MATH70045

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2024

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

Applied Probability

Date: Wednesday, May 1, 2024

Time: 10:00 – 12:30 (BST)

Time Allowed: 2.5 hours

This paper has 5 Questions.

Please Answer Each Question in a Separate Answer Booklet

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

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Allow margins for marking.

Credit will be given for all questions attempted.

Each question carries equal weight.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO

1. Let X be a time homogeneous discrete time Markov chain taking values on a countable space E with transition probability matrix $\mathbf{P} = [p_{ij}]_{i \in E, j \in E}$. Define the probability that the chain reaches state i in n steps without intermediate return to its starting point j as follows:

$$l_{ji}(n) = \mathrm{P}(X_n = i, T_j \geq n | X_0 = j), \qquad i \neq j.$$

Let also the first passage probability from i to j be denoted as

$$f_{ij}(n) = \mathrm{P}(T_j = n | X_0 = i) = \mathrm{P}(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j | X_0 = i).$$

(a) What is T_j in the definitions above?

(1 mark)

(b) Let $i \neq j$. State the value of $l_{ji}(1)$ and show that for all integers $n \geq 2$,

$$l_{ji}(n) = \sum_{r \in E : r \neq j} p_{ri} l_{jr}(n-1).$$

(5 marks)

(c) Fix a state $j \in E$ and define for any $i \in E$ by $N_i(j)$ the number of visits to the state i before visiting state j (when counting from time $n = 1$ onwards), i.e.

$$N_i(j) = \sum_{n=1}^{T_j} \mathbb{I}_{\{X_n = i\}}.$$

Let also $\rho_i(j)$ to be defined as

$$\rho_i(j) = \mathrm{E}[N_i(j) | X_0 = j].$$

What is the value of $\rho_j(j)$? Show that for $i \neq j$ we have $\rho_i(j) = \sum_{n=1}^{\infty} l_{ji}(n)$.

(5 marks)

(d) Denote $\boldsymbol{\rho}(j)$ for the the row vector consisting of elements $\rho_i(j)$ for $i \in E$. Show that $\boldsymbol{\rho}(j) = \boldsymbol{\rho}(j)\mathbf{P}$.

(5 marks)

(e) We wish to use these results to construct a stationary distribution $\boldsymbol{\pi}$. Explain how this can be made possible and in particular explain why irreducibility and positive recurrent are necessary. For this part you are **not** required to present detailed derivations.

(4 marks)

(Total: 20 marks)

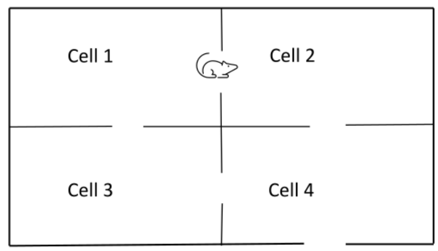

Figure 1: A mouse runs through a maze.

2. (a) A mouse runs through the maze in Figure 1. The mouse starts in either cell $\{1, 2, 3, 4\}$ and at each time moves to a neighbouring cell. As shown in the figure it can move through open doors to an adjacent cell either horizontally or vertically. We will assume that the next cell is chosen independently each time and with equal probability from those possible. The mouse will keep moving between cells until at some it escapes from Cell 4 to the outside world, which we will we will refer to as cell 0. Once the mouse has escaped it never returns to the maze.

   (i) State a sensible choice of states to model the motion of the mouse as a Markov chain. Then determine its transition matrix, draw the transition diagram and find the communicating classes associated with this Markov chain. (3 marks)

   (ii) Find all possible stationary distributions if any. (3 marks)

   (iii) For each cell $1, 2, 3, 4$ being the mouse's initial position, find the expected number of steps required to escape.
   Hint: You may use directly without proof an expression linking $\mathrm{E}[$*time of escape*$|X_0 = i]$ *and* $\mathrm{E}[$*time of escape*$|X_1 = i]$. (6 marks)

   (b) Consider a Markov chain defined on $E = \{1, 2, 3, 4\}$ with the following transition matrix:

$$\mathbf{P} = \begin{pmatrix} 0 & p & 0 & 1-p \\ 1-p & 0 & p & 0 \\ 0 & 1-p & 0 & p \\ p & 0 & 1-p & 0 \end{pmatrix}.$$

   (i) Find all possible stationary distributions if any.

   (ii) Determine the period of the chain, and whether irreducibility, positive recurence and time-reversibility holds.

   (iii) How can this model be constructed starting from the model in part (a) above? Explain how changing the model from part (a) to part (b) affects the stationarity properties of the chain and the motion of the mouse.

(8 marks)

(Total: 20 marks)

3. Consider a birth process $(N_t)_{t\geq 0}$ with birth rates $\lambda_0, \lambda_1, \ldots$ such that $\lambda_i \neq \lambda_j$ for $i \neq j$, and $N_0 = 0$. Let $p_n(t) = \mathrm{P}(N_t = n)$ for $n \in \mathbb{N} \cup \{0\}$.

(a) Derive an equation for $p_0'(t)$ in terms of $p_0(t)$. (p' denotes time derivative). (2 marks)

(b) Let $n \in \mathbb{N}$. Derive an equation for $p_n'(t)$ in terms of $p_n(t)$ and $p_{n-1}(t)$. (3 marks)

(c) Show that

$$p_0(t) = e^{-\lambda_0 t}$$

and

$$p_1(t) = \lambda_0 \left(\frac{1}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} + \frac{1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t} \right)$$

are solutions to the differential equations derived in (a) and (b). (5 marks)

(d) Denote as T_n the time of the n-th birth. Show that

$$\mathrm{P}(T_1 > t, T_2 > t + s) = p_0(t)[p_0(s) + p_1(s)] \quad \text{for } s, t > 0.$$

(5 marks)

(e) Using the results in parts (c), (d) or otherwise, derive the joint density of (T_1, T_2). (5 marks)

(Total: 20 marks)

4. (a) Let W_t be a standard Brownian motion. Consider the following stochastic processes:

$$X_t = -W_t, \quad Y_t = tW_1, \quad Z_t = W_{2t} - W_t.$$

Determine the distribution of X_t, Y_t and Z_t. Are X_t, Y_t and Z_t Brownian motions? For each of these three processes, please say "yes" or "no" and justify your answer formally.

(4 marks)

(b) Consider a general homogeneous continuous time Markov process $X = (X_t; t \geq 0)$ defined on a countable space E and let $p_{ij}(t)$ denote the transition probability from state i to state j after time t. Suppose one samples this chain X at the jump times of an independent Poisson process N with rate λ. That is, we define $Y_n = X_{J_n}$, where $(J_n)_{n\geq 0}$ denotes the sequence of jump times of N.

(i) Explain why Y_n is a Markov chain. (2 marks)

(ii) Derive an expression for the transition probabilites of the chain Y_n that uses $p_{ij}(t)$ and λ.

(5 marks)

(iii) Suppose X has a stationary distribution π. Show that this will also be a stationary distribution for Y. (3 marks)

(iv) Suppose we relax the independence assumption of the Markov process X and Poisson process N and assume instead $(N_s; s \geq t)$ is independent of $(X_s; s \leq t)$. Discuss what this means for X and N assuming they are still a Markov process with invariant distribution π and a Poisson process with rate λ respectively.
For each of the answers in (i)-(iii) above explain how would you need to modify your answers if this new assumption is used.

(6 marks)

(Total: 20 marks)

5. (a) Let μ be a probability measure on a countable space E. State a definition of the total variation norm of μ, $\|\mu\|_{TV}$. Is the the total variation norm a valid metric/distance function for probability distributions? (2 marks)

(b) Suppose X is a E-valued irreducible and aperiodic Markov chain with transition probability matrix P and initial distribution ν. Assume that X has a stationary distribution π and in addition for any $x, y \in E$ we have that $P(x,y) \geq (1-\alpha)\pi(y)$ with $\alpha \in (0,1)$.

(i) Suppose one defines a different Markov chain with transition matrix $Q = [Q(x,y)]_{x,y \in E}$. Each time this chain evolves as follows: with probability $\frac{1}{\alpha}$ the state moves from x to y with probability given by P. With probability $1 - \frac{1}{\alpha}$, we obtain an independent sample from π and set its value to y. Write down an expression for $Q(x,y)$ (the probability of moving from x to y for the new chain). Show that matrix Q is a valid stochastic matrix.
(3 marks)

(ii) Show that

$$Q^k(x,y) = \frac{P^k(x,y) - (1-\alpha^k)\pi(y)}{\alpha^k}.$$

(4 marks)

(iii) Show that for any $x \in E$

$$\|P^t(x,\cdot) - \pi(\cdot)\|_{TV} \leq \alpha^k.$$

(4 marks)

(iv) Explain why above convergence result is important (or useful) and comment on the role of α. (2 marks)

(c) Prove the convergence result in part (b)(iii) using an appropriate Markovian coupling. You are permitted to show the result with the right hand side using a different value for α.
(5 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2024

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

MATH60045/MATH70045

Applied Probability (Solutions)

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

1. (a) T_j is first hitting time of state j, i.e. $T_j = \min_{k\in\mathbb{N}}\{X_k = j\}$ $\boxed{\text{seen} \Downarrow}$

$\boxed{1,\ \text{A}}$

(b) Clearly, $l_{ji}(1) = p_{ji}$. $\boxed{\text{seen} \Downarrow}$

For $n \geq 2$, by the law of total probability and the Markov property, we get $\boxed{1,\ \text{A}}$

$$
\begin{aligned}
l_{ji}(n) &= \sum_{r\in E: r\neq j} \mathrm{P}(X_n = i, X_{n-1} = r, T_j \geq n | X_0 = j) \\
&= \sum_{r\in E: r\neq j} \mathrm{P}(X_n = i | X_{n-1} = r, T_j \geq n, X_0 = j)\mathrm{P}(X_{n-1} = r, T_j \geq n | X_0 = j) \\
&= \sum_{r\in E: r\neq j} \mathrm{P}(X_n = i | X_{n-1} = r)\mathrm{P}(X_{n-1} = r, T_j \geq n-1 | X_0 = j) \\
&= \sum_{r\in E: r\neq j} p_{ri} l_{jr}(n-1).
\end{aligned}
$$

$\boxed{4,\ \text{B}}$

$\boxed{\text{seen} \Downarrow}$

(c) Clearly, $N_j(j) = 1$ and hence $\rho_j(j) = 1$. $\boxed{1,\ \text{A}}$

Applying Tonelli's theorem and using the fact that the conditional expectation of an indicator variable is equal to the conditional probability that the indicator variable is equal to one, we get

$$
\begin{aligned}
\rho_i(j) &= \mathrm{E}[N_i(j)|X_0 = j] = \mathrm{E}\left[\sum_{n=1}^{\infty} \mathbb{I}_{\{X_n = i\}\cap\{T_j \geq n\}} \middle| X_0 = j\right] \\
&\overset{\text{Tonelli}}{=} \sum_{n=1}^{\infty} \mathrm{E}\left[\mathbb{I}_{\{X_n = i\}\cap\{T_j \geq n\}} \middle| X_0 = j\right] = \sum_{n=1}^{\infty} \mathrm{P}(X_n = i, T_j \geq n | X_0 = j) \\
&= \sum_{n=1}^{\infty} l_{ji}(n).
\end{aligned}
$$

$\boxed{4,\ \text{C}}$

(d) Showing $\boldsymbol{\rho}(j) = \boldsymbol{\rho}(j)\mathbf{P}$ is equivalent to showing that for all $i \in E$ we have $\rho_i(j) = \sum_{r\in E} \rho_r(j) p_{ri}$. Before we showed $\boxed{\text{seen} \Downarrow}$

$$
\rho_i(j) = \sum_{n=1}^{\infty} l_{ji}(n).
$$

and, by the previous question, we have $l_{ji}(1) = p_{ji}$, and for $n \geq 2$, we get

$$
l_{ji}(n) = \sum_{r\in E: r\neq j} p_{ri} l_{jr}(n-1).
$$

$\boxed{2,\ \text{A}}$

Hence, for all $i \in E$, we have

$$\begin{aligned}
\rho_i(j) &= \sum_{n=1}^{\infty} l_{ji}(n) = l_{ji}(1) + \sum_{n=2}^{\infty} l_{ji}(n) \\
&= p_{ji} + \sum_{n=2}^{\infty} \sum_{r \in E: r \neq j} p_{ri} l_{jr}(n-1) \\
&\overset{\rho_j(j)=1,\text{Tonelli}}{=} \rho_j(j) p_{ji} + \sum_{r \in E: r \neq j} p_{ri} \sum_{n=2}^{\infty} l_{jr}(n-1) \\
&= \rho_j(j) p_{ji} + \sum_{r \in E: r \neq j} p_{ri} \sum_{n=1}^{\infty} l_{jr}(n) \\
&= \rho_j(j) p_{ji} + \sum_{r \in E: r \neq j} p_{ri} \rho_r(j) = \sum_{r \in E} \rho_r(j) p_{ri}.
\end{aligned}$$

3, C

(e) To construct an invariant distribution one needs to use an expression like

sim. seen ⇓

$$\pi_i = \frac{\rho_i(j)}{\mu_j}, \quad \mu_j = \sum_{i \in E} \rho_i(j)$$

and ensure that μ_j is finite.

Irreducibility is used to establish $\rho_i(j) < \infty$ and positive recurence that $\mu_j < \infty$ by noting that $\mu_j = \mathrm{E}(T_j | X_0 = j)$.

(Although full derivation not required the latter can be seen by recalling

$$\begin{aligned}
\sum_{i \in E} N_i(j) &= \sum_{i \in E} \sum_{n=1}^{\infty} \mathbb{I}_{\{X_n = i\}} \mathbb{I}_{\{T_j \geq n\}} \\
&\overset{\text{Tonelli}}{=} \sum_{n=1}^{\infty} \mathbb{I}_{\{T_j \geq n\}} \sum_{i \in E} \mathbb{I}_{\{X_n = i\}} = \sum_{n=1}^{\infty} \mathbb{I}_{\{T_j \geq n\}} = \sum_{n=1}^{T_j} 1 = T_j.)
\end{aligned}$$

4, B

2. (a) (i) $\boxed{\text{meth seen} \Downarrow}$

Clearly each cell is a state so state space $E = \{0,1,2,3,4\}$ and the transition matrix is

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

and the transition diagram is found in Figure 1 below (end of document).

The communicating classes are $\{0\}$ (closed) and $\{1,2,3,4\}$ (open). $\boxed{3, \text{A}}$

(ii) $\boxed{\text{meth seen} \Downarrow}$

It is clear that 0 is an absorbing state and consists of the only closed communication class and hence one can check that $(1,0,0,0,0)$ is the stationary distribution. $\boxed{3, \text{A}}$

(iii) $\boxed{\text{sim. seen} \Downarrow}$

Let $T = \min\{n \geq 0 : X_n = 0\}$. Define $m_i = \mathrm{E}[T|X_0 = i]$. We note $m_0 = 0$ abd then start from state 4 and proceed recursively applying law of total conditional expectation and the Markov property:

$$\begin{aligned} m_4 =& \mathrm{E}[T|X_0 = 4] \\ =& \sum_{i \in E} \mathrm{E}[T|X_0 = 4, X_1 = i]\mathrm{P}(X_1 = i|X_0 = 4) \\ =& \sum_{i \in E} \mathrm{E}[T|X_1 = i]\mathrm{P}(X_1 = i|X_0 = 4) \end{aligned}$$

.

We have $\mathrm{E}[T|X_1 = i] = \mathrm{E}[T|X_1 = 0] + 1$ (Proof can be seen in Example Sheet Question 2-14) but here can be quoted directly as per the hint provided.

So we have

$$\begin{aligned} m_4 =& \frac{1}{3}(m_3 + 1) + \frac{1}{3}(m_2 + 1) + \frac{1}{3}(m_0 + 1) \\ =& 1 + \frac{1}{3}m_3 + \frac{1}{3}m_2 \end{aligned}$$

We proceed similarly for the rest of the states and get also

$$\begin{aligned} m_3 =& 1 + \frac{1}{2}m_1 + \frac{1}{2}m_4 \\ m_2 =& 1 + \frac{1}{2}m_1 + \frac{1}{2}m_4 \\ m_1 =& 1 + \frac{1}{2}m_3 + \frac{1}{2}m_2, \end{aligned}$$

and then a quick calculation gives $m_1 = 13, m_2 = 12, m_3 = 12, m_4 = 9$. $\boxed{6, \text{B}}$

(b) (i) $\boxed{\text{meth seen} \Downarrow}$

There is a unique closed communicating class (the whole set E), hence there is a unique invariant distribution. Let $\boldsymbol{\rho} = (\pi_1, \dots, \pi_4)$. Given the symmetry in the transitions, one can either guess and check that the uniform distribution does indeed solve $\boldsymbol{\pi} P = \boldsymbol{\pi}$, or solve for the the system $\boldsymbol{\pi}(P - I) = 0$. Either way $\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4}$. 2, A

(ii) meth seen ⇓

Period is clearly 2 for every state, given all states are accessible the chain is irreducible, positive recurrence holds as there is a unique stationary distribution (or given this is a finite chain with one closed communicating class) and for reversibility we have that it does not hold as detailed balance $\pi_i p_{ij} = \pi_j p_{ji}$ with $\pi_i = \pi_j$ would imply that P is symmetric, which is not the case. 3, A

(iii) unseen ⇓

Starting from the mouse and maze model, if one removes state 0 (e.g. by closing the door to the outside world) one can notice that both models have a transition diagram that is a square graph. A closer inspection of the connectivity of each model reveals that after removing 0 in model in (a), then a relabelling of cell 3 to cell 4 and vice verse will be an instance of the model in (b) with $p = \frac{1}{2}$. In the model in (a) states $\{1, 2, 3, 4\}$ had equal probability, but due to the absorbing state 0 this was zero and the mouse would almost surely escape. By removing the absorbing state, the mouse can never escape and we end in the stationary distribution being uniform as in the model in (b). The mouse will endlessly move around every state in the maze and at stationarity all states will have the same probability. The answer here uses $p = \frac{1}{2}$, but of course more general answers with general p are possible and clearly acceptable. 3, B

3. (a) Let $\delta > 0$, then we have

meth seen ⇓

$$p_0(t+\delta) = \mathrm{P}(N_{t+\delta} - N_t = 0, N_t = 0) = \mathrm{P}(N_{t+\delta} = 0 | N_t = 0)\mathrm{P}(N_t = 0) = (1 - \lambda_0 \delta) p_0(t) + o(\delta).$$

Then we get

$$\lim_{\delta \downarrow 0} \frac{p_0(t+\delta) - p_0(t)}{\delta} = p_0'(t) = -\lambda_0 p_0(t).$$

2, A

(b) Let $\delta > 0$ and $n \in \mathbb{N}$. Using law of total probability we have

meth seen ⇓

$$\begin{aligned} p_n(t+\delta) &= \sum_{i=0}^{n} \mathrm{P}(N_{t+\delta} = n | N_t = i)\mathrm{P}(N_t = i) \\ &= \lambda_{n-1}\delta p_{n-1}(t) + (1 - \lambda_n \delta) p_n(t) + o(\delta) \end{aligned}$$

where we used the single arrival property of a birth process to consider only state transitions from n to n and $n-1$ to n. Then

$$\lim_{\delta \downarrow 0} \frac{p_n(t+\delta) - p_n(t)}{\delta} = p_n'(t) = \lambda_{n-1} p_{n-1}(t) - \lambda_n p_n(t).$$

3, A

(c) Differentiating both given expressions gives:

meth seen ⇓

$$\begin{aligned} p_0'(t) &= -\lambda_0 e^{-\lambda_0 t} = -\lambda_0 p_0(t) \\ p_1'(t) &= \lambda_0 \left(\frac{-\lambda_0}{\lambda_1 - \lambda_0} e^{-\lambda_0} + \frac{-\lambda_1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t} \right). \quad (*) \end{aligned}$$

For $p_0(t)$ we have verified the claim. For $p_1(t)$ we start from the result in part (b) for $n = 1$ and substitute the given expresssions

$$\begin{aligned} \lambda_0 p_0(t) - \lambda_1 p_1(t) &= \lambda_0 e^{-\lambda_0 t} - \lambda_1 \lambda_0 \left(\frac{1}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} + \frac{1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t} \right) \\ &= \frac{\lambda_0(\lambda_1 - \lambda_0) - \lambda_0 \lambda_1}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} - \frac{\lambda_0 \lambda_1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t} \\ &= \frac{-\lambda_0^2}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} + \frac{-\lambda_0 \lambda_1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t} \\ &= p_1'(t), \end{aligned}$$

as derived above in $(*)$. Note that the denominator in $\frac{-\lambda_0 \lambda_1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t}$ is not used at all in this derivation and could be otherwise an arbitrary factor, so an additional verification for the solution is checking that at we have that $p_1(0) = 0$.

5, A

sim. seen ⇓

(d) We have that $\{T_1 > t\}$ is equivalent to $\{N_t = 0\}$. Similarly $\{T_2 > t + s\}$ is equivalent to $\{N_{t+s} = 0\} \cup \{N_{t+s} = 1\}$. Hence for $s, t > 0$ using law of total probability, independence of increments and and homogeneity we get:

$$\begin{aligned} \mathrm{P}(T_1 > T, T_2 > t + s) &= \mathrm{P}(N_t = 0, N_{t+s} = 0) + \mathrm{P}(N_t = 0, N_{t+s} = 1) \\ &= \mathrm{P}(N_t = 0, N_{t+s} - N_t = 0) + \mathrm{P}(N_t = 0, N_{t+s} - N_t = 1) \\ &= \mathrm{P}(N_t = 0)\mathrm{P}(N_s = 0) + \mathrm{P}(N_t = 0)\mathrm{P}(N_s = 1) \\ &= p_0(t) p_0(s) + p_0(t) p_1(s), \end{aligned}$$

which is the required result.

5, D

$\boxed{\text{unseen} \Downarrow}$

(e) Let f_{T_1,T_2} denote the bivariate density. Using the previous parts we have that for $t_2 > t_1 > 0$

$$\begin{aligned}
\mathrm{P}(T_1 > t_1, T_2 > t_2) &= p_0(t_1)\left(p_0(t_2 - t_1) + p_1(t_2 - t_1)\right) \\
&= e^{-\lambda_0 t_1}\left(e^{-\lambda(t_2-t_1)} + \lambda_0\left(\frac{1}{\lambda_1 - \lambda_0}e^{-\lambda_0(t_2-t_1)} + \frac{1}{\lambda_0 - \lambda_1}e^{-\lambda_1(t_2-t_1)}\right)\right) \\
&= e^{-\lambda_0 t_2} + \frac{\lambda_0}{\lambda_1 - \lambda_0}e^{-\lambda_0 t_2} + \frac{\lambda_0}{\lambda_0 - \lambda_1}e^{-(\lambda_0-\lambda_1)t_1}e^{-\lambda_1 t_2} \\
&= \frac{\lambda_1}{\lambda_1 - \lambda_0}e^{-\lambda_0 t_2} + \frac{\lambda_0}{\lambda_0 - \lambda_1}e^{-(\lambda_0-\lambda_1)t_1}e^{-\lambda_1 t_2}.
\end{aligned}$$

Note that

$$\mathrm{P}(T_1 > t_1, T_2 > t_2) = \int_{t_1}^{\infty}\int_{t_1}^{\infty} f_{T_1,T_2}(t_1,t_2)dt_1dt_2,$$

so using the fundamental theorem of calculus

$$\begin{aligned}
f_{T_1,T_2} &= (-1)^2\frac{\partial^2}{\partial t_1 \partial t_2}\mathrm{P}(T_1 > t_1, T_2 > t_2) \\
&= \frac{\lambda_0}{\lambda_0 - \lambda_1}e^{-(\lambda_0-\lambda_1)t_1}e^{-\lambda_1 t_2}(-\lambda_1)(-(\lambda_0 - \lambda_1))) \\
&= \lambda_0\lambda_1 e^{-(\lambda_0-\lambda_1)t_1}e^{-\lambda_1 t_2}
\end{aligned}$$

for $t_2 > t_1 > 0$ and 0 otherwise. $\boxed{5, \mathrm{D}}$

4. (a) All X_t, Y_t and Z_t are distributed according to $N(0,t)$ meth seen ⇓

The above observation ticks the correct distribution for the increment X_t is a Brownian motion as $X_0 = 0$ and shares the same properties of increments as W_t and continuity of sample paths.

Y_t is not a Brownian motion, despite satisfying $Y_0 = 0$ and continuity of sample paths. It is straightforward to consider increments $Y_{t_n} - Y_{t_{n-1}} = (\sqrt{t_n} - \sqrt{t_{n-1}})W_1$ and $Y_{t_n+1} - Y_{t_n} = (\sqrt{t_{n+1}} - \sqrt{t_n})W_1$ and notice that independence fails.

Similarly Z_t is not a Brownian motion despite satisfying $Z_0 = 0$ and continuity of sample paths. It is straightforward to consider increments $Z_{t_n} - Z_{t_{n-1}} = W_{2t_n} - W_{t_n} - W_{2t_{n-1}} + W_{t_{n-1}}$ and $Z_{t_n+1} - Z_{t_n} = W_{2t_{n+1}} - W_{t_{n+1}} - W_{2t_n} + W_{t_n}$ and notice that independence fails again. 4, A

meth seen ⇓

(b) (i) The (strong) Markov property definition for X implies this directly as

$$\mathrm{P}(X_{t_{n+1}} = i_{n+1}|X_{t_n} = i_n) = \mathrm{P}(X_{t_{n+1}} = i_{n+1}|X_{t_n} = i_n, \ldots, X_{t_0} = i_0).$$

for any time sequence satifying $t_0 < t_1 < \ldots < t_n < t_{n+1}$, so setting $J_n = t_n$ and the independence of (J_n) to X immediately gives

$$\mathrm{P}(Y_{n+1} = i_{n+1}|Y_n = i_n) = \mathrm{P}(Y_{n+1} = j|Y_n = i, Y_{n-1} = i_{n-1}, \ldots, Y_0 = i_0).$$

Note that indepence is used here to ensure that the random (t_n) sequence does not depend on the history/information of the chain. 2, A

sim. seen ⇓

(ii) Let q_{ij} denote the transition probabilities for Y. Based on definitions and (the conditional) law of total probability we have that

$$\begin{aligned} q_{ij} &= \mathrm{P}(Y_{n+1} = j|Y_n = i) \\ &= \int_0^\infty \mathrm{P}(X_{J_n+\tau} = j, H_{n+1} = \tau|X_{J_n} = i)d\tau \\ &= \int_0^\infty \mathrm{P}(X_\tau = j|X_0 = i)f_H(\tau)d\tau \\ &= \int_0^\infty p_{ij}(\tau)\lambda\exp(-\lambda\tau)d\tau, \end{aligned}$$

where H_n is the n-th interarrival time and we have used homogeneity of X and (H_n) being an i.i.d $Exp(\lambda)$ sequence. 5, C

meth seen ⇓

(iii) We have

$$\begin{aligned} \sum_{i\in E}\pi_i q_{ij} &= \int_0^\infty (\sum_{i\in E}\pi_i p_{ij}(\tau))\lambda\exp(-\lambda\tau)d\tau \\ &= \pi_j\int_0^\infty \lambda\exp(-\lambda\tau)d\tau \\ &= \pi_j. \end{aligned}$$

3, B

(iv) Given N is a standard Poisson process (so no change in the properties) and X a Markov process, this means that we allow the dynamics of $(N_s; s \geq t)$ to affect the dynamics of $(X_s; s \geq t)$, i.e. we allow the increment $N_{t_n} - N_{t_{n-1}}$ (or equivalently H_n) to influence the law of $X_{t_n} - X_{t_{n-1}}$, but the latter will be independent to $N_{t_{n+1}} - N_{t_n}$ and similarly for the rest of increments defined on any $t_0 < t_1 < \ldots < t_n < t_{n+1}$. unseen $\Downarrow$

For Part (i) let B_n denote any event defined on $(X_s; s \leq J_n)$ (i.e. $\sigma(X_s; s \leq J_n)$) or alternatively let $B_n = (X_{J_n} = i_n, X_{t_m} = k_m, \ldots, X_{t_1} = k_1, X_0 = i_0)$ for any $t_1 < \ldots < t_m < J_n$. Note that given $J_{n+1} = J_n + H_{n+1}$ and H_{n+1} is independent from B_n.
The Markov property of X implies that

$$\mathrm{P}(Y_{n+1} = i_{n+1} | Y_n = i_n, J_n = \kappa, B_n) = \mathrm{P}(Y_{n+1} = i_{n+1} | Y_n = i_n, J_n = \kappa).$$

and we can use the law of total probability in both sides to get

$$\begin{aligned}
\mathrm{P}(Y_{n+1} = i_{n+1} | Y_n = i_n, B_n) &= \int_0^\infty \mathrm{P}(Y_{n+1} = i_{n+1} | Y_n = i_n, J_n = \kappa, B_n) f_{J_n}(\kappa) d\kappa \\
&= \int_0^\infty \mathrm{P}(Y_{n+1} = i_{n+1} | Y_n = i_n, J_n = \kappa) f_{J_n}(\kappa) d\kappa \\
&= \mathrm{P}(Y_{n+1} = i_{n+1} | Y_n = i_n)
\end{aligned}$$

Setting $B_n = (Y_n = i_n, Y_{n-1} = i_{n-1}, \ldots, Y_0 = i_0)$ gives the required result for Part (i).
(Verbose answers that clarify the Markov property of X is unaffected and contain a similar argument to marginalisation w.r.t J_n and H_{n+1} mentioning independence of the two should be accepted. The answer is written more generally and one could have chosen B_n to be $(Y_n = i_n, Y_{n-1} = i_{n-1}, \ldots, Y_0 = i_0)$ from the start with some justification using the dependence in the sample path of X).

For part (ii) the arguments in the answer remain essentially unchanged as H_{n+1} is independent from B_n. The only issue is that in case homogeneity is lost then

$$\begin{aligned}
q_{ij}(n, n+1) &= \mathrm{P}(Y_{n+1} = j | Y_n = i) \\
&= \int_0^\infty \int_0^\infty \mathrm{P}(X_{J_n + \tau} = j, H_{n+1} = \tau, J_n = \kappa | X_{J_n} = i) d\kappa d\tau \\
&= \int_0^\infty \int_0^\infty \mathrm{P}(X_{\kappa + \tau} = j | X_\kappa = i) f_{J_n}(\kappa) f_H(\tau) d\kappa d\tau \\
&= \int_0^\infty \int_0^\infty p_{ij}(\kappa, \kappa + \tau) \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t} \lambda \exp(-\lambda \tau) d\tau.
\end{aligned}$$

(Last line not really necessary in the answer as long as basic argument is captured.)
For Part (iii) the argument/method remains unchanged as $p_{ij}(s,t)$ will be π invariant, so one would need to proceed similarly to the previous computation and integrate both $f_{J_{n(\kappa)}} f_H(\tau)$ to 1. 6, D

5. (a) $\|\mu\|_{TV} = \sup_{A\subseteq E} |\mu(A)|$ (other choices also possible here). Total variation is a valid metric as i) $\|\mu-\nu\|_{TV} = 0$ if $\mu = \nu$, ii) $\|\mu-\nu\|_{TV} > 0$ if $\mu \neq \nu$, iii) $\|\mu-\nu\|_{TV} = \|\nu-\mu\|_{TV}$ and iv) $\|\mu-\nu\|_{TV} \leq \|\mu-\pi\|_{TV} + \|\pi-\nu\|_{TV}$.

$\boxed{\text{unseen} \Downarrow}$

$\boxed{2, \text{M}}$

(b) (i) $Q(x,y) = \frac{P(z,y)-(1-\alpha)\pi(y)}{\alpha}$. Since $P(x,\cdot) \geq (1-\alpha)\pi(\cdot)$ we have all entries of Q are nonnegative and in addition $\sum_{y\in E} Q(x,y) = \frac{1}{\alpha}\sum_{y\in E} P(x,y) - \frac{1-\alpha}{\alpha}\sum_{y\in E}\pi(y) = 1$ so $Q(x,\cdot)$ is a probability measure for all x.

$\boxed{3, \text{M}}$

(ii) From Part (i) we have verified this for $k=1$. We proceed with an induction argument from $k-1$ to k:

$$Q^k = Q^{k-1}\cdot Q \quad \text{that is}$$

$$\begin{aligned}
Q^k(x,y) &= \sum_{z\in E} \frac{P^{k-1}(x,z) - (1-\alpha^{k-1})\pi(z)}{\alpha^{k-1}} \cdot \frac{P(z,y)-(1-\alpha)\pi(y)}{\alpha} \\
&= \frac{1}{\alpha^k}\sum_{z\in E}\{P^{k-1}(x,z)P(z,y) - (1-\alpha^{k-1})\pi(z)P(z,y) \\
&\quad - P^{k-1}(x,z)(1-\alpha)\pi(y) + (1-\alpha^{k-1})\pi(z))(1-\alpha)\pi(y)\} \\
&= \frac{1}{\alpha^k}\left(P^k(x,y) - \left(1-\alpha^{k-1}+1-\alpha-1+\alpha^{k-1}+\alpha-\alpha^k\right)\pi(y)\right) \\
&= \frac{1}{\alpha^k}\left(P^k(x,y) - \left(1-\alpha^k\right)\pi(y)\right),
\end{aligned}$$

which is the required expression.

$\boxed{4, \text{M}}$

(iii) We start by using the above expresion to get

$$P^k(x,y) - \pi(y) = \alpha^k\left(Q^k(x,y) - \pi(y)\right).$$

Then we have

$$\|P^k(x,\cdot) - \pi(\cdot)\|_{TV} = \frac{1}{2}\alpha^k \sum_{y\in E}\left|Q^k(x,y) - \pi(y)\right| \leq \alpha^k$$

where we used that the maximum total variation norm between two probability measures is 1.

$\boxed{4, \text{M}}$

(iv) This results quantifies the distance of the probability law of X_k to the stationary (and here limiting) distribution. We have an explicit rate of convergence and the chain converges exponentially/geometrically fast to stationarity with the speed of convergence depending on the parameter α that controls the similarity of P with π. If α is close to 0 they are very similar based on $P(x,\cdot) \geq (1-\alpha)\pi(\cdot)$. Higher similarity leads to faster rate of convergence. Note that both this bound and the convergence result hold for any initial condition x.

$\boxed{2, \text{M}}$

(c) Consider the following Markovian coupling (X_t, Y_t), where we let $\tau = \min_{s\in\mathbb{N}}\{s : X_s = Y_s\}$ and then

- for $t \leq \tau$ X_t and Y_t evolve independently as a Markov chains according to P with initial distributions $\nu = \delta_x$ and π respectively;
- for $t > \tau$ $X_t = Y_t$ and they evolve as a double copy of a Markov chain according to P.

Given this is the law of X_t and Y_t and they form a coupling we have

$$\|Law(X_t) - Law(Y_t)\|_{TV} = \|\delta_x P^t - \pi\|_{TV} \leq \mathrm{P}(X_t \neq Y_t).$$

Using the fact that $P(x,y) \geq (1-\alpha)\pi(y)$ means that $\beta = \min_{x,y}\{P(x,y)\} = (1-\alpha)\min_y \pi(y)$. This means that for any initial states x_0, y_0 for $t \leq \tau$ and an arbitrary target state $z \in E$ we have $\mathrm{P}(X_1 = z) \geq \beta$ and similarly $\mathrm{P}(Y_1 = z) \geq \beta$. Hence due to independence of the chains before the coupling time we get that $\mathrm{P}(X_1 \neq Y_1) \leq (1-\beta^2)$ and iterating in time gives $\mathrm{P}(X_t \neq Y_t) = (1-\beta^2)^t$, so

$$\|\delta_x P^t - \pi\|_{TV} \leq (1-\beta^2)^t.$$

$\boxed{5, \mathrm{M}}$

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: 100 of 80 marks

Total Mastery marks: 20 of 20 marks

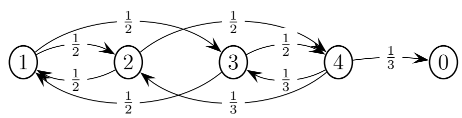

Figure 1: Transition Diagram for Question 2(a).

MATH60045 Applied Probability

Question	Marker's comment
2	Dear students, it has been a pleasure to see your progress and mark this question. Many of you did very well and I have no special notes for you except perhaps to say that in 2-a-iii, in the heat of the moment linear algebra was not everyone's forte and we decided to deduct just a minimal one mark as long as the starting point for calculating the mean escape time was derived correctly.
3	Q3 was answered quite well by a reasonable portion of the cohort. Parts (a) and (b) were generally well done, with lost marks coming most commonly from a lack of precision. Many students successfully checked the form of p_0(t) and p_1(t) given in (c) -- however a majority of students lost a mark here for failing to check that p_1(0)=0, as required. Parts (d) and (e) proved more challenging: in part (d), students often lost marks for failing to provide sufficient justification for the steps of their derivation; in part (e), a good portion of the cohort attempted the correct method, calculating the mixed derivative of the survivor function given in (d), though very few achieved full marks for an error-free final solution.
4	As a whole, Q4 proved challenging for most members of the cohort. In part (a), most students lost at least one mark, most often for failing to note that Z_t does not have independent increments, and so is not a Brownian motion. Part (b-i) was completed reasonably well by most; in part (b-ii), most students picked up one or two marks, but many dropped marks for missing key steps of the derivation (e.g. use of time homogeneity). Most students got marks in part (c-iii) for the correct high-level approach, substituting their answer from the previous part; full marks were available here for a completely correct answer using a correct expression for the transition probabilities. Many students left part (b-iv) unattempted; those that did attempt this often got one or two marks for sensible high-level statements about the impact of relaxing the independence assumption, but only a handful of students provided detailed discussion of how the work in parts (b-i), (b-ii) and (b-iii) would need revised

MATH70045 Applied Probability

Question	Marker's comment
2	As for MATH60048
3	Q3 was answered quite well by a reasonable portion of the cohort. Parts (a) and (b) were generally well done, with lost marks coming most commonly from a lack of precision. Many students successfully checked the form of p_0(t) and p_1(t) given in (c) -- however a majority of students lost a mark here for failing to check that p_1(0)=0, as required. Parts (d) and (e) proved more challenging: in part (d), students often lost marks for failing to provide sufficient justification for the steps of their derivation; in part (e), a good portion of the cohort attempted the correct method, calculating the mixed derivative of the survivor function given in (d), though very few achieved full marks for an error-free final solution
4	As a whole, Q4 proved challenging for most members of the cohort. In part (a), most students lost at least one mark, most often for failing to note that Z_t does not have independent increments, and so is not a Brownian motion. Part (b-i) was completed reasonably well by most; in part (b-ii), most students picked up one or two marks, but many dropped marks for missing key steps of the derivation (e.g. use of time homogeneity). Most students got marks in part (c-iii) for the correct high-level approach, substituting their answer from the previous part; full marks were available here for a completely correct answer using a correct expression for the transition probabilities. Many students left part (b-iv) unattempted; those that did attempt this often got one or two marks for sensible high-level statements about the impact of relaxing the independence assumption, but only a handful of students provided detailed discussion of how the work in parts (b-i), (b-ii) and (b-iii) would need revised