Institute of Actuaries of India

Subject CT4 –Models

September 2018 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)

- Define objectives for modeling process.
- Plan modeling process and how it will be validated.
- Collect and analyze the necessary data.
- Define parameters and its values.
- Define the model capturing the essence of the real world system.
- Involve experts on the real world system and get feedback to validate.
- Decide on the software/simulation package, choose reliable random number generator.
- Code and debug the computer program.
- Test output for reasonableness.
- Check sensitivity of the model.
- Ensure that any relevant professional guidance or standards have been complied with.
- Document the result and communicate appropriately. [4]

ii)

- A stochastic model allows for random variation in its input parameters.
- A stochastic model gives distribution of outcomes depending upon the distribution of inputs scenarios.
- A model that does not contain any random component is deterministic in nature.
- The output from deterministic model is single point estimate for a given set of input parameters and hence it is just a simplified case of stochastic model. [2]

iii)

- Stochastic model can reflect reality as accurately as possible by recognizing random nature of the variables involved.
- A stochastic model can provide information about the distribution of the outcomes along with various statistics such as probability and variance to allow the user to take appropriate decisions.
- Some problems are too complicated to be solved using analytical methods, stochastic models allows the use of Monte Carlo Simulations which is an extremely powerful method of solving complicated problems.
- Stochastic model requires lot of assumptions about the distribution of the input variables. There may not be sufficient or credible data to derive these assumptions.
- Both building and running stochastic model is time consuming, a simple model may be more useful in some circumstances.
- Expertise is required to build and update stochastic model, availability of expertise may be an issue or it may be expensive especially for small companies.

[3]

[9 Marks]

Solution 2:

i) A Poisson process is a counting process in continuous time $\{N_{t,}\ t \geq 0\}$, where N_t records the number of occurrences of a type of event within the time interval from 0 to t.

Events occur singly and may occur at any time;

the probability that an event occurs during the short time interval from time t to time t + h is approximately equal to λh for small h, where the parameter λ is the rate of the Poisson process.

OR

A Poisson process is an integer valued process in continuous time $\{N_t, t\square 0\}$, where

$$\Pr[N_{t+h} - N_t = 1|F_t] = \lambda * \mathrm{h} + \mathrm{o(h)}$$

$$\Pr[N_{t+h} - N_t = 0|F_t] = 1 - \lambda * \mathrm{h} + \mathrm{o(h)}$$

$$\Pr[N_{t+h} - N_t \#0,1|F_t] = \mathrm{o(h)}$$

and o(h) is such that $\lim_{h\to 0} \frac{o(h)}{h} = 0$

OR

A Poisson process with rate λis a continuous-time integer-valued process $N_t, t \geq 0$ with the following properties:

$N_t, t = 0$

N_t has independent increments

N_t has Poisson distributed stationary increments

P$[N_t - N_s = n] = \frac{[\lambda(t-s)]^n e^{-\lambda(\mathrm{t-s})}}{n!}$, s<t, n=0,1,… [2]

ii) Probability of n faulty toys

$$P(D = n) = \frac{\lambda^n e^{-\lambda}}{n!}$$

Where λ = 4 per hour (one faulty toy in 15 min, hence, 4 faulty toys per hour)

$\mathrm{P(D = 0)} = \frac{4^0 e^{-4}}{0!}$

= 0.018 [1]

iii) The exponential distribution of the holding times of Poisson process has memoryless property. Hence the next faulty toy is likely to be found in 15 min. [1]

iv) The supervisor spends 1.5 minutes checking each toy, so 40 toys checked per hour. This means 4 faulty toys per hour, or 2/3 faulty toy per 10 minutes. So

$$\Pr[D \geq 2] = 1 - \Pr[D < 2] = 1 - e^{-2/3} - (2e^{-2/3}/3) = 0.145.$$

[2]

[6 Marks]

Solution 3:

i) Type I – The duration at which the observations will be censored are specified in advance.

Type II – Observations continues until a pre-determined number/proportion of individuals have experienced the event of interest. [2]

ii) Informative Censor: mortality investigation of life office data where the individuals who lapsed their policies are expected to be in better health than the individuals who continue their policies.

[1]

iii) Examples of interval censoring:

- When we know only the calendar year of death/withdrawal
- When we know only the calendar year of birth [1]

iv)

- Right Censoring – All students who enrolled on 2014 and not qualified till 2017 ie study ends or the student qualify or student discontinue the membership or got married
- Random Censoring – Time when a student qualifies or getting married is unknown.
- Left Censoring is not present as the exact date of joining the institute known. However presence of left censoring may be felt if students are joining with examination credits from other institutes or educational bodies. Left censoring would happen only if the previous examination credits conferred exemptions.
- Type I censoring is present as there is end date of study and there is no presence of Type II censoring.

[4]

[8 Marks]

Solution 4:

i) **Force of Mortality**: The force of mortality μ_x is the instantaneous rate of mortality at age x. It is defined by the equation

$$\mu_x = \lim_{h \to 0^+} \frac{1}{h} * P[\, T \ll x + h \, | T > x]$$

Initial rate of mortality:

q_x is the initial rate of mortality and it is the probability that a life alive at exact age x (the initial time) dies before exact age x+1.

Central rate of mortality:

The Central rate of mortality is defined by the equation

$$m_x = \frac{q_x}{\int_0^1 t\text{þ}x\, dt}$$

The quantity m_x is the probability of dying between exact ages x and x+1 per person year lived between exact ages x and x+1. The denominator $\int_0^1 t\text{þ}x\, dt$ is interpreted as the expected amount of time spent alive between ages x and x+1 by a life alive at age x, and the numerator is the probability that life dying between exact ages x and x+1. [3]

ii) Let μ' be the shedding rate for the first twenty days.
The tree sheds 50% of leaves in the first 20 days.

Hence ${}_{20}p_0 = 0.5$

Then EXP (- μ' * 20) = 0.5

Solving the above equation we get μ' = 0.034668
Let μ'' be the shedding rate for the next ten days.
The tree sheds 50% of leaves in the next 10 days.
Hence ${}_{30}p_0 = 0.25$
However, ${}_{30}p_0 = {}_{20}p_0 * {}_{10}p_{20}$

Then EXP (- μ'' * 10) = 0.25/0.5
Solving the above equation we get μ'' = 0.06931

[3]

iii) If the tree sheds leaves according to Gompertz Law, then $\mu_x = B c^x$
Given that $\mu_{30} = 0.01256$ and $\mu_{40} = 0.02198$

Then μ_{40}/μ_{30} = (c ^ 10) = (0.02198 / 0.01256)

c = 1.057557

Substituting c in μ_{30} we get B = 0.00234356
And g = EXP(- B/ $\log_e c$)

g = 0.958987

The number of days further required for the tree to shed 95% of its leaves can be found using the formula ${}_xp_0 = 0.05$

The LHS can be written as ${}_{30}p_0 * {}_xp_{30}$ and the value of ${}_{30}p_0$ is 0.25 . Then
${}_xp_{30} = 0.05/0.25 = 0.2$

$\Rightarrow g^{(c^{30})} * g^{(c^x - 1)} = 0.2$ Using Gompertz Law

$\Rightarrow g^{(c^x)} = 0.240057362$

$\Rightarrow x = 63.05$ or 63 days [6]

[12 Marks]

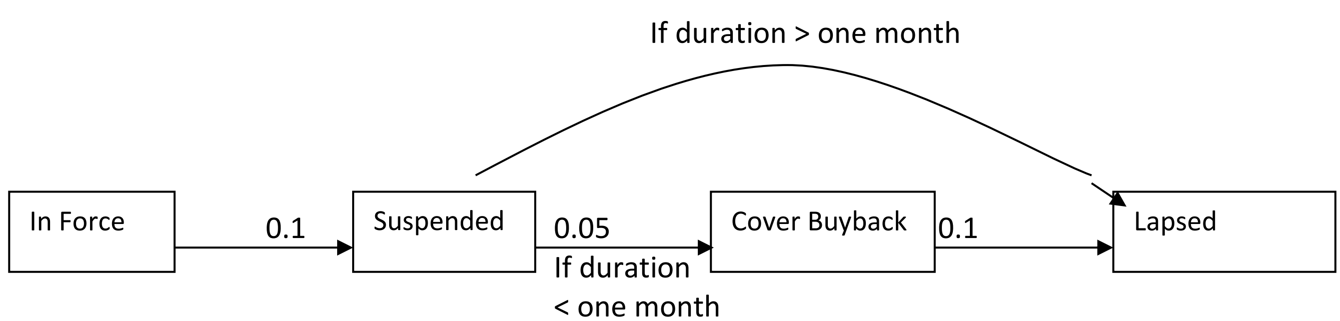
Solution 5:

i) A time inhomogeneous model should be used.
Transition probabilities out of the suspended state depend upon the duration of the stay in suspended state. For example, the probability of stay in suspended state is exp(-duration*.05) if duration < one month, however the probability of stay in suspended stay reduces to Zero once the buyback premiums are not paid within the stipulated time and the policy compulsorily move to lapse state.

[2]

ii) There are three states viz In-Force, Suspended, Lapsed.
However, a model with these state spaces would not satisfy the Markov property because a policy can only be reinstated once, so if it is In-Force we would need to know if the policy has any claim in the past.
A Markov model could be obtained by including one more state i.e. Cover Buyback which does not require information about past claims and the future transitions will depend only on the state currently occupied and duration. [3]

iii)



[2]

[7 Marks]

Solution 6:

i) Role of Graduation in producing life table

- Produce smooth set of results that are suitable for a particular purpose
- Remove the random sampling errors
- Use the information available from adjacent ages to improve the reliability of estimates. This is particularly important at the older ages where exposure numbers are small and data are sparse.

Procedure for graduating the rates:

- Choose a method of graduation. The preferable method is the graduation by parametric formula as the availability of data is large.
- Then select a statistical model to provide the graduated rates. This step involves selecting a graduation formula and determine the parametric values and then calculate the graduated rates.
- Test the graduated rates – the graduated rates must be compared with original data to see if they are acceptably close according to some graduation tests. Some general checks like whether mortality of male at a particular age is higher than the mortality of female at that age can also be performed in this stage.
- The process will be repeated with different models and/or with different number of parameters till we reach a suitable level of graduated rates. The final choice will be influenced by the level of goodness of fit (adherence to data) required against the smoothness required in the rates produced. As the purpose is to produce a national life table for general use, the balance will be tilted towards goodness of fit. [5]

ii) Shape of the mortality curve over the range of ages and the level of mortality rates. [1]

iii) This is a test for over graduation. The above test detects grouping of signs of deviations. It does this by analyzing the relationship between the deviations at nearby ages taking into account the magnitude of the values. This test will address the problem of the inability of the chi square test to detect excessive clumping of deviations of the same sign. [2]

iv)

Age	E	A	z_x	z_{x+1}	$A = z_x - \bar{z}$	$B = z_{x+1} - \bar{z}$	A2	AB
60	38.3	36	-0.37165	-1.01415	-0.40343	-1.04593	0.162758	0.421964
61	40.45	34	-1.01415	-0.69317	-1.04593	-0.72496	1.093978	0.758262
62	42.52	38	-0.69317	-0.68311	-0.72496	-0.7149	0.525569	0.518275
63	44.56	40	-0.68311	0.6332	-0.7149	0.601412	0.511083	-0.42995
64	47.63	52	0.6332	-0.45398	0.601412	-0.48577	0.361697	-0.29215
65	51.25	48	-0.45398	0.221781	-0.48577	0.189993	0.23597	-0.09229
66	55.35	57	0.221781	0.585212	0.189993	0.553424	0.036098	0.105147
67	60.45	65	0.585212	0.726943	0.553424	0.695155	0.306278	0.384716
68	63.22	69	0.726943	0.755509	0.695155	0.723721	0.483241	0.503099

69	67.78	74	0.755509	0.643078	0.723721	0.61129	0.523772	0.442403
70	71.56	77	0.643078		0.61129		0.373676	0

z_x is calculated using the formula z_x = (A-E) / SQRT (E) derived from the Poisson Distribution assumption.

Now r(j) = $\sum_{i=1}^{m-j}(z_x - \bar{z}) * (z_{x+1} - \bar{z})$/ { (m-j)/m*$\sum_{i=1}^{m}(z_x - \bar{z})$ ^ 2 }

⇒ r_1 = 0.55296

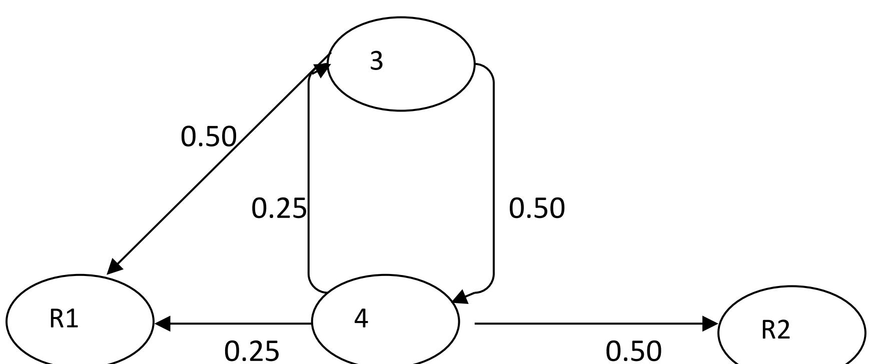
And the T ratio is r_1 * Sqrt (11) = 1.83

The T Ratio is positive which suggest that the rates are over graduated and value 1.83 < 1.96 and hence we do not reject the Null Hypothesis that the graduated rates are the true rates underlying the observed data. [5]

[13 Marks]

Solution 7:

i) We can draw the transition matrix using the absorbing groups R1 and R2 as follows:



Let α_i defines the probability of absorption in R1 from current state i

$\alpha_i = P(absorption\ in\ R1 \mid X_0 = i)$

By definition, $\alpha_{R1} = 1$ and $\alpha_{R2} = 0$

Now,

$$\alpha_3 = \frac{1}{2}\alpha_{R1} + \frac{1}{2}\alpha_4$$

$$= \frac{1}{2} + \frac{1}{2}\alpha_4$$

And,

$$\alpha_4 = \frac{1}{4}\alpha_{R1} + \frac{1}{4}\alpha_3 + \frac{1}{2}\alpha_{R2}$$

$$= \frac{1}{4} + \frac{1}{4}\alpha_3$$

Solving both the above equation,

$$\alpha_3 = \frac{5}{7}$$

Therefore if X_0 = 3, then the probability of absorption in R1 is $\frac{5}{7}$. [5]

ii) Let T be the first time the chain visits R1 or R2

$$t_i = E[T / X_0 = i]$$

By definition,

$$t_{R1} = t_{R2} = 0$$

$$t_3 = 1 + \frac{1}{2}t_{R1} + \frac{1}{2}t_4$$

$$= 1 + \frac{1}{2}t_4$$

And,

$$t_4 = 1 + \frac{1}{4}t_{R1} + \frac{1}{4}t_3 + \frac{1}{2}t_{R2}$$

$$= 1 + \frac{1}{4}t_3$$

Solving both the above equations,

$t_3 = \frac{12}{7}$, Therefore if X_0 = 3, it will take on average $\frac{12}{7}$ steps until the chain gets absorbed in R1 or R2.

[5]

[10 Marks]

<u>Solution 8:</u>

i) <u>Principle of Correspondence</u>:

A life alive at time t should be included in the exposure at age x at time t if and only if, were that life to die immediately, he or she would be counted in the death data d_x at age x.

[2]

ii) Central death rate for the year 2016 is

State I - 264/{0.5*(15000+20000)} = 0.015086
State II – 445/ { 0.5*(32000+27000)} = 0.015085

[2]

iii) The mortality experience of the life office is consistent with the census data for both Male & Female and for both Rural & Urban.

Let $\mu_{1\,r}$= Death rate for Rural population of State 1
$\mu_{1\,u}$= Death rate for Urban population of State 1

Then $0.6 * \mu_{1\,r} + 0.4 * \mu_{1\,u} = 0.015086$

$\Rightarrow 0.6*1.4* \mu_{1\,u} + 0.4 * \mu_{1\,u} = 0.015086$
$\Rightarrow \mu_{1\,u} = 0.012166$ & $\mu_{1\,r} = 0.017032$

Let $\mu_{1\,ro}$= Death rate for Rural Male population of State 1
$\mu_{1\,r1}$= Death rate for Rural Female population of State 1

Then $0.4* \mu_{1\,ro} + 0.6* \mu_{1\,r1} = 0.017032$

$\Rightarrow 0.4* 1.2* \mu_{1\,r1} + 0.6* \mu_{1\,r1} = 0.017032$
$\Rightarrow \mu_{1\,r1} = 0.015771$ &$\mu_{1\,ro} = 0.018925$

Central death rate for Rural Female in State I is 0.015771

Let $\mu_{1\,uo}$= Death rate for Urban Male population of State 1
$\mu_{1\,u1}$= Death rate for Urban Female population of State 1

Then $0.6* \mu_{1\,uo}+ 0.4* \mu_{1\,u1}= 0.012166$

$\Rightarrow 0.6* 1.2* \mu_{1\,uo}+ 0.4* \mu_{1\,u1}= 0.012166$
$\Rightarrow \mu_{1\,u1}= 0.010862$ & $\mu_{1\,ro} = 0.013035$

Central death rate for Urban Male in State I is 0.013035

Let $\mu_{2\,r}$= Death rate for Rural population of State 2
$\mu_{2\,u}$= Death rate for Urban population of State 2

Then $0.4 * \boldsymbol{\mu_{2\,r}} + 0.6 * \boldsymbol{\mu_{2\,u}} = 0.015085$

$\Rightarrow 0.4*1.4* \mu_{2\,u} + 0.6 * \boldsymbol{\mu_{2\,u}} = 0.015085$

$\Rightarrow \mu_{2\ u} = 0.013004$ & $\mu_{2\ r} = 0.018206$

Let $\mu_{2\ ro}$= Death rate for Rural Male population of State 2
$\mu_{2\ r1}$= Death rate for Rural Female population of State 2

Then $0.6^* \mu_{2\ ro} + 0.4^* \mu_{2\ r1} = 0.018206$

$\Rightarrow 0.6^*\ 1.2^* \mu_{2\ r1} + 0.4^* \mu_{2\ r1} = 0.018206$
$\Rightarrow \mu_{2\ r1} = 0.016255$ & $\mu_{2\ ro} = 0.019506$

Central death rate for Rural Male in State 2 is 0.019506

Let $\mu_{2\ uo}$= Death rate for Urban Male population of State 2
$\mu_{2\ u1}$= Death rate for Urban Female population of State 2

Then $0.4^* \mu_{1\ uo}+ 0.6^* \mu_{1\ u1}= 0.013004$

$\Rightarrow 0.4^*\ 1.2^* \mu_{1\ uo}+ 0.6^* \mu_{1\ u1}= 0.013004$
$\Rightarrow \mu_{1\ u1}= 0.012041$ & $\mu_{1\ ro} = 0.01449$

Central death rate for Urban Female in State 2 is 0.012041 [5]

iv) It is evident from the above that the death rate for State 2 is higher than State 1. The company would do better if it varies the premium according to the mix of Rural/Urban and within Rural/Urban further divided into Male/Female. This is because the level of heterogeneity within the data may affect the overall level of homogeneity assumed for pricing the product.
If the company does not differentiate the premium then it will lead to selection against the office resulting in :

- High Claim ratio – because of lower premium in one segment than the exact premium based on experience.
- Loss of business to competitors – because of higher premium in one segment than the exact premium based on experience.

[4]

[13 Marks]

Solution 9:

i) The sequence of events may by summarized as follows:

Duration	total Samosas left	Sold	Not sold
t (j)	n (j)	d (j)	c (j)
20	50	5	0
40	45	7	0
60	38	0	5
80	33	10	0
100	23	0	3
120	20	7	0
140	13	8	0
150	5	0	0

The hazard of samosas being sold is thus
5/50 at duration 20 minutes
7/45 at duration 40 minutes
10/33 at duration 80 minutes
7/20 at duration 120 minutes
8/13 at duration 140 minutes

The Nelson-Aalen estimate of the survival function, S(t) is then

Duration in minutes	Nelson-Aalen estimate S(t)
$0 \ll t < 20$	1
$20 \ll t < 40$	Exp(-5/50) = 0.904837
$40 \ll t < 80$	Exp(- (5/50 + 7/45)) = 0.774486
$80 \ll t < 120$	Exp(- (5/50 + 7/45 + 10/33)) = 0.572017
$120 \ll t < 140$	Exp(- (5/50 + 7/45 + 10/33 + 7/20)) = 0.403094
$140 \ll t < 150$	Exp(- (5/50 + 7/45 + 10/33 + 7/20 + 8/13)) = 0.217845

The Nelson-Aalen estimate is a step function. We need t for which S(t) = 0.4.
Therefore, 140 minutes is the estimated time required for him to sell 60% of samosas.

[6]

ii) Comments:

- The estimation is based on only one day's experience. Several days experience is required for better prediction.

- The student wasted eight samosas – Five thrown out due to inedible condition and three sold out for free. This may not occur in future as the student may be careful while buying or selling the samosas. The method does, however, take account of Censored data.
- Selling samosas in a railway station heavily depends upon the time. During peak hours he may be able to sell the required numbers. So the starting time of selling samosas will affect his estimated time.

[3]

[9 Marks]

Solution 10:

i) Consider the interval from age x to x+t+h. By the Markov property, we have:

$$_{t+h}p_x^{13} = {}_tp_x^{11} * {}_hp_{x+t}^{13} + {}_tp_x^{12} * {}_hp_{x+t}^{23} + {}_tp_x^{13} * {}_hp_{x+t}^{33} + {}_tp_x^{14} * {}_tp_{x+h}^{43}$$
$$_{t+h}p_x^{13} = {}_tp_x^{11} * {}_hp_{x+t}^{13} + {}_tp_x^{12} * {}_hp_{x+t}^{23} + {}_tp_x^{13} * (1 - {}_hp_{x+t}^{34}) + {}_tp_x^{14} * 0$$

For small h we can write:

$$_hp_{x+t}^{ij} = h.\mu_{x+t}^{ij} + o(h)$$

Then,

$$_{t+h}p_x^{13} = {}_tp_x^{11} * h.\mu_{x+t}^{13} + {}_tp_x^{12} * h.\mu_{x+t}^{23} + {}_tp_x^{13} * (1 - h.\mu_{x+t}^{34}) + o(h)$$

Rearranging,

$$\frac{_{t+h}p_x^{13} - {}_tp_x^{13}}{h} = {}_tp_x^{11} * \mu_{x+t}^{13} + {}_tp_x^{12} * \mu_{x+t}^{23} - {}_tp_x^{13} * \mu_{x+t}^{34} + \frac{o(h)}{h}$$

Letting h$\rightarrow$ 0 gives:

$$\frac{d}{dt}\,{}_tp_x^{13} = {}_tp_x^{11} * \mu_{x+t}^{13} + {}_tp_x^{12} * \mu_{x+t}^{23} - {}_tp_x^{13} * \mu_{x+t}^{34}$$

[5]

ii) We can write,

$$_{t+h}p_x^{11} = {}_tp_x^{11} * {}_hp_{x+t}^{11}$$

From the law of total probability

$$_hp_{x+t}^{11} + {}_hp_{x+t}^{12} + {}_hp_{x+t}^{13} + {}_hp_{x+t}^{14} + o(h) = 1$$

$$_hp_{x+t}^{11} = 1 - ({}_hp_{x+t}^{12} + {}_hp_{x+t}^{13} + {}_hp_{x+t}^{14}) + o(h)$$

So,

$$_{t+h}p_x^{11} = {}_tp_x^{11} * (1 - ({}_hp_{x+t}^{12} + {}_hp_{x+t}^{13} + {}_hp_{x+t}^{14})) + o(h)$$

$$_{t+h}p_x^{11} = {}_tp_x^{11} * (1 - (h * \mu_{x+t}^{12} + h * \mu_{x+t}^{13} + h * \mu_{x+t}^{14})) + o(h)$$

$$\frac{{}_{t+h}p_x^{11} - {}_tp_x^{11}}{h} = -{}_tp_x^{11}(\mu_{x+t}^{12} + \mu_{x+t}^{13} + \mu_{x+t}^{14}) + \frac{o(h)}{h}$$

Letting h→ 0 gives:

$$\frac{d}{dt}\,{}_tp_x^{11} = -{}_tp_x^{11}(\mu_{x+t}^{12} + \mu_{x+t}^{13} + \mu_{x+t}^{14})$$

$$\frac{1}{{}_tp_x^{11}}\frac{d}{dt}\,{}_tp_x^{11} = -(\mu_{x+t}^{12} + \mu_{x+t}^{13} + \mu_{x+t}^{14})$$

$$\frac{d}{dt} In\,{}_tp_x^{11} = -(\mu_{x+t}^{12} + \mu_{x+t}^{13} + \mu_{x+t}^{14})$$

$$In\,{}_tp_x^{11} = -\int_0^t \mu_{x+t}^{12} + \mu_{x+t}^{13} + \mu_{x+t}^{14}\,dt$$

$${}_tp_x^{11} = exp^{-\int_0^t \mu_{x+t}^{12}+\mu_{x+t}^{13}+\mu_{x+t}^{14}dt}$$

[4]

iii) The generator matrix is given as

$$\begin{matrix} -0.20 & 0.15 & 0.02 & 0.03 \\ 0.5 & -0.7 & 0.15 & 0.05 \\ 0 & 0 & -0.3 & 0.3 \\ 0 & 0 & 0 & 0 \end{matrix}$$

The force of leaving Level 1 is -0.2.

$${}_4p_x^{11} = exp^{-\int_0^4 0.2dt}$$

$${}_4p_x^{11} = exp^{-.2*4}$$

Hence $${}_4p_x^{11} = 44.93\%$$

[1]

iv) Considering the possible transfer out of state 2 to state 4 directly or via state 1:

$$P_{24}\,(s,t) = \int_0^{t-s} e^{-\int_s^{s+w} 0.7\,du} * 0.05 * P_{44}\,(s+w,t)\,dw$$
$$+ \int_0^{t-s} e^{-\int_s^{s+w} 0.7\,du} * 0.5 * P_{14}\,(s+w,t)\,dw$$
$$= \int_0^{t-s} e^{-\int_s^{s+w} 0.7\,du} * 0.05\,dw$$
$$+ \int_0^{t-s} e^{-\int_s^{s+w} 0.7\,du} * 0.5 * P_{14}\,(s+w,t)\,dw$$

[3]

[13 Marks]

****