Lecture 22: Random Walks

1 Duality in Random Walks

Essentially, if X is an exchangeable sequence of random variables, then $(X_1, X_2, \cdots, X_n)$ has the same joint distribution as $(X_n, X_{n-1}, \cdots, X_1)$. In particular, an iid sequence of random variables is exchangeable.

Proposition 1.1. *Suppose $\{X_n : n \in \mathbb{N}\}$ is a sequence of iid random variables with positive mean. Let $S_n = \sum_{k=1}^{n} X_i$ be a random walk with step size $X_n z$. If*

$$N = \min\{n \in \mathbb{N} : S_n > 0\}$$

Then $\mathbb{E}[N] < \infty$.

Proof. From duality principle we obtain that

$$\{N > n\} = \{S_i \leq 0, i \in [n]\} = \left\{\sum_{k=0}^{i-1} X_{n-k} \leq 0, i \in [n]\right\} = \{S_n \leq S_{n-i}, i \in [n]\}.$$

It follows that

$$\mathbb{E}[N] = \sum_{n \in \mathbb{N}_0} \Pr\{N > n\} = \sum_{n \in \mathbb{N}_0} \Pr\{S_n \leq S_{n-i}, i \in [n]\}.$$

We define the renewal instants to be when random walk hits a new low. (Why are these renewal instants?) Hence, n is a renewal instant after 0 if $\{S_n \leq S_i : i \in [n]\}$. Hence, we have

$$\begin{aligned}\mathbb{E}[N] &= \sum_{n \in \mathbb{N}_0} \Pr\{\text{renewal happens at time } n\} = \sum_{n \in \mathbb{N}_0} \Pr\{\text{inter-renewal length} \geq n\} \\ &= 1 + \mathbb{E}[\text{Number of renewals that occur}]\end{aligned}$$

Since $\mathbb{E}X > 0$, it follows from strong law of large numbers that $S_n \to \infty$. Hence, the expected number of renewals that occur is finite. Thus $\mathbb{E}[N] < \infty$. □

Definition 1.2. The number of distinct values of $(S_0, \cdots, S_n)$ is called **range**, denoted by R_n.

Proposition 1.3.

$$\lim_{n \in \mathbb{N}} \frac{\mathbb{E}[R_n]}{n} = \Pr\{S_n \neq 0, \forall n \in \mathbb{N}\}$$

Proof. We define indicator function

$$I_k = 1_{\{S_k \neq S_{k-i}, i \in [k]\}}.$$

Then, we can write range R_n in terms of indicator I_k as

$$R_n = 1 + \sum_{k=1}^{n} I_k$$

Let $T = \{n > 0 : S_n = 0\}$. Then, $\lim_{k\in\mathbb{N}} \Pr\{T > k\} = \Pr\{S_n \neq 0, \forall n \in \mathbb{N}\}$. Further, using the duality principle, we can write

$$\mathbb{E}[R_n] = 1 + \sum_{k=1}^{n} \Pr\{S_i \neq 0, i \in [k]\} = \sum_{k=0}^{n} \Pr\{T > k\} \tag{1}$$

Result follows by dividing both sides by n and taking limits. □

Theorem 1.4 (Simple Random Walk). *For a simple random walk, where* $\Pr\{X_1 = 1\} = p$ *the following holds*

$$\lim_{n\in\mathbb{N}} \frac{\mathbb{E}[R_n]}{n} = \begin{cases} 2p - 1, & p > \frac{1}{2} \\ 2(1-p) - 1, & p \leq \frac{1}{2}. \end{cases}$$

Proof. When $p = \frac{1}{2}$, this random walk is recurrent and thus

$$\Pr\{\text{No Return to } 0\} = 0 = \lim_{n\in\mathbb{N}} \frac{\mathbb{E}[R_n]}{n}.$$

When $p > \frac{1}{2}$, let $\alpha = \Pr\{\text{return to } 0 | X_1 = 1\}$. Since $\mathbb{E}X > 0$, we know that $S_n \to \infty$ and hence $\Pr\{\text{return to } 0 | X_1 = -1\} = 1$. We can write unconditioned probability of return of random walk to 0 as

$$\Pr\{\text{Return to } 0\} = \alpha p + 1 - p.$$

Conditioning on X_2 yields

$$\Pr\{S_n = 0 \text{ for some } n | X_1 = 1\} = p \Pr\{S_n = 0 \text{ for some } n | S_2 = 2\} + (1 - p).$$

Further noticing that

$$\Pr\{S_n = 0 \text{ for some } n | S_2 = 2\} = \alpha \Pr\{S_{n+m} = 0 \text{ for some } n | S_m = 1 \text{ for some } m\},$$

we conclude $\alpha = \alpha^2 p + 1 - p$. Solving for α yields $\alpha = \frac{1-p}{p}$, and hence the result follows. We can show similarly for the case when $p < 1/2$. □

Proposition 1.5. *In the symmetric random walk, the expected number of visits to state k before returning to origin is equal to 1 for all $k \neq 0$.*

Proof. For $k > 0$, let N_j be the hitting time to state j for random walk S_n. Further, let Y denote the number of visits to state k before the first return to origin. That is,

$$Y = \sum_{n=1}^{\infty} I_n,$$

where $I_n = 1_{\{S_k=n, N_0>n\}}$. Thus, using duality principle and recurrence of symmetric random walk, we can write

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{n=1}^{\infty} \Pr\{S_i > 0, i \in [n], S_n = k\} \\ &= \sum_{n=1}^{\infty} \Pr\{S_n - S_{n-i} > 0, i \in [n], S_n = k\} \\ &= \sum_{n=1}^{\infty} \Pr\{N_k = n\} = \Pr\{S_n = k \text{ for some } n\} = 1.\end{aligned}$$

□

1.1 GI/GI/1 Queueing Model

Consider a $GI/GI/1$ queue. Customers arrive in accordance with a renewal process having an arbitrary interarrival distribution F, and the service distribution is G.

Proposition 1.6. *Let D_n be the delay in the queue of the n^{th} customer in a GI/GI/1 queue with independent inter-arrival times X_n and service times Y_n. We also define a random walk S_n with steps $U_n = Y_n - X_{n+1}$ for all $n \in \mathbb{N}$. Then, we can write*

$$\Pr\{D_{n+1} \geq c\} = \Pr\{S_j \geq c, \text{ for some } j \in [n]\}. \tag{2}$$

Proof. The following recursion for D_n is easy to verify

$$D_{n+1} = (D_n + Y_n - X_{n+1})1_{\{D_n+Y_n-X_{n+1}\geq 0\}} = \max\{0, D_n + U_n\}.$$

Iterating the above relation with $D_1 = 0$ yields

$$\begin{aligned}D_{n+1} &= \max\{0, U_n + \max\{0, D_{n-1} + U_{n-1}\}\} \\ &= \max\{0, U_n, U_n + U_{n-1} + D_{n-1}\}.\end{aligned}$$

We can define a random walk S_n with steps U_n to write

$$D_{n+1} = \max\{0, S_n - S_{n-1}, S_n - S_{n-2}, \dots, S_n - S_0\}.$$

Using the duality principle, we can rewrite delay as

$$D_{n+1} = \max\{0, S_1, S_2, \dots, S_n\}.$$

□

Corollary 1.7. *If $\mathbb{E}U_n \geq 0$, then for all c, we have $\Pr\{D_\infty \geq c\} \triangleq \lim_{n\in\mathbb{N}} \Pr\{D_n \geq c\} = 1$.*

Proof. It follows from Proposition 1.6 that $\Pr\{D_{n+1} \geq c\}$ is nondecreasing in n. Hence, by MCT the limit exists and is denoted by $\Pr\{D_\infty \geq c\} = \lim_{n\in\mathbb{N}} \Pr\{D_n \geq c\}$. Therefore, by continuity of probability, we have from (2), that

$$\Pr\{D_\infty \geq c\} = \Pr\{S_n \geq c \text{ for some } n\}. \tag{3}$$

If $E[U_n] = E[Y_n] - E[X_{n+1}]$ is positive, then by strong law of large numbers the random walk S_n will converge to positive infinity with probability 1. The above will also be true when $E[U_n] = 0$, then the random walk is recurrent. □

Remark 1.8. Hence, we get that $E[Y_n] < E[X_{n+1}]$ implies the existence of a stationary distribution.

Proposition 1.9 (Spitzer's Identity). *Let $M_n = \max\{0, S_1, S_2, \dots, S_n\}$ for $n \in \mathbb{N}$, then*

$$\mathbb{E}M_n = \sum_{k=1}^{n} \frac{1}{k}\mathbb{E}S_k^+.$$

Proof. We can decompose M_n as

$$M_n = 1_{\{S_n>0\}}M_n + 1_{\{S_n\le 0\}}M_n.$$

We can rewrite first term in decomposition as,

$$1_{\{S_n>0\}}M_n = 1_{\{S_n>0\}}\max_{i\in[n]} S_i = 1_{\{S_n>0\}}(X_1 + \max\{0, S_2 - S_1, \dots, S_n - S_1\})$$

Hence, taking expectation and using exchangeability, we get

$$\mathbb{E}1_{\{S_n>0\}}M_n = \mathbb{E}1_{\{S_n>0\}}X_1 + \mathbb{E}1_{\{S_n>0\}}M_{n-1}.$$

Since X_i, S_n has the same joint distribution for all i,

$$\mathbb{E}S_n^+ = \mathbb{E}[S_n 1_{\{S_n>0\}}] = \mathbb{E}\sum_{i=1}^{n} X_i 1_{\{S_n>0\}}] = n\mathbb{E}[X_1 1_{\{S_n>0\}}].$$

Therefore, it follows that

$$\mathbb{E}[1_{\{S_n>0\}}M_n] = \mathbb{E}[1_{\{S_n>0\}}M_{n-1}] + \frac{1}{n}\mathbb{E}[S_n^+].$$

Also, $S_n \le 0$ implies that $M_n = M_{n-1}$, it follows that

$$1_{\{S_n\le 0\}}M_n = 1_{\{S_n\le 0\}}M_{n-1}.$$

Thus, we obtain the following recursion,

$$\mathbb{E}[M_n] = \mathbb{E}[M_{n-1}] + \frac{1}{n}\mathbb{E}[S_n^+].$$

Result follow from the fact that $M_1 = S_1^+$. ☐

Remark 1.10. Since $D_{n+1} = M_n$, we have $\mathbb{E}[D_{n+1}] = \mathbb{E}[M_n] = \sum_{k=1}^{n} \frac{1}{k} E[S_k^+]$.

2 Martingales for Random Walks

Proposition 2.1. *A random walk S_n with step size $X_n \in [-M, M] \cap \mathbb{Z}$ for some finite M is a recurrent DTMC iff $\mathbb{E}X = 0$.*

Proof. If $\mathbb{E}X \neq 0$, the random walk is clearly transient since, it will diverge to $\pm\infty$ depending on the sign of $\mathbb{E}X$. Conversely, if $\mathbb{E}X = 0$, then S_n is a martingale. Assume that the process starts in state i. We define

$$A = \{-M, -M+1, \cdots, -2, -1\}, \qquad A_j = j + [M],\ j > i.$$

Let N denote the hitting time to A or A_j by random walk S_n. Since N is a stopping time, by optional stopping theorem, we have

$$\mathbb{E}_i[S_N] = \mathbb{E}_i[S_0] = i.$$

Thus we have

$$i = \mathbb{E}_i[S_N] \geq -M\mathbb{P}_i\{S_N \in A\} + j(1 - \mathbb{P}_i\{S_N \in A_j\}).$$

Rearranging this, we get a bound on probability of random walk S_n hitting A over A_j as

$$\mathbb{P}_i\{S_n \in A \text{ for some } n\} \geq \mathbb{P}_i\{S_N \in A\} \geq \frac{j-i}{j+M}.$$

Taking limit $j \to \infty$, we see that for any $i \geq 0$, we have $\mathbb{P}_i\{S_n \in A \text{ for some } n\} = 1$. Similarly, taking $B = \{1, 2, \cdots, M\}$, we can show that for any $i \geq 0$, $\mathbb{P}_i\{S_n \in B \text{ for some } n\} = 1$. Result follows from combining the above two arguments to see that for any $i \geq 0$,

$$\mathbb{P}_i\{S_n \in A \cup B \text{ for some } n\} = 1.$$

□

Proposition 2.2. *Consider a random walk S_n with mean step size $\mathbb{E}[X] \neq 0$. For $A, B > 0$, let P_A denote the probability that the walk hits a value greater than A before it hits a value less than $-B$. Then,*

$$P_A \approx \frac{1 - e^{-\theta B}}{e^{\theta A} - e^{\theta B}}.$$

Approximation is an equality when step size is unity and A and B are integer valued.

Proof. Now For $A, B > 0$, we wish to compute the probability P_A that the walk hits at least A before it hits a value $\leq -B$. Let $\theta \neq 0$ s.t

$$E[e^{\theta X}] = 1$$

Now let $Z_n = e^{\theta S_n}$. We can see that Z_n is a martingale with mean 1. Define N as

$$N = \min\{S_n \geq A \text{ or } S_n \leq -B\}$$

From Doob's Theorem, $E[e^{S_N}] = 1$. Thus we get

$$1 = E[e^{\theta S_N}|S_N \geq A]P_A + E[e^{\theta S_N}|S_N \leq -B](1 - P_A)$$

We can obtain an approximation for P_A byneglecting the overshoots past A or $-B$. Thus we get

$$E[e^{\theta S_N}|S_N \geq A] \approx e^{\theta A}$$
$$E[e^{\theta S_N}|S_N \leq -B] \approx e^{-\theta B}$$

Hence we get, □

As an assignment, show that

$$E[N] \approx \frac{AP_A - B(1-P_A)}{E[X]}$$

Example 2.3. Gambler Ruin Consider a simple random walk with probability of increment $= p$. As an exercise, show that $E\left[(q/p)^X\right] = 1$ and thus $e^\theta = q/p$. If A and B are integers, then there is no overshoot and hence, our approximations are exact. Thus

$$P_A = \frac{(q/p)^B - 1}{(q/p)^{A+B} - 1}$$

Suppose $E[X] < 0$ and we wish to know if the random walk ever crosses A. Then

$$\begin{aligned} 1 = & E[e^{\theta S_N} | S_N \geq A] P[\text{process crossed } A \text{ before } -B] \\ & + E[e^{\theta S_N} | S_N \leq -B] P[\text{process crossed } -B \text{ before } A] \end{aligned}$$

Now $E[X] < 0$ implies $\theta > 0$ (Why?). Hence we have

$$1 \geq e^{\theta A} P[\text{process crossed } A \text{ before } -B]$$

Taking B to ∞ yields

$$P[\text{Random walk ever crosses A}] \leq e^{-\theta A}$$

3 Application to G/G/1 Queues and Ruin

3.1 The G/G/1 Queue

For the G/G/1 queue, the limiting distribution of delay is

$$P[D_\infty \geq A] = P[S_n \geq A \text{ for some } n]$$

where

$$S_n = \sum_{k=1}^{n} U_k, \quad U_k = Y_k - X_{k+1}$$

Here Y_i is the service time of the ith customer and X_i is the interarrival duration between customer $i-1$ and customer i. Thus when $E[U] = E[Y] - E[X] < 0$, letting $\theta > 0$ such that

$$E[e^{\theta U}] = E[e^{\theta(Y-X)}] = 1$$

We get

$$P[D_\infty \geq A] \leq e^{-\theta A}$$

Now the exact distribution of D_∞ can be calculated when services are exponential. Hence assume $Y_i \sim exp(\mu)$. Once again,

$$\begin{aligned} 1 = & E[e^{\theta S_N} | S_N \geq A] P[S_n \text{ crossed } A \text{ before } -B] \\ & + E[e^{\theta S_N} | S_N \leq -B] P[S_n \text{ crossed } -B \text{ before } A] \end{aligned}$$

Let us compute $E[e^{\theta S_N}|S_N \geq A]$ first. Let us condition this on $N = n$ and $X_{n+1} - \sum_{i=1}^{n-1}(Y_i - X_{i+1}) = c$. By the memoryless property, the conditional distribution of Y_n given $Y_n > c + A$ is just $c + A$ plus an exponential with rate μ. Thus we get

$$\begin{aligned} E[e^{\theta S_N}|S_N \geq A] &= E[e^{\theta(A+Y)}] \\ &= \frac{\mu e^{\theta A}}{\mu - \theta} \end{aligned}$$

Now substituting back, we get

$$\begin{aligned} 1 = &\frac{\mu e^{\theta A}}{\mu - \theta} P[S_n \text{ crossed } A \text{ before } -B] \\ &+ E[e^{\theta S_N}|S_N \leq -B] P[S_n \text{ crossed } -B \text{ before } A] \end{aligned}$$

Now as $\theta > 0$, let $B \to \infty$ to get

$$1 = \frac{\mu e^{\theta A}}{\mu - \theta} P[S_n \text{ ever crosses } A]$$

And hence

$$P[D_\infty \geq A] = \frac{\mu - \theta}{\mu} e^{-\theta A}$$

3.2 A Ruin Problem

Suppose claims made to an insurance company follow a renewal process with iid interarrival times $\{X_i\}$. Let the values of the claims also be iid and independent of the renewal process $N(t)$ of their occurence. Let Y_i be the ith claim value. Thus the total value of claims till time t is $\sum_{k=1}^{N(t)} Y_i$. Now let us suppose the insurance company receives money at constant rate c per unit time, $c > 0$. We wish to compute the probability of the insurance company, starting with capital A, will eventually be wiped out or **ruined**. Thus we require

$$p = P\left\{\sum_{k=1}^{N(t)} Y_i > ct + A \text{ for some } t \geq 0\right\}$$

As an assignment, show that the company will be ruined if $E[Y] \geq cE[X]$. So let us assume that $E[Y] < cE[X]$. Also the ruin occurs when a claim is made. After the nth claim, the company's fortune is

$$A + c\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} Y_k$$

Letting $S_n = \sum_{k=1}^{n} Y_i - cX_i$ and $p(A) = P[S_n > A \text{ for some } n]$. As S_n is a random walk, we see that

$$p(A) = P[D_\infty > A]$$

Now the results from the G/G/1 queue apply.

4 Blackwell Theorem on the Line

Let S_n denote a random walk where $0 < \mu = E[X] < \infty$. Let

$$U(t) = \#\{n : S_n \le t\} = \sum_{n=1}^{\infty} I_n$$

Where $I_n = 1$ if $S_n \le t$ and zero else. Observe that if X_n are nonnegative, then $U(t) = N(t)$. Let $u(t) = E[U(t)]$. Now we prove an analog of Blackwell Renewal Theorem.

Theorem 4.1. ***(Blackwell renewal theorem)*** *If $\mu > 0$ and X_i are not lattice, then*

$$u(t + a) - u(t) \to a/\mu \quad t \to \infty \quad \textit{for } a > 0$$

Let us define a few concepts. We say an **ascending ladder variable of ladder height** S_n occurs at time n when

$$S_n > \max(S_0, S_1, \cdots, S_{n-1})$$

where $S_0 = 0$. We may deduce that since X_i are iid random variables, then the random variables $(N_i, S_{N_i} - S_{N_{i-1}})$ are iid; where N_i denotes the time between the $(i-1)$th and ith random variable. We may analogously define descending ladder variables. Now let $p(p_*)$ denote the probability of ever achieving an ascending/descending ladder variable.

$$p = P\{S_n > 0 \text{ for some } n\}, \quad p_* = P\{S_n < 0 \text{ for some } n\}$$

At each ascension/descension there is a probability p (resp p_*) of achieving another one. Hence the number of ascensions/descensions is geometrically distributed. The number of ascending ladder variables (ascensions) will have finite mean iff $p < 1$. Now as $E[X] > 0$, by SLLN, we deduce that $w.p.1$, there will be infinitely many ascending ladder variables but finitely many descending ones. That is $p = 1$ and $p_* < 1$.

Proof. The successive ascending ladder heights are a renewal process. Let $Y(t)$ be the excess time. Now given the value of $Y(t)$, the distribution of $U(t+a)-U(t)$ is independent of t. (Why?). Hence let us denote

$$E[U(t + a) - U(t)|Y(t)] = g(Y(t))$$

for some function g. Now taking expectations yields

$$u(t + a) - u(t) = E[g(Y(t))]$$

Now since $Y(t) \to^d Y_\infty$ where Y_∞ has the equilibrium distribution, we have $E[g(Y(t))] \to E[g(Y_\infty)]$. The result would be true if we show g is continuous and bounded. We leave that as an exercise. For now, we deduce that the limit exists. Let

$$h(a) = \lim_{t\to\infty} u(t + a) - u(t)$$

This also implies $h(a + b) = h(a) + h(b)$. Thus for some constant c,

$$h(a) = ca$$

Now to get c, let N_t denote the first n for which $S_n > t$. If X_i are upper bounded by M, then

$$t < \sum_{i=1}^{N_t} X_i \leq t + M$$

Taking expectations, and using Wald's Lemma, yields

$$t < E[N_t]\mu \leq t + M$$

Thus

$$\frac{E[N_t]}{t} \to \frac{1}{\mu}$$

If X_i are unbounded, use the truncation arguments done while proving Elementary renewal theorem. Now $U(t)$ can be expressed as

$$U(t) = N_t - 1 + N_t^*$$

where N_t^* is the number of times $S_n \leq t$ after having crossed t. Since N_t^* is not greater than the number of points occuring after N_t when the random walk is less than S_{N_t}, we get

$$E[N_t^*] \leq E[\text{number of } n \text{ such that } S_n < 0]$$

Hence if we argue that RHS of above is finite, then

$$\frac{u(t)}{t} \to \frac{1}{\mu}$$

From the first proposition in Random walks, we have $E[N] < \infty$ where N is the first value of n for which $S_n > 0$. At time N, with positive probability $1 - p^*$, no future value of random walk will fall below S_N. Thus,

$$E[\text{number of } n \text{ where } S_n < 0] \leq \frac{E[N|X_1 < 0]}{1 - p^*} < \infty$$

Now follow the steps illustrated in the Blackwell renewal theorem (original) proof to arrive at the desired result.

□