INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

20th June 2019

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

Time allowed: 3 Hours 15 Minutes (14.45 – 18.00 Hours)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. *Please read the instructions inside the cover page of answer booklet and instructions to examinees sent along with hall ticket carefully and follow without exception.*
2. *Mark allocations are shown in brackets.*
3. *Attempt all questions, beginning your answer to each question on a separate sheet.*
4. *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 your answer book and this question paper to the supervisor separately. You are not allowed to carry the question paper in any form with you.

Q. 1) The Government has carried out the mortality investigation of the people getting admitted to Government hospitals. The Government asked all public hospitals to perform the study on the number of deaths happened in hospitals during the sample period 1^{st} Jan 2015 till 1^{st} Jan 2018. Patients were followed from 1^{st} Jan 2015, until they died, discharged while alive or reached 1^{st} Jan 2018. During the investigation period, some of patients died, some patients survived, some of them got discharged from hospital and some of them are still in hospital. Describe the types of censoring present in this investigation. **[2]**

Q. 2) Uber cabs are waiting in a queue for passengers to come out from the airport. Passengers for those cabs arrive according to a Poisson process with an average of 60 passengers per hour. A cab departs as soon as two passengers have been collected or 3 minutes have expired since the first passenger has got in the cab. Suppose you get in the cab as first passenger. What is your average waiting time for the departure? **[3]**

Q. 3) **i)** List the four ways of measuring the tail weight of a particular distribution. (1)

ii) Compare the two Pareto distributions with parameters $\alpha = 0.2$ and $\alpha = 0.3$ (with equal λ) respectively using 'Limiting density Ratio" method. (2)
[3]

Q. 4) A water research company has developed two indigenous methods of testing the purity level in water used for manufacturing medicines. Water samples collected from 100 consecutive days are tested for purity under both methods and the results are as follows:

Under Method I

The purity is accepted for 75 days and out of these days the results are found to be incorrect for 10 days. Out of the remaining 25 days, the results are found to be incorrect for 15 days.

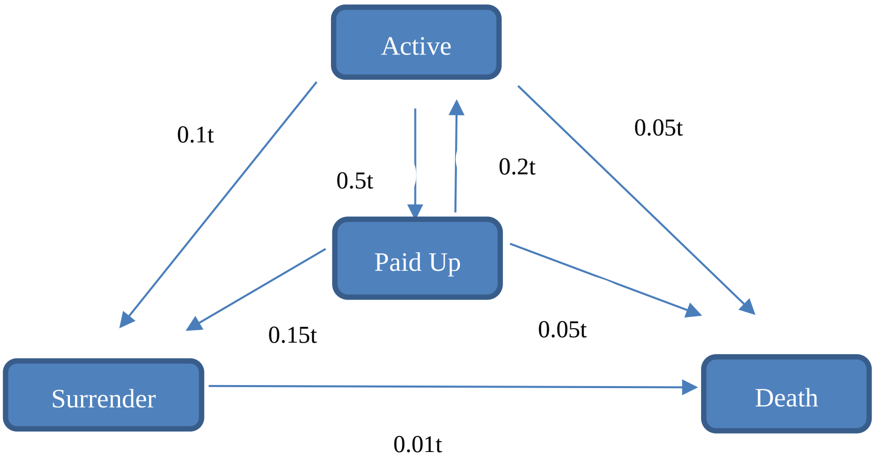
Under Method II

The purity is accepted for 70 days and out of these days, the results are found to be incorrect for 15 days. Out of the remaining 30 days, the results are found to be incorrect for 15 days.

i) Construct the confusion matrices for both methods and calculate the F_1 Score and False Positive Rate for each matrix and suggest which method is more suitable. (5)

ii) Explain the method of s-fold cross-validation technique used in predictive modelling and state its advantages. (3)
[8]

Q. 5) Consider the following time-inhomogeneous Markov jump process {X(t) : t >_ 0} of a life insurance company with policy states defined as active [A], surrender [S], paid up [P] and death [D] statuses. Transition rates as shown below:



i) Write down the Chapman-Kolmogrov equations and differentiate it to obtain the forward and backward equations. (4)

ii) State Kolmogorov's forward differential equation for probability $P_{PS}(s,t)$ and $P_{SS}(s,t)$. (4)

iii) A policyholder aged 40 surrendered his policy at 3rd policy anniversary. Calculate the probability that the policyholder is still alive at the 5th policy anniversary. (2)

[10]

Q. 6) The country of Oceania is blessed with lot of oil resources and manage almost all its expenditure from selling oil to other countries. As a result citizens of Oceania seldom pay taxes to Government of Oceania. The Country is divided into four parts namely East, West, North and South for administrative convenience. The country has a total population of One Lakh and the population is uniformly distributed to all the four states. The people in State East pays tax very rarely and pays according to a Poisson process at a rate of one payment per 25 years. However, they are paying One Crore Oceania $ as tax. The people of State North and South are paying taxes according to Poisson process at a rate of one payment per 5 years. However, the tax amount is either 10 Lakhs $ or 20 Lakhs $ with equal probabilities. The people of State West are also paying taxes according to Poisson process at a rate of one payment per 5 years but the tax amount is uniformly distributed from One Lakh $ to 10 Lakh $ with taxes are rounded to the nearest Lakh.

Let T denote the aggregate annual tax collected from the citizens of Oceania. Then

i) Calculate the mean and variance of T. (4)

ii) Using Normal Distribution as an approximation to the distribution of T, Calculate the probability that the tax amount collected for a particular year exceed 2000 Crores $. (2)

iii) Happy with expected tax collection figures, the Government of Oceania intends to provide a tax relief to small tax payers. The Government is providing a flat 20% refund of tax for people who are paying tax upto 10 Lacs Oceania $ (including 10 Lacs Oceania $). Calculate the revised Mean and Variance of annual aggregate tax collected and calculate the probability that the aggregate tax amount exceed 2000 Crores Oceania $. (6)

[12]

Q. 7) A medium size life insurance company is selling Joint Life Term Insurance product to married couples. The product waives the premium on death of one life and pay the Sum Assured on death of second life. The probability of survival for thirty years for a 50 years old male life is 0.60 and for a 50 year old female life is 0.65. The company sold this policy to couple both aged 50 for a thirty years term.

Calculate the probability of paying a death benefit in the above policy using

i) The Gumbel Copula with $\alpha = 3.5$ (2)

ii) The Clayton Copula with $\alpha = 3.5$ (1)

iii) The Frank copula with $\alpha = 3.5$ (1)

iv) Comment on the suitability of the above copulas by comparing the results to those when independent deaths are assumed. (2)

[6]

Q. 8) The Cox proportional hazards model is to be used to model the rate at which people are getting married. Assuming they get married within 3 years once they start looking for their life partner. In the fitted model, the hazard depends on the time, t, since starting the search for the life partner. The covariates, their categories and the fitted parameters for each category are shown in the table below:

Covariate	Possibility	Parameter
Profession	Service	0.3
	Business	0.5
	House Maker	0
	Social Service	-0.1
Gender	Male	0
	Female	0.3
Location	Metro	0
	Non Metro	0.2
Age Band	20-25	0.7
	25-30	0.5
	30-35	0
	35-40	-0.4

i) Defining clearly all the terms you use, write down an expression for the hazard function in this model. (2)

ii) For a female social worker aged 37 living in Mumbai, who has been looking to get married for last 1 year, the probability of staying single for next 2 years is 0.3. Calculate the probability that the 24 aged male living in non-metropolitan city and doing business will stay single for next 2 years given that he is also looking for a partner from last 1 year. (3)

iii) Explain the terms “under-graduation” and “over-graduation”. (2)

A graduation of the above experience of the population of India has been carried out. The following is an extract from the results.

Age Nearest Birthday	**Actual No of Marriages**	**Graduated Rate**	**Central Exposed to risk**	
X	Qx	Ux	Ex	ExUx
20	29.00	0.000712	39100	27.85
21	28.00	0.000516	66200	34.16
22	2.00	0.000079	76500	6.05
23	24.00	0.000255	82500	21.08
24	19.00	0.000393	64000	25.15
25	14.00	0.000446	32400	14.46
26	23.00	0.001171	19700	23.07
27	16.00	0.001638	10400	17.04
28	13.00	0.000366	55300	20.22
29	3.00	0.000164	10000	1.64
30	3.00	0.001158	4400	5.10
31	10.00	0.000250	46500	11.62
32	14.00	0.000187	87000	16.25
33	5.00	0.000453	4400	2.00
34	8.00	0.000106	73400	7.81
35	12.00	0.000683	19400	13.25
36	2.00	0.000062	43900	2.74
37	7.00	0.000038	75300	2.84
38	3.00	0.000026	57300	1.47
39	2.00	0.000060	23800	1.43
40	4.00	0.000052	82900	4.31

iv) Perform the following tests to detect different aspects of the adherence of the graduation to the data. For each test, state clearly the features of the graduation that the test is able to detect, and comment on your results.
- Chi-squared test
- Signs test (6)

[13]

Q. 9) The one year mortality assumption of male population of a region is 24 deaths per 1000 deaths. Calculate the first month mortality rate assumption considering following assumptions about mortality:
a) Uniform distribution of deaths.
b) Constant force of mortality.

In your answer, draw the shape of the survival function for one year under each of the two assumptions. **[4]**

Q. 10) Yt, a second-order autoregressive process is defined as below

Yt = -2αYt-1 + α2Yt-2 + Zt
where {Zt} is a zero-mean white noise process with Var(Zt) = σ2.

i) Determine the range of values of α for which the process Y can be stationary. (2)

ii) Derive the auto covariance’s γ1 and γ2 of Y in terms of α and σ. (5)

iii) Give an example of a circumstance in which a form of exponential smoothing might be expected to outperform Box-Jenkins forecasting in the prediction of future values of the time series. (2)

[9]

Q. 11) i) Calculate the auto covariance function and the autocorrelation function for the m^th order Moving Average process.

$$X_t = \mu + \frac{1}{m+1}(e_t + e_{t-1} + \ldots + e_{t-m}),$$

where {et : t 0} is a sequence of uncorrelated, zero-mean random variables with common variance σ_e^2 (4)

ii) Determine whether the above process in part **(i)** is invertible if m = 2. (2)

iii) The data set plotted in Figure 1a (given below) represents the number of applications, xt, for travel insurance received by an insurance company's web site, measured for 60 consecutive months. Figure 1b (given below) displays the logarithm of the same data set, yt = ln(xt). A statistician decides, based on these plots, to fit a linear time series model to yt rather than to xt. State, giving reasons for your answer, whether you agree with this decision. (2)

iv) Below are the ACF and PACF for three time series. For each series, state whether it is autoregressive or moving average, and the order (p or q).

a) Figure 2
b) Figure 3
c) Figure 4 (3)

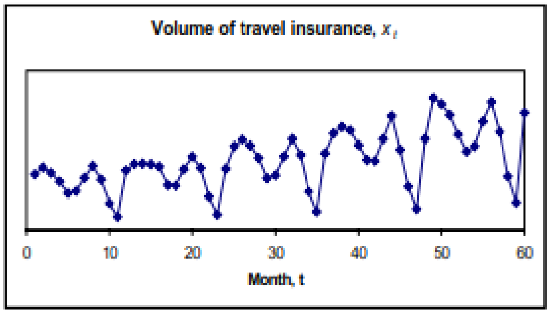


Figure 1a

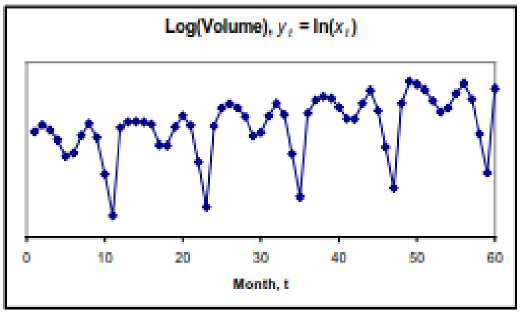
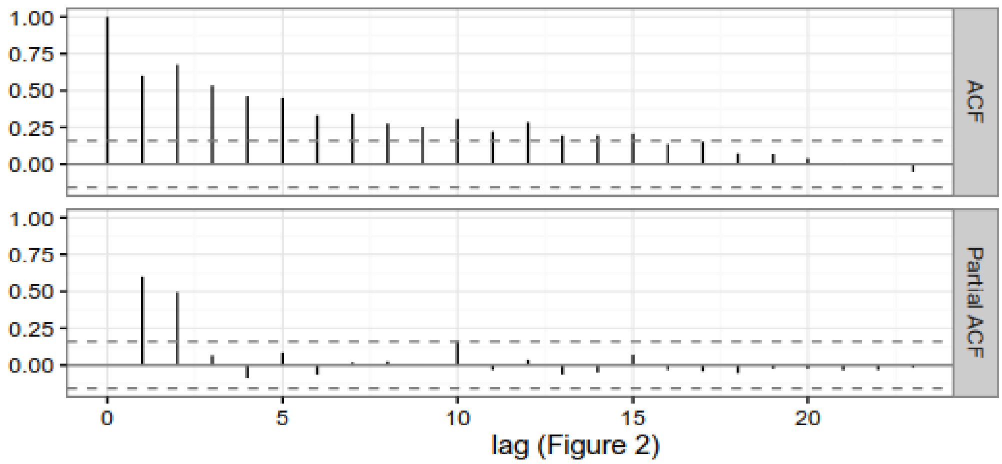
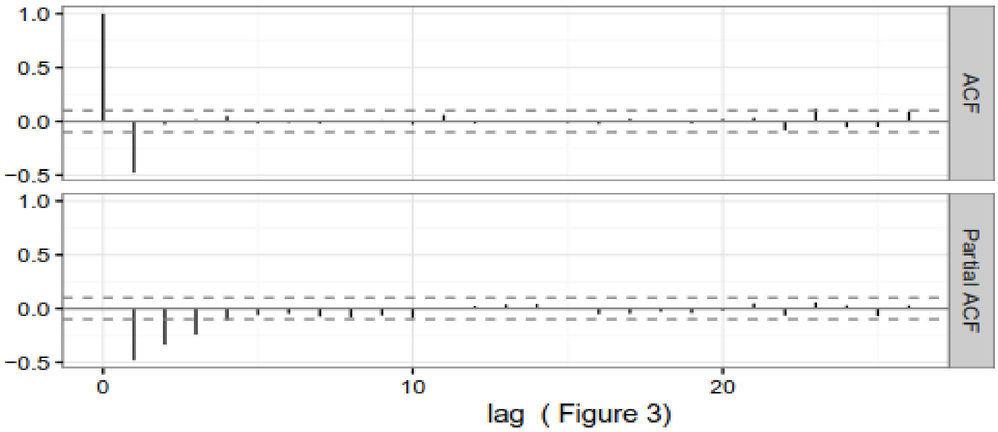
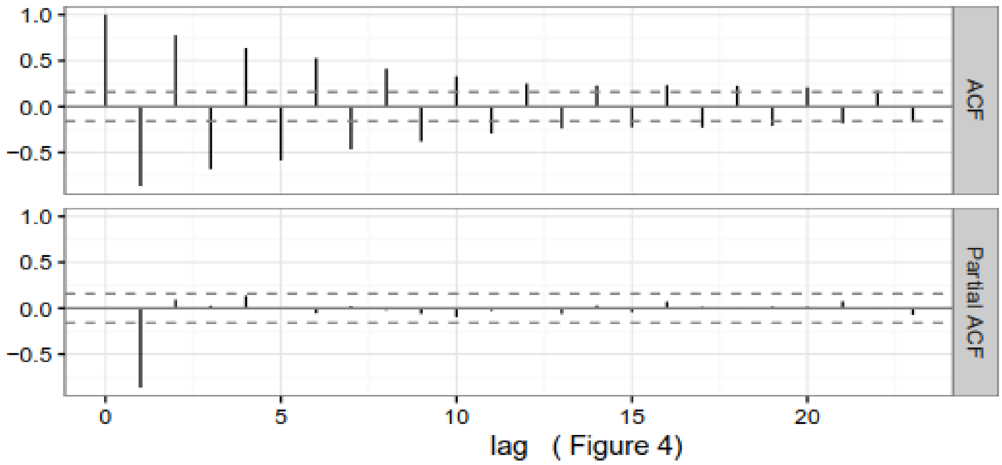


Figure 1b

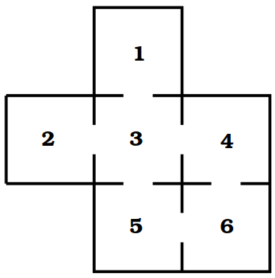






[11]

Q. 12) A rat runs through the maze shown below. At each step, it leaves the room it is in by choosing at random one of the doors out of the room.



i) Prepare the transition matrix P for this Markov chain. (2)

ii) Show that it is irreducible but not aperiodic. (2)

iii) Prove that the stationary distribution of this markov chain is

$$\pi = (\frac{1}{12}, \frac{1}{12}, \frac{4}{12}, \frac{2}{12}, \frac{2}{12}, \frac{2}{12}).$$ (3)

iv) Find the expected time to return to room 1. (1)

v) A piece of mature cheddar placed on a deadly trap in Room 5. The mouse starts in Room 1. Find the expected number of steps before reaching Room 5 for the first time, starting in Room 1. (3)

[11]

Q. 13) Consider a random walk on the finite states {-2, -1, 0, 1, 2}. If the process is in state i {i = -1,0,1} at time n, then it moves to either $i-1$ or $i+1$ at time n+1 with equal probability. If the process is in state -2 or 2 at time n, then it moves to state -1,0 or 1 at time n+1 with equal probability.

i) Write the transition probability matrix P for this random walk and draw the transition diagram. (3)

ii) Compute the stationary probability $\pi_i = (\pi_{i-2}, \pi_{i-1}, \pi_{i0}, \pi_{i1}, \pi_{i2})$ of this random walk. (3)

iii) Derive whether this random walk is an irreducible Markov chain and / or periodic. (2)

[8]

****