STATS 217: Introduction to Stochastic Processes I

Lecture 2

Recall from last time

- Given integers $A > 0, B > 0$, let

$$\tau := \min\{n \geq 0 : S_n = A \text{ or } S_n = -B\}.$$

- For $-B \leq k \leq A$, define

$$g(k) := \mathbb{E}[\tau \mid S_0 = k].$$

- Clearly, $g(-B) = 0, g(A) = 0$.
- For $-B < k < A$, we have

$$\begin{aligned} g(k) &= \frac{1}{2}\mathbb{E}[\tau \mid S_0 = k, X_1 = 1] + \frac{1}{2}\mathbb{E}[\tau \mid S_0 = k, X_1 = -1] \\ &= \frac{1}{2}\left(g(k+1) + 1\right) + \frac{1}{2}\left(g(k-1) + 1\right) \\ &= \frac{1}{2}g(k+1) + \frac{1}{2}g(k-1) + 1. \end{aligned}$$

First step analysis

- Let $(\Delta h)(k) := h(k+1) - h(k)$.
- Then, for all $-B < k < A$

$$\begin{aligned}(\Delta(\Delta g))(k-1) &= (\Delta g)(k) - (\Delta g)(k-1) \\ &= g(k+1) - g(k) - g(k) + g(k-1) \\ &= g(k+1) - (g(k+1) + g(k-1) + 2) + g(k-1) \\ &= -2.\end{aligned}$$

- "Second derivative of g is -2" so $g(k) = -k^2 + Dk + C$.
- Using boundary conditions,

$$g(k) = -(k-A)(k+B).$$

First step analysis

Therefore,

$$g(k) = \mathbb{E}[\tau \mid S_0 = k] = -(k + A)(k - B).$$

- **Answer 3:** $A = 200, B = 100,\ g(0) = 2 \times 10^4$.
- **Answer 4 (ii):** "$A = \infty$", $B = 100$, $g(0) = \infty$.
- Formally, let

$$\tau_1 = \min\{n \geq 0 : S_n = -100\},$$
$$\tau_2(\ell) = \min\{n \geq 0 : S_n = -100 \text{ or } S_n = \ell\} \quad \forall \ell \geq 1.$$

- Then, for all $\ell \geq 1$, $\tau_2(\ell) \leq \tau_1$ so that

$$100\ell = \mathbb{E}[\tau_2(\ell) \mid S_0 = 0] \leq \mathbb{E}[\tau_1 \mid S_0 = 0],$$

  and now take $\ell \to \infty$.

First step analysis

- In words, for a symmetric simple random walk starting at 0, the expected time to hit -100 is infinite! Of course, there is nothing special about -100 here.

- On the other hand, **Answer 4(i):**

$$\begin{aligned}\mathbb{P}[S_n \text{ visits } -100] &\geq \mathbb{P}[S_{\tau_2(\ell)} = -100] \\ &= \frac{\ell}{100+\ell} \\ &\to 1 \text{ as } \ell \to \infty.\end{aligned}$$

- So, a symmetric simple random walk starting at 0 visits -100 with probability 1. Again, there is nothing special about -100 here.

Summary

We have studied some aspects of the Gambler's Ruin.

- What is the probability that a symmetric simple random walk started from 0 hits 2 before -1? We saw that this is $1/3$.
- What is the expectation of the first time when the walk hits either 2 or -1? We saw that this is 2.
- Moreover, we saw that the while the probability of hitting 1 is 1, the expectation of the first time we hit 1 is infinite.

Path counting and applications

Today, we will develop tools that allow us to answer questions like the following:

- What is the probability that the first time we hit 1 is exactly 101 steps?
- What is the probability that the random walk stays non-negative for the first 2020 steps?
- What is the probability that the maximum value of the first 2020 steps of the random walk is 10?
- ...and more!

Path counting

We will need the following notation:

- $N_n(a, b)$ = number of paths from a to b with n steps.

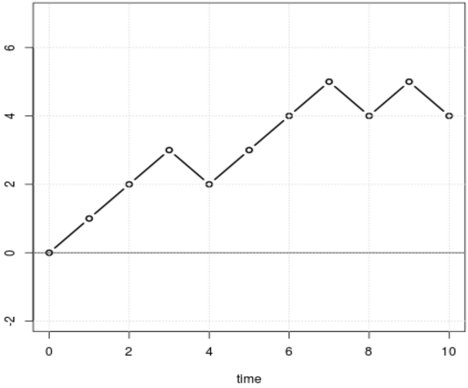

Image courtesy www.isical.ac.in

Path counting

- $N_n^0(a,b)$ =number of paths from a to b with n steps that visit 0 after time 1.
- $N_n^{\neq 0}(a,b)$ =number of paths from a to b with n steps that do not visit 0 at times $1, 2, \ldots, n-1$.

Note the following direct consequences of the definitions.

- $N_n(a,b) = N_n^{\neq 0}(a,b) + N_n^0(a,b)$.
- Also, $N_n(a,b) = N_n^0(a,b)$ if a and b have different signs.

Path counting

Let us compute $N_n(a, b)$.

- Let u denote the number of $+1$ steps and d denote the number of -1 steps.
- Since the path has n steps, we must have $u + d = n$.
- Since the path goes from a to b, we must have $u - d = b - a$.
- Hence, $u = (n + b - a)/2$ so that

$$N_n(a, b) = \binom{n}{(n + b - a)/2},$$

with the convention that $\binom{n}{r} = 0$ if r is not an integer.

Reflection principle

For any $a > 0$ and $b > 0$,

- $N_n^0(a, b) = N_n(-a, b)$.

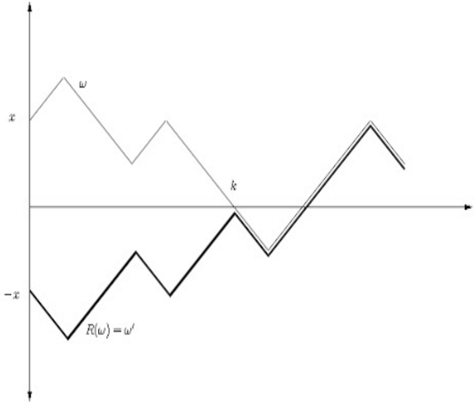

Image courtesy www.tricki.org

Reflection principle

For any $a > 0$ and $b > 0$,

- $N_n^0(a, b) = N_n(-a, b)$.
- So, $N_n^{\neq 0}(a, b) = N_n(a, b) - N_n(-a, b)$.

The point is that we already have a formula for the expressions on the right hand side.

Return time to 0

- Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0.
- Let $\tau_0 := \inf\{n \geq 1 : S_n = 0\}$.
- What is the pmf of τ_0?
- Observe that the support of τ_0 consists of even natural numbers.
- Moreover, for any $k \geq 1$

$$\mathbb{P}[\tau_0 = 2k] = N_{2k}^{\neq 0}(0,0) \cdot 2^{-2k}.$$

Return time to 0

To compute $N_{2k}^{\neq 0}(0,0)$, we can use the reflection principle.

$$
\begin{aligned}
N_{2k}^{\neq 0}(0,0) &= N_{2k-1}^{\neq 0}(1,0) + N_{2k-1}^{\neq 0}(-1,0) \\
&= 2N_{2k-1}^{\neq 0}(1,0) \\
&= 2N_{2k-2}^{\neq 0}(1,1) \\
&= 2(N_{2k-2}(1,1) - N_{2k-2}^{0}(1,1)) \\
&= 2(N_{2k-2}(1,1) - N_{2k-2}(-1,1)) \\
&= 2\left(\binom{2k-2}{k-1} - \binom{2k-2}{k}\right).
\end{aligned}
$$

Return time to 0

- Simplifying the arithmetic, we get that

$$N_{2k}^{\neq 0}(0,0) = \frac{1}{2k-1}\binom{2k}{k}.$$

- Hence,

$$\begin{aligned}\mathbb{P}[\tau_0 = 2k] &= \frac{1}{2k-1}\binom{2k}{k}2^{-2k} \\ &= \frac{1}{2k-1}\mathbb{P}[S_{2k} = 0].\end{aligned}$$

The Ballot Problem

- Consider an election with two candidates A and B.
- Suppose that a votes have been cast for A and b votes have been cast for b where $a > b$.
- After the votes have been cast, they are counted in a uniformly random order.
- Since $a > b$, after all the votes are counted, A emerges as the winner.
- What is the probability that A leads B throughout the count?

The Ballot Problem

- For $0 \leq i \leq a+b$, let S_i denote the lead of A after i votes have been counted.
- Hence, $S_0 = 0$ and $S_{a+b} = a - b$.
- Since the votes are counted in a uniformly random order, the sequence $S_0, S_1, \ldots, S_{a+b}$ is a uniformly random path from 0 to $a - b$.
- Therefore,

$$\mathbb{P}[A \text{ leads throughout}] = \frac{N_{a+b}^{\neq 0}(0, a-b)}{N_{a+b}(0, a-b)}.$$

- So, it only remains to compute $N_{a+b}^{\neq 0}(0, a-b)$.

The Ballot Problem

We need to compute $N^{\neq 0}_{a+b}(0, a-b)$.

$$
\begin{aligned}
N^{\neq 0}_{a+b}(0, a-b) &= N^{\neq 0}_{a+b-1}(1, a-b) \\
&= N_{a+b-1}(1, a-b) - N^{0}_{a+b-1}(1, a-b) \\
&= N_{a+b-1}(1, a-b) - N_{a+b-1}(-1, a-b) \\
&= \binom{a+b-1}{a-1} - \binom{a+b-1}{a} \\
&= \frac{a-b}{a+b} \cdot N_{a+b}(0, a-b).
\end{aligned}
$$

Hence,

$$\mathbb{P}[A \text{ leads throughout}] = \frac{a-b}{a+b}.$$

The Ballot Problem

- One way to reinterpret the conclusion of the Ballot problem is that for any $a > b \geq 0$ and for a simple symmetric random walk starting from $S_0 = 0$,

$$\mathbb{P}[S_i > 0 \quad \forall i = 1, \ldots, a+b-1 \mid S_{a+b} = a-b] = \frac{a-b}{a+b}.$$

- Rewritten in more convenient notation, for any integers $k, n > 0$,

$$\mathbb{P}[S_1 > 0, \ldots, S_{n-1} > 0, S_n = k] = \frac{k}{n} \cdot \mathbb{P}[S_n = k].$$

- On the homework, you will explore variants of this.