Subject CT4 — Models
Core Technical

EXAMINERS' REPORT

April 2008

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 A Stocker
Chairman of the Board of Examiners

June 2008

Comments

Comments on solutions presented to individual questions for this April 2008 paper are given below.

Question 1 — *This straightforward bookwork question was very well answered.*

Question 2 — *Answers to this question were disappointing. In part (a) many candidates did not realise that smoothness is automatically ensured when graduating with a parametric formula with a small number of parameters. In part (b) many candidates presented descriptions of the method of graphical graduation, rather than answering the question which was set.*

Question 3 — *Most candidates scored reasonably well on part (i), but few candidates could state the conditions required for a compound Poisson process to be a Poisson process in part (ii).*

Question 4 — *A reasonable attempt was made at this bookwork question by most candidates, although few made sufficient distinct points to score close to full marks.*

Question 5 — *This exposed-to-risk question was quite well answered by many candidates, who correctly identified the rate interval and the appropriate census-type formula. An encouraging number of candidates also recognised the need to adjust the age definition in order to ensure correspondence between the first marriages data and the exposed-to-risk data.*

Question 6 — *Many candidates scored well on this question. Common errors were failure to use (or incorrect use of) the continuity correction in the normal approximation to the signs test; calculating only the probability of 18 positive signs (rather than the probability of 18 <u>or more</u> signs) when using the exact binomial computation of the signs test; and calculating only the probability of 2 positive runs (rather than the probability of 2 <u>or fewer</u> positive runs) when using the exact computation of the grouping of signs test.*

Question 7 — *Only a small proportion of candidates correctly answered part (i). In part (ii) a very large number of candidates adopted a three-state solution to this problem, with state space {A, B, C}. Partial credit was given for this, and also for correctly following this three-state solution through in part (iii) to obtain the steady-state proportions of 3/11, 2/11 and 6/11 using auditors A, B and C respectively.*

Question 8 — *This question was not as well answered as some others. Some candidates failed to write the numerical values of the estimated parameters down in part (ii). There were few correct attempts at part (v). Many candidates simply calculated the ratio between the two hazards, which is incorrect. Others made unnecessary assumptions about the form of the baseline hazard (e.g. that it was constant).*

Question 9 *This straightforward calculation of the survival function was very well answered, apart from part (iv), in which only a handful of candidates realised that the Kaplan-Meier estimate of the hazard at any duration at which no event is observed to take place is 0. Given that the Kaplan-Meier estimate of the hazard is a step function, it is clear than this must be so. It was very encouraging to see the high proportion of sensible answers to part (ii). Credit was given in part (ii) to candidates who stated that the censoring was non-informative provided that the reason given was consistent with this statement.*

Question 10 *Few candidates scored highly on this question. Many candidates got no further than part (ii). Although there were a fair number of attempts to solve the differential equation in part (iii), only a minority of candidates spotted that* $P_{ON}(t) + P_{OFF}(t) = 1$.

Question 11 *This question was very well answered. Many candidates provided substantially correct answers to all parts, losing marks only for failure to include certain details in part (ii) (for example that we need to condition on the state occupied at time x+t); or for failing to point out that we need to substitute the estimated values from the data into the formula for the variance of* μ^{23} *in part (v).*

1 Sex
Age
Type of policy
Smoker/non-smoker
Level of underwriting
Duration in force
Sales channel
Policy size
Known impairments
Occupation

2 (a) Provided a formula with a small number of parameters is chosen
the resulting graduation will be acceptably smooth.

(b) The graduation should be tested for smoothness

using the third differences of the graduated rates

which should be small in magnitude and progress regularly.

A further iterative process, which involves manual adjustment of the graduation (called 'hand-polishing') is sometimes necessary to ensure smoothness.

3 (i) (a) EITHER

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

$N_0 = 0$
N_t has independent increments
N_t has stationary increments

$$P[N_t - N_s = n] = \frac{[\lambda(t-s)]^n e^{-\lambda(t-s)}}{n!} \quad s < t, n = 0, 1, 2\ldots..$$

OR

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

$N_0 = 0$
$P[N_{t+h} - N_t = 1] = \lambda h + o(h)$
$P[N_{t+h} - N_t = 0] = 1 - \lambda h + o(h)$
$P[N_{t+h} - N_t \neq 0,1] = o(h)$

(b) If N_t is a Poisson process on $t \geq 0$ and Y_i is a sequence of independent and identically distributed random variables then a compound Poisson process is defined by:

$$X_t = \sum_{i=1}^{N_t} Y_i$$

(ii) A compound Poisson process meets the conditions for being a Poisson process if Y_i is an indicator function OR if each Y_i is identically 1 (which is a special case of the indicator function)

4 *Benefits*

Systems with long time frames can be studied in compressed time, for example the operation of a pension fund (or other suitable example).

Complex systems with stochastic elements can be studied

Different future policies or possible actions can be compared.

In a model of a complex system we can usually get much better control over the experimental conditions so that we can reduce the variance of the results output from the model without upsetting their mean values

Avoids costs and risks of making changes in the real world, so we can study impact of changing inputs before making decisions.

Limitations

Model development requires a considerable investment of time and expertise.
In a stochastic model, for any given set of inputs each run gives only estimates of a model's outputs. So to study the outputs for any given set of inputs, several independent runs of the model are needed.

Models can look impressive when run on a computer so that there is a danger that one gets lulled into a false sense of confidence.

If a model has not passed the tests of validity and verification its impressive output is a poor substitute for its ability to imitate its corresponding real world system.

Models rely heavily on the data input. If the data quality is poor or lacks credibility then the output from the model is likely to be flawed.
It is important that the users of the model understand the model and the uses to which it can be safely put. There is a danger of using a model as a black box from which it is assumed that all results are valid without considering the appropriateness of using that model for the particular data input and the output expected.

It is not possible to include all future events in a model. For example a change in legislation could invalidate the results of a model, but may be impossible to predict when the model is constructed.

It may be difficult to interpret some of the outputs of the model. They may only be valid in relative, rather than absolute, terms. For example comparing the level of risk of the outputs associated with different inputs.

5 (i) Calendar year rate interval starting on 1 January each year.

(ii) The first marriages data may be described as

m_x = number of first marriages, age x on the birthday in the calendar year of marriage, during a defined period of investigation of length N years

A definition of the population data which is compatible with these data on first marriages is

$P_{x,t}$ = number of lives under observation at time t since the start of the investigation who were aged x next birthday on the 1 January immediately preceding t

Since we follow each cohort of lives through each calendar year, this exposed to risk is

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

which may be approximated as

$$E_x^c = \sum_0^{N-1} \frac{1}{2}(P_{x,t} + P_{x+1,t+1})$$

(where the summation considers just integer values of t).

This assumes that the population varies linearly across the calendar year.

However, we have data classified by age last birthday so we need to make a further adjustment.

If the number of lives aged x last birthday on 1 January in year t is $P_{x,t}$* then

$P_{x,t} = P_{x-1,t}$*

and an appropriate exposed to risk in terms of the data we have is

$$E_x^c = \sum_{t=K}^{K+N} \frac{1}{2}(P_{x-1,t}{}^{*} + P_{x,t+1}{}^{*}) \quad .$$

(iii) The age range at the start of the rate interval is (x–1, x) exact.

So, assuming that birthdays are uniformly distributed across the calendar year the average age at the start of the rate interval is x–½ and the average age in the middle of the rate interval is x.

Therefore the estimate of λ_x applies to age x.

6 (i) Since we do not know the values of the rates in the crude experience but only the signs of the deviations the tests we can carry out are limited.

We can, however, perform the signs test and the grouping of signs test.

(ii) The **signs test** looks for overall bias.
We have 25 ages, and at 18 of these the crude rates exceed the standard table rates (i.e. we have positive deviations)

If the null hypothesis is true, then the observed number of positive deviations, P, will be such that P ~ Binomial (25, ½).

EITHER

We use the normal approximation to the Binomial distribution because we have a large number of ages(>20)
This means that, approximately, P ~ Normal (12.5, 6.25).

The z-score associated with the probability of getting 18 positive deviations if the null hypothesis is true is, therefore

$$\frac{17.5-12.5}{\sqrt{6.25}}=\frac{-5}{2.5}=-2.00\ .$$

(using a continuity correction).

We use a two-tailed test, since both an excess of positive and an excess of negative deviations are of interest.

Using a 5 % significance level, we have $-2.00 < -1.96$.

This means we have just sufficient evidence to reject the null hypothesis.

OR

Using the Binomial exactly we have

Pr[*j* positive deviations] = $\binom{25}{j}0.5^{25}$.

So that the probability of obtaining 18 or more positive deviations is $\sum_{j=18}^{25}\binom{25}{j}0.5^{25}$.

This is equal to

(1 + 25 + 300 + 2,300 + 12,650 + 53,130 + 177,100 + 480,700) × 0.0000000298

= 0.02164.

We apply a 2-tailed test, so we reject the null hypothesis at the 5% level if this is less than 0.025

Since 0.02164 < 0.025

we reject the null hypothesis.

The **grouping of signs** test looks for long runs or clumps of ages with the same sign, indicating that the crude experience is different from the standard experience over a substantial age range.

The number of runs of positive signs is 2 (65–72 years and 75–84 years).

We have 25 ages and 18 positive signs in total, which means 7 negative signs.

THEN EITHER

Using the table provided under $n_1 = 18$ and $n_2 = 7$, we find that, under the null hypothesis, the greatest number of positive runs *x* for which the probability of *x* or fewer positive runs is less than 0.05 is 3.

Since we only have 2 runs, we conclude that the probability of obtaining 2 or fewer runs is much less than 0.05.

Therefore we reject the null hypothesis.

OR

Using exact computation

$$\Pr[1 \text{ positive run}] = \frac{\binom{17}{0}\binom{8}{1}}{\binom{25}{18}} = \frac{8}{480,700} = 0.0000166$$

$$\Pr[2 \text{ positive runs}] = \frac{\binom{17}{1}\binom{8}{2}}{\binom{25}{18}} = \frac{(17)(28)}{480,700} = 0.000990$$

Therefore we conclude that the probability
of obtaining 2 or fewer runs is much less than 0.05.

Therefore we reject the null hypothesis.

OR

Using the Normal approximation, the number of positive runs is distributed

$$N\left(\frac{(18)(8)}{25}, \frac{[(18)(7)]^2}{(25)^3}\right) = N\left(5.76, 1.02\right)$$

so that the *z*-score associated with the probability of getting 2 runs
is

$$\frac{2-5.76}{\sqrt{1.02}} = -3.722.$$

which is much less than -1.645 (using a 1-tailed test).

Therefore we conclude that the probability
of obtaining 2 or fewer runs is much less than 0.05.

Therefore we reject the null hypothesis.

7 (i) Required number

$$= \sum_{i=1}^{\infty} \text{ probability } i\text{th audit takes place prior to changing auditors}$$

$= 1 + 1 + 0.8 + 0.8^2 + 0.8^3 + \ldots\ldots..$

$= 1 + 1/(1 - 0.8) = 6$

(ii) The transition probabilities depend on whether it is the first year with the current auditors, so need additional states to cover this.

State space = $\{A_L, A, B_L, B, C_L, C\}$ where subscript L indicates locked in to the current auditor.

Transition matrix **A** is

	A_L	A	B_L	B	C_L	C
A_L	0	1	0	0	0	0
A	0	0.8	0.1	0	0.1	0
B_L	0	0	0	1	0	0
B	0.15	0	0	0.7	0.15	0
C_L	0	0	0	0	0	1
C	0.05	0	0.05	0	0	0.9

This is a Markov chain because the probability of future transitions is independent of history prior to arrival in current state (Markov property).

(iii) Need to find stationary distribution $\underline{\pi}$ which by definition satisfies:

$$\underline{\pi} = \underline{\pi}\mathbf{A}$$

$$0.15\pi_B + 0.05\pi_C = \pi_{A_L} \quad (1)$$

$$\pi_{A_L} + 0.8\pi_A = \pi_A \quad (2)$$

$$0.1\pi_A + 0.05\pi_C = \pi_{B_L} \quad (3)$$

$$\pi_{B_L} + 0.7\pi_B = \pi_B \quad (4)$$

$$0.1\pi_A + 0.15\pi_B = \pi_{C_L} \quad (5)$$

$$\pi_{C_L} + 0.9\pi_C = \pi_C \quad (6)$$

Combining (1) and (2), (3) and (4), and (5) and (6)

$$0.15\pi_B + 0.05\pi_C = 0.2\pi_A \qquad (1A)$$

$$0.1\pi_A + 0.05\pi_C = 0.3\pi_B \qquad (3A)$$

$$0.1\pi_A + 0.15\pi_B = 0.1\pi_C \qquad (5A)$$

(1A) – (3A) gives

$$\pi_A = 1.5\pi_B$$

(3A) – (5A) produces

$$\pi_C = 3\pi_B$$

$\sum_i \pi_i = 1$ implies

$$(1.5 + 0.3 + 1 + 0.3 + 3 + 0.3)\pi_B = 1$$

So $\begin{pmatrix} \pi_{A_L} \\ \pi_A \\ \pi_{B_L} \\ \pi_B \\ \pi_{C_L} \\ \pi_C \end{pmatrix} = \begin{pmatrix} 0.046875 \\ 0.234375 \\ 0.046875 \\ 0.15625 \\ 0.046875 \\ 0.46875 \end{pmatrix}$

And proportions using (A,B,C) are

(0.28125, 0.203125, 0.515625).

8 (i) $h(z,t) = h_0(t).\exp(\underline{\beta}.z_i^T)$

where $h(z,t)$ is the hazard at duration t

$h_o(t)$ is the baseline hazard

z_i are the covariates

$\underline{\beta}$ is the vector of regression parameters

(ii)

z_1 = 1 plays violin, 0 otherwise	$\beta_1 = 0.07$
z_2 = 1 plays trumpet, 0 otherwise	$\beta_2 = 0.14$
z_3 = 1 new tuition method, 0 otherwise	$\beta_3 = -0.05$
z_4 = 1 male, 0 otherwise	$\beta_4 = 0.02$

(iii) Baseline hazard refers to

a female,
following traditional tuition method,
playing the piano

(iv) The parameter associated with the new tuition method is -0.05. Because the parameter is negative, the hazard of dropping out is reduced by the new tuition method.
Therefore the new tuition method does appear to improve the chances of a child continuing with his or her instrument.

However the 95% confidence interval for the parameter spans zero.
So at the 5% significance level it is not possible to conclude that the new tuition method has improved the chances of children continuing to play their instrument.

(v) The hazard for a girl being taught the trumpet by the traditional method giving up is $h_0(t)\exp(0.14)$.

Therefore the probability of her still playing after 4 years is

$$S_{female}(4) = \exp\left(-\int_0^4 h_0(t)\exp(0.14)dt\right) = \exp\left(-1.150274\int_0^4 h_0(t)dt\right)$$

Since this is equal to 0.7, we have

$$\exp\left(-1.150274\int_0^4 h_0(t)dt\right) = 0.7\text{, so that}$$

$$\log_e 0.7 = -1.150274\int_0^4 h_0(t)dt\,,$$

and hence $\int_0^4 h_0(t)dt = \dfrac{\log_e 0.7}{-1.150274} = 0.310078.$

The hazard of giving up for a boy taught the piano by the new method is $h_0(t)\exp(-0.05+0.02) = h_0(t)\exp(-0.03).$

Therefore the probability of him still playing after 4 years is

$$S_{male}(4) = \exp\left(-\int_0^4 h_0(t)\exp(-0.03)dt\right) = \exp\left[-0.310078(0.970446)\right]$$

which is exp(-0.300914) = 0.74014.

ALTERNATIVELY

The hazard of giving up for a girl being taught the trumpet by the traditional method is $h_0(t)\exp(\beta_2)$.

Therefore the probability of her still playing after 4 years is

$$S_{female}(4) = \exp\left(-\int_0^4 h_0(t)\exp(\beta_2)dt\right) = \exp\left(-\exp(\beta_2)\int_0^4 h_0(t)dt\right)$$

and hence

$$\int_0^4 h_0(t)dt = \frac{\log_e[S_{female}(4)]}{-\exp\beta_2} = -\exp(-\beta_2)\log_e[S_{female}(4)]\,.$$

The hazard of a boy being taught the piano by the new method giving up is $h_0(t)\exp(\beta_3+\beta_4)$.

Therefore the probability of him still playing after 4 years is

$$S_{male}(4) = \exp\left(-\exp(\beta_3 + \beta_4)\int_0^4 h_0(t)dt\right).$$

Substituting for $\int_0^4 h_0(t)dt$ produces

$$S_{male}(4) = \exp\left(\exp(\beta_3 + \beta_4)\exp(-\beta_2)\log_e[S_{female}(4)]\right)$$

= exp[exp(-0.05+0.02)exp(-0.14)$\log_e$(0.7)]
= exp[0.970446 x 0.869358 x -0.356675)
= 0.74014.

9 (i) Type I censoring is present

because the study ends at a predetermined duration of 30 days.

Type II censoring is not present

because the study did not end after a predetermined number of patients had died

Random censoring is present

because the duration at which a patient left hospital before the study ended can be considered as a random variable.

(ii) Yes

Those patients who left hospital before 30 days had elapsed are more likely to be recovering well than those patients who remained in hospital, and so will probably be less likely to die.

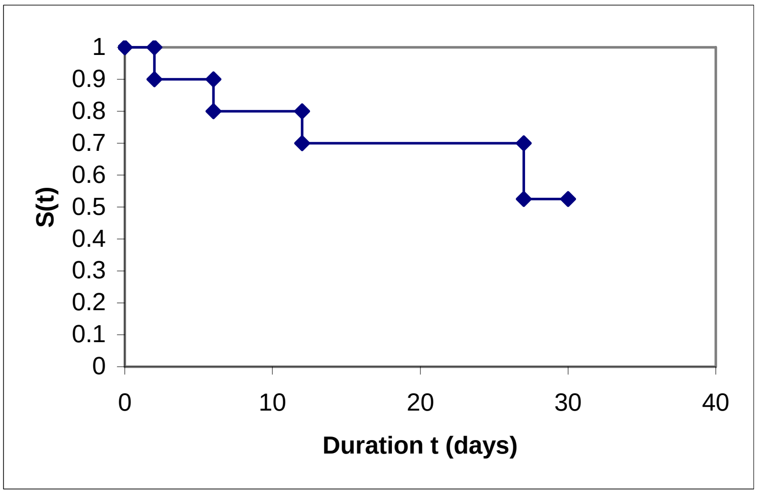
(iii) The Kaplan-Meier estimate of the survival function is estimated as follows

t_j	n_j	d_j	c_j	$\frac{d_j}{n_j}$	$1 - \frac{d_j}{n_j}$	$\prod_{t_j \le t} 1 - \frac{d_j}{n_j} = \hat{S}(t)$
0	10					
2	10	1	0	1/10	9/10	9/10 = 0.9
6	9	1	0	1/9	8/9	8/10 = 0.8
12	8	1	2	1/8	7/8	7/10 = 0.7
27	5	1	4	1/5	4/5	14/25 = 0.56

The Kaplan-Meier estimate of the survival function at duration 28 days is therefore 0.56.

(iv) The Kaplan-Meier estimate of the hazard at duration 8 days is 0.

(v) A suitable sketch is shown below.



10 (i) Operates in continuous time ($t \geq 0$)

with discrete state space {ONline, OFFline}, and transition probability does not depend on history prior to arrival in current state (Markov property).

(ii) $P'_{OFF}(t) = 0.8 * P_{ON}(t) - 0.2 * P_{OFF}(t)$

(iii) As there are only two states,

$$P_{ON}(t) + P_{OFF}(t) = 1$$

Substituting using the solution to (ii), we obtain

$$P'_{OFF}(t) + P_{OFF}(t) = 0.8$$

so that

$$d/_{dt}(e^t P_{OFF}(t)) = 0.8 * e^t$$

$$e^t P_{OFF}(t) = 0.8 * e^t + \text{constant}$$

Boundary condition $P_{OFF}(0) = 1$

So $P_{OFF}(t) = 0.8 + 0.2e^{-t}$

(iv) If O_t is a random variable denoting the amount of time spent offline and I_t is an indicator variable which takes the value 1 if offline, 0 otherwise then required expected value is

$$E[O_t \,|P_{OFF}(0) = 1] = \int_0^t E[I_s \,|P_{OFF}(0) = 1]ds = \int_0^t P_{OFF}(s)ds$$

$$\int_0^t P_{OFF}(s)ds = \int_0^t (0.8 + 0.2e^{-s})ds = \left|0.8t - 0.2e^{-t}\right|_0^t = 0.8t + 0.2(1 - e^{-t})$$

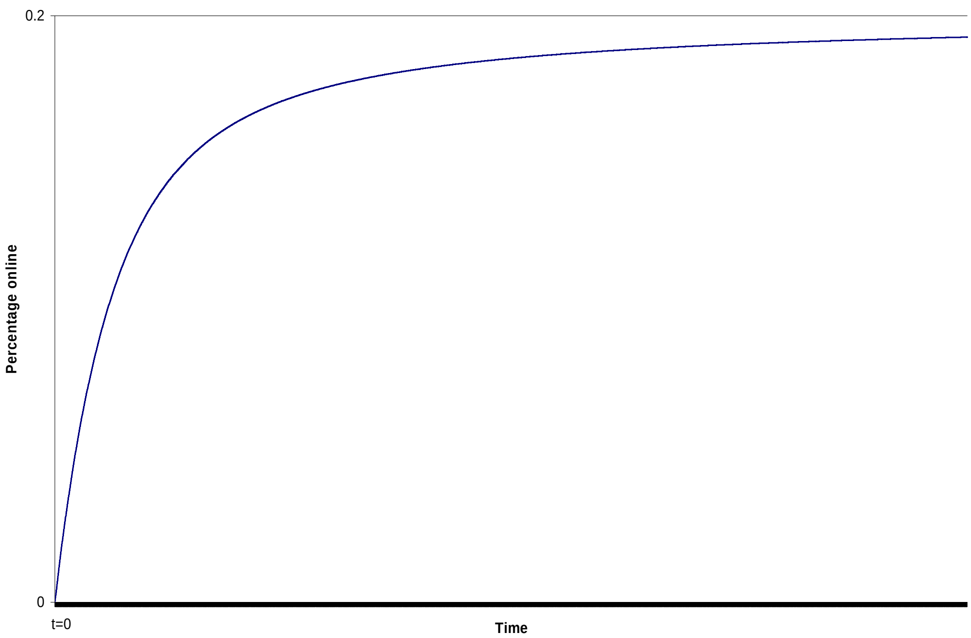
Either online or offline at any time so time spent online is:

$$t - (0.8t + 0.2(1 - e^{-t}) = 0.2t - 0.2(1 - e^{-t})$$

So proportion spent online is:

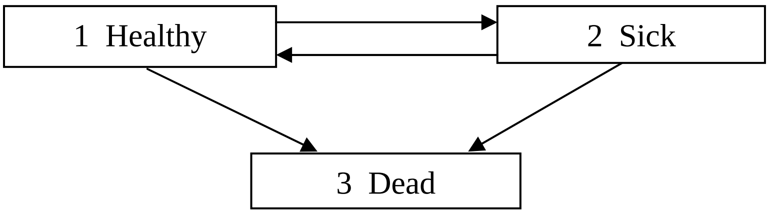
$$\frac{0.2t - 0.2(1 - e^{-t})}{t} = 0.2 - 0.2(\frac{1 - e^{-t}}{t})$$

(v) A suitable sketch is shown below.



Shape: starts at zero as given offline at that point,
asymptotes to ratio of connection to
(connection + disconnection) rates.

11 (i)



(ii) By the Markov assumption OR conditioning on the state occupied at time $x+t$

$${}_{t+dt}p_x^{23} = {}_t p_x^{21}\,{}_{dt}p_{x+t}^{13} + {}_t p_x^{22}\,{}_{dt}p_{x+t}^{23} + {}_t p_x^{23}\,{}_{dt}p_{x+t}^{33}.$$

But ${}_{dt}p_{x+t}^{33} = 1$, so

$${}_{t+dt}p_x^{23} = {}_t p_x^{21}\,{}_{dt}p_{x+t}^{13} + {}_t p_x^{22}\,{}_{dt}p_{x+t}^{23} + {}_t p_x^{23}.$$

We now assume that

$${}_{dt}p_{x+t}^{23} = \mu_{x+t}^{23}dt + o(dt) \text{ and } {}_{dt}p_{x+t}^{13\,=} = \mu_{x+t}^{13}dt + o(dt)$$

where $o(dt)$ is defined such that $\lim_{dt\to 0}\dfrac{o(dt)}{dt} = 0$.

Substituting for ${}_{dt}p_{x+t}^{23}$ and ${}_{dt}p_{x+t}^{13}$ produces

$${}_{t+dt}p_x^{23} = {}_t p_x^{22}[\mu_{x+t}^{23}dt + o(dt)] + {}_t p_x^{21}[\mu_{x+t}^{13}dt + o(dt)] + {}_t p_x^{23},$$

and, subtracting ${}_t p_x^{23}$ from both sides and taking limits gives

$$\frac{d}{dt}\,{}_t p_x^{23} = \lim_{dt\to 0}\frac{{}_t p_x - {}_{t+dt}p_x}{dt} = {}_t p_x^{21}\mu_{x+t}^{13} + {}_t p_x^{22}\mu_{x+t}^{23}$$

(iii) The likelihood, L, is proportional to

$$\exp[(-\mu^{12}-\mu^{13})v^1]\exp[(-\mu^{23}-\mu^{21})v^2](\mu^{12})^{d^{12}}(\mu^{21})^{d^{21}}(\mu^{13})^{d^{13}}(\mu^{23})^{d^{23}}$$

where v^i is the total observed waiting time in state i,
and d^{ij} is the number of transitions observed from state i to state j.

(iv) Taking the logarithm of the likelihood in the answer to part (iii) gives

$$\log L = -\mu^{23}v^2 + d^{23}\log\mu^{23} \text{ + terms not involving } \mu^{23}$$

Differentiating this with respect to μ^{23} we obtain

$$\frac{d\log L}{d\mu^{23}} = -v^2 + \frac{d^{23}}{\mu^{23}}.$$

Setting this to 0 we obtain the maximum likelihood estimator of μ^{23}

$$\hat{\mu}^{23} = \frac{d^{23}}{v^2}.$$

This is a maximum because $\dfrac{d^2(\log L)}{(d\mu^{23})^2} = -\dfrac{d^{23}}{(\mu^{23})^2}$

which is always negative.

(v) (a) Therefore, if there are 40 transitions from the Sick state to the Dead state and 140 man-years observed in the sick state, the maximum likelihood estimate of μ^{23} is $\dfrac{40}{140} = 0.2857$.

(b) The maximum likelihood estimator of μ^{23} has a variance equal to $\dfrac{\mu^{23}}{E[V]}$, μ^{23} is the true transition rate in the population and $E[V]$ is the expected waiting time in the Sick state.

Approximating μ^{23} by $\hat{\mu}^{23}$ and $E[V]$ by v^2 we estimate for the variance as $\dfrac{0.2857}{140} = 0.00204$.

A 95 per cent confidence interval around our estimate of μ^{23} is therefore $0.2857 \pm 1.96\sqrt{0.00204}$

which is 0.2857 ± 0.0885

or (0.1972, 0.3742).

END OF EXAMINERS' REPORT