INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

September 2021

CS2 - Risk Modelling and Survival Analysis
Core Principles
Paper A

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.

Sarah Hutchinson
Chair of the Board of Examiners
December 2021

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 the same error was carried forward to later parts of the answer, candidates were given full credit for the 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. The instructions applicable to this diet can be found at the beginning of the solutions contained within this document.

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

Performance was generally satisfactory. Most candidates demonstrated a reasonable understanding and application of core topics in mathematical and statistical modelling techniques.

The most poorly answered question in this paper was Question 1, on Extreme Value Theory. Candidates are reminded to relate their answers to the specific situation in the question.

Question 7, on Time Series, and Question 8, on Markov Jump Processes, were also relatively poorly answered. Candidates are reminded that when they are unable to answer one part of a question, they may still gain credit in subsequent parts by assuming a "dummy" answer.

It is important that candidates follow 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.

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.

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 Pass Mark for this exam was 58
1,264 presented themselves and 440 passed.

Solutions for Subject CS2A - September 2021
Please note the following conventions / principles that apply to this marking schedule:

Candidates **MUST** include typed workings, in addition to their typed answers, in the Word document for all numerical questions. Candidates using another software package to aid with calculations **MUST** ensure that all calculations appear in full in the Word document to ensure that they receive full marks. If sufficient workings are not displayed full marks may not be awarded.

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.

Your Word document **MUST NOT** include links to any other documents.

Q1
Collect daily returns and group into months [½]
Take the maximum loss each month and remove all other data [½]
Find the parameters for the GEV distribution [½]
using maximum likelihood estimation [½]
Calculate 1 - H(0.05), where H(x) is the cumulative distribution function of the GEV distribution [1]
which gives the probability that the maximum daily loss that month will exceed 5%

This question was very poorly answered. Many candidates described the Generalised Extreme Value distribution and the block maxima method in general, without reference to the specific situation in the question.

Q2
(i)
The likelihood function is
L = C * ((mu_H_SNR)^(N_H_SNR))
* exp (- (mu_H_SR + mu_H_SNR + mu_H_D) * T_H) [1½]

where:
mu_i_j is the transition rate from State i to State j [½]
T_H is the total observed waiting time in State H [½]
N_H_SNR is the number of transitions from State H to State SNR [½]
C is a constant independent of mu_H_SNR

The log-likelihood function is

ln L = ln C + N_H_SNR * ln mu_H_SNR
- (mu_H_SR + mu_H_SNR + mu_H_D) * T_H [1]

Differentiating with respect to mu_H_SNR gives
d ln L / d mu_H_SNR = N_H_SNR / mu_H_SNR - T_H [½]

Setting the derivative to 0 gives
mu_H_SNR^hat = N_H_SNR / T_H [½]
We have a maximum, since [½]
d^2 ln L / d (mu_H_SNR)^2 = - N_H_SNR / (mu_H_SNR)^2 < 0 [½]

(ii)
The asymptotic distribution is
Normal (mu_H_SNR, mu_H_SNR / E(T_H)) [2]

[Total 8]

Part (i) was well answered.

Answers to part (ii) were generally satisfactory, although many candidates omitted the expectation sign in the denominator of the variance.

Q3
(i)
rho_0 = 1 [½]
rho_1 = beta / (1 + beta^2) [1]
rho_k = 0 for k > 1 [½]

(ii)
From page 40 of the Golden Book,
phi_2 = (rho_2 - (rho_1)^2) / (1 - (rho_1)^2) [1]
So,
-1/3 = (0 - (rho_1)^2) / (1 - (rho_1)^2) [½]
which gives
rho_1 = 1/2 or -1/2 [1]
So,
beta / (1 + beta^2) = 1/2 or -1/2 [½]
which gives
beta = 1 or -1 [1]

ALTERNATIVE SOLUTION:
From page 41 of the Golden Book,
phi_2 = - ((1 - beta^2) * beta^2) / (1 - beta^6) [1]
So,
-1/3 = - beta^2 / (1 + beta^2 + beta^4) [1]
Hence,
beta^4 - 2 * beta^2 + 1 = 0 [½]
So,

beta^2 = 1 [1]
which gives
beta = 1 or -1 [½]

(iii)
When beta = 1 or -1, Y_t is not invertible [½]
which means that the autoregressive representation of Y_t for both values of beta is not convergent [½]
As moving average models fitted to data by statistical packages are always invertible this time series process, with these values of beta, would never be used by these packages when fitting to observed data [1]
and therefore, not be suitable for practical fitting purposes [½]

[Marks available 2½, maximum 2]

[Total 8]

Part (i) was well answered.

Answers to part (ii) were generally satisfactory. Some candidates made errors in their workings resulting in the need to solve only one quadratic equation instead of two. These candidates were awarded partial marks for follow-through.

Part (iii) was the most poorly answered question part on the whole paper. Some candidates stated that Y_t is not stationary when beta = 1 or -1, which relates to an autoregressive process, not the moving average process in the question. Of those candidates who correctly recognised that Y_t is not invertible when beta = 1 or -1, very few recognised the implications of this for the practical suitability.

Q4

The null hypothesis is that the graduated rates are the true rates underlying the observed data [½]
The alternative hypothesis is that the graduated rates are NOT the true rates underlying the observed data [½]
z_x = (Observed Deaths - Expected Deaths) / (sqrt (Expected Deaths)) [½]

Age x	Expected Deaths	z_x	(z_x)^2
55	57.31278	-0.56968	0.32454
56	61.51770	-0.57599	0.33177
57	69.82764	1.93535	3.74558
58	76.50230	0.05690	0.00324
59	79.29862	-0.59502	0.35405
60	78.67215	1.95359	3.81653
61	78.57000	0.04851	0.00235
62	82.85552	-1.63203	2.66351
63	86.20300	-0.88351	0.78059
64	87.62580	-0.49416	0.24420
65			

[1½]

The test statistic is X = sum((z_x)^2) =12.26635 [1]
+ ((105 - 1,736 y)^2) / (1,736 y) [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 - 4 = 7 [1]
The critical value of the chi-square distribution with 7 degrees of freedom at the 2.5% level is 16.01 [½]
Therefore, we are looking for y such that X < 16.01 [½]
That is:
(105 - 1,736 y)^2 < 6498.98 y [½]
i.e.
3,013,696 y^2 - 371,058.98 y + 11,025 < 0 [½]
The roots of this quadratic function are 0.0501 and 0.0730 [1]
The range of values of y required so that there is insufficient evidence, at the 97.5% confidence level, to reject the null hypothesis that the graduated rates are the true rates underlying the observed data is 0.0501 < y < 0.0730 [½]

This question was fairly well answered. The most common errors were
Using the wrong number of degrees of freedom - in particular the force of mortality at age 65, d, must be treated as a parameter.
Determining the range of values of y such that the null hypothesis is rejected, rather than accepted as specified in the question.
Errors in solving the quadratic inequality.

Q5
(i)
Right censoring is present [½]
of patients still in hospital after 30 days, or of those who leave hospital, as observation is cut short. (We only know they will die at some time after the date of censoring) [½]
Type I censoring is present [½]
as it is predetermined that observation would cease after 30 days [½]
Random censoring is present [½]
as the times at which patients leave hospital can be considered a random variable [½]
Informative censoring may be present [½]
if those who leave hospital are in better health than those who remain [½]
Non-informative censoring may be present [½]
if the fact that some patients have left hospital tells nothing about the risk of death among those who remain [½]
[Marks available 5, maximum 2]

Non-informative censoring may not be present [½]
if those who leave hospital are in better health than those who remain [½]
Informative censoring may not be present [½]
if the fact that some patients have left hospital tells nothing about the risk of death among those who remain [½]

Type II censoring is not present [½]
as the study does not continue until a predetermined number of deaths [½]
Left censoring is not present [½]
as we know the date on which each patient had their operation [½]
[Marks available 4, maximum 1]

(ii)
0.8456 [1]

(iii)
The Nelson-Aalen estimate is given by exp(-LAMBDA_HAT(t)).
So LAMBDA_HAT(t) = -ln(S(t)) [1]
LAMBDA_HAT(t) = Sum (over t_j <= t) [d_j / n_j] [1]
K-M = Kaplan-Meier estimate = Product (over t_j <= t) [1 - d_j / n_j] [1]

t (days)	LAMBDA_HAT(t)	d_j / n_j	1 - d_j / n_j	K-M
0 <= t < 5	0.0000	0.0000	1.0000	1.0000
5 <= t < 17	0.1052	0.1052	0.8948	0.8948
17 <= t < 25	0.1677	0.0625	0.9375	0.8389
25 <= t	0.3345	0.1668	0.8332	0.6990
	[½]	[½]	[½]	[½]

[Total 9]

Part (i) was very well answered, except that some candidates lost marks because they only specified a group of lives affected by a particular type of censoring, without explaining why that type of censoring is present.

Part (ii) was the best answered question part on the whole paper.

Part (iii) was fairly well answered, although many candidates lost marks for not showing sufficient workings to make clear that a valid method had been used, or for not including the ranges of values of t in their answer script.

Q6
(i)
P(Claims >= 1000) = 1 - P(Claims < 1000) [½]
= 1 - P (Z < (1000 - 250) / 300) [½]
= 1 - P (Z < 2.5) [½]
= 1 - 0.99379 = 0.00621 (or 0.621%) [½]
(using page 161 of the Golden Book)

(ii)
For the mean we have
lambda_hat / (alpha_hat - 1) = 250 [½]
and for the variance we have
(alpha_hat * lambda_hat^2) / (((alpha_hat - 1)^2) * (alpha_hat - 2)) = 300^2 [½]
Substituting gives

(alpha_hat * 250^2) / (alpha_hat - 2) = 300^2 [½]
Hence, alpha_hat = 6.55 [½]
And, lambda_hat = 250 * (alpha_hat - 1) = 1,386.36 [½]
P(Claims >= 1000) = (lambda_hat / (lambda_hat + 1000))^alpha_hat [1]
= 0.0286 (or 2.86%) [½]

(iii)
The Normal distribution is unlikely to be a good fit for the total monthly claim amounts because negative claims can't be incurred by the company [½]
and the Normal distribution assigns a non-zero probability to negative claims occurring [½]
In particular, with the given mean and standard deviation, there is a significant probability (around 20%) of claims being negative [½]
The Normal distribution is also unlikely to be a good fit because the distribution of claims incurred by the company is likely to be positively skewed [½]
and the Normal distribution is symmetric/has zero skewness [½]
Additionally, the Normal distribution is thin-tailed [½]
and therefore, not suitable for modelling situations where extreme events occur reasonably frequently [½]
which would be expected to be the case for home insurance [½]
[Marks available 4, maximum 3]

(iv)
The probability of insolvency could be underestimated [1]
which could lead to the insurance company holding insufficient capital (or taking out insufficient reinsurance) [1]
[Total 11]

Parts (i) and (ii) were very well answered.

Part (iii) was poorly answered, with many candidates referring only to one of the three aspects of negative claims, skewness and tail thickness. Alternative comments that were clear, distinct and relevant to the context of the question were also awarded credit.

Part (iv) was fairly well answered, although many candidates lost marks for not relating their answer to the information in the question that the company will face insolvency if monthly claim amounts reach or exceed 1,000.

Q7
(i)
Y_t = X_t - X_t-3 = (1 - B^3) X_t
Hence:
(1 - (alpha + beta) B + alpha * beta * B^2) Y_t = e_t [1]

(ii)
The characteristic polynomial is 1 - (alpha + beta) z + alpha * beta * z^2 [½]
with roots 1/alpha and 1/beta [½]
Hence, Y_t is stationary for abs(alpha) < 1 and abs(beta) < 1 [1]

(iii)
rho_1 - (alpha + beta) + alpha * beta * rho_1 = 0 [1]
rho_2 - (alpha + beta) * rho_1 + alpha * beta = 0 [1]

(iv)
Substituting the observed values of the auto-correlation, and letting
M = alpha + beta and N = alpha * beta gives:
0.5 − M + 0.5N = 0 [½]
0.2 − 0.5M + N = 0 [½]
The first equation gives M = 0.5 + 0.5N and substituting into the second gives:
0.2 − 0.25 − 0.25N + N = 0 [½]
So, 0.75N = 0.05 [½]
and so N = 1/15 = 0.06667 [½]
and M = 8/15 = 0.5333 [½]
This means that alpha and beta are the roots of the quadratic equation:
x^2 - 0.5333x + 0.06667 = 0 [1]
which are 1/3 (0.3333) and 1/5 (0.2) [1]

(v)
Since Y_t = X_t - X_t-3, we have that
X_550 = Y_550 + X_547 [½]
and
X_551 = Y_551 + X_548 [½]

THEN EITHER:
The forecasted values
x_550_hat = y_550_hat + x_547 [½]
and
x_551_hat = y_551_hat + x_548 [½]
where
y_550_hat = 0.53333 y_549 - 0.06667 y_548
= 0.53333 (x_549 − x_546) - 0.06667 (x_548 − x_545) [1]
and
y_551_hat = 0.53333 y_550_hat - 0.06667 (x_549 − x_546) [1]

OR:
The forecasted values
x_550_hat = 0.53333 (x_549 − x_546) - 0.06667 (x_548 − x_545) + x_547 [1½]
and
x_551_hat = 0.53333 (0.53333 (x_549 − x_546) - 0.06667 (x_548 − x_545))
- 0.06667 (x_549 − x_546) + x_548
= 0.21778 (x_549 − x_546) + 0.03556 x_545 + 0.96444 x_548 [1½]

[Total 14]

Part (i) was well answered.

Part (ii) was fairly well answered. However, as the question asks for a range of values of alpha and beta, candidates who stated that Y_t is stationary for abs(1/alpha) > 1 and abs(1/beta) > 1 did not receive full marks.

Part (iii) was poorly answered. The most common errors were
Misunderstanding what was meant by the "second" and "third" Yule-Walker equations. The statement of the first Yule-Walker equation in the question was intended to make this clear.
Expressing the equations in terms of autocovariances, rather than autocorrelations as per the question.
Candidates are reminded to read the question carefully.

Parts (iv) and (v) were very poorly answered overall, despite the fact that most candidates who answered part (iii) correctly made a reasonable attempt at them. Candidates are reminded that if they are unable to answer one part of a question, then they may still gain credit in subsequent parts by assuming a "dummy" answer.

Q8

(i)

Operates in continuous time (t >= 0) [½]
with discrete state space {ONline, OFFline} [½]
and transition probabilities do not depend on history prior to arrival in current state (Markov property) [1]

(ii)

d P_OFF_OFF(t) / dt = 0.75 * P_OFF_ON(t) - 0.25 * P_OFF_OFF(t) [2]

(iii)

P_OFF_ON(t) + P_OFF_OFF(t) = 1 [1]
Substituting this into the equation in part (ii), we obtain
d P_OFF_OFF(t) / dt + P_OFF_OFF(t) = 0.75 [1]
so that
d (exp(t) * P_OFF_OFF(t)) / dt = 0.75 exp(t) [1]
Then,
exp(t) * P_OFF_OFF(t) = 0.75 exp(t) + constant [1]
Initial condition: P_OFF_OFF(0) = 1 [1]
Therefore, constant = 0.25 [1]
So,
P_OFF_OFF(t) = 0.75 + 0.25 exp(-t) [1]

(iv)

If X_t is a random variable denoting the amount of time spent offline over the period [0, t], given that the customer is offline at time 0, then the expected value of X_t is given by:
E(X_t) = INT (0, t): P_OFF_OFF(s) ds [2]
= INT (0, t): (0.75 + 0.25 exp(-s)) ds [1]

= [0.75s - 0.25 exp(-s)]:(0, t) [1]
= 0.75t + 0.25(1 - exp(-t)) [1]
Either online or offline at any time so total time spent online is:
t - (0.75t + 0.25(1 - exp(-t))) = 0.25t - 0.25(1 - exp(-t)) [1]
So, proportion of time spent online is:
(0.25t - 0.25(1 - exp(-t))) / t = 0.25 - 0.25(1 - exp(-t)) / t [1]

[Total 18]

Parts (i) and (ii) were well answered.

Part (iii) was poorly answered, despite being a relatively standard application of the integrating factor method. Common errors included
Attempting to apply the integrating factor before applying the condition P_OFF_ON(t) + P_OFF_OFF(t) = 1.
Applying the initial condition P_OFF_OFF(0) = 0 instead of 1.

Part (iv) was very poorly answered overall, although most candidates who answered part (iii) correctly made a reasonable attempt at part (iv). Candidates who failed to solve the differential equation in part (iii) but who answered part (iv) correctly based on any plausible expression for P_OFF_OFF(s) were awarded full marks.

Q9
(i)
For the transition matrix to be valid each row should sum to 1 [½]
This holds for all values of alpha [½]
All entries of the matrix should lie between 0 and 1 inclusive [½]
Therefore:
The entries of alpha and alpha^2 require 0 <= alpha <= 1 [½]
The entries ½ - alpha and 1 - 2 * alpha require alpha <= ½ as alpha must be >=0 from above [½]
The entry 1 - 2 * alpha - alpha^2 requires alpha <= -1 + sqrt(2) as alpha must be >=0 from above [1]
Hence, overall 0 <= alpha <= sqrt(2) - 1 [½]

(ii)
If 0 < alpha <= sqrt(2) - 1 [½]
then any state can be reached from any other state and so the chain is irreducible [1]
If alpha = 0 [½]
then it's not possible to leave states A or D and so the chain is reducible [1]

(iii)
Transition matrix is:

A	0.56	0.2	0.2	0.04
B	0.2	0.3	0.3	0.2
C	0.2	0.3	0.3	0.2
D	0	0.2	0.2	0.6

[1]

For company D to provide cover to Mary for at least four years before she changes provider, Mary must renew her policy with company D at least three times [1]
The probability of renewing three times with company D is
0.6^3 = 0.216 (or 27/125) [1]

(iv)
The company covering the car on 23 December 2020 will be that securing James' business at the second renewal [1]
The probability of James being with Company A for the second renewal is the first element of the second order transition matrix, which is: [1]
0.56 * 0.56 + 0.2 * 0.2 + 0.2 * 0.2 + 0.04 * 0 = 0.3936 [½]
and hence the probability of James being with a different company for the second renewal is 0.6064 [½]

(v) Transition matrix is:

ADDA	0.6	0.2	0.2
B	0.4	0.3	0.3
C	0.4	0.3	0.3

[2]

(vi)
Observe that currently the probability of customers going from Company D to Company A is zero [1]
which suggests that there may be reasons customers of Company D do not want to use Company A [½]
There may also be reasons customers of Company A do not want to use Company D [½]
ADDA might merge its pricing system. This would change the relative pricing of an individual's cover from the different companies. To the extent that pricing is a driver of the likelihood of customers moving this might change the probabilities [1]
Economies of scale may lead to lower premiums. To the extent that pricing is a driver of the likelihood of customers moving this might change the probabilities [1]
It is not clear whether the products sold by ADDA would be the same as those previously sold by Company A or Company D. This might change the probabilities [1]
To the extent that customer service is a driver, it is not clear what the customer service of ADDA would be relative to Company A or Company D. This might change the probabilities [1]
Reduction in competition might encourage a new entrant [1]
It might be a valid assumption that customer behaviour continues unaltered after the merger [½]

[Marks available 7½, maximum 5]

[Total 20]

Answers to this question were satisfactory, with the exception of part (vi) which was poorly answered.

In part (i), many candidates did not mention that the rows of the matrix are required to sum to 1 or that this holds for all values of alpha. Many candidates also failed to test whether all the entries of the matrix are between 0 and 1 inclusive.

In part (ii), candidates were not penalised for showing alpha > 0 in the first line rather than repeating the maximum value from part (i). However candidates who stated without explanation that the chain is either reducible or irreducible received no marks.

In part (iii), the most common errors were to use the fourth rather than the third power, and to raise the whole transition matrix rather than the bottom right entry to the relevant power.

In part (iv), the most common error was to raise the transition matrix to the third rather than the second power.

In part (v), many candidates lost marks for failing to label the rows of their transition matrix.

Few candidates provided a sufficient range of comments to score highly in part (vi). Alternative comments that were clear, distinct and relevant to the context of the question were also awarded credit.

[Paper Total 100]

END OF EXAMINERS' REPORT

