The Institute of Actuaries of India

Subject CT4 – Models

16th May 2007

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.

Arpan Thanawala
Chairperson, Examination Committee

Solution 1

(a) A graduation is an attempt to obtain estimates of the true underlying rates of mortality of any observed experience.

As such it is assumed that the underlying rates progress smoothly from age to age, and the extent of smoothness is defined by the choice of function used to perform the graduation.

Graduations also have to satisfy requirements of adherence to data, such that the observed deviations between the actual and graduated mortality rates can be explained as due to purely random effects.

Different functions impose different hypotheses with regard to the progression of the true mortality rates from age to age, and their suitability can only be properly assessed after they have been fitted to the data and the adherence can be analyzed. It is therefore necessary to fit a range of feasible functions to the data and to select the graduation, which best meets, the required criteria for smoothness and adherence.

(b) Need to find values of α and β, which minimize:

$$Q = \sum E_x q_x^s [\, q_x / q_x^s - \alpha - \beta x\,]^2$$

Now

$$\frac{\delta Q}{\delta \alpha} = -2 \sum E_x q_x^s [\, q_x / q_x^s - \alpha - \beta x\,]^2$$

$$\frac{\delta Q}{\delta \beta} = -2 \sum x E_x q_x^s [\, q_x / q_x^s - \alpha - \beta x\,]^2$$

Setting both (1) and (2) to zero we obtain

$$\sum x \vartheta_x - \alpha \sum E_x q_x^s - \beta \sum x E_x q_x^s = 0$$

and

$$\sum x \vartheta_x - \alpha \sum x E_x q_x^s - \beta \sum x^2 E_x q_x^s = 0$$

(c) Reference to standard table
Shape (progression) of the standard tabular rates will be broadly reproduced in the graduation. The experience is unlikely to display exactly the same features as the standard table and hence adherence to data is likely to be poor.

Mathematical formula
Adherence to data should be improved, as the form of any other experience does not influence graduation: i.e. method is more objective, but relies on choice of formula.

It is hard to find a single mathematical function, which can adequately model mortality rates over a large age range, due to the different and complex influences on mortality at different ages. This will reduce the potential adherence to data of many functions.

Both methods can produce high levels of precision and adequate smoothness.

G Total [8]

Solution 2

i) Choose an appropriate curve, e.g.,

a) $P(t\ ) = ce^{rt}$ (any appropriate function would be acceptable)

Where $P(t)$ is the number of pension policies at time t = year-2003

r and c are constants

b) Fit to data using least squares linear regression, or alternatives, e.g.,

$$\ln P(t) = \ln(c) + rt$$

$$\rightarrow \hat{c} \text{ and } \hat{r}$$

c)

$$P(5) = \hat{c}\, e^{5\hat{r}}$$

ii) Limitations
- Assumes past trend will continue indefinitely – ignores limiting factors. If the data is from the early stage of the epidemic (which it is) then the long-term rate of growth will be grossly over-estimated.

– Assumes the population is homogeneous (ignores the existence of heterogeneous sub-groups in that, had these been projected separately, a different answer would have resulted).

iii) If we assume that the number of pension policies, remain at the maximum indefinitely then the formula can be modified by assuming

$$1/p(t) \times dp(t) / dt = r - kp(t)$$

$$\text{ie} \quad 1/p(t) = ce^{-rt} + k/r$$

Thus in the long run

$r/k = P(t) = 1,00,000$

and the model becomes $P(t) = (ce^{-rt} + 0.00001)^{-1}$

G Total [8]

Solution 3

i) *SMR = Actual /Expected*

$$= \frac{\sum E_x^c \, m_x}{\sum E_x^c \, {}^s m_x}$$

$$= \frac{\sum E_x^c \, {}^s m_x \times m_x / {}^s m_x}{\sum E_x^c \, {}^s m_x}$$

Where $\mathbf{E}_x^c$ represent the population rates in the study population, $\mathbf{m}_x$ represent the population and mortality rates in the study population and ${}^s\mathbf{m}_x$ represents the mortality rates in the standard table .

This is the weighted average with $E_x^c \, {}^s m_x / \sum E_x^c \, {}^s m_x$ as weights.

ii)

SMR (2003 and 2004) = 1.0625
SMR (2005 and 2006) = 1.0625

iii) The CMF is simply a weighted average of the ratio of the mortality rates where the weights are

$$ {}^{s}E^{c}_{x} \; {}^{s}m_{x} \,/\, \sum {}^{s}E^{c}_{x} \; {}^{s}m_{x} $$

The difference between the SMR and CMF indicates that the populations have different proportions in the two age ranges.

As the CMF<SMR, this indicates that the population 20 years ago was more heavily weighted to the older age group.

Probably one would prefer the SMR, since although the population is probably still changing it is unlikely that the population in each of the two periods would be very different from the aggregate over the two periods. However, this could be significantly different from the population 20 years earlier and hence the SMR would give a better measure.

G Total [8]

Solution 4

i) Let D_i be the number of deaths among $\pi_i N$ lives with i policies.

Then $D_i \sim \textit{Binomial}\ (q_x,\ \pi_i N)$, so

$$E[C_i] = E\left[\sum_{i=1}^{\infty} iD_i\right] = \sum_{i=1}^{\infty} iE[D_i] = \sum_{i=1}^{\infty} i\pi_i Nq_x$$

$$Var[C_i] = Var\left[\sum_{i=1}^{\infty} iD_i\right] = \sum_{i=1}^{\infty} i^2 Var[D_i] = \sum_{i=1}^{\infty} i^2 \pi_i Nq_x(1-q_x)$$

ii) $C_2 \sim$ Binomial $[q_x, \sum_{i=1}^{\infty} i\pi_i N_i]$ by assumptions and hence

(a) $$E[C_2] = \sum_{i=1}^{\infty} i\pi_i Nq_x$$

(b)

$$\mathrm{Var}[C_2] = \sum_{i=1}^{\infty} i\pi_i N q_x\ (1- q_x)$$

Hence

$$\frac{\mathrm{Var}[C_1]}{\mathrm{Var}[C_2]} = \frac{\sum_{i=1}^{\infty} i^2 \pi_i}{\sum_{i=1}^{\infty} i \pi_i}$$

iii) The variance assuming the Binomial model will be understated by the ratio in (ii)(b). Hence the Binomial variances will need to be adjusted by this ratio when calculating standard deviations in statistical tests.

G Total [8]

Solution 5

i)

Age	*Deviation*	*Standardised Deviation*	*Chi-squared Value*
45	3.65	0.81	0.655
46	-1.61	- 0.37	0.139
47	1.91	0.53	0.279
48	12.68	2.27	5.134
49	1.23	0.37	0.140
50	-3.24	-0.97	0.934
51	-8.38	-1.20	1.452
52	4.75	0.79	0.622
53	-15.98	-2.44	5.941
54	-16.31	-2.00	4.012
55	-17.27	-2.24	5.032
56	-12.30	-2.45	5.979
57	13.59	1.58	2.482
58	-15.65	-1.63	2.672
Total	52.93		35.473

a) Chi Squared test

We have 14 ages and have imposed no constraints and there are no parameters , hence DOF = 14 .

Observed $\chi^2_{14} = 35.473$

From Tables $\Pr(\chi^2_{14} > 31.32) = .005$

Therefore $\Pr(\chi^2_{14} > 35.473) < .005$

On basis of χ^2 test our hypothesis of good fit is disproved

Hence Reject H_0

b) Cumulative deviations

$$z = \sum \frac{(actual - \exp ected)}{\sqrt{\exp ected}}$$

$$Z = \frac{497 - 549.93}{\sqrt{549.93}} = -2.257$$

$p(-2.257<Z<-2.257) = 2 \times \phi(-2.257) = 2 \times (1 - \phi(2.257))$

Probability level = 0.012 X 2 = 0.024

Hence Reject H_0

c) Grouping of signs test

$n_1 = 6$; $n_2 = 8$ g=4

$E(g) = n_1 (n_2+1)/14 = 6 \text{ X } 9/14 = 3.857$

As $g > E(g)$, Accept H_0

ii) General conclusion is that standard table is unacceptable as it fails the Chi-squared overall goodness of fit test. Mortality experience is significantly lower than the standard table over the age range as a whole and particularly at higher ages. Not sufficient evidence to indicate that the shape of the A1967/70 table is inadequate.

G Total [8]

Solution 6

i)

- Investigate males and females separately.
- Possibly investigate temporary and whole life policies separately.
- Exclude duplicate policies on one life because the variance in mortality would otherwise be overestimated.
- Exclude ages where the data are scanty, i.e. very old and very young ages because these data will be unreliable.
- Exclude cases where special terms have been imposed because these are not representative of the main portfolio.
- Exclude data for earlier years because time selection will probably outweigh the advantages of a larger data set.

Note: except in the case of duplicate policies, all cases, which are suggested for exclusion, could alternatively be investigated separately.

ii) Consider an investigation over a period T years, where $T \leq 15$; assume that T is an integer, beginning on a 1 January.

Define $\theta_{x,r}$ = Deaths aged x nearest birthday and curtate duration r.

We require $E^c_{x,r} = \int_o^T {}_tP_{x,r} \; dt$

Where ${}_tP_{x,r}$ = number of lives at time t (t = 0 is the start of investigation aged x nearest birthday and having curtate duration r .

Consider $\int_o^T {}_tP_{x,r} \; dt = ½ \; ({}_0P_{x,r} + {}_1P_{x,r})$

assuming ${}_tP_{x,r}(k)$ varies linearly over the calendar year.

The data are $P_{x,r}(k)$ = number of lives at 1.1.1991 + k aged x nearest birthday and with curtate duration r.

If start of investigation is at 1.1.1991 + m, then

T -1

$$\int_0^1 {}_tP_{x,r}\ dt = \tfrac{1}{2} P_{x,r}(m) + \sum_{k=1} P_{x,r}(m+k) + \tfrac{1}{2} P_{x,r}(m+T)$$

Then our crude *m*-type select rate is

$$^{0}m_{[x+f-(r+h)]+r+h} = \theta_{x,r} \,/\, E^{c}_{x,r}$$

The rate interval for x = life year from x - ½ exact. Hence f = -½, with no assumption necessary.

The rate interval for r = policy year from duration r exact. Hence h = 0, with no assumption necessary.

Hence we have estimated $m_{[x-1/2-r]+r}$

G Total [10]

Solution 7

i) Given Ft, we know that
$N(t+s) - N(t) \sim \text{Poisson}(\lambda s)$

ii) Thus, we have
$E\,(\theta^{N(t+s)}\,|\,F_t) = \theta^{N(t)} e^{(\theta-1)\lambda s}$.

iii) In this case, from each corner, the next transition is equally likely to be at any one of the three adjacent corners. The transition probability matrix is given as

$$P = \begin{vmatrix} 0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 0 & 0 & 1/3 & 0 & 0 & 1/3 \\ 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \end{vmatrix}$$

iv) In this case, the probability is inversely proportional to the distance traveled. The assumption is that the distance is in a straight line from the start of the journey. The transition probability matrix is

P =

$$\begin{vmatrix} 0 & 1 & 1 & 1/\sqrt{2} & 1 & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1 & 0 & 1/\sqrt{2} & 1 & 1/\sqrt{2} & 1 & 1/\sqrt{3} & 1/\sqrt{2} \\ 1 & 1/\sqrt{2} & 0 & 1 & 1/\sqrt{2} & 1/\sqrt{3} & 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1 & 1 & 0 & 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{2} & 1 \\ 1 & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{3} & 0 & 1 & 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1 & 1/\sqrt{3} & 1/\sqrt{2} & 1 & 0 & 1/\sqrt{2} & 1 \\ 1/\sqrt{2} & 1/\sqrt{3} & 1 & 1/\sqrt{2} & 1 & 1/\sqrt{2} & 0 & 1 \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{2} & 1 & 1/\sqrt{2} & 1 & 1 & 0 \end{vmatrix} * 1/(3+3\sqrt{2}+\sqrt{3})$$

G Total [10]

Solution 8

i) There is an explicit dependence on the past behavior of $\{Y_j: j \le n\}$ in the probability distribution of Y_{n+1}; further, X_n is nothing but the sum of Y_j. Hence the Markov property does not hold.

ii) In part i), we show that there is explicit dependence on the past behavior of $\{Y_j: j \le n\}$ and hence the Markov property does not hold, this implies that that sequence $\{Y_n ; n \ge 1\}$ does not form a Markov chain.

iii) The transition matrix is given by:

$$\begin{vmatrix} p & 1-p & 0 & . & . & 0 \\ 0 & pe^{-\lambda} & 1-pe^{-\lambda} & 0 & . & . \\ . & 0 & pe^{-2\lambda} & 1-pe^{-2\lambda} & 0 & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & . & . & . & . & . \end{vmatrix}$$

iv)

(a) The chain is time homogeneous since the transition probabilities calculated in part i) is independent of time n.

(b) It is irreducible, since the number of errors can never go down.

(c) There are no recurrent states; hence there can be no stationary distribution.

Alternatively, if a stationary distribution π exists, it has to follow:

$$\pi_0 p = \pi_0$$
$$\pi_0 (1-p) + \pi_1 \, pe^{-\lambda} = \pi_1$$
$$\pi_1 (1-pe^{-\lambda}) + \pi_2 \, pe^{-2\lambda} = \pi_2$$

and so on.

Since $p \leq 1$, we have $\pi_0 = 0$ and then $\pi_1 = 0$, etc. Hence, no stationary distribution exists.

v) Probability of no further error is

$$(pe^{-j\lambda})^n = p^n e^{-nj\lambda}$$

G Total [10]

Solution 9

The transition probability matrix is

$$P = \begin{vmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{vmatrix}$$

i) The stationary distribution $\underline{\pi} = (\pi_1 \ \ \pi_2 \ \ \pi_3 \ \ \pi_4)$ is obtained from

$$(\pi_1 \ \ \pi_2 \ \ \pi_3 \ \ \pi_4) \begin{vmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{vmatrix}$$

$$= (\pi_1 \ \ \pi_2 \ \ \pi_3 \ \ \pi_4)$$

Thus, we have
$(\pi_1 + \pi_2)/2 = \pi_1$, $(\pi_1 + \pi_3)/2 = \pi_2$, $(\pi_2 + \pi_4)/2 = \pi_3$,
$(\pi_3 + \pi_4)/2 = \pi_4$

$$\Rightarrow \pi_1 = \pi_2 = \pi_3 = \pi_4 = 1/4$$

Thus, the long run average premium per year is given by
$(20000 + 17000 + 14000 + 12000)/4 = 15570$.

ii) Let e_0 be the expected time taken to reach 40% level starting at 0%, e_1 be the expected time taken to reach 40% level starting at 15% and e_2 be the expected time taken to reach 40% level starting at 30%. Then, we have

$$e_0 = 1 + e_0/2 + e_1/2$$

$$e_1 = 1 + e_0/2 + e_2/2$$

$e_2 = 1 + e_1/2$

$e_2 = 6$, $e_1 = 10$ & $e_0 = 12$.

Thus the expected time taken to reach 40% for the first time is 12 years.

iii) Let P_0 be the expected total premium till he reaches 40% level starting at 0%, P_1 be the expected total premium till he reaches 40% level starting at 15% and P_2 be the expected total premium till he reaches 40% level starting at 30%. Then, we have

$P_0 = 20000 + P_0/2 + P_1/2$

$P_1 = 17000 + P_0/2 + P_2/2$

$P_2 = 14000 + P_1/2$

$P_2 = 102000$ $P_1 = 176000$ $P_0 = 216000$.

Thus the expected value of the total premium paid till he reaches 40% for the first time is Rs. 216000/-.

G Total [12]

Solution 10

i) Let Sn be the crude price at time n and B_R be the event of first return to the original price after n steps. Thus,

$$P(B_R) = f_0(R).$$

Now,

$$p_0(n) = P(Sn = 0) = \sum_{k=1}^{n} P(S_n = 0|B_k)P(B_k) = \sum_{k=1}^{n} P(Sn = 0|B_k)f_0(k).$$

Note that S_n is a Markov chain with the state space being $\{\ldots,-2,-1, 0, 1, 2, \ldots\}$. This implies that $P(S_n = 0|B_k) = p_0(n-k)$. So,

$$p_0(n) = \sum_{k=1}^{n} p_0(n-k)f_0(k).$$

Also,

$$\sum_{k=1}^{\infty} s^n p_0(n) = P_0(s) - 1 = \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} p_0(n-k)f_0(k)$$

$$= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} s^{n-k}s^k p_0(n-k)f_0(k)$$

$$= \left(\sum_{k=1}^{\infty} s^k f_0(k)\right) \sum_{n=k}^{\infty} s^{n-k}p_0(n-k)$$

$$= \left(\sum_{k=1}^{\infty} s^k f_0(k)\right) \sum_{n=0}^{\infty} s^n p_0(n) = F_0(s)P_0(s)$$

$\Rightarrow P_0(s) = 1 + P_0(s)F_0(s)$.

ii) If $S_n = 0$, then the price must have moved up and down an equal number of times, i.e. there must have been an even number of price changes. So,

$$p_0(n) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \begin{Bmatrix} n \\ n/2 \end{Bmatrix} (pq)^{n/2} & \text{if } n \text{ is even} \end{cases}$$

Hence,

$$P_0(s) = \sum_{n=0}^{\infty} \begin{Bmatrix} 2n \\ n \end{Bmatrix} (pq)^n s^{2n} = \sum_{n=0}^{\infty} (2n!/n!)\ (pqs^2)^n.$$

Further, the expansion of $(1-4pqs^2)^{-1/2}$ leads to the same result as above and hence they are the same.

iii) From i) and ii),

$P_0(s)\ [1 - F_0(s)] = 1$, $F_0(s) = 1 - 1/P_0(s)$

$\Rightarrow F_0(s) = 1 - (1 - 4pqs^2)^{1/2}$

iv) The price of OPEC crude will ever return to zero (the initial value) is the sum of the probabilities of first return; i.e.

$$\sum_{n=1}^{\infty} f_0(n) = F_0(1) = 1 - (1 - 4pq)^{1/2} = P(\text{ever returns}).$$

Also note that $(1 - 4pq) = (p - q)^2$. This implies that $P(\text{ever returns}) = 1 - |p - q|$.

v) Now, when $p = 1/2$, using the result in part iv), we get
P(OPEC crude price will return to its original value) = 1
i.e. the price will return to its original value “almost surely”.

G Total [18]

****