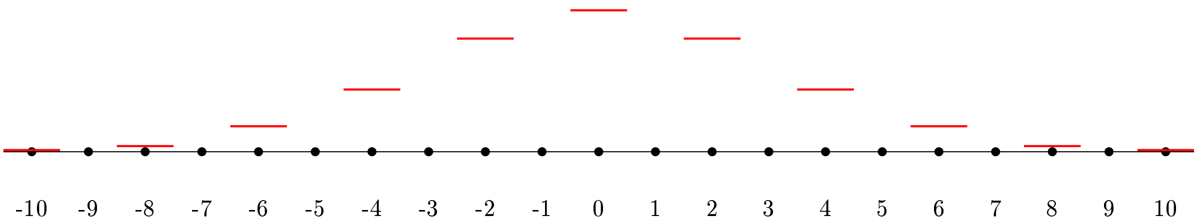
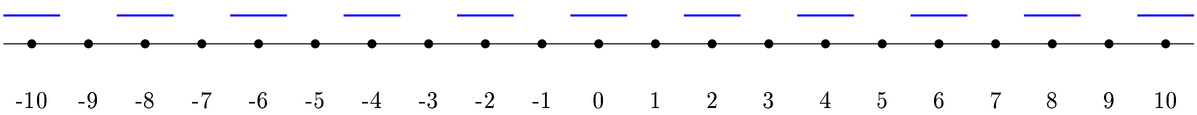
1 Unpredictable walks

Your coding assignment this week was to make a simple game, that takes an input of an integer from -10 to 10 and then simulates a 10-step simple random walk on the integers starting at 0. If the walk ends at the input integer, it's a win. Clearly, the best bet was to input 0, since that gives you the best chance of winning. The probability distribution of the final position is binomial on even integers in the range:

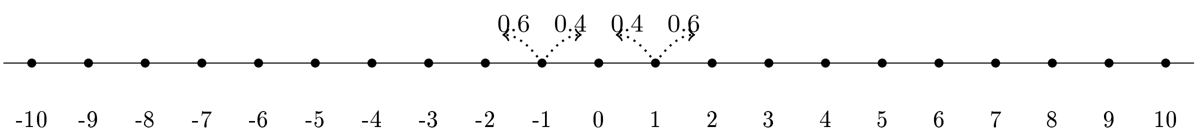


There is an entire research area devoted to *predictability* of a random walk, i.e. the question of how good is the best bet of a person who knows everything about how the walk works (can see your code), if they want to guess the position of the walk after k steps. Suppose you're writing a computer worm that's supposed to jump between computer systems, and the sysadmin in charge got hold of a copy of your code, so knows how the walk works. But to neutralize the worm, he needs to catch it in a particular system. It's this kind of a problem. You'd want your walk to give rise to probabilities that are as evenly split as possible:



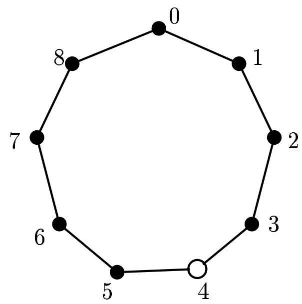
In real world, writing worms is bad manners. However, unpredictable walks can be a powerful force for good. TOR, a network that helps thousands of people get around censorship and surveillance, uses unpredictable walks. The acronym stands for "The Onion Router," because it uses public key cryptography in layers, like an onion. It's a network of relays ran by people who decide to donate some of their bandwidth so that anyone with censored access to the internet can go through their computers to see the blocked websites. A connection takes a random walk of 3 steps through the network of relays, in order to properly hide among other connections... and if someone watches the entries and exits, we would like their best bet on which traffic belongs to whom to be as bad as possible. So the goal here is to construct an unpredictable walk.

A good way to make sure that the distribution is a little more spread out than for a simple random walk is to set up probabilities so that the walk is pushed away from the center. For example, all nodes to the right give slightly higher probability of pushing the token further to the right and vice versa:

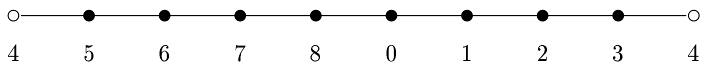


2 Hitting time

Another interesting question on the homework, was the one about a walk on a cycle. On the graph below, if 4 is an absorbing state and all other nodes behave as in a simple random walk, what is the expected time to absorption if you start the walk at 0?



Most of you noticed that we can make this into a case as in Gambler's Ruin, if we cut this circle at 4 and look at the walk as a walk on a path with absorbing states on both ends:

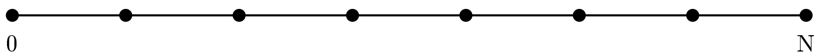


Then the expected time to reach 4 is 5×4. In fact, for a walk on a cycle like this one, the expected time to hit another node will be the product of distances in both directions.

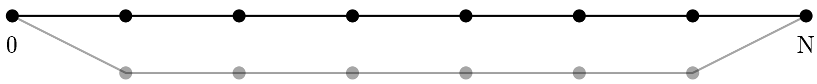
We don't need to look at 4 as an absorbing state. It could be a simple random walk, with 4 just like any other node and we could be simply looking for the expected time until we first visit it.

Definition 1 (Hitting time) In a Markov chain, the *hitting time from i to j*, written h_{ij}, is the expected number of steps for a walk starting at state i to visit state j for the first time.

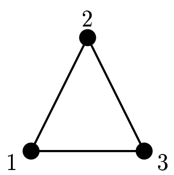
Puzzle: What is the hitting time of node N for a simple random walk that starts at 0?



We can use what we know about random walks on a cycle - create a shadow duplicate of this path as pictured below. The hitting times are the same! So the answer is N^2.



Here's another one. What is h_{13} for a simpe random walk on the triangle graph below?



We can do this recursively: you might succeed on the first step. If you don't, you've gone to node 2. You've spent one step, and you're in a situation identical as at the start because the graph, and node 3 in particular, looks the same from node 2 as from node 1. This can be written as:

$$h_{13} = \frac{1}{2} + \frac{1}{2}(1 + h_{13}),$$
$$h_{13} = 2.$$

In other words, each time you have probability 1/2 of success, and as you remember the expected number of trials to first success in that case is 2.

3 Irreducible and Regular Markov Chains

Definition 2 (Irreducible/Ergodic Markov Chains) A markov chain is *irreducible*, or, equivalently *ergodic*, if for any two states i and j, it's possible to go from state i to state j in a finite number of steps.

Definition 3 (Regular Markov Chains) A markov chain is *regular*, if there exists a natural number n such that for any two states i and j, it's possible to go from state i to state j in exactly n steps.

Clearly, any regular Markov chain is also ergodic. Is it true that any ergodic Markov chain is regular?

The answer is no. Recall a simple random walk on:

You can go between the states in a finite number of steps, but you can only come back to the same state in an even number of steps and go to the other one in an odd number. So this Walk is ergodic, but not regular.

What does being ergodic or regular mean for the transition matrix P of a Markov chain? Saying that the chain is ergodic is the same as saying that for any i, j there is a power of P in which the entry in row i and column j is bigger than 0. Saying that the chain is regular is the same as saying that there exists some power n of P such that ALL entries of P^n are bigger than 0.

For the simple random walk on two nodes, the powers of the transition matrix are:

$$P^n = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ if } n \text{ is odd,} \quad P^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ if } n \text{ is even.}$$

So again, we can see that it's ergodic but not regular.

The following two facts we already mentioned on Monday, but they should carry a lot more weight now:

- Any irreducible Markov chain has a unique stationary distribution u.
- For any regular Markov chain,

$$P^n \to \begin{bmatrix} u \\ \hline \vdots \\ \hline u \end{bmatrix} \text{ as } n \to \infty$$

  i.e. the transition matrix after n steps tends to a matrix in which every row is the stationary distribution as n grows. In other words, we eventually lose memory of where we started, the probabilities of where the token will be if we were to suddenly freeze it get closer and closer. This process of forgetting where we started is called the *mixing* of a random walk.

The second statement is not true for chains that are ergodic but not regular, it's enough to look at the simple random walk on two nodes to convince yourself of that. However a weaker statement is true: if we average out the coefficient in consecutive matrices P^n, the result will tend to the stationary distribution. This is called the *ergodic theorem.*

Theorem 1 (Ergodic Theorem)
Let P be the transition matrix of an ergodic Markov chain. Then, if we define A_n as:

$$A_n = \frac{I + P + P^2 + \cdots + P^n}{n+1},$$

we have:

$$A_n \to \begin{bmatrix} u \\ \hline \vdots \\ \hline u \end{bmatrix} \text{ as } n \to \infty$$

Finally, the following is possibly the most beautiful property of ergodic Markov chains. As we keep going, with high probability the proportion of times that a state is visited approaches its coefficient in the stationary distribution. So for example, in a simple walk on a graph each state will be visited a number of times proportional to its degree in the graph (i.e. the number of edges coming out.)

Theorem 2 (Law of Large Numbers for Markov Chains)
For an ergodic Markov chain, let $H_j^{(n)}$ be the proportion of times in n steps that the state j is visited. Then, for any $\epsilon > 0$,

$$P(|H_j^{(n)} - u_j| > \epsilon) \to 0 \text{ as } n \to \infty.$$

4 Reversible Markov Chains

Suppose that you're watching a random walk. What if someone played that random walk backwards, could you tell the difference? Remember that we're not talking about a particular instance of the walk, but two different instances of the same programming... does it ever look like it could be *programmed the same way*, if you play it backwards?

For a walk that just alternates between two nodes it certainly does. But for a walk that just goes clockwise in a circle, you'd immediately be able to tell if it suddenly started going backwards. The key here is, does it go from i to j the same proportion of steps as it goes from j to i? Here's a formalization of this idea:

Definition 4 (Reversible Markov Chains)
A Markov chain is *reversible* if there exists an initial probability distribution $s = (s_1, \ldots, s_M)$ such that for any states i, j:

$$s_i P_{ij} = s_j P_{ji}.$$

The distribution s must be the stationary distribution.

*Proof:*If we multiply s and P, then the jth coefficient of the result will be:

$$\sum_i s_i P_{ij} = \sum_i s_j P_{ji} = s_j \sum_i P_{ji} = s_j.$$

□

Example 1 A simple random walk on a graph is a reversible Markov chain.

Proof: Let D be the sum of degrees of all the nodes and d_i be the degree of node i. Then $s_i = d_i/D$. The P_{ij} entry in the transition matrix is 0 if there is no edge between i and j, and $1/d_i$ if there is. Similarly, the entry P_{ji} is 0 if there is no edge and $1/d_j$ otherwise. So if there is no edge:

$$s_i P_{ij} = 0$$

$$s_j P_{ji} = 0$$

and if an edge exists:

$$s_i P_{ij} = \frac{d_i}{D}\frac{1}{d_i} = \frac{1}{D}$$

$$s_j P_{ji} = \frac{d_j}{D}\frac{1}{d_j} = \frac{1}{D}$$

In either case,

$$s_i P_{ij} = s_j P_{ji},$$

and so the walk is reversible. □