Stat205B: Probability Theory (Spring 2003)	**Lecture: 7**

Examples

Lecturer: Jim Pitman *Scribe: Hao Zhang* nhz@cs.berkeley.edu

8 First Example

Consider a random walk on $\{0, 1, 2, ...\}$. Start at 1, there is a probability of p of going up and q of going down at each step. (For convenience, let 0 be reflecting: $p(0, 1) = 1$. but in the way we will define T_0, the behavior at 0 will not matter.)

Assume $q > p$. Let T_0 be the first time the random walk hits 0: $T_0 = \inf\{n : X_n = 0\}$. Then we know from Chapter 3 exercise 1.8-1.10 that $P(T_0 < \infty) = 1$. (In the terminology of the exercises, the random walk is the case of $P(\beta < \infty) = 1$ because up till the stopping time, it behaves the same as if there's no special treatment at 0, i.e. a random walk with downward drift).

Question: What is the area under the path until T_0? (in figure 1) Namely, let A be the area under the path, we want to compute $E_1(A)$. (the subscript "1" means it starts in state 1)

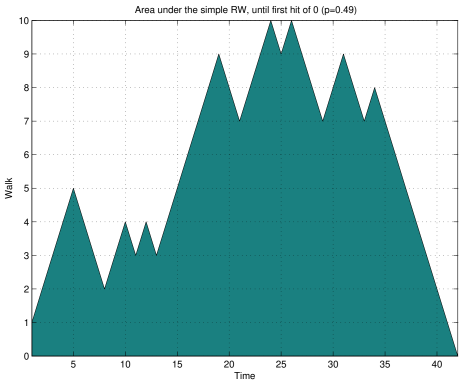

Figure 1: Simple random walk for example 1

Let's marshall some resources and try to write down a formula for A (write T for T_0 as shorthand):

$$\begin{aligned} A &= \frac{X_0 + X_1}{2} + \frac{X_1 + X_2}{2} + ... + \frac{X_{T-1} + X_T}{2} \\ &= \sum_{n=1}^{\infty} \frac{X_{n-1} + X_n}{2} 1(n < T_0) \\ &= \frac{1}{2} + \sum_{n=1}^{\infty} X_n 1(n < T_0) \qquad \text{(note } \tfrac{X_T}{2} = 0 \text{ because } X_{T_0} = 0 \text{ by definition of } T_0) \end{aligned}$$

We could try and write down a formula for $E(X_n 1(n < T_0))$ but this is hard!

Differently, let's look at the sum from another perspective, considering the number of hits to a certain level j. Formally, define $N_j = |\{n : X_n = j \text{ and } n < T_0\}|$, number of times state j is hit before T_0. Then,

$$A = \frac{1}{2} + \sum_{j=1}^{\infty} j N_j$$

The advantage of the N_j's is that we can compute $E(N_j)$ because of the connection between the "expected number of hits in an x-block" and the stationary measure, stated as Theorem 5.4.3 in Chapter 5 of Durrett.

Carry out the computation: define $\mu(j) = E_1(N_j)$ for $j \geq 1$, and let $\mu(0) = 1$. For convenience, define state 0 to be reflecting as in the beginning. Then μ is a stationary measure by theorem 5.4.3. Because the random walk occurs on a chain (an infinite chain, a special case of trees, c.f. random walk on graphs in Example 5.4.5 in Durrett), in the stationary case, probability mass has nowhere to escape but to obey the "detailed balance" condition along each edge:

$$\begin{aligned} \mu(0) &= 1 \\ \mu(j)p &= \mu(j+1)q \\ \mu(0) \cdot 1 &= \mu(1)q \end{aligned}$$

(Detailed balance also implies that the chain is reversible.) Solve for $\mu(j)$ gives:

$$\mu(j) = \frac{1}{q}(\frac{p}{q})^j$$

As a check, compute

$$E_1(T_0) = \sum_{j=1}^{\infty} \mu(j) = \sum_{j=1}^{\infty} \frac{1}{q}(\frac{p}{q})^j = \frac{1}{q(1-\frac{p}{q})} = \frac{1}{q-p}$$

which agrees with results obtained via Wald's equation.

There's yet an alternative method by a recurrence argument and a clever use of SMP (strong Markov property):

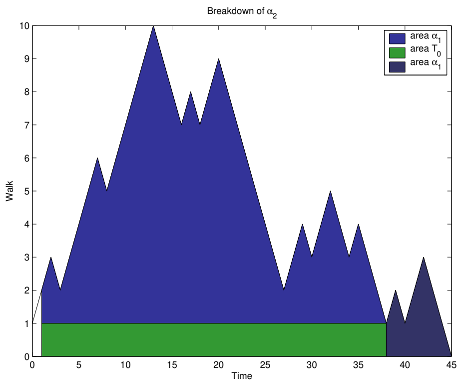

Figure 2: Breakdown of walk from state 2 for example 1

Let α_k be $E_k(A)$. Consider where the walk goes on the first step: condition on X_1 we see that $\alpha_1 =$

$q \cdot \frac{1}{2} + p \cdot (\frac{3}{2} + \alpha_2)$. In general, we have

$$\begin{aligned} \alpha_0 &= 0 \\ \alpha_1 &= q(\frac{1}{2}) + p(\frac{3}{2} + \alpha_2) \\ \alpha_2 &= q(\frac{3}{2} + \alpha_1) + p(\frac{5}{2} + \alpha_3) \\ &\dots \end{aligned}$$

However, one can break down the area of α_2 considering the first time the walk hits state 1. Up till then, the walk behaves the same as starting from 1 and hit 0. After that, it's yet an IID copy of the walk from 1 to 0. This is illustrated by figure 2. Hence we have

$$\alpha_2 = 2\alpha_1 + \frac{1}{q-p} \qquad \text{(last term is } E(T_0))$$

Combined with the previous expression for α_1, we solve for two equations in two unknowns and reach the answer:

$$\alpha_1 = \frac{1}{1-2p}(\frac{q}{2} + \frac{3p}{2} + \frac{p}{q-p})$$

9 Exercise with stationarity

Let (X_n) be a stationary sequence of r.v.'s governed by $\mathbb{P}$. Think of (X_n) as a Markov chain with $X_0 \sim \pi$ and $\pi P = \pi$. Let $T_B = \inf\{n \geq 1 : X_n \in B\}$ $(\inf \emptyset = \infty)$, the hitting time of B. Show that

$$\mathbb{P}(X_0 \in B, T_B \geq n) = \mathbb{P}(T_B = n)$$

Proof. Note that the two events are "reversal" of each other. The LHS event says that the first state is in B and for the first $2..(n-1)$ step the chain doesn't touch B. The RHS event says that the n'th step is in B, however the first $2..(n-1)$ steps are not in B. Formally, those events are:

LH event	$(X_0 \in B, X_1 \notin B, ..., X_{n-1} \notin B)$
RH event	$(X_1 \notin B, X_2 \notin B, ..., X_{n-1} \notin B, X_n \in B)$

So this would be obvious if $(X_0, ..., X_n) \stackrel{d}{=} (X_n, ..., X_0)$, i.e. if the chain is *reversible and stationary.* However, here we only assume stationarity $(X_0, ..., X_{n-1}) \stackrel{d}{=} (X_1, ..., X_n)$.

To see this, the key idea is look at those events as the "set-difference" between events of the form:

$$X_0 \notin B, X_1 \notin B, ..., X_n \notin B$$

Formally, define $A_{i,j} = \{X_i \notin B, ..., X_j \notin B\}$. Then, LHS $= A_{1,n-1} - A_{0,n-1}$ and RHS $= A_{1,n-1} - A_{1,n}$. We use "-" instead of set minus because for each expression, the latter set is contained in the earlier. This also implies the probability substracts fine, too.

Hence, all is left to show is that $P(A_{0,n-1}) = P(A_{1,n})$, which holds by the stationarity assumption. (This unusual proof method is noted in works of Mark Kac and David Freedman.) $\square$

Why is it interesting? Let's sum over n and see if anything shows up: the RHS simply sums to $P(T_B < \infty)$;

the LHS is by intuition some kind of expectation, derived formally as

$$\begin{aligned}
\sum_{n=1}^{\infty} P(X_0 \in B, T_B \geq n) &= \sum_{n=1}^{\infty} E[1(X_0 \in B)1(T_B \geq n)] \\
&= E[\sum_{n=1}^{\infty} 1(X_0 \in B)1(T_B \geq n)] \qquad \text{by MCT} \\
&= E[1(X_0 \in B)\sum_{n=1}^{\infty} 1(T_B \geq n)] \\
&= E[1(X_0 \in B)T_B]
\end{aligned}$$

Hence we have $E(T_B 1(X_0 \in B) = \mathbb{P}(T_B < \infty) \leq 1$. Relax it to get $E(T_B 1(X_0 \in B)) < \infty$, meaning that when X_0 is in B, then the chain is "positive recurrent" on B because $E(T_B) < \infty$. Or, formally we have the (weaker) statements that

$$\begin{aligned}
\mathbb{P}(T_B < \infty \mid X_0 \in B) &= 1 \qquad \text{or} \\
\mathbb{P}(X_0 \in B) &= 0
\end{aligned}$$

To summarize, we now know that for a Markov chain that has a stationary distribution π. For any set of states B, if $\pi(B) > 0$, then B must be recurrent.

This enables us to reprove the fact concerning a Markov chain which is (1) irreducible and (2) has a stationary distribution. By irreducibility we know that the stationary distribution must be everywhere positive (by argument in a previous lecture). Therefore, by the newly obtained result, we know all states are recurrent. Moreover, let $B = \{i\}$ in

$$E_{X_0 \sim \pi}(T_B 1(X_0 \in B)) = \mathbb{P}(T_B < \infty)$$

we know $B = \{i\}$ is reccurent so the RHS is 1. Consider the LHS, the r.v. inside the expectation takes non-zero value only on $\{X_0 = i\}$, moreover, on this set, T_B is simply $T_{i,i}$ the expected time of going from i to i itself. Hence, it follows by reasoning with individual paths that the LHS is

$$\pi_i E_i(T_i)$$

Therefore, $E_i(T_i) = \frac{1}{\pi_i}$, a result obtained in Durrett as Theorem 5.4.6.