University of Washington Math 523A Lecture 4

LECTURER: YUVAL PERES

Friday, April 10, 2009

1 Problems

Problem 1: In a sequence of fair coin tosses, find $\mathbb{P}(\tau_{001} < \tau_{011})$, where τ_w is the hitting time of the word w.

For example, in the sequence $010111001\ldots$, $\tau_{011} = 5$ and $\tau_{001} = 9$, so the complement of the above event occurs. Note that although the probability of any given sequence of length 3 occurring at a particular location is 1/8, the probability that one sequence occurs before another is not necessarily 1/2, but depends on the particular pair of sequences.

2 Basic results on hitting times for SRW

Last time we used the Optional Stopping Theorem to show that for a simple random walk $\{S_t\}$ on $\mathbb{Z}$, we have

$$\mathbb{P}_k[\tau_n < \tau_0] = k/n \quad \text{and} \quad \mathbb{E}_k\tau_{\{0,n\}} = k(n-k).$$

Transforming $[0, n]$ to a general interval, if $a, b, x \in \mathbb{Z}$ with $x \in [a, b]$, then

$$\mathbb{E}_x\tau_{\{a,b\}} = (x-a)(b-x). \tag{2.1}$$

For example,

$$\mathbb{E}_0\tau_{\{-n,n\}} = n^2.$$

Thus, for a simple random walk $\{Y_t\}$ on $[0, \infty)$, we have

$$\mathbb{E}_0\tau_n = n^2$$

because $\{Y_t\}$ has the same distribution as $\{|S_t|\}$, and clearly $\tau_n(|S_t|) = \tau_{\{-n,n\}}(S_t)$. (The same result is also true for a simple random walk $\{Y'_t\}$ on $[0, n]$ since the distribution of $\{Y'_t\}$ is the same as that of $\{|S_t|\}$ until the time $\tau_n(Y'_t) = \inf\{t : Y'_t = n\}$.)

Again applying (2.1), we see that

$$\mathbb{E}^S_k\tau_{\{-n,n\}} = (n+k)(n-k) = n^2 - k^2,$$

so

$$\mathbb{E}_k^Y \tau_n = n^2 - k^2. \tag{2.2}$$

(Here the superscripts indicate whether the expectation is for the process $S = \{S_t\}_{t\geq 0}$ or $Y = \{Y_t\}_{t\geq 0}$.) Another way to think about this result is as follows. Let Y^x denote the process Y started from x, and let $k \in [0, n]$. In order for Y^0 to hit n, it must first go from 0 to k, then go from k to n. Therefore we should have

$$\tau_n(Y^0) \stackrel{d}{=} \tau_k(Y^0) + \tau_n(\tilde{Y}^k),$$

where Y^0 and $\tilde{Y}^k$ are independent SRW's on $[0, \infty)$. Taking expectations, we (formally) get

$$\begin{aligned} \mathbb{E}_0 \tau_n &= \mathbb{E}_0 \tau_k + \mathbb{E}_k \tau_n \\ n^2 &= k^2 + \mathbb{E}_k \tau_n, \end{aligned}$$

which gives (2.2). To make this argument rigorous, we need to use the Strong Markov Property.

3 More on SRW hitting times via a new martingale

Here is a result related to (2.2):

Proposition 3.1. *For a simple random walk $Y = \{Y_t\}_{t\geq 0}$ on $[0, \infty)$, we have*

$$\mathbb{E}_k[\tau_n \mid \tau_n < \tau_0] = \frac{n^2 - k^2}{3}.$$

Note: Proposition 3.1 remains true if Y is replaced by S. (Why?)

Proposition 3.1 was first proved by doing a long computation, but is there a nice way to see that the answer should be one third of the expectation in (2.2)? We'll give a proof using martingales, but what martingale should we use? We've already seen the linear martingale S_t and the quadratic martingale $S_t^2 - t$, but these turn out to be insufficient. We'll use the third degree martingale,

$$M_t = S_t^3 - 3tS_t.$$

We'll see shortly how to obtain such formulas for any degree, but for now we simply check that M_t is a martingale.

Recall that $S_{t+1} = S_t + X_{t+1}$, where $\{X_t\}_{t>0}$ are IID with

$$X_t \sim \begin{cases} +1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2. \end{cases}$$

Taking $\mathcal{F}_t = \sigma\{S_0, \ldots, S_t\} = \sigma\{S_0, X_1, \ldots, X_t\}$, we note that X_{t+1} is independent of $\mathcal{F}_t$, so

$$\begin{aligned} \mathbb{E}[S_{t+1}^3 \mid \mathcal{F}_t] &= \mathbb{E}[S_t^3 + 3S_t^2 X_{t+1} + 3S_t X_{t+1}^2 + X_{t+1}^3 \mid \mathcal{F}_t] \\ &= S_t^3 + 3S_t^2 \mathbb{E}X_{t+1} + 3S_t \mathbb{E}X_{t+1}^2 + \mathbb{E}X_{t+1}^3 \\ &= S_t^3 + 0 + 3S_t + 0, \end{aligned}$$

and since S_t is an $\mathcal{F}_t$-martingale,

$$\mathbb{E}[(t+1)S_{t+1} \mid \mathcal{F}_t] = (t+1)S_t.$$

Therefore, M is adapted to the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\geq 0}$ (because M_t is defined in terms of $S_t \in \mathcal{F}_t$), and

$$\begin{aligned}
\mathbb{E}[M_{t+1} \mid \mathcal{F}_t] &= \mathbb{E}[S_{t+1}^3 - 3(t+1)S_{t+1} \mid \mathcal{F}_t] \\
&= S_t^3 + 3S_t - 3tS_t - 3S_t \\
&= S_t^3 - 3tS_t = M_t,
\end{aligned}$$

so M is an $\mathcal{F}$-martingale.

Proof of Proposition 3.1. We want to apply the Optional Stopping Theorem to M_t, so we need to verify one of the hypotheses under which it's valid. We have

$$\sup_{j\leq \tau_{\{0,n\}}} |M_j| \leq n^3 + 3n\tau_{\{0,n\}},$$

and the righthand side is integrable, so we're good to go. Starting with $S_0 = k$ we have $M_0 = k^3$, so

$$\begin{aligned}
k^3 = \mathbb{E}_k M_0 &= \mathbb{E}_k M_{\tau_{\{0,n\}}} \\
&= \mathbb{P}_k(\tau_0 < \tau_n) \cdot \mathbb{E}_k\big[M_{\tau_{\{0,n\}}} \mid \tau_0 < \tau_n\big] + \mathbb{P}_k(\tau_n < \tau_0) \cdot \mathbb{E}_k\big[M_{\tau_{\{0,n\}}} \mid \tau_n < \tau_0\big] \\
&= \mathbb{P}_k(\tau_0 < \tau_n) \cdot \mathbb{E}_k\big[0 \mid \tau_0 < \tau_n\big] + \mathbb{P}_k(\tau_n < \tau_0) \cdot \mathbb{E}_k\big[n^3 - 3n\tau_n \mid \tau_n < \tau_0\big] \\
&= 0 + \frac{k}{n} \cdot \left(n^3 - 3n \cdot \mathbb{E}_k\big[\tau_n \mid \tau_n < \tau_0\big]\right).
\end{aligned}$$

Therefore,

$$k^2 = n^2 - 3 \cdot \mathbb{E}_k\big[\tau_n \mid \tau_n < \tau_0\big],$$

which proves the theorem. □

This martingale approach is much cleaner than a more direct, computational approach such as induction. Why would we think to use the 3rd degree martingale M_t in the first place? A general principle is that it's a good idea to work with the *exponential martingale*, which contains all the polynomial martingales such as the ones we've seen so far.

4 The exponential martingale

Take a general random walk on $\mathbb{R}$, i.e. let $\{X_j\}_{j\in\mathbb{N}}$ be IID, and define $S_t = \sum_{j=1}^t X_j$. (We can also add S_0, but first take $S_0 = 0$ to simplify things.) For the induced filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\geq 0}$, we have for any $\lambda \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda S_{t+1}} \mid \mathcal{F}_t\right] &= \mathbb{E}\left[e^{\lambda S_t} e^{\lambda X_{t+1}} \mid \mathcal{F}_t\right] \\
&= e^{\lambda S_t} \mathbb{E}\left[e^{\lambda X_{t+1}} \mid \mathcal{F}_t\right] \\
&= e^{\lambda S_t} \psi(\lambda),
\end{aligned}$$

where $\psi(\lambda) = \mathbb{E}e^{\lambda X_1}$ (which equals $\mathbb{E}\left[e^{\lambda X_{t+1}} \mid \mathcal{F}_t\right]$ for any t since the X_j are IID). Therefore,

$$M_t := e^{\lambda S_t}\psi(\lambda)^{-t}$$

will be an $\mathcal{F}$-martingale for any λ, namely the **exponential martingale**.

For a simple random walk,

$$\psi(\lambda) = \frac{e^{\lambda} + e^{-\lambda}}{2} = \cosh(\lambda) = 1 + \frac{\lambda^2}{2} + \frac{\lambda^4}{24} + \dots,$$

and

$$M_t = e^{\lambda S_t}[\cosh(\lambda)]^{-t} = \sum_{k=0}^{\infty} A_{k,t}\lambda^k,$$

where the coefficient $A_{k,t}$ is some function of k, t, and S_t, hence $\mathcal{F}_t$-measurable. (Note: Since for any $s, t \in \mathbb{R}$, the Taylor series for $e^{\lambda s}[\cosh(\lambda)]^{-t}$ converges for all λ in some neighborhood of 0 which is independent of s, there is some neighborhood of 0 in which the above power series expansion holds almost surely.) Using the fact that power series converge uniformly, we can interchange summation with expectation to obtain

$$\sum_{k=0}^{\infty} A_{k,t}\lambda^k = M_t = \mathbb{E}[M_{t+1} \mid \mathcal{F}_t] = \mathbb{E}\left[\sum_{k=0}^{\infty} A_{k,t+1}\lambda^k \mid \mathcal{F}_t\right] = \sum_{k=0}^{\infty} \mathbb{E}\left[A_{k,t+1} \mid \mathcal{F}_t\right]\lambda^k.$$

Equating the coefficients of the power series on the left and right, we get

$$A_{k,t} = \mathbb{E}\left[A_{k,t+1} \mid \mathcal{F}_t\right].$$

Thus, for each k, the collection $\{A_{k,t}\}_{t\geq 0}$ is an $\mathcal{F}$-martingale.

We can compute the first few martingale coefficients $A_{k,t}$ for SRW from the Taylor expansion of M_t. Using the binomial expansion

$$(1+x)^{-t} = \sum_{k=0}^{\infty} \binom{-t}{k} x^k = 1 - tx + \frac{t(t+1)}{2!}x^2 - \frac{t(t+1)(t+2)}{3!}x^3 + \dots$$

for $|x| < 1$, we have (for λ near 0)

$$(\cosh\lambda)^{-t} = \left(1 + \frac{\lambda^2}{2} + \frac{\lambda^4}{24} + \frac{\lambda^6}{720} + \dots\right)^{-t} = 1 - \frac{t\lambda^2}{2} + C_{4,t}\lambda^4 + C_{6,t}\lambda^6 + \dots$$

for some coefficients $C_{k,t}$. Thus

$$\begin{aligned} e^{\lambda S_t}(\cosh\lambda)^{-t} &= \left(1 + \lambda S_t + \frac{\lambda^2 S_t^2}{2} + \frac{\lambda^3 S_t^3}{6} + \dots\right)\cdot\left(1 - \frac{t\lambda^2}{2} + C_{4,t}\lambda^4 + \dots\right) \\ &= 1 + \lambda S_t + \frac{\lambda^2}{2}\left(S_t^2 - t\right) + \frac{\lambda^3}{6}\left(S_t^3 - 3tS_t\right) + \dots, \end{aligned}$$

so we recover the three martingales we've seen so far, plus higher order martingales if we expand more terms. The freedom of the parameter λ allows the exponential martingale M_t to encode all the information from the martingales $A_{k,t}$.

5 Hitting times for words

Now we begin to approach the problem from the first section.

Setting: $X_1, X_2, \ldots$ are IID and take values in some finite alphabet $\mathcal{A}$, with

$$\mathbb{P}(X_i = a) = p_a \quad \forall a \in \mathcal{A}.$$

Given a word $w \in \mathcal{A}^k$, the hitting time of w is

$$\tau_w = \min\{t \geq k : (X_{t-k+1}, \ldots, X_t) = w\}.$$

Question: Find $\mathbb{E}\tau_w$.

Martingale method (Li 1980)

Think of a gambler, Gandalf, making a sequence of fair bets on w coming up, starting at time t. First Gandalf bets on the 1st digit of w, then continues to bet on each successive digit in a way that makes all the bets fair (i.e. his expected winnings are 0 at each step):

- At time t Gandalf puts a dollar on $X_t = w_1$. He gets 0 if wrong, gets $\frac{1}{p_{w_1}}$ if right. In the latter case, he bets this on $X_{t+1} = w_k$, and gets $\frac{1}{p_{w_1} p_{w_2}}$ if right.
- Continue this process $\forall j < k$: If $(X_t \ldots, X_{t+j-1}) = (w_1, \ldots, w_j)$, Gandalf receives $\prod_{i=1}^{j} \frac{1}{p_{w_i}}$ and bets that on $X_{t+j} = w_j$, and he wins $\prod_{i=1}^{j+1} \frac{1}{p_{w_i}}$ if right.

Assume that one gambler enters the game at each time step, and each gambler bets on w using the scheme above. The game stops at time τ_w. Let M_t be the net winnings (at time t) of all gamblers who entered by time t. M_t is an $\{\mathcal{F}_t\}$-martingale, where $\mathcal{F}_t = \sigma(X_1, \ldots, X_t)$. For all $t \leq \tau_w$, we have

$$M_t = \sum_{j=1}^{k} \left(\mathbf{1}_{\left\{(X_{t-j+1}, \ldots, X_t) = (w_1, \ldots, w_j)\right\}} \cdot \prod_{i=1}^{j} \frac{1}{p_{w_i}} \right) - t.$$

At time t, there are t gamblers in the game, and each has paid a dollar, hence the "$-t$" term, and the sum computes their gross winnings.

Again, we want to apply the Optional Stopping Theorem to M_t. We have $|M_t| \leq C + t$ for some constant C, so

$$\sup_{t < \tau_w} |M_t| \leq C + \tau_w.$$

Is the righthand side integrable? Yes, because if we divide time into blocks of length k, in each block the chance of not seeing w is $1 - p_{w_1} \cdots p_{w_k}$, so

$$\mathbb{P}(\tau_w > \ell) \leq (1 - p_{w_1} \cdots p_{w_k})^{\lfloor \ell/k \rfloor}.$$

This shows that τ_w is dominated by a geometric random variable, hence integrable. Thus, the OST applies, so

$$0 = \mathbb{E}M_0 = \mathbb{E}M_{\tau_w} = w * w - \mathbb{E}\tau_w,$$

where

$$w * w := \sum_{j=1}^{k} \left(\mathbf{1}_{\left\{(w_{k-j+1},\ldots,w_k)=(w_1,\ldots,w_j)\right\}} \cdot \prod_{i=1}^{j} \frac{1}{p_{w_i}} \right).$$

So the answer to our question is $\mathbb{E}\tau_w = w * w$. Taking $p_0 = p_1 = \frac{1}{2}$ and $w = 001$, $v = 000$, we have $w * w = 8$ and $v * v = 2 + 4 + 8 = 14$.