Stats 150, Fall 2019
Lecture 7, Thursday, 9/19/2019

1 First Step Analysis: Continued

The simple idea here is to derive equations by conditioning on step 1. We can find all sorts of things about Markov chains by doing exactly this.
Pitman notes that the text does not make a remark about this but rather keeps doing technique. Recall that first step analysis for a Markov chain $(X_0, X_1, X_2, \dots)$ for some random variable $Y = Y(X_0, X_1, X_2, \dots)$.
If we know $E_x Y$ and we want to compute the expectation of probability distribution $\lambda = \lambda(x), x \in S$, namely $\mathbb{E}_\lambda Y$, we would take:

$$\mathbb{E}_\lambda Y = \sum_{x \in S} \lambda(x) \mathbb{E}_x Y$$

Put simply, the expectation of a random variable Y is the expectation of the expectation of Y conditioned on some X_0. That is,

$$\mathbb{E}(Y) = \mathbb{E}\left[\mathbb{E}(Y \mid X_0)\right].$$

We may want to condition on X_1 as well, which is how we derived the harmonic equations in the previous lecture. Let's look at an example where we can do this again.

Example: Suppose we have a set of states A (we will make them absorbing as a matter of technique), and consider:

$$V_A := \min\{n \geq 0 \mid X_n \in A\},$$

and we want to find: $\mathbb{E}_x V_A$ for any initial x.
Now if $x \in A$, then we trivially have:

$$\mathbb{E}_x V_A = 0.$$

If $x \notin A$, then we define a function for mean, say $m(x) = m_A(x) := \mathbb{E}_x V_A$, where we drop the subscript A as it is understood from context. We want equations for $m(x)$.
From x, we hit $X_1 = y$ with probability $P(x, y)$. Now given $X_0 = x$ and $X_1 = y$, let $x \notin A$. Then:

$$\mathbb{E}(V_A \mid X_0 = x, X_1 = y) = 1 + \mathbb{E}_y(V_A),$$

Notice especially that this is correct if $y \in A$. If we happen to hit an absorbing state, then the second term $\mathbb{E}_y(V_A)$ is zero. Additionally, this is correct if $y \notin Q$, where $\mathbb{E}_y V_A \geq 1$ (it would take at least one step).
This means that we can write down a system of equations, relating to the mean times (for $x \notin A$):

$$m(x) = 1 + \sum_{y \in S} P(x, y) m(y)$$

If we have only a small number of states, then we have a finite number of linear equations and this number of unknowns.

Pitman notes that in the text, there is a theorem that states that as long as we can reach the boundary from the interior (in some number of steps) with positive probability, this system of equations will have a unique solution. Pitman notes that we can simply check this computationally via matrices.

1.1 Application to Simple Symmetric Random Walks

This is just the usual Gambler's ruin for a fair coin. We start with x dollars and play for $\pm\$1$ gains with equal probability until we hit either 0 or some $\$N$. Last lecture, we showed:

$$\mathbb{P}_x(\text{reach } N \text{ before } 0) = \frac{x}{N}.$$

Now set $A := \{0, N\}$ as our absorbing states, and V_A here will be the duration of the game, where

$$V_A := \min\{n \geq 0 \mid X_n \in A\}.$$

Recall that there remains the scenario of never hitting the boundary A, but we have already found before that the probability assigned to this enormously infinite number of never-ending paths is zero. To see this, notice that for any 'block' of N steps, there is a strictly positive probability that we hit a boundary state. We use this argument to form the geometric bound as we have before in a previous lecture.
It remains to solve $m(x) := \mathbb{E}_x V_A$. To find the equations, we first write out the boundary conditions. That is,

$$m(0) = m(N) = 0.$$

Now the nontrivial cases, we again break into two parts:

$$m(x) = 1 + \frac{1}{2}m(x+1) + \frac{1}{2}m(x-1),$$

for $0 < x < N$. Then we solve for this system of equations. Pitman notes it's a good idea to recall that $h(x) = \frac{1}{2}h(x+1) + \frac{1}{2}h(x)$, which implies that $h(x)$ is linear (affine). Now writing in terms of placeholder constants a, b, we have:

$$h(x) = ax + b$$

Consider now, if we insert another term:

$$m(x) = cx^2 + ax + b.$$

Then we observe:

$$\begin{aligned}\frac{1}{2}c(x+1)^2 + \frac{1}{2}c(x-1)^2 &= cx^2 + \underbrace{\frac{1}{2}c(2x) + \frac{1}{2}c(-2x)}_{=0} + c \\ &= \boxed{c(x^2+1)}.\end{aligned}$$

Now from this sort of consideration, $m(x)$ as above solves the equation

$$\frac{1}{2}m(x+1) + \frac{1}{2}m(x-1) = c + m(x)$$

Hence we conclude that our system of equations is solved by a quadratic function of the form $m(x) = cx^2 + ax + b$.

1.2 Summarizing our findings:

$$g_1(x) = ax + b \implies \frac{1}{2}g_1(x+1) + \frac{1}{2}g_1(x-1) = g_1(x)$$
$$g_2(x) = cx^2 \implies \frac{1}{2}g_2(x+1) + \frac{1}{2}g_2(x-1) = g_2(x) + c$$

These together imply:

$$g(x) = cx^2 + ax + b = (g_1 + g_2)(x) \implies \frac{1}{2}g(x+1) + \frac{1}{2}g(x-1) = g(x) + c$$

Hence we have that

$$m(x) := cx^2 + bx + a$$

solves our equations from earlier if and only if $c = -1$. Then plugging this in, we have:

$$m(x) = -x^2 + bx + a,$$

and additionally recall that $m(0) = m(N) = 0$. There's only one quadratic that satisfies these, namely:

$$m(x) = -x(x - N) = \boxed{x(N - x)}$$

In summary, with the idea to try a quadratic (which Pitman notes is not too different from noticing before that our function need to be linear), finding the exact solution is not too tricky.

Break time.

We went slowly to look at first passage times in a particular example, and Pitman notes this is simply find the matrix, write down the equations, and compute the equations. Now we would like to consider that X_1 may not be the only variable on which we would like to condition. There may be more clever techniques, where we will want to use our imagination to find a better variable on which to condition.
According to Pitman, often in a Markov chain, we can commonly try X_0, X_1, X_n. In particular, to derive the recurrence of a mean (as on our homework), we would want to use X_n.
Now we consider a more special example:

2 Independent Bernoulli (p) Trials

We want to find the mean time until we see N successes in a row, for example $N = 3$.
For example, we illustrate one sequence:

$$X_n : 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1$$

Let τ_N be the number of trials required. The distribution itself is tricky, but we want the expectation, which is relatively simple.
Of course,

$$\begin{aligned}\mathbb{E}\tau_N &= \sum_{k=N}^{\infty} k\mathbb{P}(\tau_N = k)\\ &= \sum_{k=N}^{\infty} \mathbb{P}(\tau_N \geq k),\end{aligned}$$

where neither the simple probability (first equality) nor the tail sum (second) has a simple formula for our purposes.
Hence we ask, what should we condition on? We try τ_{N-1}, which is to have $N-1$ of 1s in a row.

$$\tau_N = \tau_{N-1} + \Delta_N,$$

where

$$\Delta_N = \begin{cases} 1 \text{ with probability } p \\ 1 + \text{ a copy of } \tau_N \text{ otherwise} \end{cases}$$

Now this leaves us depressed as we must now start all over if we fail the last required trial; however, this expression above may give us enough to solve the problem!
Let $\mu_N := \mathbb{E}\tau_N$. We have:

$$\mu_N = \mu_{N-1} + 1 + q\mu_n.$$

where rearranging gives:

$$\mu_N = \frac{\mu_{N-1} + 1}{p}$$

We test this:

$$\mu_1 = \frac{1}{p},$$

by the mean of geometric (tail sums). Similarly,

$$\mu_2 = \frac{\left(\frac{1}{p} + 1\right)}{p} = \frac{1+p}{p^2},$$

and we guess:

$$\mu_N = \frac{1 + p + p^2 + \cdots + p^{N-1}}{p^N}.$$

In summary, we solved this problem by noticing that to get to N in a row, we needed to first get to $N-1$ in a row, and then reconsider the two states for the final step. Now Pitman gives his own solution.

2.1 Pitman's Solution

Define G_0 as the first $n \geq 1$ such that $X_n = 0$ (that is, wait for the first 0). In other words, G_0 is one plus the length of the first run of 1s. Then $G_0 \sim$ Geometric(q), where q is the failure probability.
We want to find $\mathbb{E}(\tau_N)$ by some suitable conditioning (which requires artistic thinking). Here, we want to condition on G_0, as G_0 is closely related to τ_N. If $G_0 > N$, then $\tau_N = N$. On the other hand, if $G_0 = g \leq N$, then

$$\tau_N = g + \hat{\tau}_N,$$

where equality in distribution gives:

$$(\hat{\tau}_N \mid G_0 = g) \stackrel{d}{=} \tau_N.$$

Therefore, by conditioning on G_0, we have:

$$\mathbb{E}\tau_N = \left[\sum_{g=1}^{N} \mathbb{P}(G_0 = g)(g + \mathbb{E}\tau_N)\right] + \mathbb{P}(G_0 > N)N$$

Now let $\mu_N := \mathbb{E}\tau_N$, so that the earlier equation gives us:

$$\mu_N = \sum_{g=1}^{N} p^{g-1} q(g + \mu_N) + p^N N,$$

which is another equation satisfied by μ_N. Pitman leaves it to us as an exercise to algebraically show the equality of these results in general.
We look at a simple $N := 2$ case. Here in this solution, we have:

$$\begin{aligned} \mu_2 &= p^0 q(1 + \mu_2) + pq(2 + \mu_2) + p^2 \cdot 2 \\ \mu_2(1 - q - pq) &= q + 2pq + 2p^2, \end{aligned}$$

and we should check this more closely.

Lecture ends here.