Actuarial Society of India

Examinations

May 2006

CT4 (104) – Survival Models

Indicative Solutions

NOTE: Some of the candidates pointed out that the current tables do not contain a(55) table. In case of question 1, if any appropriate table other than a(55) is used and if the candidate has correctly interpreted the results, full marks has been awarded.

Que 1)

(i) H_0 : The true underlying mortality of the annuitants is that of the standard table

Age (x)	Exposed to Risk (Ex)	Observed Deaths (èx)	q_x (from tables)	Ex q_x	z_x= (èx - Ex q_x) / Ex q_x(1- q_x)	z^2_x
70	600	23	0.03776	22.656	0.073676	0.00543
71	750	31	0.04170	31.275	-0.050232	0.00252
72	725	33	0.04602	33.3645	-0.064608	0.00417
73	650	29	0.05075	32.9875	-0.712583	0.50778
74	700	35	0.05595	39.165	-0.684965	0.46918
75	675	39	0.06164	41.607	-0.417228	0.17408

Degrees of freedom for ÷2 test = number of ages = 6

One tailed test as large values of Óz^2_x indicate excessive deviations.

÷$2_{6,0.95}$ = 12.59

Óz^2_x < 12.59 , so there is no evidence to reject H_0

(6)

All deviations but one are negative, which could indicate that the true mortality is lighter than a(55). This is not detected by the Chi-Squared test as the statistic is based on squared deviations.

(2)

(ii) Strengths & Weaknesses:

- It is a good test for overall goodness of fit.

Weaknesses:

- There could be a few large deviations offset by a lot of very small deviations.
- The graduation might be biased above or below the data since the Chi-squared statistic is based on squared deviations.
- Even if the graduation is not biased, there could be significant groups of consecutive runs over which it is biased up or down.

(3)

(iii) The further tests that could be done to over come the above deficiencies are:

- Standardised deviations test
- Signs test
- Cumulative deviations test

(2)

Que 2)

Proof of Gompertz' Law:

We know that $_tp_x = \exp[-_0{}^t \mu_{x+s}\, ds] = \exp[-_0{}^t Bc^{x+s}\, ds]$
We know c^{x+s} as $c^x e^{s\log c}$, so that the integral becomes

$$_0{}^t B\, c^x e^{s\log c}\, ds] = \frac{B\, c^x}{\log c}[e^{s\log c}]_0^t = \frac{B\, c^x}{\log c}[c^s]_0^t = \frac{B\, c^x}{\log c}[c^t - 1]$$

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 we find that:

$$_tp_x = \exp[\log g.c\text{^}x(c\text{^}t - 1)] = (e^{\log g})\text{^}c\text{^}x(c\text{^}t-1) = g\text{^}c\text{^}x(c\text{^}t-1).$$

Que 3)
a) The central exposed to risk takes account of the time up to exist , and thus is independent of the decrement being investigated. The initial exposed to risk counts time until the end of the rate interval if the life leaves by the cause under consideration. The initial exposed to risk can therefore count days outside the period of investigation.

(2)

b) (i)

Life A Contributions:

$E^c_{38,7}$ = 31+28+31 = 90 days

$E^c_{38,8}$ = 30+31+5 = 66 days

Life B Contributions:

$E^c_{37,7}$ = 31+28+31+30+1 = 121 days

$E^c_{38,7}$ = 30+30+31+31+30 = 152 days

$E^c_{38,8}$ = 31+29 = 60 days

Life C Contributions:

$E^c_{38,0}$ = 31+30+31+30+31 = 153 days

(5)

(ii) Life A Contribution:

$E^{c}_{38,7}$ = 31+28+31 = 90 days

$E^{c}_{38,8}$ = 30+31+5 = 66 days

Life B Contributions:

$E^{c}_{38,7}$ = 121+ 152 = 273 days
$E^{c}_{38,8}$ = 60 days

Life C Contributions:

$E^{c}_{38,0}$ = 153 days

(3)
[10]

Que 4)

a) There will be right censoring of all the lives that survive to age 45 or who withdrew before age 45.

There will be random censoring of all the lives that withdraw before death or attaining age 45.

(2)

b)

Person	Duration(months)
1	6
2	6+
3	12
4	12
5	18+
6	27
7	27+
8	27
9	30+
10	36+
11	39
12	39+
13	51+
14	54+
15	57
16	60+

17	60+
18	60+
19	60+
20	60+

So the times of death are 6,12,27,39 and 57. the initial risk set is 20

1	2	3	4	5	6
j deaths	t_j times	d_j deaths	C_j censored	$n_{j=} n_{j-1}- d_{j-1}- C_{j-1}$ j > 0 (risk set)	$S_j = (n_j- d_j)/ n_j$
0	0	0	0	20	(20-0)/20 =1
1	6	1	1	20	(20-1)/20 =0.950
2	12	2	1	18	(18-2)/18=0.889
3	27	2	3	15	(15-2)/15 =0.867
4	39	1	3	10	(10-1)/10=0.900
5	57	1	5	6	(6-1)/6 =0.833

Then the estimate of the survival function is:

Time	0<=t<6	6<=t<12	12<=t<27	27<=t<39	39<=t<57	57<=t
Probability	1	1*0.95= 0.95	0.95*0.889=0.844	0.844*0.867= 0.732	0.732*0.9 =0.659	0.659*0.833=0.549

Standard error of these estimates will be given by Greenwood's formula.

1	2	7	8
j deaths	t_j times	$d_j/(n_j- d_j) n_j$ deaths	Ó $d_j/(n_j- d_j) n_j$ t_j<=t
0	0	0/(20-0)=0	0
1	6	1/(20-1)20=0.00263	0.00263
2	12	2/(18-2)18=0.00694	0.00957
3	27	2/(15-2)15=0.01026	0.01983
4	39	1/(10-1)10=0.01111	0.03094
5	57	1/(6-1)6=0.03333	0.06427

Time	0<=t<6	6<=t<12	12<=t<27	27<=t<39	39<=t<57	57<=t
Probability	1	1*0.95=0.95	0.95*0.889=0.844	0.844*0.867=0.732	0.732*0.9=0.659	0.659*0.833=0.549
Standard error	1* 0 =0	0.95* 0.00263 =0.0487	0.844* 0.00957 =0.0826	0.732* 0.01983 =0.1031	0.659* 0.03094 =0.1159	0.95* 0.06427 =0.1392

-

Then maximum likelihood estimate of Survival Function with approximate 95% confidence intervals:

Time	0<=t< 6	6<=t<12	12<=t<27	27<=t<39	39<=t<57	57<=t
Probability	1	1*0.95=0.95	0.95*0.889=0.844	0.844*0.867 =0.732	0.732*0.9 =0.659	0.659*0.833 =0.549
Approximate 95% confidence interval	1±0 = 1	0.95 ± 0.096 =0.854 ,1*	0.844 ± 0.162 =0.682 ,1*	0.732 ± 0.202 =0.530,0.934	0.659 ± 0.227 =0.432,0.886	0.549 ± 0.0.273 =0.276,0.822

Since probability cannot exceed 1.

(10)
[12]

Que 5)
(i) Proportional Hazards:

For example, if we take the ratio μ(t,i)/ μ(t,j) , then

exp[á(x_i - 70) + â.y_i + ã. z_i] given time t, the force of mortality of a life having any particular set of characteristics(x_j , y_j & z_j) is a constant proportion of the force of mortality of a life having another particular set of characteristics (x_j , y_j & z_j).

(2)

(ii) Subdividing into homogeneous groups:

The main reason is that the models of mortality that we use assume that all the lives involved experience mortality that is consistent with a particular probability distribution, i.e. that the lives are homogeneous with respect to their mortality.

Subdividing the data according to characteristics that are considered to affect mortality aims to achieve homogeneity within the sub-groups obtained.

In this case, older lives usually exhibit higher mortality than younger lives, males usually exhibit higher mortality than females, and smokers usually exhibit higher mortality than non-smokers. Hence we would expect each of the parameters á, â and ã to be positive.

(3)

(iii) Calculation of force of mortality:

The value of the exponential equals 1, so the force of mortality during 2007 for this groups is (0.015-0.0001*7) = 0.0143.

(1)

(iv) (a) Partial Likelihood:

$$L (á, â , ã) = Ï_{deaths} \exp[á(x_j - 70) + â.y_j + ã. z_j] / Ó \exp[á(x_i - 70) + â.y_j + ã. z_j$$

Where the summation in the denominator is the sum over all the lives that could have died at the time of j's death, and the product is over all the deaths observed.

Given the assumption stated in the question, then the denominator is the same regardless of the time of death i.e.;

$$800 e^{-5á + ã} + 200 e^{-10á + â + ã} + 450 + 150 e^{-5á + â} = E (á, â , ã)$$

Hence $L (á, â , ã) = (e^{-5á + ã})^6 (e^{-10á + â + ã})^5 (e^{-5á + â}) / (E (á, â , ã))^{14}$

(v) (b) Estimating the parameters:

The values of á, â and ã would be found that together maximized the value of L (á, â , ã) . This would be done using numerical techniques.

(6)
[12]

****