Lecture 1

LECTURER: SIVA ATHREYA SCRIBE: SRIVATSA B, ATREYA CHODHURY

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 $\mathbb{P}_N : \Omega_N \to [0, 1]$ is defined as

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

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 $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$ is called a (symmetric) simple random walk on $\mathbb{Z}$, of finite length N, starting at 0.*

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

† added illustrations

Observations.

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

Proof.

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

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. □

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

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

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

□

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

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

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

Proof. Left as an exercise. □

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

Proof. Left as an exercise. □

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

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

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

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

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

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

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

$\square$

(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.

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

$\square$

(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,

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

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 $\mathbb{E}[S_T] > 0$? 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. *Let $T : \Omega \to \{0, 1, \ldots, N\}$ be a stopping time. Then*

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

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 $\mathbb{E}[S_T]$ as

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

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

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

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

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)

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

Substituting the above values in (†) and (††), we finally have

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

□

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*

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

Theorem 2. *Let $\{V_k\}_{k=1}^{N}$ be a bet sequence. Then*

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

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

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

□

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 $1 \le k < m \le N$, $\mathbb{P}(S_m = b \mid S_k = a) = \mathbb{P}(S_{m-k} = b - a)$.
4. Show that $\mathbb{E}[S_n^2] = \sum_{i=1}^n \mathbb{E}[X_i^2]$.
5. (a) Show that for any $a, b \in \mathbb{R}$,

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

   (b) Using (a), conclude that

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

   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.