Imperial College London

MATH97083 MATH70045

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May-June 2022

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

Applied Probability

Date: 17 May 2022

Time: 09:00 – 11:30 (BST)

Time Allowed: 2:30 hours

Upload Time Allowed: 30 minutes

This paper has 5 Questions.

Candidates should start their solutions to each question on a new sheet of paper.

Each sheet of paper should have your CID, Question Number and Page Number on the top.

Only use 1 side of the paper.

Allow margins for marking.

Any required additional material(s) will be provided.

Credit will be given for all questions attempted.

Each question carries equal weight.

SUBMIT YOUR ANSWERS AS SEPARATE PDFs TO THE RELEVANT DROPBOXES ON BLACKBOARD (ONE FOR EACH QUESTION) WITH COMPLETED COVERSHEETS WITH YOUR CID NUMBER, QUESTION NUMBERS ANSWERED AND PAGE NUMBERS PER QUESTION.

Recall that $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \ldots\}$.

1. (a) Consider a discrete-time, time-homogeneous Markov chain $X = (X_n)_{n\in\mathbb{N}_0}$ on the state space $E = \{1, 2, 3, 4, 5\}$. We denote the (one-step) transition matrix by $\mathbf{P} = (p_{ij})_{i,j\in E}$ and the corresponding marginal distribution of X at time n by $\boldsymbol{\nu}^{(n)}$, for $n \in \mathbb{N}_0$.

Suppose that

$$\mathbf{P} = \begin{pmatrix} 0.2 & 0.3 & 0.5 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\nu}^{(0)} = (0.2, 0.6, 0.1, 0.1, 0).$$

(i) Draw the transition diagram of this Markov chain. (2 marks)

(ii) Find $\boldsymbol{\nu}^{(1)}$ and $\boldsymbol{\nu}^{(2)}$. (3 marks)

(iii) Specify the communicating classes and, for each class, determine whether it is transient, null recurrent or positive recurrent. (4 marks)

(iv) Find all possible stationary distributions. (3 marks)

(b) Consider a discrete-time, time-homogeneous Markov chain $X = (X_n)_{n\in\mathbb{N}_0}$ on the state space $E = \{1, 2, 3, 4\}$ with (one-step) transition matrix given by

$$\mathbf{P} = \begin{pmatrix} 0 & a & 0 & 1-a \\ a & 0 & 1-a & 0 \\ 0 & 1-a & 0 & a \\ 1-a & 0 & a & 0 \end{pmatrix}, \qquad \text{for } a \in (0, 1).$$

(i) Explain whether or not this Markov chain is irreducible. (2 marks)

(ii) Find all possible stationary distributions of the Markov chain. (2 marks)

(iii) Can you formulate suitable conditions such that this Markov chain is time-reversible? (4 marks)

(Total: 20 marks)

Recall that $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \ldots\}$.

2. (a) Consider a discrete-time, time-homogeneous Markov chain $X = (X_n)_{n \in \mathbb{N}_0}$ on the state space $E = \{1, 2, 3\}$ with (one-step) transition matrix given by

$$\mathbf{P} = \begin{pmatrix} 1-(a+b) & a & b \\ c & 1-(c+d) & d \\ e & f & 1-(e+f) \end{pmatrix}$$

and suppose that all entries in $\mathbf{P}$ are strictly positive.

(i) Find a necessary condition for this Markov chain to have a stationary distribution which satisfies the detailed balance equations. (5 marks)

(ii) Give an interpretation of the necessary condition you have derived in (i) and give a concrete example of a Markov chain (by choosing the values of a, b, c, d, e, f appropriately) which satisfies this necessary condition. (5 marks)

(b) Consider a discrete-time, time-homogeneous Markov chain $X = (X_n)_{n \in \mathbb{N}_0}$ on the state space $E = \mathbb{Z}$ with one-step transition probabilities for $i, j \in E$ given by

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

for $p \in (0, 1)$. Suppose that $X_0 = 0$. Define the stochastic process $Y = (Y_n)_{n \in \mathbb{N}_0}$ with $Y_n = \max_{0 \le k \le n} X_k$. Is Y a Markov chain? Justify your answer carefully. (4 marks)

(c) (i) Suppose that $X = (X_n)_{n \in \mathbb{N}_0}$ is a stochastic process taking values in a countable state space E. Suppose that there exist independent and identically distributed random variables $(Z_n)_{n \in \mathbb{N}}$ with values in an arbitrary state space F and a function $f : E \times F \to E$ such that, for each $n \in \mathbb{N}$, we can write $X_{n+1} = f(X_n, Z_{n+1})$. We further assume that X_0 is independent of $(Z_n)_{n \in \mathbb{N}}$. Show that X is a Markov chain on E.

(4 marks)

(ii) Give a practical example of a Markov chain that can be represented as in Part (i). *Hint: You can consider a suitable (sequence of) random experiments that can be modelled as a Markov chain. Purely theoretical examples will not be accepted.*

(2 marks)

(Total: 20 marks)

3. (a) A bakery specialising on birthday and wedding cakes produces cakes according to a Poisson process of rate $\lambda = 0.5$ per hour. The bakery is open between 4 am and 6 pm on workdays.

(i) How many cakes, on average, does this bakery produce on a typical workday? (2 marks)

(ii) What is the probability that the bakery can produce at least 10 cakes on one workday? (2 marks)

(iii) Consider the event that, on a given workday, the bakery produces 4 cakes between 4 am and 7 am. Conditional on this event, find the joint conditional cumulative distribution function of the random variables X and Y, where X denotes the time when the first cake was completed and Y denotes the time when the fourth cake was completed. (5 marks)

(b) Suppose that commuters arrive at a train station in the morning of a regular work day.

You are interested in modelling the number of arrivals between 6:30 am and 8:30 am by a stochastic process. You know that the expected number of arrivals between 6:30 am and 7:30 am is 10 commuters and the expected number of arrivals between 6:30 am and 8:30 am is 25 commuters.

(i) Explain why a Poisson process with rate $\lambda > 0$ is not suitable for this application. (1 mark)

(ii) Write down an inhomogeneous Poisson process which could be used to model the number of arrivals between 6:30 am and 8:30 am and justify your model choice. Remember to specify the rate/intensity function of the process explicitly. (4 marks)

(iii) In your model, find the probability that exactly three commuters arrive between 6:30 am and 7:00 am. (2 marks)

(iv) Work again with your model proposed in (ii). If exactly three commuters have arrived between 6:30 am and 7:00 am, based on your model, how many commuters, in total, do you expect to arrive at the train station between 6:30 am and 8:30 am. (2 marks)

(v) Discuss any limitations of your modelling assumptions. (2 marks)

(Total: 20 marks)

4. (a) Consider a continuous-time, time-homogeneous, minimal Markov chain $X = (X_t)_{t\geq 0}$ on the state space $E = \{1, 2, 3\}$ with generator given by

$$\mathbf{G} = \begin{pmatrix} -7 & 3 & 4 \\ 1/10 & -1/5 & 1/10 \\ 1 & 2 & -3 \end{pmatrix}.$$

(i) Draw the transition diagram of this Markov chain. (2 marks)

(ii) Find the (one-step) transition matrix of the embedded jump chain. (3 marks)

(iii) For each state in the state space, justify whether it is recurrent or transient for X. (2 marks)

(b) Suppose that $B = (B_t)_{t\geq 0}$ denotes a standard Brownian motion. Let $N = (N_t)_{t\geq 0}$ denote a Poisson process with rate $\lambda > 0$. Suppose that B and N are independent of each other. Define the stochastic processes $Y = (Y_t)_{t\geq 0}$ with $Y_t = t + N_t$ and $X = (X_t)_{t\geq 0}$ with $X_t = B_{Y_t}$ for all $t \geq 0$.

(i) Find $\mathrm{E}(Y_t)$ and $\mathrm{Var}(Y_t)$, for $t \geq 0$. (3 marks)

(ii) Find $\mathrm{E}(X_t)$ and $\mathrm{Var}(X_t)$, for $t \geq 0$. (5 marks)

(iii) Find $\mathrm{Cov}(X_s, X_t)$, for $s, t \geq 0$. (5 marks)

(Total: 20 marks)

5. This question refers to the additional reading material: D. J. Daley, D. G. Kendall, Stochastic Rumours, IMA Journal of Applied Mathematics, Volume 1, Issue 1, March 1965, Pages 42–55.

We refer to the continuous-time stochastic Markov chain model for the spread of a rumour in a population of size $N+1$ (for $N \in \mathbb{N}$) described in this paper as the $\mathbf{DK\text{-}model}$.

(a) Describe the DK-model and state and justify the corresponding infinitesimal transition probabilities.

(8 marks)

(b) Describe the random walk model associated with the DK-model and state and justify the corresponding transition probabilities. (6 marks)

(c) Let $P_{x,y}$ denote the probability that the two-dimensional random walk associated with the DK-model passes through the state (x, y), where x and y represent the number of ignorants and spreaders, respectively. Derive the 'forward' equations for $P_{x,y}$, i.e. express $P_{x,y}$ as a function of $P_{x+1,y-1}, P_{x,y+2}, P_{x,y+1}$ and justify your derivations, paying attention also to any constraints needed on x, y. (6 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2022

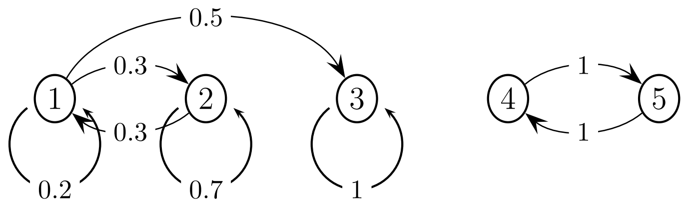
This paper is also taken for the relevant examination for the Associateship.

MATH60045/70045/97083

Applied Probability (Solutions)

Setter's signature	Checker's signature	Editor's signature
....	....	....

1. (a) (i) The transition diagram for this Markov chain is given by sim. seen ⇓



2, A

(ii) From lectures, we know that sim. seen ⇓

$$\begin{aligned}\boldsymbol{\nu}^{(1)} = \boldsymbol{\nu}^{(0)}\mathbf{P} &= (0.2 \cdot 0.2 + 0.6 \cdot 0.3, 0.2 \cdot 0.3 + 0.6 \cdot 0.7, 0.2 \cdot 0.5 + 0.1 \cdot 1, 0, 0.1)\\ &= (0.04 + 0.18, 0.06 + 0.42, 0.1 + 0.1, 0, 0.1)\\ &= (0.22, 0.48, 0.2, 0, 0.1).\end{aligned}$$

Moreover, we have 1, A

$$\begin{aligned}\boldsymbol{\nu}^{(2)} = \boldsymbol{\nu}^{(0)}\mathbf{P}^2 = \boldsymbol{\nu}^{(1)}\mathbf{P}&\\ &= (0.22 \cdot 0.2 + 0.48 \cdot 0.3, 0.22 \cdot 0.3 + 0.48 \cdot 0.7, 0.22 \cdot 0.5 + 0.2 \cdot 1, 0.1, 0)\\ &= (0.044 + 0.144, 0.066 + 0.336, 0.11 + 0.2, 0.1, 0)\\ &= (0.188, 0.402, 0.31, 0.1, 0).\end{aligned}$$

2, A

(iii) From the transition matrix (or the diagram), we read off that there are three communicating classes: $T_1 = \{1, 2\}, C_1 = \{3\}, C_2 = \{4, 5\}$. We note that T_1 is not closed, hence transient. Both C_1 and C_2 are finite and closed and hence positive recurrent (by a result from lectures). sim. seen ⇓

4, A

(iv) We know from a result from lectures that the stationary distribution is not unique in this case since we have two positive recurrent communication classes. We recall that the elements of the stationary distribution associated with transient states are 0. sim. seen ⇓

Next, we solve two systems of equations (one is trivial) to find the stationary distributions associated with C_1 and C_2.

Using the notation from lectures, we note that any π_{C_1} satisfies $\pi_{C_1} = \pi_{C_1}\mathbf{P}_{C_1}$ since $\mathbf{P}_{C_1} = 1$.

Also, $\pi_{C_2} = \pi_{C_2}\mathbf{P}_{C_2} \iff \pi_{C_2} = \pi_{C_2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which implies that the first and second component of π_{C_2} are identical.

Finally, we need to ensure that the components of the stationary distribution are non-negative and sum up to one. Hence, we obtain that all stationary distributions of the Markov chain are of the form

$$\pi = (0, 0, a, b, b), \quad \text{for any } a, b \geq 0 \text{ such that } a + 2b = 1.$$

[Alternative representations of the above solution are possible, e.g.:

$$\pi = a \cdot (0, 0, 1, 0, 0) + b \cdot (0, 0, 0, 0.5, 0.5), \quad \text{for any } a, b \geq 0 \text{ such that } a + b = 1.$$

] 3, A

(b) (i) Yes, this Markov chain is irreducible since all states communicate with each other. To see that, note that we assume that a and hence also $1-a$ are in $(0,1)$. By definition, each state communicates with itself. Let us show that state 1 and 2 communicate: We note that $p_{12} = a > 0$, hence 2 is accessible from 1. Also, since $p_{23} = 1-a > 0, p_{34} = a > 0, p_{41} = 1-a > 0$, we get (by Chapman-Kolmogorov) that $p_{21}(3) \geq p_{23}p_{34}p_{41} > 0$, which implies that 1 is accessible from 2. Analogous arguments can be applied to show communication between the remaining states since $1 \to 2, 2 \to 3, 3 \to 4, 4 \to 1$. sim. seen ⇓ 2, A

(ii) Since the Markov chain is irreducible and the state space is finite, we know from a result from the lectures that there is a unique stationary distribution. We observe that the transition matrix is doubly-stochastic. Hence, according to lectures, the discrete uniform distribution on E is the unique stationary distribution. I.e. $\boldsymbol{\pi} = (0.25, 0.25, 0.25, 0.25)$ is the unique stationary distribution. sim. seen ⇓ 2, B

(iii) We note that the Markov chain is irreducible and has a unique stationary distribution denoted by $\boldsymbol{\pi} = (0.25, 0.25, 0.25, 0.25)$. sim. seen ⇓
Suppose that the marginal distributions of this Markov chain are given by $\boldsymbol{\nu}^{(n)} = \boldsymbol{\pi}$ for all $n \in \mathbb{N}_0$. 2, B
According to lectures, the Markov chain is time-reversible if and only if the detailed balance equations hold:

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j \in E.$$

Here we need that 1, B

$$0.25 p_{ij} = 0.25 p_{ji} \quad \forall i, j \in E \iff p_{ij} = p_{ji} \quad \forall i, j \in E.$$

The latter condition is satisfied since $\mathbf{P}$ is a symmetric matrix. 1, B

2. (a) (i) Since all entries of the transition matrix are positive, all states communicate with each other and, hence, the Markov chain is irreducible. Since, in addition, the state space is finite, we know that there exists a unique stationary distribution which we denote by $\boldsymbol{\pi}$ (with all entries being strictly positive). The detailed balance equations hold iff $\pi_i p_{ij} = \pi_j p_{ji}$ for all $i, j \in E$. **unseen ⇓**

In order for the detailed balance equations to hold, we require the following conditions:

$$\pi_1 a = \pi_2 c, \quad \pi_1 b = \pi_3 e, \quad (\pi_2 c = \pi_1 a,) \quad \pi_2 d = \pi_3 f, \quad (\pi_3 e = \pi_1 b,) \quad (\pi_3 f = \pi_2 d),$$

where we put the redundant equations in brackets. The detailed balance equations for $i = j$ are trivially satisfied. **2, B**

Rearranging these three equations leads to

$$\pi_2 = \frac{a}{c}\pi_1, \qquad \pi_3 = \frac{b}{e}\pi_1, \qquad \pi_3 = \frac{d}{f}\pi_2.$$

We plug the first and second equation into the third equation and get

$$\frac{b}{e}\pi_1 = \frac{d}{f}\frac{a}{c}\pi_1 \iff 1 = \frac{eda}{bfc} \iff ade = bfc.$$

Hence a necessary condition for the stationary distribution to satisfy the detailed balance equations is that $ade = bfc$. **3, B**

(ii) Interpretation: We note that the necessary condition is equivalent to $p_{12}p_{23}p_{31} = p_{13}p_{32}p_{21}$. **2, C**

I.e. the probabilities of moving around the loop 1-2-3 in either direction are equal. **2, D**

[This condition is called a cycle condition.]

For the example, we need to ensure that all entries in the matrix are positive and that $ade = bfc$. A valid choice is given by $a = b = c = d = e = f = 1/3$. Then

$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix},$$

and $ade = bfc = 1/27$. **1, B**

(b) The process Y is not a Markov chain. We show this by considering two different paths which lead to the same state and show that the corresponding conditional probabilities are not identical. $\boxed{1, \mathrm{B}}$

Consider the sample path $(X_0, X_1, X_2, X_3, X_4) = (0, -1, 0, 1, 2)$ and the corresponding maximum $(Y_0, Y_1, Y_2, Y_3, Y_4) = (0, 0, 0, 1, 2)$. Then

$$\mathrm{P}(Y_5 = 3 | (Y_0, Y_1, Y_2, Y_3, Y_4) = (0, 0, 0, 1, 2)) = p.$$

Now we consider another sample path of the maximum process given by $(Y_0, Y_1, Y_2, Y_3, Y_4) = (0, 1, 2, 2, 2)$. Such a path could be achieved, by two different sample paths of X: $(X_0, X_1, X_2, X_3, X_4) = (0, 1, 2, 1, 2)$ or $(X_0, X_1, X_2, X_3, X_4) = (0, 1, 2, 1, 0)$. I.e. we could have either $X_4 = 2$ or $X_4 = 0$. Hence

$$\mathrm{P}(Y_5 = 3 | (Y_0, Y_1, Y_2, Y_3, Y_4) = (0, 1, 2, 2, 2)) < p.$$

$\boxed{2, \mathrm{D}}$

Since the path to $Y_4 = 2$ impacts the conditional probability of the next transition, the process Y does not satisfy the Markov property. $\boxed{1, \mathrm{B}}$

(c) (i) From the definition we note that X is a stochastic process on E. It remains to show the Markov property. unseen $\Downarrow$

When iterating the recurrence equation, we get

$$X_n = f(X_{n-1}, Z_n) = f(f(X_{n-2}, Z_{n-1}), Z_n) = \dots .$$

I.e. for all $n \in \mathbb{N}$, there exists a function g_n such that $X_n = g_n(X_0, Z_1, \dots, Z_n)$. 2, A

Hence, for any $j, i, i_{n-1}, \dots, i_0 \in E$, we have

$$\begin{aligned}
&\mathrm{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&= \mathrm{P}(f(X_n, Z_{n+1}) = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&= \mathrm{P}(f(i, Z_{n+1}) = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathrm{P}(f(i, Z_{n+1}) = j),
\end{aligned}$$

since the event $\{X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\}$ can be expressed in terms of $X_0, Z_1, \dots, Z_n$, which are independent of Z_{n+1}. Similarly, $\mathrm{P}(X_{n+1} = j | X_n = i) = \mathrm{P}(f(X_n, Z_{n+1}) = j | X_n = i) = \mathrm{P}(f(i, Z_{n+1}) = j | X_n = i) = \mathrm{P}(f(i, Z_{n+1}) = j)$ due to the independence of Z_{n+1} and X_n (since X_n can be expressed as a function of $X_0, Z_1, \dots, Z_n$). Hence, for any $n \in \mathbb{N}$ and for any $j, i, i_{n-1}, \dots, i_0 \in E$, $\mathrm{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathrm{P}(X_{n+1} = j | X_n = i)$. I.e. the Markov property holds. 2, B

(ii) We can recall the following example from the problem class: Let X_n be the maximum reading obtained in the first n rolls of a fair die (for $n \in \mathbb{N}$). Let D_n be the score of the die at time n. Then D_n is a uniform random variable on $\{1, \dots, 6\}$. Then, for $n \in \mathbb{N}$, we have the representation $X_n = \max\{D_1, \dots, D_n\} = \max\{X_{n-1}, D_n\}$, i.e. the function f is given by $f(x, y) = \max\{x, y\}$. The initial value can be chosen as $X_0 = 1$, independent of $(D_n)_{n \in \mathbb{N}}$. seen $\Downarrow$ 2, A

[Many other examples are possible and, if correct, will be given full credit.]

3. (a) (i) Let $N = (N_t)_{t\geq 0}$ denote the Poisson process of rate $\lambda = 0.5$ per hour describing the number of cakes produced by the bakery. The time is recorded in hours and $t = 0$ corresponds to 4 am. We note that a workday is 14 hours long. meth seen ⇓

We note that $N_t \sim \text{Poi}(\lambda t)$ and that $\text{E}(N_t) = \lambda t$. Hence the average number of cakes produced on a weekday is given by

$$\text{E}(N_{14}) = 14 \cdot 0.5 = 7.$$

2, A

(ii) As above, we note that $N_{14} \sim \text{Poi}(7)$. Hence, the probability that the bakery can produce at least 10 cakes on a workday is given by meth seen ⇓

$$\text{P}(N_{14} \geq 10) = \sum_{k=10}^{\infty} \text{P}(N_{14} = k) = \sum_{k=10}^{\infty} \frac{7^k}{k!} e^{-7}.$$

2, A

(iii) Here we condition on the event that $N_3 = 4$. Let J_i denote the time of the ith jump of the Poisson process. From lectures we know that the conditional joint density of the first four jump times is given by unseen ⇓

$$f_{(J_1,J_2,J_3,J_4)|N_3=4}(t_1, t_2, t_3, t_4|N_3 = 4) = \frac{4!}{3^4},$$

if $0 < t_1 < t_2 < t_3 < t_4 \leq 3$ and 0 otherwise. We need to integrate out over t_2 and t_3 to find the corresponding joint conditional density of J_1 and J_4:

$$f_{(J_1,J_4)|N_3=4}(t_1, t_4|N_3 = 4) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_{(J_1,J_2,J_3,J_4)|N_3=4}(t_1, t_2, t_3, t_4|N_3 = 4) dt_2 dt_3$$
$$= \int_{t_1}^{t_4} \int_{t_1}^{t_3} \frac{4!}{t^3} dt_2 dt_3 = \frac{4!}{3^4} \int_{t_1}^{t_4} (t_3 - t_1) dt_3 = \frac{4!}{3^4} \frac{t_3^2}{2} - t_1 t_3 \bigg|_{t_3=t_1}^{t_4}$$
$$= \frac{4!}{3^4}(t_4^2/2 - t_1 t_4 - t_1^2/2 + t_1^2) = \frac{4!}{3^4}\left(\frac{t_4^2}{2} - t_1 t_4 + \frac{t_1^2}{2}\right) = \frac{4}{27}\left(t_4^2 - 2t_1 t_4 + t_1^2\right),$$

if $0 < t_1 < t_4 \leq 3$ and 0 otherwise. 3, D

We can now compute the conditional cumulative distribution function as follows: For $0 < t_1 < t_4 \leq 3$, we have

$$F_{(J_1,J_4)|N_3=4}(t_1, t_4|N_3 = 4) = \int_0^{t_1} \int_{u_1}^{t_4} \frac{4}{27}\left(u_4^2 - 2u_1 u_4 + u_1^2\right) du_4 du_1$$
$$= \int_0^{t_1} \frac{4}{27}\left(\frac{1}{3}u_4^3 - u_1 u_4^2 + u_1^2 u_4\right)\bigg|_{u_4=u_1}^{t_4} du_1$$
$$= \int_0^{t_1} \frac{4}{27}\left(\frac{1}{3}t_4^3 - u_1 t_4^2 + u_1^2 t_4 - \frac{1}{3}u_1^3 + u_1^3 - u_1^3\right) du_1$$
$$= \int_0^{t_1} \frac{4}{27}\left(\frac{1}{3}t_4^3 - u_1 t_4^2 + u_1^2 t_4 - \frac{1}{3}u_1^3\right) du_1$$
$$= \frac{4}{81}t_1 t_4^3 - \frac{2}{27}t_1^2 t_4^2 + \frac{4}{81}t_1^3 t_4 - \frac{1}{81}t_1^4$$
$$= \frac{1}{81}\left[4t_1 t_4^3 - 6t_1^2 t_4^2 + 4t_1^3 t_4 - t_1^4\right].$$

Note that $X = J_1$ and $Y = J_4$. The conditional joint distribution function of X and Y is given as follows:
For $0 < x < y \leq 3$, we have

$$\mathrm{P}(X \leq x, Y \leq y | N_3 = 4) = \frac{1}{81}\left[4xy^3 - 6x^2y^2 + 4x^3y - x^4\right].$$

1, D

For $0 < x < 3 \leq y$, we have

$$\begin{aligned}\mathrm{P}(X \leq x, Y \leq y | N_3 = 4) &= \frac{1}{81}\left[4x3^3 - 6x^2 3^2 + 4x^3 3 - x^4\right] \\ &= \frac{4}{3}x - \frac{2}{3}x^2 + \frac{4}{27}x^3 - \frac{1}{81}x^4,\end{aligned}$$

For $0 < 3 < x < y$, we have

$$\mathrm{P}(X \leq x, Y \leq y | N_3 = 4) = 1.$$

For any other cases we get that $\mathrm{P}(X \leq x, Y \leq y | N_3 = 4) = 0$. 1, D

(b) (i) Let $N = (N_t)_{t \geq 0}$ denote a stochastic process, where N_t describes the total number of commuters which have arrived at the train station in the time interval $[0, t]$. Here the time $t = 0$ corresponds to 6:30 am and we measure time in hours. [Other time units (e.g. minutes) can be considered; the units should be clearly stated.] unseen ⇓

Suppose that N is a Poisson process with rate $\lambda > 0$. Then $N_t \sim \mathrm{Poi}(\lambda t)$ and $\mathrm{E}(N_t) = \lambda t$.

We are given the information that the expected number of arrivals between 6:30 am and 7:30 am is 10 commuters. This would imply that $\mathrm{E}(N_1) = \lambda \cdot 1 = 10 \iff \lambda = 10$. Similarly, the expected number of arrivals between 6:30 am and 8:30 am is given as 25, which implies that $\mathrm{E}(N_2) = \lambda \cdot 2 = 25 \iff \lambda = 12.5$. Hence there is no constant $\lambda > 0$ which satisfies both equations simultaneously, hence the Poisson model is not suitable. 1, B

(ii) Given the fact that a constant rate λ does not seem suitable for our application, we can consider an inhomogeneous Poisson process, where the rate $\lambda(t)$ can vary with time. Various specifications for the rate function are possible. Given the additional information, we might expect that $\lambda(t)$ should be increasing in time (during the time period considered). unseen ⇓ 1, B

Arguably the easiest functional form is given by a linear function. So suppose that $\lambda(t) = a + bt$ for constants $a, b \in \mathbb{R}$. Here we shall measure time in hours.

1, C

Let $m(t) = \int_0^t \lambda(s)ds = as + \frac{b}{2}s^2|_0^t = at + \frac{b}{2}t^2$. If N is a non-homogeneous Poisson process with rate function $\lambda(t)$, then $N_t \sim \mathrm{Poi}(m(t))$. Hence $\mathrm{E}(N_t) = m(t)$.

Given the information in the question, we set

$$\mathrm{E}(N_1) = a + \frac{b}{2} \stackrel{!}{=} 10, \qquad \mathrm{E}(N_2) = 2a + 2b \stackrel{!}{=} 25.$$

Solving for a, b leads to $a = 7.5, b = 5$. Hence, we suggest using an inhomogeneous Poisson process with rate $\lambda(t) = 7.5 + 5t$ (measured in hours) in this application. 2, C

(iii) Since $N_{0.5} \sim \mathrm{Poi}(m(0.5))$ with $m(0.5) = 15/4 + 5/2^3 = 35/8$, the probability that exactly three commuters arrive between 6:30 am and 7 am is given by $\boxed{\text{meth seen} \Downarrow}$

$$\mathrm{P}(N_{0.5} = 3) = \frac{m(0.5)^3}{3!} e^{-m(0.5)} = \frac{(35/8)^3}{3!} e^{-35/8}.$$

[Note that the results might differ depending on the model choice.] $\boxed{2, \mathrm{B}}$

(iv) Let $s = 0.5$ (corresponding to 7:00 am) and $t = 2$ (corresponding to 8:30 am). Then we can use the conditional expectation $\boxed{\text{unseen} \Downarrow}$

$$\mathrm{E}(N_t|N_s) = \mathrm{E}(N_t - N_s + N_s|N_s) = \mathrm{E}(N_t - N_s|N_s) + \mathrm{E}(N_s|N_s) = \mathrm{E}(N_t - N_s) + N_s$$

to compute the forecast, where we applied well-known properties of the conditional expectation and the fact that N has independent increments.
Then $\mathrm{E}(N_t|N_s = 3) = m(t) - m(s) + 3$.
[Students may start their answer from the next line.]
Given $N_{0.5} = 3$, in total, we expect to see

$$3 + \mathrm{E}(N_2 - N_{0.5}) = 3 + m(2) - m(0.5) = 3 + 25 - 35/8 = 23 + 5/8 = 23.625$$

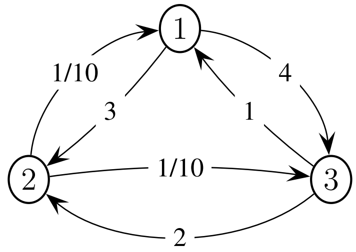
arriving at the train station between 6:30 am and 8:30 am. $\boxed{2, \mathrm{D}}$

(v) Possible limitations include: $\boxed{\text{meth seen} \Downarrow}$

- The single-arrival property might not be satisfied, e.g. people might be arriving in pairs or groups (e.g. coming from a connecting bus).
- The choice of a linear rate function might be too simplistic. E.g. as we get closer to 8:30 am the rate might start slowing down. This could be the case if we consider a train station which is say a 45-minute commute away from many workplaces and a standard workday is assumed to start at 9am.
- It might be better to consider a stochastic rate function to allow for more flexibility in the model. For instance, the rate might depend on weather or other traffic conditions.

[At least two limitations should be mentioned for full marks.] $\boxed{2, \mathrm{C}}$

4. (a) (i) The transition diagram is given by — meth seen ⇓



2, A

(ii) According to lectures, the transition probabilities of the corresponding embedded jump chain are given by $p_{ij} = -g_{ij}/g_{ii}$ for all $i, j \in E$ provided that $g_{ii} \neq 0$. Hence, the transition matrix of the embedded jump chain is given by — meth seen ⇓

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{3}{7} & \frac{4}{7} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}.$$

3, A

(iii) We observe that the jump chain has one communicating class, i.e. it is an irreducible Markov chain. Since the class is finite and closed, we conclude that the class and hence all states are positive recurrent. — meth seen ⇓

We know that if a state is recurrent (transient) for the jump chain, then it is recurrent (transient) for the continuous-time Markov chain. So we conclude that all states are recurrent for the continuous-time Markov chain. — 2, A

(b) (i) Let $t \geq 0$. We recall that $\mathrm{E}(N_t) = \mathrm{Var}(N_t) = \lambda t$ since $N_t \sim \mathrm{Poi}(\lambda t)$. Using the linearity of the expectation and the fact that deterministic constants do not impact the variance we get $\mathrm{E}(Y_t) = t + \lambda t$ and $\mathrm{Var}(N_t) = \lambda t$. — meth seen ⇓ — 3, A

(ii) Let $t \geq 0$. We use the law of total expectation to deduce that — unseen ⇓

$$\begin{aligned}
\mathrm{E}(X_t) &= \mathrm{E}(B_{t+N_t}) = \mathrm{E}(\mathrm{E}(B_{t+N_t}|N_t)) \\
&= \sum_{n=0}^{\infty} \mathrm{E}(B_{t+N_t}|N_t = n)\mathrm{P}(N_t = n) = \sum_{n=0}^{\infty} \mathrm{E}(B_{t+n}|N_t = n)\mathrm{P}(N_t = n) \\
&\overset{\text{indep. of } B \text{ and } N}{=} \sum_{n=0}^{\infty} \mathrm{E}(B_{t+n})\mathrm{P}(N_t = n) = \sum_{n=0}^{\infty} 0\mathrm{P}(N_t = n) = 0,
\end{aligned}$$

since the Brownian motion has zero mean. — 3, C

[1 mark for computation, 2 marks for justifications]

Similarly, we have

$$\begin{aligned}
\mathrm{E}(X_t^2) &= \mathrm{E}(B_{t+N_t}^2) = \mathrm{E}(\mathrm{E}(B_{t+N_t}^2|N_t)) \\
&= \sum_{n=0}^{\infty} \mathrm{E}(B_{t+N_t}^2|N_t = n)\mathrm{P}(N_t = n) = \sum_{n=0}^{\infty} \mathrm{E}(B_{t+n}^2|N_t = n)\mathrm{P}(N_t = n) \\
&\overset{\text{indep. of } B \text{ and } N}{=} \sum_{n=0}^{\infty} \mathrm{E}(B_{t+n}^2)\mathrm{P}(N_t = n) \\
&= \sum_{n=0}^{\infty} (t+n)\mathrm{P}(N_t = n) = t + \mathrm{E}(N_t) = t + \lambda t,
\end{aligned}$$

since $E(B_x^2) = x$ for all $x \geq 0$. Hence, $\mathrm{Var}(X_t) = t + \lambda t$. — 2, C

[1 mark for computation, 1 mark for justifications]

(iii) Without loss of generality, we assume that $0 \leq s < t$. Recall that $\mathrm{Cov}(X_s, X_t) = \mathrm{E}(X_s X_t) - \mathrm{E}(X_s)\mathrm{E}(X_t) = \mathrm{E}(X_s X_t)$ since $\mathrm{E}(X_s) = \mathrm{E}(X_t) = 0$. Then $\boxed{\text{unseen} \Downarrow}$

$$\begin{aligned}
\mathrm{E}(X_s X_t) &= \mathrm{E}(\mathrm{E}(X_s X_t | N_s, N_t)) \\
&= \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} \mathrm{E}(B_{s+N_s} B_{t+N_t} | N_s = n, N_t = m)\mathrm{P}(N_s = n, N_t = m) \\
&= \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} \mathrm{E}(B_{s+n} B_{t+m} | N_s = n, N_t = m)\mathrm{P}(N_s = n, N_t = m) \\
&\overset{\text{indep. of } B \text{ and } N}{=} \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} \mathrm{E}(B_{s+n} B_{t+m})\mathrm{P}(N_s = n, N_t = m) \\
&= \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} (s+n)\mathrm{P}(N_s = n, N_t = m),
\end{aligned}$$

where we used the result from lectures that $\mathrm{E}(B_x B_y) = \min\{x, y\}$. Here we have that $s < t$, which also implies that $N_s < N_t$, hence, $s + n < s + m$. Applying the law of total probability, leads to

$$\begin{aligned}
\mathrm{E}(X_s X_t) &= \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} (s+n)\mathrm{P}(N_s = n, N_t = m) \\
&= \sum_{n=0}^{\infty} (s+n)\mathrm{P}(N_s = n) = s + \lambda s.
\end{aligned}$$

Hence, for $0 \leq s < t$, $\mathrm{Cov}(X_s, X_t) = s + \lambda s$. We can conclude that, for $s, t \geq 0$, $\mathrm{Cov}(X_s, X_t) = (1 + \lambda)\min\{s, t\}$. $\boxed{5, \mathrm{D}}$

[**3 marks for the calculations, 2 marks for the justifications**]
[Students may apply the law of total covariation for an alternative (quicker) solution, but need to give proper justifications as well.]

5. (a) We refer to the continuous-time stochastic Markov chain model for the spread of a rumour in a population of size $N+1$ (for $N \in \mathbb{N}$) described in this paper as the **DK-model**.

The model considers a closed homogeneously mixing population of $N+1$ individuals, which are divided into three (exclusive and exhaustive) groups: $X_t \geq 0$, $Y_t \geq 0$ and $Z_t \geq 0$ denote the number of *ignorants* (people who are not aware of the rumour), *spreaders* (people who are spreading the rumour) and *stiflers* (people who know the rumour, but have ceased communicating it after meeting somebody who has already heard the rumour) at time $t \geq 0$, respectively. A person aware of the rumour will spread the rumour until he meets another person knowing the rumour. It is assumed that $X_t + Y_t + Z_t = N+1$ for all $t \geq 0$ and the initial condition is given by $Y_0 = 1, X_0 = N, Z_0 = 0$. $\boxed{2, \mathrm{M}}$

It is assumed that the infinitesimal transition probabilities of the MC are proportional to the frequencies of meetings between members of the various groups and are given as follows: Let $t \geq 0, \delta > 0$ and $x, y, z \in \{0, 1, \ldots\}$ such that $x + y + z = N+1$, then

$$\mathrm{P}(X_{t+\delta} = x-1, Y_{t+\delta} = y+1, Z_{t+\delta} = z | X_t = x, Y_t = y, Z_t = z) = xy\delta + o(\delta), \tag{1}$$

$$\mathrm{P}(X_{t+\delta} = x, Y_{t+\delta} = y-2, Z_{t+\delta} = z+2 | X_t = x, Y_t = y, Z_t = z) = \frac{1}{2}y(y-1)\delta + o(\delta), \tag{2}$$

$$\mathrm{P}(X_{t+\delta} = x, Y_{t+\delta} = y-1, Z_{t+\delta} = z+1 | X_t = x, Y_t = y, Z_t = z) = yz\delta + o(\delta)[= y(N+1-x-y)\delta + o(\delta)], \tag{3}$$

$$\mathrm{P}(X_{t+\delta} = x, Y_{t+\delta} = y, Z_{t+\delta} = z | X_t = x, Y_t = y, Z_t = z) = 1 - (xy\delta + \frac{1}{2}y(y-1) - yz)\delta + o(\delta), \tag{4}$$

and all other infinitesimal transition probabilities are given by $o(\delta)$. $\boxed{3, \mathrm{M}}$

We note that the infinitesimal transition probabilities, assuming that they are proportional to the frequencies of meetings between members of the various groups, can be derived/justified as follows:

* In the case of (1), there are x ignorants, y spreaders and z stiflers, for the number of ignorants to decrease by one and the number of spreaders to increase by one, we need one ignorant and one spreader to meet, for which there are $\binom{x}{1}\binom{y}{1} = xy$ possibilities.
* Similarly, in the case of (2), there are x ignorants, y spreaders and z stiflers, for the number of spreaders to decrease by two and the number of stiflers to increase by two, we need that two spreaders meet, for which there are $\binom{y}{2} = y(y-1)/2$ possibilities.
* In the case of (3), there are x ignorants, y spreaders and z stiflers, for the number of spreaders to decrease by one and the number of stiflers to increase by one, we need one spreader and one stifler to meet, for which there are $\binom{y}{1}\binom{z}{1} = zy$ possibilities.
* In the case of (4), we balance the previous probabilities,
* and set all remaining infinitesimal transition probabilities to $o(\delta)$. $\boxed{3, \mathrm{M}}$

(b) The random walk model associated with the DK model is the discrete-time jump chain on the integer lattice in the plane associated with the continuous-time Markov chain.

The (non-trivial) one-step transition probabilities are given as follows:

$$p_{(x,y),(x-1,y+1)} = \mathrm{P}(X_n = x-1, Y_n = y+1 | X_{n-1} = x, Y_{n-1} = y) = \frac{x}{N - 0.5(y-1)},$$
$$g_{(x,y),(x,y-2)} = \mathrm{P}(X_n = x, Y_n = y-2 | X_{n-1} = x, Y_{n-1} = y) = \frac{0.5(y-1)}{N - 0.5(y-1)},$$
$$g_{(x,y),(x,y-1)} = \mathrm{P}(X_n = x, Y_n = y-1 | X_{n-1} = x, Y_{n-1} = y) = \frac{N+1-x-y}{N - 0.5(y-1)},$$

where it is assumed that $y > 0, x \geq 0, x + y \leq N + 1$. $\boxed{2, \mathrm{M}}$

Justification: From lectures, we know that the transition probabilities of the jump chain can be obtained by setting $p_{ij} = -g_{ij}/g_{ii}$, where g_{ij} are the corresponding elements of the generator for states $i, j \in E$.

First we compute the sum of the rates given in equations (1), (2), (3): $a := xy + 0.5y(y-1) + yz = xy + 0.5y(y-1) + y(N+1-x-y) = y(N - 0.5(y-1))$ and we have the correspondence that $g_{(x,y),(x,y)} = -a$ by a result from lectures.

Moreover, $g_{(x,y),(x-1,y+1)} = xy, g_{(x,y),(x,y-2)} = 0.5y(y-1), g_{(x,y),(x,y-1)} = y(N+1-x-y)$.

Hence we get that $p_{(x,y),(x-1,y+1)} = g_{(x,y),(x-1,y+1)}/a, p_{(x,y),(x,y-2)} = g_{(x,y),(x,y-2)}/a, p_{(x,y),(x,y-1)} = g_{(x,y),(x,y-1)}/a$, which leads to the one-step transition probabilities stated above. $\boxed{4, \mathrm{M}}$

(c) Using the notation from the question, from (b) and the law of total probability (by conditioning on the previous step), we note that

$$\begin{aligned} P_{x,y} &= P_{x+1,y-1}p_{(x+1,y-1),(x,y)} + P_{x,y+2}p_{(x,y+2),(x,y)} + P_{x,y+1}p_{(x,y+1),(x,y)} \quad (5) \\ &= P_{x+1,y-1}\frac{x+1}{N-0.5(y-2)} + P_{x,y+2}\frac{0.5(y+1)}{N-0.5(y+1)} + P_{x,y+1}\frac{N-x-y}{N-0.5y}, \quad (6) \end{aligned}$$

where we used the one-step transition probabilities mentioned in (b). Similarly, we have $\boxed{3, \mathrm{M}}$

$$P_{x,1} = P_{x,3}p_{(x,3),(x,1)} + P_{x,2}p_{(x,2),(x,1)} = P_{x,3}\frac{1}{N-1} + P_{x,2}\frac{N-x-1}{N-0.5},$$
$$\begin{aligned} P_{x,0} &= P_{x,2}p_{(x,2),(x,0)} + P_{x,1}p_{(x,1),(x,0)} \\ &= P_{x,2}\frac{0.5}{N-0.5} + P_{x,1}\frac{N-x}{N}. \end{aligned}$$

We note that (5) is only to be used for $y \geq 2$ and we use the convention that $P_{xy} = 0$ when $x + y$ exceeds $N + 1$, then the forward equations can be used for all $0 \leq x \leq N-1$, $0 \leq y \leq N+1-x$. $\boxed{2, \mathrm{M}}$ $\boxed{1, \mathrm{M}}$

Review of mark distribution:
Total A marks: 32 of 32 marks
Total B marks: 20 of 20 marks
Total C marks: 12 of 12 marks
Total D marks: 16 of 16 marks
Total marks: 80 of 80 marks
Total Mastery marks: 20 of 20 marks

Please record below, some brief but non-trivial comments for students about how well (or otherwise) the questions were answered. For example, you may wish to comment on common errors and misconceptions, or areas where students have done well. These comments should note any errors in and corrections to the paper. These comments will be made available to students via the MathsCentral Blackboard site and should not contain any information which identifies individual candidates. Any comments which should be kept confidential should be included as confidential comments for the Exam Board and Externals. If you would like to add formulas, please include a separate pdf file with your email.

ExamModuleCode	**QuestionNumber**	**Comments for Students**
Applied Probability_MATH60045 MATH97083 MATH70045	1	Many students did very well on this question, with nearly everyone getting high marks on part (a). In part (b)(i) there was a tendency to ignore the fact that a is in (0,1) as the chain is not irreducible otherwise. Parts (b)(ii) and (b)(iii) were done well on the whole.
Applied Probability_MATH60045 MATH97083 MATH70045	2	Question 2 was much more of a challenge to students, than question 1. In (a)(i). a mark was lost if you did not recognise the conditions under which there is a unique stationary distribution. Some people found the derivation of the required necessary condition for time reversibility (ade=bcf) very easily while other struggled over several pages. In (a)(ii) some did not appreciate the meaning of "interpretation" or gave examples in which the transition matrix has one or more zero entries. Part (b) required you show that Y was not a Markov chain. The best answers used two different sample paths to show that the Markov property did not hold. Part (c)(i) proved a challenge for most people and a variety of approaches were attempted. Vague answers earned reduced marks. Some scripts were barely legible because of terrible handwriting or poor scans. Part (c)(ii) was done well although some people ignored the hint to give a practical example.
Applied Probability_MATH60045 MATH97083 MATH70045	3	Students found Q2(a)(iii) very challenging, but the uestion can be solved fairly simply with some careful reasoning. In part (b)(v), not many students provided two good model criticisms; even though it was descriptive and so perhaps less appealing, this question should have been treated with more care.
Applied Probability_MATH60045 MATH97083 MATH70045	4	On the whole Q4 was done well. Many students lost a mark by adding jump probabilities rather than transition intensities to the diagram in (a)(i); this left them illogically answering part (ii) before part (i). In part (b), the most common mistake for (ii) and (iii) was to state answers in terms of Y_t or N_t.
Applied Probability_MATH60045 MATH97083 MATH70045	5	Part c was done surprisingly well, with many students writing down precise answers with no workings; others showed nice working. Parts a and b, which were more descriptive, scored less well; a common mistake was adopting the crude notation used in the paper, rather than translating to the necessary time-indexed notation used in the model solutions.