UNIVERSITY OF LEICESTER

MA7404
All Candidates

Semester 1 Examinations 2021

DO NOT OPEN THE QUESTION PAPER UNTIL INSTRUCTED TO DO SO BY THE CHIEF INVIGILATOR	
School	Mathematics and Actuarial Science
Module Code	MA7404
Module Title	Markov processes
Exam Duration	Two hours + 45 minutes upload time
CHECK YOU HAVE THE CORRECT QUESTION PAPER	
Number of Pages	4
Number of Questions	4
Instructions to Candidates	Answer all questions
FOR THIS EXAM YOU ARE ALLOWED TO USE THE FOLLOWING:	
Calculators	Yes
Books/Statutes provided by the University	Formulae and tables for actuarial examinations
Are students permitted to bring their own Books/Statutes/Notes?	Yes
Additional Stationery	Yes

1. An insurance company receives claims of sizes $X_1, X_2, \dots$. If claim X_i exceeds £1000, the insurance company pays £1000, and the remaining part $X_i - 1000$ is covered by reinsurance. Hence, the total payments made by the insurance company during the next week is $S = Y_1 + Y_2 + \cdots + Y_N$, where N is the (random) number of claims to be received next week, and $Y_i = \min\{X_i, 1000\}$ for $i = 1, 2, \dots, N$. The company assumes that N can take values from 0 to K with equal probabilities, X_i are independent from each other and from N, identically distributed, and follow the uniform distribution on $[0, B]$. Parameters K and B are unknown and should be estimated from the past data.

   (i) [5 marks] The numbers of claims during the last 10 weeks were

   $$4, 0, 5, 6, 1, 0, 3, 2, 5 \text{ and } 4.$$

   Use the method of maximum likelihood to estimate parameter K.

   (ii) [10 marks] The sizes of the last 10 claims were

   $$700, 800, 500, 100, 1000+, 900, 800, 1000+, 200 \text{ and } 300,$$

   where $1000+$ indicates that the claim size is over £1000 but the exact size is unknown.

   (a) **[5 marks]**. Which of the following methods would be the best to use to estimate parameter B: method of moments, method of percentiles at level $\alpha = 0.75$, or method of percentiles at level $\alpha = 0.95$? Please justify your answer.

   (b) **[5 marks]**. Use the method selected in Part (a) to estimate the parameter B.

   (iii) [10 marks]

   (a) **[4 marks]**. Estimate the mean and variance of N.

   (b) **[4 marks]**. Estimate the mean and variance of Y_i.

   (c) **[2 marks]**. Estimate the mean and variance of S.

   Total: 25 marks

2. **(i) [10 marks]** A company receives claims from two policies. Let N and M be the numbers of claims from these policies, respectively. The company assume that there are three equally likely scenarios: (1) $N = 0$, $M = 2$, (2) $N = 1$, $M = 0$, and (3) $N = 2$, $M = 1$.

   (a) **[4 marks]**. Estimate Pearson correlation coefficient $Corr(N, M)$.

   (b) **[4 marks]**. Estimate Kendall coefficient of concordance.

   (c) **[2 marks]**. Based on the results in parts (a) and (b), discuss where N and M has direct dependence, inverse dependence, or independent.

   (ii) [10 marks] The sizes of claims from both policies follow the Gamma distribution, but with different parameters $\alpha > 0$ and $\gamma > 0$. Using the limiting density ratios test, determine whether the tail of the Gamma distribution becomes heavier or lighter if

   (a) **[5 marks]**. λ is fixed and α increases;

   (b) **[5 marks]**. α is fixed and λ increases.

(iii) [5 marks] The claim sizes X and Y are dependent with copula $C(u,v) = \min(u,v)$. Calculate the coefficient of upper tail dependence of X and Y.

Total: 25 marks

3. An insurance company is considering investing their capital into a financial instrument whose current price is $P_0 = 1$ and future price after t weeks is modelled as

$$P_t = \exp(0.1t + 0.2X_t), \quad t = 0, 1, 2, \dots,$$

where X_t is the symmetric simple unrestricted random walk on integers starting at $X_0 = 0$.

(i) [5 marks] State whether each of the following statements is true or false. The stochastic process X_t:

(a) **[1 mark]**. is stationary;

(b) **[1 mark]**. is weakly stationary;

(c) **[1 mark]**. has independent increments;

(d) **[1 mark]**. has stationary increments;

(e) **[1 mark]**. is a Markov chain.

(ii) [5 marks] Answer the same questions as in Part (i) for the stochastic process P_t.

(iii) [5 marks] Estimate the probability that $X_7 = 5$.

(iv) [5 marks] Find the expectation and variance of X_7.

(v) [5 marks] Estimate the probability that $P_7 < 2$.

Total: 25 marks

4. To model COVID-19 sickness and mortality, a life insurance company uses a time-inhomogeneous sickness-death model, with states H (healthy), S (sick), C (critically sick) and D (dead), and age-dependent transition rates $\sigma(t)$, $\rho(t)$, $\mu(t)$, $v(t)$ and $w(t)$ for transitions $H \to S$, $S \to C$, $C \to D$, $C \to S$, and $S \to H$, respectively. Here t is an age measured in years. All other transition rates are assumed to be 0.

(i) [10 marks] Write down

(a) **[5 marks]** the generator matrix and the transition graph.

(b) **[5 marks]** the Kolmogorov forward equation for transition probability $p_{HH}(s,t)$ $\left(\frac{\partial p_{HH}(s,t)}{\partial t} = \dots\right)$.

(ii) [5 marks] Assuming that $\rho(t) = 0.1\sqrt{t}$, and $w(t) = 200/t$, estimate the probability that a person aged 50, currently in state S, will remain there continuously for the next year.

(iii) [5 marks] Is the transition probability $p_{CC}(50,51)$ greater than, less than, or equal to the answer computed in Part (ii)? Justify your answer.

(iv) [5 marks] To apply this model, you need to be able to classify patients as those considered to be “sick” and those considered to be “critically sick”. You have a sample of $10,000$ patients you want to classify. The doctors can classify for you any 100 patients on your choice, but no more because they are too busy. They are, however, ready to give you all the information about all patients, such as their temperature, blood pressure, etc. Explain how you can use the nearest-neighbour method in supervised machine learning to classify the remaining patients.

Total: 25 marks