INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

22nd November 2024

CS2 - Risk Modelling and Survival Analysis (Paper A)

Time allowed: 3 Hours 15 Minutes

Total Marks: 100

Q. 1) A Markov chain on a triangle moves from one vertex to another with probability of 0.5. Find the probability that in 'p' number of steps, the Markov chain returns to its vertex from where it started.

A. $\frac{1}{3}+\frac{2}{3}\,(-0.25)^{p}$
B. $\frac{1}{3}+\frac{2}{3}\,(-0.5)^{p}$
C. $\frac{1}{2}+\frac{1}{3}\,(-0.5)^{p}$
D. $\frac{1}{3}+\frac{1}{2}\,(-0.75)^{p}$
E. $\frac{1}{2}+\frac{2}{3}\,(-0.25)^{p}$ **[3]**

Q. 2) A group of lives of 120 aged 40 years has 30% sick and rest healthy lives. For the sick lives, the force of decrement is 0.05 and for healthy it is 0.10, both for age x >=40 years. Find the mortality rate (q_{65}) for a life selected randomly from those who live up to 65 years:

A. 0.05384
B. 0.04878
C. 0.15872
D. 0.19707
E. 0.02557 **[4]**

Q. 3) A run-off life insurance product is getting lapsed at 0.4 per day. Assume policies are independent to each other, and that there are only 2 policies in-force. The last 2 policies have operational cost of Rs 591 per day. The insurance company is deliberating on a cost estimate for outsourcing the operation of these policies to a vendor which will incur negligible cost to continue these policies. The proposal may be considered if it results into lower overall cost to the company. Meanwhile, one of the policies has lapsed. Estimate the maximum cost the company may pay to the vendor to take up one policy left.

A. 1057.64
B. 918.24
C. 1024.12
D. 1287.24
E. 1124.68 **[4]**

Q. 4) In a genetic study, it was found that a gene appears in two forms – D or d. An organism has a pair of genes – DD (dominant) or Dd (mixed; order does not matter – Dd or dD) or dd (residual).
In any generation, a gene is derived from its previous generation (i.e. combination of two pairs of genes) with equal probability.
An organism is always combined with a mixed gene type progressively through generations.
Assuming the starting organism has mixed gene type, compute the probability that the n^{th} generation organism is with residual gene type.

A. $\{(2^{2n-1}-1)+0.5\}(2)^{-n}$
B. $\{(2^{n-1}-1)+1.5\}(2)^{-n}$
C. $\{(2^{n-1}-1)+1.5\}(2)^{-n-1}$
D. $\{(2^{n-2}-1)+1.5\}(2)^{-n}$
E. $\{(2^{n-2}-1)+0.5\}(2)^{-n}$

[4]

Q. 5) Consider sample autocorrelation estimates using 9504 data points. **[3]**

Lag	1	2	3	4	5
Co-efficient	-0.0211	0.2001	-0.0156	-0.0021	0.02001

Assuming normal distribution of co-efficients, select the lags at which co-efficient(s) is(are) statistically significant at 5%.

A. 1 and 3
B. 1 and 2
C. 3 and 4
D. 3, 4 and 5
E. Can not be determined as standard errors are not given.

Q. 6) Consider the number of policies and number of deaths as follows:

Age	**Policies in force as on date 'D'**				
	2010	**2011**	**2012**	**2013**	**2014**
48	65444	69000	72017	73065	73841
49	58741	59877	60587	61578	62547
50	58354	59544	60587	61587	63544
51	56874	58970	59575	59874	60587

Age	**Deaths**				
	2010	**2011**	**2012**	**2013**	**2014**
48	54	59	57	60	61
49	60	62	65	59	58
50	61	62	63	64	63
51	62	58	59	64	64

Calculate the force of mortality ($\mu_{x+0.5}$) at age 49 and 51, assuming that average number of policies in force throughout is same as in force policies as on date 'D' and Poisson model of mortality.

A. 0.00100 and 0.00103 respectively
B. 0.00101 and 0.00104 respectively
C. 0.00100 and 0.00104 respectively
D. 0.00102 and 0.00104 respectively
E. 0.00101 and 0.00103 respectively **[3]**

Q. 7) An insurer sells insurance cover to users of mobile phones, and each policy covers one mobile phone.
There were 19849 such policies which are in-force in a certain year. 100% repair cost is covered.
The number of claims is distributed as follows:

No. of claims in that year	*No. of policies*

0	5744
1	5477
2	4500
3	3254
4	874

Assuming Poisson distribution (with parameter 'λ') of no. of claims, calculate the 92% linear confidence interval for λ based on above:

A. (1.377, 1.402)
B. (1.383, 1.412)
C. (1.397, 1.402)
D. (1.402, 1.412)
E. (1.369, 1.415)

[3]

Q. 8) An kitchen appliance manufacturer has developed has developed fast chargers for its portable and chargeable electric kettle. Currently, the manufacturer is testing the life of batteries using such chargers.
'charge life' is defined as 'the number of times the appliance can be charged using such chargers'. It has conducted some tests and has found the following:

- 80% of the electric kettles lasted 410 charges.
- 60% of the electric kettles lasted 820 charges.

The fraction of electric kettles that would last for 1100 charges if battery life follows a Weibull distribution is:

A. 0.48097
B. 0.41093
C. 0.39282
D. 0.40912
E. 0.30887

[3]

Q. 9) Consider the a moving average process defined by:
$y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3}$,
where σ_t is a zero mean white noise process with variance σ^2.

Consider the following statements:

I. y_t has zero mean
II. Autocorrelation function will have zero value at lag 6.
III. y_t has variance σ^2.
IV. Autocorrelation function will have value one at lag 0.

Select the correct option:
A. I and II
B. I and III
C. II and IV
D. I, III and IV
E. Only III

[3]

Q. 10) Consider a survival distribution function at age x:
$S_0(x) = P[X>x]$ **[2]**
Such that $S_0(0) = 1$

$S_0(\infty) = \lim_{x \to \infty} S_0(x) = 0$
Select the correct statement:

A. Survival distribution function is non-increasing function of x as it may be possible to have higher probability of surviving for longer time period
B. Survival distribution function is non-increasing function of x as it is not possible to have higher probability of surviving for longer time period
C. Survival distribution function is increasing function of x as it may be possible to have higher probability of surviving for longer time period
D. Survival distribution function is increasing function of x as it is not possible to have higher probability of surviving for longer time period
E. None of the above

Q. 11) Consider a survival distribution function at age x:
$S_0(x) = P[X>x]$
Such that $S_0(0) = 1$
$S_0(\infty) = \lim_{x \to \infty} S_0(x) = 0$
Consider density and the respective cumulative function as $f_x(t)$ and $F_x(t)$ respectively such that:

$S_0(x) = 1 - \int_0^x f_0(z)dz$, and

Let $F_0(x) = 1-\left(1-\frac{x}{117}\right)^{\frac{1}{5}}$.

Assuming constant force of mortality and limiting age as 117, find the complete expectation of life at age 82.

A. 27.3365
B. 30.6587
C. 28.9872
D. 29.1667
E. 28.2148 **[3]**

Q. 12) Consider the following AR process with the errors having zero mean and unit variance
$y_t = 0.28 + 0.79\, y_{t-1} + \varepsilon_t$

Consider the following statements:
I. The unconditional mean of y_t is 0.3544
II. The unconditional variance of of y_t is 2.6602
III. The value of autocovariance function at lag 3 is 1.3561
IV. The value of autocorrelation function at lag 3 is 0.4930

Select the correct option:
A. I, III and IV
B. I and III
C. II, III and IV
D. II and IV
E. I, II and III

[4]

Q. 13) A transport company has for many years maintained a fleet comprising a large number of cars, each of which runs for the same mileage everyday under broadly the same road and traffic conditions. Every 4 weeks, tests of engine performance are carried out and in the event of failure to pass this test, the car is taken out of the service and permanently removed from the fleet. Cars are also taken from the fleet if they are damaged in an **[2]**

accident of defined severity (such accidents are not uncommon). You are asked, on the basis of records to be maintained over one year, to estimate the expected fleet life of a car and the separate accident and engine failure removal rates in each year of duration from entry into the fleet (year to year changes in the model can be ignored). Each car is reserved for the use of one particular driver throughout its life, and the management wishes to know whether cars have a longer life if handled by older drivers. The period of investigation runs from 1 January 2022 to 31 December 2022.
What records would you require?

A. Date of entering service of the driver, date of birth of driver, date of permanent removal from service of the driver, reason for permanent removal of the car
B. Date of entering service of the driver, date of birth of driver, date of permanent removal from service of the car, reason for permanent removal of the car
C. Date of entering service of the car, date of birth of driver, date of permanent removal from service of the driver, reason for permanent removal of the driver
D. Date of entering service of the driver, date of birth of driver, date of permanent removal from service of the driver, reason for permanent removal of the driver
E. Date of entering service of the car, date of birth of driver, date of permanent removal from service of the car, reason for permanent removal of the car

Q. 14) Consider a survival distribution function at age x:
$S_0(x) = P[X>x]$
Such that $S_0(0) = 1$
$S_0(\infty) = \lim_{x \to \infty} S_0(x) = 0$
Consider density and the respective cumulative function as $f_x(t)$ and $F_x(t)$ respectively such that:

$S_0(x) = 1 - \int_0^x f_0(z)dz$, and

Let $F_0(x) = 1-\left(1 - \frac{x}{120}\right)^{\frac{1}{6}}$.

Assuming constant force of mortality and limiting age as 120, find the complete expectation of life at age 30.

A. 76.5483
B. 77.1428
C. 79.2454
D. 76.6973
E. 78.7842

[3]

Q. 15) Suppose that $\phi_0, \phi_1, \phi_2,...$ are independent random variables. Let $P_n = \{1,... ,N\}$.

Consider a simple random walk process P_n. Assume it starts at 0 and for states ϕ_i for i = 1, 2, … ,

$P(\phi_1 = 1) = p$ and $P(\phi_1 = -1) = q = 1- p$.

Compute $E[e^{\alpha P_n}]$for $\alpha \in R$.

A. $(pe^{\alpha} + qe^{-\alpha})^{-n}$
B. $(pe^{\alpha} + qe^{\alpha})^{n}$
C. $(pe^{\alpha} + qe^{-\alpha})^{n}$
D. $(pe^{-\alpha} + qe^{-\alpha})^{n}$
E. $(pe^{-\alpha} + qe^{\alpha})^{n}$

[3]

Q. 16) Consider the following statements:

I. ARIMA process given by $X_t = X_{t-1} + e_t$ has variance of $var(X_0) + \sigma^2$.
II. ARIMA process given by $X_t = X_{t-1} + e_t$ is stationary.
III. If closing share price on day t is given by Z_t, such that $Z_t = Z_{t-1}e^{(\mu+e_t)}$, then this $\ln(Z_t)$ is equivalent to a random walk process with drift and is I(2) process.
IV. Suppose Y is an AR(2) process, then Y holds Markov property.

Choose the correct statements:
A. II and III are correct
B. Only II is correct
C. I and III are correct
D. I, II and III are correct
E. None of the statements is correct **[4]**

Q. 17) Consider the following statements w.r.t. the two-factor Lee-Carter model.

A. The parameter $\varepsilon_{x,t}$ is an independently distributed normal random variable with [mean, variance] as [0, to be estimated variance] and a_x is mean of the time-averaged logarithms of the central mortality rate at age x
B. The parameter $\varepsilon_{x,t}$ is an independently distributed normal random variable with [mean, variance] as [0, a^2] and a_x is mean of the time-averaged logarithms of the central mortality rate at age x
C. The parameter $\varepsilon_{x,t}$ is an independently distributed lognormal random variable with [mean, variance] as [0, a^2] and a_x is median of the time-averaged logarithms of the central mortality rate at age x
D. The parameter $\varepsilon_{x,t}$ is an independently distributed poison random variable with [mean, variance] as [**λ, λ**] and a_x is mean of the time-averaged logarithms of the initial mortality rate at age x
E. The parameter $\varepsilon_{x,t}$ is an independently distributed normal random variable with [mean, variance] as [0, a] and a_x is median of the time-averaged logarithms of the central mortality rate at age x **[2]**

Q. 18) Consider a backward shift operator D such that: $Dy_t=y_t-1$

The expression: $(1-D)(1-D^m)y_t$ is simplified to:

A. $y_t - y_{t-1} - y_{t-m} - y_{t-m-1}$
B. $y_t - y_{t+1} - y_{t-m} + y_{t-m-s}$
C. $y_t - y_{t-1} + y_{t-m} - y_{t-m-1}$
D. $y_t + y_{t-1} - y_{t-m} - y_{t-m-1}$
E. None of them

[2]

Q. 19) Consider a survival distribution function at age x:
$S_0(x) = P[X>x]$
Such that $S_0(0) = 1$
$S_0(\infty) = \lim_{x \to \infty} S_0(x) = 0$
Consider density and the respective cumulative function as $f_x(t)$ and $F_x(t)$ respectively such that:

$S_0(x) = 1 - \int_0^x f_0(z)dz$, and

Let $F_0(x) = 1-\left(1-\frac{x}{120}\right)^{\frac{1}{6}}$. **[4]**

Let $f_0(x) = f_0(x+t) / S_0(x)$ and let $\dot{e}_x$ be the expectation of T_x.

Assuming constant force of mortality and limiting age as 120, find the standard deviation of T_x at age 30.

A. 21.395
B. 24.547
C. 23.214
D. 22.458
E. 25.096

Q. 20) Consider the number of policies and number of deaths as follows:

Age last birthday	Policies in force as on date 'D'				
	2010	**2011**	**2012**	**2013**	**2014**
48	65444	69000	72017	73065	73841
49	58741	59877	60587	61578	62547
50	58354	59544	60587	61587	63544
51	56874	58970	59575	59874	60587

Age last birthday	Deaths				
	2010	**2011**	**2012**	**2013**	**2014**
48	54	59	57	60	61
49	60	62	65	59	58
50	61	62	63	64	63
51	62	58	59	64	64

Calculate the estimates of initial mortality rate at ages 48 and 50, assuming that average number of policies in force throughout is same as in force policies as on date 'D' and Poisson model of mortality.

A. 0.000732 and 0.001038 respectively
B. 0.000721 and 0.001039 respectively
C. 0.000823 and 0.001030 respectively
D. 0.000922 and 0.001031 respectively
E. 0.000913 and 0.001032 respectively **[4]**

Q. 21) A run-off life insurance product is getting lapsed at 0.3 per day. Assume policies are independent to each other, and that there are only 2 policies in-force. Calculate the Kolmogorov forward equations for $\frac{d}{dt}P_{21}(t)$:

A. $0.6P_{22}(t) - 0.3P_{21}(t)$
B. $0.3P_{22}(t) - 0.6P_{21}(t)$
C. $0.6P_{22}(t) + 0.3P_{21}(t)$
D. $0.3P_{22}(t) + 0.6P_{21}(t)$
E. $0.3P_{22}(t) - 0.3P_{21}(t)$ **[2]**

Q. 22) Consider a survival distribution function at age x:
$S_0(x) = P[X>x]$ **[4]**
Such that $S_0(0) = 1$

$S_0(\infty) = \lim_{x \to \infty} S_0(x) = 0$
Consider density and the respective cumulative function as $f_x(t)$ and $F_x(t)$ respectively such that:

$S_0(x) = 1 - \int_0^x f_0(z)dz$, and
Let $F_0(x) = 1-\left(1 - \frac{x}{117}\right)^{\frac{1}{5}}$.

Let $f_0(x) = f_0(x+t) / S_0(x)$ and let $\dot{e}_x$ be the expectation of T_x.

Assuming constant force of mortality and limiting age as 117, find the standard deviation of T_x at age 82.

A. 11.688
B. 7.684
C. 10.058
D. 9.701
E. 12.325

Q. 23) In 2024, an insurer follows a Poisson distribution ($\lambda = 2.4$) in number of claims for a 1-year policy and Pareto distribution in claim cost ($\alpha = 4$, $\theta = 1200$). Assume mutual independence in claims. In 2025, the cost of inflation is expected to inflate by 20% uniformly.
Find the deductible that the insurer can impose to ensure that the aggregate claim amount remains the same in 2025 also.

A. 82.78
B. 92.26
C. 81.44
D. 96.23
E. 90.23 **[3]**

Q. 24) In 2024, an insurer follows a Poisson distribution ($\lambda = 2.4$) in number of claims for a 1-year policy and Pareto distribution in claim cost ($\alpha = 4$, $\theta = 1200$). Assume mutual independence in claims. In 2025, the cost of inflation is expected to inflate by 20% uniformly.
Calculate the expected number of claims if the deductible is imposed.

A. 1.909
B. 2.423
C. 1.114
D. 1.882
E. 2.816 **[2]**

Q. 25) In 2024, an insurer follows a Poisson distribution ($\lambda = 2.4$) in number of claims for a 1-year policy and Pareto distribution in claim cost ($\alpha = 4$, $\theta = 1200$). Assume mutual independence in claims. In 2025, the cost of inflation is expected to inflate by 20% uniformly.
The insurer is planning to impose a maximum covered loss of 'M' along with a deductible of 30, such that aggregate claims is same as in 2025 also. Calculate 'M'.

A. 1424.5
B. 1492.9 **[3]**
C. 1693.8

D. 1596.2
E. 1422.9

Q. 26) Let 'S' be the aggregate claims for portfolio of a number of policies issued by an insurer, given that:

- Claim frequency, N with E[N] = 1, var[N] = 0.9
- Claim severity, Y with E[Y] = 10, var[Y] = 60

N & Y will respective distribution among any of the distributions – Binomial, negative Binomial and Poisson.

Calculate the standard deviation of S.

A. 12.25
B. 11.28
C. 14.24
D. 11.97
E. 12.90 **[2]**

Q. 27) Let 'S' be the aggregate claims for portfolio of a number of policies issued by an insurer, given that:

- Claim frequency, N with E[N] = 1, var[N] = 0.9
- Claim severity, Y with E[Y] = 10, var[Y] = 60

N & Y will respective distribution among any of the distributions – Binomial, negative Binomial and Poisson.

Find the distribution of 'N'.

A. Poisson with $\lambda = 1$
B. Binomial (10, 0.1)
C. Binomial (10, 0.2)
D. Negative Binomial (10, 0.2)
E. Negative Binomial (10, 0.1) **[2]**

Q. 28) An insurance in a motor insurance portfolio has the distribution for this portfolio as exponential distribution with mean of 850. The insurer has opted for a proportional reinsurance while retaining a 73% proportion. The distribution of insurer's net claim amount random variable:

A. $(1\text{-}850)^{-1}$
B. $(1\text{-}1164.4)^{-1}$
C. $(1\text{-}620.5)^{-1}$
D. $(1\text{-}896.1)^{-2}$
E. $(1\text{-}560.7)^{-2}$ **[2]**

Q. 29) An equity analyst has calculated the value of dependency parameter $\alpha = 2$, by fitting Gumbel copula to equity investment returns of 3 three equity instruments.
The probabilities of making loss over next time period 'T' on those three instruments is calculated as 4%, 7% and 18%. Calculate the probability that all three instruments make losses over time period 'T'.

A. 0.029428
B. 0.018947
C. 0.038823
D. 0.010958
E. 0.022328

[4]

Q. 30) In a clinical trial, the scientists were given a task to calculate the probability of a type of cancer in a person who was positively tested in a diagnostic test of cancer but was initially thought to have a 1% risk of cancer. The diagnostic test's accuracy is 80% of cancerous tumours and 90% of benign ones.
Calculate the probability of cancer which will be calculated by 95% of the scientists:

A. 8.3%
B. 7.5%
C. 9.2%
D. 10.8%
E. 14.1% **[4]**

Q. 31) Given an equation: $Y = \beta_0 + \beta_1 X + \varepsilon$.
Find the incorrect option:

A. β_1 is mean increase in Y with one-unit increase in X.
B. ε, the error term, is assumed to have independence of X.
C. β_0 is expected value of Y.
D. Method of least may be used to estimate β_0 and β_1.
E. None of these **[2]**

Q. 32) A data analyst named 'DA' runs a regression on a dataset of 'n' data points. DA then calculates a 90% confidence interval (a, b) on y for a given set of predictors. She also calculates a 90% prediction interval (p, q) on y for the same set of predictors. There are three statements given as: Select the correct options:

I. $\lim_{n \to Inf} (b - a) = 0$
II. $\lim_{n \to Inf} (q - p) = 0$
III. q-p > b-a

A. II and III
B. I and III
C. III only
D. II only
E. I only **[3]**

Q. 33) Determine incorrect statement about clustering:

A. It is used to find out structure in a dataset.
B. It is unsupervised learning
C. It finds homogenous groups among a dataset's observations.
D. It helps reduce dimensionality of dataset.
E. Number of clusters needs to be pre-specified. **[2]**

Q. 34) A simple linear regression model has sum of squares of residuals as $\sum_{i}^{25} e^2 = 874.54$ and $R^2 = 0.87$. Calculate the sum of squares for this model.

A. 6727.23
B. 8747.21
C. 9877.14
D. 7577.46
E. 4456.54 **[2]**

****