{C}) \le -I(\mathcal{C}) \,\forall\,$closed sets $\mathcal{C}$
3. $\underline{\lim\limits_{n\to\infty}} \frac{1}{n} \log \mathbf{P}_n(\mathcal{O}) \ge -I(\mathcal{O}) \,\forall\,$open sets $\mathcal{O}$

where, $A \subseteq \mathbb{R}$, $I(A) = \inf\limits_{y\in A} I(y)$.

Theorem 8.0.4. *$\{\mathcal{P}_n\}_{n\ge 1}$ satisfied LDP with rate n then , $I(\cdot)$ is unique.*

Theorem 8.0.5 (Varadhan's lemma)**.** *If, $\{\mathcal{P}_n\}_{n\ge 1}$ satisfies LDP with rate n and rate function $I(\cdot)$, let $F_n(x) = \mathbf{P}_n((-\infty, n])$ for some continuous and bounded above function $F : \mathbb{R} \to \mathbb{R}$, we have*

$$\int e^{nF(x)} dF_n(x) \xrightarrow[n\to\infty]{} \sup_{x\in\mathbb{R}} [F(x) - I(x)]$$

Applications

For, $\theta \in S^1$, $t \in \mathbb{R}$, $u : S^1 \times \mathbb{R}_+ \to \mathbb{R}$,

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2}\frac{\partial^2 u}{\partial \theta^2} + V(\theta)u \\ u(0,\theta) &= 1 \end{aligned}$$

then,

$$\frac{1}{t}\log u(t,\theta) \xrightarrow[t\to\infty]{} \lambda_1 = \sup_{f\in\cdots}\left\{\int V(\theta)f(\theta)d\theta - \frac{1}{8}\int \frac{(f'(\theta)^2)}{f(\theta)}d\theta\right\}$$

we can represent this as follows,

$$u(t,\theta) = \mathbf{E}\, e^{\int_0^t V(\theta_s)ds}, \ \{\theta_s\} - \text{brownian motion on}\, S^1$$

Exercises

1. For any $a \in \mathbb{R}$, show that,

$$\frac{\lceil an \rceil}{n} \xrightarrow[n\to\infty]{} a \ \text{ and } \ \frac{n - \lceil an \rceil}{n} \xrightarrow[n\to\infty]{} 1 - a$$

Discrete Time Martingales

Week 4 **January 27, 2023**

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

$$\mathbb{E}(Z_n | Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1) = z_{n-1} \quad \forall\ n \geq 2 \tag{7.7}$$

Things to understand- conditional expectation for discrete and conditional random variable [3]. 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 7.7?

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 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!

Week 5 **February 3, 2023**
Lecturer: Siva Athreya Scribe: Atreya Choudhury, Ankan Kar

We define $f : D \subseteq \mathbb{R}^{n-1} \to \mathbb{R}$ where

$$f(z_1, z_2, \ldots, z_{n-1}) = \mathbf{E}[Z_n | Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \ldots, Z_1 = z_1]$$

Define $Y_n : \Omega \to \mathbb{R}$ where

$$Y_n(\omega) := f(Z_1(\omega), Z_2(\omega), \ldots, Z_{n-1}(\omega)) \tag{7.8}$$

You can check that $\{Y_n\}$ is a random variable.

Property 7.0.1. *Some properties of* $\{Y_n\}$

1. $A := \{Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \ldots, Z_1 = z_1\}$

$$\omega \in A \implies Y_n(\omega) = f(z_1, z_2, \ldots, z_{n-1})$$

2. $L := \{Y_n \leq c\} = \{f(Z_1, Z_2, \ldots, Z_{n-1}) \leq c\}$

$$L \in \mathcal{A}_{n-1} \equiv \textit{observable events upto } n-1$$

(7.8) $\iff$ $\{Y_n\}$ *has the above two properties*

If $\{Z_n\}$ *is martingale,* $Y_n = Z_{n-1}$

Lemma 7.0.5. *Let* $\{Y_n\}_{n\geq 1}$ *be martingale. Then,*

$$\forall\, 1 \leq i \leq n,\ \mathbf{E}[Z_n|Z_i, Z_{i-1}, \ldots, Z_1] = Z_i$$

Proof. We fix i and prove by induction on n.
We look at n = i+1. By martingale property,

$$\mathbf{E}[Z_{i+1}|Z_i, Z_{i-1}, \ldots, Z_1] = Z_i$$

Let $k > 0$ and the statement hold for $n = i + k$. We look at $n = i + k + 1$

$$\begin{aligned}
&\mathbf{E}[Z_{i+k+1}|Z_i, Z_{i-1}, \ldots, Z_1] \\
&= \mathbf{E}[\mathbf{E}[Z_{i+k+1}|Z_{i+k}, Z_{i+k-1}, \ldots, Z_1]|Z_i, Z_{i-1}, \ldots, Z_1] \\
&= \mathbf{E}[Z_{i+k}|Z_i, Z_{i-1}, \ldots, Z_1] \qquad \text{[using (7.0.5)]} \\
&= Z_i
\end{aligned}$$

where the last equality is obtained from the induction hypothesis □

The property used in the first equality is called the Tower property. We now formally state and prove the same.

Property 7.0.2 (Tower Property)**.**

$$\mathbf{E}[\mathbf{E}[X|Y, Z]|Y] = E[X|Y]$$

Proof.

$$\mathbf{E}[\mathbf{E}[X|Y, Z]|Y] = E[h(Y, Z)|Y] = k(Y)$$

Let $y \in \mathbb{R}$ such that $\mathbf{P}(Y = y) > 0$

$$\begin{aligned}
k(y) &= E[h(Y, Z)|Y] \\
&= \sum_{\substack{m\in\mathrm{Range}(Y) \\ t\in\mathrm{Range}(Z)}} h(m, t)\mathbf{P}(Y = m, Z = t|Y = y) \\
&= \sum_{t\in\mathrm{Range}(Z)} h(y, t)\mathbf{P}(Z = t|Y = y) \\
&= \sum_{t\in\mathrm{Range}(Z)} \sum_{k\in\mathrm{Range}(X)} k\mathbf{P}(X = k|Y = y, Z = t)\mathbf{P}(Z = t|Y = y) \\
&= \sum_{t\in\mathrm{Range}(Z)} \sum_{k\in\mathrm{Range}(X)} k\frac{\mathbf{P}(X = k, Y = y, Z = t)}{\mathbf{P}(Y = y, Z = t)} \frac{\mathbf{P}(Z = t, Y = y)}{\mathbf{P}(Y = y)} \\
&= \sum_{k\in\mathrm{Range}(X)} \sum_{t\in\mathrm{Range}(Z)} k\frac{\mathbf{P}(X = k, Y = y, Z = t)}{\mathbf{P}(Y = y)} \\
&= \sum_{k\in\mathrm{Range}(X)} k\frac{\mathbf{P}(X = k, Y = y)}{\mathbf{P}(Y = y)} \\
&= \mathbf{E}[X|Y = y]
\end{aligned}$$

□

$\{Z_n\}$ is a Martingale
$E[Z_n|Z_i, Z_{i-1}, ..., Z_1] = Z_i$ where $1 \leq i \leq n$
$E[Z_n] = E[Z_1]$

7.1 Stopping time and Stopped process

Definition 7.1.1. *Let $(\Omega, A, \mathbf{P})$ be a probability space on which $\{Z_n\}_{n\geq 1}$ is defined.*
$\mathscr{A}_k$= events determined by $Z_1, Z_2, ..., Z_k$.
*$T : \Omega \longrightarrow \mathbb{N} \cup \{\infty\}$ is called a **stopping time** for $\{Z_n\}_{n\geq 1}$ if $\{T = k\} \in \mathscr{A}_k$, i.e. $\mathbf{1}_{T=k}$ = "function" of $Z_1, Z_2, ..., Z_k$.*

Definition 7.1.2. *for any stopping time T, we define the **stopped process**:*
$$Z_n^T(w) = Z_{n\wedge T(w)}(w) = \begin{cases} Z_n & \text{if } n < T \\ Z_T & \text{if } n \geq T \end{cases}$$

Theorem 7.1.1. *Given a sequence of random variables $\{Z_n\}_{n\geq 1}$ and $T : \Omega \longrightarrow \mathbb{N}\cup\{\infty\}$, a stopping time of $\{Z_n\}_{n\geq 1}$. Then $\{Z_n^T\}_{n\geq 1}$ is a martingale iff $\{Z_n\}_{n\geq 1}$ is a martingale*

Idea of the proof: $\mathbf{E}(Z_n^T|Z_{n-1}^T, ..., Z_1^T) = \mathbf{E}(Z_{n-1}^T)$
Take $Z_1 = z_1, ..., Z_{n-1} = z_{n-1} \rightarrow$ determine if T has happened by time n-1 or not
$\rightarrow$ if $T \geq n$, $Z_n^T = Z_n$
if $T < n$, $Z_n^T = z_{n-1}$ $\square$
Let $\{X_i\}, X, Y, Z$ be discrete random variables.

$$\mathbf{E}[Y|X = x_1] = \sum_{k\in Range(Y)} k\mathbf{P}(Y = k|X = x_1) \tag{7.9}$$

$$\mathbf{E}[Y|X_1 = x_1, ..., X_n = x_n] = \sum_{k\in Range(Y)} k\mathbf{P}(Y = k|X_1 = x_1, ..., X_n = x_n) \tag{7.10}$$

where $\mathbf{E}[Y|X_1 = x_1, ..., X_n = x_n] \equiv f(x_1, x_2, ..., x_n)$
$f : \prod_{i=i}^n Range(X_i) \rightarrow \mathbb{R}$

$$\mathbf{E}[Y|X_1, ..., X_n](\omega) = \sum_{x\in Range(X_i)} k\mathbf{E}(Y = k|X_1 = x_1, ..., X_n = x_n)\mathbf{1}_{(X_1=x_1,...,X_n=x_n)}(\omega) \tag{7.11}$$

where $\mathbf{E}[Y|X_1, ..., X_n] \equiv \mathbf{E}[Y|\mathscr{A}_n]$, i.e. events observable by time n.

7.2 Tower Property

Let $\mathscr{A}_n \subset \mathscr{A}_m$, $n \leq m$ then $\mathbf{E}[E[Y|\mathscr{A}_m]|\mathscr{A}_n] = \mathbf{E}[Y|\mathscr{A}_n]$

7.3 Markov property and Strong Markov Property

Property for $\{X_n\}$ random walk on (Γ, y).

$\Omega = \mathbb{V}^{\mathbb{Z}_+}$.

$X_n : \Omega \to \mathbb{V}$.

$X_n(\omega) = \omega(n)$.

$\mathscr{A}_n$= events determined by $X_1, ..., X_n$.

$\mathbf{P}^x(X_0 = x_0, X_1 = x_1, ..., X_n = x_n) = \mathbf{1}_x(x_0) \prod_{i=0}^n \mathscr{P}(x_{i-1}, x_i)$

$\mathscr{P}(x, y) = \frac{\mu_{xy}}{\mu_y}$

$\xi \to$ random variable that is determinable by $\mathscr{A}_n$ i.e. $\xi = g(X_1, X_2, ..., X_n$ for some g.

$\forall k \geq 1$, $\theta_k : \Omega \to \mathbb{V}^{\mathbb{Z}_+}$, $\theta_k(\omega) = (\omega(k), \omega(k+1), ...)$

Let $\eta : \Omega \to \mathbb{R}$ be any random variable.

$\mathbf{E}[\xi"\eta \text{ after time } n"|\mathscr{A}_n] = \mathbf{E}[\xi \mathbf{E}^{X_n}[\eta \text{ after time } n"]]$

Markov Property:

$$\mathbf{E}[(\xi) \times (\eta.\theta_n)|\mathscr{A}_n] = \mathbf{E}[\xi \mathbf{E}^{X_n}[\eta]] \tag{7.12}$$

Strong Markov Property:

T is a stopping time of $\{X_n\}_{n \geq 1}$.

$\mathscr{A}_n \equiv$events determined by time T.

if ξ is determinable by time T, then

$$\mathbf{E}[(\xi) \times (\eta.\theta_T)|\mathscr{A}_T] = \mathbf{E}[\xi \mathbf{E}^{X_T}[\eta]] \tag{7.13}$$

Bibliography

[1] Luca Avena, Markus Heydenreich, Frank den Hollander, Evgeny Verbitskiy, and Willem van Zuijlen. Random walks. *Lecture Notes from the Mathematical Institute of Leiden University in The Netherlands*, 2016.

[2] Martin T Barlow. *Random walks and heat kernels on graphs*, volume 438. Cambridge University Press, 2017.

[3] Siva Athreya, Deepayan Sarkar, and Steve Tanner. *Probability and Statistics with Examples using R*. 2016. Unfinished Book, Last Compilation April 25th 2016, available at http://www.isibang.ac.in/~athreya/psweur/index.html.