Stat205B: Probability Theory (Spring 2003) **Lecture: 2**

Hitting Probabilities

Lecturer: James W. Pitman *Scribe: Brian Milch*

2.1 Hitting probabilities

Consider a Markov chain with a countable state space S and a transition matrix P. Suppose we want to find $\mathbb{P}_i(\text{chain hits } A \text{ before } C)$ for some $i \in S$ and disjoint subsets A and C of S. We define the **boundary** $B := A \cup C$. It is obvious that $\mathbb{P}_i(\text{chain hits } A \text{ before } C)$ for $i \in B^c$ is determined by the transition probabilities $P(j,k)$ for $j \notin B, k \in S$. Since this probability does not involve $P(b,j)$ for $b \in B, j \in S$, there is no loss of generality in assuming, as we will from now on, that B is an **absorbing boundary**, meaning that $P(b,b) = 1$ for all $b \in B$.

If we let $\tau := \inf\{n : X_n \in B\}$, where B is absorbing, then we can restate our problem as finding the **hitting probability** $\mathbb{P}_i(X_\tau \in A)$ for some $A \subseteq B$. Define:

$$h_A(i) \triangleq \mathbb{P}_i(X_\tau \in A)$$

We can condition on X_1 and decompose to get:

$$\begin{aligned} \mathbb{P}_i(X_\tau \in A) &= \sum_j \mathbb{P}_i(X_1 = j, X_\tau \in A) \\ h_A(i) &= \sum_j P_{ij} h_A(j) \qquad (*) \end{aligned}$$

To see that we're handling the boundary cases correctly, note that if $i \in B$, then $P_{ij} = 0$ for $j \neq i$. And the values of h_A on the boundary are:

$$h_A(j) = \begin{cases} 1 & \text{if } j \in A \\ 0 & \text{if } j \in B \setminus A \end{cases}$$

The assumption that B is absorbing makes the equations $(*)$ hold for all $i \in S$, which simplifies the discussion.

Another way to write $(*)$ is: $h_A = Ph_A$. Thus h_A is a harmonic function, as defined in the previous lecture. Also, h_A satisfies the **boundary condition**: $h = \mathbf{1}_A$ on B, meaning $h_A(i) = \mathbf{1}_A(i)$ for $i \in B$. Thus the function $h = h_A$ solves the *boundary value problem*

$$\begin{aligned} h &= Ph \\ h &= 1_A \text{ on } B \end{aligned} \qquad \text{(BVP)}$$

One of the homework problems (an easy application of martingale theory) is to show that if $\mathbb{P}_i(\text{hit } B \text{ eventually}) = 1$, then $h = h_A$ is the *unique* solution of this BVP. This can also be deduced from the more general characterization of h_A given in the following theorem, by consideration of both h_A and h_{B-A}.

What if $\mathbb{P}_i(\text{hit } B \text{ eventually}) < 1$? Then consider the **escape probability**:

$$\begin{aligned} e_B(i) &\triangleq \mathbb{P}_i(\text{never hit B}) \\ &= \mathbb{P}_i(X_n \notin B, \forall n \geq 0) \end{aligned}$$

For example, consider a biased random walk with $P(i, i+1) = p$ and $P(i, i-1) = q$. Suppose $p > q$, and let $B = \{0\}$. Then $e_B(1) > 0$.

By conditioning on X_1, we can show $e_B = Pe_B$: that is, escape probabilities are also harmonic functions. Clearly, $e_B = 0$ on B. Now, note that the set of solutions of $h = Ph$ is a vector space: if we add harmonic functions or take scalar multiples, we still get harmonic functions. We have seen that:

- $h = h_A$ solves $h = Ph$ and $h = \mathbf{1}_A$ on B
- $h = e_B$ solves $h = Ph$ and $h = 0$ on B

Therefore for any constant c, $h = h_A + ce_B$ solves $h = Ph$ and $h = \mathbf{1}_A$ on B. So when e_B is not identically zero, h_A is not the unique h such that $h = Ph$ and $h = \mathbf{1}_A$ on B. The problem now is how to characterize the hitting probability vector h_A among all solutions of the BVP.

Theorem 2.1 *$h = h_A$ is the minimal non-negative solution of (BVP). That is, if h solves (BVP) and $h(i) \geq 0$ for all i in S, then $h(i) \geq h_A(i)$ for all $i \in S$.*

Proof: First note that

$$\begin{aligned} h_A(i) &= \mathbb{P}_i(X_n \in A \text{ for some } n) \quad \text{by absorbing boundary assumption} \\ &= \lim_{n\to\infty} \mathbb{P}_i(X_n \in A) \quad \text{because } X_n \in A \text{ implies } X_{n+1} \in A \\ &= \lim_{n\to\infty} (P^n \mathbf{1}_A)_i \end{aligned}$$

Suppose $h \geq 0$ and h satisfies (BVP). Then:

$$\begin{aligned} h &\geq \mathbf{1}_A \\ P^n h &\geq P^n \mathbf{1}_A \quad \forall n \quad \text{because } P_{ij} \geq 0 \end{aligned}$$

But $P^n h = h$ because h is harmonic. So for all $i \in S$, we have:

$$\begin{aligned} h(i) &\geq (P^n \mathbf{1}_A)_i \quad \forall n \\ h(i) &\geq \lim_{n\to\infty} (P^n \mathbf{1}_A)_i = h_A(i). \end{aligned}$$

■

2.1.1 Example: Simple random walk on $\{0, 1, 2, \ldots\}$ with 0 absorbing

Let $P(0, 0) = 1$, and for $i \geq 1$, let $P(i, i+1) = p$ and $P(i, i-1) = q$. Let $h_0(i) = \mathbb{P}_i(\text{ever hit } 0)$. Then $h = h_0$ solves:

$$\begin{aligned} h(i) &= qh(i-1) + ph(i+1) \\ h(0) &= 1 \end{aligned}$$

Rather than solving these equations directly, we can use a shortcut based on the observation that $\mathbb{P}_1(\text{ever hit } 0) = \mathbb{P}_i(\text{ever hit } i-1)$ for all $i \geq 1$. So by the strong Markov property:

$$h(2) = h(1)^2$$
$$h(i) = (h(1))^i$$

So we just have to compute $h(1)$.

$$h(1) = q \times 1 + p \times h(1)^2$$
$$x = q + px^2$$

where $x = h(1)$. The roots of this equation are $x \in \{1, q/p\}$. So we have two non-negative harmonic functions that solve (BVP):

$$h(i) \equiv 1$$
$$h(i) = \left(\frac{q}{p}\right)^i$$

By the theorem proved above, h_0 is the smaller one. That is:

$$h_0(i) = \left(\min\left(1, \frac{q}{p}\right)\right)^i = \begin{cases} 1 & \text{if } p \leq q \\ \left(\frac{q}{p}\right)^i & \text{if } p > q \end{cases}$$

2.1.2 Example: Extinction probabilities for branching process

Suppose we are given an offspring distribution $p_0, p_1, p_2, \ldots$ with $\sum p_i = 1$, $p_1 < 1$, and $p_i \geq 0$. A branching process models a population where each individual has a number of offspring with this distribution (that is, the probability of an individual having 1 offspring is p_1, etc.). The chain is:

$$X_0 = \text{initial number of individuals}$$
$$X_n = \text{number of individuals in generation } n$$

Clearly the 0 state is absorbing. So $h_0(i) = \mathbb{P}_i(\text{population dies out}) = \mathbb{P}_i(X_n = 0 \text{ eventually})$.

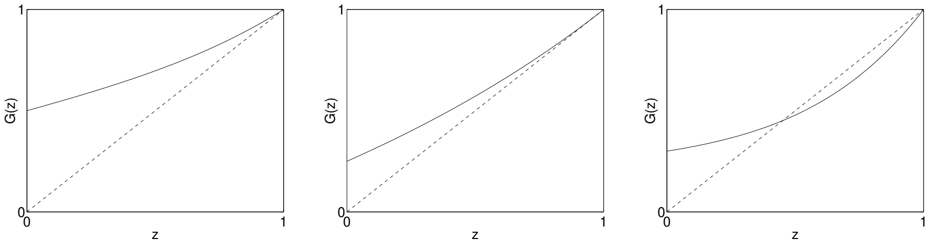
Note that if there are i individuals in a generation, then their descendants form i independent trees. These i trees must all die out in order for the entire population to die out. So $h_0(i) = h_0(1)^i$. Now, by summing over X_1, we find that:

$$h_0(1) = \sum_{j=0}^{\infty} p_j (h_0(1))^j$$

We now introduce a probability generating function:

$$G(z) \triangleq \sum_{j=0}^{\infty} p_j z^j = E\left(z^X\right)$$

where $\Pr(X = j) = p_j$. Note that $z = h_0(1)$ solves $z = G(z)$: in fact, $h_0(1)$ is the minimal solution in $[0, 1]$ of $z = G(z)$.



subcritical $\mu < 1$ | critical $\mu = 1$ | supercritical $\mu > 1$

Figure 2.1: Plots of the curve $G(z)$ (solid) and the 45-degree line (dashed) for branching processes with various values of μ.

Thus we can find $h_0(1)$ by plotting $G(z)$ versus z over the interval $[0, 1]$, and observing where the plot intersects the 45-degree line $G(z) = z$. Note that $G(0) = p_0 > 0$, and $G(1) = \sum p_j = 1$. Also, $z \mapsto G(z)$ is convex. As illustrated in Figure 2.1, the number of roots depends on the quantity:

$$\mu \triangleq \sum_{j=0}^{\infty} jp_j = G'(1)$$

In the **subcritical** case where $\mu < 1$, the only root is at $z = 1$, so $h_0(1) = 1$. In the **critical** case where $\mu = 1$, the $G(z)$ curve is tangent to the line $z = G(z)$ at $z = 1$. Again, $z = 1$ is the only root, so $h_0(1) = 1$. But in the **supercritical** case where $\mu > 1$, there are two roots, and $h_0(1)$ is the smaller one.

The conclusion is that for $\mu \leq 1$, the branching process dies out almost surely. For $\mu > 1$, the branching process dies out with probability $(h_0(1))^i$, where $h_0(1)$ is the smaller of the two solutions of $z = G(z)$.

The case $p_1 = 1$ is special: then $\mu = 1$, but $G(z) = z$ for all $z \in [0, 1]$, so $h_0(1) = 0$ (the smallest root).

2.2 Strong Markov property

Section 5.2 of Durrett deals with the strong Markov property. Suppose we have a chain $X_0, X_1, \ldots$ with sequence space Ω. Recall that a random time $\tau : \Omega \to \{0, 1, 2, \ldots, \infty\}$ is a **stopping time** if $\{\tau = n\} \in \sigma(X_1, X_2, \ldots, X_n)$. For example, $\tau = \inf\{n : X_n \in A\}$ is a stopping time.

Theorem 2.2 (Strong Markov property) *For any Markov chain X_n with transition matrix P, and any stopping time τ:*

- *$(X_0, X_1, \ldots, X_\tau)$ and $(X_\tau, X_{\tau+1}, \ldots)$ are conditionally independent given X_τ on $\{\tau < \infty\}$;*
- *given $(X_0, \ldots, X_\tau)$ with $X_\tau = j$, the process $(X_\tau, X_{\tau+1}, \ldots)$ is a Markov chain with transition matrix P started in state j.*

Sketch of proof: We first show that the theorem holds when τ is a fixed (not random) time. Then we show that it is true in general by conditioning on the value of τ, and summing over all possible τ values. See the textbook for details.