INSTITUTE OF ACTUARIES OF INDIA

CT4 – Models

May 2009 Examination

Indicative Solution

Soln.1

a). If data on exact times of death are available, the two-state model uses all the information available (i.e. the times of death), while the Binomial model represents a restricted view of the process, since it represents only the year of death and not the time of death.

b). The Binomial model requires estimation of *q* and some assumption about the distribution of deaths with age in order to calculate μ; the two-state model does not.

c). The two-state model is extended very simply to processes with more than one decrement (i.e. to a multiple-state model), and to processes with increments and decrements. The Binomial model is not.

[3]

Soln.2

(i) Gompertz Law is a suitable model for human mortality for middle to older ages say 35 and over.

There is evidence that the Gompertz Law breaks down at very advanced ages and therefore 35 to 90 years is acceptable.

(ii) Since $_tP_x = \exp\left(-\int_0^t \mu_{x+s}\,ds\right)$

Putting $\mu_x = Bc^x$

$$_tP_x = \exp\left(-\int_0^t Bc^{x+s}\,ds\right)$$

We can write $c^{x+s} as \longrightarrow c^x e^{s \log c}$

$$\int_0^t Bc^{x+s}\,ds = \int_0^t Bc^x e^{s\log c}\,ds = \frac{Bc^x}{\log c}\left[e^{s\log c}\right]_0^t$$

$$\frac{Bc^x}{\log c}\left[e^{s\log c}\right]_0^t = \frac{Bc^x}{\log c}\left[c^s\right]_0^t = \frac{Bc^x}{\log c}\left[c^t - 1\right]$$

If we introduce the auxiliary parameter g defined by log g = -B/ log c, the value of the integral is –log g $c^x(c^t - 1)$ and hence

$$_tP_x = \exp\left(\log\ gc^x\left[c^t - 1\right]\right) = \left(e^{\log g}\right)^{c^x(c^t-1)} = g^{c^x(c^t-1)}$$

[4]

Soln.3

Life Censoring

Data in this study would be left censored if the censoring mechanism prevent us from knowing when the policyholder joined the company.
This is not present because the policy issue date is given.

Right Censoring

Data would be right censored if the censoring mechanism cuts short observations in progress, so that we are not able to discover if and when the policy is surrendered.

Data in this study would be right censored if the policy is terminated before the maturity date for reasons than surrender.

Interval Censoring

Data in this study would be interval censored if the observational plan only allows us to say that the duration of policy at the time of surrender fell within some interval of time.
Here we know the calendar year of surrender and the policy issue date, so we will know that the duration of the policy falls within one year rate interval. Interval censoring is present.

Informative Censoring

Censoring in this study would be informative if the censoring event divided individuals into two groups whose subsequent experience was thought to be different.

Here the censoring event of surrendering the policy might be suspected to be informative, as those who are likely to surrender the policy to be in better health than those who do not surrender the policy.

[7]

Soln. 4

a). A stochastic model allows for the randomness of the input parameters.

Stochastic model have following advantage over deterministic model:

- a stochastic model provides the distribution of the results (probabilities and variances) and not just a single best estimate.

- Stochastic model correctly reflects the random nature of the variables involved as against deterministic one.

- Stochastic model allow to use Monte Carlo simulation which is a powerful technique to solve complex problem.

b).

(i) Assume that the functions p_{ij} (s,t) are continuously differentiable, the transition rates are defined by differentiation with respect to t.

$$\sigma_{ij}(s) = \left[d/d_t \left(p_{ij}(s,t) \right) \right]_{t=s}$$

(ii) Since

$$\sum_{j \in s} p_{ij}(s,t) = 1 \text{ where S - State Space}$$

$$\text{We have } \sum_{j \in s} \sigma_{ij}(s) = \sum_{j \in s} \left[d/d_t \; p_{ij}(s,t) \right]_{t=s}$$

$$= d/d_t \sum_{j \in s} p_{ij}(s,t)$$

$$= d/d_t (1)$$

$$= 0$$

[8]

<u>Soln. 5</u>

a). We have that s_1, the state of the system in period 1, is given by s_1=[0.55, 0.45] with s_2=[0.67, 0.33] and s_3=[0.70, 0.30]

Assuming the researcher is correct then s_1 and s_2 are linked by $s_2 = s_1$ (P) where P is the transition matrix as to how buyer changes buying preference. We also have that s_3 and s_2 are linked by $s_3 = s_2$ (P).

Now we have that P will be a 2 by 2 matrix, and that the elements of each row of P add to one, so that we can write

$$P = \begin{vmatrix} x_1 & 1-x_1 \\ x_2 & 1-x_2 \end{vmatrix}$$

Where x_1 and x_2 are unknown.

Using $s_2 = s_1$(P) we have

$$[0.67, 0.33] = [0.55, 0.45] \begin{vmatrix} x_1 & 1-x_1 \\ x_2 & 1-x_2 \end{vmatrix}$$

$$[0.70, 0.30] = [0.67, 0.33] \begin{vmatrix} x_1 & 1-x_1 \\ x_2 & 1-x_2 \end{vmatrix}$$

Hence, expanding, we have

$0.67 = 0.55x_1 + 0.45x_2$ (1)

$0.33 = 0.55(1 - x_1) + 0.45(1 - x_2)$ (2)

$0.70 = 0.67x_1 + 0.33x_2$ (3)

$0.30 = 0.67(1 - x_1) + 0.33(1 - x_2)$ (4)

Equation (2), when rearranged, becomes equation (1) and similarly equation (4), when rearranged, becomes equation (3). Hence we have two simultaneous equations (equations (1) and (3)) in two unknowns.

From equation (1)

$x_1 = (0.67 - 0.45x_2)/0.55$

so substituting for x_1 in equation (3) we get

$0.70 = 0.67[(0.67 - 0.45x_2)/0.55] + 0.33x_2$

i.e. $(0.70)(0.55) = (0.67)(0.67) - (0.67)(0.45)x_2 + (0.33)(0.55)x_2$

i.e. $x_2 = [(0.67)(0.67)-(0.70)(0.55)]/[(0.67)(0.45)-(0.33)(0.55)]$

i.e. $x_2 = 0.5325$ and
$x_1 = (0.67 - 0.45x_2)/0.55 = 0.7825$

Hence our estimate of the transition matrix P is equal to

$$P = \begin{vmatrix} 0.7825 & 0.2175 \\ 0.5325 & 0.4675 \end{vmatrix}$$

b). The market shares for period 4 are given by

$s_4 = s_3(P)$

$$[0.70, 0.30] \begin{vmatrix} 0.7825 & 0.2175 \\ 0.5325 & 0.4675 \end{vmatrix}$$

i.e. $s_4 = [0.7075, 0.2925]$

and note here that the elements of s_4 add to one (as required).
Hence the estimated market shares of two products in period 4 are 70.75% and 29.25%.

c). If the actual market shares are 71% and 29% then this compares well with the shares estimated above and so there would seem no reason to revise the estimate of the transition matrix.

[8]

Soln.6

a).

$$A = \begin{array}{c} \\ \text{Full} \\ \text{Partial} \\ \text{None} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{F} & \mathbf{P} & \mathbf{N} \end{array} \\ \begin{bmatrix} .4 & .35 & .25 \\ .45 & .4 & .15 \\ .8 & .15 & .05 \end{bmatrix} \end{array}$$

b).

$$A^2 = \begin{array}{c} \mathbf{F} \\ \mathbf{P} \\ \mathbf{N} \end{array} \begin{bmatrix} .5175 & .3175 & .165 \\ .48 & .34 & .18 \\ .4275 & .3475 & .225 \end{bmatrix}$$

Based on above matrix answer is 0.18

c).

$$\begin{array}{c} \mathbf{F} \\ \mathbf{P} \\ \mathbf{N} \end{array} \begin{bmatrix} .49 & .33 & .18 \\ .49 & .33 & .18 \\ .49 & .33 & .18 \end{bmatrix}$$

d). If we are looking far enough into the future (a few weeks or longer), it doesn't matter what kind of assignment we have today. We have a 49% chance of having a full assignment, a 33% chance of having a partial assignment and an 18% chance of not having an assignment.

[8]

Soln. 7

i. $h(x,t) = h_0(t).\exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k)$

where $h(x,\ t)$ is the hazard at duration t, $h_0(t)$ is some unspecified baseline hazard, $x_1\ ....x_k$ are covariates and $\beta_1 .....\ \beta_k$ are their associated parameters.

ii. Women who were 'unable to care for themselves' at the time of diagnosis, who were given the 'existing' treatment, and whose tumors were of the 'Inflammatory' type.

iii. The value of the parameter associated with the new treatment is {-0.2}. This implies that the ratio of the hazards of death for two otherwise identical patients, one of whom is given the new treatment and the other the existing treatment is exp (-0.2) = 0.819 i.e. we estimate that the risk of death associated with new treatment is 18.1% lower than that of existing treatment. Thus the new treatment appears to decrease the risk of death.

Further, the standard error associated with the parameter is 0.15. The approximate 95% confidence interval is therefore -0.2 ± 1.96(0.15) = (-0.494, 0.094), which includes 0. Therefore, the value of the parameter is not significantly different from zero at the 5% level, so it is not possible to say with the available data whether the new treatment affects the risk of death.

iv. The hazard for women with ‘Invasive’ type tumors who were ‘able to care for themselves’ at the time of diagnosis is $h_0(t)\exp(-0.60 + 0.45)$.

The hazard for women with ‘Inflammatory’ type tumors who were ‘unable to care for themselves’ at the time of diagnosis is $h_0(t)$, since this is the baseline category. The ratio is thus

$$\frac{h_0(t)\exp(-0.60 + 0.45)}{h_0(t)} = \exp(-0.15) = 0.8607$$

So the risk of death is 14% lower for women with ‘Invasive’ type tumors who were ‘able to care for themselves’ at the time of diagnosis.

[11]

Soln. 8

(i) $H0$: the true underlying mortality of the annuitants is that of the standard table.

Age:	*Exposed to risk:*	*Standard Mortality rates*	*Expected Deaths*	*Actual deaths*	*A-E*	*Stand. dev*	*Chi-squared*
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8] =([6]/[7])^2
60	600	0.0109	6.54	7	0.46	2.5434	0.0327
61	750	0.0117	8.78	8	-0.78	2.9449	0.0693
62	725	0.0118	8.56	8	-0.56	2.9076	0.0364
63	650	0.0121	7.87	7	-0.87	2.7874	0.0963
64	700	0.0128	8.96	7	-1.96	2.9741	0.4343
65	675	0.0139	9.38	8	-1.38	3.0417	0.2066
Total	4100		50.08	45.00			0.876

Degrees of freedom for chi-square test = 6 (number of ages)

Table value of Chi-square (6, 95%) = 12.59 which means sum of chi squared (0.876) is less than 12.59. So there is no evidence to reject Ho.

All deviations but one are negative, which could indicate that the true mortality is lighter than the standard table. This is not detected by the Chi-squared test as the statistic is based on squared deviations. Also, the Chi-square distribution provides a good approximation provided the numbers in each group are not too small.

(ii) If the true mortality is lighter than the mortality assumption used for pricing (normally expressed as a percentage of standard table), the company will charge inadequate premiums and will suffer a loss on the policies. However, with the given exposure, it is early to make any conclusion based on current experience.

(iii) For testing adherence to data, the test statistic and process would remain unchanged, but the number of degrees of freedom will get reduced.

In fitting the relationship two parameters have been estimated so the number of degrees of freedom will be reduced from 6 to 4, with a further reduction of degrees of freedom (say 2 or 3) for the constraints imposed by the choice of the standard table.

(Need to award marks based on justification given for each cause of reduction in degrees of freedom).

(iv)

Reasons why crude rates will require graduation:

- by graduation, rates at extreme ages could be estimated reliably referring rates at nearby ages.
- overall low data volumes mean that crude rates are likely to be subject to relatively large sampling errors and therefore will not progress smoothly with age, so need graduation of crude rates
- Volume of data is too small to attempt a direct graduation so ruling out use of parametric formula.
- Also, large sampling errors would make graphical graduation imprecise and in any case computationally inefficient.

So suggest we graduate via some simple relationship to a standard table, many of which exist based on large volumes of data relating to similar lives. This approach also allows us to compare our population with the population underlying the standard table.

[13]

Soln.9

The solutions to all subsections are given in a tabular format below:

Deaths Vs matching initial Exposed -to-risk	Rate interval	Deaths-d50	Initial exposed-to-risk - E50	Central exposed-to-risk - $E50^{C}$	Estimate of q_{50+f}	Estimate of m_{50+f}	Value of "f"
${}^{A}d_{50} \sim {}^{2}E_{50}$	Calendar Year (Starting at 1st January where lives aged x nearest birthday)	60	12300	12270	0.004878	0.00489	(-1/2<f<1/2);take f=0 assuming that birthdays are uniformly spread over calendar year
${}^{B}d_{50} \sim {}^{3}E_{50}$	Policy Year (Starting at policy anniv. where lives aged x last birthday)	66	11600	11567	0.00569	0.0057059	(0<f<1);take f=1/2, assuming that birthdays are uniformly spread over policy year
${}^{C}d_{50} \sim {}^{1}E_{50}$	Policy Year (Starting at policy anniv. where lives aged x nearest birthday)	62	12200	12169	0.005082	0.0050949	(-1/2<f<1/2);take f=0, assuming that birthdays are uniformly spread over policy year

[12]

Soln. 10

a) Take *S* = {0, 1, 2,3,4, 5} as the discrete state space of the random process X_t with X_t = *i* *(i=0,1…..5)* indicating that the discount is 10% * X_t at the end of the year *t*. We set X_0 = 0 as there is no discount initially. The X_t forms a Markov Chain since we can find the relevant transition probabilities. Namely, if X_t = *i* then either X_{t+1} = 0 with probability *p* (the claim was made during the year *t*+1) or X_{t+1}= min(*i*+1,5) with probability 1- *p* (no claims during the year *t* +1).

b) We already know that P_{i0} = *p* for any *i* ∈ *S*, P_{55} = 1- *p*, and $P_{i,i+1}$ = 1- *p* if 0 ≤ *i* ≤ 4.

Hence

$$P = \begin{pmatrix} p & 1-p & 0 & 0 & 0 & 0 \\ p & 0 & 1-p & 0 & 0 & 0 \\ p & 0 & 0 & 1-p & 0 & 0 \\ p & 0 & 0 & 0 & 1-p & 0 \\ p & 0 & 0 & 0 & 0 & 1-p \\ p & 0 & 0 & 0 & 0 & 1-p \end{pmatrix}$$

c) The probability that X_4 = 2 given that X_0 = 0 can be found by considering all paths to year 4 from state 0 to state 2. We obtain

$$P\{X_4 = 2 \mid X_0 = 0\} = P_{00}\, P_{00}\, P_{01}\, P_{12} + P_{01}\, P_{10}\, P_{01}\, P_{12}$$

$$= p^2(1-p)^2 + (1-p)p(1-p)^2$$

$$= p(1-p)^2[p+1-p]$$

$$= p(1-p)^2$$

d)

The answer is *P* { X_{16} = 0} = *p*.

Indeed, we have

$$P\{X_{16} = 0\} = \sum_{i=0}^{5} P\{X_{15} = i\} P\{X_{16} = 0 \mid X_{15} = i\}$$

$$= \sum_{i=0}^{5} P\{X_{15} = i\} P_{i0} = p \sum_{i=0}^{5} P\{X_{15} = i\} = p$$

Since p_{io} = *p for all i and* $\sum_{i=0}^{5} P\{X_{15} = i\}$ = *1*

[13]

Soln. 11

a). A Poisson process with rate λ is a continuous –time integer –valued process N_t , $t \geq 0$ with the following properties ;

- $N_0 = 0$
[1/2]
- N_t has independent increments
[1/2]
- N_t has Poisson distributed stationary increments:

$$P[Nt-Ns = n] = [\lambda\,(t-s)]^n e^{-\lambda(t-s)}/n!\,, \qquad s<t,\ n = 0,1, \ldots\ldots$$

b). As per the independent and stationary increments property of the Poisson process:

$$\Pr(N3 = 8; N7 = 12) = \Pr(N3 = 8; N7 - N3 = 4)$$

$$= \Pr(N3 = 8)\Pr(N7 - N3 = 4)$$

$$= \frac{15^8 \exp(-15)}{8!}\,\frac{20^4 \exp(-20)}{4!}$$

$$= \frac{15^8\, 20^4 \exp(-35)}{8!\,4!}$$

$$= 2.67 \times 10^{-07}$$

c).

As per the independent, stationary increments property of Poisson process:

$$\Pr(N_{k1} = r | N_{k2} = n) = \Pr(N_{k1} = r, N_{k2} = n)/\Pr(N_{k2} = n)$$

As in the previous problem:

$$\Pr(N_{k2} = n) = (\lambda k2)^n \exp(-\lambda k2)/n!$$

$$\Pr(N_{k1} = r; N_{k2} = n) = \Pr(N_{k1} = r; N_{k2} - N_{k1} = n-r)$$

$$= \Pr(N_{k1} = r)\Pr(N_{k2} - N_{k1} = n-r)$$

$$= \frac{(\lambda k1)^r \exp(-\lambda k1)}{r!}\,\frac{(\lambda(k2-k1))^{n-r} \exp(-\lambda(k2-k1))}{n-r!}$$

$$= \frac{\lambda^n\,(k1)^r\,(k2-k1)^{n-r} \exp(-\lambda k2)}{r!\quad n-r!}$$

$$\Pr(N_{k1} = r | N_{k2} = n) = \frac{n!}{r!\,n-r!}\,\frac{(k1)^r\,(k2-k1)^{n-r}}{(k2)^n}$$

$$= \frac{n!}{r!\,n-r!}\,\frac{(k1)^r\,(k2-k1)^{n-r}}{(k2)^r\,(k2)^{n-r}}$$

$$= \frac{n!}{r!\,n-r!}\,\frac{(k1)^r}{(k2)^r}\,\frac{(k2-k1)^{n-r}}{(k2)^{n-r}}$$

$$= \binom{n}{r} p^r (1-p)^{n-r} \quad , \quad \text{where } p=k1/k2$$

d).

$$E(Y_t) = E(E(\sum_{k=1}^{N_t} X_k / N_t))$$

$$= E(E(\sum_{k=1}^{N_t} E(X_k)/N_t))$$

$$= E(mN_t)$$

$$= m \lambda t$$

$Var(Y_t) = Var(E(Y_t/N_t)) + E(Var(Y_t/N_t))$ as X_k s are IID variables

Since, $Var(Y_t/N_t=n) = n S^2$

So,
$Var(Y_t) = Var(N_t m) + E(N_t S^2) = m^2 Var(N_t) + S^2 E(N_t) = \lambda t (m^2 + S^2)$

[13]

[Total 100 Marks]

****