1 Random Walks

A random walk is a special kind of Markov chain. In a random walk, the states are all integers. Negative numbers are (sometimes) allowed. Say you start in a state \( a \). The one-step transitions are that, with probability \( p \), you move to state \( a + 1 \) and with probability \( q = 1 - p \), you move to state \( a - 1 \). The largest move you can make per transition is one step in either direction, and there is no probability of remaining in the same state.

A couple of interesting facts:

The simple random walk is *temporally homogeneous*:

\[
P(S_n = j | S_0 = a) = P(S_{m+n} = |S_m = a)
\]

What this means is that starting in state \( a \) and being in state \( j \) after \( n \) transitions has the same probability as being in state \( a \) after the first \( m \) transitions, and then being in state \( j \) \( n \) transitions after that.

The simple random walk has the *Markov property*:

\[
P(S_{m+n} = j | S_0, S_1, \ldots, S_m) = P(S_{m+n} = j | S_m)
\]

This means that the probability of getting to state \( j \) in \( n \) transitions depends only on the state you’re currently in. Knowing anything or everything that occurred prior to that state gives no additional information.

2 Absorbing Probabilities

Suppose we have a random walk which is restricted to the range \([a, b]\). In other words, you start at some state in that range, and once your walk reaches either state \( a \) or \( b \), the walk ends. Here, \( a \) and \( b \) are called *absorbing* states: once the walk reaches either state, it will never leave that state.
In the lecture notes, the professor derives a formula for finding the probability that the walk ends at state $b$ rather than state $a$, given that you started in state $h$:

$$w_h = \left(\frac{q}{p}\right)^b - \left(\frac{q}{p}\right)^a \quad \text{for} \ p \neq q$$

$$w_h = \frac{h - a}{b - a} \quad \text{for} \ p = q$$

There are similar equations given for the probability you end in state $a$:

$$u_h = \left(\frac{q}{p}\right)^b - \left(\frac{q}{p}\right)^h \quad \text{for} \ p \neq q$$

$$u_h = \frac{b - h}{b - a} \quad \text{for} \ p = q$$

Here’s an exercise dealing with these probabilities:

A gambler, playing roulette, makes a series of $\$1$ bets. He wins a dollar with probability 9/19 and loses a dollar with probability 10/19. He starts with 8 dollars, and determines that he’ll quit when he’s broke, or when he’s reached $\$10$. What are the absorption probabilities?

We know that $p = 9/19$ and $q = 10/19$, so $q/p = 10/9$. Our lower bound is $a = 0$ and the upper is $b = 10$. Finally, our starting state is $h = 8$. Plugging these figures into the formulae above:

$$w_h = \frac{(10/9)^8 - (10/9)^0}{(10/9)^{10} - (10/9)^0} = .7083$$

$$u_h = \frac{(10/9)^{10} - (10/9)^8}{(10/9)^{10} - (10/9)^0} = .2917$$

Note that these sum to 1. This isn’t surprising! It’s provable (and no, I’m not going to prove it...) that a random walk of this set up will eventually reach one of its absorbing states.

### 3 Mean number of steps taken until walk stops

Formula in the lecture notes:

$$m_h = \frac{w_h(b - h) + u_h(a - h)}{p - q}$$

Let’s see how long, on average, our gambler will be playing:
\[
m_{90} = \frac{.7083(10 - 8) + .2917(0 - 8)}{9/19 - 10/19} \approx 17.4
\]

4 Maximum height of the walk

This all has been working towards the test statistic used in BLAST. In BLAST, you start at \( h = 0 \), there’s an absorbing state at \( a = -1 \), but there’s no upper absorbing state: the number can get as big as you want! However, the instant you get to -1, it’s game over. The concern is with \( Y_{\text{max}} \), the largest number your walk reaches before it ends. Class notes show that

\[
\mathbb{P}(Y_{\text{max}} \geq y) = 1 - (1 - (1 - e^{-\lambda})e^{-\lambda y})^m
\]

where \( \lambda = \log (q/p) \).