Imperial College London

M3S4

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May-June 2019

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

Applied Probability

Date: Wednesday 15 May 2019
Time: 14.00 - 16.00
Time Allowed: 2 Hours
This paper has 4 Questions.
Candidates should use ONE main answer book for Questions 1 and 2, and ONE main answer book for Question 3 and 4.
Supplementary books may only be used after the relevant main book(s) are full.
All required additional material will be provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Calculators may not be used.

© 2019 Imperial College London

Imperial College London

M4/5S4

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May-June 2019

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

Applied Probability

Date: Wednesday 15 May 2019
Time: 14.00 - 16.30
Time Allowed: 2 Hours 30 Minutes
This paper has 5 Questions.
Candidates should use ONE main answer book for Questions 1, 2 and 5, and ONE main answer book for Question 3 and 4.
Supplementary books may only be used after the relevant main book(s) are full.
All required additional material will be provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Calculators may not be used.

© 2019 Imperial College London

1. (a) (i) For two states s and t of a Markov chain, define what it means to say that s is accessible from t, and what it means to say that s and t communicate with each other.

(ii) Briefly give a practical or theoretical reason for why it would be useful to identify the communicating classes of a Markov chain (you can cite any facts about them without proof).

(b) Consider a Markov chain with state space $E = \{1, 2, 3, 4, 5\}$ and transition matrix given by

$$\begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{pmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{array}$$

(i) Draw the transition diagram.

(ii) Identify the communicating classes along with their corresponding periods (you don't need to justify your answer).

(iii) For each communicating class, pick a state contained in it, and work out its first passage time. Use this information to classify each communicating class as either recurrent or transient.

2. (a) Suppose that $Z_0, Z_1, \ldots$ are independent, identically distributed random variables such that $Z_i = 0$ with probability $\frac{2}{3}$ and $Z_i = 1$ with probability $\frac{1}{3}$.

In each of the following cases, determine only the state-space, E, and the transition matrix of the Markov chain $(X_n)_{n\geq 0}$ (no proofs required).

(i) $X_n = Z_n$

(ii) $X_n = 3Z_n + 1$

(iii) $X_n = Z_{4n}$

(iv) Set $X_0 = Z_0$, and for $n > 0$ set $X_n = X_{n-1} + Z_n$.

(b) A particular track on the London Underground serves three different lines: the District line, Circle line and Metropolitan line. The trains run on this track in the following pattern:

* A District line train is equally likely to be immediately followed by a train from any of the 3 lines, but 2 consecutive District line trains are always immediately followed by a Metropolitan line train only.

* A Circle line train is equally likely to be immediately followed by a train from the two other lines only (so is never immediately followed by another Circle line train).

* A Metropolitan line train is equally likely to be immediately followed by a train from any of the 3 lines.

(i) You are interested in modelling the lines of the trains running on the track as a time-homogeneous Markov chain with four states labelled $\{1, 2, 3, 4\}$. Present the transition matrix of this process, clearly define what each state represents in your model.

(ii) Work out the limiting probabilities of your model to show what fraction of trains running through the track belong to each of the three lines.

(iii) An expert from the Transport for London (TfL) highlights the following limitation with our model. Tracks on the underground can enter a state of malfunction due to signal failures, and moreover the probability of entering this state increases each time a train runs on the track. Briefly outline a way you can modify or enhance the model to capture this phenomena.

3. (a) (i) Define what it means to say that a random variable X has the memoryless property, and show that the exponential distribution is memoryless.

(ii) If X_1 and X_2 are two independent exponential random variables with respective mean $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$ show that:

$$P\{X_1 < X_2\} = \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

(iii) Alice enters a bank served by two clerks. She finds that there are only two other customers (Kat and Sarah) each being served by one of the clerks. She is told that she will be served by one of the clerks as soon as one of the two customers leaves. If the amount of time any clerk spends with a customer is exponentially distributed with mean $\frac{1}{\lambda}$, then what is the probability that out of the three customers Alice is the last to leave?

(b) (i) What does it mean to say that a stochastic process $\{X(t), t \geq 0\}$ is a compound Poisson process?

(ii) Suppose that students visit a cafe in groups, at a Poisson rate of $\lambda = 5$ per hour. If the number of students per group takes on the values 1,2,3,4 with respective probabilities $\frac{1}{4}, \frac{1}{3}, \frac{1}{6}, \frac{1}{4}$, then what is the expected value and variance of the number of students visiting the cafe in a three-hour period?

4. (a) (i) What are the four conditions that a counting process $\{N(t), t \geq 0\}$ must satisfy to be a nonhomogeneous Poisson process with intensity function $\lambda(t)$?

(ii) Define the mean value function of a nonhomogeneous Poisson process with intensity function $\lambda(t)$.

(b) Consider a nonhomogeneous Poisson process $N = (N_t)_{t\geq 0}$ with intensity function given by

$$\lambda(t) = \begin{cases} 5+5t, & \text{for } 0 \leq t \leq 3, \\ 20, & \text{for } 3 \leq t \leq 5, \\ t(t-1), & \text{for } 5 \leq t \leq 9 \\ 72, & \text{for } 9 < t \end{cases}$$

(i) Find $\mathbb{P}(N_2 = 1)$.

(ii) Find $\mathbb{P}(N_8 - N_6 = 0)$.

(c) A train company is interested in modelling the arrival of travellers to a particular train station in London. Briefly outline why using a splitting (also known as thinning) nonhomogeneous Poisson process might be an effective model to use for this application.

5. (a) Let $X_1, \ldots, X_n$ be a sequence of independent Bernoulli random variables with $P(X_i = 1) = p_i$. Let $X = \sum_{i=1}^{n} X_i$ and let $\mu = E[X]$.

   (i) State Markov's inequality.

   (ii) Prove that:

$$M_X(t) \leq \exp\{(e^t - 1)\mu\}$$

   (iii) Use the above inequality to prove that the following form of the Chernoff bound holds for any $\delta > 0$:

$$P(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

(b) Briefly give an example of an application (practical or theoretical) from probability or statistics that uses large deviations bounds on the sum of independent Bernoulli random variables.

DISCRETE DISTRIBUTIONS

	range $\mathbb{X}$	*parameters*	*pmf* f_X	*cdf* F_X	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	*mgf* M_X
$Bernoulli(\theta)$	$\{0,1\}$	$\theta \in (0,1)$	$\theta^x(1-\theta)^{1-x}$		θ	$\theta(1-\theta)$	$1-\theta+\theta e^t$
$Binomial(n,\theta)$	$\{0,1,...,n\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$		$n\theta$	$n\theta(1-\theta)$	$(1-\theta+\theta e^t)^n$
$Poisson(\lambda)$	$\{0,1,2,...\}$	$\lambda \in \mathbb{R}^+$	$\frac{e^{-\lambda}\lambda^x}{x!}$		λ	λ	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,...\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\frac{1}{\theta}$	$\frac{(1-\theta)}{\theta^2}$	$\frac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(n,\theta)$	$\{n,n+1,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{x-1}{n-1}\theta^n(1-\theta)^{x-n}$		$\frac{n}{\theta}$	$\frac{n(1-\theta)}{\theta^2}$	$\left(\frac{\theta e^t}{1-e^t(1-\theta)}\right)^n$
or	$\{0,1,2,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n+x-1}{x}\theta^n(1-\theta)^x$		$\frac{n(1-\theta)}{\theta}$	$\frac{n(1-\theta)}{\theta^2}$	$\left(\frac{\theta}{1-e^t(1-\theta)}\right)^n$

The PDF of the *multivariate normal distribution* is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{K/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\},$$

for $\boldsymbol{x} \in \mathbb{R}^K$ with $\boldsymbol{\Sigma}$ a $(K \times K)$ variance-covariance matrix and $\boldsymbol{\mu}$ a $(K \times 1)$ mean vector.

The location/scale transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = \frac{1}{\sigma} f_X\left(\frac{y-\mu}{\sigma}\right) \qquad F_Y(y) = F_X\left(\frac{y-\mu}{\sigma}\right)$$

$$M_Y(t) = e^{\mu t}M_X(\sigma t) \qquad \mathrm{E}[Y] = \mu + \sigma\mathrm{E}[X] \qquad \mathrm{Var}[Y] = \sigma^2\mathrm{Var}[X]$$

The *gamma function* is given by $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$.

CONTINUOUS DISTRIBUTIONS

		parameters	*pdf*	*cdf*	E[X]	Var[X]	*mgf*
$Uniform(\alpha,\beta)$ (stand. model $\alpha=0,\beta=1$)	(α,β)	$\alpha<\beta\in\mathbb{R}$	$\frac{1}{\beta-\alpha}$	$\frac{x-\alpha}{\beta-\alpha}$	$\frac{(\alpha+\beta)}{2}$	$\frac{(\beta-\alpha)^2}{12}$	$\frac{e^{\beta t}-e^{\alpha t}}{t(\beta-\alpha)}$
$Exponential(\lambda)$ (stand. model $\lambda=1$)	$\mathbb{R}^+$	$\lambda\in\mathbb{R}^+$	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)$
$Gamma(\alpha,\beta)$ (stand. model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$		$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(\frac{\beta}{\beta-t}\right)^\alpha$
$Weibull(\alpha,\beta)$ (stand. model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}$	$1-e^{-\beta x^\alpha}$	$\frac{\Gamma(1+1/\alpha)}{\beta^{1/\alpha}}$	$\frac{\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma\left(1+\frac{1}{\alpha}\right)^2}{\beta^{2/\alpha}}$	
$Normal(\mu,\sigma^2)$ (stand. model $\mu=0,\sigma=1$)	$\mathbb{R}$	$\mu\in\mathbb{R},\sigma\in\mathbb{R}^+$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$		μ	σ^2	$e^{\{\mu t+\sigma^2t^2/2\}}$
$Student(\nu)$	$\mathbb{R}$	$\nu\in\mathbb{R}^+$	$\frac{(\pi\nu)^{-\frac{1}{2}}\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left\{1+\frac{x^2}{\nu}\right\}^{(\nu+1)/2}}$		0 (if $\nu>1$)	$\frac{\nu}{\nu-2}$ (if $\nu>2$)	
$Pareto(\theta,\alpha)$	$\mathbb{R}^+$	$\theta,\alpha\in\mathbb{R}^+$	$\frac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}$	$1-\left(\frac{\theta}{\theta+x}\right)^\alpha$	$\frac{\theta}{\alpha-1}$ (if $\alpha>1$)	$\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$ (if $\alpha>2$)	
$Beta(\alpha,\beta)$	$(0,1)$	$\alpha,\beta\in\mathbb{R}^+$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$		$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	

M345S4 2018-2019 Exam Solutions

Q1

a)
i) [4 Marks, A, seen]
State j is accessible from state i if if the chain may visit state j with positive probability starting from i. In other words there exists $m \geq 0$ such that $p_{ij}(m) > 0$.

We say that j and i communicate with each other if i is accessible from j and j is accessible from i.

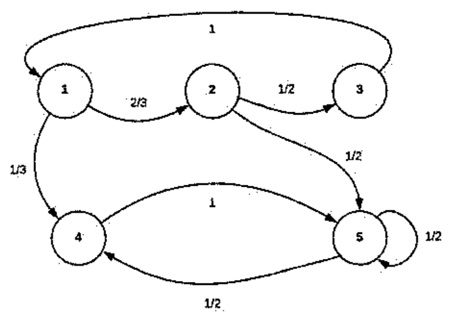
ii) [4 Marks, A, seen]

Identifying communicating classes simplifies the analysis of Markov chains, since we have proven in the module that states in the same communicating class share some important properties.

In particular, they share periodicity, (null) recurrence and transience. We also proved in the module the Decomposition Theorem, which says any Markov chain can be broken down along its communicating classes.

This means that rather than computing the periodicity of each state, you can work it out for one state per class, which can be computationally helpful in analysing large Markov chains.

b)
i) [5 Marks, B, bookwork]



ii) [3 Marks, C, bookwork]
$C1 = \{1,2,3\}$, period 3, and $C2 = \{4,5\}$ with period 1. ($C1$ has period 3 because, by inspection, the set $\{n : p_{11}^n > 0\} = \{3,6,9,\ldots\}$, while $C2$ has period 1 because $p_{55}^1 = \frac{1}{2}$).

iii) [4 Marks, D, bookwork]
State 1 has $f_{11}(1) = 0, f_{11}(2) = 0, f_{11}(3) = \frac{2}{3}\frac{1}{2} = \frac{1}{3}$ and then $f_{11}(n) = 0$ for all $n \geq 4$. Exactly similar calculations for state 2 and 3, yield $f_{22} = f_{33} = f_{11}$. A correct solution needs to consider only one of state 1, 2 or 3.
Hence we have that $f_{11} = \frac{1}{3} < 1$, so it is transient implying that all of class 1 is transient.

For state 4 we have that $f_{44}(1) = 0, f_{44}(2) = 1 \cdot \frac{1}{2}$, and in general $f_{44}(n) = \frac{1}{2}^{n-1}$.
So $f_{44} = \frac{1}{2}(1 + \frac{1}{2} + \frac{1}{4} + \ldots) = 1$.
State 5 has $f_{55}(1) = \frac{1}{2}, f_{55}(2) = \frac{1}{2}$ and then $f_{55}(n) = 0$ for all $n \geq 3$. So we have that $f_{55} = \frac{1}{2} + \frac{1}{2} = 1$. Again for a correct solution only one of state 4 or 5 need to be considered.
From $f_{44} = 1$ or $f_{55} = 1$ we can conclude that the whole of the second class is recurrent.

Q2

a)
i) [1 Mark, A, seen]
$E = \{0, 1\}$

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

ii) [2 Marks, A, seen]
$E = \{1, 4\}$

$$\begin{array}{c c} & \begin{array}{cc} 1 & 4 \end{array} \\ \begin{array}{c} 1 \\ 4 \end{array} & \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} \end{array}$$

iii) [2 Marks, D, unseen]
$E = \{0, 1\}$

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

iv) [2 Marks, D, unseen]
$E = \{0, 1, 2, 3, \ldots\}$

$$\begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \\ \end{array} & \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 & \cdots \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 & \cdots \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} & \cdots \\ 0 & 0 & 0 & \frac{2}{3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}$$

b) i) [5 Marks, B, bookwork]
Define the states as E = { D1, D2, C, M, }.
Where D1 is the state of a District line train arriving that wasn't preceded by another District line train. D2 is the state of a District line train arriving that was preceded by another District line train. While C and M correspond to a Circle and Metropolitan line arriving respectively.

$$
\begin{array}{c}
\\ D1 \\ D2 \\ C \\ M
\end{array}
\begin{array}{c}
\begin{array}{cccc} D1 & D2 & C & M \end{array} \\
\begin{pmatrix}
0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
0 & 0 & 0 & 1 \\
\frac{1}{2} & 0 & 0 & \frac{1}{2} \\
\frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3}
\end{pmatrix}
\end{array}
$$

ii) [6 Marks, A, bookwork]
We have to solve the steady-state equations:

$$\pi P = \pi$$
$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$$

Solving the equations (for example by Gaussian elimination) we get the following values:

$\pi_1 = \frac{9}{35}$

$\pi_2 = \frac{3}{35}$

$\pi_3 = \frac{8}{35}$

$\pi_4 = \frac{15}{35}$

Hence approximately $\frac{9}{35} + \frac{3}{35} = \frac{12}{35}$ of the trains are from the District line, $\frac{8}{35}$ are from the Circle line, and $\frac{15}{35}$ are from the Metropolitan line.

iii) [2 Marks, C, unseen]
One idea is to introduce a new null recurrent state to correspond to the track malfunctioning, and adjust the model so that each of the original four states can transition into the new state with a non-zero probability. This captures the fact that the signal failure can occur after any train passes through the track, irrespective of the line of the train.

Moreover the process can then be modelled as a time *inhomogeneous* Markov chain, so that the probability of getting into the new state from the original four increases with each iteration. This captures the fact that the chance of a signal failure on the track increases with each passing train.

Q3

a)

i) [4 Marks, A, seen]

Memoryless property: $P\{X > s+t | X > t\} = P\{X > s\}$. (other equivalent versions acceptable). If X is an exponentially distributed random variable with parameter λ, then $P\{X > s+t\} = e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t} = P\{X > s\}P\{X > t\}$. Therefore the exponential distribution satisfies the memoryless property.

ii) [3 Marks, C, seen]

Let X_1, X_2 be exponentially distributed with parameter λ_1 and λ_2 respectively.

$$P\{X_1 < X_2\} = \int_0^\infty P\{X_1 < X_2 \mid X_1 = x\}\{X_1 = x\}\, dx$$

Where we used the law of probability.

$$\int_0^\infty P\{X_1 < X_2 \mid X_1 = x\}\{X_1 = x\}\, dx$$

$$= \int_0^\infty P\{X_1 < X_2 \mid X_1 = x\}\lambda_1 e^{-\lambda_1 x}\, dx$$

$$= \int_0^\infty P\{x < X_2\}\lambda_1 e^{-\lambda_1 x}\, dx$$

$$= \int_0^\infty e^{-\lambda_2 x}\lambda_1 e^{-\lambda_1 x}\, dx$$

$$= \int_0^\infty \lambda_1 e^{-(\lambda_1+\lambda_2)x}\, dx$$

$$= \frac{\lambda_1}{\lambda_1+\lambda_2}$$

iii) [4 Marks, D, unseen]

The answer is $\frac{1}{2}$. This is reasoned as follow. When Alice starts to get served by one of the clerks (due to one of the customers leaving, say Sarah) then, due to the memoryless property of the exponential, it would be as if Kat was starting her service at this point with Alice. The answer then follows by symmetry.

b)

i) [3 Marks, A, seen]
A stochastic process $\{X(t), t \geq 0\}$ is said to be a compound Poisson process if it can be represented as:

$$X(t) = \sum_{i=1}^{N(t)} Y_i, \quad t \geq 0$$

Where $\{N(t), t \geq 0\}$ is a Poisson process and $\{Y_i, i \geq 1\}$ is a family of independent and identically distributed random variables that is also independent of $\{N(t), t \geq 0\}$.

ii) [6 Marks, B, bookwork]
Letting Y_i denote the number of students in the ith group, we have:

$E[Y_i] = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{4} = \frac{29}{12}$

$E[Y_i^2] = 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{3} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{4} = \frac{85}{12}$

Hence, letting $X(3)$ denote the number of students visiting during a three hour period, from the results proved in the lecture we have that:

$E[X(3)] = 5 \cdot 3 \cdot \frac{29}{12} = \frac{145}{4}$.

$Var(X(3)) = 5 \cdot 3 \cdot \frac{85}{12} = \frac{425}{4}$.

Q4

a)
i) [6 Marks, A, seen]
The process $N(t)$ must satisfy the following conditions:

- $N(0) = 0$,
- $\{N(t), t \geq 0\}$ has independent increments,
- $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$
- $P(N(t+h) - N(t) \geq 2) = o(h)$.

ii) [2 Marks, A, seen]

The mean value function $m(t)$ is defined as $m(t) = \int_0^t \lambda(u)du$.

b)

i) [4 Marks, B, bookwork]
We know that $N_2 \sim \text{Poi}(m(2))$, where

$$m(2) = \int_0^2 \lambda(u)du = \int_0^2 (5+5u)du = 5u + \frac{5}{2}u^2 \Big|_0^2 = 20$$

Hence $\mathbb{P}(N_2 = 1) = 20\exp(-20)$.

ii) [4 Marks, C, bookwork]
Here we have that $N_8 - N_6 \sim \text{Poi}(m(8) - m(6))$, where

$$m(8) - m(6) = \int_6^8 \lambda(u)du = \int_6^8 u(u-1)du = \frac{u^3}{3} - \frac{u^2}{2}\Big|_6^8 = \frac{254}{3}$$

Hence $\mathbb{P}(N_8 - N_6 = 0) = \exp(-\frac{254}{3})$.

c) [4 Marks, D, unseen]
An answer that gets full marks must highlight why the application makes use of both the multi-type nature of splitting processes, and the variability of the arrival rate of non-homogeneous Poisson processes.

An example would be saying that you can have three types of commuters, say students, city workers, and tourists. For each of these you can have an arrival rate depending on the time of day for example: morning rush hour, lunch, evening rush hour, and late night.

This model allows us to capture the fact that for example the city worker type has big arrival rates in the morning and late rush hour intervals (but a small arrival rate in the intervals between or after), while the tourist type has a more uniform arrival rate throughout all intervals of the day.

Q5

a)
i) [3 Marks, A, seen]
Markov's inequality: Let X be a random variable that assumes only non-negative values. Then, for all $a > 0$:

$$P(X \geq a) \leq \frac{E[X]}{a}$$

ii) [7 Marks, B, bookwork]

First we need to compute M_{X_i}

$M_{X_i}(t) = E[e^{tX_i}]$

$p_i e^t + (1 - p_i)$

$= 1 + p_i(e^t - 1)$

$\leq e^{p_i(e^t-1)}$

Since X is a sum of n random variable, the mgf of X will be the product of the mgfs of the n random variables. So we have:

$M_X(t) = \prod_{i=1}^{n} M_{X_i}(t)$

$\leq \prod_{i=n}^{n} exp\{p_i(e^t - 1)\}$

$= exp\{\sum_{i=1}^{n} p_i(e^t - 1)\}$

$= exp\{(e^t - 1)\mu\}$.

iii) [6 Marks, C, bookwork] Applying Markov's inequality we have:

$P(X \geq (1 + \delta)\mu) = P(e^{tX} \geq e^{t(1+\delta)})$

$\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}}$

Now we can use what was proved in the previous question to get:

$\leq \frac{exp\{e^t-1)\mu\}}{exp\{t(1+\delta)\mu\}}$

Now we can optimise the above for t by differentiating to work out the minimum, which gives the value $t = ln(1 + \delta) > 0$.

Setting this in the above we obtain:

$$P(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

b) [4 Marks, D, seen]
There are two particular examples mentioned in the mastery material.

The first is that computing tail-bounds helps in parameter estimation in statistics (the particular application discussed in the material is to do with genetic sequencing).

In this application samples can be thought of as trials in a Poisson process with an unknown parameter, then we can use tail bounds to workout how likely the value of parameter falls within a certain range based on the samples collected.

The second is that computing tail-bounds help in analysing the run-time performance of randomised algorithms in applied probability. Other examples are acceptable (there are many that can be thought of) so long as they highlight the application's use of computing tail bounds.