Institute of Actuaries of India

Subject CT4 – Models

Nov 2012 Examinations

INDICATIVE SOLUTIONS

Question 1:

i. The Cox model proposes the following form of hazard function for the **i** th life (where, in keeping with statistical habit, we denote hazards by λ rather than μ):

$$\lambda(t, z_i) = \lambda_0(t)\exp(\beta z^T)$$

(Note: here the "**T**" denotes the transpose of the vector z_i , not a lifetime.)

β is a $1 \times p$ vector of regression parameters

Through the scalar product βz^T the influence of each factor in z_i enters the hazard multiplicatively.

$\lambda_0(t)$ is the baseline hazard

In this simple formulation only $\lambda_0(t)$ depends on time, but the model can also be formulated with time-dependent covariates.

ii. The force of mortality at time t a smoker who is having cancer type B is:

$$\lambda_{Smoker,Type\,B}(t) = \lambda_0(t)e^{(.017+.026)}$$

$$= \lambda_0(t)e^{.043}$$

The force of mortality at time t a smoker who is having cancer type A is:

$$\lambda_{Smoker,Type\,A}(t) = \lambda_0(t)e^{.017}$$

$$= \lambda_0(t)e^{.017}$$

The ratio of these two quantities is:

$$\frac{\lambda_{Smoker,Type\,B}(t)}{\lambda_{Smoker,Type\,A}(t)} = e^{.026}$$

$$= e^{.026}$$
$$= 1.02634$$

So, according to the model, we estimate that force of mortality for a smoker who is having cancer type B is 2.63% higher than that of a smoker who is having cancer type A.

[5]

Question 2:

Outline the key differences between deterministic models and stochastic models. Give an example each of deterministic and stochastic models.

- Stochastic model recognizes inherent randomness of input parameters
- Output of a deterministic model is a single set of results whereas a stochastic model would produce a distribution of results
- A deterministic model could be thought of as a special (simplified) case of a stochastic model
- Greater complexity in a stochastic model when compared to a deterministic model
- Deterministic models can be constructed quickly whereas construction and run-times of stochastic models are likely to be greater.
- Deterministic model with well thought out scenario tests is often used as a proxy where constructing stochastic models is envisaged to be time-consuming.
- 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".
- The results for a deterministic model can often be obtained by direct calculation
- If a stochastic model is sufficiently tractable, it may be possible to derive the results one wishes by analytical methods.
- Many practical problems, however, are too complicated for analytical methods to be used easily, and Monte Carlo simulation is an extremely powerful method for solving complicated problems.

Examples of deterministic models:
- pricing an insurance product based on point estimates of input parameters
- benefit illustration based on a fixed (e.g. 10%) investment return

Examples of stochastic models:
- assessing the cost of guarantee in an insurance product by running simulated asset return paths
- assessing longevity risk in an annuity portfolio by simulating future mortality improvement scales

[5]

Question 3:

i.   Left censoring

Data are left censored if the censoring mechanism prevents us from knowing when entry into the state that we wish to observe took place. An example arises in medical studies in which patients are subject to regular examinations. Discovery of a condition tells us only that the onset fell in the period since the previous examination; the time elapsed since onset has been left censored.

In the investigation above, the fund managers wish to study the exact duration of customers in-force. However, for the policies sold prior to 2002, only the year of entry may be imputed and the exact date of policy entry is no known. Therefore, the data prior to 2002 may be considered left censored.

ii. Random censoring

If censoring is random, then the time C_i (say) at which observation of the ith lifetime

is censored is a random variable. The observation will be censored if $C_i < T_i$ where T_i

is the (random) lifetime of the ith life. In such a situation, censoring is said to be random.

Random censoring arises when individuals may leave the observation by a means other than the decrement under study, and where the time of leaving is not known in advance.

As such, in the given investigation there are only two decrements: withdrawal which is the decrement being studied and maturities which are known in advance. Therefore, one should expect that random censoring should not be present. However, random censoring may exist if there are any data issues. For example, if there are any customers whose files are 'lost' such that it is not known whether they are still in-force or whether they have withdrawn prior to maturity, then such records may need to be excluded from the investigation and this may lead to come random censoring.

iii. Type II censoring

If observation is continued until a predetermined number of deaths has occurred, then "Type II censoring" is said to be present. This can simplify the analysis, because then the number of events of interest is non-random.

The study design involving Type II censoring is often used to speed up the production of results in situations where an minority of subjects may have very long lifetimes, but interest really focuses on the majority with lifetimes within the normal range.

In the current investigation of withdrawals from the investment fund, there is no pre-determined limit beyond which the investigation must cease. Instead, the investigation is carried out for all withdrawals and maturities over the period 2008-2012. Therefore, there is no Type-II censoring present although there is Type I censoring at the point where the investigation ceases.

iv. Informative censoring

Informative censoring occurs when the data being censored provides some information in respect of the population that remains at risk. For example, Ill-health retirements from pension schemes, because these are likely to be in worse than average health than the continuing members. So the mortality rates of those who

remain in the pension scheme are likely to be lower than the mortality rates of the lives that left through ill-health retirement.

There is no specific information in the observational plan for the investigation of investment fund withdrawals that would suggest any informative censoring present. Therefore, it may be assumed that all censoring is non-informative such as the end of the investigation period in 2012, which would affect all customers equally, regardless of their propensity to withdraw at a future point in time.

v.   Interval censoring

Data are interval censored if the observational plan only allows us to say that an event of interest fell within some interval of time. An example arises in actuarial investigations, where we might know only the calendar year of death. Both right and left censoring can be seen as special cases of interval censoring.

For policies sold prior to 2002, only the calendar year of inception and calendar date of exit is known. Therefore this would give rise to interval censoring. Equally, both right and left censoring are present in the data and as mentioned above, these can be seen as special cases of interval censoring. Therefore, it can be said that interval censoring is present in the data.

[10]

Question 4:

i. Using Markov Property,

$$_{t+h}p_x^{24} = \sum_{k=2}^{5} {}_tp_x^{2k}\, {}_hp_{x+t}^{k4}$$

We are assuming here that the probability of transition into any state at any given point in time depends on the current state occupied and current age only.

Hence,

$$_{t+h}p_x^{24} = {}_tp_x^{22}\left(\eta h + o(h)\right) + {}_tp_x^{23}\left(\delta h + o(h)\right) + {}_tp_x^{24}$$

(Assuming $_hp_{x+t}^{24} = \eta h + o(h),\ _hp_{x+t}^{34} = \delta h + o(h)$)

$o(h)$ is a quantity such that $\lim_{h \to 0+} \dfrac{o(h)}{h} = 0$

We can rearrange this to give:

$$\frac{_{t+h}p_x^{24} - {}_tp_x^{24}}{h} = {}_tp_x^{22}\eta + {}_tp_x^{23}\delta + \frac{o(h)}{h}$$

Allowing h tends to zero gives:

$$\frac{\delta}{\delta t}\, {}_tp_x^{24} = {}_tp_x^{22}\eta + {}_tp_x^{23}\delta$$

ii. The expressions are:

$$\frac{\delta}{\delta t}\, {}_tp_x^{22} = -{}_tp_x^{22}(\eta + \beta)$$

and

$$\frac{\delta}{\delta t} {}_t p_x^{23} = -{}_t p_x^{23} (\delta + \gamma) + {}_t p_x^{22} \beta$$

iii.   Since,

$$\frac{\delta}{\delta t} {}_t p_x^{22} = -{}_t p_x^{22} (\eta + \beta)$$

$$\frac{\delta}{\delta t} \ln({}_t p_x^{22}) = -(\eta + \beta)$$

Integrating gives

$$\ln({}_t p_x^{22}) = -t(\eta + \beta) + c$$

where c is the constant of integration. Using the initial condition ${}_0 p_x^{22} = 1$ we find that $c = 0$

Hence,

$${}_t p_x^{22} = e^{-t(\eta + \beta)}$$

[11]

Question 5:

i.   The possible scenarios starting from 30-30 are as follows:

WW

WLWW, LWWW

WLLWWW, LWLWWW, LWWLWW, WLWLWW

So on……

Hence, the probability $P_D(30,30)$ is

$$p^2 + 2p^3 q + 4p^4 q^2 + \ldots\ldots$$

$$= p^2 (1 + 2p\,q + 4p^2 q^2 + \ldots\ldots)$$

$$= \frac{p^2}{1 - 2pq}$$

Alternatively:

Probability of Djokovic winning from 30-30 would involve:

-   Djokovic winning the next point and then probability of him winning from 40-30
-   Djokovic losing the next point and then probability of him winning from 30-40

$$P_D(30,30) = pP_D(40,30) + qP_D(30,40)$$

Likewise:

$$P_D(40,30) = p + qP_D(30,30)$$

$$P_D(30,40) = pP_D(30,30)$$

$$\therefore P_D(30,30) = p\big(p + qP_D(30,30)\big) + qpP_D(30,30)$$

$$\therefore P_D(30,30) = p^2 + pqP_D(30,30) + qpP_D(30,30)$$

$$\therefore P_D(30,30) = \frac{p^2}{1 - 2pq}$$

ii.   The possible scenarios starting from 40-30 can lead to a direct win or win via deuce

The probability $P_D(40,30)$ is

$$p + q(p^2 + 2p^2q^2 + 4p^3q^3 + \ldots\ldots)$$

The possible scenarios starting from 30-40 can lead to a win only through deuce

The probability $P_D(30,40)$ is

$$p(p^2 + 2p^3q + 4p^4q^2 + \ldots\ldots)$$

$$= p^3 + 2p^4q + 4p^5q^2 + \ldots\ldots$$

$$= p^3(1 + 2pq + 4p^2q^2 + \ldots\ldots)$$

Alternatively,:

$$P_D(40,30) = p + qP_D(30,30) = p + \frac{qp^2}{1 - 2pq}$$

$$P_D(30,40) = pP_D(30,30) = \frac{p^3}{1 - 2pq}$$

iii.   Djokovic can also win the game without losing a single point Murray taking only one point

The sequence that could be followed is 15-0, 30-0, 40-0 and game.

Hence the probability of winning the game is

p^4

Djokovic can also win the game losing only one point. The possible scenarios are:

15-0, 30-0, 40-0, 40-15 and game
15-0, 30-0, 30-15, 40-15 and game
15-0, 15-15, 30-15, 40-15 and game
0-15, 15-15, 30-15, 40-15 and game
The probability of each of the scenario is p^4q

Hence, probability of all the four scenarios $4p^4q$

Djokovic can win the game after losing at least two points.

Djokovic can reach to game by following any of the three possible routes

30-30(deuce)
40-30 without routing via 30-30 (Djokovic being 40)

30-40 without routing via 30-30 (Djokovic being 30)

The probability of reaching 30-30 is $p^2 q^2$ and the same can be reached by six possible ways

15-0, 30-0, 30-15, 30-30
15-0, 15-15, 30-15, 30-30
15-0, 15-15, 15-30, 30-30
0-15, 0-30, 15-30, 30-30
0-15, 15-15, 15-30, 30-30
0-15, 15-15, 30-15, 30-30

Since, the probability $P_D(30,30)$ is $\dfrac{p^2}{1-2pq}$

Hence, Probability of Djokovic winning a game through 30-30 (or deuce) is

$\dfrac{6p^4 q^2}{1-2pq}$ .... (1)

The probability of reaching 40-30 without via 30-30 is $p^3 q^2$ and the same can be reached by four possible ways

15-0, 30-0, 40-0, 40-15, 40-30
15-0, 30-0, 30-15, 40-15, 40-30
15-0, 15-15, 30-15, 40-15, 40-30
0-15, 15-15, 30-15, 40-15, 40-30

Since, the probability $P_D(40,30)$ is

$$p + q(p^2 + 2p^2 q^2 + 4p^3 q^3 + \ldots \ldots)$$

Hence, Probability of Djokovic winning a game through 40-30 (without going via 30-30) is

$$4p^3 q^2(p + qp^2 + 2p^3 q^2 + 4p^4 q^3 + \ldots \ldots) \ldots \ldots \ldots (2)$$

The probability of reaching 30-40 without via 30-30 is $p^2 q^3$ and the same can be reached by four possible ways

15-0, 15-15, 15-30, 15-40, 30-40
0-15, 15-15, 15-30, 15-40, 30-40
0-15, 0-30, 15-30, 15-40, 30-40
0-15, 0-30, 0-40, 15-40, 30-40

Since, the probability $P_D(30,40)$ is

$$p^3 \left(1 + + 2p\,q + 4p^2\,q^2 + \ldots \ldots \right)$$

Hence, Probability of Djokovic winning a game through 30-40 (without going via 30-30) is

$$4p^2\,q^3 \left(p^3 + 2p^4\,q + 4p^5\,q^2 + \ldots \ldots \right) \ldots \ldots \ldots \ldots \ldots (3)$$

Adding (2) and (3) we get the probability of winning the game either through 40-30or 30-40(without going via 30-30) as given below

$$4p^3\,q^2 \left(p + qp^2 + 2p^3\,q^2 + 4p^4\,q^3 + \ldots \ldots \right) + 4p^2\,q^3 \left(p^3 + 2p^4\,q + 4p^5\,q^2 + \ldots \ldots \right)$$

$$= 4p^4\,q^2 \left(1 + 2pq + 4p^2\,q^2 + 8p^3\,q^3 + \cdots \right)$$

$$= 4p^4\,q^2 \frac{1}{1 - 2pq}$$

$$= \frac{4p^4\,q^2}{1 - 2pq} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (4)$$

Adding (1) and (4) we get the required probability of Djokovic winning the game after losing at least two points of $\dfrac{10p^4\,q^2}{1 - 2pq}$

Hence, the total probability of Djokovic wining the game, $P_D(0,0)$, is

$$p^4 + 4p^4 q + \frac{10p^4 q^2}{1 - 2pq}$$

[15]

Question 6:

i. Let state 0 indicate the "Low" level of sales and state 1 indicate the "High" level of sales, where each transition of the process goes from one year to the next.

a) The one-step transition matrix for the "never introduce new product" strategy is

$$\text{State} \quad 0 \quad 1$$

$$\mathbf{P} = \begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{bmatrix}.$$

b) The one-step transition matrix for the "introduce new product every year" strategy is

$$\text{State} \quad 0 \quad 1$$

$$\mathbf{P} = \begin{matrix} 0 \\ 1 \end{matrix} \begin{vmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{vmatrix}.$$

c) The one-step transition matrix for the CEO's proposal is

$$\text{State} \quad 0 \quad 1$$

$$\mathbf{P} = \begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

ii.   The steady-state probabilities (π_0, π_1) can be found by solving the following equations:

$$\pi_0 = \pi_0 \, p_{00} + \pi_1 \, p_{10}$$

$$\pi_1 = \pi_0 \, p_{01} + \pi_1 \, p_{11}$$

$$\pi_0 + \pi_1 = 1$$

Where one of the first two equations (say, the second) is redundant and so can be ignored.

For $a)$ the "never introduce new product" strategy, this system of equations becomes

$$\pi_0 = \frac{3}{4} \, \pi_0 + \frac{1}{2} \pi_1$$

So

$$\pi_0 = 2\pi_1$$

Since

$$\pi_0 + \pi_1 = 1$$

This   implies   that
$$\pi_0 = \frac{2}{3} \quad and \quad \pi_1 = \frac{1}{3}$$

Similarly, for the other two options

For **b)** the "introduce new product every year" strategy,

$$\Pi_0 = \frac{1}{3} \text{ and } \Pi_1 = \frac{2}{3}$$

For **c)** the CEO's proposal,

$$\Pi_0 = \frac{1}{2} \text{ and } \Pi_1 = \frac{1}{2}$$

iii. Given that the company's yearly profits are Rs. 10 Cr. when sales are "High" (State = 1) and Rs. 5 Cr. when sales are "Low" (State = 0). Therefore, the long run expected average profit is $5\pi_0 + 10\pi_1$. Taking the cost of introducing new products (Rs. 1 Cr.) in to account, the following calculations are performed to calculate average profits under all three options.

For **a)** the "never introduce new product" strategy, the long-run expected average profit is

$$\text{Profit} = 5(2/3) + 10(1/3) = \text{Rs. } 20/3 \text{ Cr.}$$

For **b)** the "introduce new product every year" strategy, the long-run expected average profit is

$$\text{Profit} = 5(1/3) + 10(2/3) - 1 = \text{Rs. } 22/3 \text{ Cr.}$$

For **c)** the CEO's proposal, the long-run expected average profit is

$$\text{Profit} = (5-1)(1/2) + 10(1/2) = \text{Rs. } 21/3 \text{ Cr.}$$

Therefore, in order to maximize the long run expected average profit, the optimal strategy is "introduce new product every year

[13]

Question 7:

i. To test for smoothness, we can calculate the third differences of the graduated rates. These second differences should be smooth in progression and the third differences should be small in magnitude.

The third differences are calculated below:

Age	Mortality rate	First difference	Second difference	Third difference
41	0.002495	0.000399	0.000067	-0.000001

42	0.002894	0.000466	0.000066	-0.000002
43	0.003360	0.000532	0.000064	-0.000002
44	0.003892	0.000596	0.000062	-0.000002
45	0.004488	0.000658	0.000060	-0.000001
46	0.005146	0.000718	0.000059	-0.000004
47	0.005864	0.000777	0.000055	-0.000004
48	0.006641	0.000832	0.000051	-0.001136
49	0.007473	0.000883	-0.001085	0.001127
50	0.008356	-0.000202	0.000042	0.000005
51	0.008154	-0.000160	0.000047	0.000004
52	0.007994	-0.000113	0.000051	0.000007
53	0.007881	-0.000062	0.000058	0.000005
54	0.007819	-0.000004	0.000063	0.000008
55	0.007815	0.000059	0.000071	0.000006
56	0.007874	0.000130	0.000077	0.000011
57	0.008004	0.000207	0.000088	0.000008
58	0.008211	0.000295	0.000096	
59	0.008506	0.000391		
60	0.008897			

From the results above, we can make the following conclusions:

- The first differences are positive for ages 41-49, negative from 51-54 and positive again from age 55 onwards. This indicates mortality rates increase until age 50, then reduce over early 50s age group before increasing again from mid-50 onwards

- The second differences appear to progress in an orderly fashion until there is a sharp kink at age 49.

- The third differences are small except for ages 48-49.

The above suggests that the mortality rates are smooth over the ages 41-50 and 51-60. However, at age 50, there is a discontinuity in the mortality rates.

This might be a concern for a pricing actuary as the premium rates may result in a discontinuity at age 50 as well, if the above mortality rates are used. This could potentially result in premiums for life insurance for older ages in the age band 50-55 being lower than younger policyholders in the similar age band. Since the convention is to charge higher premiums for older ages in life insurance, this might be difficult to justify to potential customers.

ii. Chi-square test

Purpose: A chi square test can be used to decide whether the observed numbers of individualswho fall into specified categories are consistent with a model that predicts the expectednumbers in each category. It is a test for overall *goodness of fit.* In the case of the insurance company, it can be used to determine whether the graduated mortality rates provide a good fit to the underlying experience (or crude) mortality rates.

Conclusion: From the Tables, the upper 95% point for the χ^2_{17} distribution is 27.59.

The observed value of the test statistic is only 14.87 which is much below this, so we can conclude that the graduated rates represent the underlying mortality well.

In fact, the test statistic is significantly lower than the critical Chi-square value - this may indicate under-graduation, which can be tested by considering the standardised deviations or by applying the smoothness test.

Standardised deviations test

Purpose: We can use the standardised deviations test to detect overall goodness of fit. Where the test reveals a problem this may be due to under/overgraduation, heterogeneity or duplicates. Under the standardised deviations test, if the graduated rates are not a good fit, the distribution will not be "centred correctly". If there is heterogeneity within the age groups or deaths are not independent, the variance will be (respectively) smaller or greater than we would expect if our underlying model were correct. If we have undergraduation, then we would expect the standardised deviations to be tightly bunched. Conversely, if we have overgraduation, we would expect the standardised deviations to be too spread out.

Conclusion: We can draw the following conclusions from standardised deviation test results:

- Overall shape: The number of values in each of the ranges conforms broadly with the expected values for the normal distribution.

- Outliers: there are no standardised deviations lying outside the range (-3,3). Therefore, it is safe to conclude that there are no obvious outliers in the graduated rates.

- Absolute deviations: Roughly half the standardised deviations should fall in the range (-2/3, 2/3). This indicates that the deviations are not overly concentrated towards the tale.

- Symmetry: There are exactly 10 positive and ten negative deviations, suggesting no overall bias in the graduated rates.

The above conclusions suggest that the graduated rates are neither over- nor under-graduated and also represent the data appropriately. Together with the result of the chi-square test, it appears that the graduated rates are indeed a good fit to the data. Therefore, the kink observed in the smoothness test may well be a genuine feature of the underlying mortality as opposed to being a result of under-graduation.

[14]

Question 8:

i. Likelihood expression

The likelihood of observing d deaths if the true value of the hazard rate is μ is:

$$L(\mu) = \frac{(\mu E_x^c)^d e^{-\mu E_x^c}}{d!}; \text{ where:}$$

μ is the rate parameter representing the force of mortality; and

E_x^c denotes the total observed waiting time (or central exposed to risk) at age x

ii. Distribution parameters for MLE

We can maximise the log-likelihood to find the maximum likelihood estimator of μ:

$$\log L(\mu) = d(\log \mu + \log E_x^c) - \mu E_x^c - \log d!$$

Differentiating with respect to μ:

$$\frac{\partial}{\partial \mu} \log L(\mu) = \frac{d}{\mu} - E_x^c$$

which is zero when:

$$\hat{\mu} = \frac{d}{E_x^c}$$

This is a maximum since $\frac{\partial^2}{\partial \mu^2} \log L(\mu) = -\frac{d}{\mu^2} < 0.$

$\hat{\mu}$ is the maximum likelihood estimate of $\hat{\mu}$. It is the realised value of the maximum likelihood estimator $\tilde{\mu} = \frac{D}{E_x^c}$

It is a property of all maximum likelihood estimators that they are asymptotically normally distributed. Therefore, the parameters of the statistical distribution can be fully expressed in terms of the mean and variance of $\tilde{\mu}$.

We can thus deduce the expectation and variance of $\tilde{\mu}$ as follows:

The observed number of deaths D has a Poisson (μE_x^c) distribution, which means that it has a mean μE_x^c and variance μE_x^c.

Therefore, $E[\tilde{\mu}] = E\left(\dfrac{D}{E_x^c}\right) = \dfrac{\mu E_x^c}{E_x^c} = \mu$; and

$$Var[\tilde{\mu}] = Var\left(\dfrac{D}{E_x^c}\right) = \dfrac{\mu E_x^c}{(E_x^c)^2} = \dfrac{\mu}{E_x^c}$$

Thus, the parameters for the statistical distribution of the MLE are:

$$\tilde{\mu} \sim Normal\left(\mu, \dfrac{\mu}{E_x^c}\right), \text{ where } \hat{\mu} = \dfrac{d}{E_x^c}$$

iii. **Mortality investigation**

a. Using the results from part (ii), the best estimate of μ is given by $\dfrac{D}{E_x^c}$

From the mortality investigation,

$$D = 182 + 217 + 180 = 579$$

$$E_x^c = 1814 + 1943 + 1732 = 5489$$

Therefore, $\hat{\mu} = {}^{579}/_{5489} = 0.105484$

i.e. the best estimate constant force of mortality is equal to 10.5%

b. We can use the parameters for the normal distribution to estimate the 99.5[th] percentile value:

$$\hat{\mu} + 2.576 \times \sqrt{\frac{\hat{\mu}}{E_x^c}} = 0.105484 + 2.576 \times \sqrt{\frac{0.105484}{5489}} = 0.116775$$

i.e. the 99.5th percentile value for constant force of mortality is equal to 11.7%

c. Under the Poisson model,

$$q_x = 1 - P[D = 0|v = 1] = 1 - e^{-\mu_x}$$

Therefore, we can determine:

The best estimate of $q_x = 1 - \exp(-0.105484) = 0.100111$; and

The 99.5th percentile $q_x = 1 - \exp(-0.116775) = 0.110215$

iv.   **Interpretation of results**

The results above suggest that the best estimate mortality assumption for the insurance company should be 10.01% compared with a 11.02% probability of death at the 99.5th percentile level.

From the insurance company's perspective, this implies that there is a 1-in-200 chance that the true mortality may be approximately 10% higher than the best estimate assumption.

The insurance company may allow for this uncertainty by either introducing a margin for prudence in its mortality basis or carrying out a sensitivity test with a 10% increase in mortality and holding sufficient risk capital to cover for the possibility that its actual mortality experience may turn out to be worse than the best estimate.

[15]

Question 9:

i.   **Advantages of central exposed to risk**.

Two advantages of central exposed to risk over initial exposed to risk are:

1. The central exposed to risk is simpler to calculate from the data typically available compared to the initial exposed to risk. Moreover, central exposed to risk has an intuitive appeal as the total observed waiting time and is easier to understand than the initial exposed to risk.

2. It is difficult to interpret initial exposed to risk in terms of the underlying process   being modelled if the number of decrements under study increase or the situations become more elaborate. On the contrary, the central exposed to risk is more versatile and it is easy to extend the concept of central exposed to risk to cover more elaborate situations.

ii. **Calculation of exposed to risk**.

   Rita

   Rita turned 30 on 1 October 2009, when she was already married. She died on 1 January 2010, 3 months after her 30th birthday.

   Thus, Rita's contribution to central exposed to risk = 3 months

   And contribution to initial exposed to risk = 1 year

   Sita

   Sitaturned 30 on 1 September 2011, when she was already married. Time spent under investigation, aged 30 last birthday by Sita was 1 September 2011 – 31 August 2012.

   Thus, Sita's contribution to both central and initial exposed to risk is 1 year.

   Nita

   Nita turned 30 on 1 December 2009 and married 2 months later. Therefore, she joined the investigation of married women on 1 February 2010. She divorced 9 months later, when she would be censored from the investigation of married women.

   Thus, Nita's contribution to both central and initial exposed to risk is 9 months.

   Gita

   Gita got married on 1 June 2011, at which time she was already past her 31st birthday. Therefore, she has spent no time during the investigation period as a married woman at age 30 last birthday.

   Thus, her contribution to both central and initial exposed to risk is nil.

iii. **Total exposed to risk**.

   Hence, total exposed to risk is:

   Central exposed to risk = 0.25 + 1 + 0.75 + 0 = 2 years.

   Initial exposed to risk = 1 + 1 + 0.75 + 0 = 2.75 years

   From the results above, it can be seen that the central exposed to risk is 2 years and the initial exposed to risk is 2.75 years. The approximation would suggest that the initial exposed to risk should be 2.5 years. However, this is

not a good approximation for the data provided as the approximation is based on the assumption that deaths would be evenly spread and thus can be assumed to occur half way through the year, on average. This also relies on an implicit assumption of a reasonably large data set. In the data above, there were only 4 lives, which is not statistically significant. Moreover, there was only one death, which occurred 3 months after the 30thbirthday. As a result of the statistical sparseness in the data, the approximation is seen not to work very well.

[12]