Institute of Actuaries of India

Subject CT4 – Models

September 2016 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)

Review the regulatory guidance.

- Define the scope of the model, for example which factors need to be modeled stochastically.
- Plan the development of the model, including how the model will be tested and validated.
- Consider alternative forms of model, and decide and document the chosen approach. Where appropriate, this may involve discussion with experts on the underlying stochastic processes.
- Collect any data required, for example historic losses or policy data.
- Choose parameters. For economic factors should be able to calibrate to market data. For other factors e.g. expenses, claim distributions need to look at internal data adjusted for credibility
- Existing worst case scenarios. Understand the rationale, especially to gauge views on the probability of events occurring.
- Decide on the software to be used for the model.
- Write the computer programs.
- Debug the program, for example by checking the model behaves as expected for simple, defined scenarios.
- Review the reasonableness of the output. May include:
    - median outcomes (how do these compare with business plans)
- what probability is assigned to worst case scenarios
- Test the sensitivity of the model to small changes in parameters.
- Calculate the capital requirement.
- Communicate findings to management. Document.

Other suitable points were given credit, including:

- Validate data.
- Run model on historic data to compare models predictions with previous observations.
- Review parameters that have greatest effect on outputs.
- Present range of capital requirements for differing parameter inputs.

[4]

ii)

- A stochastic model is one that recognizes the random nature of the input components. A model that does not contain any random component is deterministic in nature.
- In a deterministic model, the output is determined once the set of fixed inputs and the relationships between them have been defined. By contrast, in a stochastic model the output is random in nature − like the inputs, which are random variables.
- A deterministic model is really just a special (simplified) case of a stochastic Model
- Whether one wishes to use a deterministic or a stochastic model depends on whether one is interested in the results of a single "scenario" or in the distribution of results of possible "scenarios".
- A deterministic model will give one the results of the relevant calculations for a single scenario; a stochastic model gives distributions of the relevant results for a distribution of scenarios.
- If the stochastic model is investigated by using "Monte Carlo" simulation, then this provides a collection of a suitably large number of different deterministic models, each of which is considered equally likely.
- The results for a deterministic model can often be obtained by direct calculation, integrate functions or to solve differential equations. However, If a stochastic model is sufficiently tractable, it may be possible to derive the results one wishes by analytical methods. If this can be done it is often preferable to, and also often much quicker than, Monte Carlo simulation. One gets precise results and can analyses the effect of changes in the assumptions more readily.

[4]

[8 Marks]

Solution 2:

i) b−aqx+a = P(Tx< b | Tx> a) [by definition]

= P(a <Tx< b) /1 − P(Tx< a) [conditional probability]

=(P(Tx< b) − P(Tx< a)) / P(Tx> a)

= bP(Tx< 1) − aP(Tx< 1)/ 1 − aP(Tx< 1) [from the uniform assumption]

= (b − a)qx/ 1 – aqx [as P(Tx< 1) = qx].

[2]

ii) ${}_{b-a}q_{x+a}$ = P(Tx< b | Tx> a) [by definition]

= P(a <Tx< b)/ 1 − P(Tx< a) [conditional probability]

= (P(Tx> a) − P(Tx> b))/ P(Tx> a)

$$=\frac{\frac{1-qx}{1-(1-a)qx}-\frac{1-qx}{1-(1-b)qx}}{\frac{1-qx}{1-(1-a)qx}}$$ (by definition of Balducci assumption and P(Tx< 1) = qx)

= (b − a)qx /(1 − (1 − b)qx)

[2]

iii) ${}_{b-a}q_{x+a}$ = P(Tx< b | Tx> a) [by definition]

= P(a <Tx< b) P(Tx> a) [conditional probability]

= (P(Tx> a) − P(Tx> b))/ P(Tx> a)

=((1 − qx)^a − (1 − qx)^b)/ (1 − qx)^a (by defn of constant force of mortality)

[2]

iv) The force of mortality under the uniform distribution of deaths increases with duration over age x to x+1, and that under Balducci assumption decreases with duration. It remains constant over age x to x+1, under the constant force of mortality.

The proof can be done in other methods as well. Part (i) could be as follows

Uniform distribution of deaths means that ${}_tq_x = tq_x$ (for 0 < t < 1). Hence:

$$ {}_{b-a}q_{x+a} = 1 - {}_{b-a}p_{x+a} = 1 - \frac{{}_bp_x}{{}_ap_x} = 1 - \frac{1 - {}_bq_x}{1 - {}_aq_x} = 1 - \frac{1 - bq_x}{1 - aq_x} = \frac{(b-a)q_x}{1 - aq_x}. $$

Also, part (iii) could be done as below.

Assuming a constant force of mortality, μ, we can write: ${}_tp_x = e^{-\mu t}$ (for 0 < t < 1)

We then have

$$ {}_{b-a}q_{x+a} = 1 - {}_{b-a}\, p_{x+a} = 1 - \frac{{}_b p_x}{{}_a p_x} = 1 - \frac{e^{-\mu b}}{e^{-\mu a}} = \frac{e^{-\mu a} - e^{-\mu b}}{e^{-\mu a}} = \frac{(p_x)^a - (p_x)^b}{(p_x)^a}\ , $$

when the result follows directly using$p_x = 1 - q_x$.

[2]

[8 Marks]

Solution 3:

i)

a) The state space is the set of values which it is possible for each random variable X_t to take.

b) The time set is the set J, the times at which the process contains a random variable X_t

c) A sample path is a joint realization of the variables X_t for all t in J that is a set of values for X_t (at each time in the time set) calculated using the previous values for Xt in the sample path.

d) The jump chain (or embedded chain) of a Markov jump process is the sequence of states that the process enters. The time spent in each state is ignored. The jump chain is a Markov chain in its own right.

[2]

ii) (a+b)

- **Discrete State Space, Discrete Time**

  (a) Simple random walk, Markov chain, or any other suitable example

  (b) Any reasonable example. For example: No Claims Discount systems, Credit Rating at end of each year

- **Discrete State Space, Continuous Time**

  (a) Poisson process, Markov jump process, for example

  (b) Any reasonable example. For example: Claims received by an insurer, Status of pension scheme member

- **Continuous State Space, Discrete Time**

  (a) General random walk, time series, for example

  (b) Any reasonable example. For example: Share prices at end of each trading day, Inflation index

- **Continuous State Space, Continuous Time**

  (a) Brownian motion, diffusion or Itô process, for example. Compound Poisson process if the defined state space is continuous.

  (b) Any reasonable example. For example: Share prices during trading period, Value of claims received by insurer

[2+4=6]
[8 Marks]

Solution 4:

i) Since the model assumes simultaneous influence of many parameters and the parametric function underlying the hazard need not necessarily be known or inferred, the Cox proportional hazards model is a good model for the purpose.

[1]

ii) $\lambda(t) = \lambda 0(t)\exp(\beta A * A + \beta G * G + \beta T_1 * T_1 + \beta T_2 * T_2 + \beta N * N)$,where

$\lambda(t)$ is the estimated hazard
$\lambda 0$ (t) is the baseline hazard
'β' are the regression parameters
'A' is the age-covariate
'G' is gender and is 1 if male and 0 if female
'T1 and T2' are the drug-covariates, T1 is 1 if heroin and 0 otherwise, T2 is 1 if cocaine and 0 if otherwise
'N' is number of previous treatments **[2]**

iii) The assumptions underlying the model are

- The general shape of the 'hazard' function is same for all individuals
- The ratio of 'hazard' for two individuals is proportional to the co-variates and is time-independent
- The factors influencing the hazard operate individually and no interaction exists

[1]

iv) The first decrement occurs at t=2. The partial likelihood contribution is given by

L(β) = $L(\beta) = \frac{e^{\beta}}{(5+5e^{\beta})}$

The likelihood at t=5 is given as

$$L(\beta) = \frac{e^{\beta}}{(4 + 4e^{\beta})}$$

There are 2 decrements at t=6.

Using Breslow's approximation, the partial likelihood from the 2 is given by

$L(\beta) = \frac{e^{\beta}}{(4+2e^{\beta})} * \frac{1}{(4+2e^{\beta})}$

The likelihood at t=9 is given by

$$L(\beta) = \frac{1}{(2 + e^{\beta})}$$

There are 2 possible ways of solving the above sub-question. The solution above assumes the indicator for 'male' is 1 and female is 0. It can also be solved assuming 'male' is 0 and female is 1 – In this case the solution will have the following likelihoods.

At t = 1. L(β) = $L(\beta) = \frac{1}{(5+5e^{\beta})}$

At t = 5, $L(\beta) = \frac{1}{(4+4e^{\beta})}$

At t = 6, $L(\beta) = \frac{1}{(2+4e^{\beta})} * \frac{e^{\beta}}{(2+4e^{\beta})}$

At t =9 $L(\beta) = \frac{e^{\beta}}{(1+2e^{\beta})}$

[4]
[8 Marks]

Solution 5:

i) 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 dx at age x.

[1]

ii) **a)** The corresponding E^c_x is the lives aged 'x' nearest birthday.

Let Px,t be the number of policies in force aged x nearest birthday at time t. Then, if t is measured in years since the 1stJanuary in the year of investigation, a consistent exposed-to risk would be

$$E_x^c = \int_0^1 P_{x,t}$$

which, assuming that birthdays are uniformly distributed across the calendar year, may be approximated as

$$E_x^c = \frac{1}{2}[P_{x,0} + P_{x,1}]$$

d_x/E^c_x would, assuming that deaths occur uniformly over the rate-interval, estimate the force of mortality for age at the midpoint of the rate interval (x-1/2, x+1/2) i.eμ_x .

[2]

b) The definition relates to age next birthday on the 1st of January. correspondingE^c_x is the lives aged 'x' next birthday on the 1st of January .

Let Px,t be the number of lives aged x next birthday on the 1st of January. As the age label does not change during the year of investigation a consistent exposed-to risk would be

$$E_c^x = \int_0^1 P_{x,t}$$

which, assuming that additions and exits are uniformly distributed across the calendar year, may be approximated as

$$E_c^x = \frac{1}{2}[P_{x,0} + P_{x,1}]$$

d_x/E^c_x would estimate the force of mortality for age at the midpoint of the rate interval (x-1, x) i.e$\mu_{x-1/2}$.

[2]

c) The definition relates to age last birthday on the previous policy anniversary. The corresponding E^c_x is the lives aged 'x' last birthday on the previous policy anniversary. Let LetPx,t be the number of lives aged 'x' last birthday on the previous policy anniversary at time t.

Then, if t is measured in years since the 1st January in the year of investigation, a consistent exposed-to risk would be

$$E_c^x = \int_0^1 P_{x,t}$$

which, assuming that birthdays are uniformly distributed across the calendar year and policy anniversaries are uniformly distributed across the calendar year, may be approximated as

$$E_c^x = \frac{1}{2}[P_{x,0} + P_{x,1}]$$

d_x/E^c_x would estimate the force of mortality for age at the midpoint of the rate interval (x, x+1) i.e$\mu_{x+1/2}$.

[2]

iii) Initial exposed to risk,(E_x) is where the lives resulting in death are exposed for the full year in the year of death. Unlike in central exposed to risk, where lives resulting in death are only exposed till the time of death. **[1]**

iv) If the exact age at death is known and is 'x+t_i' for life 'i' then $E_x = E^c_x + \sum(1 - ti)$. If the exact age at death is not known then, assuming deaths occur on an average mid-way during the year $E_x = E^c_x + \frac{d}{2}$ where'd' is the number of deaths.

[2]
[10 Marks]

Solution 6:

i) The duration exposed for each of the lives is tabulated below.

Life	Date of birth	Status on 31st December 2015	Date of death/lapse/surrender	Date of entry (Later of age 45 or start date of investigation)	Date of exit (earlier lapse/surrender/death or end date of investigation)	Duration exposed
A	1st January 1970	In-force		1st January 2015	31st Dec 2015	12 months
B	1st November 1969	Lapsed	1st May 2015	1st January 2015	1st May 2015	4 months
C	1st May 1970	In-force		1st May 2015	31st Dec 2015	8 months
D	1st October 1969	Died	1st May 2015	1st January 2015	1st May 2015	4 months
E	1st February 1969	Lapsed	1st November 2014	-	-	-
F	1st April 1970	Died	1st October 2015	1st April 2015	1st October 2015	6 months

The likelihood function for each of the lives is derived using the formula

For lives that are censored = $e^{-\alpha*)t)^{\beta}}$ where'tʼ is the duration exposed

For deaths = $\alpha\beta(t)^{\beta-1}e^{-\alpha*t^{\beta}}$ where'tʼ is the duration exposed

The likelihood function for each of the lives is as given below

$$A = e^{-\alpha*1^{\beta}}$$
$$B = e^{-\alpha*1/3^{\beta}}$$
$$C = e^{-\alpha*2/3^{\beta}}$$

$$D = \alpha\beta(\frac{1}{3})^{\beta-1}e^{-\alpha*1/3^{\beta}}$$

$$E = nil$$

$$F = \alpha\beta(\frac{1}{2})^{\beta-1}e^{-\alpha*1/2^{\beta}}$$

[6]

ii) The parameters would be estimated by using maximium likelihood estimation method. Since there are two parameters, the estimation involves solving two simultaneous equations, which would involve iterative methods.

[1]

iii) If exponential function represents the survival function, then the likelihood is given by

$$A = e^{-\mu*1}$$
$$B = e^{-\mu*1/3}$$
$$C = e^{-\mu*2/3}$$
$$D = \mu e^{-\mu*1/3}$$
$$E = nil$$
$$F = \mu e^{-\mu*1/2}$$

[1.5]

iv) The likelihood function is given as the product of all partial likelihood and is

$\mathrm{L} = \mu^2 * \mathrm{e}^{-\mu*17/6}$.

The value of 'µ' which maximizes the likelihood function will also maximize the log of the function.

Hence taking logs

$$\log \mathrm{L} = 2 \log \mu - \mu * \frac{17}{6}$$

Differentiating this with respect to µ and setting it equal to zero, gives the MLE of µ.

$\mu = 2/(1+1/3+2/3+1/3+1/2) = 0.705882$

[1.5]

v) Censoring present in the investigation.

Right censoring of B, C and E – C on reaching the end of investigation period, B and E lapsing before reaching age 46.

Type I censoring as lives are censored at end of investigation period.

Censoring could be informative, if lives in good health lapse their policies.

[2]

[12 Marks]

Solution 7:

i) State space: { Identical , Edge **H**anu , Edge **B**heem , Game Hanu, Game Bheem }

Transition matrix:

$$\begin{array}{cccccc} & Identical & \text{Edge Hanu} & Edge\ Bheem & Game\ Hanu & Game\ Bheem \\ Identical & 0 & 0.6 & 0.4 & 0 & 0 \\ Edge\ Hanu & 0.4 & 0 & 0 & 0.6 & 0 \\ Edge\ Bheem & 0.6 & 0 & 0 & 0 & 0.4 \\ Game\ Hanu & 0 & 0 & 0 & 1 & 0 \\ Game\ Bheem & 0 & 0 & 0 & 0 & 1 \end{array}$$

The chain is Markov because the probability of moving to the next state does not depend on history prior to entering that state (because the probability of each player winning a point is constant)

[3]

ii) The chain is reducible because it has two absorbing states Game H and Game B. States Game H and Game B are absorbing so have no period. The other three states each have a period of 2 so the chain is not aperiodic.

[2]

iii) The game either ends after 2 points or it returns to Identical. The probability of it returning to Identical after two points is:

Prob H wins 1 st point × Prob B wins 2 nd point

+ Prob B wins 1 st point × Prob H wins 2 nd point

= 0.6×0.4 + 0.4×0.6 = 0.48. ([This can also be obtained by calculating the square of the transition matrix.]

Need to find number of such cycles N such that:

$0.48^N < 1- 0.9$

so that

N>ln 0.1 / (ln (0.48) > 3.14

But the game can only finish every two points so we require 4 cycles, that is 8 points.

[3]

iv) a) Define H_X to be the probability that H ultimately wins the game when the current state is X.

We require $H_{dentiacal}$.

By definition $H_{Game\ H} = 1$ and $H_{Game\ B} = 0$.

Conditioning on the first move out of state Edge H:

$H_{edgeH} = 0.6* H_{Game\ H} + 0.4 * H_{identical} = 0.6 + 0.4 *H_{identical}$

Similarly:

$H_{edge\ B} = 0.6* H_{Identical}$and

$H_{identical}= 0.6* H_{edgeH} + 0.4* H_{edgeB} = 0.6* H_{edge\ H}+ 0.24*H_{identical}$

So

$H_{\ Identical} = o.6/\ 0.4\ H_{edge\ H}$

$H_{\ Edge\ H} = 0.6+ 0.4 * 0.6/0.76\ H_{\ edgeH}$and

$H_{edgeH} = 0.8769$

$H_{identical}= 0.6923$

<u>Alternate method</u>

Probability H wins after 2 points = 0.6*0.6 =0.36

Probability that H wins from identical

= ∑Probability A wins after i points have been played

= Probability A wins after 2 points + Probability A wins after 4 points +…..

(as period 2)

$= 0.36 + 0.48 * 0.36 + 0.48^{2} * 0.36 +…….$

= 0.36/(1-0.48) as a geometric progression

= 0.6923

= 0.6923 **[4]**

b) This is higher than 0.6 because Hanu has to win at least two points in a row to win the game.

[1]
[13 Marks]

Solution 8:

i) Graduation refers to the process of using statistical techniques to improve the estimates provided by the crude rates.

The aims of graduation are

- To produce a smooth set of rates that are suitable for a particular purpose
- To remove random sampling errors
- To use the information available from adjacent ages

[2]

ii) The three desirable features of graduation are

(a) smoothness
(b) adherence to data and
(c) suitability for the purpose in hand.

[1]

iii) If the graduation process results in rates that are smooth but show little adherence to the data, then we say that the data may be overgraduated. Here the graduated rates may not be representative of the underlying experience.

Undergraduation refers to the case where insufficient smoothing has been carried out. This will tend to produce a curve of inadequate smoothness, but better adherence to data. In this case, the graduated rates will follow the crude rates very closely, but will show an irregular progression over the range of ages.

[2]

iv) The three methods of graduation are

- graduation by parametric formula
- graduation by reference to a standard table
- graphical graduation

Graduation by parametric formula
Advt

- if a small number of parameters is used, the resulting graduation will be acceptably smooth

- when comparing several experiences, the same parametric formula can be fitted to all of them. Differences between the parameters, given their standard errors, give insight into the difference between the experiences
- the approach is well suited to the production of standard tables from large amounts of data

Disadvt

- it can be difficult to find a suitable curve which fits the experience at all ages. Partly this is because of the different features that predominate at different ages (infant mortality, the accident hump and exponential mortality after middle ages).
- Care is needed when extrapolating. The fit of the curve will probably be best where there is most data, but results where data are scanty may be poor and require adjustment.

Appropriate

- It is used when standard table is being produced
- When using a large experience, where a suitable formula can be found to fit all ages is available

Graduation by reference to a standard table

Advt

- Provided a simple function is chosen and the standard table is smooth to begin with, the resultant rates are automatically smooth.
- it can be useful to fit relatively small data sets when a suitable standard table exists
- the standard table can be very good at deciding the shape of the graduation at extreme ages where data are sparse

DisAdvt

- it can be difficult to find a suitable standard table for the data. (choosing an inappropriate table could impart the wrong shape to the entire graduation)
- it is not suitable for the preparation of standard tables

Appropriate

- if data are scanty AND a suitable table exists

Graphical Graduation

Advt

- It can be used for small data where no suitable standard tables exists
- It can allow for known features of the experience for example the accident hump

DisAdvt

- it is hard to achieve accuracy
- It takes a skilled practitioner
- It is very difficult to achieve adequate smoothness
- not a suitable method for the production of standard tables based on large amounts of data
- the method is subjective.

Appropriate

- No suitable standard table is likely to exist AND the experience is small.

[5]

v) The graduated rates are derived using the formulae given and the estimated parameters, as given below:

Age	Male	Female
20	0.000574	0.000337
25	0.000849	0.000499
28	0.001074	0.000632
33	0.001587	0.000935
66	0.020633	0.012362
71	0.030221	0.018202

[4]

vi) The appropriateness of graduation can be tested using chi-square test. The exposed to risk at each age can be taken as 1,000 lives. Based on this exposed-to-risk, the actual and expected deaths can be estimated.

The difference between the observed deaths and expected deaths as per the graduated rates are summarised below.

As the actual deaths for females at age 20 and 25 are below 5, these 2 cells are combined.

	Age	Actual	Expected	(A-E)	(A-E)^2	(A-E)^2/E

Males	20	5.69	0.5742	5.1158	26.1714	45.5792
	25	7.94	0.8491	7.0909	50.2805	59.2142
	28	11.06	1.0737	9.9863	99.7256	92.8777
	33	15.41	1.5875	13.8225	191.0625	120.3570
	66	21.51	20.6325	0.8775	0.7700	0.0373
	71	29.98	30.2212	-0.2412	0.0582	0.0019
Females	20-25	8.05	0.8358	7.2142	52.0441	62.2658
	28	6.51	0.6315	5.8785	34.5567	54.7212
	33	9.04	0.9351	8.1049	65.6887	70.2443
	66	12.55	12.3615	0.1885	0.0355	0.0029
	71	17.44	18.2025	-0.7625	0.5813	0.0319

The observed χ2 statistic is higher than χ2 with 11 degrees of freedom. Hence we can conclude that the graduation process is not appropriate.

[3]
[17 Marks]

Solution 9:

i) $P_{10}(t+h) = (1- \rho h).P_{10}(t) + \sigma h P_{11}(t)$

$= d/dtP10\ (t) = - \rho\ P_{10}(t) + \sigma h P_{11}(t)$ ..... eq (1)

And

$P_{11}\ (t+h) = \rho h\ P_{10}(t) + (\ 1- \sigma h\ )P_{11}(t)$

$= d/dtP_{11}(t) = \rho P_{10}(t) - \sigma P_{11}(t)$ ..... eq (2)

[(2) also follows from the fact that $P_{10}\ (t)+ P_{11}(t) =1$]

[3]

ii) $P_{10}(t)+ P_{11}(t) =1$

so from (1) $d/dt\, P_{10}(t) + (\sigma + \rho) \,.\, P_{10}(t) = \sigma$

$d/dt Exp((\sigma + \rho)t)\, P_{10}(t) = \sigma * Exp((\sigma + \rho)t) + C$

$P_{10}(t) =)= \sigma / (\sigma + \rho) * \exp(\sigma + \rho) * t) + C * \exp - (\sigma + \rho) * t$

$P_{10}(t) =)= \sigma / (\sigma + \rho) * (1 - \exp - (\sigma + \rho) * t)$ since P10(0) = 0.

Therefore P11(t) = $1 - \sigma / (\sigma + \rho)) * (1 - \exp - (\sigma + \rho) * t) = \frac{(\rho + \sigma \exp(-(\sigma + \rho) * t)}{\sigma + \rho}$

[5]

iii)

a) The generator matrix is now

$$\begin{matrix} -\rho & \rho & 0 \\ \sigma & -(\sigma + \rho) & \rho \\ 0 & 2\sigma & -2\sigma \end{matrix}$$

[2]

b) Hence the Forward Equations are

$d/dt(P_0(t) = -\rho * p_0(t) + \sigma\, p_1(t)$

$d/dt(P_1(t)) = \rho\, p_0 - (\sigma + \rho)\, p_{1(t)+} 2\, \sigma p_2(t)$

$d/dt(P_2(t)) = \rho\, p_1(t) - 2\, \sigma p_2(t)$

[3]

c) Simply substituting in the suggested values gives the required result. **[1]**

d) The implication is that the given distribution is stationary. By the standard properties of Markov processes, it follows that it is the equilibrium distribution, so that the long-term probabilities of being in each of the three states are known.

[2]

[16 Marks]

****