INSTITUTE OF ACTUARIES OF INDIA

CT4 – MODELS

OCTOBER 2009 EXAMINATION

INDICATIVE SOLUTION

Soln 1:

a)

- In graduation there is trade off between smoothness & adherence to data. A satisfactory graduation must provide a good balance between the two.

- 'Under Graduation' occurs when too much emphasis is given to goodness of fit. The 'under graduated' rates adhere more closely to crude rates at the cost of smoothness.

- 'Over Graduation' is the reverse when too much emphasis is given to smoothness at the cost of adherence to data (crude rates).

b). Four advantages

- Can give good results even when data is scanty
- Easy to make special allowance for features such as discontinuances.
- Naturally gives weight to be given to those ages where most data is available
- It allows scope for individual judgment

Four Disadvantages

- Requires high degree of skill
- Individual judgment can lead to bias and prejudice
- Can be difficult to achieve high degree of smoothness
- It is unclear as to how many degree of freedom are to be used when testing over all adherence to data by χ^2 test

[7]

Soln 2:

(i) The maximum likelihood estimate of a Poisson mean is the sample mean. Here it is

$$\frac{(1000*0+1200*1+600*2+200*3)}{3000}=1$$

The variance of the sample mean is the variance of a single observation divided by the sample size. For the Poisson distribution, the variance equals the mean, so the MLE of the variance is the sample mean. The estimated variance of the sample mean is 1/3000.

(ii) The 90% confidence interval is

$$1 \pm 1.645\sqrt{1/3000}$$

and lower end point of the confidence interval is $1-1.645\sqrt{1/3000}$ = 0.970.

(iii)

The Poisson model is an approximation to the two-state model.

While the two-state model can be specified so as to allow increments (i.e. lives entering the population), this is not possible for Poisson model.

The estimation of transition rates in the two-state model involves the measurement of two random variables – the observed number of decrements and the exposed to risk that give rise to these decrements.

The Poisson model assumes that the exposed to risk remains constant and estimation of the transition rates in the model only involves the measurement of the observed number of decrements.

[7]

Soln 3:

(i) Now, for all n and *i*, $\sum_{j=1}^{N} p_{ij}^{(n)} = 1$.

Hence

$$\lim_{n\to\infty} \sum_{j=1}^{N} p_{ij}^{(n)} = 1$$

$$\sum_{j=1}^{N} \lim_{n\to\infty} p_{ij}^{(n)} = 1$$

$$\sum_{j=1}^{N} \pi_j = 1$$

Also, for all n and I, $0 \le p_{ij}^{(n)} \le 1$.

Hence,

$$0 \le \lim_{n\to\infty} p_{ij}^{(n)} \le 1$$

$$0 \le \pi_j \le 1$$

(ii) Now, .

$$\pi_k = \lim_{n\to\infty} p_{ik}^{(n)}$$

$$= \lim_{n\to\infty} \left(\sum_{j=1}^{N} p_{ij}^{(n-1)} p_{jk}\right)$$

$$= \sum_{j=1}^{N} \left(\lim_{n\to\infty} p_{ij}^{(n-1)}\right) p_{jk}$$

$$= \sum_{j=1}^{N} \pi_j p_{jk}$$

(iii) Follows immediately from (ii). If $\lim_{n\to\infty} p_{ij}^{(n)}$ exists and only depends on j, for each j, then it is a stationary probability distribution for a Markov chain. Furthermore it will limit to this distribution, irrespective of the initial distribution of the chain.

[7]

Soln 4:

We know following

$$\begin{cases} P_0(t+\delta t) = (1-\lambda\delta t - o(\delta t))P_0(t) + (\mu\delta t + o(\delta t))P_1(t) + o(\delta t) \\ P_1(t+\delta t) = (1-\mu\delta t - o(\delta t))P_1(t) + (\lambda\delta t + o(\delta t))P_0(t) + o(\delta t). \end{cases}$$

Rearranging the terms, one gets

$$\begin{cases} \dfrac{P_0(t+\delta t) - P_0(t)}{\delta t} = -\lambda P_0(t) + \mu P_1(t) + (P_1(t) - P_0(t))\dfrac{o(\delta t)}{\delta t} \\ \dfrac{P_1(t+\delta t) - P_1(t)}{\delta t} = \lambda P_0(t) - \mu P_1(t) + (P_0(t) - P_1(t))\dfrac{o(\delta t)}{\delta t}. \end{cases}$$

Letting δt goes to zero, we get

$$\begin{cases} \dfrac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t) \\ \dfrac{dP_1(t)}{dt} = \lambda P_0(t) - \mu P_1(t). \end{cases}$$

Solving the above differential equations, we have

$$P_1(t) = \frac{1}{\lambda+\mu}(\mu e^{-(\lambda+\mu)t} + \lambda)$$

and

$$P_0(t) = \frac{1}{\lambda+\mu}(\mu - \mu e^{-(\lambda+\mu)t})$$

[8]

Soln 5:

(i) The general model is $\lambda(t;z_i) = \lambda_0(t).\exp(\beta z_i^{\,T})$

Where

$\lambda(t;z_i)$ is the hazard function at time t for the life

$\lambda_0(t)$ is the baseline hazard function at time t

Z_i is a row vector with the values of the covariates for the life

β is a row vector of parameters

Initially there is just one covariate so Z_i and β are just scalers.

Our aim is to estimate the value of the parameter β based on the data given. We can estimate β by maximizing the partial likelihood, which equals

$$L(\beta) = \prod_{j=1}^{k} \frac{\exp(\beta z_i^T)}{\sum_{i \in R(t_j)} \exp(\beta z_i^T)}$$

Where k is the number of deaths, t_j is the jth lifetime and $R(t_j)$ denotes the set of lives at risk at time t_j^-.

(ii)

Relative risk of a female non-smoker compared to male smoker

$$= \frac{e^0}{e^{\beta_1+\beta_2}} = e^{-\beta_1-\beta_2}$$

Which has partial derivatives $e^{-0.2}$ = 0.818731 at $\hat{\beta}_1 = 0.05 and \hat{\beta}_2 = 0.15$

The mean relative risk $\hat{\beta}$ = 0.818731

The variance of the relative risk is

$$= \begin{pmatrix} \frac{\partial L}{\partial \beta_1} & \frac{\partial L}{\partial \beta_2} \end{pmatrix} \text{Cov}(\beta_1, \beta_2) \begin{pmatrix} \frac{\partial L}{\partial \beta_1} \\ \frac{\partial L}{\partial \beta_2} \end{pmatrix}$$

$$= \begin{pmatrix} e^{-0.2} & e^{-0.2} \end{pmatrix} \begin{pmatrix} 0.0002 & 0.0001 \\ 0.0001 & 0.0003 \end{pmatrix} \begin{pmatrix} e^{-0.2} \\ e^{-0.2} \end{pmatrix}$$

$=0.0007\ e^{-0.4} = 0.000469$

The standard deviation is = 0.021662

Confidence limit is $\hat{\beta} \pm 1.96$ * standard deviation

Confidence limit 0.818731 $\pm$ 1.96 * 0.021662

The upper limit is = 0.818731 + 1.96 * 0.021662 = 0.8612

[9]

Soln 6:

a).

(i) $_{t+h}p_x^{11} = {}_t p_x^{11} \times {}_h p_{x+t}^{11}$

As we know that

$${}_h p_{x+t}^{11} + {}_h p_{x+t}^{12} + {}_h p_{x+t}^{13} = 1$$

$${}_{t+h} p_x^{11} = {}_t p_x^{11} \left(1 - {}_h p_{x+t}^{12} - {}_h p_{x+t}^{13}\right)$$

$$= {}_t p_x^{11} \left(1 - (\mu h + o(h)) - (\nu h + o(h))\right)$$

$$\frac{{}_{t+h}p_x^{11} - {}_t p_x^{11}}{h} = -(\mu + \nu)\, {}_t p_x^{11} + \frac{o(h)}{h}$$

$$\frac{\delta}{\delta t}\,{}_t p_x^{11} = \lim_{h \to 0} \frac{{}_{t+h}p_x^{11} - {}_t p_x^{11}}{h}$$

$$= -(\mu + \nu)\,{}_t p_x^{11}$$

$$\frac{\frac{\delta}{\delta t}\,{}_t p_x^{11}}{{}_t p_x^{11}} = \frac{\delta}{\delta t} \log {}_t p_x^{11}$$

so this equation gives

$$\frac{\delta}{\delta t} \log {}_t p_x^{11} = -(\mu + \nu)$$

Integrating wrt to "s" within 0 and t

$$\int_0^t \frac{\delta}{\delta s} \log {}_s p_x^{11}\, \delta s = -\int_0^t (\mu + \nu)\delta s$$

$$\log {}_t p_x^{11} - \log {}_0 p_x^{11} = -(\mu + \nu)t$$

$$\log {}_0 p_x^{11} = 0$$

$${}_t p_x^{11} = \exp(-(\mu + \nu)t)$$

(ii) $${}_{t+h}p_x^{12} = {}_t p_x^{11} \times {}_h p_x^{12} + {}_t p_x^{12} \times 1$$

$$= {}_t p_x^{11}(\mu h + oh) + {}_t p_x^{12}$$

$$\frac{\delta}{\delta t}\,{}_t p_x^{12} = \lim_{h \to 0} \frac{{}_{t+h}p_x^{12} - {}_t p_x^{12}}{h}$$

$$= \lim_{h \to 0} \left(\mu \times {}_t p_x^{11} + \frac{o(h)}{h} \right)$$

$$= \mu \times {}_t p_x^{11}$$

$$\frac{\delta}{\delta t}\,{}_t p_x^{12} = \mu \times \exp(-(\mu+\nu)t)$$

Integrating wrt t

$$\int_0^t \frac{\delta}{\delta s}\,{}_s p_x^{12}\,\delta s = \left[\frac{\mu}{\mu+\nu}\exp[-(\mu+\nu)s\right]_0^t$$

$${}_t p_x^{12} = \frac{\mu}{\mu+\nu}[1-\exp[-(\mu+\nu)t]$$

b). Assume that the probability of two or more transitions in the short time h is $o(h)$,then we can express the transition probabilities in terms of the forces of transition:

1) ${}_t p_x^{12} = \mu h + o(h)$

2) ${}_t p_x^{13} = \nu h + o(h)$

The Markov assumption: only current state affects the probabilities of future events.

[10]

Soln 7:
Using integrated form of Kolmogorov backward equations, we can write

$$p_{ij}(s,t) = \delta_{ij} e^{\int_s^t \sigma_{ii}(u)du} + \int_s^t \sum_{l \neq i} e^{\int_s^u \sigma_{ii}(r)dr} \sigma_{il}(u) p_{lj}(u,t)du$$

In this case i and j can just take two values A or D. In total there are four backward equations, we solve each in turn.

(i)

$$P_{DD}(s,t) = \delta_{DD} e^{\int_s^t \sigma_{DD}(u)du} + \int_s^t e^{\int_s^u \sigma_{DD}(r)dr} .\sigma_{DA}(u).P_{AD}(u,t)du$$

$$= e^{\int_s^t 0du} + \int_s^t e^{\int_s^u 0dr} .0.P_{AD}(u,t)du$$

$$= e^0$$

$$= 1$$

(ii)

$$P_{AD}(s,t) = \delta_{AD} e^{\int_s^t \sigma_{AD}(u)du} + \int_s^t e^{\int_s^u \sigma_{AA}(r)dr} .\sigma_{AD}(u).P_{DD}(u,t)du$$

$$= 0 + \int_s^t e^{-\int_s^u \mu(r)dr} .\mu(u).P_{DD}(u,t)du$$

$$= \int_s^t e^{-\int_s^u \mu(r)dr} .\mu(u).1.du$$

from earlier result. So

$$P_{AD}(s,t) = \int_s^t e^{-\int_s^u \mu(r)dr} .\mu(u)du$$

$$= \int_s^t \frac{d}{du}\left(-e^{-\int_s^u \mu(r)dr} \right).du$$

$$= -e^{-\int_s^t \mu(r)dr} - (-e^{-\int_s^s \mu(r)dr})$$

$$= -e^{-\int_s^t \mu(r)dr} + 1$$

(iii)

$$P_{DA}(s,t) = \delta_{DA} e^{\int_s^t \sigma_{DA}(u)du} + \int_s^t e^{\int_s^u \sigma_{DD}(r)dr} .\sigma_{DA}(u).P_{AA}(u,t)du$$

$$= 0 + \int_s^t e^{-\int_s^u \mu(r)dr} .0.P_{AA}(u,t)du$$

$$= \int_s^t 0.du$$

$$= 0$$

(iv)

And, finally,

$$P_{AA}(s,t) = \delta_{AA} e^{\int_s^t \sigma_{AA}(u)du} + \int_s^t e^{\int_s^u \sigma_{AA}(r)dr} .\sigma_{AD}(u).P_{DA}(u,t)du$$

$$= e^{-\int_s^t \mu(u)du} + \int_s^t e^{-\int_s^u \mu(r)dr} .\mu(u).P_{DA}(u,t)du$$

$$= e^{-\int_s^t \mu(u)du} + \int_s^t e^{-\int_s^u \mu(r)dr} .\mu(u).0.du$$

$$= e^{-\int_s^t \mu(u)du}$$

[12]

Soln 8:

Process of fitting a time homogenous simple stochastic Markov model to a set of observations:

(i) Estimating transition probabilities

- The first thing to fix when setting up a Markov model is the state space.
- Once the state space is determined, however, the Markov model must be fitted to the data by estimating the transition probabilities pij
- Denote by X_1, X_2 ...... , X_N the available observations and define:
  - n_i as the number of times t (1 ≤ t ≤ N − 1) such that $X_t = i$;
  - n_{ij} as the number of times t (1 ≤ t ≤ N − 1) such that X_t = i and $X_{t+1} = j$.
- Thus n_{ij} is the observed number of transitions from state *i* to *j*, n_i the observed number of transitions from state *i*.
- Then the best estimate of p_{ij} is $\hat{p}_{ij} = n_{ij} / n_i$
- If a confidence interval is required for a transition probability, the fact that the conditional distribution of N_{ij} given N_i is Binomial (N_i, p_{ij}) means that a confidence interval may be obtained by standard techniques.

(ii) Assessing the fit

- The next step is to ensure that the fit of the model to the data is adequate, or in other words to check that the Markov property seems to hold.

- Denote by n_{ijk} the number of times t ($1 \le t \le N - 2$) such that $X_t = i$, $X_{t+1} = j$ and $X_{t+2} = k$.

- **If** the Markov property holds we expect n_{ijk} to be an observation from a Binomial distribution with Parameters n_{ij} and p_{jk} .

- A simple but effective test, therefore, is the chi-square goodness-of- fit test based on the test statistic:

$$\chi^2 = \sum_i \sum_j \sum_k (n_{ijk} - n_{ij} \hat{p}_{jk})^2 / n_{ij} \hat{p}_{jk}$$

- The number of degrees of freedom of the distribution of χ^2 is given by the formula r-q+s-1, where:

  s denotes the number of states i in the state space such that $n_i > 0$

  q denotes the number of pairs (i, j) for which $n_{ij} > 0$ and

  r denotes the number of triplets (i, j,k) for which $n_{ij}\, n_{jk} > 0$

(iii) Simulation

- An additional method in frequent use for assessing goodness of fit is to run some simulations of the fitted chain and to compare graphs of the resulting trajectories with a graph of the process actually observed. This method often highlights deficiencies that are missed by the chi-square test.

- The Markov property means that the conditional distribution of X_{t+1} given the history of *X* up until time *t* is only dependent on X_t. The random walk, which has independent increments, is one such example; another might be a process which can only make transitions of the form x $\rightarrow$ x + 1 or x $\rightarrow$ x -1, with respective probabilities θ_x and 1-θ_x .
- A commercial simulation packages, which are able to simulate Markov chains without difficulty, even standard spreadsheet software can easily cope with the practical aspects of estimating transition probabilities and performing a simulation.

[13]

Soln 9:

a).

- A population under investigation in a mortality study is homogenous if all the lives have similar mortality characteristics.

- The problem with lack of homogeneity is that the investigation will reveal only the average experience of the group.

- The insurer will be too expensive for the good risks and hence lose their business. At the same time the insurer will be too cheap for the bad risks and hence end up in writing unprofitable business.

- This will be a problem if other insurers after taking into account of the factors causing the lack of homogeneity and price correctly.

b).

Factors usually considered are:

- Age
- Sex
- Type of policy
- Smoker/non-smoker status
- Level of Underwriting
- Duration in force
- Other that might be used are; sales channel, policy size, occupation etc

c).

We may make assumptions as under

A – Birthdays are uniformly distributed over the calendar year.
B – Birthdays are uniformly distributed over the policy year.
C – Policy anniversaries are uniformly distributed over the calendar year.

We define θ_x as the number of deaths aged x according to the given age definition.

- The event that causes an age level to change is the passing of 1st January. Hence the rate interval is a calendar year. The rate year for label x commences on 1st January at which a life is aged x next birthday. For a life to be counted in θ_x, the person must be between age x to x+1 at the previous 1st January i.e. at the start of the rate year. Using assumption A, the person is aged x-1/2 on average at the commencement of calendar year. In this case, we would be estimating $q_{x-1/2}$.

- It is the policy anniversary nearest to the death that determines the age label for a death and so this age definition leads to a policy year rate interval. The rate year for label x starts six months before the policy anniversary at which a life is aged x last birthday. For a life to be counted in θ_x the person must be aged between x and x+1 exact at the nearest policy anniversary. Using assumption B, the person will be aged x+1/2 on average on the nearest policy anniversary. For death nearest to that policy anniversary it must lie within 6 months on either side and must be between x and x+1 exact at death. So rate year starts at exact age x and we are estimating q_x.

- The age at entry is fixed as soon as the life takes out a policy and it is the duration that changes and affects the value of age level. Duration is defined as curtate duration at death and so the policy anniversary determines the value of the age level. Rate interval is the policy year beginning at the policy anniversary at which a life is aged x nearest birthday. Exact age at entry lies between y –1/2 and y +1/2 if y is assumed to be age at entry and using assumptions B, age at entry is y on average. At the last policy anniversary assuming curtate duration r, the exact duration was r with no assumption. At the start of the year, the average age is y + r = x and we estimate q_x.

[13]

Soln 10:

(i) Those lives who have gone back to their relatives during the period of investigation or who are still alive and resident in the old care home at the end of the period of investigation are right censored.

Duration is recorded to the nearest month so there is interval censoring.

Being released from old care home is a form of random censoring.

(ii) We can summarise the data to obtain the statistics necessary to complete the estimation.

j	t_j .	R_j	C_j	d_j .
	months	Risk set	No. censored	No of deaths
		$R_j - d_j - C_j$		
1	6	14	1	3
2	7	10	1	1
3	10	8	2	1
4	13	5	0	1
5	16	4	1	1
6	20	2	1	1

Then estimates of survival probabilities are

j	t_j	$S_j = S_{j-1} * (\frac{R_{j-1} - d_{j-1}}{R_{j-1}})$
1	6	1
2	7	1 x (14 - 3) / 14 = .7857
3	10	.7857 x (10 - 1) /10 = .7071
4	13	.7071 x (8 - 1) / 8 = .6188
5	16	.6188 x (5 - 1) / 5 = .4950
6	20	.4950 x (4 - 1) / 4 = .3712
7	23	.3712 x (2 - 1) / 2 = .1856

So estimates of the survival function is:

$$S(t) = \begin{cases} 1 & 0<=t<=6 \\ .79 & 6<t<=7 \\ .71 & 7<t<=10 \\ .62 & 10<t<=13 \\ .50 & 13<t<=16 \\ .37 & 16<t<=20 \\ .19 & 20<t<=23 \end{cases}$$

(iii)

- Non-informative censoring: Time to censoring i.e. leaving for reasons other than death is independent of time to death.

- If we believe that more healthy lives will tend to leave and thus have lighter mortality, this assumption is violated. [OR could say less healthy lives may tend to leave — same conclusion.]

- However if we believe that being released is unrelated to state of health, then assumption is OK.

- Lives are independent i.e. time to censoring, i.e. being released, or time to death determined independently for each life.

[14]
Total [100] Marks

****