INSTITUTE AND FACULTY OF ACTUARIES

EXAMINATION

17 April 2024 (am)

Subject CS2 – Risk Modelling and Survival Analysis Core Principles

Paper A

Time allowed: Three hours and twenty minutes

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

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

1 Suppose that the size of an insurance claim, X, has the following density function:

$$f(x) = \frac{1}{3x\sqrt{2\Pi}} e^{-z^2/2}$$

for $x > 0$ where $z = \frac{\ln(x) - 2.6}{3}$.

The insurance coverage pays claims subject to a deductible of £2,000 per claim.

(i) State the distribution of X and its parameters. [1]

(ii) Calculate the expected claim amount paid by the insurer per claim assuming that the claims occur at the end of the year with inflation of 20% p.a. [10]

[Total 11]

2 The following table shows mortality data and graduated mortality rates of male and female members of a pension scheme:

Gender	*Age (x years)*	*Exposed-to-risk (years)*	*Observed deaths*	*Graduated rates*
Female	70	1,007	153	0.14227
Female	71	978	166	0.16530
Female	72	1,111	204	0.19205
Female	73	1,500	326	0.22313
Female	74	1,200	306	0.25924
Female	75	2,001	599	0.30119
Male	70	908	338	0.36788
Male	71	998	439	0.42742
Male	72	1,009	497	0.49659
Male	73	877	508	0.57695
Male	74	859	571	0.67032

The data was graduated using the following parametric formula:

$$\log(\mu_x) = \alpha + \beta I_F + \delta x$$

where I_F is an indicator variable taking the value 1 for female members and 0 for male members.

(i) Calculate the values of α, β and δ, by considering the graduated rates at ages 70 and 71. [4]

(ii) Comment on your results. [1]

(iii) Perform a Chi-square test at a 5% significance level to assess the overall appropriateness of these graduated rates for both male and female members, stating your null and alternative hypotheses. [5]

(iv) State two limitations of the Chi-square test in this context and in each case suggest an alternative test that would address that limitation. [2]

[Total 12]

3 An analyst is trying to fit the following time series model:

$$y_t = a_2 y_{t-2} + a_3\, y_{t-3} + \varepsilon_t$$

(i) Show that the Yule–Walker equations for lags 1 and 2 are $\rho_1 = a_2\rho_1 + a_3\rho_2$ and $\rho_2 = a_2 + a_3\rho_1$. [2]

(ii) Derive expressions for the autocorrelation function for lags 1 and 2. [2]

(iii) Derive the values of the partial autocorrelation function, ϕ_k, for lags $k = 1, 2, 3, 4$ and 5, if $a_2 = 0.3$ and $a_3 = -0.2$. [6]

[Total 10]

4 An insurer has a portfolio of independent policies under which the number of claims follows a Poisson process with 300 claims expected per annum. Claim amounts are exponentially distributed with mean £3,000. Let S denote aggregate annual claims from the portfolio. A check is made for ruin only at the end of the year. The insurer includes a loading of 12% in the premiums for all policies. Expenses are ignored.

(i) Estimate the initial capital required, U, using a Normal approximation to the distribution of S, in order that the probability of ruin at the end of the first year is 5%. [6]

The insurer is considering purchasing proportional reinsurance from a reinsurer that includes a loading, β, in its reinsurance premiums. The proportion of each claim to be retained by the direct insurer is 80%.

Let S_1 denote the aggregate annual claims paid by the direct insurer net of reinsurance and U_1 denote the required initial capital after purchasing the reinsurance. The insurer uses a Normal approximation to the distribution of S_1.

(ii) Calculate the maximum reinsurance loading β for which U_1 is less than U given the same 5% ruin probability. [9]

[Total 15]

5 Apple trees growing in an orchard are classified by the farmer in one of three states:

F healthy trees with fruit
N unhealthy trees that give no fruit
D dead trees.

Every summer the farmer examines each tree and records its state. Year-to-year changes in tree classification are modelled using a Markov chain with the following 1-year transition probabilities:

	F	*N*	*D*
F	0.96	0.03	0.01
N	0.24	0.73	0.03
D	0	0	1

(i) Calculate the probability that a healthy tree will be rated unhealthy in 2 years. [2]

(ii) Calculate the percentage of healthy trees that are dead in 2 years. [2]

(iii) Calculate the probability that a healthy tree will never be rated unhealthy. [4]

(iv) Explain how this model could be adapted if the farmer wanted to take into account the age of each tree. [4]

[Total 12]

6 X and Y are random variables with X~N(0, 1) and Y~N(0, 1). For the purposes of parts (i) to (iv), the dependence between X and Y is modelled using a Gumbel copula with a copula parameter equal to 1.5.

(i) Calculate $P(X < -1.64, Y < -2.33)$. [3]

(ii) Calculate $P(X > 1.64, Y > 2.33)$. [3]

(iii) Calculate $P(-1.64 < X < 1.64, -2.33 < Y < 2.33)$. [3]

(iv) Comment briefly on the key properties of the Gumbel copula using your answers to parts (i) and (ii). [3]

[Total 12]

7 Let X_t be the zero-mean time series process defined by:

$$X_t = aX_{t-1} + be_{t-1} + X_{t-1}e_{t-1} + e_t$$

where e_t is a sequence of independent and identically distributed N(0, σ^2) random variables and a and b are positive constants with $a^2 + \sigma^2 < 1$.

(i) Demonstrate that $E(X_t e_t) = \sigma^2$. [3]

(ii) Demonstrate that $E(X_t e_{t-k}) = a^{k-1}((a+b)\sigma^2 + \sigma^4)$ for $k \geq 1$. [9]

By calculating further similar expectations it can be shown that the autocorrelation function, ρ, of X_t is of the form:

$$\rho_k = Aa^k + Bka^k$$

for $k \geq 1$, for suitable positive constants A and B.

Let Y_t be the zero-mean time series process defined by:

$$Y_t = aY_{t-1} + e_t$$

(iii) Explain the similarities and differences in shape between the autocorrelation functions of X_t and Y_t. [4]

[Total 16]

8 The following matrix is the generator matrix of a three-state continuous time Markov chain:

$$G = \begin{pmatrix} a & 2 & 1 \\ 2.5 & -3 & b \\ 0 & c & 0 \end{pmatrix}$$

(i) Calculate the values of a, b and c. [2]

(ii) Derive the probability that the chain will leave state 2 sometime before 0.5 time units, given that the chain is in state 2 at time 0. [4]

(iii) Derive the transition matrix of the corresponding Markov jump chain. [6]

[Total 12]

END OF PAPER

Institute and Faculty of Actuaries

EXAMINERS' REPORT

CS2 Risk Modelling and Survival Analysis

Paper A

Core Principles

April 2024

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.

For some candidates, this may be their first attempt at answering an examination using open books and online. The Examiners expect all candidates to have a good level of knowledge and understanding of the topics and therefore candidates should not be overly dependent on open book materials. In our experience, candidates that spend too long researching answers in their materials will not be successful either because of time management issues or because they do not properly answer the questions.

Many candidates rely on past exam papers and examiner reports. Great caution must be exercised in doing so because each exam question is unique. As with all professional examinations, it is insufficient to repeat points of principle, formula or other text book works. The examinations are designed to test "higher order" thinking including candidates' ability to apply their knowledge to the facts presented in detail, synthesise and analyse their findings, and present conclusions or advice. Successful candidates concentrate on answering the questions asked rather than repeating their knowledge without application.

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.

Sarah Hutchinson
Chair of the Board of Examiners
June 2024

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

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.

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.

In cases where an error was carried forward to later parts of the answer, candidates were given full credit for those later parts.

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.

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.

B. Comments on *candidate performance in this diet of the examination.*

The overall performance in this subject continues to improve although the average mark and the proportion of candidates passing is still below those the examiners hope for. In open book statistics and modelling assessments, the key skill sought by the examiners is the ability to apply knowledge of the statistical techniques in the syllabus to problems. Doing this successfully requires understanding of the principles and techniques covered in the Core Reading as well as practice in the application of those principles and techniques to a range of scenarios. In general, the problem questions in CS2 Paper A seek a well-structured response that draws upon understanding of stochastic processes, survival models, time series, risk models or statistical distributions. Candidates who are able to evidence both an understanding of the principles and a thought-through approach to the particular problem's scenario generally score the majority of marks available for a question even where they do not arrive at the correct final numerical answer.

The last part of many questions asks for comments on the results or on strengths and weaknesses of the approach being examined. Many candidates miss marks by being too brief in these questions. Comments that relate to the particular technique and ones that draw upon the problem scenario or earlier numerical answers can both be rewarded. Very often each successful point made is worth ½ mark so we would encourage candidates to look to make at least two commentary points (bullet points if you like) for every one mark offered for these questions.

C. Pass Mark

The Pass Mark for this exam was 55.
1286 presented themselves and 463 passed.

The expected Pass Mark was determined when the examination papers were written through a "bottom-up" summation of the marks deemed to indicate a minimally competent candidate. This mark was then confirmed during marking with both a quantitative and qualitative analysis of question performance.

Solutions for Subject CS2A – April 2024

Q1

(i)

Claim amount follows Lognormal distribution with [½]

parameters mu = 2.6 and sigma = 3. [½]

(ii)

Gross claim amount Y = 1.2X [½]

Claim amount for the insurer

Z = Max(0, Y 2000) [1]

Using Page 18 of golden book,

E(Z) = {1.20* exp(mu +.5* sigma ^2) * ([phi(U1) - [phi(L1)])} – {2000 * phi(U0) - [phi(L0)]) } [2]

U1 = ((log (inf) – mu) / sigma) - sigma [1]

L1 = ((log (2000/1.20) – mu) / sigma) – sigma [1]

U0 = ((log (inf) – mu) / sigma) [½]

L0 = ((log (2000/1.20) – mu) / sigma) [½]

Phi corresponds to standard normal distribution

phi(U1) = 1 = phi(U0)

phi(L1) = 0.08169

phi(L0) = 0.94588 [1]

E(Z) = 1454.36049 * (1 - 0.08169) - 2000* (1 – 0.94588) [1]

= 1335.5537 – 108.24 [1]

= £1227.314 [½]

[Total 11]

Commentary:
Part (i) of this question was well answered and part (ii) was generally poorly answered with an average score of 4 of the 10 marks for this second part. The most common errors resulted from a lack of clear structure for dealing with the both the deductible and inflation and then translation of that structure into an equation for the expected value.

Q2

(i)

δ is the slope of the graduated rate (on log scale) for each gender.

That is $\hat{\delta} = \frac{log(0.42742)-log(0.36788)}{71-70}$ [1]

= 0.15 [½]

α is the intercept of the graduated rate (on log scale) for each gender.

That is $\hat{\alpha} = log(0.36788) - \hat{\delta} \times 70$ [1]
$= -11.5$ [½]

β is the gap between the graduated rates (on log scale) for women and men.

That is $\hat{\beta}$ = log(0.14227) - log(0.36788) [½]
= -0.95 [½]

(ii)

The graduated rate increases with age for men and women since $\hat{\delta} > 0$. [½]

The graduated rate for men is higher than that for women since $\hat{\beta} < 0$. [½]

Note: Alternative valid comment receives ½ mark each up to a maximum of 1 mark.

(iii)

The null hypothesis is that the graduated rates for men and women are the true rates underlying the observed data, and the alternative hypothesis is that the graduated rates for men and women are not the true rates underlying the observed data. [½]

Note that both H0 and H1 needed

z_x = (Observed Deaths – Expected Deaths) / (sqrt (Expected Deaths)) [½]

Gender	Age	Expected deaths	z_x	(z_x)^2
F	70	143.26589	0.813251	0.661378
F	71	161.6634	0.34107	0.116329
F	72	213.36755	-0.6413	0.411267
F	73	334.695	-0.47528	0.225886
F	74	311.088	-0.28847	0.083217
F	75	602.68119	-0.14995	0.022485
M	70	334.03504	0.216942	0.047064
M	71	426.56516	0.602071	0.362489
M	72	501.05931	-0.18135	0.032886
M	73	505.98515	0.089572	0.008023
M	74	575.80488	-0.20024	0.040095

[1]

The test statistic is X = sum((z_x)^2) = 2.01 [1]

Under the null hypothesis, X has a chi-square distribution with m degrees of freedom, where m is the number of groups less one for each parameter fitted.

So, in this case m = 11 - 3 = 8. [1]

The critical value of the chi-square distribution with 8 degrees of freedom at the 5% level is 15.51. [½]

Since 2.01 < 15.51, we have no reason to reject the null hypothesis. [½]

(iv)

There could be a few large deviations offset by very many small deviations. [½]

This can be addressed via standardised deviations test. [½]

The graduation might be biased above or below the data by a small amount. [½]

This can be addressed via signs test. [½]

Note: Alternative valid limitations are accepted: ½ mark for stating a limitation + ½ mark for suggesting an alternative test that would address it.

[Total 12]

Commentary:
This question was generally well answered. This continues the trend in recent sittings of survival analysis type questions being well answered. The answer given for part (ii) above are perhaps the most obvious and common but other valid comments on the interpretation of the three parameters were given marks. In part (iii) both the null and alternative hypotheses were required and candidates are reminded of the need to set out such tests in full including the test statistic, degrees of freedom and a one sentence conclusion in addition to the numerical results.

Q3

(i)

$cov(y_t, y_{t-1}) = a_2 cov(y_{t-2}, y_{t-1}) + a_3 cov(y_{t-3}, y_{t-1}) + 0$ [½]

i.e. $\rho_1 = a_2\rho_1 + a_3\rho_2$ [½]

$cov(y_t, y_{t-2}) = a_2 cov(y_{t-2}, y_{t-2}) + a_3 cov(y_{t-3}, y_{t-2}) + 0$ [½]

i.e.

$\rho_2 = a_2 + a_3\rho_1$ [½]

(ii)

$\rho_1 = a_2\rho_1 + a_3\rho_2 = a_2\rho_1 + a_3(a_2 + a_3\rho_1)$ [1]

Therefore

$$\rho_1 = \frac{a_2 a_3}{1 - a_2 - a_3^2}$$ [½]

and

$$\rho_2 = a_2 + \frac{a_2 a_3^2}{1 - a_2 - a_3^2}$$ [½]

(iii)

As $a_2 = 0.3$ and $a_3 = -0.2$, it now follows that

$\rho_1 = \frac{-0.3*0.2}{1-0.3-0.2^2}$ =-0.091 [1]

and

ρ_2=0.3+$\frac{0.3*0.2^2}{1-0.3-0.2^2_3}$=0.3182 [1]

For the PACF:

[1]

from the tables $\phi_1 = \rho_1 = \frac{-0.3*0.2}{1-0.3-0.2^2}$ =-0.091

and $\phi_2 = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = 0.3125$ [1]

$\phi_3 = -0.2$ from the definition [1]

And $\phi_4 = \phi_5 = 0$ since the process is AR(3) [1]

[Total 10]

Commentary:
This question was reasonably well answered and certainly better than many of the time series questions found in recent years. The question is a relatively simple application of autoregressive models.

Q4

(i)

N follows Poisson process (300) , X follows exponential distribution

S follows a Compound distribution

E(S) = E(N) * E(X) = 300*3000 = 900,000 [1]

Var(S) = E(N) * var(X) + var(N) * E(X)^2 [½]

= 300*[3000^2 *2] = 5.4* 10^9 [½]

Let 'p' denote premium to be received , p = 1.12* E(S) [½]

Prob(u + p < S) = .05 [1]

Prob(u + p > S) = .95

Prob(Z< (u+ p- E(S)) / sd(S)) = 0.95 [½]

(u+ p- E(S)) / sd(S) = 1.64 [½]

u = 1.64* sqrt(5.4* 10^9) – 1.12*E(S) + E(S) [1]

u = £12,515 [½]

(ii)

Let 'p1' denote premiums to be received after reinsurance.

Proportion retained = alpha = 80%

p1 = 1.12* E(S) - (1- alpha) *(1+ beta) *E(S) [1]

E(S1) = alpha *E(S) [½]

Var(S1) = alpha ^ 2* Var (S) [½]

Prob(u1 + p1 < S1) = .05 [1]

Prob(u1 + p1 > S1) = .95

Prob(Z< (u1+ p1- E(S1)) / sd(S1)) = .95 [1]

(u1+ p1- E(S1)) / sd(S1) = 1.64 [½]

Substituting the values of p1 and E(S1) and Var(S1) , we get

(u1+ 1.12* E(S) - (1- alpha) *(1+ beta) *E(S) – alpha * E(S)) / (alpha * sd(S)) = 1.64 [1]
Rearranging

u1 = 1.64* alpha * sd(S) + (1- alpha) *(1+ beta) *E(S) + alpha * E(S) -1.12* E(S) [1]

u1 = 96411.92 + 180000*(1+ beta) + 720000 – 1008000 [½]

Therefore,
u1 – u <0 implies [½]
96411.92 + 180000*(1+ beta) + 720000– 1008000- 12515 <0 [½]

beta < 0.134 [½]
Therefore, reinsurance loading should not exceed 13.4%. [½]

[Total 15]

Commentary:
This question was poorly answered with the average response scoring just 5 of the 15 marks available. Compound distributions have regularly been examined in CS2 although often in shorter questions at the beginning of the paper rather than a longer question such as this. This question is an excellent example of the importance of a thought-through structure for answering. The presence of reinsurance terms, premium loading and ruin probability need to be translated into probability statements. Once this has been done and the expected value and variance of the compound distributions calculated, proceeding to the final answer is relatively straightforward. Candidates who used 1.6449 rather than 1.64 for the Normal distribution in part (i) arrived at a final answer of £12,875 and scored full marks.

Q5
(i)
We first need the two year transition matrix

$$\begin{pmatrix} 0.96 & 0.03 & 0.01 \\ 0.24 & 0.73 & 0.03 \\ 0 & 0 & 1 \end{pmatrix} \%*\% \begin{pmatrix} 0.96 & 0.03 & 0.01 \\ 0.24 & 0.73 & 0.03 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.9288 & 0.0507 & 0.0205 \\ 0.4056 & 0.5401 & 0.0543 \\ 0 & 0 & 1 \end{pmatrix}$$

so probability = 0.0507

(ii)
From the matrix in (i) percentage = 2.05%
Alternative solution to (i) and (ii) without calculating the two year matrix but direct from the probabilities :
In (i) Required probability = P_FF * P_FN + P_FN * P_NN + P_FD * P_DN [1]
= 0.96 * 0.03 + 0.03 * 0.73 + 0 [½]
= 0.0507 [½]
Note: This probability can also be obtained from the square of the one-year transition matrix.
In (ii) Required percentage = P_FF * P_FD + P_FN * P_ND + P_FD * P_DD [1]
= 0.96 * 0.01 + 0.03 * 0.03 + 0.01 * 1 [½]
= 2.05% [½]

(iii)
For tree to never be unhealthy it must eventually go straight from healthy to dead
Pr(this happens in year1) = 0.01 [½]
Pr(this happens in year2) = 0.96 * 0.01

Pr(this happens in year 3) = 0.96^2 * 0.01

…

Pr(this happens in year n) = 0.96^(n-1) * 0.01 [1]

The required probability is the sum of the above probabilities over all n [1]

$$\frac{0.01}{1-0.96}$$ [1]

= 0.25 [½]

(iv)
For an age-related model:
we need to move from a time homogeneous to in-homogeneous model [1]
e.g. through a multi-state model with age related transition intensities [1]
that could be estimated through MLE with transition and waiting time data [1]
which would need more than annual survey data to do accurately [1]
with annual data some census approximation to the waiting time is needed [1]
example of how this might look related to age of tree [1]

[Marks Available 6, Maximum 4 (must include 1st point)]

[Total 12]

Commentary:
The first three calculation parts of this question were well answered. In parts (i) and (ii) the answers could be obtained via matrix multiplication or direct from probabilities and full marks were available for either route. Part (iv) was poorly answered. A very wide range of comments attracted marks but many candidates wrote very little for this part.

Q6
(i)
$F_N(-1.64)=0.05$; $F_N(-2.33)=0.01$ [1]
C(0.05,0.01) = Exp[-{(- ln 0.05) ^1.5 + (-ln 0.01) ^1.5}^0.6667] [1]
= 0.00224 [1]
(ii)
The required probability is 1 - C(1,0.99) - C(0.95,1) + C(0.95,0.99) [1]
We have C(1,0.99) = 0.99, C(0.95,1) = 0.95 [½]
and C(0.95,0.99) = 0.94723 by calculation [½]
The required probability is therefore 1-.99-.95+ 0.94723 [½]
= 0.00723 [½]
(iii)
The required probability is C(0.95,0.99) - C(0.05,0.99)- C(0.95,0.01) + C(0.05,0.01) [1]

From part (i), C(0.05,0.01) = 0.00224
From part (ii), C(0.95,0.99) = 0.94723
By calculation, C(0.05,0.99) = 0.04998 [½]
By calculation, C(0.95,0.01) = 0.00996 [½]
The required probability is therefore .94723 - .04998 - .00996+ .00224 [½]
=0.88953 [½]
(iv)
The Gumbel copula has zero lower tail dependence [1]
but positive upper tail dependence [1]
This is consistent with the probability in part (ii) being higher than in part (i) [1]

[Total 12]

Commentary:
The first two parts of this question were well answered but the last two parts were generally less well answered. For the CS2 syllabus area on risk distributions the examiners have noted in recent years that Copula and Extreme Value Theory questions have been less well answered and candidates are reminded of the importance of covering these areas as well as the earlier parts of this syllabus section.

Q7
(i)
E(X_t * e_t) = E((a * X_t-1 + b * e_t-1 + X_t-1 * e_t-1) * e_t) + E(e_t ^ 2) [½]
= E(a * X_t-1 + b * e_t-1 + X_t-1 * e_t-1) * E(e_t) + E(e_t ^ 2) [½]
since a * X_t-1 + b * e_t-1 + X_t-1 * e_t-1 and e_t are independent [½]
= 0 + var(e_t) [½]
since E(e_t) = 0 [½]
= sigma ^ 2 [½]
(ii)
E(X_t * e_t-1)
= a * E(X_t-1 * e_t-1) + b * E(e_t-1 ^ 2) + E(X_t-1 * e_t-1^2) + E(e_t * e_t-1) [1]
Where
E(X_t-1 * e_t-1) = E(X_t * e_t) = sigma ^ 2 [½]
E(e_t-1 ^ 2) = E(e_t ^ 2) = sigma ^ 2 [½]
E(X_t-1 * e_t-1^2) = E((a * X_t-2 + b * e_t-2 + X_t-2 * e_t-2 + e_t-1) * e_t-1^2) [½]
= E(a * X_t-2 + b * e_t-2 + X_t-2 * e_t-2) * E(e_t-1^2) + E(e_t-1^3) [½]
= a * E(X_t-2) * E(e_t-1^2) + b * E(e_t-2) * E(e_t-1^2) + E(X_t-2 * e_t-2) * E(e_t-1^2) + E(e_t-1^3) [½]
= 0 + 0 + E(X_t-2 * e_t-2) * E(e_t-1^2) + 0 [½]
since E(X_t-2) = 0 [½]
= sigma ^ 4 [½]
E(e_t * e_t-1) = E(e_t) * E(e_t-1) = 0 [½]
Hence E(X_t * e_t-1) = (a + b) * sigma ^ 2 + sigma ^ 4
For k > 1, E(X_t * e_t-k)

= *a* * *E*(*X_t*-1 * *e_t*-*k*) + *b* * *E*(*e_t*-1 * *e_t*-*k*) + *E*(*X_t*-1 * *e_t*-1* *e_t*-*k*) + *E*(*e_t* * *e_t*-*k*) [½]
where
E(*e_t*-1 * *e_t*-*k*) = *E*(*e_t*-1) * *E*(*e_t*-*k*) = 0 [½]
E(*X_t*-1 * *e_t*-1* *e_t*-*k*)
= *E*((*a* * *X_t*-2 + *b* * *e_t*-2 + *X_t*-2 * *e_t*-2 + *e_t*-1) * *e_t*-1* *e_t*-*k*) [½]
= *E*((*a* * *X_t*-2 + *b* * *e_t*-2 + *X_t*-2 * *e_t*-2) * *e_t*-*k*) * *E*(*e_t*-1) + *E*(*e_t*-1 ^ 2) * *E*(*e_t*-*k*) [½]
= 0 [½]
E(*e_t* * *e_t*-*k*) = *E*(*e_t*) * *E*(*e_t*-*k*) = 0 [½]
Hence *E*(*X_t* * *e_t*-*k*) = *a* * *E*(*X_t*-1 * *e_t*-*k*)
The result follows by induction [½]
(iii)
The autocorrelation function of *Y_t* is *rho_k* = *a* ^ *k* [1]
Both autocorrelation functions decay exponentially for large values of *k* [1]
since both *a* ^ *k* and *k* * *a* ^ *k* tend to zero as *k* tends to infinity [½]
The autocorrelation function of *X_t* is higher than implied by exponential decay for small values of *k* [½]
as a result of the term in *k* * *a* ^ *k* [½]
whereas the exponential decay of the autocorrelation function of *Y_t* starts immediately [½]

[Total 16]

Commentary:
This second, longer time series question was less well answered than the other, shorter time series question. The demonstration in part (i) was straightforward and well answered. In part (ii) the route to a solution of first considering k = 1 and then later k > 1 was not widely used. Part (iii) is another question that would have benefitted from more structured answers with consideration of Y and then of X.

Q8
(i)
From the definition such that sum_j(mu_ij) =0: [1]
a=-(2+1) =-3; b =0.5 and c =0 [1]
(ii)
The required probability is the probability that the holding time in state 2 is less than 0.5 [1½]
The holding time is exponentially distributed with parameter 3 [1]
The required probability is therefore 1 – exp(-3 * 0.5) [1]
= 0.7768698 [½]
(iii)
P_31 = P_32 = 0 and P_33 = 1 [½]
since state 3 is absorbing [½]
P_11 = P_22 = 0 [1]

since the next jump from state 1 must be to a state other than state 1, and similarly for state 2 [½]
The remaining entries of P are obtained as P_ij = mu_ij / (-mu_ii) [1]
This gives P_12 = 2/3, P_13 = 1/3, P_21 = 5/6, P_23 = 1/6 [1½]

Hence $P = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 5/6 & 0 & 1/6 \\ 0 & 0 & 1 \end{pmatrix}$ [1]

[Total 12]

Commentary:
This question was generally well answered. It is a relatively straightforward application of the Core Reading material on Markov chains.

[Paper Total 100]

END OF EXAMINERS' REPORT

Institute and Faculty of Actuaries

www.actuaries.org.uk