Stochastic Processes
1401-2
Dr. Peyvandi
Department of Mathematical Sciences
Amirhossein Mahmoudi - 98108779

Task 1

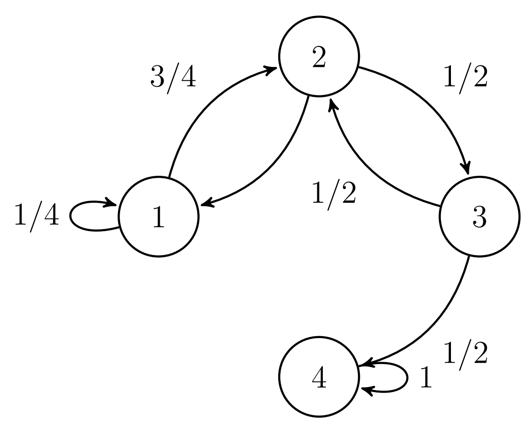
1. **Question:** Consider the voter model on an $n \times n$ table. Each cell from the table is colored red or blue, and at each step, a cell is randomly selected and its color is randomly changed to the color of one of its neighbors. Let $X(t)$ and $Y(t)$ be the state of the table and the number of red cells in step t, respectively. Also suppose that initially only the cell in the lower left corner of the table is red and the rest of the cells are blue.

   (a) Is $(Y(t))_t$ a Markov chain? (It is possible that the answer depends on n).

   (b) Note that $(X(t))_t$ is a Markov chain. Find all stable distributions of $(X(t))_t$.

   Answer:

   (a) To determine if $(Y_t)_t$ is a Markov chain, we need to consider if the future state Y_{t+1} depends only on the current state Y_t and not on the sequence of events that preceded it. For $(Y_t)_t$, which tracks the number of red cells at each step, the future state does depend on the current configuration of the table because the color of a randomly selected cell is changed based on the color of its neighbors. However, the exact configuration of the table is not encapsulated in Y_t alone, as it only provides the count of red cells, not their distribution. Therefore, without additional information about the arrangement of the cells, $(Y_t)_t$ by itself does not satisfy the Markov property and is not a Markov chain. For small tables, it might be possible to track the configuration of red cells in such a way that Y_t encapsulates all the necessary information to predict Y_{t+1} , thus satisfying the Markov property. for example for $n = 2$ there is the following graph for the Markov Chain:



   (b) The state $(X_t)_t$, however, does represent the entire configuration of the table at each step, and the next state depends only on the current configuration. Hence, $(X_t)_t$ is a Markov chain.

Finding all stable distributions of $(X_t)_t$ involves identifying the distributions that remain unchanged in the process over time. For the voter model, the stable distributions are the absorbing states of the system, which occur when all cells become the same color, resulting in no further changes. Since the model is defined on a finite $n \times n$ table and there's a non-zero probability that eventually one color will take over the entire table, the only stable distributions are the ones where all cells are red or all cells are blue.

2. **Question:** Consider playing top to random on decks numbered from 1 to n: at each step, we remove the top card of the deck and place it randomly in the deck. Suppose that X_t and Y_t are respectively the position of the cards and the number of the top card after the draw. Also, suppose that the deck of cards is sorted from 1 to n at the beginning.

   (a) Is $(Y_t)_t$ a Markov chain?

   (b) Note that $(X_t)_t$ is a Markov chain. Find all stable distributions of $(X_t)_t$.

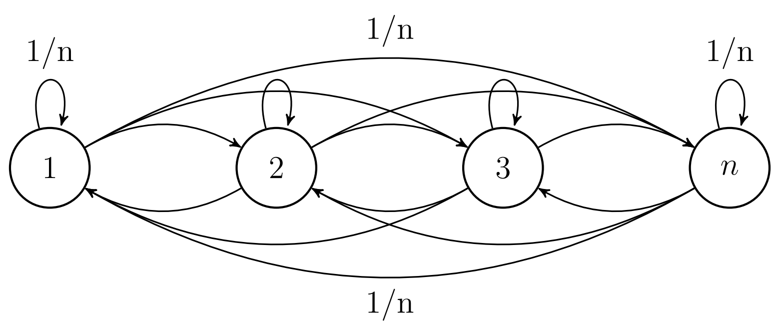
   (c) Show that one step after the card number n is seen on top of the deck, the deck will be fully randomized.

   Answer:

   (a) To prove that $(Y_t)_t$ is a Markov chain using the definition of the Markov property, we need to show that the probability of moving to the next state depends only on the current state and not on the past states.
   The Markov property states that for any times $s < t$, the future state Y_{t+1} is independent of the past states $Y_1, Y_2, ..., Y_s$ given the current state Y_t. Formally, this can be written as:

   $$P(Y_{t+1} = y | Y_1 = y_1, Y_2 = y_2, ..., Y_t = y_t) = P(Y_{t+1} = y | Y_t = y_t)$$

   For the top-to-random shuffle, Y_t represents the number of the top card after the t-th shuffle. When we perform the top-to-random shuffle, we remove the top card and place it randomly in the deck. The key point here is that the new position of the top card is determined solely by the random process of insertion and does not depend on the sequence of cards before the shuffle.
   Thus, the probability of any particular card y being on top at the next step $t+1$ depends only on the current top card y_t and not on the order of cards in the previous steps. This satisfies the Markov property, as the future state Y_{t+1} is conditionally independent of the past states given the current state Y_t.
   Therefore, $(Y_t)_t$ is indeed a Markov chain under the top-to-random shuffle process. The graph of this MC is a full with all edges begin $\frac{1}{n}$:



(b) The stable distributions of $(X_t)_t$ are the uniform distributions, where each card is equally likely to be in any position. As the shuffling continues indefinitely, the probability of any card being in any position tends towards $\frac{1}{n}$.

(c) Once the n-th card is seen on top, it indicates that every card has had the chance to be randomly inserted into the deck, leading to a uniform distribution of the card positions. Therefore, after the n-th card has been shuffled into the deck, the deck is considered to be fully randomized.

3. **Question:** A random walker on Z moves to the right with a probability of p and to the left with a probability of $1 - p$.

(a) Let A_n be the event that the random walker starts their movement from the number 1 and sees n before reaching 0. Also suppose that the event A is that the walker starts their movement from the number 1 and finally reaches 0. Show that:

$$\lim_{n\to\infty} P(A_n) = 1 - P(A).$$

(b) Using part (a), calculate the probability of the walker returning to their starting point.

(c) Consider the event A defined in part (a). Show:

$$P(A) = pP(A)^2 + (1 - p).$$

By solving this equation for $p = \frac{1}{2}$, it is concluded that the random walker is symmetric on $\mathbb{Z}$. Note that this equation has two solutions for $p \neq \frac{1}{2}$.

(d) Calculate the return probability for an asymmetric random walker. Using the law of large numbers, show that the probability of an asymmetric random walker returning to their starting point on $\mathbb{Z}$ is not 1. Using this fact and part (b), find the return probability.

Answer:

(a) To show that $\lim_{n\to\infty} P(A_n) = 1 - P(A)$, we consider the event A_n as the probability of reaching n before 0, starting from 1. As n approaches infinity, the probability of never reaching 0 approaches 1, and thus the probability of eventually reaching 0 is the complement of that, which is $1 - P(A)$.

(b) The walker can either:

- Move to 2 with probability p and then return to 1 with probability $P(A)$, or
- Move to 0 with probability $1 - p$ and then return to 1 with probability $P(A)$.

Let $P(B)$ be the probability of returning to 1. Then we have:

$$P(B) = p \cdot P(A) + (1 - p) \cdot P(A)$$

Since $P(A)$ is common in both terms, we can factor it out:

$$P(B) = P(A) \cdot (p + 1 - p)$$

$$P(B) = P(A)$$

So, the probability of the walker returning to the starting point 1 is the same as the probability $P(A)$ of the walker eventually reaching 0 when starting from 1. This is because, regardless of the first step, the walker must eventually pass through 1 to reach 0, and this passage through 1 is what we're interested in for $P(B)$.

(c) The probability $P(A)$ can be expressed as the sum of two mutually exclusive events: moving to the right (towards 2) and then eventually returning to 0, and moving directly to the left (towards 0). The first event has a probability of $pP(A)^2$ since the walker must reach 2 and then return to 1 twice (paths from 2 to 1 and then 1 to 0 which are $P(A)$). The second event has a probability of $1-p$. Thus, we have:

$$P(A) = pP(A)^2 + (1-p)$$

Solving this quadratic equation for $p \neq \frac{1}{2}$ yields two solutions, indicating two different probabilities for moving left or right that result in the same overall probability of reaching 0.

(d) Consider an asymmetric random walk where the probability of stepping to the right is p and to the left is $1-p$, with $p \neq \frac{1}{2}$. The expected step size is $E[X] = p-(1-p) = 2p-1$, which is not zero, indicating a drift in one direction.

According to the law of large numbers, for a large number of steps n, the position S_n of the random walker will be approximately $n(2p-1)$, diverging from the origin as n increases if $p \neq \frac{1}{2}$.

To solve the recurrence relation $P(A) = p \cdot P(A)^2 + (1-p)$ completely, we'll treat it as a quadratic equation in terms of $P(A)$. The equation can be rewritten as:

$$p \cdot P(A)^2 - P(A) + (1-p) = 0$$

This is a standard quadratic equation of the form $ax^2+bx+c=0$, where $a=p$, $b=-1$, and $c=1-p$. To find the roots of this equation, we use the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

Substituting our values into the quadratic formula gives us:

$$P(A) = \frac{-(-1) \pm \sqrt{(-1)^2 - 4 \cdot p \cdot (1-p)}}{2 \cdot p}$$

$$P(A) = \frac{1 \pm \sqrt{1-4p(1-p)}}{2p}$$

Since $P(A)$ represents a probability, it must be between 0 and 1. Therefore, we discard the negative root because it would give us a negative probability, which is not possible. We are left with:

$$P(A) = \frac{1 + \sqrt{1-4p(1-p)}}{2p}$$

This is the return probability for the asymmetric random walker on $\mathbb{Z}$. It's important to note that if $p = 0.5$, the square root term becomes zero, and the return probability

$P(A)$ becomes 1, which aligns with the result for a symmetric random walk. However, for $p \neq 0.5$, the return probability is less than 1, indicating the walker is less likely to return to the starting point.

4. **Question:** In a 6 x 6 table, we blacken each cell with probability p. Calculate the mathematical expectation of the perimeter of the black area! (Notice that a part of the perimeter of the black area may be placed on the border of the table and they are considered in the perimeter). For what value of p does the expected value of the perimeter become maximum?

   Answer:

   To calculate the expected perimeter E, we consider two types of borders: the outer borders of the table and the internal borders between cells.

   For the outer borders, there are 24 such borders (6 on each side of the square grid), and each contributes to the perimeter if the corresponding cell is blackened, which happens with probability p. Therefore, the expected contribution from these outer borders is $24p$.

   For the internal borders, there are 60 such borders (30 horizontal and 30 vertical, excluding the grid's outer borders), and each contributes to the perimeter if exactly one of the two adjacent cells is blackened. The probability of this happening is $2p(1-p)$, as either cell can be blackened while the other is not. Therefore, the expected contribution from the internal borders is $60 \times 2p(1-p)$.

   The total expected perimeter E is the sum of the expected contributions from the outer and internal borders:

   $$E = 24p + 120p(1-p)$$

   To maximize E, we take the derivative with respect to p and set it to zero:

   $$\frac{dE}{dp} = 24 - 240p + 120$$

   Solving for p when the derivative is zero gives us:

   $$24 - 240p + 120 = 0$$

   $$-240p = -144$$

   $$p = \frac{144}{240}$$

   $$p = \frac{3}{5}$$

   So, the value of p that maximizes the expected perimeter E is $\frac{3}{5}$ or 0.6.

5. **Question:** Consider n pieces of string. If we randomly tie two ends together out of the $2n$ free ends of the strings, what is the probability that as a result of this work, a large ring of length n will be formed?

   Answer:

   The probability of forming a single large loop without creating any smaller loops along the way can be calculated by considering the sequence of choices that lead to the formation of a single large loop from n strings.

   Let $P(n)$ be the probability of forming a single large loop. The first end of a string can be tied to any of the $2n-2$ other ends that are not its own. After this first tie, we have $2n-2$ ends left. The next end can be tied to any of the $2n-4$ ends that would not form a loop, and so on, until the last two ends are tied together, completing the single large loop.

   The total number of ways to tie all the ends without forming a loop at each step is:

$$(2n-2) \times (2n-4) \times \cdots \times 2$$

   The total number of ways to tie all the ends in any manner is:

$$(2n-1) \times (2n-3) \times \cdots \times 1$$

   Thus, the probability $P(n)$ is:

$$P(n) = \frac{(2n-2) \times (2n-4) \times \cdots \times 2}{(2n-1) \times (2n-3) \times \cdots \times 1} = \frac{(2n-2)!!}{(2n-1)!!}$$

6. **Question:** Show that the random walk on an infinite 3-regular tree is transient. Consider a random walk on the sequence $\{s_0, s_1, \ldots\}$, which represents the state of being at depth i relative to the origin, with transition probabilities: $s_{0,1} = 1$ and $s_{i,i+1} = \frac{2}{3}$, $s_{i,i-1} = \frac{1}{3}$. We need to show that this Markov chain is transient. Let $X_n = i$ mean that we are at depth i at time n. Complete this by showing that $\sum P(X_n = i | X_0 = i) < \infty$.

   Answer:

   To show that the random walk on an infinite 3-regular tree is transient, we need to demonstrate that the expected number of visits to any given state is finite. That is, for any state i, the series $\sum_{n=0}^{\infty} P(X_n = i | X_0 = i)$ converges.

   Given the transition probabilities $s_{i,i+1} = \frac{2}{3}$ and $s_{i,i-1} = \frac{1}{3}$, we can consider the probability of returning to the origin after n steps. The tree's 3-regularity implies a biased random walk favoring movement away from the origin.

   Using Upper Bound Convergence:

   Since we can only return to a point with an even number of steps, we consider paths of length $2k$ from a vertex to itself.

   For a vertex at depth i, the number of paths of length $2k$ that return to the same vertex can be bounded above by considering the number of ways to choose k steps that move towards the root and k steps that move away from the root. This is given by the binomial coefficient $\binom{2k}{k}$.

The probability of each of these paths is $\left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^k$ because there are k steps moving away from the root (with probability $\frac{2}{3}$) and k steps moving towards the root (with probability $\frac{1}{3}$). Therefore, the probability of returning to the vertex at depth i in $2k$ steps is:

$$P(X_{2k} = i | X_0 = i) \leq \binom{2k}{k} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^k$$

Using Stirling's approximation for the binomial coefficient, we have:

$$\binom{2k}{k} \sim \frac{4^k}{\sqrt{\pi k}}$$

Thus, the upper bound for $P(X_{2k} = i | X_0 = i)$ becomes:

$$P(X_{2k} = i | X_0 = i) \leq \frac{4^k}{\sqrt{\pi k}} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^k = \frac{1}{\sqrt{\pi k}} \left(\frac{8}{27}\right)^k$$

Since $\frac{8}{27} < 1$, the series $\sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi k}} \left(\frac{8}{27}\right)^k$ is convergent by the ratio test. Therefore, the expected number of visits to any state i is finite, which implies that the random walk on an infinite 3-regular tree is transient.

Using Generating Function:

We can calculate the expected number of visits to the origin using the generating function technique. Let $G(z)$ be the generating function for the number of visits to the origin. The coefficients of $G(z)$ give us the probabilities $P(X_n = i | X_0 = i)$. We need to show that $G(1)$, the sum of these probabilities, is finite. The generating function $G(z)$ for the number of visits to the origin satisfies the relation due to the Markov property and transition probabilities:

$$G(z) = 1 + \frac{2}{3} z G(z)^2$$

Solving this quadratic equation for $G(z)$ and evaluating at $z = 1$ gives us the expected number of visits to the origin. If $G(1)$ is finite, then the random walk is transient.

We have the generating function equation for a random walk on an infinite 3-regular tree:

$$G(z) = 1 + \frac{2}{3} z G(z)^2$$

Substituting $z = 1$, we get:

$$G(1) = 1 + \frac{2}{3} G(1)^2$$

This is a quadratic equation in terms of $G(1)$. Rearranging it, we have:

$$\frac{2}{3} G(1)^2 - G(1) + 1 = 0$$

Multiplying through by 3 to clear the fraction:

$$2G(1)^2 - 3G(1) + 3 = 0$$

Using the quadratic formula where $a = 2$, $b = -3$, and $c = 3$, we find:

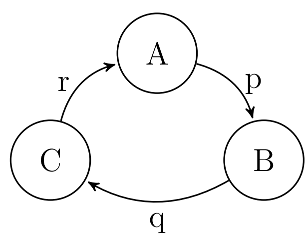
$$G(1) = \frac{-(-3) \pm \sqrt{(-3)^2 - 4 \cdot 2 \cdot 3}}{2 \cdot 2}$$

$$G(1) = \frac{3 \pm \sqrt{9 - 24}}{4}$$

$$G(1) = \frac{3 \pm \sqrt{-15}}{4}$$

Since the square root of a negative number is not real, this implies that $G(1)$ does not have a real solution, supporting the claim that the random walk is transient.

7. **Question:** Consider a particle performing a probabilistic circular walk on a graph consisting of three interconnected nodes, denoted as A, B, and C. The particle transitions from node A to B with probability p, from B to C with probability q, and from C back to A with probability r, all in a clockwise direction. Given these transition probabilities, determine the particle's long-term average velocity and the predominant direction of rotation.



Answer:

The long-term behavior can be described by the stationary distribution of the Markov chain representing the walk. The stationary distribution gives the long-term probability of the walker being at each node.

To find the stationary distribution, we can solve the system of linear equations given by the transition matrix of the Markov chain. The corrected transition matrix P for this 3-node graph is:

$$P = \begin{bmatrix} 0 & p & 1-r \\ 1-p & 0 & q \\ r & 1-q & 0 \end{bmatrix}$$

The stationary distribution π is a probability vector that satisfies $\pi P = \pi$ and $\sum \pi_i = 1$, where π_i is the probability of being at node i in the long term.

As for the speed, it is not typically defined for a random walk on a graph because the concept of speed implies a continuous movement over time, whereas a random walk consists of discrete steps. However, if we consider ”speed” as the rate of steps taken, it would simply be the number of steps per unit of time, which is constant if the walker takes one step at a time.

The rotation direction is determined by the probabilities p, q, and r. If they are all equal, the walker has no preferred direction in the long term. If one probability is higher than the others, the walker will have a higher chance of moving towards the node with the higher transition probability, giving a preferred direction of rotation.

To calculate the exact stationary distribution and any long-term directional bias, we would need the specific values of p, q, and r.