EXAMINATION

April 2006

Subject CT4 — Models (includes both 103 and 104 parts)
Core Technical

EXAMINERS' REPORT

Introduction

The attached subject report has been written by the Principal Examiner with the aim of helping candidates. The questions and comments are based around Core Reading as the interpretation of the syllabus to which the examiners are working. They have however given credit for any alternative approach or interpretation which they consider to be reasonable.

M Flaherty
Chairman of the Board of Examiners

June 2006

© Faculty of Actuaries
© Institute of Actuaries

Comments

Comments on solutions presented to individual questions for this April 2006 paper are given below.

103 Part

Question A1	*This was well answered overall.* *Most candidates scored better on part (b); marks were lost on part (a) because answers were imprecise.*
Question A2	*This was reasonably well answered overall.* *Marks were lost because candidates did not show sufficient steps.*
Question A3	*This was reasonably well answered overall* *In part (ii), many candidates included more states than required. (See end of solution for further comments.)*
Question A4	*This was poorly answered.* *Very few candidates scored highly on this question. Most failed to provide sufficient, distinct points.*
Question A5	*This was very well answered.* *Marks were lost on part (ii) when candidates failed to consider all the conditions applying, and part (v) where many candidates calculated P^3.*
Question A6	*This was poorly answered, although the better candidates did manage to score highly.*

104 Part

Question B1	*This was well answered overall.* *The most common mistake was to use only one variable for self-esteem.*
Question B2	*This was reasonably well answered overall.* *In part (i), many candidates discussed premium setting and anti-selection, which was not relevant to the question asked.*
Question B3	*This was very poorly answered, with very few candidates scoring highly.* *Some alternative approaches to part (i) received credit, although care was needed over the ranges fro which μ_x was constant. Most candidates attempted part (ii), although few used the solution to part (i).*
Question B4	*This was very poorly answered.* *Most solutions offered lacked a coherent explanation.*
Question B5	*This was very well answered.* *Marks were most frequently lost in part (i), because of insufficient explanation of the types of censoring present.*
Question *B6*	*This was reasonably well answered overall.* *In part (ii), many candidates carried out a signs test. The use of the Normal approximation to the Binomial was not acceptable in this case, and candidates who used this lost marks. (See end of solution (ii)(c) for further comments.)*

103 Solutions

A1 (a) For a process to be strictly stationary, the joint distribution of $X_{t_1}, X_{t_2}, ..., X_{t_n}$ and $X_{t+t_1}, X_{t+t_2}, ..., X_{t+t_n}$ are identical for all $t, t_1, t_2, ..., t_n$ in J and all integers n.

This means that the statistical properties of the process remain unchanged over time.

(b) Because strict stationarity is difficult to test fully in real life, we also use the less stringent condition of weak stationarity.

Weak stationarity requires that the mean of the process, $E[X_t] = m(t)$, is constant and the covariance, $E[(X_s - m(s))\,(X_t - m(t))]$, depends only on the time difference $t - s$.

A2 Condition on the state occupied at time t to consider the survival probability ${}_{t+dt}p_0^{AB}$ (this requires the Markov property):

$${}_{t+dt}p_0^{AB} = {}_t p_0^{AA} \cdot {}_{dt}p_t^{AB} + {}_t p_0^{AB} \cdot {}_{dt}p_t^{BB} + {}_t p_0^{AC} \cdot {}_{dt}p_t^{CB} + {}_t p_0^{AD} \cdot {}_{dt}p_t^{DB}$$

Observe that ${}_{dt}p_t^{CB} = {}_{dt}p_t^{DB} = 0$

From the law of total probability:

$${}_{dt}p_t^{BB} = 1 - {}_{dt}p_t^{BA} - {}_{dt}p_t^{BC} - {}_{dt}p_t^{BD}$$

Substituting for ${}_{dt}p_t^{BB}$

$${}_{t+dt}p_0^{AB} = {}_t p_0^{AA} \cdot {}_{dt}p_t^{AB} + {}_t p_0^{AB}.(1 - {}_{dt}p_t^{BA} - {}_{dt}p_t^{BC} - {}_{dt}p_t^{BD})$$

For small dt:

$$\begin{aligned}
{}_{dt}p_t^{BA} &= \mu_t^{BA}.dt + o(dt) \\
{}_{dt}p_t^{BC} &= \mu_t^{BC}.dt + o(dt) \\
{}_{dt}p_t^{BD} &= \mu_t^{BD}.dt + o(dt) \\
{}_{dt}p_t^{AB} &= \mu_t^{AB}.dt + o(dt)
\end{aligned}$$

Where $o(dt)$ covers the possibility of more than one transition in time dt and

$$\lim_{dt \to 0} \frac{o(dt)}{dt} = 0$$

Substituting in:

$${}_{t+dt}p_0^{AB} = {}_tp_0^{AA}.\mu_t^{AB}.dt + {}_tp_0^{AB}(1 - \mu_t^{BA}.dt - \mu_t^{BC}.dt - \mu_t^{BD}.dt) + o(dt)$$

$$\frac{\partial}{\partial t}{}_tp_0^{AB} = \lim_{dt \to 0^+} \frac{{}_{t+dt}p_0^{AB} - {}_tp_0^{AB}}{dt} = {}_tp_0^{AA}.\mu_t^{AB} - {}_tp_0^{AB}(\mu_t^{BA} + \mu_t^{BC} + \mu_t^{BD})$$

A3 (i) This is not a Markov chain because it does not possess the Markov property, that is transition probabilities do not depend only on the current state.

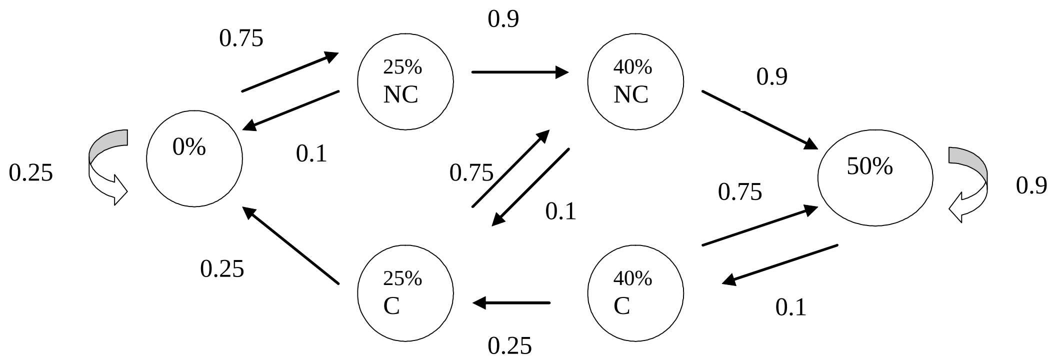
Specifically, if you are in the 25% discount level, the transition probability to state 0% is 0.25 if a claim was made last year and 0.1 if the previous year was claim free.

(ii) (a) Split the 25% and 40% discount states to include whether the previous year was claim free.

New state space:

0% discount
25%NC (no claim last year)
25%C (at least one claim last year)
40%NC (no claim last year)
40%C (at least one claim last year)
50%

(b)



	New state						
Old State		0%	25%C	25%NC	40%C	40%NC	50%
	0%	0.25	0	0.75	0	0	0
	25%C	0.25	0	0	0	0.75	0
	25%NC	0.1	0	0	0	0.9	0
	40%C	0	0.25	0	0	0	0.75
	40%NC	0	0.1	0	0	0	0.9
	50%	0	0	0	0.1	0	0.9

(iii) In theory, the insurer should just use 2 NCD states according to whether the policyholder made a claim in the previous year. This is because the company believes the claims frequency is the same for drivers who have not made a claim for 1, 2, 3… years (i.e. it remains at 0.1 whether the driver has been claims-free for 1 or 10 years).

However there may be other reasons for adopting this scale:

- Marketing or competitive pressures.

- It may discourage the policyholder from making small claims, or encourage careful driving, to preserve their discount.

General comments:

The following, more general comments about the appropriateness of an NCD model also received credit:

- *It is appropriate to award a no-claims discount because there is empirical evidence that drivers who have made a recent claim are more likely to make a further claim.*

- *More factors should be taken into account (with a suitable example such as how long the policyholder has been driving).*

A4 (i) Systems with long time frames such as the operation of a pension fund can be studied in compressed time.

Different future policies or possible actions can be compared to see which best suits the requirements or constraints of a user.

Complex situations can be studied.

Modelling may be the only practicable approach for certain actuarial problems.

(ii) A model is described as stochastic if it allows for the random variation in at least one input variable.

Often the output from a stochastic model is in the form of many simulated possible outcomes of a process, so distributions can be studied.

A deterministic model can be thought of as a special case of a stochastic model where only a single outcome from the underlying random processes is considered.

Sometimes stochastic models have analytical/closed form solutions, such that simulation is not required, but they are still stochastic as they allow for factors to be random variables.

(iii)

- If the distribution of possible outcomes is required then stochastic modelling would be needed, or if only interested in a single scenario then deterministic.
- Budget and time available — stochastic modelling can be considerably more expensive and time consuming.
- Nature of existing models.
- Audience for the results and the way they will be communicated.

The following factors may favour a stochastic approach:

- The regulator may require a stochastic approach.
- Extent of non-linear variation — for example existence of options or guarantees.
- Skewness of distribution of underlying variables, such as cost of storm claims.
- Interaction between variables, such as lapse rates with investment performance.

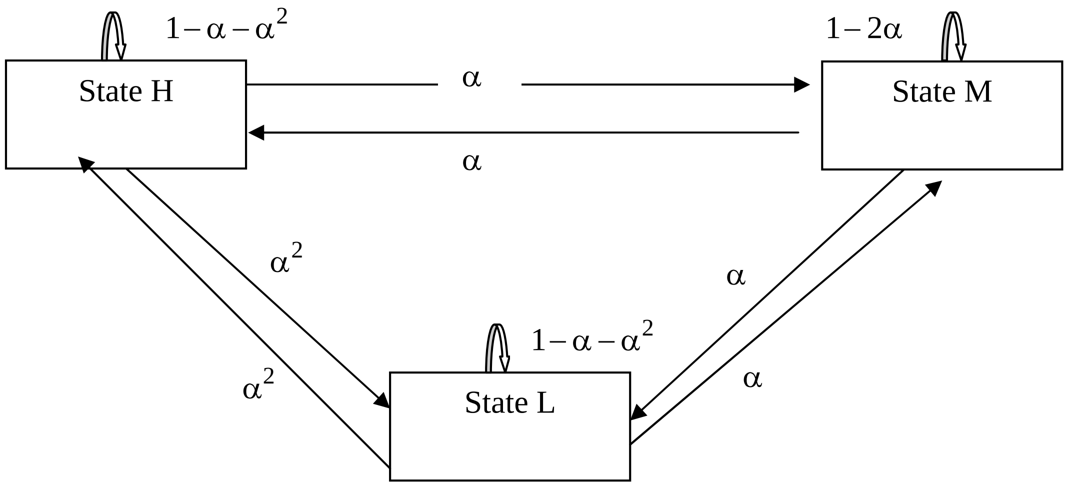
The following may favour a deterministic approach:

- Lack of credible historic data on which to fit distribution of a variable.
- If accuracy of result is not paramount, for example if a simple model with deliberately cautious assumptions is chosen so as not to underestimate costs.

(iv) A deterministic result on best estimate assumptions could be compared with the mean and median outcomes from a stochastic approach.

A deterministic model may also be used to calculate the expected or median outcome, with a stochastic approach being used to estimate the volatility around the central outcome.

A5 (i) Transition graph given below.



(ii) Transition probabilities must lie in [0,1]. Thus we need $\alpha \geq 0$, $1 - 2\alpha \geq 0$ and $1-\alpha-\alpha^2 \geq 0$.

The solution of the quadratic is the interval $\left[-\frac{1}{2}-\frac{\sqrt{5}}{2}, -\frac{1}{2}+\frac{\sqrt{5}}{2}\right]$, so all conditions are satisfied simultaneously for $\alpha \in [0,\frac{1}{2}]$.

(iii) The chain is both irreducible, as every state can be reached from every other state, and aperiodic, as the chain may remain at its current state for all H, M, L.

(iv) From the result in (iii), a stationary probability distribution exists and it is unique. Let $\pi = (\pi_H, \pi_M, \pi_L)$ denote the stationary distribution. Then, π can be determined by solving $\pi P = \pi$.

For $\alpha = 0.2$, the transition matrix becomes

$$P = \begin{pmatrix} 0.76 & 0.2 & 0.04 \\ 0.2 & 0.6 & 0.2 \\ 0.04 & 0.2 & 0.76 \end{pmatrix}$$

So that the system $\pi P = \pi$ reads

$$0.76\,\pi_H + 0.2\,\pi_M + 0.04\,\pi_L = \pi_H \qquad (1)$$
$$0.2\,\pi_H + 0.6\,\pi_M + 0.2\,\pi_L = \pi_M$$
$$0.04\,\pi_H + 0.2\,\pi_M + 0.76\,\pi_L = \pi_L \qquad (2)$$

Discard the second of these equations and use also that the stationary probabilities must also satisfy

$$\pi_H + \pi_M + \pi_L = 1 \qquad (3)$$

Subtracting (2) from (1) gives $\pi_H = \pi_L$.

Substituting into (1) we obtain $\pi_H = \pi_M$, thus (3) gives that $\pi_H = \pi_M = \pi_L = 1/3$. The proportion of employees who are in state L in the long run is 1/3.

(v) The second order transition matrix is

$$P^2 = \begin{pmatrix} 0.76 & 0.2 & 0.04 \\ 0.2 & 0.6 & 0.2 \\ 0.04 & 0.2 & 0.76 \end{pmatrix} \cdot \begin{pmatrix} 0.76 & 0.2 & 0.04 \\ 0.2 & 0.6 & 0.2 \\ 0.04 & 0.2 & 0.76 \end{pmatrix}$$

$$= \begin{pmatrix} 0.6192 & 0.28 & 0.1008 \\ 0.28 & 0.44 & 0.28 \\ 0.1008 & 0.28 & 0.6192 \end{pmatrix}$$

The relevant entries are those in the last column, so that the answers are:

(a) 0.1008
(b) 0.28
(c) 0.6192.

A6 (i) (a) A continuous-time Markov process $X_t, t \geq 0$ with a discrete state space S is called a Markov jump process.

(b) In the case where the probabilities $P\left(X_t = j \mid X_s = i\right)$ for *i*, *j* in S and $0 \leq s < t$ depend only on the length of time interval $t - s$, the process is called time-homogeneous.

(ii) A model with time-inhomogeneous rates has more parameters, and there may not be sufficient data available to estimate these parameters.

Also, the solution to Kolmogorov's equations may not be easy (or even possible) to find analytically.

(iii) $P'(t) = P(t).A(t)$

where

$$A(t) = \begin{pmatrix} -\sigma(t) - \mu(t) & \sigma(t) & \mu(t) \\ \rho(t) & -\rho(t) - \nu(t) & \nu(t) \\ 0 & 0 & 0 \end{pmatrix}$$

(iv) (a) $\Pr(\text{Waiting time} > T - w \mid X_w = S) = \exp\left[-\int_w^T (\rho(t) + \nu(t))dt \right]$

(b) Given there is a transition from state H at time *w*, the probabilities that this is into state S or D are given by the relative transition rates at time *w*.

So Probability into state S = $\dfrac{\sigma(w)}{\mu(w) + \sigma(w)}$

(c) This is the probability that the individual is in state H at time *w*, multiplied by the sum of transition rates out of state H at time *w*, that is:

$$P_{HH}(\tau, w).(\mu(w) + \sigma(w)) \cdot dw$$

(v) Expressing time in years,

$$\Pr(X_T = S, \text{Waiting time} < 1/2 \,|\, X_\tau = H)$$

$$= \int_{T-1/2}^{T} \Pr(\text{Transition from state H at } w) \times \Pr(\text{Transition to S}) \times \Pr(\text{stays in S to time T})\, dW$$

$$= \int_{T-1/2}^{T} P_{HH}(\tau, w).(\mu(w)+\sigma(w)).\frac{\sigma(w)}{\mu(w)+\sigma(w)}.\exp\left[-\int_{w}^{T}(\rho(t)+\nu(t))dt\right].dw$$

$$= \int_{T-1/2}^{T} P_{HH}(\tau, w).\sigma(w).\exp\left[-\int_{w}^{T}(\rho(t)+\nu(t))dt\right].dw$$

(vi) (a) This is likely to improve the predictive power of the model because:

- There is empirical evidence that recovery rates depend on the duration of the sickness.
- The limit of 6 months on sick pay may cause some durational effects around this point.

However this would make the model more complicated to analyse, and increase the volume of data required to fit parameters reliably.

(b) For individuals in employment mortality rates are likely to be low, and may be ignorable. It is less likely that mortality out of state S could be excluded.

104 Solutions

B1 $h(t) = h_0(t)\exp[\beta_1 F + \beta_2 M + \beta_3 H]$

where

$h(t)$ is the estimated hazard,

$h_0(t)$ is the baseline hazard,

F is a variable taking the value 1 if the life is female, and 0 otherwise,

M is a variable taking the value 1 if the life has "medium" self-esteem and 0 otherwise,

H is a variable taking the value 1 if the life has "high" self-esteem and 0 otherwise, and

β_1, β_2 and β_3 are parameters to be estimated.

B2 (i) (a) The models of mortality we use assume that we can observe a group of lives with the same mortality characteristics. This is not possible in practice.

However, data can be sub-divided according to certain characteristics that we know to have a significant effect on mortality.

This will reduce the heterogeneity of each group, so that we can at least observe groups with similar, but not the same, characteristics.

(b) Sub-dividing data using many factors can result in the numbers in each class being too low.

It is necessary to strike a balance between homogeneity of the group and retaining a large enough group to make statistical analysis possible.

Sufficient data may not be collected to allow sub-division.

This may be because marketing pressures mean proposal forms are kept to a minimum.

(ii) The following are factors often used:

Sex
Age
Type of policy
Smoker/Non-smoker status
Level of underwriting
Duration in force
Sales channel
Policy size
Occupation (or social class) of policyholder
Known impairments
Geographical region

B3 (i) Consider the year of age between y and $y + 1$. We know that

$$_t p_y = \exp\left[-\int_0^t \mu_{y+s} ds \right].$$

If t=1 and $\mu_{y+s} = \mu_y$ (a constant), evaluating the integral produces

$$p_y = \exp\left[-\mu_y \right].$$

Now, conditioning on survival to age x, survival to age y + 1 implies survival from age x to age y and then survival for a further year:

$$_{y+1-x} p_x = p_y \cdot {}_{y-x} p_x .$$

Thus

$$p_y = \frac{_{y+1-x} p_x}{_{y-x} p_x},$$

which, since, in general $_t p_x = S_x(t)$, may be written

$$p_y = \frac{S_x(y+1-x)}{S_x(y-x)}.$$

Therefore

$$\exp(-\mu_y) = \frac{S_x(y+1-x)}{S_x(y-x)},$$

so that

$$\mu_y = \log\left[\frac{S_x(y-x)}{S_x(y+1-x)}\right] = \log[S_x(y-x)] - \log[S_x(y+1-x)].$$

(ii) (a) Using the result from part (i) and putting $x = 50$, $y = 50$ gives

$$\mu_{50} = \log\left[\frac{S_{50}(0)}{S_{50}(1)}\right] = -\log[S_{50}(1)]$$

Since we have censored data, because of the possibility of policy lapse, we should estimate $S_{50}(1)$ using the Kaplan-Meier or Nelson-Aalen estimator and hence obtain an estimate of μ_{50}.

(b) $_5q_{50} = 1 - {_5p_{50}}$,

and, since

$_5p_{50} = S_{50}(5)$,

$_5q_{50}$ can be estimated directly as $1 - S_{50}(5)$,

where $S_{50}(5)$ is the Kaplan-Meier or Nelson-Aalen estimator of the probability of a life aged 50 years surviving for a further 5 years.

B4 (i) We have a policy-year rate interval.

(ii) The age classification of the lapsing data is "age last birthday on the policy anniversary prior to lapsing".

This can be calculated by adding the policyholder's age last birthday when the policy was taken to out to the number of annual premiums paid minus 1 (assuming that the first premium was paid at policy inception).

Define $P_{x,t}$ as the "number of policies in force aged x last birthday at the preceding policy anniversary" at time t. This corresponds with the lapsing data.

Then, if t is measured in years since 1 January 2003, a consistent exposed-to-risk would be

$$E_x^c = \int_0^1 P_{x,t}\,dt\,,$$

which, assuming that policy anniversaries are uniformly distributed across the calendar year,

may be approximated as

$$E_x^c = \frac{1}{2}[P_{x,0} + P_{x,1}]\,.$$

But we do not observe $P_{x,t}$ directly. Instead we observe $P_{x,t}{}^*$ the number of policies in force at time t, classified by age last birthday at time t.

But the range of exact ages that could apply to a life aged x last birthday on the policy anniversary prior to lapsing is $(x, x + 2)$.

Assuming that birthdays are uniformly distributed across the policy year, half of these lives will be aged x last birthday and half will be aged x+ 1 last birthday.

Hence,

$$P_{x,t} = \frac{1}{2}[P_{x,t}{}^* + P_{x+1,t}{}^*]\,.$$

Therefore, by substituting this into the approximation above, the appropriate exposed-to-risk is

$$E_x^c = \frac{1}{2}\left[\frac{1}{2}[P_{x,0}{}^* + P_{x+1,0}{}^*] + \frac{1}{2}[P_{x,1}{}^* + P_{x+1,1}{}^*]\right].$$

(iii) Both assumptions might be unreasonable because:

policies might be taken out in large numbers just before the end of the tax year,

policies might tend to be taken out just before birthdays,

under group schemes, many policy anniversaries might be identical.

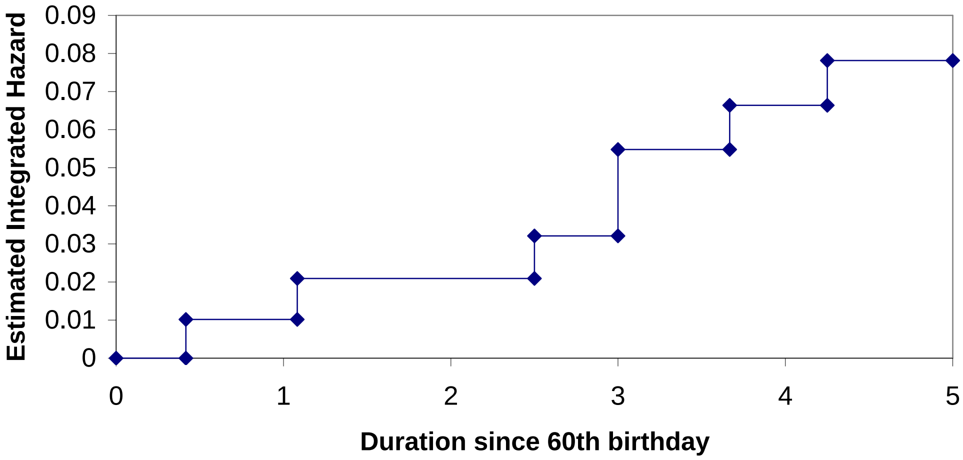
B5 (i) The following types of censoring will be present:

- Right censoring because some policyholders cancel their policy before the end of the period.
- Type I censoring because the investigation stops at a fixed time.
- Random censoring because some lives cancel their policy at an unknown time.
- Informative censoring because those who cancel their policy tend to be in better health.

(ii) (a) The calculations are as follows:

t_j *(years)*	n_j	d_j	c_j	$\frac{d_j}{n_j}$	$\Lambda_j = \sum \frac{d_j}{n_j}$
$0 \le t < \frac{5}{12}$	100	0	2	0	0
$\frac{5}{12} \le t < 1\frac{1}{12}$	98	1	4	1/98	0.0102
$1\frac{1}{12} \le t < 2\frac{6}{12}$	93	1	2	1/93	0.0210
$2\frac{6}{12} \le t < 3$	90	1	1	1/90	0.0321
$3 \le t < 3\frac{8}{12}$	88	2	0	2/88	0.0548
$3\frac{8}{12} \le t < 4\frac{3}{12}$	86	1	1	1/86	0.0664
$4\frac{3}{12} \le t$	84	1	1	1/84	0.0783

(b)



(iii) *Either*

Using the results of the calculation in (ii), the survival function can be estimated by $S(t) \approx \exp(-\Lambda_t)$.

And so, for $t \geq 4\ ^3/_{12}$, we have

$$S(t) = \exp(-0.0783) = 0.925$$

which is the probability of survival to 65.

Or

Using the Kaplan-Meier estimate of $S(t) = \prod_{t_j < t} \left(1 - \frac{d_j}{n_j}\right)$,

we get, for $t \geq 4\ ^3/_{12}$:

$$S(t) = \left(1-\frac{1}{98}\right)\cdot\left(1-\frac{1}{93}\right)\cdot\left(1-\frac{1}{90}\right)\cdot\left(1-\frac{2}{88}\right)\cdot\left(1-\frac{1}{86}\right)\cdot\left(1-\frac{1}{84}\right)$$

$$= 0.9243$$

B6 (i) The null hypothesis is that the crude rates come from a population in which true underlying rates are the graduated rates.

The test statistic is $X = \sum_x z_x^2$

Under the null hypothesis X has a χ^2 distribution with m degrees of freedom, where m is the number of age groups less one for each parameter fitted. So in this case m = 15 – 3 = 12, ie $X \sim \chi^2_{12}$

The observed value of X is 12.816.

The critical value of the χ^2_{12} distribution at the 5% level is 21.03

This is greater than the observed value of X

and so we have insufficient evidence to reject the null hypothesis.

(ii) (a) The obvious problem with the graduation is one of overall bias. The graduated rates are consistently too high, resulting in too many negative deviations.

(b) This is not detected by the χ^2 test because the test statistic is the sum of the squared deviations and so information on the sign and some information on the size of the individual deviations is lost. The χ^2 test would detect large bias, but in this case the graduated and crude rates are close enough that the statistic is below the critical value.

(c) *Signs test*

Let P be the number of positive deviations.

Under the null hypothesis, $P \sim \text{Binomial}(15, 0.5)$.

We have 3 positive deviations. The probability of getting 3 or fewer positive signs (if the null hypothesis is true) is:

$$\left(\frac{1}{2}\right)^{15} \times \left(\binom{15}{0} + \binom{15}{1} + \binom{15}{2} + \binom{15}{3}\right)$$

$$= \left(\frac{1}{2}\right)^{15} \times (1 + 15 + 105 + 455)$$

$$= 0.0176$$

This is less than 0.025 (this is a two-tailed test)

and so we reject the null hypothesis.

Cumulative deviations test

Our test statistic is $\dfrac{\sum_x \left(E_x \hat{q}_x - E_x \overset{\circ}{q}_x\right)}{\sqrt{\sum_x E_x \overset{\circ}{q}_x \left(1 - \overset{\circ}{q}_x\right)}}$

Under the null hypothesis, this has Normal(0, 1) distribution.

Using the data in the question, we have

Age x	$E_x\left(\hat{q}_x - \overset{\circ}{q}_x\right)$	$E_x\,\overset{\circ}{q}_x\left(1-\overset{\circ}{q}_x\right)$
40	-6.40146	38.2751
41	-3.0025	42.84188
42	-7.92472	42.7289
43	-7.62982	36.46509
44	-6.08904	32.93758
45	2.63525	42.20447
46	4.2237	30.62388
47	-3.49218	51.31917
48	-4.9133	61.63457
49	-9.1832	51.99181
50	-7.488	58.25669
51	-8.24226	51.00139
52	1.10244	45.70533
53	-8.55647	72.14466
54	-7.87508	61.5123
Total	**-72.837**	**719.643**

$$\Rightarrow \frac{\sum_x\left(E_x q_x - E_x\,\overset{\circ}{q}_x\right)}{\sqrt{\sum_x E_x\,\overset{\circ}{q}_x\left(1-\overset{\circ}{q}_x\right)}} = \frac{-72.837}{\sqrt{719.643}} = -2.715$$

This is a two-tailed test.

Since $|-2.715| > 1.96$, we reject the null hypothesis.

Comments:

Candidates also received credit for using the standardised deviations test to show that there were too many deviations in the (-2, -1) range.

(iii) The problem is that the graduated rates are too high. There doesn't appear to be a problem with the overall shape.

So we should be able to adjust the parameters rather than change the underlying equation.

The problem persists across the whole age range, so the first adjustment to try would be to decrease the value of α.

END OF EXAMINERS' REPORT