Statistical and Risk Modelling - 2

Q1

Answer:

(i)

A stochastic model refers to the type of model wherein the input variables follow a time dependent random phenomenon.

The advantages of a stochastic model over a deterministic model are as follows:

a. A stochastic model allows the randomness of the input variables to be incorporated in the model which is more practical relative to a deterministic model which incorporates a deterministic i.e., a constant value at a particular point of time.
b. Stochastic models allow the use of advanced numerical methodologies like Monte Carlo simulation, which is an extremely powerful technique for solving complex problems.

(ii)

For a time in-homogeneous model, the transition rates are a function of time t. It is certainly possible to improve the fit by using a time in-homogeneous model in this instance. However, if the age profile is represented by a density function f(a), then the overall average rate at which a healthy employee falls sick is = $\int f(a)\sigma(a)da$, roughly constant for all time t. The same of course applies to the overall average rate of recovery.

(iii)

Required Proof:

Given that the non-overlapping increments of the stochastic process are independent, $X_{t+u} - X_t$ for every u > 0 are independent of past values X_m and are non-overlapping.

Thus,

$$Y = P[X_t = a \,|X_{s_1} = x_1, X_{s_2} = x_2, \dots, X_{s_n} = x_n, X_s = x]$$

Because, $X_s = x \rightarrow X_t = X_t - X_s + x$

$$Y = P[X_t - X_s + x = a \,|X_{s_1} = x_1, X_{s_2} = x_2, \dots, X_{s_n} = x_n, X_s = x]$$

Since, $X_t - X_s$ are independent of past values of X, therefore $X_t - X_s + x$ will also be independent of past values of X_m.

Therefore,

$$Y = P[X_t - X_s + x = a \,|X_s = x]$$

$$Y = P[X_t = a \,|X_s = x] \dots \dots .. (X_s = x)$$

Thus, independent increments satisfy Markov Property.

Q2

Answer:

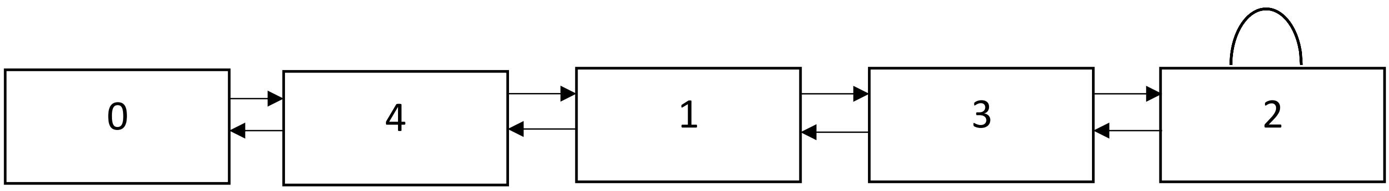
(i)

Consider a Markov chain taking values in the set S = {i:i = 0,1,2,3,4}, where i represents the number of umbrellas in the place where the actually currently is (at home or office).

If i = 1 and it rains, then it takes the umbrella, moves to the other place where there are already 3 umbrellas, and including the one he brings, therefore, he will now have four umbrellas. Thus, $p_{1,4} = p$ because p is the probability of raining

If i = 1, but it does not rain, then, he does not take the umbrella, goes to the other place and find 3 umbrellas, thus, $p_{1,3} = 1 - p = q$.

However, if i = 0, he must move to another place where four umbrellas are kept with probability $p_{0,4} = p$. Similarly for other states, the process is depicted by the following diagram.



(ii)

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & q & p \\ 0 & 0 & q & p & 0 \\ 0 & q & p & 0 & 0 \\ q & p & 0 & 0 & 0 \end{bmatrix}$$

(iii)

Using the condition of stationarity, $\pi P = \pi$,

Where, P is the transition probability matrix.

The equations formed include,

$$\pi_0 = \pi_4 * q \qquad -1$$

$$\pi_1 = \pi_3 * q + \pi_4 * p \qquad -2$$

$$\pi_2 = \pi_2 * q + \pi_3 * p \qquad -3$$

$$\pi_3 = \pi_1 * q + \pi_2 * p \qquad -4$$

$$\pi_4 = \pi_0 + \pi_1 * p \qquad -5$$

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \qquad -6$$

Solving the above equations we get,

$$\pi_0 = \frac{q}{4+q}\ , \pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4+q}$$

(iv)

He will get wet every time he happens to be in state 0 and it rains.

The chance he is in stage 0 is π_0. The chance it rains is p. Hence, probability that he gets wet is,

Probability of getting wet = Probability of being in state zero and it's raining.

$$\text{Probability of getting wet} = \frac{\mathrm{q}}{4+\mathrm{q}} * p$$

(v)

Given that p = 0.6,

$$\text{Probability of getting wet} = 0.0545$$

If he wants the chance to be less than 1%, then clearly he needs more umbrellas. So, suppose he needs N umbrellas. Setting up the Markov chain as above and generalizing it is clear that:

$$\pi_1 = \pi_2 = \pi_3 = \pi_4 = \cdots = \pi_N = \frac{\pi_0}{q}$$

$$But, \pi_0 + \pi_1 + \pi_2 + \pi_3 + \cdots + \pi_N = 1$$

Thus,

$$\pi_N * q + N * \pi_N = 1$$

$$\pi_N = \frac{1}{N+q}$$

Now, the probability of getting wet = $p * \pi_0 = \frac{p+q}{N+q}$

For, Probability of getting wet $< 1/100$

$$\frac{p+q}{N+q} < \frac{1}{100}$$

$$N > 100 * p * q - q$$

Substituting values of p and q,

N > 23.6

Thus, it is not worth to keep 24 umbrellas instead of four to reduce the probability of getting wet from 6% to 1%. On the days he did not take the umbrella but it starts raining, he may buy a cheap umbrella from a nearby local shop, or take a cab or any other mode of transport to office or borrow (share) an umbrella from some friend or colleague using the same route, or wait for the rain to subside, etc.

Q3

Answer:

(i)
(ii)

Process	State Space	Time Set
Counting Process	Discrete	Either
General Random Walk	Either	Discrete
Poisson Process	Discrete	Continuous
Markov Chain	Discrete	Discrete
Markov Jump Process	Discrete	Continuous

(iii)
(a) Poisson Process
(b) Markov Jump Process
(c) Poisson Process
(d) Markov Jump Process

Q4

Answer:

(i)

Past history of the sportsmen is required to decide where to go further in the Markov chain. If a sportsman is at A and his performance reduces, we need to know what level of performance he was at the previous year to determine whether he or she drops one or two levels in the chain.

(ii)

The B state spaces needs to be split in two states

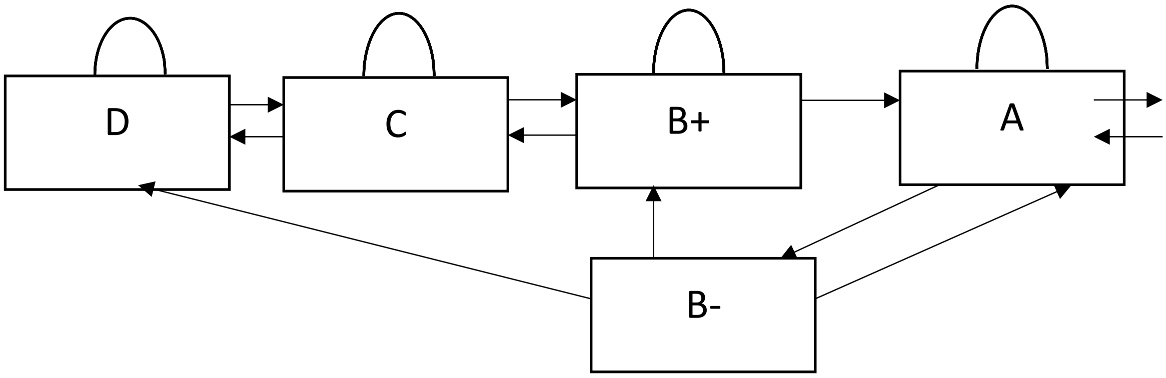
Let, B+ be the level with no reduction in performance parameter last year.

B- be the level with reduction in performance parameter last year.

These levels are for modelling purposes and not the actual levels.

(iii)

The transition graph is as follow:



(iv)

The transition probability matrix is given follow:

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.3 & 0.2 & 0.5 & 0 & 0 \\ 0 & 0.3 & 0.2 & 0 & 0.5 \\ 0.3 & 0 & 0.2 & 0 & 0.5 \\ 0 & 0 & 0 & 0.3 & 0.7 \end{bmatrix}$$

(v)

Using the condition of stationarity, $\pi P = \pi$,

The equations formed are:

$$0.5\pi_D + 0.3\pi_C + 0.3\pi_{B-} = \pi_D \quad -1$$

$$0.5\pi_D + 0.2\pi_C + 0.3\pi_{B+} = \pi_c \quad -2$$

$$0.5\pi_C + 0.2\pi_{B+} + 0.2\pi_{B-} = \pi_{B+} \quad -3$$

$$0.3\pi_A = \pi_{B-} \quad -4$$

$$0.5\pi_{B+} + 0.5\pi_{B-} + 0.7\pi_A = \pi_A \quad -5$$

$$\pi_A + \pi_{B+} + \pi_{B-} + \pi_C + \pi_D = 1 \quad -6$$

Solving the above equations we get,

$$\pi_A = 0.42445$$

$$\pi_{B-} = 0.12733$$

$$\pi_{B+} = 0.12733$$

$$\pi_C = 0.15280$$

$$\pi_D = 0.16808$$

(vi)

The long run average contract value is,

$$100\% * \pi_D + 120\% * \pi_C + 150\% * (\pi_{B+} + \pi_{B-}) + 175\% * \pi_A$$

Substituting the values,

$$Long\ run\ average\ contract\ value = 1.4762\ million\ USD$$

(vii)

Let m_i be the number of transitions (and years) taken to reach level A from any current level i. Then,

$$m_D = 1 + 0.5m_D + 0.5m_C \quad -1$$

$$m_C = 1 + 0.2m_C + 0.5m_B + 0.3m_D \quad -2$$

$$m_{B+} = 1 + 0.2m_{B+} + 0.5m_A + 0.3m_C \quad -3$$

$$m_{B-} = 1 + 0.3m_D + 0.5m_A + 0.2m_{B+} \quad -4$$

$$m_A = 0 \quad -5$$

Solving the above equations, we get,

$$m_c = 7.12\ years$$

$$Thus, m_D = 2 + m_C = 9.12\ years$$

Q5

Answer:

(i)

There is an explicit dependence on the past behaviour of $\{Y_j : j \leq n\}$, in the probability distribution of Y_{n+1}; further X_n is nothing but the sum of Y_j. Hence, the Marco property does not hold.

(ii)

In the above part, we show that there is explicit dependence on the past behaviour of $\{Y_j : j \leq n\}$ and hence, the Markov property does not hold, this implies that the sequence $\{Y_n ; n \geq 1\}$ does not form a Markov chain.

(iii)

The transition matrix is given as follows:

$$P = \begin{bmatrix} p & 1-p & 0 & . & . & 0 \\ 0 & pe^{-\lambda} & 1-pe^{-\lambda} & 0 & . & . \\ . & 0 & pe^{-2\lambda} & 1-pe^{-2\lambda} & 0 & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & . & . & . & . & . \end{bmatrix}$$

(iv)

(a) The chain is time homogeneous since the transition probabilities calculated in part (i) is independent of time ‘n’.

(b) It is irreducible, since the number of errors can never go down.
(c) There are no recurrent states; hence there can be no stationary distribution.

(v)

Probability of no further error is,

$$\left(pe^{-j\lambda}\right)^n = p^n * e^{-nj\lambda}$$

Q6

Answer:

(a)

A stochastic model allows for the randomness of the input parameters.

The advantages of a stochastic model over a deterministic model are as follows:

a. A stochastic model allows the randomness of the input variables to be incorporated in the model which is more practical relative to a deterministic model which incorporates a deterministic i.e., a constant value at a particular point of time.
b. Stochastic models allow the use of advanced numerical methodologies like Monte Carlo simulation, which is an extremely powerful technique for solving complex problems.

(b)

The following subparts are not there as proof.

Q7

Answer:

(a)

The process may be expressed as a Markov chain by considering the following states. Each state indicates the current number of successive defeats.

Let,

1 = 0 – Not defeated last time

2 = 1 – Defeated in the last match

3 = 2 – Defeated in the last 2 matches

4 = 3 – Defeated in the last 3 matches

5 = 4 – Captain sacked

The transition matrix is as follows:

$$P = \begin{bmatrix} 0.7 & 0.3 & 0 & 0 & 0 \\ 0.7 & 0 & 0.3 & 0 & 0 \\ 0.7 & 0 & 0 & 0.3 & 0 \\ 0.7 & 0 & 0 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b)

A Markov chain is said to be irreducible if any state j can be reached from any state i. The above process is not irreducible as the captain, once sacked, can never become the captain again.

(c)

(Calculation was done in rough)

The probability of remaining the captain for exactly four matches is 0.0081

The probability of remaining the captain for exactly five matches is 0.00567

The probability of remaining the captain for exactly seven matches is 0.00567

The probability of remaining the captain for exactly nine matches is 0.0056241

(d)

Solving the sub part as a stationary distribution approach, we get,

N1 = 174.94

N2 = 171.60

N3 = 160.49

N4 = 123.46

Thus, E[N] = 174.94 matches.

Q8

Answer:

(i)

The score currently stands at tie. Whoever wins the next point will move into a lead. If the player in lead wins the subsequent point as well, he would win the tiebreaker. However, if the player in lead loses the next point, the score would be back to tie.

Since, the probability of moving to the next state does not depend on the history prior to entering the state, Markov property holds.

The state space is defined as follows.

Let, T is equal to Tie

L_F is equal to Federer leads

L_N is equal to Nadal leads

G_F is equal to Federer wins.

G_N is equal to Nadal wins.

(ii)

The transition matrix is set out below:

$$P = \begin{bmatrix} 0 & 0.55 & 0.45 & 0 & 0 \\ 0.45 & 0 & 0 & 0.55 & 0 \\ 0.55 & 0 & 0 & 0 & 0.45 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(iii)

The chain is reducible as it has two absorbing states – G_F and G_N.

The chain is aperiodic.

(iv)

The probability of returning to tie after two points is given by = 0.55*0.45 + 0.45*0.55 = 0.495

We need to find the number of such cycles of returning to tie such that,

$$0.495^N = 1 - 0.95$$

$$N = 4.26$$

Since, the game can finish in cycles of two points, the required number of cycles is 5 i.e, 10 points.

(v)

After two points:

a. Nadal may have won the tie-breaker
b. Federer may have won the tie-breaker
c. Tie-breaker may have come back to tie

Let F_T be the probability that Federer wins and let N_T be the probability that Nadal wins.

We have,

$$N_T = \ 0.2025 + 0.495 * N_T$$

Solving we get,

$$N_t = 0.401$$

Thus,

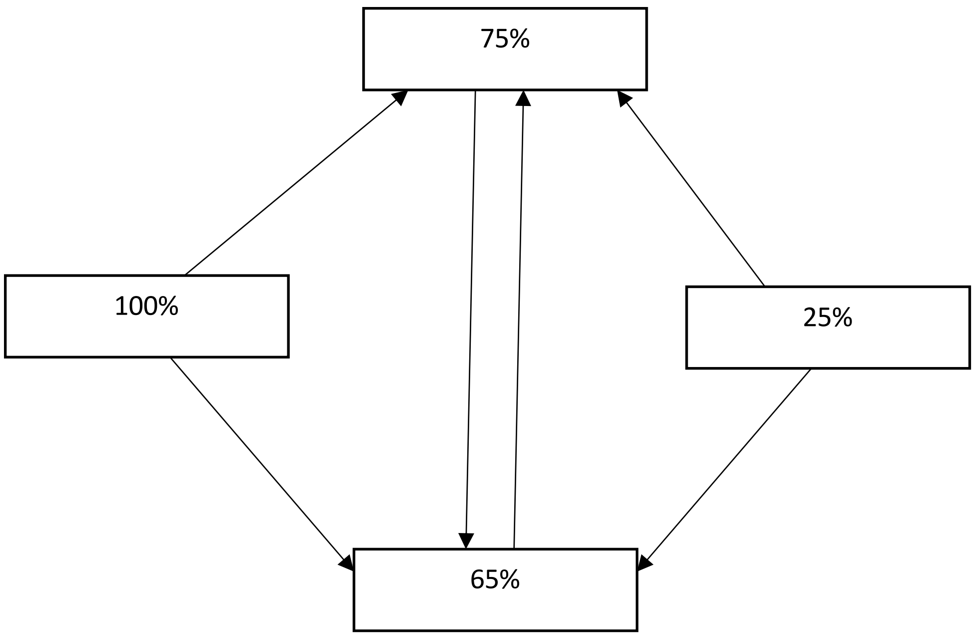
$$F_T = 0.599$$

Q9

Answer:

(a)

The state transition diagram is set out below:



(b)

Using the stationary distribution condition,

(Calculation in rough)

x1 = 0, x2 = 0.5423, x3 = 0.3559 and x4 = 0.1017

(c)

The steady state proportion of customers having a 65% target allocation for XYZ Ltd is 35.59%.

It is noted that initially this proportion is 45% and it quickly drops to:

37.5% in one year's time and 36.2% in two years' time.

Q10

Answer:

(Not: This question is similar to Question 2)

Q11

Answer:

(i)

If each room is represented by the state, then the transition matrix P for this Markov chain is as follows:

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{bmatrix}$$

(ii)

The chain is irreducible, because it is possible to go from any state to any other state. However, it is not aperiodic, because for any even n, P(6,1) will be zero and for any odd n, P(6,5) will also be zero. This means that there is no power of P that would have all its entries strictly positive.

(iii)

For P to be stationary,

$$\pi p = \pi$$

We need to satisfy the above condition.

(iv)

We find that the mean recurrence time (i.e., the expected time to return) for the room 1 is $\frac{1}{\pi(1)} = 12$

(v)

Expected time to return to room 1 after calculating turns out to be 7.

Q12

Answer:

We wish to find $m_{2,4}$,

Now,

$$m_{i4} = 0, if\ i = 4\ and\ 1 + summation(p_{ij} m_{j4}), if\ i\ is\ noy = t\ equal\ to\ 4$$

Thus, $m_{(24)} = 1 + (1 * m_{34})$

$$m_{44} = 0$$

$$m_{34} = 1 + \left(\frac{1}{3} * m_{44}\right) + (\frac{2}{3} * m_{24})$$

$$m_{34} = 1 + 0 + \frac{2}{3} * (1 + m_{34})$$

$$m_{34} = 5$$

Hence,

$$m_{24} = 1 + 5 = 6\ steps$$

(ii)

$$P(1,2,3,2,3,4) = P(X_0 = 1) * (p_{12}) * (p_{23}) * (p_{32}) * (p_{23}) * (p_{34})$$
$$= \frac{3}{4} * \frac{3}{5} * 1 * \frac{2}{3} * 1 * \frac{1}{3} = \frac{1}{10}$$

Q13

Answer:

(i)
(a)

Data: RSRRSSSSRRSRSSRSRSRRSSSRRSSRRR
Prr = 6/14 = 3/7

Prs = 8/14 = 4/7

Psr = 8/15

Pss = 7/15

(b)

Probability tat it will rain on 3rd July 2020 – 0.142222 + 0.130612 + 0.130612 + 0.078717 = 0.482164

(ii)
(a)

Here, mu = 3

‘Some policies’ means ‘1 or more policies’ i.e., 1 minus the ‘zero policies’ probability

P(X >0) = 1 – P(X = 0)

It follows Poisson distribution

Therefore, the probability of 1 or more policies is given by = 1 – P(X = 0) = 0.95021

(b)

The probability of selling 2 or more, but less than 5 policies is:

P(2 <= X < 5) = P(X = 2) + P(X = 3) + P(X = 4) = 0.61611

(c)

Average number of policies sold per day: 3/5 = 0.6

So, on a given day, P(X) = 0.32929.