

INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

21st July 2022

**Subject CS2A – Risk Modelling and Survival
Analysis (Paper A)**

Time allowed: 3 Hours 30 Minutes (9.30 - 13.00 Hours)

Total Marks: 100

- Q. 1)** A stand-alone health insurance company incurs claims in a year as per the following discrete distribution.

No. of claims	Probability
0	60%
1	20%
2	20%

Claim Amount	Probability
10000	80%
100000	20%

Calculate the probability distribution of the aggregate total claims for the health insurer. [6]

- Q. 2)** An investigation studied the mortality of persons aged between exact ages 50 and 51 years. The investigation began on 1 January 2019 and ended on 31 December 2019. The following table gives details of 10 lives involved in the investigation.

<i>Life</i>	<i>Date of 50th birthday</i>	<i>Date of death</i>
1	1 February 2018	–
2	1 April 2018	1 October 2019
3	1 June 2018	–
4	1 September 2018	–
5	1 November 2018	15 March 2019
6	1 January 2019	–
7	1 May 2019	15 December 2019
8	1 July 2019	1 October 2019
9	1 August 2019	–
10	1 December 2019	–

Persons with no date of death given were still alive when the investigation ended.

- i) Calculate the central exposed to risk using the data for the 10 lives in the sample. (3)
- ii) Calculate the Maximum Likelihood Estimate of the hazard of death at age 50 last birthday. (1)
- iii) Hence, or otherwise, estimate q_{50} . (2)
- [6]

- Q. 3)** For a population of a certain country, estimates of survival probabilities are as under:

$${}_1p_{55} = 0.9925$$

$${}_2p_{55} = 0.9843.$$

Using the Gompertz's law, calculate the values of the parameters for the above survival probabilities. [4]

- Q. 4)** The losses on certain type of insurance policy X follow Burr distribution with density function:

$$f(x) = \frac{0.5*\alpha*0.4^\alpha*x^{0.5}}{(0.4+x^{0.4})^{\alpha+1}} \text{ where '}\alpha\text{' is an unknown constant.}$$

An insurer has an excess of loss reinsurance arrangement with retention limit of 234. Claims observed are as below:

- **Claims below 234 (10 claims)** : 201, 230, 158, 198, 199, 205, 141, 155, 197, 217
- **Claims above 234 (7 claims)**: Amounts unknown

- i) Estimate 'α' using method of maximum likelihood. (5)
 - ii) Estimate 'α' applying the method of percentiles to the median. (3)
 - iii) Explain briefly as to why do you find the estimate of α different under (i) and (ii) above. (2)
- [10]**

- Q. 5)**
- i) Describe the utility of ROC (Receiver's Operating Characteristics) curve in the context of a classification problem. (1)
 - ii) How is the Area under the curve analysed? (2)
 - iii) Illustrate the use of Recall, F1 score and Precision taking an appropriate example. (3)
- [6]**

- Q. 6)** In order to model a particular seasonal data set an Actuary is considering using a model of the form $[(1-B^3)1 - (\alpha+\beta)B + \alpha\beta B^2]X_t = e_t$, where B is the backward shift operator and e_t is a white noise process with variance σ^2 .
- i) For what value of "s" can the seasonal difference series $Y_t = X_t - X_{t-s}$ be stationary? (1)
 - ii) Determine the range of values of α and β for which (i) above can be stationary. (2)

After appropriate seasonal differencing, assume $\alpha=0.4$ and $\beta=0.2$

- iii) Write down the Yule-Walker equations. (2)
 - iv) Calculate the first two values of the auto-correlation function ρ_1 and ρ_2 . (5)
 - v) Forecast the next two observations \hat{x}_{101} and \hat{x}_{102} based on the observed values x_1, x_2, \dots, x_{100} of X_t . (5)
- [15]**

- Q. 7)** Claims under a certain line of insurance business follow Pareto distribution with mean 1000 and standard deviation 1200. The insurer has 500 policies in-force and is planning to take excess-of-loss reinsurance. The expected frequency of claims is 25%. Two different reinsurance arrangements are being considered. (1) Retention limit of 3300 with a premium of 3000 and (2) Retention limit of 2200 with a premium of 3900.
- i) Calculate the probability of the reinsurer getting involved under each of the two arrangements mentioned above. (4)

- ii) Calculate the Expected value of the reinsurance payout under both the arrangements. (6)
- iii) Suggest with a brief explanation, the most suitable reinsurance arrangement for the insurer out of the two arrangements. (2)

[12]

- Q. 8)** With the success of a first gold medal in the Javelin throw event during the 2021 Olympics, the Bharat Olympics Committee (BOC) has started showing keen interest in the development of future javelin throwers for the country.

In search of the new talent, the Committee inducts new athletes for training if they throw a javelin for a distance of at least 50m in 3 attempts. Such inducted athletes participate in two events each year. Out of these inducted athletes, the committee considers an athlete as qualified for extensive Olympics level training if they threw a minimum distance of 80m in three attempts. The Central Government has decided to bear the cost of training support for such qualified athletes. The Government has asked BOC to assess the cost of providing such support and therefore, BOC wishes to know the average time it can expect to take for athletes to qualify for this extensive training.

The Committee has maintained a list of 23 athletes who have thrown at least 50m distance in 3 attempts by the start of 2014. The progress of those athletes has been recorded up to and including the last event of 2020. The following data records the number of events which had been held before the qualification of an athlete in this cohort:

Qualified 6, 8, 8, 9, 9, 9, 11, 11, 13, 13, 13

Injured and stopped participating 4, 5, 8, 11, 14

The remaining seven athletes were still participating in the events at the end of 2020.

- i) Determine the median number of events taken to qualify for those athletes who qualified for Olympics level training during the period of observation. (2)
- ii) Calculate the Kaplan-Meier estimate of the survival function, $S(t)$, for the “hazard” of qualifying, where t is the number of events since 1 January 2014. (5)
- iii) Using the result from (ii), estimate the median number of events to qualify for the athletes. (2)
- iv) Explain the reasons for difference between the results in (i) and (iii) above. (3)

[12]

- Q. 9)** An insurance company has two business segments ‘A’ and ‘B’. The probability of making a loss in segment ‘A’ is 0.05 and that in segment ‘B’ is 0.40. Calculate the probability that both operations make a loss, using the generator function of a Clayton copula. You can use the value of Kendall’s Tau for the copula as 0.08. (4)

[4]

- Q. 10)** The most popular T20 cricket league Hindustan Premier League (HPL) has seen two new entrants LSG and GT reaching the finals for the year 2022. It was a very closely fought contest, with each team scoring exactly 221 runs in their allotted 20 overs. Now the winner for the tournament has to be decided through a modified version of super over, in order to identify the winner quickly. In this modified version of super over, the players of each team will be given 6 balls initially, in turns, to hit the wickets by bowling. Bowler of each team is given chances in alternate fashion, one after the other. The team which hits the wickets will

be given one point and in case, it misses the wickets, a point will be awarded to the opposing team. At the end of allotted 6 balls, the team which has maximum points (either owing to hitting the wickets in its chances or due to other team missing the wickets in their chances), is adjudged the winner of HPL. In case there is a tie, each team is asked to bowl additional balls and the team which has a clear lead of 2 points over the other, is adjudged the winner.

At the end of 20 balls, both LSG and GT stand tied at 10-10 each. If LSG hits the wicket in the next ball, then the score will be 11-10 in its favour and if GT misses the wicket in the following ball, then LSG would win the tournament. In case LSG misses the wicket in the next ball (Score 10-11 in GT's favour) and GT hits the wicket the following ball after that, then GT would win the tournament.

In case LSG wins the next point and GT wins the point after that, then the match is back tied.

The probability of LSG winning a point is 60%. The probability of GT winning a point, therefore, is 40%.

You are asked to model the tied position in this case as a Markov chain.

- i) Clearly specify the five different states of this Markov chain. (2.5)
 - ii) Create a transition matrix based on the states obtained in (i) above. (2.5)
 - iii) Calculate the number of points that must be played before there is more than 90% chance of the match being completed. (4)
 - iv) Calculate the probability that LSG eventually wins the match and hence, the tournament. (4)
 - v) Compare the result in (iv) with the probability of LSG winning a point and comment. (2)
- [15]**

Q. 11) i) Which of the below methods can be used in predicting the continuous dependent variable?

- I. Linear regression
- II. Logistic regression

- A. Only I
- B. Only II
- C. Both I and II
- D. None of these (1)

ii) Which of the following statements are true?

- I. Lasso regression uses subset feature selection
- II. Ridge regression uses subset feature selection
- III. Lasso regression has a non-linear solution
- IV. Ridge regression has a closed-form solution

- A. I, II and III only
- B. I and IV only

- C. I and III only
 D. II, III and IV only (2)

iii) Which of the following statements are FALSE?

- I. Linear models for time series are stationary when they include functions of time.
 II. All moving average processes are stationary.
 III. All random walk processes are stationary.

- A. I and II only
 B. II and III only
 C. I and III only
 D. I, II and III only (2)

iv) Match the following:

Type of Process	Model
(I) State Space Discrete & Time Discrete	(a) Markov Jump process
(II) State Space Discrete & Time Continuous	(b) Ito Process
(III) State Space Continuous & Time Discrete	(c) Random Walk
(IV) State Space Continuous & Time Continuous	(d) ARIMA time series

- A. (I) – (b), (II) – (d), (III) – (a), (IV) – (c)
 B. (I) – (a), (II) – (d), (III) – (c), (IV) – (b)
 C. (I) – (c), (II) – (a), (III) – (d), (IV) – (b)
 D. (I) – (d), (II) – (b), (III) – (c), (IV) – (a) (2)
 [7]

Q. 12) An Actuary has conducted investigations into the mortality of the following classes of lives:

- a) female members of a medium-sized pension scheme
 b) young population of a county in order to study accident hump
 c) entire population of a large developed country to prepare a new standard table

The Actuary wishes to graduate the crude rates.

State an appropriate method of graduation for each of the three classes of lives and, for each class, briefly explain your choice. [3]
