MAS275 Probability Modelling

Exercises

Note: these questions are intended to be of variable difficulty. In particular:
Questions or part questions labelled (*) are intended to be a bit more challenging. Questions or part questions labelled (P) involve computer programming, so are intended for those who have taken MAS115 and are interested in using the ideas from that course to get more of a feel for what is going on in this one.

Introductory Markov chains questions

1. Two ponds, A and B, support a population of three ducks between them. At each time step, one duck (each duck is chosen with probability $1/3$ independently of which pond it is in and the number of ducks currently in that pond), will move from one pond to the other. Writing X_n for the number of ducks in pond A after n time steps and modelling $(X_n)_{n\in\mathbb{N}}$ as a Markov chain with state space $\{0, 1, 2, 3\}$,
   (a) write down the transition matrix;
   (b) if it is known that X_n takes each of its possible values with probability $1/4$, calculate the distribution of X_{n+1} and deduce the probability that pond B has no ducks at time $n+1$.
2. A particle moves on a circle through points which have been labelled $1, 2, 3, 4, 5$ (in a clockwise order). At each step it has a probability p of moving to the right (clockwise) and $1-p$ of moving to the left (anti-clockwise). Let X_n denote its location on the circle after the nth step.
   (a) Find the transition matrix of the Markov chain $(X_n)_{n\in\mathbb{N}}$.

(b) If the initial probability distribution is uniform and $p = 0.3$, calculate
$$P(X_0 = 1, X_1 = 2, X_2 = 3, X_3 = 4, X_4 = 5).$$

3. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{1, 2, 3\}$ has transition matrix P where
$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

   (a) If the initial distribution $\boldsymbol{\pi}^{(0)} = (\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3})$, calculate
$$P(X_0 = 3, X_1 = 2, X_2 = 1).$$

   (b) If the initial distribution $\boldsymbol{\pi}^{(0)} = (\frac{1}{2}\ \frac{1}{6}\ \frac{1}{3})$, what do you notice about $\boldsymbol{\pi}^{(1)}$?

4. Consider a Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{1, 2, 3\}$ and transition probability matrix P where
$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \end{pmatrix}$$

   Calculate P^2 and hence calculate

   (a) $P(X_3 = 2 \mid X_1 = 2)$;

   (b) $P(X_5 = 3 \mid X_3 = 1)$;

   (c) the distribution of X_2 if the distribution of X_0, $\boldsymbol{\pi}^{(0)}$, is $(\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3})$.

Further Markov chains questions

5. Let $S := \{000, 001, 010, 011, 100, 101, 110, 111\}$ where each state represents one corner of a cube in the obvious way, as shown in Figure 1.

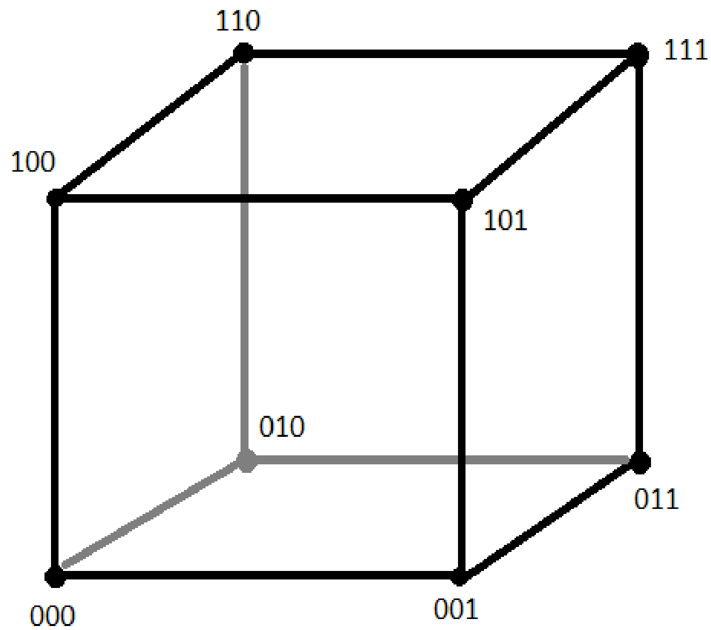

Figure 1: Labelling for question 5

Write down the transition matrix of a symmetric random walk on the cube (that is, the Markov chain in which from any corner, independently of the past, the next transition is equally likely to be to any one of the three adjacent corners).

6. (Bernoulli-Laplace model for diffusion.) Two containers each contain N particles, and of the total of $2N$ particles, N are white and N are black. At each time step a particle chosen at random from Container A changes place with a particle chosen independently at random from Container B.

   (a) Let $N = 2$. Find the transition matrix of the Markov chain which tracks the number of black particles in Container A.

(b) For general N, find expressions for the transition probabilities of, and sketch the transition matrix of, the Markov chain which tracks the number of black particles in Container A. [HINT: Think about what you did in (a). If you're still stuck, try the $N = 3$ case first.]

7. Two gamblers A and B are repeatedly playing two different games. Gambler A pays a non-returnable stake of one unit, and then either wins a prize of two units with probability 0.6 or wins no prize with probability 0.4; gambler B pays a non-returnable stake of one unit, and then either wins a prize of three units with probability 0.45 or wins no prize with probability 0.55. Either player stops if they run out of money. Sketch the transition matrices of the Markov chains representing the fortunes of the two gamblers.

8. In a tennis match, a stage has been reached in a tie break where the first player to gain a clear lead of two points wins the set. Independently, player 1 has probability p $(0 < p < 1)$ of winning each point regardless of who is serving. You can think of there as being five states of the game:

   - scores level;
   - player 1 one point ahead;
   - player 2 one point ahead;
   - player 1 has won;
   - player 2 has won.

   Modelling the sequence of states as a Markov chain on these five states, give the transition matrix.

9. This question aims to modify question 8 to take account of changes in service. The position of server changes from one player to the other whenever the total number of points which have been played is odd.

Independently, player 1 has probability p $(0 < p < 1)$ of winning each point when he is server, and player 2 has probability r $(0 < r < 1)$ of winning each point when he is server. Playing each point thus transfers the game from one "state" into another. Describe the eight possible states of the game, and sketch the transition matrix of the corresponding Markov chain.

10. Let P be the 3×3 stochastic matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}.$$

Show that $P^2 \neq P$ but that $P^3 = P$. Deduce that $P = P^3 = P^5 = \ldots$ and $P^2 = P^4 = P^6 = \ldots$, and that P^n does not converge as $n \to \infty$.

11. Find a diagonalisation of the 3×3 stochastic matrix

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix},$$

which is the transition matrix of a symmetric random walk on a triangle. Hence comment on the behaviour of P^n as $n \to \infty$.

12. Find the unique stationary distribution of the Markov chain with state space $\{1, 2, 3\}$ and the transition matrix in question 11.

13. A Markov chain has transition matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix},$$

as in question 3. Show that the distribution of $\boldsymbol{\pi}^{(0)}$ in question 3(b) is the unique stationary distribution of this Markov chain.

14. Consider the Markov chain with state space the non-negative integers and transition probabilities

$$p_{ij} = \begin{cases} \frac{1}{i+1} & j = i+1 \\ \frac{i}{i+1} & j = i-1 \\ 0 & \text{otherwise,} \end{cases}$$

so that the top left of the transition matrix looks like

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \dots \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 & \dots \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Write down the general form of equilibrium equation, and show (by substituting the given form into the equations) that these equations are satisfied by

$$\pi_j = k\frac{(j+1)}{j!} \quad \text{for } j = 0, 1, 2, \dots$$

for some constant k.

Renewal processes

15. In a renewal process, write down expressions for u_4 and u_5 in terms of $f_1, f_2, f_3, \dots$ which are analogous to those given in lectures for u_1, u_2 and u_3.

16. For each of the following sequences of f_n (for $n \geq 1$) determine whether the associated renewal process is transient, positive recurrent or null recurrent:

   (a) $f_n = \left(\frac{1}{3}\right)^n$.

(b) $f_n = \frac{6}{\pi^2 n^2}$.

(c) $f_n = \frac{4}{5}\left(\left(\frac{1}{2}\right)^n - \left(\frac{-1}{3}\right)^n\right)$.

[HINT: You may use the fact that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. In (b), trying to find the generating function $F(s)$ is not recommended.]

17. Recall that a random variable X with a Poisson distribution with parameter λ has probability function $f_n = P(X = n) = \frac{e^{-\lambda}\lambda^n}{n!}$ for integers $n \geq 0$.

    (a) Show that the generating function of the Poisson distribution with parameter λ is $F(s) = e^{\lambda(s-1)}$.

    (b) Use the facts about generating functions stated in the notes to show that

    i. if X and Y are independent random variables with Poisson distributions with parameters λ and μ respectively, then $X+Y$ has a Poisson distribution with parameter $\lambda + \mu$;

    ii. if X has a Poisson distribution with parameter λ then $E(X) = \lambda$.

18. In a renewal process, let

$$\begin{aligned} f_1 &= a \\ f_2 &= b \\ \text{and} \quad f_n &= 0 \quad \text{for } n = 3, 4, 5, \ldots \end{aligned}$$

where a and b are positive numbers with $a + b = 1$. Find the generating functions $F(s)$ and $U(s)$, and hence an explicit expression for u_n. Comment on its behaviour as n tends to infinity.

19. (a) In a transient renewal process, let R be the total number of renewals, excluding the one at time 0. Show that the distribution of R is given by $P(R = r) = f^r(1-f)$, where f is as defined in section 2.1. (Note:

the fact that $\sum_{r=0}^{\infty} P(R = r) = 1$ shows that R takes some finite value with probability 1, i.e. that renewals eventually stop, as stated in the notes.)

(b) In a simple random walk with $p < \frac{1}{2}$, starting from zero, find the probabilities that

   i. the walk never visits zero;

   ii. the walk visits zero exactly twice.

   (Note: We can find the value of f using

$$U(1) = \frac{1}{1 - F(1)}$$

   where $U(s)$ for this process is given in Theorem 5.)

20. A variant on a simple random walk is a Markov chain on $\mathbb{Z}$ which starts in state 0 and at each time step goes up by 2 with probability p and down by 1 with probability $1 - p$, where $0 < p < 1$, so that the transition probabilities are

$$p_{ij} = \begin{cases} p & j = i + 2 \\ 1 - p & j = i - 1 \\ 0 & \text{otherwise.} \end{cases}$$

    Find the period d of the renewal process formed by returns to state 0 of this process, and give a formula for u_{dm} for $m \geq 1$. (You do not need to try to simplify it.)

    (*) You might like to guess which values of p make this renewal process recurrent. Can you prove it? (HINT: it might help to look up *Stirling's approximation* for factorials.)

21. In a sequence of independent tosses of a fair coin, say that a renewal occurs after alternate heads, starting with the second. (For example, if the sequence starts

$$HTHTTHH\ldots$$

then renewals occur after the third and seventh tosses.)

(a) Treating this as a renewal process (without delay),

i. explain why

$$f_n = \begin{cases} 0 & n = 1 \\ (n-1)\left(\frac{1}{2}\right)^n & n \geq 2; \end{cases}$$

ii. find the mean time until the first renewal. (HINT: Question **??**(c) in the series handout should be useful here.)

(b) Show by induction that for $n \geq 1$ the probability that there has been an odd number of heads in the first n tosses is $\frac{1}{2}$. Hence show that in the above renewal process $u_n = \frac{1}{4}$ for all $n \geq 2$.

22. In a sequence of independent tosses of a fair coin, say that a renewal occurs after alternate heads, starting with the first. (For example, if the sequence starts

$$HTHTTHHTH\ldots$$

then renewals occur after the first, sixth and ninth tosses.) Treating this as a delayed renewal process,

(a) explain why f_n and u_n are the same as in question 21, for all n;

(b) explain why $b_n = (1/2)^n$ for all $n \geq 1$, and hence give the generating function of the delay, $B(s)$;

(c) use Theorem 7 to find the generating function $V(s)$, and use this to find v_n for all $n \geq 1$. Check that your answers for small n match what you would expect from the description of the process.

23. (a) In repeatedly tossing a fair coin, find the expected number of tosses required until the first completed occurrence of each of the following sequences:

i. HT;

ii. HH.

(b) In a sequence of tosses of a fair coin, suppose that the sequence $TTHH$ has just occurred. Use a delayed renewal process to find the expected time until the sequence HHH is next completed.

24. A monkey typing on a keyboard produces a sequence of letters, where the ith letter can be any of the 26 letters in the Latin alphabet, each with probability 1/26, and the letters are independent of each other. Find the expected number of letters required before the monkey produces each of the following 13 letter strings:

    (a) $ABCDEFGHIJKLM$;

    (b) $TOBEORNOTTOBE$.

25. (P) Try to write some code to simulate sequences of coin tosses and to estimate the expected values found in question 23a. (You will need to think about how many runs you need.) Do the simulation results match the theoretical values? Try estimating expected times until the appearance of some other similar sequences; do you see any patterns? Why didn't I suggest doing this with question 24?

Long term behaviour of Markov chains

26. In the "tennis" Markov chain with service mattering (question 9) find the classes, the closed classes and the periods of the classes.

27. (a) Explain why in the random walk on a cube (question 5) the states are periodic with period 2.

    (b) Give an example of a symmetric random walk on a graph where the states are aperiodic. (Quite small ones exist.)

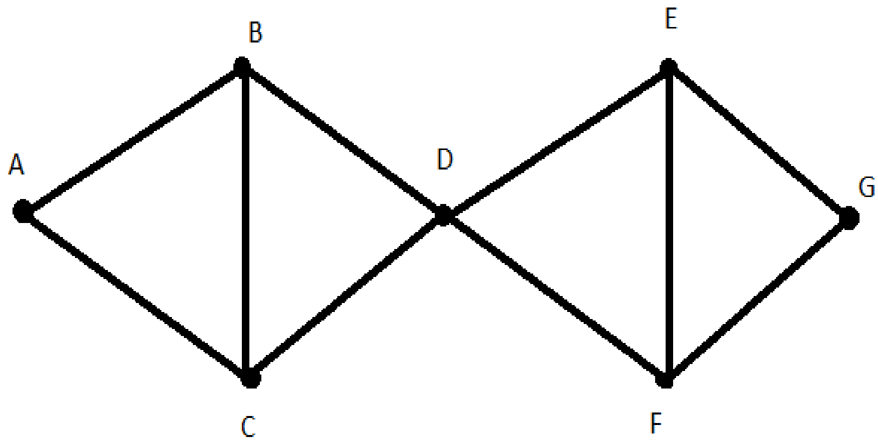

Figure 2: Graph for question 31

28. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{A, B, C\}$ has transition matrix P where

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix},$$

as in question 10. (This is a random walk on a graph with three vertices A, B, C and two edges AB and BC.) Assume the chain starts at time 0 in state A, and consider the non-delayed renewal process where renewals are returns to A.

(a) Explain why, using renewal theory notation, $u_n = 0$ if n is odd and $u_n = \frac{1}{2}$ if n is even and at least 2. (Recall that $u_0 = 1$ for any renewal process.)

(b) Find the generating function $U(s)$, and hence find the generating function $F(s)$.

(c) Find the expected time until the first return to state A.

29. Find (a) the classes, (b) the closed classes, (c) the periods of the classes, in the Markov chain with the following transition matrix. (Label the states from 1 to 8.)

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

30. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{1, 2, 3\}$ has transition matrix P where

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

(a) Verify that this Markov chain is irreducible and aperiodic.

(b) Find the unique stationary distribution of the chain.

(c) Using the theory on the limiting behaviour of Markov chains, give an approximate value for $P(X_{20} = 3 | X_0 = 1)$.

31. Consider the symmetric random walk on the graph consisting of seven nodes A, B, C, D, E, F and G linked by edges AB, AC, BC, BD, CD, DE, DF, EF, EG and FG. (See Figure 2.)

(a) Find the unique stationary distribution of this random walk. (HINT: you may find symmetry arguments useful.)

(b) Use the theory on long term behaviour of Markov chains to describe the limiting behaviour of the probabilities that the walk is in each of its possible states at time n, as $n \to \infty$.

(c) (*) The *degree* of a vertex in a graph is the number of neighbours it has. Do you notice a link between the degrees of the vertices in this graph and the stationary distribution of the symmetric random walk? Does something like this work for all (finite) graphs, as defined in section 1.5?

32. In a positive recurrent aperiodic renewal process, show that the distribution given by

$$\boldsymbol{\pi} = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \dots \end{pmatrix}$$

where

$$\pi_i = \frac{1}{\mu}(1 - f_1 - f_2 - \ldots - f_i),$$

for $i = 0, 1, 2, \ldots$, is indeed a stationary distribution for the forward recurrence time process, which is a Markov chain with transition matrix

$$P = \begin{pmatrix} f_1 & f_2 & f_3 & f_4 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

by showing that $\boldsymbol{\pi} P = \boldsymbol{\pi}$. (HINT: First find an expression for a general component of $\boldsymbol{\pi} P$, and do not worry that these vectors and matrices may be infinite in extent.)

33. Let (X_n) be a Markov chain in equilibrium, with transition matrix P and stationary distribution $\boldsymbol{\pi}$. Noting that

$$P(X_n = j, X_{n+1} = i) = \pi_j p_{ji},$$

and using the definition of conditional probability, find an expression for $P(X_n = j | X_{n+1} = i)$ for $n = 0, 1, 2, \ldots$ and $i, j \in S$.

Deduce that the Markov chain is **reversible**, in the sense that its behaviour backwards in time is the same as its behaviour forwards in time, if and only if

$$\pi_i p_{ij} = \pi_j p_{ji}$$

for all $i, j \in S$.

34. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{1, 2, 3\}$ has transition matrix P where

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

   (a) Find the classes of the chain, and state which classes are closed.

   (b) Show that although this chain is not irreducible it has a unique stationary distribution.

   (c) Conditional on the chain starting in state 2, give approximate probabilities for the chain being in each state at time n, for large n. (HINT: consider the chain on the "reduced" state space $\{2, 3\}$, and use standard results.)

35. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{1, 2, 3, 4\}$ has transition matrix P where

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

   (a) Find the classes of the chain, and give their periods.

   (b) Find all stationary distributions of the chain.

36. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{1, 2, 3, 4, 5\}$ has transition matrix P where

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(a) Show that the chain is irreducible, and find its period. (Hint: it may be easiest to find the period of state 1.)

(b) Let $\boldsymbol{\pi}^{(n)}$ be the row vector representing the distribution of the state of the chain at time n. If the chain is known to start in state 1, then find $\boldsymbol{\pi}^{(1)}$, $\boldsymbol{\pi}^{(2)}$ and $\boldsymbol{\pi}^{(3)}$. Hence give $\boldsymbol{\pi}^{(n)}$ for all $n \geq 1$, and deduce that $\boldsymbol{\pi}^{(n)}$ does not converge to a stationary distribution as $n \to \infty$.

37. (a) A renewal process has $f_1 = 1/2$, $f_2 = f_3 = 1/4$, and $f_n = 0$ for $n \geq 4$. Use the renewal theorem to find the limit of u_n as $n \to \infty$.

(b) A renewal process has $f_2 = 1/3$, $f_6 = 2/3$, and $f_n = 0$ for $n \notin \{2, 6\}$. Use the renewal theorem to find the limit of u_{2n} as $n \to \infty$.

38. (P)

(a) Try writing some code to simulate the Kruskal count card trick, discussed in lectures, and to estimate the probability of the trick working. Questions you might like to think about include how the assumptions (such as assigning a score of 5 to court cards) affect the probability of the trick working.

(b) Write some code to simulate Markov chains on three states with a given transition matrix. We have seen that under appropriate conditions on the transition matrix the probabilities that the chain is in each of its states tend to the values given by the stationary distribution. Investigate this convergence for some examples. How is the speed of convergence affected by the choice of matrix and by the starting point? What happens for Markov chains with more states?

Google PageRank

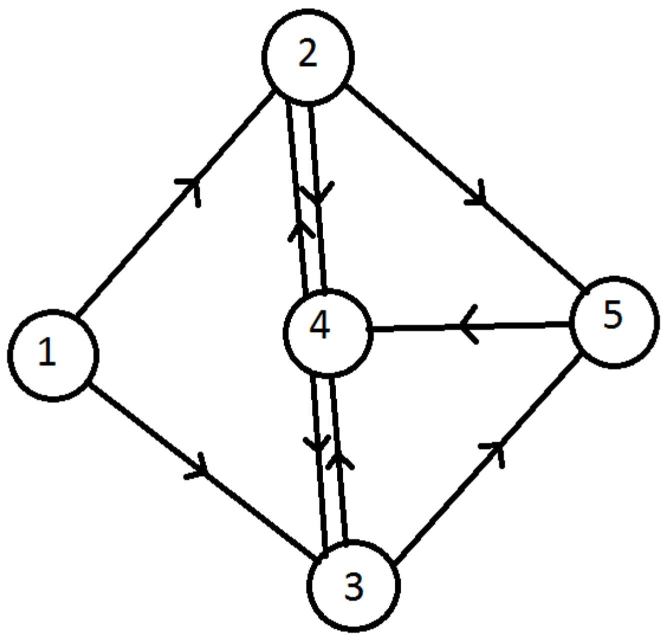

Figure 3: Mini web for question 39

39. Consider the mini web of pages represented in Figure 3, where double arrows represent pairs of pages with links in both directions, and single arrows represent pairs of pages with a link in the indicated direction only.

(a) Write down the transition matrix of the random surfer model on this network. Is it irreducible?

(b) Find the unique stationary distribution of the random surfer model on this network.

(c) Find the transition matrix of the PageRank model on this network with damping parameter $d = 6/11$. Verify that

$$\frac{1}{55}\begin{pmatrix} 5 & 11 & 11 & 17 & 11 \end{pmatrix}$$

gives the PageRanks of the pages (with this damping factor).

Absorption probabilities and times

40. Identify the closed classes in the Markov chain with the following transition matrix, and find the probabilities of absorption in each of the closed classes starting in each of the states which are not in closed classes. (Label the states from 1 to 7.)

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{2}{3} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

41. In the "tennis" model (question 9 and 26), with the states labelled as in the solutions, let q_i be the probability that Player 1 wins the set (i.e. absorption in state 7 occurs) starting in state i. Write down the

equations linking q_1, q_2, q_3, q_4, q_5 and q_6. By eliminating all but q_1 and q_2, derive two simultaneous linear equations for q_1 and q_2. Deduce that

$$q_1 = q_2 = \frac{p(1-r)}{p+r-2pr}.$$

42. In the random walk on the graph of question 31, find the probability, starting at B, of visiting G for the first time before visiting A for the first time.

43. In the Markov chain of question 40, find the expected number of steps to absorption in the closed classes, starting in each of the states which are not in closed classes.

44. In the "tennis" model (questions 9, 26 and 41), show that, starting with the scores level and regardless of who is serving, the expected number of points played until the game terminates is

$$e_1 = e_2 = \frac{2}{p+r-2pr}.$$

45. Each day until it reaches its destination, independently of its previous itinerary, a parcel sent by rail is equally likely to travel to any of the towns adjacent to the town where it is currently located on the network represented by the graph in which there are six nodes (towns) A, B, C, D, E and F linked by the following edges: AB, BC, BD, CD, CE, CF and DF.

    If its starting town is A and its destination is D, find the expected length of time until it arrives.

46. Use the Markov chain method to find the answers to question 23a.

47. Another possibly counterintuitive result about sequences of independent tosses of a fair coin is that the probability that the four character se-

quence $TTHH$ appears before the three character sequence HHH is 7/12. Try some of the following:

(a) (*) Use a method involving absorption probabilities in a Markov chain to show this. (HINT: the Markov chain needs to be a bit more complicated than in question 46; you need to keep track of the length of the current run of heads, and also some information about the length of the run of tails which preceded it. You also need to think carefully about what the initial state is.)

(b) (*) Can you use a renewal theory based argument to prove this? (HINT: the result of question 23b is useful.)

(c) (P) Try modifying code from question 25 to estimate this probability. Does 7/12 seem a likely value?

Poisson processes

48. Customers arrive at a facility according to a Poisson process having rate $\lambda = 4$. Let $N_{0,t}$ be the number of customers that have arrived up to time t. Determine the following probabilities

    (a) $P(N_{0,1} = 3)$

    (b) $P(N_{0,1} = 3, N_{0,3} = 12)$

    (c) $P(N_{0,1} = 3 | N_{0,3} = 12)$

    (d) $P(N_{0,3} = 12 | N_{0,1} = 3)$

49. Cases of a rare disease arrive at a hospital as a Poisson process with rate 0.7 per week.

    (a) What is the probability that there are more than two cases in the first two weeks of the year?

(b) What is the probability that there are two or more cases in the first week and eight or fewer in the next twelve weeks?

50. In the Poisson process with rate 1 per minute, find an approximate probability that there will be fewer than 50 occurrences in the first hour. (Recall that a Poisson random variable with a sufficiently large parameter λ can be approximated by a Normal random variable whose mean and variance are both λ.)

    Find the mean and variance of the length of time until the fiftieth occurrence, and explain why this random variable has an approximately normal distribution. Hence find an alternative approximation to the probability defined above.

    (Note: you can find the exact probability using the `ppois` function in R.)

51. Consider a constant rate Poisson process with rate λ, and let $S_2 = T_1 + T_2$ be the time of the second occurrence after time zero. By considering the probability that $N_{0,t} \leq 1$, find the probability density function of S_2.

52. Suppose that supernovas in our Galaxy occur as a Poisson process with rate $\lambda = 0.005$, where time is measured in years, and that each supernova in the Galaxy is observable from Earth with probability 0.8, independently of other supernovas and the times at which they occur.

    (a) Using the thinning property of Poisson processes, what distribution would the number of observable supernovas in a period of length t have?

    (b) Find the probability that there is at least one observable supernova in the Galaxy in a period of 100 years.

    (c) If two observable supernovas occur in the Galaxy between the years 2012 and 2612, find the probability that both occur before 2212.

53. A shop opens for eight hours, denoted by $(0, 8]$ with time measured in hours, and arrivals at the shop are modelled by the Poisson process with rate

$$\lambda(t) = \begin{cases} 1 + 2t & \text{for } 0 \le t \le 4; \\ 17 - 2t & \text{for } 4 \le t \le 8. \end{cases}$$

Find the probabilities of the following events.

(a) Neither the first hour nor the last hour of the opening period have any arrivals.

(b) Each of the eight one-hour periods which make up the opening period contains at least two arrivals.

(c) Given that there were five arrivals in the first two hours, at least two of them were in the first hour.

54. A model for the breaking of a world athletics record is that the clock times of new records form a Poisson process with variable rate $\lambda(t) = 1/t$.

(a) For $0 < u < v$, find the probability that there are no new records in the time interval $(u, v]$. Explain why this is consistent with the notion that the time of the best performance in the time interval $(0, v]$ is uniformly distributed on this interval.

(b) Find the distribution function of the length of time until the next new record, starting at clock time $u > 0$. Show that this length of time has infinite expected value.

55. An infinite forest contains both oak and ash trees, whose locations are given by independent spatial Poisson processes with densities $1/50$ and $1/40$ respectively, with area measured in square metres.

(a) If the region A has area a square metres, give an expression for the probability that both species of tree are present in A.

(b) Given a region B with area 4000 square metres, give an approximate value for the probability that there are more oak trees than ash trees in B.

56. A botanist is interested in the patterns of the location of trees of a certain species in a 1km by 1km square area of a forest. A typical density of trees of a certain species in this type of forest is λ per square kilometre.

    (a) The 1km by 1km square is divided into n^2 small squares, each $\frac{1}{n}$km by $\frac{1}{n}$km. Assuming that each of the n^2 small squares contains one tree of this species with probability θ (and otherwise does not contain one), and that the different small squares are independent of each other, what would be the distribution of the number of trees of this species in the large square? What value of θ would you use?

    (b) By considering the case where n is large, what distribution might you use as a model for the number of trees of this species in the large square?

Revision questions

57. An ordinary six-sided dice is rolled repeatedly. Consider the renewal process where a renewal occurs on the nth roll if it completes the sequence $1, 2, 3, 4, 5, 6$.

    (a) If u_n is the probability of a renewal occurring on the nth roll, explain why $u_n = 0$ if $1 \leq n \leq 5$, and give the value of u_n for $n \geq 6$.

    (b) Use the above to find the generating function $U(s)$, and hence use Theorem 4 to find the generating function $F(s)$.

    (c) Find the expected time until the first renewal.

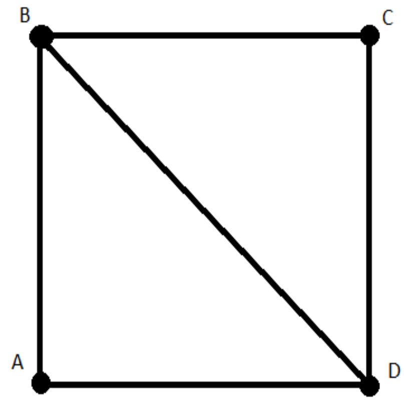

Figure 4: Graph for questions 58 and 59

58. A Markov chain $(X_n, n \in \mathbb{N})$ with state space $\{A, B, C, D\}$ has transition matrix P where

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

(This is a symmetric random walk on the graph with edges AB, AD, BC, BD, CD, i.e. a square with one diagonal edge added. See Figure 4)

(a) Find the unique stationary distribution of the chain.

(b) Describe the limiting behaviour of the probabilities $P(X_n = i)$, for $i = A, B, C, D$, as $n \to \infty$.

59. In the random walk of question 58,

(a) find the probability that, starting in state A, the walk reaches C before D;

(b) find the expected time until the walk reaches C, if it starts in A.

60. A Poisson process has variable rate $\lambda(t) = 2t$ for $t > 0$.

(a) Find the probability that there are exactly two occurrences in $(0, 1]$.

(b) Find the probability that, conditional that there are exactly two occurrences in $(0, 1]$, exactly one occurs in $(0, 3/4]$.

(c) If each occurrence is retained with probability $1/3$, independently of all other occurrences, then describe the process consisting of the retained occurrences, and give the probability that there are no retained occurrences in $(0, 1]$.