

INSTITUTE AND FACULTY OF ACTUARIES

EXAMINATION

29 September 2020 (am)

Subject CS2A – Risk Modelling and Survival Analysis Core Principles

Time allowed: Three hours and fifteen minutes

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator from the approved list.

If you encounter any issues during the examination please contact the Examination Team on T. 0044 (0) 1865 268 873.

- 1** Consider a two-dimensional Gaussian copula function, $C_{\text{Gauss}}(u_1, u_2)$, with parameter $\rho = 0$:
- (i) Give the solution to the copula function $C_{\text{Gauss}}(1, 1)$. [1]
 - (ii) Give the solution to the copula function $C_{\text{Gauss}}(1, 0.2)$. [1]
 - (iii) Give the solution to the copula function $C_{\text{Gauss}}(0.2, 0.2)$. [1]
 - (iv) Outline how your answers to parts (i), (ii) and (iii) would change if $\rho = 1$. [2]

An insurer uses copulas to model the dependencies between various types of claims. A statistical analysis of the insurer's claims shows that a Gumbel copula is a better representation of historic claims data than a Gaussian copula.

- (v) Discuss why the use of a Gaussian copula in a claims model could result in solvency issues for this insurer. [3]
- [Total 8]

- 2** (i) Explain why Extreme Value Theory (EVT) models can be useful. [2]

A sports scientist is interested in analysing the probability that the javelin world record may be broken next year and is intending to use EVT to do this.

The sports scientist has obtained data for the distances of all javelin throws from all javelin competitions in 2019. The total number of throws recorded was 3,000. The sports scientist has carried out an EVT analysis using the Generalised Pareto Distribution by selecting only those throws that exceeded 50 metres. This resulted in the longest 150 throws being selected for the analysis.

The following parameters were obtained from the EVT analysis:

$$\begin{aligned}\beta &= 15, \\ \gamma &= 3.\end{aligned}$$

- (ii) Determine the percentage of javelin throws that would be expected to exceed 70 metres next year. [4]
 - (iii) Comment on the limitations of this analysis. [4]
- [Total 10]

- 3** Consider the following time series model:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + 0.7e_{t-1} + e_t$$

where e_t is a white noise process with variance σ^2 .

- (i) Express this model in terms of the backwards shift operator B , simplifying your expression as far as possible. [2]
 - (ii) Identify the values of p , d and q for which the model is an ARIMA(p, d, q) model. [1]
 - (iii) Which **one** of the following options represents the correct values of $\text{Cov}(e_t, \nabla Y_t)$ and $\text{Cov}(e_{t-1}, \nabla Y_t)$ where $\text{Cov}()$ is the covariance function and ∇ is the difference operator?
 - A $\text{Cov}(e_t, \nabla Y_t) = \sigma^2$, $\text{Cov}(e_{t-1}, \nabla Y_t) = 1.2\sigma^2$
 - B $\text{Cov}(e_t, \nabla Y_t) = 1.7\sigma^2$, $\text{Cov}(e_{t-1}, \nabla Y_t) = 2.55\sigma^2$
 - C $\text{Cov}(e_t, \nabla Y_t) = 1.8\sigma^2$, $\text{Cov}(e_{t-1}, \nabla Y_t) = 1.6\sigma^2$
 - D $\text{Cov}(e_t, \nabla Y_t) = 3.4\sigma^2$, $\text{Cov}(e_{t-1}, \nabla Y_t) = 3.4\sigma^2$
 - (iv) Which **one** of the following options represents the correct values of the autocovariance function, γ , of ∇Y_t ?
 - A $\gamma_0 = 0.5\gamma_1 + 1.84\sigma^2$, $\gamma_1 = 0.5\gamma_0 + 0.7\sigma^2$, $\gamma_k = 0.5\gamma_{k-1}$ for $k > 1$
 - B $\gamma_0 = 0.5\gamma_1 + 2.92\sigma^2$, $\gamma_1 = 0.5\gamma_0 + 2.86\sigma^2$, $\gamma_k = 0.5\gamma_{k-1}$ for $k > 1$
 - C $\gamma_0 = 0.5\gamma_1 + 3.49\sigma^2$, $\gamma_1 = 0.5\gamma_0 + 1.19\sigma^2$, $\gamma_k = 0.5\gamma_{k-1}$ for $k > 1$
 - D $\gamma_0 = 0.5\gamma_1 + 5.78\sigma^2$, $\gamma_1 = 0.5\gamma_0 + 5.78\sigma^2$, $\gamma_k = 0.5\gamma_{k-1}$ for $k > 1$
 - (v) Which **one** of the following options represents the correct values of γ_0 , and γ_k for $k \geq 1$?
 - A $\gamma_0 = 2.92\sigma^2$, $\gamma_k = 0.5^{k-1}(2.16\sigma^2)$ for $k \geq 1$
 - B $\gamma_0 = 5.44\sigma^2$, $\gamma_k = 0.5^{k-1}(3.91\sigma^2)$ for $k \geq 1$
 - C $\gamma_0 = 5.8\sigma^2$, $\gamma_k = 0.5^{k-1}(5.76\sigma^2)$ for $k \geq 1$
 - D $\gamma_0 = 11.56\sigma^2$, $\gamma_k = 0.5^{k-1}(11.56\sigma^2)$ for $k \geq 1$
- [2]
[Total 10]

- 4** Zebras in a large African wildlife park have recently become susceptible to a particular disease called Zebra rabies. Zebras often die as a direct result of contracting this disease. An investigation to monitor deaths due to this disease was carried out between 1 January 2018 and 1 January 2019.

A researcher was interested in the rate at which zebras die once contracting this disease and decided to monitor the health of each zebra on the first day of each month. All zebras in the wildlife park were tagged to ensure that they were identifiable.

14 zebras were diagnosed with rabies during 2018. The data recorded on these 14 zebras are set out below:

<i>Reference tag</i>	<i>Date of diagnosis</i>	<i>Date of death</i>	<i>Reason for death</i>
1	1 Jan 2018	1 Jun 2018	Rabies
3	1 Jan 2018	1 Dec 2018	Rabies
4	1 Apr 2018	1 Jul 2018	Killed by lion
7	1 Apr 2018	1 Jun 2018	Rabies
8	1 Apr 2018	1 Dec 2018	Rabies
10	1 Jul 2018	1 Sep 2018	Rabies
11	1 Aug 2018	1 Oct 2018	Rabies
12	1 Aug 2018	1 Nov 2018	Rabies
19	1 Sep 2018	1 Oct 2018	Rabies
20	1 Oct 2018	1 Nov 2018	Rabies

Two zebras (reference tags 9 and 21) escaped from the wildlife park on 1 December 2018 having been diagnosed with rabies on 1 July 2018 and 1 November 2018, respectively.

In addition, the following two zebras that contracted the disease were still alive at the end of the investigation:

<i>Reference tag</i>	<i>Date of diagnosis</i>
13	1 Aug 2018
25	1 Dec 2018

- (i) Discuss whether or not the different types of censoring present in the above investigation are likely to be informative. [3]
 - (ii) Determine the Kaplan–Meier estimate of the survival function, where the decrement of interest is death due to rabies. [8]
- [Total 11]

- 5** An investigation was undertaken into the mortality of policyholders for a large life insurance company. The crude mortality rates were graduated using a formula of the form:

$$\mu_x = \exp(ax + bx^2).$$

An extract of the results is set out below. All data have been collated between 1 January 2018 and 1 January 2019 on an ‘age last birthday’ basis:

<i>Age</i>	<i>Exposed-to-risk (years)</i>	<i>Observed deaths</i>	<i>Graduated rates</i>
50	23,308	70	0.00368
51	19,316	58	0.00379
52	16,914	54	0.00391
53	21,082	90	0.00402
54	14,820	70	0.00415
55	24,084	96	0.00428
56	28,076	114	0.00441
57	22,958	86	0.00455
58	24,960	102	0.00469
59	21,134	86	0.00485
60	18,374	94	0.00500

- (i) Perform the signs test to explore the hypothesis that the graduated mortality rates are the true rates underlying the observed data from the life insurance company. [5]
 - (ii) State the main limitation of the signs test in assessing the suitability of the graduated rates. [1]
 - (iii) Perform a chi-square goodness-of-fit test to explore further the hypothesis that the graduated mortality rates are the true rates underlying the observed data from the life insurance company. [4]
 - (iv) Comment on your answers to parts (i) and (iii). [2]
- [Total 12]

- 6** An online professional examination paper is sat by a large number of candidates. After the examination, each script is marked electronically by two examiners in sequence.

All scripts are initially sent to a First Marking folder (F) for marking by a first examiner.

Once the first examiner has marked a script, that script is immediately sent to a Second Marking folder (S) for marking by a second examiner.

Once both examiners have marked a script the marks are assessed to determine whether the script is in the ‘borderline’ category. A script is ‘borderline’ if one of the marks is below the pass mark and the other one is above.

If the script is **not** placed in the ‘borderline’ category, it is automatically moved to the Completed folder (C). If the script is placed in the ‘borderline’ category, it is sent to the Borderline folder (B). A third examiner will then mark the ‘borderline’ script and forward it to the Completed folder once finished.

The first and second examiners mark scripts at a rate μ per day. The third examiner marks ‘borderline’ scripts at a rate β per day. A proportion b of the scripts are deemed to be ‘borderline’.

The Kolmogorov forward equation for $p^{SS}(t)$ is $\frac{d}{dt}p^{SS}(t) = -\mu p^{SS}(t)$.

- (i) Which **one** of the following options represents the Kolmogorov forward equation for $p^{SB}(t)$?

- A $\frac{d}{dt}p^{SB}(t) = b\mu p^{SS}(t) - \mu p^{SB}(t)$
B $\frac{d}{dt}p^{SB}(t) = b\mu p^{SS}(t) - \beta p^{SB}(t)$
C $\frac{d}{dt}p^{SB}(t) = b\mu p^{BB}(t) - \mu p^{SB}(t)$
D $\frac{d}{dt}p^{SB}(t) = b\mu p^{BB}(t) - \beta p^{SB}(t)$

[2]

In a certain examination marking process, $\mu = 0.2$, $\beta = 0.3$ and $b = 0.25$.

- (ii) Which **one** of the following options represents the correct probability that a script now awaiting second marking will be found in the Borderline folder in t days’ time?

- A $0.5\exp(-0.2t) - 0.5\exp(-0.3t)$
B $0.05t\exp(-0.3t)$
C $0.5\exp(-0.2t)$
D $0.05t\exp(-0.2t)$

[4]

- (iii) Which **one** of the following options represents the correct probability that a script now awaiting second marking will be found in the Completed folder in 10 days' time?
- A 0.7970
 B 0.8219
 C 0.8398
 D 0.8647

[2]

The Examination Board, who are responsible for overseeing the examination process, wants to investigate ways of speeding up the script marking process. One of the members of the Board has suggested that more examiners could be added to the pool of first and second examiners.

- (iv) Comment on the Board member's suggestion. [4]
 [Total 12]

- 7 The transition matrix, P , describes a Markov chain for the state space $S = \{1, 2, 3, 4\}$:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- (i) Explain whether this Markov chain is irreducible. [3]
 (ii) Determine all of the stationary probability distributions, π , for the Markov chain defined above, if any such probability distributions exist. [3]

The hitting probability, h_{ij} , is defined as the probability of **ever** reaching state j , starting from the initial state i .

- (iii) Determine the hitting probabilities for state 4 (h_{i4}) for all states i in S . [6]
 [Total 12]

- 8** A computer manufacturing company is monitoring the quality of the electronic circuit boards it produces. It has two factories, which are located in Oxford and London. The Oxford factory was opened only 12 months ago whereas the London factory has been producing circuit boards for 10 years.

The owner of the company wants to know whether the quality of production differs materially between the two factories. In addition, the owner wants to test whether a new manufacturing process improves the production quality. Over the last 6 months, half of the circuit boards in both factories have been made using the new process.

An analyst has been employed to test whether the location and the manufacturing process affect the quality of the circuit boards. The quality is measured by recording how long each circuit board lasts before developing a fault. The analyst decides to use the Cox proportional hazards model in the analysis.

The two covariates used are:

$$\text{Location: London} = 0, \text{Oxford} = 1, \quad \text{Process: Old} = 0, \text{New} = 1.$$

- (i) State the type of circuit board that represents the baseline hazard. [1]

During the investigation, 6,000 boards were tested, equally split between the four possible permutations of location and process. The first board to fail was one made in Oxford using the new process. No previous censoring events had occurred.

- (ii) State the term in the partial likelihood expression that relates to this first failure, clearly defining all notation used. [3]

The following coefficients were obtained from the investigation using the Cox proportional hazards model:

$$\text{Location} = 0.01, \quad \text{Process} = -0.30.$$

- (iii) Comment on the coefficient values obtained. [2]

- (iv) Identify the additional information that would be required before being able to reach a conclusion on the effects of location and process on the quality of the circuit boards produced. [1]

Prior to the introduction of the new process, the probability that a circuit board made in London failed in the first year was estimated to be 20%. The company plans to make a batch of 10,000 circuit boards next year in London.

- (v) Which **one** of the following options represents the best estimate of the probability that a circuit board made in London using the new process does **not** fail in the first year?

- A 0.5927
- B 0.6965
- C 0.7982
- D 0.8476

[3]

(vi) Which **one** of the following options represents the best estimate of the expected absolute difference in the number of circuit board failures in the first year if the new process is used, compared to if the old process is used?

- A 18
- B 476
- C 1,035
- D 2,073

[2]

[Total 12]

- 9** An insurance company offers annual home insurance policies in partnership with a bank. The distribution deal involves taking part in the bank's loyalty scheme called '1234'. Under '1234', a customer gets a discount when buying or renewing the policy according to how many bank accounts they hold, as follows:

<i>Number of bank accounts</i>	<i>Discount</i>
One	5%
Two	10%
Three	15%
Four or more	25%

An analysis of the data suggests that the transition matrix for the number of bank accounts held at annual intervals is as follows:

$$\begin{array}{l} \text{One} \\ \text{Two} \\ \text{Three} \\ \text{Four+} \end{array} \left(\begin{array}{cccc} 0.5 & 0.2 & 0.2 & 0.1 \\ 0.2 & 0.4 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0 & 0.2 & 0.2 & 0.6 \end{array} \right)$$

A customer takes out a policy in January 2017 at the 10% discount level.

- (i) Calculate the probability that the customer remains at the 10% discount level for all of their renewals up to and including 2020. [2]
 - (ii) Calculate the probability that the customer receives a discount of at least 15% in 2019. [2]
 - (iii) Which **one** of the following options represents the correct stationary distribution, π , of the transition matrix above?
 - A $\pi_1 = \frac{34}{45}\pi_3$, $\pi_2 = \frac{8}{9}\pi_3$, π_3 , $\pi_4 = \frac{41}{45}\pi_3$
 - B $\pi_1 = \frac{34}{45}\pi_3$, $\pi_2 = \frac{8}{9}\pi_3$, π_3 , $\pi_4 = \pi_3$
 - C $\pi_1 = \pi_3$, $\pi_2 = \frac{8}{9}\pi_3$, π_3 , $\pi_4 = \frac{41}{45}\pi_3$
 - D $\pi_1 = \pi_3$, $\pi_2 = \pi_3$, π_3 , $\pi_4 = \pi_3$
 - (iv) Which **one** of the following options represents the correct average long-term level of discount?
 - A 13.60%
 - B 13.75%
 - C 14.19%
 - D 14.45%
 - (v) Comment on the commercial implications for the insurer of the bank's loyalty scheme. [3]
- [Total 13]

END OF PAPER

INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

September 2020

Subject CS2A –Risk Modelling and Survival Analysis Core Principles

Introduction

The Examiners' Report is written by the Chief Examiner with the aim of helping candidates, both those who are sitting the examination for the first time and using past papers as a revision aid and also those who have previously failed the subject.

The Examiners are charged by Council with examining the published syllabus. The Examiners have access to the Core Reading, which is designed to interpret the syllabus, and will generally base questions around it but are not required to examine the content of Core Reading specifically or exclusively.

For numerical questions the Examiners' preferred approach to the solution is reproduced in this report; other valid approaches are given appropriate credit. For essay-style questions, particularly the open-ended questions in the later subjects, the report may contain more points than the Examiners will expect from a solution that scores full marks.

The report is written based on the legislative and regulatory context pertaining to the date that the examination was set. Candidates should take into account the possibility that circumstances may have changed if using these reports for revision.

Mike Hammer
Chair of the Board of Examiners
December 2020

A. General comments on the aims of this subject and how it is marked

1. The aim of the Risk Modelling and Survival Analysis Core Principles subject is to provide a grounding in mathematical and statistical modelling techniques that are of particular relevance to actuarial work, including stochastic processes and survival models.
2. Some of the questions in this paper admit alternative solutions from those presented in this report, or different ways in which the provided answer can be determined. All mathematically correct and valid alternative solutions received credit as appropriate.
3. In cases where the same error was carried forward to later parts of the answer, candidates were given full credit for the later parts.
4. In higher order skills questions, where comments were required, well-reasoned comments that differed from those provided in the solutions also received credit as appropriate.
5. As a change to previous practice, this diet of the exam was delivered as an open-book, online examination where candidates submitted their typed solutions in Word documents. As such, a number of new marking principles were established and communicated to candidates both prior to the exam taking place and in the candidate instructions provided with the exam paper. Candidates are advised to take careful note of all instructions that are provided with the exam in order to maximise their performance in future CS2A examinations. The instructions applicable to this diet can be found at the beginning of the solutions contained within this document.

B. Comments on candidates’ performance in this diet of the examination.

1. Performance was generally satisfactory, with most candidates demonstrating a reasonable understanding and application of core topics in mathematical and statistical modelling techniques.
2. Topics that were not particularly well-answered in this paper include Copulas (e.g. Q1), Extreme Value Theory (e.g. Q2) and Informative/Non-Informative Censoring (e.g. Q4(i)), despite these being examined through reasonably straightforward application questions. Candidates are reminded that it is very important to be familiar with all aspects of the syllabus.
3. It is important that candidates heed all of the instructions provided with the examination paper. A number of candidates lost marks because they did not include workings for numerical questions despite being forewarned about this in the instructions.
4. Higher order skills questions were generally answered poorly. Candidates should recognise that these are generally the questions which differentiate those candidates with a good grasp and understanding of the subject.

5. The comments that follow the questions in the marking schedule below, concentrate on areas where candidates could have improved their performance. Candidates approaching the subject for the first time are advised to concentrate their revision in these areas.

C. Pass Mark

The combined Pass Mark for the CS2 exam was 56.
1,363 presented themselves and 476 passed.

Solutions for Subject CS2 Paper A September 2020

Please note that the following principles apply to the CS2A solutions. These principles were set out in the instructions provided to candidates along with the examination paper:

1. Candidates **DO NOT** need to include any workings for multiple-choice questions. One, and only one, answer option should be included in the Word document for each multiple-choice question. Partial marks will **NOT** be awarded for candidates' workings.
2. Candidates **MUST** include workings for all numerical questions that are not multiple-choice. A correct answer will **NOT** score full marks unless workings are also shown.
3. Candidates should type their workings and answers into the Word document using standard keyboard typing. Candidates **DO NOT** need to use notation that requires specialised equation editing e.g. the “Equation Editor” functionality in Word.
4. Candidates **MUST** type all their workings and answers into the Word document. Calculations pasted in from another application (e.g. Excel or R) will **NOT** be accepted.

1

- (i) When $\rho = 0$, $C(u_1, u_2) = u_1 * u_2$

Therefore: $C(1, 1) = 1 * 1 = 1$. [1]

- (ii) $C(1, 0.2) = 1 * 0.2 = 0.2$. [1]

- (iii) $C(0.2, 0.2) = 0.2 * 0.2 = 0.04$. [1]

- (iv) When $\rho = 1$, $C(u_1, u_2) = \min(u_1, u_2)$.

Therefore: $C(1, 1) = \min(1, 1) = 1$ and $C(1, 0.2) = \min(1, 0.2) = 0.2$.

Hence (i) and (ii) will remain the same. [1]

However, (iii) will be different; $C(0.2, 0.2) = \min(0.2, 0.2) = 0.2$. [1]

- (v) The Gaussian copula has zero upper tail dependence (for $\rho < 1$)... [½]

...whereas the Gumbel copula has positive upper tail dependence (for $\alpha > 1$). [½]

Given that the analysis suggests that the Gumbel copula is a better fit... [½]

...by using the Gaussian copula, the insurer is likely to be underestimating the probability/risk of multiple extreme claims occurring simultaneously/in tandem. [½]

The insurer’s solvency could be threatened if this occurs in practice and the insurer has not made adequate provision for such events. [1]

[Total 8]

This question was one of the least well-answered in the entire paper. Candidates are advised to be familiar with all aspects of the syllabus.

Parts (i) to (iii) were poorly answered with many candidates unsuccessfully attempting to use the integral form of the bivariate Gaussian copula to derive their answers.

Part (iv) was very poorly answered. Some candidates lost marks because they correctly stated that the co-monotonic copula applied here but did not use its formula to derive the new values of the copula functions.

In parts (i) to (iv), a number of candidates additionally lost marks for not showing their workings.

Part (v) was also very poorly answered. Many candidates only discussed one of the two copulas which restricted the number of marks they could be awarded. Candidates are reminded to take note of the command verbs used in higher order skills questions to maximise their performance in future exams.

2

- (i) By fitting a distribution across the whole data range, the single distribution chosen may be a good overall fit of the data but could be a poor fit where there is little data, e.g. in the tails which are of primary concern. [1]

EVT can be useful where we are particularly interested in the tail of a distribution and need to model that part accurately. [1]

- (ii) $P(X > 70 \text{ given } X > 50) = 1 - G(20)$, where $20 = 70 - 50$. [1]

EITHER:

Using the GPD CDF; $G(20) = 1 - (1 + 20/(15*3))^{-3} = 0.668184$

Therefore $P(X - 50 \leq 20 \text{ given } X > 50) = 0.668184$.

So, $P(X > 70 \text{ given } X > 50) = 1 - 0.668184 = 0.331816$.

OR:

$$\begin{aligned} P(X > 70 \text{ given } X > 50) &= [(gamma * beta) / (gamma * beta + 20)]^{\gamma} \\ &= [(3 * 15) / (3 * 15 + 20)]^3 \\ &= 0.331816. \end{aligned}$$

[2]

The sports scientist has selected $(150 / 3,000) = 5\%$ of the data. [½]

Given that only 5% exceed the threshold;

$$\begin{aligned} P(X > 70) &= P(X > 70 \text{ given } X > 50) * P(X > 50) \\ &= 0.332 * 0.05 = 0.016591 = 1.6591\% \end{aligned}$$

[½]

- (iii) There are a number of limitations with this analysis:

Not all throws are independent.

OR:

An example of a source of non-independence, e.g. each thrower will make multiple throws.

[1]

Not all throws are identically distributed.

OR:

An example of a source of non-identical distribution, e.g. changing weather conditions, different abilities of throwers.

[1]

There could be different throwers next year, compared to the year analysed. [1]

There could be trends in the distances thrown over the years (e.g. improvements in training techniques, improvements in javelin technology (e.g. lighter javelins)).

[1]

Changes to rules and regulations might influence the distances thrown. [1]

Alternative thresholds should be analysed. [1]

The sample size is not particularly large. [1]

The generalized Pareto distribution is a limiting distribution and the actual distribution of the exceedances over any finite threshold will be different. [1]

[Marks available 8, maximum 4]

[Total 10]

Overall, this question was the least well-answered question in the whole paper. Candidates are advised to be familiar with all aspects of the syllabus.

Part (i) was well-answered, although in a number of cases, candidates’ answers were often vague and lacking detail despite this being a reasonably straightforward knowledge question.

Part (ii) was very poorly answered. The two most common mistakes were for candidates to use 70 in the GPD CDF rather than 20 and also for candidates to forget to calculate the unconditional probability of $X > 70$.

Part (iii) was also very poorly answered with most candidates focusing their comments solely on the limitations concerning the threshold used rather than expanding out their comments to the wider limitations. Alternative comments that were clear, distinct and relevant to the context of the question were also awarded credit.

3

(i) We have:

$$(1 - 1.5B + 0.5B^2)Y_t = (1 + 0.7B)e_t \quad [1]$$

so that:

$$(1 - B)(1 - 0.5B)Y_t = (1 + 0.7B)e_t \quad [1]$$

(ii) ARIMA(1, 1, 1), i.e. p = 1, d = 1, q = 1 [1]

(iii) Answer: A. [2]

$$\nabla Y_t = 0.5 \nabla Y_{t-1} + e_t + 0.7 e_{t-1}$$

$$\text{Cov}(e_t, \nabla Y_t) = 0.5 \text{Cov}(e_t, \nabla Y_{t-1}) + \text{Cov}(e_t, e_t) + 0.7 \text{Cov}(e_t, e_{t-1}) = \sigma^2$$

since e_t is independent of both e_{t-1} and ∇Y_{t-1} .

$$\text{Cov}(e_{t-1}, \nabla Y_t) = 0.5 \text{Cov}(e_{t-1}, \nabla Y_{t-1}) + \text{Cov}(e_{t-1}, e_t) + 0.7 \text{Cov}(e_{t-1}, e_{t-1})$$

$$= (0.5 + 0.7)\sigma^2$$

$$= 1.2\sigma^2$$

(iv) Answer: A. [3]

$$\text{Cov}(\nabla Y_t, \nabla Y_t) = 0.5 \text{Cov}(\nabla Y_t, \nabla Y_{t-1}) + \text{Cov}(\nabla Y_t, e_t) + 0.7 \text{Cov}(\nabla Y_t, e_{t-1})$$

$$\gamma_0 = 0.5\gamma_1 + \sigma^2 + 0.7(1.2\sigma^2)$$

$$\gamma_0 = 0.5\gamma_1 + 1.84\sigma^2$$

$$\text{Cov}(\nabla Y_{t-1}, \nabla Y_t) = 0.5 \text{Cov}(\nabla Y_{t-1}, \nabla Y_{t-1}) + \text{Cov}(\nabla Y_{t-1}, e_t) + 0.7 \text{Cov}(\nabla Y_{t-1}, e_{t-1})$$

$$\gamma_1 = 0.5\gamma_0 + 0.7\sigma^2$$

For $k > 1$:

$$\text{Cov}(\nabla Y_{t-k}, \nabla Y_t) = 0.5 \text{Cov}(\nabla Y_{t-k}, \nabla Y_{t-1}) + \text{Cov}(\nabla Y_{t-k}, e_t) + 0.7 \text{Cov}(\nabla Y_{t-k}, e_{t-1})$$

$$\gamma_k = 0.5\gamma_{k-1}$$

(v) Answer: A.

[2]

Using the equations from part (iv):

$$\gamma_0 = 0.5(0.5\gamma_0 + 0.7\sigma^2) + 1.84\sigma^2$$

$$\gamma_0 = 0.25\gamma_0 + 2.19\sigma^2$$

$$\gamma_0 = 2.92\sigma^2$$

Therefore, the variance of the new process, $\nabla Y_t = 2.92\sigma^2$

From above, $\gamma_1 = 0.5\gamma_0 + 0.7\sigma^2$

$$\gamma_1 = 0.5(2.92\sigma^2) + 0.7\sigma^2$$

$$\gamma_1 = 2.16\sigma^2$$

and, $\gamma_k = 0.5\gamma_{k-1}, k > 1$

$$\gamma_k = 0.5^{k-1} \gamma_1, k > 1$$

$$\gamma_k = 0.5^{k-1}(2.16\sigma^2), k > 1$$

[Total 10]

Overall, this was one of the best answered questions in the entire paper.

Answers to part (i) were generally mixed with many candidates only providing the first line of the solution.

Part (ii) was well-answered, although quite a number of candidates suggested that $p = 2$.

Parts (iii) to (v) were very well-answered; part (iii) in particular.

4

- (i) If we take the view that zebras close to death are unlikely to be able to escape then we could say that the two escapees on 1 Dec 2018 are likely to have been in better health than the zebras that remained in the wildlife park and hence informative censoring is present.

OR:

If we take the view that zebras close to death are more likely to behave erratically and escape then we could say that the two escapees on 1 Dec 2018 are likely to have been in poorer health than the zebras that remained in the wildlife park and hence informative censoring is present.

OR:

If we take the view that all the zebras are equally likely to escape then we could say that non-informative censoring is present in the two escapees on 1 Dec 2018.

[1]

If we take the view that weaker zebras are more likely to fall prey to predators then we could say that the zebra that was killed by a lion on 1 Jul 2018 is likely to have been in poorer health than the others in the wildlife park and hence informative censoring is present.

OR:

If we take the view that predators are more likely to kill healthy-looking zebras then we could say that the zebra that was killed by a lion on 1 Jul 2018 is likely to have been in better health than the others in the wildlife park and hence informative censoring is present.

OR:

If we take the view that all the zebras are equally likely to be killed by a lion then we could say that non-informative censoring is present in the zebra killed by a lion on 1 Jul 2018.

[1]

If we assume that the end date of the investigation was set without reference to the health of the zebras at that time, then non-informative censoring is likely to be present in the zebras that were right censored at the end of the investigation.

OR:

If the end date of the investigation was not known in advance and was set because the remaining two zebras were close to death at that time, then informative censoring is likely to be present in the zebras that were right censored at the end of the investigation. [1]

(ii)

Time	Reason for leaving	Reference tag
1	Death from rabies	19
1	Death from rabies	20
1	Censored	21
1	Censored	25
2	Death from rabies	7
2	Death from rabies	10
2	Death from rabies	11
3	Death from rabies	12
3	Censored	4
5	Death from rabies	1
5	Censored	9
5	Censored	13
8	Death from rabies	8
11	Death from rabies	3

[2]

j	t_j	n_j	d_j	$\lambda_j = d_j / n_j$	$1-\lambda_j$	Product over all $i \leq j (1-\lambda_i)$
1	1	14	2	1/7	6/7	0.8571
2	2	10	3	3/10	7/10	0.6000
3	3	7	1	1/7	6/7	0.5143
4	5	5	1	1/5	4/5	0.4114
5	8	2	1	½	½	0.2057
6	11	1	1	1	0	0

[½]

[1]

[½]

[½]

[½]

[1]

Time period	$S(t)$	$S(t)$
$0 \leq t < 1$	1.0000	1
$1 \leq t < 2$	0.8571	6/7
$2 \leq t < 3$	0.6000	3/5
$3 \leq t < 5$	0.5143	18/35
$5 \leq t < 8$	0.4114	72/175
$8 \leq t < 11$	0.2057	36/175
$11 \leq t$	0	0

[2]

[Total 11]

Part (i) was very poorly answered with many candidates struggling to adequately display a clear understanding of informative and non-informative censoring. The two most common mistakes were for candidates to definitively state that the different types of censoring were non-informative censoring without providing any justification and also for candidates to provide a definition of informative / non-informative censoring without applying it to the different types of censoring specified in the question. Candidates are reminded to take note of the command verbs used in higher order skills questions to maximise their performance in future exams.

Part (ii) was well-answered. The two most common mistakes were for candidates to exclude the two zebras, censored at the end of the investigation, from the analysis and for candidates to limit the final value of the survival function to $t < 12$. A number of candidates also lost marks for not showing their workings.

5

- (i) Null hypothesis: There is no bias in the graduated rates (or the standardised deviations).

OR:

The graduated rates are the true rates underlying the observed data. [½]

$$z_x = (\text{Observed Deaths} - \text{Expected Deaths}) / (\text{sqrt}(\text{Expected Deaths}))$$

Age x	Expected Deaths	z_x
50	85.77344	-1.70314
51	73.20764	-1.77739
52	66.13374	-1.49205
53	84.74964	0.57032
54	61.50300	1.08347
55	103.07952	-0.69730
56	123.81516	-0.88209
57	104.45890	-1.80606
58	117.06240	-1.39215
59	102.49990	-1.62974
60	91.87000	0.22222

[1½]

Under the null hypothesis, the standardised deviations are distributed as Binomial (11, 0.5). [½]

There are eleven deviations in total of which three are positive. [½]

According to pages 187 and 188 of the Golden Book, the likelihood of getting three or fewer positive deviations is between 0.0730 (when n = 12) and 0.1719 (when n = 10).

OR:

By explicit calculation, the likelihood of getting three or fewer positive deviations is 0.1133 [1]

Which exceeds 2.5% (two tailed test). [½]

So, there is insufficient evidence to reject the null hypothesis. [½]

- (ii) The signs test does not take into account the magnitude of the deviations. [1]

- (iii) Null hypothesis: The graduated rates are the true rates underlying the observed data. [½]

Age x	z_x	(z_x)^2
50	-1.70314	2.90068
51	-1.77739	3.15913
52	-1.49205	2.22621
53	0.57032	0.32527
54	1.08347	1.17391
55	-0.69730	0.48622
56	-0.88209	0.77807
57	-1.80606	3.26187
58	-1.39215	1.93808
59	-1.62974	2.65607
60	0.22222	0.04938

[1]

The test statistic is $X = \text{sum}((z_x)^2) = 18.95489$.

[½]

Under the null hypothesis, X has a chi-square distribution with m degrees of freedom, where m is the number of age groups less one for each parameter fitted.
So, in this case $m = 11 - 2 = 9$.

[1]

The critical value of the chi-square distribution with 9 degrees of freedom at the 5% level is 16.92.

[½]

$18.95489 > 16.92$; therefore, there IS enough evidence to reject the null hypothesis at 5% level.

[½]

- (iv) Although the signs test did not find systematic bias in the signs of the deviations, the chi-square goodness-of-fit test concluded that the magnitude of the deviations was too large under the null hypothesis.

[1]

Therefore, it is a good practice to always consider multiple statistical tests when assessing any hypothesis.

OR:

The chi-square goodness-of-fit test has therefore addressed a weakness of the signs test.

[1]

The deviations are large in magnitude in the age ranges of 50-52 and 57-59.

[½]

The significant result of the chi-square goodness of fit test suggests that alternative graduation methods should be considered.

[1]

**[Marks available 3½, maximum 2]
[Total 12]**

Parts (i), (ii) and (iii) were very well-answered.

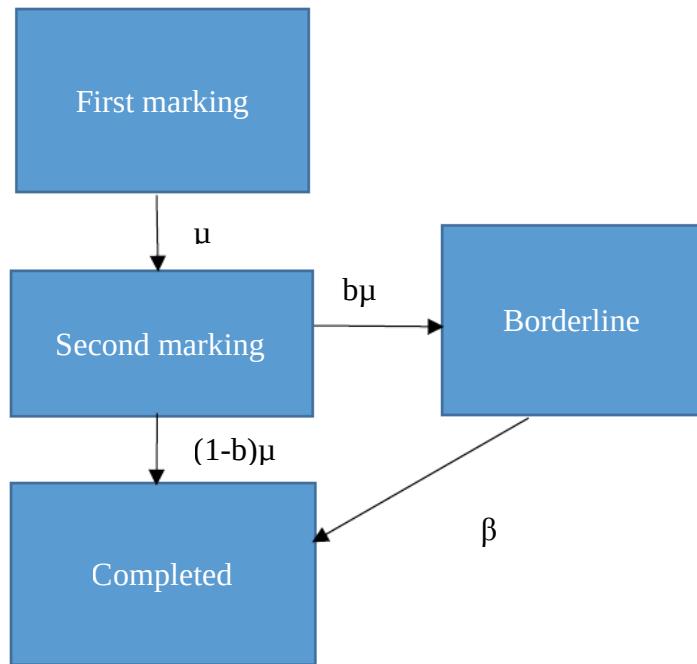
In part (i), most candidates recognised the need for a two-tailed test but one of the more common errors was for candidates to approximate the Binomial distribution with a Normal distribution. The number of age groups is not sufficiently large enough to allow this approximation.

In part (iii), most candidates correctly identified the number of degrees of freedom to use in the chi-square test although some candidates lost marks for not justifying their choice.

A number of candidates also lost marks for not showing their workings in parts (i) and (iii).

Part (iv) was poorly answered with most candidates focusing their comments solely on the contrasting conclusions of the two tests in parts (i) and (iii). Alternative comments that were clear, distinct and relevant to the context of the question were also awarded credit.

6



$$\frac{d}{dt} p^{SS}(t) = -\mu p^{SS}(t) \quad (1)$$

(i) Answer: **B** [2]

$$\frac{d}{dt} p^{SB}(t) = b\mu p^{SS}(t) - \beta p^{SB}(t) \quad (2)$$

(ii) Answer: **A** [4]

Solving equation (1) above and using the data given in the question we have:

$$p^{SS}(t) = \exp(-0.2t)$$

Then, in equation (2) we have:

$$\frac{d}{dt} p^{SB}(t) = (0.25)(0.2) \exp(-0.2t) - 0.3p^{SB}(t)$$

$$\frac{d}{dt} p^{SB}(t) = 0.05 \exp(-0.2t) - 0.3p^{SB}(t)$$

$$\frac{d}{dt} p^{SB}(t) + 0.3p^{SB}(t) = 0.05 \exp(-0.2t)$$

Using the integrating factor $\exp(0.3t)$ we have:

$$\frac{d}{dt} [\exp(0.3t)p^{SB}(t)] = 0.05 \exp(-0.2t) \exp(0.3t) = 0.05 \exp(0.1t).$$

Integrating produces:

$$\exp(0.3t)p^{SB}(t) = \frac{0.05}{0.1} \exp(0.1t) + k$$

Using the initial condition that $p^{SB}(0) = 0$ we have:

$$0 = \frac{0.05}{0.1} + k = 0.5 + k$$

Therefore $k = -0.5$.

$$\text{So } p^{SB}(t) = 0.5 \exp(-0.2t) - 0.5 \exp(-0.3t)$$

(iii) Answer: **B** [2]

For $t = 10$:

$$p^{SS}(t) = \exp(-0.2t) = 0.1353$$

$$\begin{aligned} p^{SB}(t) &= 0.5 \exp(-0.2t) - 0.5 \exp(-0.3t) \\ &= 0.5(0.1353) - 0.5(0.0498) \\ &= 0.0428. \end{aligned}$$

$$\text{Hence } p^{SC}(t) = 1 - 0.1353 - 0.0428 = 0.8219$$

(iv) Adding more examiners to the pool of first and second examiners would increase μ and hence speed up the time to complete first and second marking. [1]

However, μ might not increase in proportion with the number of first/second examiners if the new examiners mark more slowly because they are less experienced. [1]

However, introducing more examiners could increase the proportion b that are deemed to be “borderline” because of more variation in the marking. [1]

Therefore, any reduction in time taken, due to increasing μ , would be offset by any increase in time, due to proportion b increasing, and the overall reduction in time would depend on the relative changes in each. [1]

If more scripts are deemed to be “borderline”, then β might also reduce due to a “bottleneck” of scripts awaiting review by a third examiner. [1]

Additionally, the Board of Examiners could decide to add more examiners to the pool of third examiners (hence increasing β) if the reduction in time due to the Board member’s suggestion was not sufficient. [1]

The availability of examiners and overall cost of adding more to the pool would need to be considered by the Board before proceeding with these suggestions. [1]

The impact on the existing examiners (e.g. the effect of a reduction in workload) would also need to be considered before proceeding. [½]

[Marks available 7½, maximum 4]

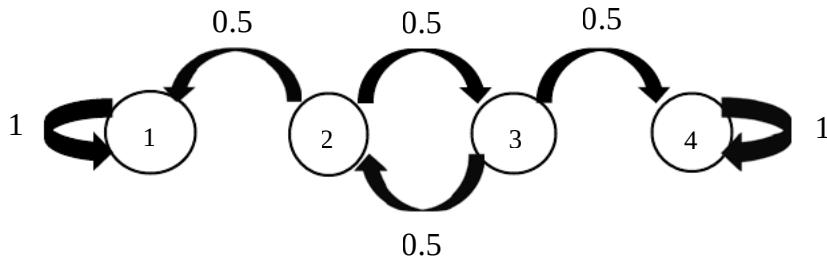
[Total 12]

Part (i) was very well-answered but parts (ii) and (iii) much less so.

Part (iv) was poorly answered with many candidates simply stating that the process would speed up. Candidates are reminded that higher order skills questions are generally the questions which differentiate those candidates with a good grasp and understanding of the subject. Alternative comments that were clear, distinct and relevant to the context of the question were also awarded credit.

7

(i)



A Markov Chain is irreducible if, given any pair of states i, j there exists an integer n with $p_{ij}(n) > 0$. [1]

For this Markov Chain there exist pairs of states (e.g. $i = 1$ and $j = 2$) where $p_{ij}(n) = 0$ for all n . [1]

Therefore, this Markov Chain is not irreducible. [1]

(ii) For π_i to be a stationary probability distribution of P :

$$\pi_{i_1} = \pi_{i_1} + 0.5 * \pi_{i_2}$$

$$\pi_{i_2} = 0.5 * \pi_{i_3}$$

$$\pi_{i_3} = 0.5 * \pi_{i_2}$$

$$\pi_{i_4} = 0.5 * \pi_{i_3} + \pi_{i_4} \quad [1]$$

Therefore $\pi_{i_2} = \pi_{i_3} = 0$ [½]

and $\pi_{i_1} + \pi_{i_4} = 1$ [½]

$\pi_i = \{x, 0, 0, 1 - x\}$ where $0 \leq x \leq 1$ (since $\pi_{i_1}, \pi_{i_4} \geq 0$) [1]

(iii) $h_{14} = 0$

$$h_{44} = 1$$

$$h_{24} = \frac{1}{2}h_{34} + \frac{1}{2}h_{14} = \frac{1}{2}h_{34}$$

$$h_{34} = \frac{1}{2}h_{24} + \frac{1}{2}h_{44} = \frac{1}{2}h_{24} + \frac{1}{2}$$

Therefore, solving:

$$h_{34} = \frac{1}{2}(\frac{1}{2}h_{34}) + \frac{1}{2}$$

$$h_{34} = 2/3$$

$$\begin{aligned} \text{So, } h_{24} &= \frac{1}{2}(2/3) \\ &= 1/3 \end{aligned} \quad [6]$$

Therefore: $h_{i4} = \{0, 1/3, 2/3, 1\}$

Alternative solution:

$$h_{14} = 0$$

$$h_{44} = 1$$

Starting from State 2, to reach State 4 the chain must move to State 3 and back to State 2 n times, where n is a non-negative integer, then move to State 3, then move to State 4. This has probability $0.5^{(2n+2)}$.

$$\text{Hence } h_{24} = 0.5^2 + 0.5^4 + 0.5^6 + \dots$$

This is the sum of a geometric progression with $a = r = 0.5^2 = 0.25$.

$$\text{Hence } h_{24} = 0.25 / (1 - 0.25) = 1/3.$$

A similar argument shows that $h_{34} = 2/3$. [6]

Therefore: $h_{i4} = \{0, 1/3, 2/3, 1\}$

[Total 12]

Part (i) was very well-answered but parts (ii) and (iii) less so.

In part (ii), few candidates derived the general form of the stationary probability distribution together with the condition on x. A more common response was for candidates to simply give one or two valid stationary distributions that satisfy the general form.

In part (iii), the majority of candidates who made a good attempt at this question, derived their solutions using the geometric progression approach. A number of candidates lost marks for not showing their workings in part (iii).

8

(i) Circuit boards made in “London” using the “Old” process. [1]

(ii) $\exp(\beta_L + \beta_P) / (1500(\exp(\beta_L + \beta_P) + \exp(\beta_L) + \exp(\beta_P) + 1))$ [2]

where:

β_L = regression coefficient for the Location covariate [½]

β_P = regression coefficient for the Process covariate [½]

(iii) The Oxford factory seems to produce circuit boards with (slightly) higher failure rates than the London factory. [1]

The New process seems to produce a (significant) reduction in failure rates when compared to the Old process. [1]

(iv) We need information on the standard errors for both coefficient estimates before we can reach a firm conclusion about the comments in part (iii) above. [1]

(v) Answer: **D** [3]

$$\text{Prob } (\text{No Fail} | \text{London \& Old}) = 1 - 0.2 = 0.8 = \exp [-\int \lambda_0(t) dt]$$

$$\text{Prob } (\text{No Fail} | \text{London \& New}) = \exp [-\int \lambda_0(t) \exp[\beta_P] dt]$$

$$= 0.8^{\exp[-0.30]}$$

$$= 0.8476$$

(vi) Answer: **B** [2]

There are 2,000 ($= 10,000 * 0.2$) failures expected with the old process

but only 1,524 ($= 10,000 * (1 - 0.8476)$) expected if the new process is used.

Hence there are 476 fewer expected failures in the first year if the new process is used.

[Total 12]

Part (i) was very well-answered.

Part (ii) was poorly answered with many candidates struggling to properly define the partial likelihood expression. Some candidates lost marks for not clearly defining the regression coefficients.

Part (iii) was well-answered although some candidates provided conclusions about the failure rates that were directionally opposite to the correct solutions.

Part (iv) was very poorly answered with many candidates suggesting that the baseline hazard function was the additional information required. Candidates who stated that log-likelihoods, test statistics, test results or p-values were required, provided that it was clear that they were referring to significance testing the regression coefficients, were awarded full marks.

Parts (v) and (vi) were reasonably well-answered.

9

- (i) For this to be the case, the customer's 2018, 2019 and 2020 renewals all have to be at the 10% discount level.

The probability of remaining in state with 10% discount is 0.4. [½]

Therefore, the required probability is $0.4^3 = 0.064$. [1½]

- (ii) This is the probability of being in states Three or Four or more in 2019.

We can obtain these probabilities from the matrix P^2 :

$$\begin{pmatrix} 0.5 & 0.2 & 0.2 & 0.1 \\ 0.2 & 0.4 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0 & 0.2 & 0.2 & 0.6 \end{pmatrix} \cdot \begin{pmatrix} 0.5 & 0.2 & 0.2 & 0.1 \\ 0.2 & 0.4 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0 & 0.2 & 0.2 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.33 & 0.24 & 0.26 & 0.17 \\ 0.24 & 0.28 & 0.3 & 0.18 \\ 0.22 & 0.24 & 0.3 & 0.24 \\ 0.08 & 0.24 & 0.26 & 0.42 \end{pmatrix}$$

Given that the customer starts in 10% discount level in 2017, the probability of a 15% discount in 2019 is 0.3 and a 25% discount in 2019 is 0.18.

OR:

Given that the customer starts in 10% discount level in 2017, the probability of a 15% discount in 2019 is:

$$0.2 * 0.2 + 0.4 * 0.3 + 0.3 * 0.4 + 0.1 * 0.2 = 0.3$$

Given that the customer starts in 10% discount level in 2017, the probability of a 25% discount in 2019 is:

$$0.2 * 0.1 + 0.4 * 0.1 + 0.3 * 0.2 + 0.1 * 0.6 = 0.18 \quad [1]$$

Required probability is 0.48. [1]

- (iii) Answer: A [3]

For this we require the long-term probabilities for the chain.

Stationary distribution π satisfies $\pi = \pi P$.

$$\frac{1}{2}\pi_1 + \frac{1}{5}\pi_2 + \frac{1}{5}\pi_3 = \pi_1 \quad (1)$$

$$\frac{1}{5}\pi_1 + \frac{2}{5}\pi_2 + \frac{1}{5}\pi_3 + \frac{1}{5}\pi_4 = \pi_2 \quad (2)$$

$$\frac{1}{5}\pi_1 + \frac{3}{10}\pi_2 + \frac{2}{5}\pi_3 + \frac{1}{5}\pi_4 = \pi_3 \quad (3)$$

$$\frac{1}{10}\pi_1 + \frac{1}{10}\pi_2 + \frac{1}{5}\pi_3 + \frac{3}{5}\pi_4 = \pi_4 \quad (4)$$

from (2)-(3)

$$\frac{1}{10}\pi_2 - \frac{1}{5}\pi_3 = \pi_2 - \pi_3$$

$$\pi_2 = \frac{8}{9}\pi_3$$

substitute in (1)

$$\frac{1}{2}\pi_1 = \frac{8}{45}\pi_3 + \frac{1}{5}\pi_3 = \frac{17}{45}\pi_3$$

$$\pi_1 = \frac{34}{45}\pi_3$$

substitute in (4)

$$\frac{2}{5}\pi_4 = \left(\frac{34}{450} + \frac{8}{90} + \frac{1}{5}\right)\pi_3$$

$$\pi_4 = \frac{41}{45}\pi_3$$

(iv) Answer: **C** [3]

$$\text{As } \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$$

$$\left(\frac{34}{45} + \frac{40}{45} + \frac{45}{45} + \frac{41}{45}\right)\pi_3 = 1$$

$$\pi_3 = \frac{9}{32} = 0.28125$$

Hence:

$$\pi_1 = \frac{17}{80} = 0.2125$$

$$\pi_2 = \frac{1}{4} = 0.25$$

$$\pi_4 = \frac{41}{160} = 0.25625$$

Hence the average discount is:

$$0.2125*5\% + 0.25*10\% + 0.28125*15\% + 0.25625*25\% = 14.19\%$$

- (v) The level of discount the insurance company gives in its home insurance is not completely within its control. [1]

The arrangement could lead to significant volumes of business for the insurance company through access to the bank’s customers, especially if the loyalty scheme is popular. [1]

Customers may open a number of small accounts in order to benefit from a higher discount. [1]

Bank customers with lots of accounts may be more loyal and this might result in better persistency on the home insurance, or even better claims experience. [1]

The insurer might need to raise its “full” prices to compensate for the introduction of the discount, and hence might become uncompetitive for those purchasing insurance outside the discount scheme. [1]

If the insurer does not raise its “full” prices to compensate for the introduction of the discount, it could find large numbers of poorer risks opening a number of bank accounts to take advantage of the larger discounts. This could lead to solvency issues for the insurer. [1]

[Marks available 6, maximum 3]
[Total 13]

Part (i) was poorly answered. The two most common mistakes were for candidates to mistakenly suggest that four years of renewals were required rather than three and/or for candidates to determine the second row and second column entry of the transition matrices, P^3 or P^4 .

Part (ii) was slightly better answered. Again some candidates mistakenly suggested that three years of renewals were required rather than two.

A number of candidates lost marks for not showing their workings in parts (i) and (ii).

Parts (iii) and (iv) were very well-answered.

Part (v) was very poorly answered with many candidates commenting on the commercial implications for the bank rather than the insurer. Candidates are reminded of the need to read the question carefully. Alternative comments that were clear, distinct and relevant to the context of the question were also awarded credit.

END OF EXAMINERS’ REPORT