

INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

28th November 2023

**Subject CS2B – Risk Modelling and Survival Analysis
(Paper B)**

Time allowed: 1 Hour 45 Minutes (14.45 – 16.30 Hours)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

- 1. Mark allocations are shown in brackets.*
- 2. Attempt all questions beginning your answer to each question on a new page.*
- 3. Attempt all sub-parts of the question in one document only, unless otherwise instructed to do so.*
- 4. All the detailed guidelines are available on exam screen.*
- 5. Do save your work in solution template on a regular basis.*
- 6. If Any, Data set file(s) accompanying the question paper is available for download on the exam screen.*
- 7. You need to import the same into R studio as soon as you begin the exam.*
- 8. Ensure to copy and paste R codes and output at regular intervals onto the solution template.*
- 9. Please check if you have received complete Question Paper and no page is missing. If so, kindly get new set of Question Paper from the Invigilator.*

AT THE END OF THE EXAMINATION

Please return this question paper to the supervisor separately. You are not allowed to carry the question paper in any form with you. You are requested to save and submit the work before leaving the examination premises.

Q. 1)

- i) Simulate a time series of length 500 observations in a line chart having following arguments:
- Seed Value = 100
ar = 0.9
ma = 0.2
order = c(1,1,1) (3)
- ii) Comment on the general features of the chart. (2)
- iii) Plot the sample Autocorrelation Function (ACF) and sample Partial Autocorrelation function (PACF) of the data and paste the above graphs. (2)
- iv) Determine the best least squares linear fit, adding it to your chart in part (i) and paste the new chart. (3)
- v) Explain whether this least square linear trend can be removed such that stationary distribution is appropriate for the residuals. (2)
- vi) Fit an AR(1), AR(3) and ARMA(1,1) model to the time series. (3)
- vii) Determine the corresponding 95% Confidence Interval for the AR(1) model fitted above. (2)
- viii) Mention the best fit model from your observations in part (vi). Calculate the predicted values using the best model fit above for 10 steps ahead. (2)
- ix) Construct the Autocorrelation Function (ACF) and sample Partial Autocorrelation function (PACF) for the residuals of the best fitted model above and plot the graph. (2)
- x) Comment on the graphical outputs of part (iii) and part (ix). (2)
- xi) Perform the Ljung and Box Portmanteau test for the residuals of the model with lag of 4, 6, 12 and comment whether the model is an appropriate fit. (5)
- [28]**

Q. 2) Consider the data file “*Claims.csv*” of claim amounts in 000’s in respect of a commercial property portfolio over a period of 10 years.

- i) Calculate the block maxima for these claims using block sizes of 5 claims and paste the output. (2)
- ii) Plot a histogram of the block maxima and label the diagram. (2)
- iii) Plot an empirical density function in the same graph as in part (ii) above. (2)
- iv) Fit a Weibull distribution on the block maxima and calculate the estimates of the parameters of the Weibull distribution. (3)
- v) Calculate the mean, standard deviation and skewness of the block maxima. (3)
- vi) Write a function in R to compute the log-likelihood of the above distribution.

Hint : Use the following PDF function to compute the log likelihood.

$$\frac{1}{\beta} \left(1 + \frac{\gamma(x-\alpha)}{\beta} \right)^{-\left(1+\frac{1}{\gamma}\right)} \exp \left(- \left(1 + \frac{\gamma(x-\alpha)}{\beta} \right)^{\frac{1}{\gamma}} \right) \quad (2)$$

- vii)** Find the maximum likelihood estimate of the parameters of the above distribution. Use mean, standard deviation and skewness calculated above as the initial estimates of the parameters. (2)
- viii)** Superimpose a Generalized Extreme Value distribution on the block maxima using the estimates of the parameters calculated above in part (vii). (2)
- ix)** Calculate the hazard rate for the block maxima and plot the graph of Hazard rate against block maxima and comment on the graph. (4)
- x)** Calculate the mean residual life for the block maxima and plot the graph of mean residual life against block maxima. Comment on the graph. (5)
- [27]

Q. 3) You are working in Country “XYZ”, for an organisation which works standard tables for mortality rates. Your colleague has collected the data and calculated the mortality rates. Two sets of graduation on the data have been performed.

The data file “*Std_Table.csv*” contains mortality data from a recent mortality investigation. The deaths and exposure data are all in respect of period 1 April 2022 to 31 March 2023 inclusive. The file contains the following variable.

Age : Age last birthday x
 Deaths : Number of observed deaths at age x
 Exposure : Central exposed to risk at age x
 Graduation 1 : Central mortality rate at age x, m_x , derived from graduation 1
 Graduation 2 : Central mortality rate at age x, m_x , derived from graduation 2

Due to unavailability of your colleague, your head of department has asked you to perform following steps and conduct the tests.

- i)** Calculate individual standardise deviation Z_x for each of the graduations. Display first 10 values of the data. (3)
- ii)** Calculate the third differences of graduated quantities for both the graduations and comment on the smoothness of each graduation. Paste first 10 values of the third differences column in your data. (3)
- iii)** Comment on the goodness of fit using chi square test for each of the graduations given that members of degree of freedom are as follows:

Graduation 1	Graduation 2
46	47

(3)

- iv) Determine number of positive and negative derivatives for each graduation. (3)
- v) Carry out a 2-tailed sign test using a null hypothesis that $P \sim \text{Binomial}(m, 1/2)$ using R, where m is the age. (3)
- vi) Determine the number of groups of positive deviations for each of the graduation. (3)
- vii) Determine p value of the grouping of signs test for each graduation.
Hint: Function $\text{choose}(n, r)$ in R for binomial coefficient may be used. (3)
- viii) Perform a serial correlation test for each of the graduation at lag 1. (3)
- ix) Perform cumulative deviations test on of the graduation and comment on the bias of the graduation rates. (3)
- x) Based on the finding of your above tests, comment on graduation rates to be published. (3)

[30]

Q. 4) A bank in the country Actuarial is recruiting employees for different sections. Initially, it is compulsory to work for Marketing section for one year. After working for one year, they can either move to Administration or Training sections. On exactly April 1st they have the choice of moving to any other sections. Probability of moving from one section to another is dependent on the age last birthday.

From the historical data it is evident that the transition probability of a person aged 'x' last birthday on April 1st is as per the following table:

	Marketing	Administration	Training
Marketing	$1 - 0.0025 * x - 0.0075 * x$	$0.0025 * x$	$0.0075 * x$
Administration	$0.003 * x$	$1 - 0.003 * x - 0.004 * x$	$0.004 * x$
Training	$0.001 * x$	$0.003 * x$	$1 - 0.001 * x - 0.003 * x$

- i) Create a vector with the state space of the Markov Chain using R Code. You should paste the results in your answer. (1)
- ii) Consider two persons aged 30 last birthday and aged 40 last birthday respectively who are currently in marketing section. Construct transition matrices of the section change probabilities. You should paste the results to your answer. (3)
- iii) Create a Markov Chain objects with the state space equal to your vector in part (i) and transition matrices from part (ii). You should paste the results to your answer.
(R package for Markov Chains may be loaded and used for the same). (3)
- iv) Calculate the probability of transition for the following employees who are currently in Marketing section:
 - a) an employee with age 30 last birthday moving into Training section in three years.
 - b) an employee with age 40 last birthday moving into Training section in five years. (2)

- v) Plot the transition probability matrices in part (iv) for age 40 years. (1)
- vi) Set a seed of 250. Simulate sequences of 250 states using *markovchainSequence* function for both persons in part (ii). Compute the frequency of the states. (3)
- vii) Use a bar chart to plot the relative frequency of the states (for age 30) and label the chart appropriately. (2)

[15]
