

Institute of Actuaries of India

Subject CT4 – Models

November 2013 Examinations

INDICATIVE SOLUTIONS

Solution 1 :

- Multiple state model

In the multiple-state model, the MLE is consistent and asymptotically unbiased; the variance of the estimator is also only available asymptotically. Simulation experiments suggest that the results are reasonable if $dx \geq 10$.

- Poisson model

In the Poisson model, the MLE is consistent and unbiased. Its mean and variance are available exactly in terms of the true μ , but are estimated from the data by the same expressions as estimate the asymptotic mean and variance in the two-state model.

- Binomial model

In the “naive” binomial model, in which N identical lives are observed for exactly one year, the MLE is consistent and unbiased, and the exact mean and variance can be obtained in terms of the true qx . In practice, the data rarely conform to the “naive” model, so only approximate results are available.

In the case of human mortality where μ is small, there are few reasons to prefer any one of these models on the basis of the statistical properties of the MLEs alone. Therefore, we can consider any of the above models for modelling human mortality. However, other considerations such as availability of data, purpose of investigation etc. may be more relevant in this case to finalise the model to consider.

(4 Marks)**Solution 2 :**

- i) For any u such that $s < u < t$ we have,

$$P_{ij}(s, t) = \sum_k P_{ik}(s, u)P_{kj}(u, t)$$

(1 Mark)

- ii) The Chapman-Kolmogorov equations can be written in the form

$$P_{ij}(s, t + h) = P_{ij}(s, t)P_{jj}(s, t + h) + \sum_{k \neq j} P_{ik}(s, t)P_{kj}(t, t + h)$$

Now we can apply the definition of transition rate. We know that for a small time interval h , we have:

$$P_{jj}(t, t + h) = 1 + h\mu_{jj}(t) + o(h)$$

where the $o(h)$ term accounts for the fact that there is a very small chance of more than one transition during the period $(t, t + h)$. Similarly

$$P_{kj}(t, t+h) = h\mu_{kj}(t) + o(h)$$

Therefore,

$$P_{ij}(s, t+h) = P_{ij}(s, t)(1 + h\mu_{jj}(t)) + \sum_{k \neq j} P_{ik}(s, t)h\mu_{kj}(t) + o(h)$$

$$\frac{P_{ij}(s, t+h) - P_{ij}(s, t)}{h} = \frac{\sum_k P_{ik}(s, t)h\mu_{kj}(t) + o(h)}{h}$$

Taking limits $h \rightarrow 0$, then gives the desired result.

(4 Marks)
[Total Marks-5]

Solution 3 :

1 marks to be given for points which generally cover a brief comparison between deterministic and stochastic models

3 marks to be given for specific instances

Markers are encouraged to give one mark for every sensible example provided by the candidate where deterministic models could be more appropriate than stochastic models.

Examples could be:

- Modelling products where the risk profile does not justify investments in stochastic models (e.g. unit linked products without investment guarantees)
- for projects where time constraints wouldn't permit developing stochastic models
- where statistically credible data is not available for calibrating a stochastic model

(4 Marks)

Solution 4:

i.

Let X_n be the number of shoes at the front door. Then $X_n \in \{0, 1, 2, 3, \dots, x\}$ is a Markov chain.

The probability of the Samit jogging without shoes would be equal to:

50% x Probability that there are no shoes at the front door when he leaves (in a steady state)
+
50% x Probability that there are no shoes at the back door when he leaves (in a steady state)

The transition matrix is:

$$\begin{bmatrix} .75 & .25 & & & & & \\ .25 & .5 & .25 & & & & \\ & .25 & .5 & .25 & & & \\ & & .25 & .5 & .25 & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ & & & .25 & .5 & .25 & \\ & & & & .25 & .75 & \end{bmatrix}$$

$$\pi_0 = .75\pi_0 + .25\pi_1 \dots\dots\dots (1)$$

$$\pi_1 = .5\pi_1 + .25\pi_0 + .25\pi_2 \dots\dots\dots(2)$$

$$\pi_2 = .5\pi_2 + .25\pi_1 + .25\pi_3 \dots\dots\dots (3)$$

$$\pi_{x-1} = .5\pi_{x-1} + .25\pi_x + .25\pi_{x-2} \dots\dots\dots(x)$$

$$\pi_x = .75\pi_x + .25\pi_{x-1} \dots\dots\dots(x+1)$$

From eqn (1) and (x+1) we get

$$\pi_1 = \pi_0$$

And

$$\pi_{x-1} = \pi_x$$

From other equations we get ,

$$\pi_1 = \pi_2 = \pi_3 = \dots = \pi_{x-2} = \pi_{x-1}$$

Since,

$$\pi_0 + \pi_1 + \pi_2 + \dots + \pi_{x-2} + \pi_{x-1} + \pi_x = 1$$

$$(x + 1)\pi_0 = 1$$

$$\pi_0 = \frac{1}{(x + 1)}$$

Likewise,

$$\pi_x = \frac{1}{(x + 1)}$$

(2 Marks)

ii.

The probability of Samit jogging without shoes is therefore $\frac{1}{(x+1)}$

(being equal to $0.5 \cdot \frac{1}{(x+1)} + 0.5 \cdot \frac{1}{(x+1)}$)

(5 Marks)
[Total Marks-7]

Solution 5 :

- i) The occurrence of hurricanes in a county can be modeled as a Poisson process. Since, twenty hurricanes have touched down in a county within the last twenty years the mean number of hurricane per year 1.

Next year will be a "hurricane year" if one or more hurricane occur during the year.

The probability of the same is $P(X \geq 1) = 1 - P(X = 0)$ where X follows Poisson distribution with mean 1.

Which is $1 - \exp(-1) = 0.632$

[2 Marks]

- ii) The probability of two "hurricane year" in the next three years is equivalent to there will be two years when in each year at least one hurricane is expected to touch down.

The probability of the same is

$$\binom{3}{2} \times 0.632^2 \times (1 - 0.632) \\ = 0.441$$

[3 Marks]

- iii) The expected number of hurricanes over ten years is 10.

[1 Mark]

- iv) 0.632 is the probability of "hurricane year" in any year. The expected number of 'hurricane years' in ten years time $0.632 \times 10 = 6.32$

[1 Mark]

[Total Marks-7]

Solution 6 :

- i) The one step transition probability matrix

$$\begin{bmatrix} .5 & .5 & & & \\ .5 & & .5 & & \\ .5 & & & .5 & \\ .5 & & & & .5 \\ .5 & & & & & .5 \end{bmatrix}$$

(1 Mark)

- ii) The two- step transition probability matrix

$$\begin{bmatrix} .5 & .5 & & & \\ .5 & & .5 & & \\ .5 & & & .5 & \\ .5 & & & & .5 \\ .5 & & & & & .5 \end{bmatrix} \begin{bmatrix} .5 & .5 & & & \\ .5 & & .5 & & \\ .5 & & & .5 & \\ .5 & & & & .5 \\ .5 & & & & & .5 \end{bmatrix}$$

$$\begin{bmatrix} .5 & .25 & .25 & & \\ .5 & .25 & & .25 & \\ .5 & .25 & & & .25 \\ .5 & .25 & & & .25 \\ .5 & .25 & & & .25 \end{bmatrix}$$

The steady state equation is $\pi P = \pi$

$$\begin{aligned} .5(\pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5) &= \pi_1 \\ .5\pi_1 &= \pi_2 \\ .5\pi_2 &= \pi_3 \\ .5\pi_3 &= \pi_4 \\ .5(\pi_4 + \pi_5) &= \pi_5 \end{aligned}$$

(2 Marks)

iii)

The normalizing condition

$$\begin{aligned} \sum \pi &= 1 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 &= 1 \end{aligned}$$

$$\pi_1 = \frac{1}{2}, \quad \pi_2 = \frac{1}{4}, \quad \pi_3 = \frac{1}{8}, \quad \pi_4 = \frac{1}{16}, \quad \pi_5 = \frac{1}{16}$$

(4 Marks)

[Total Marks-7]

Solution 7 :

i. Definitions:

- a** $F_x(t) = P[T_x \leq t] \quad (0 \leq x \leq \omega)$
is the distribution function of T_x .
- b** $S_x(t) = P[T_x > t] = 1 - F_x(t) \quad (0 \leq x \leq \omega)$
is the survival function of T_x .

(2 Marks)

ii. In actuarial terms,

$${}_tq_x = F_x(t); \text{ and}$$

${}_tq_x$ represents the probability that a person aged x will die within the next t years. For example, ${}_{20}q_{40}$ represents the probability that a 40 year old dies before his/her 60th birthday.

$${}_tp_x = 1 - {}_tq_x = S_x(t)$$

${}_t p_x$ represents the probability that a person aged x will live for at least another t years. For example, ${}_{20} p_{40}$ represents the probability that a 40 year old will live to see his/her 60th birthday.

(2 Marks)

iii. The probability density function, $f_x(t)$ can be expressed as:

$f_x(t) = \frac{d}{dt} F_x(t)$. Then:

$$f_x(t) = \frac{d}{dt} P[T_x \leq t]$$

$$= \lim_{h \rightarrow 0^+} \frac{1}{h} \times (P[T_x \leq t + h] - P[T_x \leq t])$$

$$= \lim_{h \rightarrow 0^+} \frac{1}{h} \times (P[T \leq x + t + h | T > x] - P[T \leq x + t | T > x])$$

$$= \lim_{h \rightarrow 0^+} \frac{P[T \leq x + t + h] - P[T \leq x] - (P[T \leq x + t] - P[T \leq x])}{S(x) \times h}$$

$$= \lim_{h \rightarrow 0^+} \frac{P[T \leq x + t + h] - P[T \leq x + t]}{S(x) \times h}$$

Now multiply and divide by $S(x + t)$ and we have:

$$f_x(t) = \frac{S(x + t)}{S(x)} \times \lim_{h \rightarrow 0^+} \frac{1}{h} \frac{P[T \leq x + t + h] - P[T \leq x + t]}{S(x + t)}$$

$$= S_x(t) \times \lim_{h \rightarrow 0^+} \frac{1}{h} P[T \leq x + t + h | T > x + t]$$

$$= S_x(t) \times \mu_{x+t}$$

Or, in actuarial notation, for a fixed age x between 0 and ω :

$$f_x(t) = {}_t p_x \mu_{x+t} \quad (0 \leq t < \omega - x)$$

(3 Marks)

iv. Expectation of life in Edgestone

Given that nearly all deaths occur due to people falling off the cliff, assume that the only cause of death in Edgestone is due to an accident.

Further assume that people of all ages are equally prone to accidents – in particular that of falling off the cliff in Edgestone.

Therefore, it follows that we can assume that the mortality rate for all ages are equal, i.e. survival probability over a given year is age independent.

The probability of death in any given year, for any given age, $q_x = 1/8 = 0.125$.

Probability of death over 18 years = $1 - (1 - 0.125)^{18} = 0.909605$

Probability of both girls dying in 18 years = $0.909605^2 = 0.827381$

Probability that at least one of the girls survive to see her 18th birthday = $1 - 0.827381 = 0.172619$

(4 Marks)
[Total Marks-11]

Solution 8 :

- i. Census approximation is used when the exact dates of entry and exit from observation are not recorded, but rather in-force or census data is available as at a particular date for each year (for example total number of policies at the end of each year).

The census approximation requires that:

$$E_x^c = \int_K^{K+N+1} P_{x,t} dt \cong \sum_{t=K}^{K+N} 0.5 \times (P_{x,t} + P_{x,t+1})$$

This is based on the underlying assumption that $P_{x,t}$ is linear between census dates.

(2 Marks)

- ii. Exposed to risk for 1 January 2011 – 31 December 2011, Age 45 last birthday

Company A:

IF policies @ 01 Jan 2011, Aged 45 last birthday = $0.5 \times (5,920 + 5,993) = 5,956.5$

IF policies @ 31 Dec 2011, Aged 45 last birthday = $0.5 \times (5,911 + 5,988) = 5,949.5$

$E_{45}^c = 0.5 \times (5,956.5 + 5,949.5) = 5,953$

Company B:

IF policies @ 31 Mar 2010, Aged 45 last birthday = 3,939

IF policies @ 31 Mar 2011, Aged 45 last birthday = 3,921

$E_{45}^c = 0.25 \times 3,939 + 0.75 \times 3,921 = 3,925.5$

Company C

IF policies @ 01 Jan 2011, Aged 45 last birthday = 9,237

IF policies @ 31 Dec 2011, Aged 45 last birthday = 9,252

$E_{45}^c = 0.5 \times (9,237 + 9,252) = 9,244.5$

(8 Marks)
[Total Marks-10]

Solution 9 :

Null Hypothesis: The graduated rates for estimating mortality are consistent with the standard table rates.
The key test statistics and ISDs are set out below:

Age	Ex	Graduated rates	Standard table	Graduated dx	Standard dx	Deviation	z	z ²
40	100,000	0.001504	0.001803	150.4	180.3	-29.9	-2.2268	4.9585
41	100,000	0.001674	0.001959	167.4	195.9	-28.5	-2.0362	4.1462
42	100,000	0.001874	0.002140	187.4	214.0	-26.6	-1.8183	3.3064
43	100,000	0.002106	0.002350	210.6	235.0	-24.4	-1.5917	2.5334
44	100,000	0.002378	0.002593	237.8	259.3	-21.5	-1.3352	1.7827
45	100,000	0.002696	0.002874	269.6	287.4	-17.8	-1.0500	1.1024
46	100,000	0.003066	0.003197	306.6	319.7	-13.1	-0.7327	0.5368
47	100,000	0.003494	0.003567	349.4	356.7	-7.3	-0.3865	0.1494
48	100,000	0.003985	0.003983	398.5	398.3	0.2	0.0100	0.0001
49	100,000	0.004539	0.004444	453.9	444.4	9.5	0.4506	0.2031
50	100,000	0.005154	0.004946	515.4	494.6	20.8	0.9353	0.8747
								19.5937

i. Standardised deviations test

There are 11 independent sample values with ISDs tabulated above. Assuming a normal approximation for standardised deviations, the expected and actual distribution of ISDs can be set out as follows:

Range	$(-\infty, -3)$	$(-3, -2)$	$(-2, -1)$	$(-1, 0)$	$(0, 1)$	$(1, 2)$	$(2, 3)$	$(3, \infty)$
Actual	0	2	4	2	3	0	0	0
Expected	0	0.2	1.5	3.7	3.7	1.5	0.2	0

Observations:

- Overall, the distribution of ISDs is not what we might expect if normally distributed.
- There are 8 negative deviations and only 3 positive. We would expect the number of positive and negative deviations to be evenly spread. This indicates a possible negative bias in the graduation.
- There are a disproportionately higher number of deviations than expected in the range $(-3, -1)$ whereas none in the range $(1, 3)$.
- There are no deviations lying outside the range $(-3, 3)$ indicating no outliers.

- The standardised deviations do not appear to conform to the standard normal distribution, thus indicating that the graduated mortality rates do not conform to the standard table rates.

Chi-square test

The test statistic for the Chi-square test as tabulated above = 19.6

Under the null hypothesis this test statistic has a χ^2 distribution with 11 degrees of freedom.

This is a one-sided test. The upper 5% point of χ^2_{11} is 19.7

Since the value of test statistic is less than 19.7, we accept the null hypothesis under this test.

Grouping of signs test

The number of positive signs is 3 and the number of negative signs is 8. There is only one group each of positive runs and negative runs – therefore it is clear from the deviations that there is grouping of deviations of the same sign.

Consequently, we reject the null hypothesis under the grouping of signs test as well.

Conclusions

From the above tests, we can conclude that:

- Overall, the graduated rates indicate a reasonable fit to the standard table rates based on the results of the Chi-square test
- However, the graduation is negatively biased below age 47 and positively biased above age 48.
- Moreover, there is a clear single group of negative bias and a single group of positive bias which would be a concern.

Therefore, based on the above, it may be recommended that further investigation is necessary into the graduated rates to ensure that the above flaws are remedied.

(10 Marks)

Solution 10:

i. Kaplan-Meir estimate of the survival function:

Old version:

j	t_j	d_j	n_j	$\hat{\lambda}_j = d_j/n_j$	$(1 - \hat{\lambda}_j)$	$c_j = \prod_{k=1}^j (1 - \hat{\lambda}_k)$	$1 - c_j$
1	1	1	49	0.02041	0.97959	0.97959	0.02041
2	3	2	46	0.04348	0.95652	0.93700	0.06300
3	7	1	40	0.02500	0.97500	0.91358	0.08642
4	8	1	38	0.02632	0.97368	0.88953	0.11047
5	15	1	30	0.03333	0.96667	0.85988	0.14012
6	18	2	26	0.07692	0.92308	0.79374	0.20626
7	22	1	20	0.05000	0.95000	0.75405	0.24595
8	25	1	16	0.06250	0.93750	0.70692	0.29308
9	28	2	12	0.16667	0.83333	0.58910	0.41090

10	31	1	7	0.14286	0.85714	0.50495	0.49505
----	----	---	---	---------	---------	---------	---------

From the above, $\hat{F}_{KM}(t) = 0.24595$ for $22 \leq t < 25$

Therefore, the proportion of laptops that need replacement during warranty period ~25%

New version:

j	t_j	d_j	n_j	$\hat{\lambda}_j = d_j/n_j$	$(1 - \hat{\lambda}_j)$	$c_j = \prod_{k=1}^j (1 - \hat{\lambda}_k)$	$1 - c_j$
1	2	1	48	0.02083	0.97917	0.97917	0.02083
2	3	1	46	0.02174	0.97826	0.95788	0.04212
3	9	1	39	0.02564	0.97436	0.93332	0.06668
4	17	1	30	0.03333	0.96667	0.90221	0.09779
5	21	1	25	0.04000	0.96000	0.86612	0.13388
6	25	1	20	0.05000	0.95000	0.82281	0.17719
7	26	2	18	0.11111	0.88889	0.73139	0.26861
8	29	1	13	0.07692	0.92308	0.67513	0.32487
9	31	1	10	0.10000	0.90000	0.60762	0.39238

From the above, $\hat{F}_{KM}(t) = 0.13388$ for $21 \leq t < 25$

Therefore, the proportion of laptops that need replacement during warranty period ~13%

(6 Marks)

- ii. The Chief Engineer has commented on the total lifetime of the laptops. However, the warranty is applicable only for 2 years, therefore it is more relevant to consider the expected failure within the warranty period only rather than looking at the total lifetime on the laptops.

Further, the Chief Engineer has not made any allowance for the difference in cost of replacement between the old and new versions. It is possible that the newer versions are more expensive to replace, therefore even with the same expected lifetime and failure rates, the replacement for newer versions could be more expensive for the company.

(3 Marks)

- iii. From Greenwood's formula:

$$\text{var}[\hat{F}_{KM}(24)] \approx (1 - \hat{F}_{KM}(24))^2 \sum_{t_j \leq 24} \frac{d_j}{n_j(n_j - d_j)}$$

j	t_j	d_j	n_j	$\frac{d_j}{n_j(n_j - d_j)}$
1	2	1	48	0.0004433
2	3	1	46	0.0004831
3	9	1	39	0.0006748
4	17	1	30	0.0011494
5	21	1	25	0.0016667
				0.0044172

$$\text{var}[\hat{F}_{KM}(24)] \approx (1 - 0.13398)^2 \times 0.0044172 = 0.0043380$$

Assuming a Normal approximation is appropriate, we can construct a 95% confidence interval as follows:

$$(0.13398 \pm 1.96 \times \sqrt{0.0043380}) = (0.48\%, 26.30\%)$$

Average cost of replacement = 26.30% x Rs 35,000 = Rs 9,205.

Therefore, the mark-up required is Rs 9,205 on a Rs 50,000 laptop, i.e. 18.41%.

(6 Marks)

[Total Marks-15]

Solution 11:

i)

a) Let p be the probability of winning and P_{300} be the probability of eventually accumulating the target amount given that the Ajoy starts with Rs.300.

$$P_i = p P_{i+1} + (1 - p)P_{i-1}$$

$$\text{and } P_0 = 0 \text{ and } P_{800} = 1$$

$$P_{100} = p P_{200} + (1 - p)P_0$$

$$P_{100} = p P_{200}$$

$$P_{200} = p P_{300} + (1 - p)P_{100}$$

$$P_{200} - P_{100} = \frac{P_{100}}{p} - P_{100} = \frac{(1 - p)P_{100}}{p}$$

Let us also define $\rho = \frac{1-p}{p}$

$$P_{200} = p P_{300} + (1 - p)P_{100}$$

$$P_{300} - P_{200} = \frac{(1 - p)(P_{200} - P_{100})}{p}$$

Expanding on the similar line,

$$P_{200} - P_{100} = \rho P_{100}$$

$$P_{300} - P_{200} = \rho(P_{200} - P_{100}) = \rho^2 P_{100}$$

$$P_{300} - P_{200} = \rho(P_{200} - P_{100}) = \rho^2 P_{100}$$

$$P_{400} - P_{300} = \rho(P_{300} - P_{200}) = \rho^3 P_{100}$$

.....

$$P_{800} - P_{700} = \rho(P_{700} - P_{600}) = \rho^{8-1}P_{100}$$

Taking the sum of the equations we get

$$P_{800} - P_{100} = \rho(P_{700} - P_{100}) = (\rho + \rho^2 + \rho^3 + \dots + \rho^7) P_{100}$$

$$P_{800} = (1 + \rho + \rho^2 + \rho^3 + \dots + \rho^{m-1}) P_{100}$$

$$P_{800} = \frac{1 - \rho^m}{1 - \rho} P_{100}$$

Now we know that $P_{800} = 1$

$$\text{Hence, } P_{100} = \frac{1 - \rho}{1 - \rho^m}$$

$$\text{and } P_{300} = \frac{1 - \rho^3}{1 - \rho^8} = 0.207$$

(5 Marks)

b)

Since Ajoy is starting with 300 and he places a bet of the amount required to achieve the target sum of Rs. 800 subject to maximum at his disposal, the eventual stake depends on the available amount at the point of time, and the minimum amount required to achieve the target sum.

Hence, the possible stakes at various point of time are different. To begin with, Rs. 300 is the stake when Ajoy has Rs. 300 at his disposal.

If he wins, he moves up to Rs.600, with stake required not more than Rs.200.

If he loses he would have lost all his money and can't place any more bets and game ends.

It wins from Rs.600 by placing a bet of Rs.200 he achieves the target sum of Rs.800 and game ends.

If he loses from Rs.600, he goes down Rs.400.

Now from Rs.400 he can place a maximum possible bet of Rs.400 to achieve the target sum.

Whether he wins or loses, in both cases the game ends as Ajoy either achieves his target sum of Rs.800 or loses his entire kitty.

These are two possible combinations of stakes during the series of bets before the Ajoy either lost all his money or achieves the target sum of Rs.800.

Given p is the probability of winning, there are two possible ways he can reach the desired sum without losing all his money as explained above.

In case of first one, where he reaches Rs.800 by two consecutive win, the probability is p^2

In case of the second one, where he reaches the desired sum by losing the second bet but the winning the third one, the probability is $p(1 - p)p$.

Hence, the probability of winning by opting the second option is $p^2 + p^2(1 - p)$

which is 0.314.

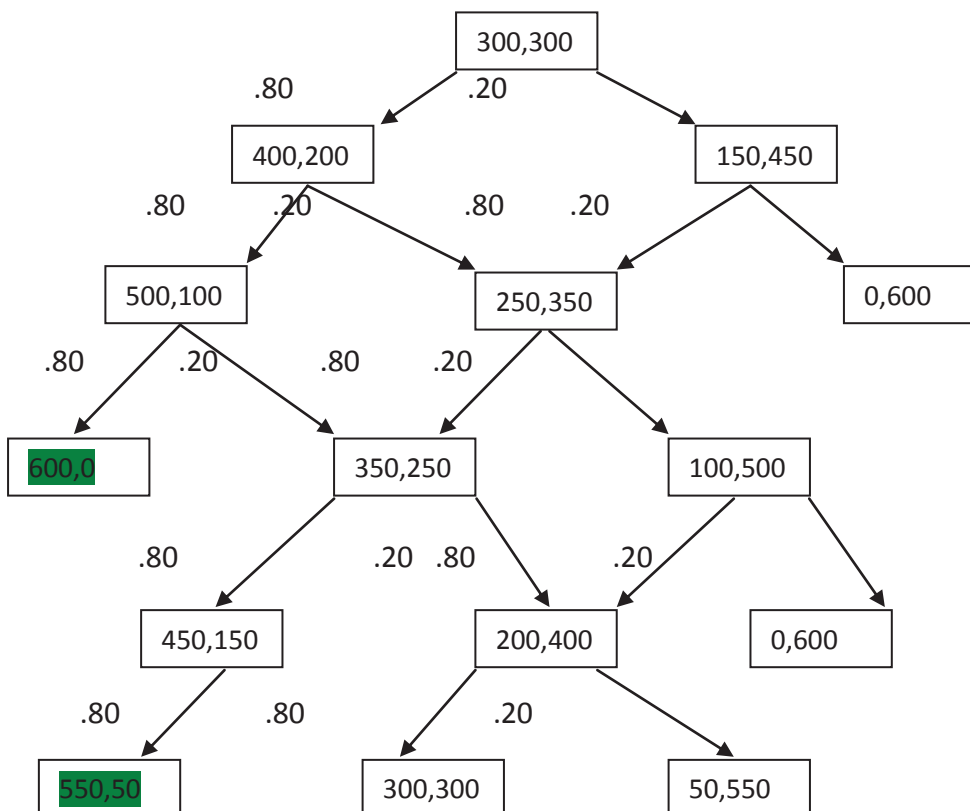
(6 Marks)

ii. Hence, second option gives better chance of winning.

(1 Mark)

iii. Starting against friend A, the game progresses in the following manner

((x,y) represents the amount of money(in INR) with Ajoy and his friend A at various point of the game and arrow captures probability of moving from one stage to another with either Ajoy winning or A.)



We are required to calculate the probability of reaching the highlighted cells, from which Ajoy can continue playing against B to have a chance of reaching the desired amount of Rs 800.

The probability of reaching stage (600,0) is 0.80^3

The probability of reaching stage (500,50) is $3 \times 0.80^4 \times 0.20$

Since with B Ajoy can play up to maximum of three rounds,

Starting from

600,0

There is only one possible way Ajoy can reach 800 by winning two consecutive bets with a probability of 0.90^2 .

Hence, on this route the probability of reaching the desired sum is $0.80^3 \times 0.90^2$

Starting from

550,50

also there is only one possible way Ajoy can reach 800 by winning three consecutive bets with a probability of 0.90^3 .

Hence, on this route the probability of reaching the desired sum is

$$3 \times 0.80^4 \times 0.20 \times 0.90^3$$

Hence, the overall probability of getting in to the stadium is 0.594

(8 Marks)
[Total Marks-20]
