

# 01:640:478 - Homework 2

Pranav Tikkawar

October 6, 2024

## 1. Question 1.

Find the stationary probability for the markov chains with transitions matrices:

i)  $\begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{bmatrix}$

ii)  $\begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}$

**Solution:**

i) Let  $\pi = (\pi_1, \pi_2, \pi_3)$  be the stationary probability vector. Then, we have the following equations:

$$\pi_1 = 0.5\pi_1 + 0.2\pi_2 + 0.1\pi_3$$

$$\pi_2 = 0.4\pi_1 + 0.5\pi_2 + 0.3\pi_3$$

$$\pi_3 = 0.1\pi_1 + 0.3\pi_2 + 0.6\pi_3$$

$$\pi_1 + \pi_2 + \pi_3 = 1$$

Solving the above equations, we get

$$\pi = (.234043, .404255, .351702)$$

ii) Let  $\pi = (\pi_1, \pi_2, \pi_3)$  be the stationary probability vector. Then, we have the following equations:

$$\pi_1 = 0.5\pi_1 + 0.3\pi_2 + 0.2\pi_3$$

$$\pi_2 = 0.4\pi_1 + 0.4\pi_2 + 0.2\pi_3$$

$$\pi_3 = 0.1\pi_1 + 0.3\pi_2 + 0.6\pi_3$$

$$\pi_1 + \pi_2 + \pi_3 = 1$$

Solving the above equations, we get

$$\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

## 2. Question 2.

There are two coins. Coin 1 comes up heads with probability 0.6, and coin 2 with probability 0.5. A coin is continually flipped until it comes up tails, at which time it is put aside and we start flipping the other one.

i) What proportion of flips use coin 1?

ii) If we start the process with coin 1 what is the probability that coin 2 is used on the fifth flip?

iii) What proportion of flips land heads?

**Solution:**

i) To find the long run proportion of flips that use coin 1, we can use the following markov chain. Let  $X_n$  be the state of the markov chain at time  $n$ . Then, we have the following transition matrix:

$$P = \begin{bmatrix} .6 & .4 \\ .5 & .5 \end{bmatrix}$$

Let  $\pi = (\pi_0, \pi_1)$  be the stationary probability vector. Then, we have the following equations:

$$.6\pi_0 + .5\pi_1 = \pi_0$$

$$.4\pi_0 + .5\pi_1 = \pi_1$$

$$\pi_0 + \pi_1 = 1$$

Solving the above equations, we get

$$\pi = \left(\frac{5}{9}, \frac{4}{9}\right)$$

ii) To determine this probability, we need the probability of all the ways that starting with coin 1, then leads to coin 2 on the fifth flip. We can utilize the TPM to find this probability. We can find the probability of going from state 1 to state 2 in 4 steps. We get  $P_{01}^{(4)} = .4444$ .

iii) To find the proportion of flips that land we would utilize the long run probability and the likelihood of landing heads. We get that like probability to land on heads is

$$\pi_0(.6) + \pi_1(.5)$$

$$\frac{5}{9}(.6) + \frac{4}{9}(.5) = \frac{5}{9}$$

Thus the proportion is  $\frac{5}{9}$

3. Question 3. A university has made e-cycles available to the students to travel between campuses. In the first month of operation, it was found that 25% of students are using e-cycles while 75% are using the campus busses. Suppose that each month 10% of students using e-cycles go back to using the busses, while 30% of students using the buses switch to the e-cycles

- i) We can describe this using a 2-state Markov chain. Write the states and the transition probability matrix  $P$
- ii) Compute  $P^3$ , the three step transition probability matrix
- iii) What will be the fraction of students using e-cycles in the fourth month?
- iv) What will be the fraction of students using e-cycles in the long run?

**Solution:**

- i) The transition matrix is given by:

$$P = \begin{bmatrix} .9 & .1 \\ .3 & .7 \end{bmatrix}$$

Where the states are  $S = \{E, B\}$ , where E is the state of using e-cycles and B is the state of using buses.

- ii) We can find the three step transition probability matrix by computing  $P^3$ :

$$P^3 = \begin{bmatrix} .804 & .196 \\ .588 & .412 \end{bmatrix}$$

- iii) To find the fraction of students using e-cycles in the fourth month, we can use the three step transition probability matrix. then multiply it by the initial state vector. We get:

$$\begin{bmatrix} .25 & .75 \end{bmatrix} \begin{bmatrix} .804 & .196 \\ .588 & .412 \end{bmatrix} = \begin{bmatrix} .642 & .358 \end{bmatrix}$$

Thus, the fraction of students using e-cycles in the fourth month is .648.

- iv) The long run fraction of students using e-cycles can be found by finding the stationary probability vector. Let  $\pi = (\pi_1, \pi_2)$  be the stationary probability vector. Then, we have the following equations:

$$\begin{aligned} .9\pi_1 + .3\pi_2 &= \pi_1 \\ .1\pi_1 + .7\pi_2 &= \pi_2 \\ \pi_1 + \pi_2 &= 1 \end{aligned}$$

Solving the above equations, we get

$$\pi = (.75, .25)$$

Thus, the long run fraction of students using e-cycles is .75.

#### 4. Question 4.

Three out of every four trucks on the road are followed by a car, while only one out of every five cars is followed by a truck. Assuming the vehicles are appearing (to the observer) according to a 2-state Markov chain with state space car, truck, find the fraction of trucks on the road. **Solution:**

We can describe this using a 2-state Markov chain. Let  $X_n$  be the state of the markov chain at time  $n$ . Then, we have the following transition matrix:

$$P = \begin{bmatrix} .25 & .75 \\ .2 & .8 \end{bmatrix}$$

With states  $S = \{\text{car}, \text{truck}\}$ . Let  $\pi = (\pi_1, \pi_2)$  be the stationary probability vector. Then, we have the following equations:

$$\begin{aligned} .25\pi_1 + .2\pi_2 &= \pi_1 \\ .75\pi_1 + .8\pi_2 &= \pi_2 \\ \pi_1 + \pi_2 &= 1 \end{aligned}$$

Solving the above equations, we get

$$\pi = \left(\frac{4}{19}, \frac{15}{19}\right)$$

Thus the fraction of trucks on the road is  $\frac{15}{19}$ .

5. Question 5.

Consider three urns, one colored red, one white, and one blue. The red urn contains 1 red and 4 blue balls; the white urn contains 3 white balls, 2 red balls, and 2 blue balls; the blue urn contains 4 white balls, 3 red balls, and 2 blue balls. At the initial stage, a ball is randomly selected from the red urn and then returned to that urn. At every subsequent stage, a ball is randomly selected from the urn whose color is the same as that of the ball previously selected and is then returned to that urn. In the long run, what proportion of the selected balls are red? What proportion are white? What proportion are blue?

**Solution:**

Let  $X_n$  be the state of the markov chain at time  $n$ . Then, we have the following transition matrix:

$$P = \begin{bmatrix} \frac{1}{5} & 0 & \frac{4}{5} \\ \frac{3}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{3}{9} & \frac{4}{9} & \frac{2}{9} \end{bmatrix}$$

Where the states are  $S = \{\text{red}, \text{white}, \text{blue}\}$ . Where the color corresponds to the color of the ball last selected. Let  $\pi = (\pi_1, \pi_2, \pi_3)$  be the stationary probability vector. Then, we have the following equations:

$$\begin{aligned} \frac{1}{5}\pi_1 + \frac{2}{7}\pi_2 + \frac{3}{9}\pi_3 &= \pi_1 \\ 0\pi_1 + \frac{3}{7}\pi_2 + \frac{4}{9}\pi_3 &= \pi_2 \\ \frac{4}{5}\pi_1 + \frac{2}{7}\pi_2 + \frac{2}{9}\pi_3 &= \pi_3 \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

Solving the above equations, we get

$$\pi = \left(\frac{25}{89}, \frac{28}{89}, \frac{36}{89}\right)$$

Thus the proportion of selected balls that are red is  $\frac{25}{89}$ , the proportion of selected balls that are white is  $\frac{28}{89}$ , and the proportion of selected balls that are blue is  $\frac{36}{89}$ .

## 6. Question 6.

Consider a branching process having  $\mu < 1$ . Show that if  $X_0 = 1$ , then the expected number of individuals that ever exist in this population is given by  $1/(1 - \mu)$ . What if  $X_0 = n$ ?

**Solution:**

Let  $Z_n$  be the number of individuals in the population at time  $n$ . Then, we have the following equations:

$$\begin{aligned} E[Z_1] &= \mu \\ E[Z_2] &= \mu + \mu^2 \\ E[Z_3] &= \mu + \mu^2 + \mu^3 \\ &\vdots \\ E[Z_n] &= \mu + \mu^2 + \mu^3 + \cdots + \mu^n \end{aligned}$$

Thus, we have that

$$E[Z_n] = \frac{1 - \mu^{n+1}}{1 - \mu}$$

Thus, if  $X_0 = 1$ , then the expected number of individuals that ever exist in this population is given by  $1/(1 - \mu)$ . If  $X_0 = n$ , then the expected number of individuals that ever exist in this population is given by  $n/(1 - \mu)$ .

7. Question 7. There are four types of bases found in a DNA molecule: adenine (A), cytosine (C), guanine (G), and thymine (T). The Kimura 2-parameter Model is a Markov model of base substitution and is given by the following transition probability matrix (in A,G,C,T order along rows and columns):

$$\begin{bmatrix} 1 - p - 2r & p & r & r \\ p & 1 - p - 2r & r & r \\ q & q & 1 - p - 2q & p \\ q & q & p & 1 - p - 2q \end{bmatrix}$$

Show that the chain is reversible and find the stationary distribution.

**Solution:**

To show that the chain is reversible we need to show that  $\pi_i P_{ij} = \pi_j P_{ji}$  for all  $i, j$ . Let  $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$  We can get the equations:

$$\begin{aligned} \pi_1 &= 1 - p - 2r\pi_1 + p\pi_2 + q\pi_3 + q\pi_4 \\ \pi_2 &= p\pi_1 + 1 - p - 2r\pi_2 + q\pi_3 + q\pi_4 \\ \pi_3 &= r\pi_1 + r\pi_2 + 1 - p - 2q\pi_3 + p\pi_4 \\ \pi_4 &= r\pi_1 + r\pi_2 + p\pi_3 + 1 - p - 2q\pi_4 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \end{aligned}$$

Solving the above equations, we get

$$\pi = \left( \frac{q}{2q + 2r}, \frac{q}{2q + 2r}, \frac{r}{2q + 2r}, \frac{r}{2q + 2r} \right)$$

We can then verify that the chain is reversible by checking that  $\pi_i P_{ij} = \pi_j P_{ji}$  for all  $i, j$ . Thus, the chain is reversible and the stationary distribution is  $\left( \frac{q}{2q+2r}, \frac{q}{2q+2r}, \frac{r}{2q+2r}, \frac{r}{2q+2r} \right)$ .

8. Question 8. A DNA nucleotide has any of four values. A standard model for a mutational change of the nucleotide at a specific location is a Markov chain model that supposes that in going from period to period the nucleotide does not change with probability  $1 - 3\alpha$ , and if it does change then it is equally likely to change to any of the other three values, for some  $0 < \alpha < 1/3$ .
- Model this as a Markov chain consisting of four states 1, 2, 3, 4 corresponding to the four values. What is the transition probability matrix
  - Show that  $P_{1,1}^n = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n$
  - Find the long-run proportion of time the chain is in each state by solving the system of equations:  $\sum_j \pi_j P_{ji} = \pi_i, \sum_j \pi_j = 1$
  - Find the limiting probabilities for the chain using the equation you proved in the second part (and a little bit more). What do you observe?

**Solution:**

- i) The transition matrix is given by:

$$P = \begin{bmatrix} 1 - 3\alpha & \alpha & \alpha & \alpha \\ \alpha & 1 - 3\alpha & \alpha & \alpha \\ \alpha & \alpha & 1 - 3\alpha & \alpha \\ \alpha & \alpha & \alpha & 1 - 3\alpha \end{bmatrix}$$

- ii) We can prove this by utilizing a proof by induction:

Base case:  $n = 1$

$$P_{1,1}^1 = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha) = 1 - 3\alpha$$

Thus we need to assume that  $P_{1,1}^n = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n$

Inductive step:

$$P_{1,1}^{n+1} = \sum_{i=1}^4 P_{1,i}^n P_{i,1}$$

Since  $P_{1,2} = P_{1,3} = P_{1,4} = \alpha$ , we get

$$\begin{aligned} P_{1,1}^{n+1} &= (1 - 3\alpha)P_{1,1}^n + 3\alpha P_{1,2}^n \\ P_{1,1}^{n+1} &= (1 - 3\alpha) \left( \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n \right) + 3\alpha \left( \frac{1}{4} + \frac{1}{4}(1 - 4\alpha)^n \right) \\ P_{1,1}^{n+1} &= \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^{n+1} \end{aligned}$$

Thus,  $P_{1,1}^n = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n$

- 2 iii) Let  $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$  be the stationary probability vector. Then, we have the

following equations:

$$\begin{aligned}(1 - 3\alpha)\pi_1 + \alpha\pi_2 + \alpha\pi_3 + \alpha\pi_4 &= \pi_1 \\ \alpha\pi_1 + (1 - 3\alpha)\pi_2 + \alpha\pi_3 + \alpha\pi_4 &= \pi_2 \\ \alpha\pi_1 + \alpha\pi_2 + (1 - 3\alpha)\pi_3 + \alpha\pi_4 &= \pi_3 \\ \alpha\pi_1 + \alpha\pi_2 + \alpha\pi_3 + (1 - 3\alpha)\pi_4 &= \pi_4 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1\end{aligned}$$

Solving the above equations, we get

$$\pi = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$$

iv) To find the limiting probabilities for the chain, we can use the equation we proved in part ii. We get

$$\lim_{n \rightarrow \infty} P_{1,1}^n = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n = \frac{1}{4}$$

Thus, the limiting probabilities for the chain are  $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ . We observe that this is the same as the stationary probabilities.