Institute of Actuaries of India

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

March 2022 Examination

INDICATIVE SOLUTION

Introduction
The indicative solution has been written by the Examiners with the aim of helping candidates. The solutions given are only indicative. It is realized that there could be other points as valid answers and examiner have given credit for any alternative approach or interpretation which they consider to be reasonable.

Solution 1:

i)

a) A stochastic process X_n is stationary if the joint distributions of the $X_{t1}, X_{t2}, \ldots, X_{tm}$ and $X_{t1+k}, X_{t2+k}, \ldots, X_{tm+k}$ are identical for all $t_1, t_2, \ldots, t_m$, $k+t_1, k+t_2, \ldots, k+t_m$ and all integers m.

b) The process is weakly stationery if the expectations E[Xt] are constant with respect to t and the covariances Cov(Xt , Xt +k) depend only on the lag k.
Comment: Variance is constant as well, thus give full credit in that case.

c) If t and t+u are in the set of permissible values, then the increment for time u will be Xt+u –Xt

d) For a discrete process the Markov property requires that: P[X t =x |Xt1 =x1, Xt 2=x2 Xtm= xm]=P[Xt =x | X tm=xm] for all times t1 <t2 <....<tm <t and all states x1 , x2 ,...., x m, x

e) A discrete time martingale X_n satisfies two conditions:

1.E $[|X_n|] < \infty$ for all n
2.E $[X_n | X_0, X_1, \ldots, X_m] = X_m$ for all m<n. (5)

ii)
a) Consider S n as given
P[S o = m]=1,
P[S 1 = m+1]=p, P[S 1 = m-1]=1-p,
P[S 2 = m+2]=p^2, P[S 2 = m]=2p(1-p), P[S 2 = m-2]=$(1-p)^2$
Price of share is non stationary [1]
b)

$$E(S_n) = E(S_0) + E\left(\sum_{i=1}^{n} X_i\right) = m + nE(X_i) = m + n(2p-1)$$

which is clearly dependent on n
V(S n) = 4np(1-p) Since $V(X_i)$=4p(1-p), i=1,2....n
Cov(S_r, S_{r+u}) r for each r=0,1,2....u=0,1,2...
Covariance does not depend on u
Therefore, price of share is not weakly stationary [2]
c) Process has independent increments [1]
d) Since process has independent increments it has Markov property [1]
e) E[S n+1 /S n] = S n +2p-1
The process is not Martingale, if p =0.5 then it can be Martingale [1]

(6)
[11 Marks]

Solution 2:

i) Row vector Pi is called as stationary probability distribution for a Markov chain with transition matrix P if the following conditions hold for all j in S:

- π = π P where π is row vector i.e. $\pi_j = \sum_{i \in S} \pi_i P_{ij}$
- $\pi_{i} >= 0$
- $\sum_{j \in S} \pi_j = 1$

[1.5]

In general a Markov chain need not have a stationary probability distribution, and if it exists it need not be unique.

[0.5]

(2)

ii) Solution : (D). (2)

iii)

$$P^2 = \begin{pmatrix} 0.960 & 0.040 & 0.000 & 0.000 \\ 0.06 & 0.910 & 0.030 & 0.000 \\ 0.010 & 0.000 & 0.985 & 0.005 \\ 0.020 & 0.000 & 0.000 & 0.980 \end{pmatrix}$$

$$* \begin{pmatrix} 0.960 & 0.040 & 0.000 & 0.000 \\ 0.06 & 0.910 & 0.030 & 0.000 \\ 0.010 & 0.000 & 0.985 & 0.005 \\ 0.020 & 0.000 & 0.000 & 0.980 \end{pmatrix}$$

$$= \begin{pmatrix} 0.924 & 0.075 & 0.001 & 0.000 \\ 0.113 & 0.831 & 0.057 & 0.000 \\ 0.020 & 0.000 & 0.970 & 0.010 \\ 0.039 & 0.001 & 0.000 & 0.960 \end{pmatrix}$$

[2]

$q_k = P[X_0 = k]$, k=1,2,3,4
q" = [0.5,0.25,0.15,0.10]

q'P^2 = [0.497,0.245,0.160,0.098]

[2]

Expected number if agents in each grade at the beginning of week 3 = [149, 74, 48, 29] [1]

(5)

iv)

$$(\pi_1\ \pi_2, \pi_3, \pi_4) = (\pi_1\ \pi_2, \pi_3, \pi_4) * \begin{pmatrix} 0.960 & 0.040 & 0.000 & 0.000 \\ 0.060 & 0.910 & 0.030 & 0.000 \\ 0.010 & 0.000 & 0.985 & 0.005 \\ 0.020 & 0.000 & 0.000 & 0.980 \end{pmatrix}$$

[1]

$\Pi_1 = 0.96\ \Pi_1 + 0.06\Pi_2 + 0.01\ \Pi_3 + 0.02\ \Pi_4$
$\Pi_2 = 0.04\ \Pi_1 + 0.91\Pi_2$
$\Pi_3 = 0.03\Pi_2 + 0.985\ \Pi_3$
$\Pi_4 = 0.005\ \Pi_3 + 0.98\ \Pi_4$

[2]

$0.04\ \Pi_1 - 0.09\ \Pi_2 = 0$

$0.03\ \Pi_2 - 0.005\ \Pi_3 = 0$

$0.02\ \Pi_4 - 0.005\ \Pi_3 = 0$

$\Pi_3 = 4\ \Pi_4$
$\Pi_2 = 2\ \Pi_4$

$\Pi_1 = 4.5\ \Pi_4$

$\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4 = 1$

$\Pi_4 = 1/11.5 = 0.0870$
$\Pi_1 = 4.5\ \Pi_4 = 0.3913$
$\Pi_2 = 2\ \Pi_4 = 0.1739$
$\Pi_3 = 4\ \Pi_4 = 0.3478$

Thus, no of employees in steady states are $[\pi_1\ \pi_2, \pi_3, \pi_{4]}$ * 300 = [117,52,104,26] [2]

(5)

[14 Marks]

Solution 3:

i) The base line hazard represents a female of age 30 with surgical procedure C adopted (2)

ii) male aged 40 with surgical procedure A
$Z_1 = 40\text{-}30 = 10$
$Z_2 = 1$
$Z_3 = 1$
$Z_4 = 0$
female aged 45 with surgical procedure B
$Z_1 = 45\text{-}30 = 15$
$Z_2 = 0$
$Z_3 = 0$
$Z_4 = 1$

[1]

$\lambda(t) = \lambda_0(t)e^{\beta Z^T}$ for life 1/ $\lambda(t) = \lambda_0(t)e^{\beta Z^T}$ for life 2

= exp^(10*0.003+1*0.021+1*-0.025+0) / exp^(15*0.003+0+0+1*0.015)
=exp^0.026 / exp^0.06
= 0.966571505 < 1

Survival period of male will be greater [2]

(3)

iii) For male, $Z_5 = 10$, $Z_6 = 0$

For female, $Z_5 = 0$, $Z_6 = 15$

[1]

$\lambda(t) = \lambda_0(t)e^{\beta Z^T}$ for life 1/ $\lambda(t) = \lambda_0(t)e^{\beta Z^T}$ for life 2

= exp^(10*0.003+1*0.021+1*-0.025+0+10*0.0021+0) / exp^(15*0.003+0+0+1*0.015+0+15*0.00033)
=exp^0.047 / exp^0.06495
= 0.98221 < 1
Survival period of male will be greater using model 2 as well [1]

(2)

iv) The likelihood ratio statistic is -2(L1 – L2) = - 2(- 64.34 – (-64.21)) = 0.26 has an asymptotic chi-square distribution with 6 - 4 =2 degrees of freedom.

Since 0.26<5.991 where is the 5% prob. Distribution point, there is no evidence to reject the null hypothesis that beta_5, $\text{beta}_6 = 0$.

(2)

Considering the above model 1 is preferable as it works with less no. of parameters without using any additional information.

[9 Marks]

Solution 4:

i) It is intuitively sensible to think that mortality is a smooth function of age. A purely practical reason for smoothing mortality data is that we will use the life table to compute financial equities, such as premiums for life insurance contracts.

To test smoothness, we need to calculate the third differences of the graduated rates {qx} or {ìx+½}. The criterion of smoothness usually used is that the third differences of the graduated quantities {qx} or {ìx+½} should;

(a). be small in magnitude compared with the quantities themselves; and
(b). progress regularly (3)

ii) Rationale of applying various statistical test for testing the adherence to data

a) Chi Square Test

Advantages:

- The chi square test is based on statistic, which represents the sum of square of the discrepancies between the actual and expected values for one group (with an appropriate weighting factor applied).
- A high value of statistic indicates that the over all discrepancy is quite large and would lead us to reject the model. A low value indicates that the observed data fit the model well. Thus, easy to interpret the results

Disadvantages:

- Important to identify number of degrees of freedom basis the use of the test
- Any specific trend which may have large financial impact, large deviations, bias over or below the average value for all age groups or any specific age group. All of this is not identified applied at aggregate level

[1]

b) Standardized Deviation Test

Advantages:

- If the graduated rates are not good fit the distribution will not be "centred correctly" and thus require to perform this test
- If there is heterogeneity within the age groups or deaths are not independent, the variance will be smaller or greater than what we expect would be if underlying model is correct and will require to perform this test
- In case of under graduation standardized deviations are expected to be tightly bunched. Conversely in case of over-graduation they would be too spread. Then Chi square test doesn't work

Disadvantages:

- Deviation is a good all round test that detects most of the problems that might be present in a graduation

[1]

c) Sign Test

Advantages:

If the graduated rates do not tend to be higher or lower than the crude rates on average, we would expect roughly half the graduated rate values to be above the crude rate and half below. So if there are *m* age groups. The number of above (or below) should have a B (m, ½) distribution. An excessively high number of positive and negative deviations will indicate that the rates are biased.

Disadvantages:

This is a qualitative test than a quantitative

[1]

d) Cumulative Deviation Test

Advantages:

Detects overall bias or long runs of deviations of the same sign

Disadvantages:

Many methods of graduation result in a cumulative deviation of zero as part of the fitting process, in which case this test cannot be applied

[1]

e) Grouping of sign Test

Advantages:

Under the null hypothesis that mortality is according to the graduated rates, the sign of individual deviations is independent and equally likely to be positive or negative. The numbers of positive and negative deviations might be reasonable according to Sign Test but we should still have doubt if the number of groups of positive signs (and consequently number of negative groups) is small.

Disadvantages:

Test may result into different conclusion depending upon positives or negatives are counted for grouping

[1]

f) Serial Correlation Test

Advantages:

If graduated rates are neither over-graduated nor under-graduated, we would expect the individual standardized deviations at consecutive ages to behave as if they were independent. However, if the graduated rates are over-graduated the graduated curve will tend to stay the same side of the crude rates for relatively long periods and, although there will be random variations in the numbers of deaths, we would expect the values of consecutive deviations to have similar values *ie* they will be positively correlated.

Disadvantage:

The grouping of sign test is more accurate in case of over-graduation as compared to serial correlation test as possibility of cancelling out the opposite correlation is there

[1]

(6)

iii) Roughness in the mortality might be due to random errors due to less credible data or could be intrinsic characteristic of the group (eg. Heterogeneity). Therefore, if this intrinsic roughness were removed in smoothening process, adopted rates would not reflect the true feature of the data. (1)

iv) If rates are considered not to be completely reflecting the true characteristic of the target group and hence judgments to be applied to arrive at mortality rates for pricing assurance and annuity products, then margin would be on higher side of the rates for assurance and on lower side of the rates for annuity. (2)

[12 Marks]

<u>Solution 5:</u>

i) Solution (D) - main sources of error in mortality forecasts (1)

ii) Solution : (A) - Mortality experience groupings (1)

iii) Solution: (C) - Terminologies related to mortality (1)

iv) Solution: (A) - Purpose of graduation (1)

[4 Marks]

Solution 6:

i) Let Ti be the time until default of bond i where i =1, 2, 3. We want to calculate the joint probability:

P(T1 ≤1,T2 ≤1,T3 ≤1) = C[u1,u2,u3], where ui = P[Ti ≤1] = 0.2 for i =1,2,3.

Using an independence copula

P(T1 ≤1,T2 ≤1,T3 ≤1) = C[u1,u2,u3] = u1*u2*u3 (1 Mark)
= 0.2 * 0.2 * 0.2 = 0.008 = 0.8% (2 Marks) (3)

ii) Let Ti be the time until default of bond i where i =1, 2, 3. We want to calculate the joint probability:

P(T1 ≤1,T2 ≤1,T3 ≤1) = C[u1,u2,u3], where ui = P[Ti ≤1] = 0.2 for i =1,2,3.

Using a Gumbel copula with parameter α = 2

P(T1 ≤1,T2 ≤1,T3 ≤1) = C[u1,u2,u3]

$c[u_1, u_2, u_3] = e^{\left\{-[(-\ln u_1)^2+(-\ln u_2)^2+(-\ln u_3)^2]^{\frac{1}{2}}\right\}}$ (1 Mark)

$= e^{\left\{-[3(-\ln 0.2)^2]^{\frac{1}{2}}\right\}}$

$= 0.0616 = 6.16\%$ (2 Marks) (3)

iii) Let Ti be the time until default of bond i where i =1, 2, 3. We want to calculate the joint probability:

P(T1 ≤1,T2 ≤1,T3 ≤1) = C[u1,u2,u3], where ui = P[Ti ≤1] = 0.2 for i =1,2,3.

Using a Clayton copula with parameter α = 2

P(T1 ≤1,T2 ≤1,T3 ≤1) = C[u1,u2,u3]

$$c[u_1, u_2, u_3] = [max(u_1^{-2} + u_2^{-2} + u_3^{-2} - 2,0)]^{-1/2}$$
$$= [\max(0.2^{-2} + 0.2^{-2} + 0.2^{-2} - 2,0)]^{-\frac{1}{2}}$$
$$= [\max(3 \times 0.2^{-2} - 1,0)]^{-\frac{1}{2}}$$
$$= 0.1170 = 11.70\%$$

Alternatively if the students used $max(u_1^{-2} + u_2^{-2} + u_3^{-2} - 1,0)$ as the material discussed only about the two asset cases, then also marks are awarded in full without any penalization.

$$c[u_1, u_2, u_3] = [max(u_1^{-2} + u_2^{-2} + u_3^{-2} - 1,0)]^{-1/2}$$
$$= [\max(0.2^{-2} + 0.2^{-2} + 0.2^{-2} - 1,0)]^{-\frac{1}{2}}$$
$$= [\max(3 \times 0.2^{-2} - 1,0)]^{-\frac{1}{2}}$$

$= 0.1162 = 11.62\%$ (= (3)

iv) The independent Copula exhibits no tail dependence. This is very unlikely as the three companies are from the same industry. So independent copula may be more inappropriate.

(3)

The Gumbel copula exhibits (non-zero) upper-tail dependence, the degree of which can be varied by adjusting the single parameter α. But, it exhibits no lower tail dependence. Hence, the Gumbel copula is appropriate if we believe that the three investments are likely to behave similarly as the term approaches ten years but not at early durations. This is unlikely to be the case. If one bond defaults early on, then it may be indicative of problems in the industry sector or the economy and so the other investments may also be likely to default early on.

On the other hand, Clayton copula exhibits (non-zero) lower-tail dependence, the degree of which can be varied by adjusting the single parameter α. Hence it is very much possible that due to some external economic or industry related factors, if one bond defaults in the early periods, the effect is felt on the others as well. Hence Clayton Copula is more appropriate to model the defaults.

[12 Marks]

Solution 7:

i)

$$\int_{15000}^{50000} f(x)dx = \Phi\left(\frac{\ln 50{,}000 - 10}{0.9}\right) - \Phi\left(\frac{\ln. 15000 - 10}{0.9}\right)$$
$$= \Phi(0.910865) - \Phi(-0.42688)$$
$$= 0.8188 - 0.3347$$
$$= 0.4841$$

(2)

ii)

$$\int_{0}^{15000} x \cdot f(x)dx = e^{10+\frac{1}{2}(0.9)^2}\left[\phi\left(\frac{\ln 15000 - 10}{0.9} - 0.9\right) - \phi(-\infty)\right]$$

$$= e^{10.405}[\Phi(-1.32688)]$$
$$= 3047.279$$

(2)

iii)

$$\int_{50000}^{\infty} x^2 f(x)dx = e^{21.62} + 2(0.9)^2\left[\phi(\infty) - \phi\left(\frac{\ln 50000 - 10}{0.9} - 2(0.9)\right)\right]$$
$$= e^{21.62}[1 - \phi(-0.88914)]$$
$$= e^{21.62}(0.8130)$$
$$= 1{,}993{,}222{,}739$$

(2)

[6 Marks]

Solution 8:

i) Solution :(C) (2)

ii) Solution: (C) (2)

iii)

- Catastrophic risks are the ones where large number of people are exposed to the risk of a large loss by reason of the occurrence of a peril
- Low frequency and high severity events
- The ‘true’ distribution of catastrophic data is more leptokurtic (more peaked with fatter tails) and more skewed than the normal distribution
- Parameter estimates are inappropriately influenced by the main bulk of the data in the middle of the distribution
- Extreme Value Theory needs to be applied instead of Central Limit Theorem
- Models like Block Maxima or Peaks over threshold needs to be implemented (3)

- The distributions that can be evaluated include the GEV family of distributions (Gumbel, Weibull and Freshet) or Generalized Pareto (GPD) distributions as they are good at modelling the extreme values

[7 Marks]

Solution 9:

i) Solution: (C) (3)

ii) Solution: (C) (4)

iii) Solution: (A) (2)

iv) Solution : (B) (2)

v) Solution: (A) (2)

[13 Marks]

Solution 10:

i)

$$\begin{aligned} y_t = \alpha_1 y_{t-1} &+ \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \epsilon_t \\ \mathrm{Cov}(y_t, y_{t-1}) &= \mathrm{Cov}(\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \epsilon_t, y_{t-1}) \\ \gamma_1 &= \alpha_1 \gamma_0 + \alpha_2 \gamma_1 + \alpha_3 \gamma_2 + 0 \end{aligned}$$

Dividing both sides by γ_0 gives autocorrelation on LHS at lag 1

$$p_1 = \alpha_1 + \alpha_2 p_1 + \alpha_3 p_2 - - - - - - - - (1)$$
$$\Rightarrow p_1(1 - \alpha_2) = \alpha_1 + \alpha_3 \rho_2$$
$$\mathrm{cov}(y_t, y_{t-2}) = \mathrm{cov}(\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \epsilon_t, y_{t-2})$$
$$\gamma_2 = \alpha_1 \gamma_1 + \alpha_2 \gamma_0 + \alpha_3 \gamma_1 + 0$$

Dividing both sides by γ_0

$$p_2 = \alpha_1 p_1 + \alpha_2 + \alpha_3 p_1$$
$$\Rightarrow \quad p_2 = (\alpha_1 + \alpha_3) p_1 + \alpha_2 - - - - - - - - (2)$$

Substituting (2) in (1)

$$\begin{aligned} p_1(1 - \alpha_2) &= \alpha_1 + \alpha_3[(\alpha_1 + \alpha_3) p_1 + \alpha_2] \\ &= \alpha_1 + \alpha_3(\alpha_1 + \alpha_3) p_1 + \alpha_3 \alpha_2 \\ \rho_1\big(1 - \alpha_2 - \alpha_3(\alpha_1 + \alpha_3)\big) &= \frac{\alpha_1 + \alpha_3 \alpha_2}{\rho_1} = \frac{\alpha_1 + \alpha_3 \alpha_2}{1 - \alpha_2 - \alpha_3(\alpha_1 + \alpha_3)} \\ &= \left.\frac{(\alpha_1 + \alpha_3)[\alpha_1 + \alpha_2 \alpha_3}{1 - \alpha_2 - \alpha_3(\alpha_1 + \alpha_3)}\right] + \alpha_2 \end{aligned}$$

(5)

ii)

$$\rho_1 = \frac{0.3 + 0.09}{1 - 0.3 - 0.3(06)}$$
$$\frac{0.39}{0.52} = 0.75$$

$$\begin{aligned} p_2 &= 0.6p_1 + \alpha_2 \\ &= 0.6(075) + 0.3 \\ &= 0.45 + 0.3 \\ &= 0.75 \end{aligned}$$

(2)

iii) PACF at Lag' $= ACF$ at Lag $' = 0.75$

PACF at Lag2 $= \frac{\rho_2 - p_1^2}{1 - p_1^2} = \frac{0.75 - 0.75^2}{1 - 0.75^2} = 0.428571$ (2)

iv) The process is a AR(3) process and hence the PACF will be significant for lags up to 3. The PACF for Lags 1 and 2 is 0.75 and 0.428 are significant. From Lag 4 onwards, the PACF values will become insignificant (1.5 Marks)

There is no MA order in the process. In an AR process, the ACF gradually falls to zero. Here ACF at lag 1 and lag 2 are 0.75 but they will gradually fade off after a few more lags. Higher-lag autocorrelations will satisfy the Yule-Walker equation $\rho_k = 0.3\ (\rho_{k-1} + \rho_{k-2} + \rho_{k-3})$. So the values will tail off quite quickly to zero, always taking positive values. (1.5 Marks) (3)

[12 Marks]

****