# INSTITUTE OF ACTUARIES OF INDIA

# EXAMINATIONS

**26th November 2020**

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

**Time allowed: 3 Hours 30 Minutes (09.30 – 13.00 Hours)**

**Total Marks: 100**

**Q. 1)** Find out the correct option for the below questions.

**i)** The process $X_t = 2 + e_t - 5\, e_{t-1} + 6\, e_{t-2}$ is

a) Invertible
b) Non-invertible
c) Both Invertible and non invertible
d) Cannot be determined
e) None of the above

(2)

**ii)** The process $12X_t = 10X_{t-1} - 2\, X_{t-2} + 12e_t - 11e_{t-1} + 2e_{t-2}$ is

a) Stationary and Non-invertible
b) Stationary and Invertible
c) Non-Stationary and Non-invertible
d) Non-Stationary and Invertible
e) Stationary and cannot be determined Invertibility

(2)

**iii)** The probability that a random loss exceeds the mean loss amount if the loss distribution is Pareto with $\alpha=3$, and $\lambda=2000$

a) 0.2396
b) 0.2963
c) 0.2369
d) 0.2936
e) None of the above

(2)

**iv)** Claims from a particular portfolio have a generalized Pareto distribution with $\alpha=6$, $\lambda=2000$ and $k=4$. A proportional reinsurance is in force with a retained proportion of 80%. The mean and variance of the amount paid by the insurer and reinsurer is

a) Insurer(64,9216), Reinsurer(32,576)

b) Insurer(64,9216), Reinsurer(64,576)

c) Insurer(128,9126), Reinsurer(64,576)

d) Insurer(128,9216), Reinsurer(32,576)

e) None of the above

(2.5)

**v)** An insurer believes that claims from a particular portfolio would follow a Pareto distribution with parameters $\alpha=2$, and $\lambda=900$. If the insurer does not want to pay for 20% claims from that portfolio what would be the policy deductible

a) 176.09
b) 200.43
c) 124.57
d) 106.23
e) None of the above

(2.5)

**vi)** Which is not the required criteria for an insurable risk

a) An occurrence probability should be very high
b) Individual risk events should be independent
c) There should be an ultimate liability on insurer
d) Moral hazards should be eliminated as far as possible (1)

**vii)** An insurer believes that the claims from its particular portfolio in coming year will be log normally distributed with a mean size of INR 5000 and a standard deviation of INR 7500. What percentage of claim that will be above INR 25000 in next year.

a) 1.3%
b) 1.8%
c) 2.1%
d) 3.5%
e) None of the above (3)

**viii)** A motor insurer operates a no claims discount system with the following levels of discount {0%, 10%, 20%, 35%, 60%}. The rules governing a policyholder's discount level, based upon the number of claims made in the previous year, are as follows:

- Following a year with no claims, the policyholder moves up one discount level, or remains at the 60% level (maximum level).
- Following a year with one claim, the policyholder moves down one discount level, or remains at 0% level (minimum level).
- Following a year with two or more claims, the policyholder moves down two discount levels (subject to a limit of the 0% discount level).

The number of claims made by a policyholder in a year is assumed to follow probability distribution as follows:

0 claim during a year = 0.75
1 claim during the year = 0.24
2 or more claim during the year = 0.01

What is the transition matrix for the no claims discount system.

a)
$$\begin{bmatrix} 0.25 & 0.75 & 0 & 0 & 0 \\ 0.25 & 0 & 0.75 & 0 & 0 \\ 0.01 & 0.24 & 0 & 0.75 & 0 \\ 0 & 0 & 0.25 & 0 & 0.75 \\ 0 & 0 & 0 & 0.25 & 0.75 \end{bmatrix}$$

b)
$$\begin{bmatrix} 0.24 & 0.75 & 0.01 & 0 & 0 \\ 0.24 & 0.01 & 0.75 & 0 & 0 \\ 0.01 & 0.24 & 0 & 0.75 & 0 \\ 0 & 0.01 & 0.24 & 0 & 0.75 \\ 0 & 0 & 0.01 & 0.24 & 0.75 \end{bmatrix}$$

c)
$$\begin{bmatrix} 0.25 & 0.75 & 0 & 0 & 0 \\ 0.25 & 0 & 0.75 & 0 & 0 \\ 0.01 & 0.24 & 0 & 0.75 & 0 \\ 0 & 0.01 & 0.24 & 0 & 0.75 \\ 0 & 0 & 0.01 & 0.24 & 0.75 \end{bmatrix}$$

d)
$$\begin{bmatrix} 0.25 & 0.75 & 0 & 0 & 0 \\ 0.25 & 0 & 0.75 & 0 & 0 \\ 0 & 0.25 & 0 & 0.75 & 0 \\ 0 & 0 & 0.25 & 0 & 0.75 \\ 0 & 0 & 0 & 0.25 & 0.75 \end{bmatrix}$$

(3)

**ix)** For transition probabilities derived in (viii) above, find out the stationary distribution:

a) (0.0422, 0.1264, 0.3782, 0.1134, 0.3411)

b) (0.0112, 0.0302, 0.0818, 0.2192, 0.6576)

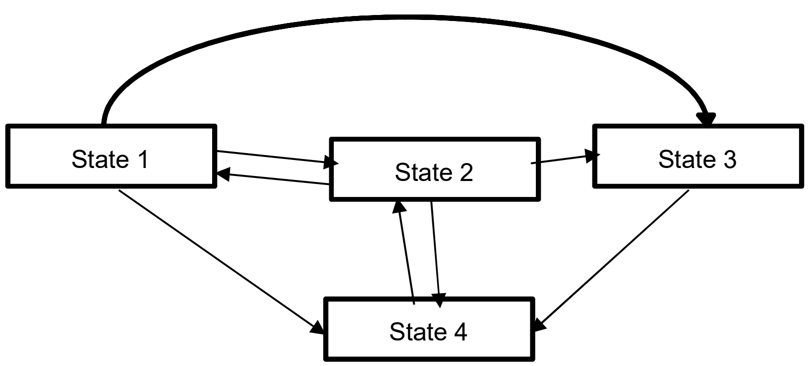
c) (0.2287, 0.2160, 0.1850, 0.0926, 0.2777)

d) (1, 0, 0, 0, 0)

e) None of the above

(5)

**x)** Select the correct forward differential equation using first principles for $\frac{\partial {}_tp_x^{24}}{\partial t}$ for following multiple state model in which S(t), the state occupied at time t by a life initially aged x, is assumed to follow a continuous- time Markov process.




Let $\mu_{x+t}^{ij}$ denote the force of transition at age x+t (t >= 0) from state i to state j, and let

${}_tp_x^{ij} = P\ (S(t) = j \mid S(0) = i)$

a) $\frac{\partial {}_tp_x^{24}}{\partial t} = {}_tp_x^{21}\mu_{x+t}^{14} + {}_tp_x^{23}\mu_{x+t}^{34} + {}_tp_x^{24}\mu_{x+t}^{44} - {}_tp_x^{22}\mu_{x+t}^{24}$

b) $\frac{\partial {}_tp_x^{24}}{\partial t} = {}_tp_x^{23}\mu_{x+t}^{34} + {}_tp_x^{21}\mu_{x+t}^{14} + {}_tp_x^{24}\mu_{x+t}^{42} - {}_tp_x^{22}\mu_{x+t}^{24}$

c) $\frac{\partial_t p_x^{24}}{\partial t} = {}_tp_x^{23}\mu_{x+t}^{34} + {}_tp_x^{21}\mu_{x+t}^{14} + {}_tp_x^{21}\mu_{x+t}^{12} - {}_tp_x^{24}\mu_{x+t}^{42}$

d) $\frac{\partial_t p_x^{24}}{\partial t} = {}_tp_x^{21}\mu_{x+t}^{14} + {}_tp_x^{23}\mu_{x+t}^{34} + {}_tp_x^{22}\mu_{x+t}^{24} - {}_tp_x^{24}\mu_{x+t}^{42}$

e) None of the above (5)

**xi)** Match the below examples with the correct stochastic process.

| **Example** | **Stochastic process** |
|---|---|
| (i) Wrestlers weight every year on 1st Jan | (A) Discrete time discrete space |
| (ii) Monthly car accidents in Mumbai city | (B) Continuous time discrete space |
| (iii) Temperature of incubator while hatching chicken eggs | (C) Discrete time continuous space |
| (iv) Trend of Air pollution index on meter | (D) Continuous time continuous space |

(2)

**[30]**

**Q. 2)** The following time series model is used for the daily increase rate in Covid-19 patients ($Y_t$) in the country of Indiana

$Y_t = 0.4Y_{t-1} + 0.2Y_{t-2} + Z_t + 0.025$

Where $\{Z_t\}$ is a sequence of uncorrelated identically distributed random variables whose distributions are normal with mean zero.

**i)** Find out the correct option for a, b and c when this model is considered as an ARIMA (a, b, c) model.

a) (0,2,1)
b) (1,1,0)
c) (2,0,0)
d) (2,1,1)
e) None of the above (2)

**ii)** Determine whether $\{Y_t\}$ is a stationary process. (1.5)

**iii)** Assuming an infinite history, calculate the expected value of the rate of increase of Covid-19 over this. (1)

**iv)** Calculate the autocorrelation function of $\{Y_t\}$ (4)

**v)** Explain how the equivalent infinite-order moving average representation of $\{Y_t\}$ may be derived. (1.5)

**vi)** If $R_t$ follows MA(1) and $S_t = 0.8+0.5t+R_t$ then prove that the standard deviation of first difference of $S_t$ will be higher than that of $R_t$. (6)

**[16]**

**Q. 3)** The life insurance company performing the analysis current experience (5 years) with population for its maximum selling age group of 30-39 years

**i)** Based on below data please test the hypothesis that the company's experience is in line with population.

| Age | Exposed to risk | Actual Deaths | Population mortality |
|---|---|---|---|
| 30 | 2,921 | 5 | 0.001605 |
| 31 | 4,040 | 6 | 0.001652 |
| 32 | 3,855 | 7 | 0.001712 |
| 33 | 4,640 | 7 | 0.001787 |
| 34 | 4,822 | 9 | 0.001875 |
| 35 | 5,583 | 10 | 0.001980 |
| 36 | 6,013 | 13 | 0.002104 |
| 37 | 5,911 | 12 | 0.002247 |
| 38 | 5,657 | 13 | 0.002412 |
| 39 | 5,219 | 14 | 0.002603 |

**a)** Calculate the test statistic (4)

**b)** State degrees of freedom (0.5)

**c)** State upper tail value of the Chi-square distribution with calculated degrees of freedom at 95% (0.5)

**d)** Conclusion on the hypothesis tested (1)

**ii)** Perform signs test by stating n, p and x. State the p-value and conclusion whether the Null hypothesis of positive deviations is having binomial distribution. (4)

**[10]**

**Q. 4)** **i)** In a city during rainy season the days were observed to be raining (R) or not raining (S) on any particular day. The data was recorded for a month of Jun2020 as below:

Week 1: RSRRSSS
Week 2: SRRSRSS
Week 3: RSRSRRS
Week 4: SSRRSSR
Week 5: RR

It was assumed that the rain prediction is dependent only on previous day and hence decided to fit Markov chain.

**a)** Calculate transition probabilities for the Markov chain. (2)

**b)** Determine the probability that it will rain on 3rd July 2020. (3)

**ii)** A life insurance agent sells on an average three life insurance policies per week. Use Poisson's law to calculate the probability that in a given week he will sell:

**a)** Some policies. (2)

**b)** 2 or more policies but less than 5 policies. (2.5)

**c)** Assuming that there are 5 working days per week, what is the probability that in a given day he will sell one policy? (2.5)

**[12]**

**Q. 5)** In the country of Indiana the testing kits for pandemic Covid-19 are limited in numbers. So the Government has decided to test only those citizens who qualify certain criteria.

Citizens of the country are given a questionnaire and asked for having physical symptoms of Covid-19. The questionnaire asked some simple questions like international travel history, coming in contact with any positive tested person, any co-morbid conditions. The symptom checked were fever, cough, difficulty in breathing, and any other respiratory problem persisting since last 3 days. An algorithm is then applied to estimate the number of infected persons in the country to augment medical services

In the test, both the questionnaire and the symptoms test are given to 2000 patients whose disease status is known. Some results from the test are shown in the table below

| True Status | Prediction from Symptom | | Prediction from questions | |
|---|---|---|---|---|
| | Has disease | Does not have disease | Has disease | Does not have disease |
| Has disease | 900 | 100 | 800 | 200 |
| Does not have disease | 300 | 700 | 200 | 800 |

**i)** Find out the correct option for (symptoms, questionnaire) test

a) Precision: (0.75,0.80) ; Recall: (0.90,0.80), F1 Score: (0.818, 0.80)

b) Precision: (0.90,0.80) ; Recall: (0.75,0.80), F1 Score: (0.90, 0.80)

c) Precision: (0.75,0.90) ; Recall: (0.75,0.80), F1 Score: (0.80, 0.90)

d) Precision: (0.75,0.90) ; Recall: (0.90,0.75), F1 Score: (0.818, 0.90)

e) None of the above (3)

**ii)** Show that the F1 score expresses the true positives as a proportion of the true positives plus the average of those incorrectly classified. (2)

**iii)** Comment on your result from (i) and on the usefulness of the test. (3)

**[8]**

**Q. 6)** Individual claims under a certain type of insurance policy are for either 1 or 2. The insurer is considering entering into an excess of loss reinsurance arrangement with retention $1 + k$ (where $k < 1$). Let $X_i$ denote the amount paid by the insurer (net of reinsurance) on the i th claim.

**i)** Calculate and simplify expressions for the mean and variance of $X_i$. (3)

The number of claims in a year follows a Poisson distribution with mean 500 (assume $\alpha = 0.2$). The insurer wishes to set the retention so that the probability that aggregate claims in a year will exceed 700 is less than 1%.

**ii)** Show that setting $k = 0.334$ gives the desired result for the insurer. (5)

**[8]**

**Q. 7)** In country of Originia with dense population in metro cities and having 30 geographical states, the study was undertaken by using the recent data collected for the latest pandemic in the world. The data available is only for one full year from 1st infection was observed till vaccine is prepared.

**i)** Explain why lives might be lost to the investigation if we are carrying out study of multistate model (states such as normal, infected, recovered, death). (2)

**ii)** Explain the above in context of informative and non-informative censoring. (2)

**iii)** Any other type of censoring is demonstrated by this data. (1)

In order to perform survival analysis, you have been asked to investigate the impact of covariates.

**iv)** List down possible covariates to perform survival analysis. (3)

**v)** Explain possible lifetime distributions for proportional hazard function. Discuss or debate its advantages or disadvantages for performing survival analysis for population of Originia. (8)

**[16]**

**********************