EXAMINATION

13 April 2005 (am)

Subject CT4 (103) — Models (103 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 6 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 (i) Define each of the following examples of a stochastic process

(a) a symmetric simple random walk
(b) a compound Poisson process

[2]

(ii) For each of the processes in (i), classify it as a stochastic process according to its state space and the time that it operates on. [2]
[Total 4]

2 You have been commissioned to develop a model to project the assets and liabilities of an insurer after one year. This has been requested following a change in the regulatory capital requirement. Sufficient capital must now be held such that there is less than a 0.5% chance of liabilities exceeding assets after one year.

The company does not have any existing stochastic models, but estimates have been made in the planning process of "worst case" scenarios.

Set out the steps you would take in the development of the model. [6]

3 Let Y_1, Y_3, Y_5, …, be a sequence of independent and identically distributed random variables with

$$P\left(Y_{2k+1} = 1\right) = P\left(Y_{2k+1} = -1\right) = \frac{1}{2}, \qquad k = 0,\ 1,\ 2,...$$

and define $Y_{2k} = Y_{2k+1} / Y_{2k-1}$ for k = 1, 2, ….

(i) Show that $\{Y_k : k = 1, 2, ...\}$ is a sequence of independent and identically distributed random variables.

Hint: You may use the fact that, if X, Y are two variables that take only two values and $E\left(XY\right) = E\left(X\right)E(Y)$, then X, Y are independent. [4]

(ii) Explain whether or not $\{Y_k : k = 1, 2, ...\}$ constitutes a Markov chain. [1]

(iii) (a) State the transition probabilities $p_{ij}(n) = P\left(Y_{m+n} = j \mid Y_m = i\right)$ of the sequence $\{Y_k : k = 1, 2, ...\}$.

(b) Hence show that these probabilities do not depend on the current state and that they satisfy the Chapman-Kolmogorov equations.

[3]
[Total 8]

4 Marital status is considered using the following time-homogeneous, continuous time Markov jump process:

- the transition rate from unmarried to married is 0.1 per annum
- the divorce rate is equivalent to a transition rate of 0.05 per annum
- the mortality rate for any individual is equivalent to a transition rate of 0.025 per annum, independent of marital status

The state space of the process consists of five states: Never Married (NM), Married (M), Widowed (W), Divorced (DIV) and Dead (D).

P_x is the probability that a person currently in state x, and who has never previously been widowed, will die without ever being widowed.

(i) Construct a transition diagram between the five states. [2]

(ii) Show, by general reasoning or otherwise, that P_{NM} equals P_{DIV}. [1]

(iii) Demonstrate that:

$$P_{NM} = \frac{1}{5} + \frac{4}{5} \times P_M$$

$$P_M = \frac{1}{4} + \frac{1}{2} \times P_{DIV}$$

[2]

(iv) Calculate the probability of never being widowed if currently in state NM. [2]

(v) Suggest two ways in which the model could be made more realistic. [1]

[Total 8]

5 A No-Claims Discount system operated by a motor insurer has the following four levels:

Level 1: 0% discount
Level 2: 25% discount
Level 3: 40% discount
Level 4: 60% discount

The rules for moving between these levels are as follows:

- Following a year with no claims, move to the next higher level, or remain at level 4.
- Following a year with one claim, move to the next lower level, or remain at level 1.
- Following a year with two or more claims, move back two levels, or move to level 1 (from level 2) or remain at level 1.

For a given policyholder the probability of no claims in a given year is 0.85 and the probability of making one claim is 0.12.

$X(t)$ denotes the level of the policyholder in year t.

(i) (a) Explain why $X(t)$ is a Markov chain.
(b) Write down the transition matrix of this chain.
[2]

(ii) Calculate the probability that a policyholder who is currently at level 2 will be at level 2 after:

(a) one year
(b) two years
(c) three years
[3]

(iii) Explain whether the chain is irreducible and/or aperiodic. [2]

(iv) Calculate the long-run probability that a policyholder is in discount level 2.
[5]
[Total 12]

6 An insurance policy covers the repair of a washing machine, and is subject to a maximum of 3 claims over the year of coverage.

The probability of the machine breaking down has been estimated to follow an exponential distribution with the following annualised frequencies, λ:

$$\lambda = \begin{cases} 1/10 & \text{If the machine has not suffered any previous breakdown.} \\ 1/5 & \text{If the machine has broken down once previously.} \\ 1/4 & \text{If the machine has broken down on two or more occasions.} \end{cases}$$

As soon as a breakdown occurs an engineer is despatched. It can be assumed that the repair is made immediately, and that it is always possible to repair the machine.

The washing machine has never broken down at the start of the year (time $t = 0$).

$P_i(t)$ is the probability that the machine has suffered i breakdowns by time t.

(i) Draw a transition diagram for the process defined by the number of breakdowns occurring up to time t. [1]

(ii) Write down the Kolmogorov equations obeyed by $P_0'(t), P_1'(t)$ and $P_2'(t)$. [2]

(iii) (a) Derive an expression for $P_0(t)$ and

(b) demonstrate that $P_1(t) = e^{-\frac{t}{10}} - e^{-\frac{t}{5}}$. [3]

(iv) Derive an expression for $P_2(t)$. [3]

(v) Calculate the expected number of claims under the policy. [3]

[Total 12]

END OF PAPER

EXAMINATION

13 April 2005 (am)

Subject CT4 (104) — Models (104 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 7 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 (i) Write down the equation of the Cox proportional hazards model in which the hazard function depends on duration t and a vector of covariates $\mathbf{z}$. You should define all the other terms that you use. [2]

(ii) Explain why the Cox model is sometimes described as "semi-parametric". [1]
[Total 3]

2 Show that if the force of mortality μ_{x+t} $(0 \le t \le 1)$ is given by

$$\mu_{x+t} = \frac{q_x}{1 - tq_x},$$

this implies that deaths between exact ages x and $x + 1$ are uniformly distributed. [4]

3 An investigation of mortality over the whole age range produced crude estimates of q_x for exact ages x from 2 years to 93 years inclusive. The actual deaths at each age were compared with the number of deaths which would have been expected had the mortality of the lives in the investigation been the same as English Life Table 15 (ELT15). 53 of the deviations were positive and 39 were negative.

Test whether the underlying mortality of the lives in the investigation is represented by ELT15. [5]

4 A life insurance company has investigated the recent mortality experience of its male term assurance policy holders by estimating the mortality rate at each age, q_x. It is proposed that the crude rates might be graduated by reference to a standard mortality table for male permanent assurance policy holders with forces of mortality $\mu^s_{x+\frac{1}{2}}$, so that the forces of mortality $\mathring{\mu}_{x+\frac{1}{2}}$ implied by the graduated rates $\mathring{q}_x$ are given by the function:

$$\mathring{\mu}_{x+\frac{1}{2}} = \mu^s_{x+\frac{1}{2}} + k,$$

where k is a constant.

(i) Describe how the suitability of the above function for graduating the crude rates could be investigated. [2]

(ii) (a) Explain how the constant k can be estimated by weighted least squares.

(b) Suggest suitable weights.
[4]

(iii) Explain how the smoothness of the graduated rates is achieved. [1]
[Total 7]

5 A study of the mortality of 12 laboratory-bred insects was undertaken. The insects were observed from birth until either they died or the period of study ended, at which point those insects still alive were treated as censored.

The following table shows the Kaplan-Meier estimate of the survival function, based on data from the 12 insects.

t (weeks)	*S(t)*
$0 \le t < 1$	1.0000
$1 \le t < 3$	0.9167
$3 \le t < 6$	0.7130
$6 \le t$	0.4278

(i) Calculate the number of insects dying at durations 3 and 6 weeks. [6]

(ii) Calculate the number of insects whose history was censored. [1]
[Total 7]

6 An investigation into mortality collects the following data:

θ_x = total number of policies under which death claims are made when the policyholder is aged x last birthday in each calendar year

$P_x(t)$ = number of in-force policies where the policyholder was aged x nearest birthday on 1 January in year t

(i) State the principle of correspondence. [1]

(ii) Obtain an expression, in terms of the $P_x(t)$, for the central exposed to risk, E_x^c, which corresponds to the claims data and which may be used to estimate the force of mortality in year t at each age x, μ_x. State any assumptions you make. [4]

(iii) Comment on the effect on the estimation of the fact that the θ_x relate to claims, rather than deaths, and the $P_x(t)$ relate to policies, not lives. [4]
[Total 9]

7 An investigation took place into the mortality of pensioners. The investigation began on 1 January 2003 and ended on 1 January 2004. The table below gives the data collected in this investigation for 8 lives.

Date of birth	*Date of entry into observation*	*Date of exit from observation*	*Whether or not exit was due to death (1) or other reason (0)*
1 April 1932	1 January 2003	1 January 2004	0
1 October 1932	1 January 2003	1 January 2004	0
1 November 1932	1 March 2003	1 September 2003	1
1 January 1933	1 March 2003	1 June 2003	1
1 January 1933	1 June 2003	1 September 2003	0
1 March 1933	1 September 2003	1 January 2004	0
1 June 1933	1 January 2003	1 January 2004	0
1 October 1933	1 June 2003	1 January 2004	0

The force of mortality, μ_{70}, between exact ages 70 and 71 is assumed to be constant.

(i) (a) Estimate the constant force of mortality, μ_{70}, using a two-state model and the data for the 8 lives in the table.

(b) Hence or otherwise estimate q_{70}.

[7]

(ii) Show that the maximum likelihood estimate of the constant force, μ_{70}, using a Poisson model of mortality is the same as the estimate using the two-state model. [5]

(iii) Outline the differences between the two-state model and the Poisson model when used to estimate transition rates. [3]

[Total 15]

END OF PAPER

EXAMINATION

April 2005

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

15 June 2005

EXAMINERS' COMMENTS

Comments on solutions presented to individual questions for this April 2005 paper are given below:

103 Part

Question A1 *This was reasonably well answered.*
Descriptive (rather than formulaic) answers to part (i) were given equal credit. Very few candidates correctly identified the state space for the compound Poisson process in part (ii).

Question A2 *This was reasonably well answered.*
Marks were lost by candidates who did not provide sufficient detail or did not provide enough distinct points. Some candidates attempted to define the model they would adopt, rather than the stages in the modelling process.

Question A3 *This was very poorly attempted by most candidates.*
Very few candidates provided any real attempt at part (i). The examiners were looking here for a demonstration of pairwise (not mutual) independence, and the hint should have made this clear.
In part (ii), most candidates wrongly stated that the sequence was Markov.
Many candidates did not attempt part (iii); this may be because of the failure to make any progress in part (i), although it should be noted that subsequent parts of the question did not depend on correctly answering part (i).

Question A4 *This was well answered overall.*
In part (i), some candidates did not allow for re-marriage from the divorced or widowed states, which then caused them problems in part (ii).
Candidates lost marks in part (iii) if they did not provide sufficient explanation of their steps.

Question A5 *This was very well answered, with the majority of candidates scoring highly.*

Question A6 *Overall this was not well answered, but the better candidates did score well.*
Many candidates produced good answers to part (i) to (iv). In part (iii), a number of candidates did not verify that the boundary conditions were satisfied.
Some candidates struggled with part (v) and a significant number did not attempt this part of the question.

104 Part

Question B1 *This was well answered overall.*
Most candidates answered part (i) well, but many then struggled to express clearly what was required in part (ii).

Question B2 *This was very poorly answered.*
Many candidates did not seem to know how to start this, with a significant number starting with the uniform distribution assumption and working backwards.

Question B3 *This was well answered overall. Many candidates included a continuity correction. This was not necessary, as there were 92 ages, but candidates who did so received full credit if they used it correctly.*

Question B4 *This was not well answered.*
In part (i) significant numbers of candidates talked about general goodness of fit tests. This did not receive credit, as it was the appropriateness of the linear form of the function that we were looking for, before doing the graduation. Goodness-of-fit tests come later, after the graduation has been done, and were not part of this question.
In parts (i) and (ii), many candidates considered the graduated rates rather than the crude rates, for example plotting $\overset{\circ}{m}_{x+\frac{1}{2}}$ *against* $\mu^s_{x+\frac{1}{2}}$ *and this was penalised.*

Question B5 *This was well answered.*
Some candidates assumed that there was no censoring until the end of the investigation. This led to a non-integer number of deaths, which should have indicated an error, but few of these candidates realised this.

Question B6 *Most candidates correctly answered part (i).*
As with similar questions in previous years, part (ii) was not well answered. Many candidates lost marks by not providing sufficient explanation of their working.
In part (iii), most candidates mentioned the "variance ratio" and gave the formula from the gold book, but many did not provide a good explanation of what this meant in practice.

Question B7 *This was reasonably well answered overall.*
In part (i), candidates were asked to "estimate", so some indication of how they reached their answer was required for full credit.

103 Part

A1 (i) (a) Let $Y_1, Y_2, \ldots, Y_j, \ldots$, be a sequence of independent and identically distributed random variables with

$$P\left(Y_j = 1\right) = P\left(Y_j = -1\right) = \frac{1}{2}$$

and define

$$X_n = \sum_{j=1}^{n} Y_j$$

Then $\left\{X_n\right\}_{n=1}^{\infty}$ constitutes a symmetric simple random walk.

(b) Let N_t be a Poisson process, $t \geq 0$ and let $Y_1, Y_2, \ldots, Y_j, \ldots$, be a sequence of i.i.d. random variables. Then a compound Poisson process is defined by

$$X_t = \sum_{j=1}^{N_t} Y_j, \qquad t \geq 0.$$

(ii) (a) A simple random walk operates on discrete time and has a discrete state space (the set of all integers, Z).

(b) A compound Poisson process operates on continuous time.

It has a discrete or continuous state space depending on whether the variables Y_j are discrete or continuous respectively.

A2

- Review the regulatory guidance.
- Define the scope of the model, for example which factors need to be modelled stochastically.
- Plan the development of the model, including how the model will be tested and validated.
- Consider alternative forms of model, and decide and document the chosen approach. Where appropriate, this may involve discussion with experts on the underlying stochastic processes.

- Collect any data required, for example historic losses or policy data.

- Choose parameters. For economic factors should be able to calibrate to market data. For other factors e.g. expenses, claim distributions need to discuss with staff.

- Existing "worst case" scenarios. Discuss with staff who made the estimates, especially to gauge views on the probability of events occurring.

- Decide on the software to be used for the model.

- Write the computer programs.

- Debug the program, for example by checking the model behaves as expected for simple, defined scenarios.

- Review the reasonableness of the output. May include:

  – median outcomes (how do these compare with business plans)
  – what probability is assigned to "worst case" scenarios

- Test the sensitivity of the model to small changes in parameters.

- Calculate the capital requirement.

- Communicate findings to management. Document.

Other suitable points were given credit, including:

- Validate data.
- Run model on historic data to compare model's predictions with previous observations.
- Review parameters that have greatest effect on outputs.
- Present range of capital requirements for differing parameter inputs.

A3 (i) It is clear that Y_{2k} can only take two values, ±1, with probabilities

$$P\left(Y_{2k}=1\right)=P\left(Y_{2k+1}=Y_{2k-1}=+1\right)+P\left(Y_{2k+1}=Y_{2k-1}=-1\right)=\frac{1}{2}$$

and

$$P\left(Y_{2k}=-1\right)=$$
$$P\left(Y_{2k+1}=+1,Y_{2k-1}=-1\right)+P\left(Y_{2k+1}=-1,Y_{2k-1}=+1\right)=\frac{1}{2}$$

so that they have the same distribution as Y_{2k+1}.

To show that Y_{2k}, Y_{2k+1} are independent, we observe first that

$E(Y_{2k}) = E(Y_{2k+1}) = 0.$

Next,

$$E(Y_{2k}Y_{2k+1}) =$$
$$\frac{1}{2}E(Y_{2k}Y_{2k+1} \mid Y_{2k-1} = 1) + \frac{1}{2}E(Y_{2k}Y_{2k+1} \mid Y_{2k-1} = -1)$$

But

$$E(Y_{2k}Y_{2k+1} \mid Y_{2k-1} = 1) = 1\times 1 + 0\times(-1) = 1,$$

and similarly $E(Y_{2k}Y_{2k+1} \mid Y_{2k-1} = -1) = -1,$ which yields that

$$E(Y_{2k}Y_{2k+1}) = \frac{1}{2}\times 1 + \frac{1}{2}\times(-1) = 0.$$

Since

$$E(Y_{2k}) = E(Y_{2k+1}) = E(Y_{2k}Y_{2k+1})$$

it now follows from the hint that Y_{2k}, Y_{2k+1} are independent.

For the proof to be complete, we need to show that Y_{2k}, Y_{2m} are also independent for all *k, m*. This is obvious from the statement for all *k, m* except when *m* = *k* + 1 or *m* = *k* - 1. For this case, we could either argue as above or simply state that it is obvious by symmetry.

(ii) The sequence $\{Y_k : k = 1,2,...\}$ is not Markov; for instance

$$P(Y_{2k+1} = -1 \mid \; Y_{2k} = 1) = \frac{1}{2}$$

but

$$P(Y_{2k+1} = -1 \mid Y_{2k} = 1,\ Y_{2k-1} = 1) = 0.$$

(iii) (a) Since the Y_k are pairwise independent, we see that for all *i, j, m, n,*

$$p_{ij}(n) = P(Y_{m+n} = j \mid \; Y_m = i) = \frac{1}{2}.$$

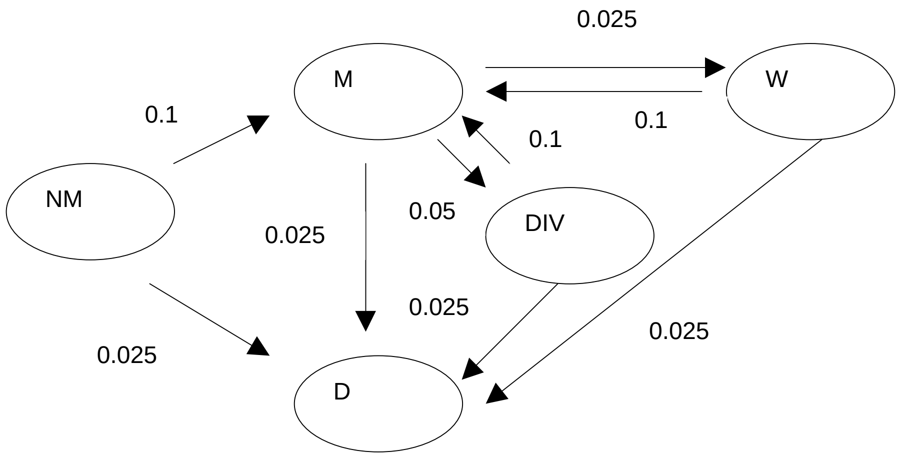
(b) The probabilities do not depend on the current state as they are all ½

Using the result in (a) we therefore see that

$$\sum_{k\in\{-1,1\}} p_{ik}(n)p_{kj}(r) = \frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}=\frac{1}{2}$$
$$= p_{ij}(n+r).$$

which shows that the Chapman – Kolmogorov equations are satisfied although $\{Y_k : k = 1,2,...\}$ is not Markov.

A4 (i)



(ii) The transitions out of the divorced state are to the same states, and with the same transition probabilities, as the transitions out of state NM.
Therefore the probability of ever reaching state W is the same from both states.

Alternatively, this could be shown by producing the equation conditioning on the first move out of DIV, as in part (iii), and showing this is identical to that for P_{NM} .

(iii) Conditioning on the first move out of each state:

$$P_{NM} = \frac{0.025}{0.125}\times P_D + \frac{0.1}{0.125}\times P_M$$

$$P_M = \frac{0.025}{0.1}\times P_D + \frac{0.05}{0.1}\times P_{DIV} + \frac{0.025}{0.1}\times P_W$$

As $P_D = 1$ and $P_W = 0$, these give

$$P_{NM} = \frac{0.025}{0.125} + \frac{0.1}{0.125} \times P_M = \frac{1}{5} + \frac{4}{5} \times P_M$$

$$P_M = \frac{0.025}{0.1} + \frac{0.05}{0.1} \times P_{DIV} = \frac{1}{4} + \frac{1}{2} \times P_{DIV}$$

as required.

(iv) Using $P_{NM} = P_{DIV}$ in the above equations gives:

$$P_{NM} = \frac{1}{5} + \frac{4}{5} \times \left(\frac{1}{4} + \frac{1}{2} \times P_{NM} \right)$$

$$\Rightarrow \left(1 - \frac{2}{5}\right) \times P_{NM} = \frac{2}{5}$$

$$\Rightarrow P_{NM} = \frac{2}{3}$$

(v)

- Make mortality and marriage rates age dependent.
- Divorce rate dependent on duration of marriage.
- Divorce rate dependent on whether previously divorced.
- Make mortality rate marital status-dependent.

Other sensible suggestions received credit.

A5 (i)(a) It is clear that $X(t)$ is a Markov chain; knowing the present state, any additional information about the past is irrelevant for predicting the next transition.

(b) The transition matrix of the process is

$$P = \begin{pmatrix} 0.15 & 0.85 & 0 & 0 \\ 0.15 & 0 & 0.85 & 0 \\ 0.03 & 0.12 & 0 & 0.85 \\ 0 & 0.03 & 0.12 & 0.85 \end{pmatrix}$$

(ii)(a) For the one year transition, $p_{22} = 0$,
as can be seen from above (or is obvious from the statement).

(b) The possible transitions, and relevant probabilities are:

$2 \rightarrow 1 \rightarrow 2: \quad 0.15 \times 0.85 = 0.1275$

$2 \rightarrow 3 \rightarrow 2: \quad 0.85 \times 0.12 = 0.102$

The required probability is 0.1275 + 0.102 = 0.2295

Alternatively

The second order transition matrix is

$$P^2 = \begin{pmatrix} 0.15^2+0.85\times0.15 & 0.85\times0.15 & 0.85^2 & 0 \\ 0.15^2+0.85\times0.03 & 0.85\times0.15+0.85\times0.12 & 0 & 0.85^2 \\ 0.03\times0.15+0.12\times0.15 & 0.85\times0.03\times2 & 0.85\times0.12\times2 & 0.85^2 \\ 0.03\times0.15+0.12\times0.03 & 0.12^2+0.85\times0.03 & 0.85\times0.03+0.85\times0.12 & 0.12\times0.85+0.85^2 \end{pmatrix}$$

$$= \begin{pmatrix} 0.15 & 0.1275 & 0.7225 & 0 \\ 0.048 & 0.2295 & 0 & 0.7225 \\ 0.0225 & 0.051 & 0.204 & 0.7225 \\ 0.0081 & 0.0399 & 0.1275 & 0.8245 \end{pmatrix}$$

Hence the required probability is 0.2295.

(c) The possible transitions, and relevant probabilities are:

$2\to1\to1\to2:$ $0.15\times0.15\times0.85=0.019125$
$2\to3\to1\to2:$ $0.85\times0.03\times0.85=0.021675$
$2\to3\to4\to2:$ $0.85\times0.85\times0.03=0.021675$

The required probability is
0.019125 + 0.021675 + 0.021675 = 0.062475

Alternatively

The relevant entry from the third-order transition matrix equals

$0.15\times0.1275+0.85\times0.051=0.062475.$

(iii) The chain is irreducible as
any state is reachable from any other.

It is also aperiodic;
If currently at either state 1 or 4, it can remain there. This is not true for states 2 and 3, however these are also aperiodic states since the chain may return e.g. to state 2 after 2 or 3 transitions.

(iv) In matrix form, the equation we need to solve is $\pi P = \pi$,
where π is the vector of equilibrium probabilities.

This reads

$$
\begin{array}{lllll}
0.15\pi_1 + 0.15\pi_2 + 0.03\pi_3 & & & = \pi_1 & (1) \\
0.85\pi_1 + & +0.12\pi_3 + 0.03\pi_4 & & = \pi_2 & (2) \\
+0.85\pi_2 & & +0.12\pi_4 & = \pi_3 & (3) \\
 & 0.85\pi_3 + 0.85\pi_4 & & = \pi_4 & (4)
\end{array}
$$

Discard the first of these equations and use also that $\sum_{i=1}^{4} \pi_i = 1$. Then, we obtain first from (4) that $0.85\pi_3 = 0.15\pi_4$ or, that $\pi_4 = 17\pi_3/3$

Substituting in (3) this gives

$$0.85\pi_2 + 0.12 \times \frac{17}{3}\pi_3 = \pi_3 \Rightarrow \pi_3 = 2.65625\pi_2$$

(2) now yields that

$$
\begin{aligned}
0.85p_1 &= p_2 - 0.12p_3 - 0.03p_4 \\
&= \frac{1}{2.65625}p_3 - 0.12p_3 - 0.17p_3 = 0.0865p_3,
\end{aligned}
$$

so that finally we get $\pi_1 = 0.10173\pi_3$.

Using now that the probabilities must add up to one, we obtain
$\pi_1 + \pi_2 + \pi_3 + \pi_4 = (0.10173 + 0.3765 + 1 + 5.666)\pi_3 = 1$,
or that $\pi_3 = 0.13996$.

Solving back for the other variables we get that
$\pi_1 = 0.01424$, $\pi_2 = 0.05269$, $\pi_4 = 0.79311$

The long-run probability that the motorist is in discount level 2 is therefore 0.05269.

A6 (i)

$$\text{No Breakdowns} \xrightarrow{1/10} \text{One Breakdown} \xrightarrow{1/5} \text{Two Breakdowns} \xrightarrow{1/4} \text{Three Breakdowns} \cdots$$

(ii) $$P_0'(t) = -\frac{1}{10} \times P_0(t)$$

$$P_1'(t) = \frac{1}{10} \times P_0(t) - \frac{1}{5} \times P_1(t)$$

$$P_2'(t) = \frac{1}{5} \times P_1(t) - \frac{1}{4} \times P_2(t)$$

(iii)(a) Dividing the first equation by $P_0(t)$:

$$\frac{d}{dt}\left[\ln P_0(t)\right] = -\frac{1}{10}$$

Hence, using the boundary condition $P_0(0) = 1$

$$P_0(t) = e^{\frac{-t}{10}}$$

(b) Substitute into the second equation above to obtain

$$P_1'(t) = \frac{1}{10} e^{-\frac{t}{10}} - \frac{1}{5} * P_1(t)$$

Using an integrating factor $e^{\frac{t}{5}}$, we get

$$e^{\frac{t}{5}} \times \left[P_1'(t) + \frac{1}{5} P_1(t)\right] = \frac{1}{10} \times e^{-\frac{t}{10}+\frac{t}{5}}$$

$$\Rightarrow \frac{d}{dt}\left[e^{\frac{t}{5}} \times P_1(t)\right] = \frac{1}{10} \times e^{\frac{t}{10}}$$

$$\Rightarrow e^{\frac{t}{5}} \times P_1(t) = e^{\frac{t}{10}} + const$$

$$\Rightarrow P_1(t) = e^{-\frac{t}{10}} + const \times e^{-\frac{t}{5}}$$

$$\Rightarrow P_1(t) = \exp^{-\frac{t}{10}} - \exp^{-\frac{t}{5}}$$

using boundary condition $P_1(0) = 0$

Alternatively

Differentiate the suggested solution and verify it obeys the second equation.

And that the boundary condition is satisfied.

(iv) Proceeding in a similar way with the equation for $P_2(t)$

$$P_2'(t) = \frac{1}{5}\exp^{-\frac{t}{10}} - \frac{1}{5}\exp^{-\frac{t}{5}} - \frac{1}{4} * P_2(t)$$

$$\frac{d}{dt}\left[\exp^{\frac{t}{4}} \times P_2(t)\right] = \frac{1}{5} \times (\exp^{\frac{3}{20}t} - \exp^{\frac{1}{20}t})$$

$$\exp^{\frac{t}{4}} \times P_2(t) = \frac{4}{3} \times \exp^{\frac{3}{20}t} - 4 \times \exp^{\frac{1}{20}t} + \frac{8}{3}$$

$$P_2(t) = \frac{4}{3}[\exp^{-\frac{t}{10}} - 3 \times \exp^{-\frac{t}{5}} + 2 \times \exp^{-\frac{t}{4}}]$$

(v) $$\text{Expected Claims} = 1 \times P_1(1) + 2 \times P_2(1) + 3 \times \sum_{i=3}^{\infty} P_i(1)$$

$$= P_1(1) + 2 \times P_2(1) + 3 \times \left(1 - P_0(1) - P_1(1) - P_2(1)\right)$$

$$P_0(1) = \exp^{-1/10} = 0.905$$

$$P_1(1) = \exp^{-1/10} - \exp^{-1/5} = 0.0861$$

$$P_2(1) = \frac{4}{3}[\exp^{-\frac{1}{10}} - 3 \times \exp^{-\frac{1}{5}} + 2 \times \exp^{-\frac{1}{4}}] = 0.00832896$$

Substituting these values gives:

Expected Claims = 0.1049

104 Part

B1 (i) If the hazard for life i is $\lambda(t; z_i)$, then

$$l\ (t; z_i) = l\ _0(t)\ \ \exp(b\, z_i{}^T)\ ,$$

where $\lambda_0(t)$ is the baseline hazard,

and β is a vector of regression parameters.

(ii) The model is semi-parametric because is possible to estimate β from the data without estimating the baseline hazard.

Therefore the baseline hazard can have any shape determined by the data.

B2 Since

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

$$ {}_t q_x = 1 - {}_t p_x = 1 - \exp\left(-\int_0^t \mu_{x+s} ds \right).$$

Substituting for μ_{x+s} produces

$$ {}_t q_x = 1 - \exp\left[-\int_0^t \frac{q_x ds}{1 - s q_x} \right]$$

Performing the integration we have

$$\begin{aligned} {}_t q_x &= 1 - \exp\left(-\left[-\log(1 - s q_x) \right]_0^t \right) \\ &= 1 - \exp\left(-\left[-\log(1 - t q_x) + \log 1 \right] \right) \\ &= 1 - \exp\left(-\left[-\log(1 - t q_x) \right] \right) \\ &= 1 - \exp\left(\log(1 - t q_x) \right) \\ &= 1 - (1 - t q_x) \\ &= t q_x . \end{aligned}$$

This is the assumption of a uniform distribution of deaths and implies that deaths between exact ages x and x + 1 are uniformly distributed.

B3 The null hypothesis is that the observed rates are a sample from a population in which English Life Table 15 represents the true rates.

If the null hypothesis is true, then the observed number of positive deviations, *P*, will be such that *P* ~ Binomial (92, ½).

We use the normal approximation to the Binomial distribution because we have > 20 ages

This means that, approximately, *P* ~ Normal (46, 23).

The *z*-score associated with the probability of getting 53 positive deviations if the null hypothesis is true is, therefore

$$\frac{53-46}{\sqrt{23}} = \frac{7}{4.79} = 1.46 .$$

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 -1.96 < 1.46 < +1.96.

(Alternatively, the p-value of the test statistic could be calculated.)

This means we have insufficient evidence to reject the null hypothesis.

B4 (i) The suitability of a linear relationship between $\mu^s_{x+\frac{1}{2}}$ and $\mathring{\mu}_{x+\frac{1}{2}}$ could be investigated by plotting $-\log(1-q_x)$ against $-\log(1-q_x^s)$ or by plotting $\mu_{x+\frac{1}{2}}$ against $\mu^s_{x+\frac{1}{2}}$ and looking for a linear relationship.

An approximately linear relationship will suffice.

If data are scarce, too close a fit is not to be expected, especially at extreme ages.

(ii) (a) We can work with either q_x^s or $\mu^s_{x+\frac{1}{2}}$.

The value of k which minimises either

$$\sum_x w_x (q_x - \overset{\circ}{q}_x)^2$$

or

$$\sum_x w_x \left(\mu_{x+\frac{1}{2}} - \overset{\circ}{\mu}_{x+\frac{1}{2}} \right)^2$$

should be found (note that the summations are over all relevant ages x)

At each age there will be a different sample size or exposed to risk, E_x. This will usually be largest at ages where many term assurances are sold (e.g. ages 25 to 50 years) and smaller at other ages.

(b) The estimation procedure should pay more attention to ages where there are lots of data. These ages should have a greater influence on the choice of k than other ages.

This implies weights $w_x \propto E_x$.
A suitable choice would be

$$w_x = \frac{1}{\operatorname{var} q_x} \text{ or } w_x = \frac{1}{\operatorname{var} \mu_{x+\frac{1}{2}}} \text{ or } w_x = E_x$$

(iii) The graduated forces of mortality are a linear function of the forces in the standard table.

Since the forces in the standard table should already be smooth, a linear function of them will also be smooth.

B5 (i) Consider the durations t_j at which events take place.
Let the number of deaths at duration t_j be d_j and the number of insects still at risk of death at duration t_j be n_j.

At $t_j = 1$, $S(t)$ falls from 1.0000 to 0.9167.

Since the Kaplan-Meier estimate of $S(t)$ is

$$S(t) = \prod_{t_j \le t} (1 - \lambda(t_j)) ,$$

we must have $0.9167 = 1 - \lambda(1)$,

so that $\lambda(1) = 0.0833$.

Since $\lambda(1) = \dfrac{d_1}{n_1}$, then we have $\dfrac{d_1}{n_1} = 0.0833$,

and, since all 12 insects are at risk of dying at $t_j = 1$, we must therefore have $d_1 = 1$ and $n_1 = 12$.

Similarly, at $t_j = 3$, we must have $0.7130 = 0.9167(1 - \lambda(3))$

so that $\lambda(3) = \dfrac{0.9167 - 0.7130}{0.9167} = 0.222 = \dfrac{d_3}{n_3}$.

Since we can have at most 11 insects in the risk set at $t_j = 3$, we must have $d_3 = 2$ and $n_3 = 9$.
Similarly, at $t_j = 6$, we must have $0.4278 = 0.7130(1 - \lambda(6))$,

so that $\lambda(6) = \dfrac{0.7130 - 0.4278}{0.7130} = 0.400 = \dfrac{d_6}{n_6}$.

Since we can have at most 7 insects in the risk set at $t_j = 6$, we must have $d_6 = 2$ and $n_6 = 5$.

Therefore 2 insects died at duration 3 weeks and 2 insects died at duration 6 weeks.

Alternatively

Some candidates worked back to produce a table in the usual format, as follows; this received full credit.

t	$S(t) = \prod(1- \lambda_t)$	λ_t	n_t	d_t	c_t
0	1.0000	0	12	0	
1	0.9167	0.0833	12	1	2
3	0.7130	0.22	9	2	2
6	0.4278	0.4	5	$\underline{2}$	$\underline{3}$
				5	7

(ii) Summing up the number of deaths we have
total deaths = $d_1 + d_3 + d_6 = 1 + 2 + 2 = 5$.

Since we started with 12 insects, the remaining 7 insects' histories were right-censored.

B6 (i) The principle of correspondence states that a life alive at time t should be included in the exposure at age x at time t if and only if were that life to die immediately, he or she would be counted in the deaths data θ_x at age x.

(ii) $P_x(t)$ is the number of policies under observation aged x nearest birthday on 1 January in year t.

To correspond with the claims data, we wish to have policies classified by age last birthday.

Let the number of policies aged x last birthday on 1 January in year t be $P'_x(t)$. Then, assuming that birthdays are evenly distributed,

$$P'_x(t) = \frac{1}{2}\left[P_x(t) + P_{x+1}(t)\right].$$

The central exposed to risk is then given by

$$E_x^c = \int_0^1 P'_x(t)dt.$$

Using the trapezium approximation this is

$$E_x^c \approx \frac{1}{2}\left[P'_x(t) + P'_x(t+1)\right],$$

and, substituting for the $P'_x(t)$ in terms of $P_x(t)$ from the equation above produces

$$E_x^c \approx \frac{1}{2}\left[\frac{1}{2}\left[P_x(t) + P_{x+1}(t)\right] + \frac{1}{2}\left[P_x(t+1) + P_{x+1}(t+1)\right]\right].$$

(iii) The principle of correspondence still holds, because we are dealing with claims and policies: one policy can only lead to one claim.

However, because one life may have more than one policy it is possible that two distinct death claims are the result of the death of the same life.

Therefore claims are not independent, whereas deaths are.

The effect of this is to increase the variance of the number of claims (compared to the situation in which each life has one and only one policy) by the ratio

$$\frac{\sum_i i^2 \pi_i}{\sum_i i \pi_i},$$

where π_i is the proportion of the lives in the investigation owning i policies (i = 1, 2, 3, ...).

Typically the ratio will vary for each age x.

B7 (i)(a) The two-state estimate of μ_{70} is $\frac{d_{70}}{v_{70}}$, where v_{70} is the total time the members of the sample are under observation between exact ages 70 and 71 years.

$v_{70} = \sum_i v_{70,i}$,

where $v_{70,i}$ is the duration that sample member i is under observation between exact ages 70 and 71 years.

For each sample member, $v_{70,i}$ = ENDDATE – STARTDATE

where ENDDATE is the earliest of the date at which the observation of that member ceases and the date of the member's 71st birthday,
and STARTDATE is the latest of the date at which observation of that member begins and the date of the member's 70th birthday.

The table below shows the computation of v_{70}.

i	*Date obs. begins*	*Date of 70th birthday*	*Date obs. ends*	*Date of 71st birthday*	$v_{70,i}$ *(years)*
1	1/1/2003	1/4/2002	1/1/2004	1/4/2003	0.25
2	1/1/2003	1/10/2002	1/1/2004	1/10/2003	0.75
3	1/3/2003	1/11/2002	1/9/2003	1/11/2003	0.5
4	1/3/2003	1/1/2003	1/6/2003	1/1/2004	0.25
5	1/6/2003	1/1/2003	1/9/2003	1/1/2004	0.25
6	1/9/2003	1/3/2003	1/1/2004	1/3/2004	0.3333
7	1/1/2003	1/6/2003	1/1/2004	1/6/2004	0.5833
8	1/6/2003	1/10/2003	1/1/2004	1/10/2004	0.25

Therefore $v_{70} = \sum_i v_{70,i}$ = 3.167.

We observed two deaths (members 3 and 4), so

$$\hat{\mu}_{70} = \frac{2}{3.167} = 0.6316.$$

(b) $\hat{q}_{70} = 1 - \exp(-\hat{\mu}_{70})$

$$= 1 - \exp(-0.6316) = 1 - 0.5318 = 0.4682.$$

(ii) The contributions to the Poisson likelihood made by each member are proportional to the following

Member	
1	$\exp(-0.25\mu_{70})$
2	$\exp(-0.75\mu_{70})$
3	$\mu_{70}\exp(-0.5\mu_{70})$
4	$\mu_{70}\exp(-0.25\mu_{70})$
5	$\exp(-0.25\mu_{70})$
6	$\exp(-0.3333\mu_{70})$
7	$\exp(-0.5833\mu_{70})$
8	$\exp(-0.25\mu_{70})$

The total likelihood, L, is proportional to the product

$$L \propto [\exp(-3.167\mu_{70})](\mu_{70})^2.$$

Then

$$\log L = -3.167\mu_{70} + 2\log\mu_{70}$$

so that

$$\frac{d\log L}{d\mu_{70}} = -3.167 + \frac{2}{\mu_{70}}.$$

Setting this equal to zero and solving for μ_{70} produces the maximum likelihood estimate,
which is 2/3.167 = 0.6316

Since $\frac{d^2\log L}{d\mu_{70}^2} = -\frac{2}{\mu_{70}^2}$, which is always negative, we definitely have a maximum.

This is the same as the estimate from the two-state model.

(iii) The Poisson model is not an exact model, since it allows for a non-zero probability of more than n deaths in a sample of size n.

The variance of the maximum likelihood estimator for the two-state model is only available asymptotically, whereas that for the Poisson model is available exactly in terms of the true μ.

The two-state model extends to processes with increments, whereas the Poisson model does not.

The Poisson model is a less satisfactory approximation to the multiple state model when transition rates are high.

EXAMINATION

14 September 2005 (am)

Subject CT4 (103) — Models (103 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 7 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 An insurance company has a block of in-force business under which policyholders have been given options and investment-related guarantees. A stochastic model has been developed which projects option and guarantee costs. You have used the model to estimate, for the Company Board, the probability of the insurance company having insufficient assets to honour the payouts under the policies. A Board member has asked whether there are any factors which could cause this probability to be inaccurate.

Outline the items you would mention in your response. [5]

2 (i) In the context of a stochastic process denoted by $\{X_t : t \in J\}$, define:

(a) state space
(b) time set
(c) sample path

[2]

(ii) Stochastic process models can be placed in one of four categories according to whether the state space is continuous or discrete, and whether the time set is continuous or discrete. For each of the four categories:

(a) State a stochastic process model of that type.

(b) Give an example of a problem an actuary may wish to study using a model from that category.

[4]
[Total 6]

3 A die is rolled repeatedly. Consider the following two sequences:

I B_n is the largest number rolled in the first n outcomes.
II C_n is the number of sixes rolled in the first n outcomes.

For each of these two sequences:

(a) Explain why it is a Markov chain.
(b) Determine the state space of the chain.
(c) Derive the transition probabilities.
(d) Explain whether the chain is irreducible and/or aperiodic.
(e) Describe the equilibrium distribution of the chain. [7]

4 A life insurance company prices its long-term sickness policies using a three-state Markov model in continuous time. The states are healthy (H), ill (I) and dead (D). The forces of transition in the model are $\sigma_{HI} = \sigma$, $\sigma_{IH} = \rho$, $\sigma_{HD} = \mu$, $\sigma_{ID} = v$ and they are assumed to be constant over time.

For a group of policyholders observed over a 1-year period, there are:

23 transitions from State H to State I;
15 transitions from State I to State H;
3 deaths from State H;
5 deaths from State I.

The total time spent in State H is 652 years and the total time spent in State I is 44 years.

(i) Write down the likelihood function for these data. [3]

(ii) Derive the maximum likelihood estimate of σ. [2]

(iii) Estimate the standard deviation of $\tilde{\sigma}$, the maximum likelihood estimator of σ. [2]

[Total 7]

5 Claims arrive at an insurance company according to a Poisson process with rate λ per week.

Assume time is expressed in weeks.

(i) Show that, given that there is exactly one claim in the time interval $[t, t + s]$, the time of the claim arrival is uniformly distributed on $[t, t + s]$. [3]

(ii) State the joint density of the holding times $T_0, T_1, \ldots, T_n$ between successive claims. [1]

(iii) Show that, given that there are n claims in the time interval $[0, t]$, the number of claims in the interval $[0, s]$ for $s < t$ is binomial with parameters n and s/t. [3]

[Total 7]

 PLEASE TURN OVER

6 A Markov jump process X_t with state space $S = \{0, 1, 2, \ldots, N\}$ has the following transition rates:

$\sigma_{ii} = -\lambda$ for $0 \le i \le N-1$

$\sigma_{i,i+1} = \lambda$ for $0 \le i \le N-1$

$\sigma_{ij} = 0$ otherwise

(i) Write down the generator matrix and the Kolmogorov forward equations (in component form) associated with this process. [3]

(ii) Verify that for $0 \le i \le N-1$ and for all $j \ge i$, the function

$$p_{ij}(t) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}$$

is a solution to the forward equations in (i). [2]

(iii) Identify the distribution of the holding times associated with the jump process. [2]

[Total 7]

7 A time-inhomogeneous Markov jump process has state space {A, B} and the transition rate for switching between states equals $2t$, regardless of the state currently occupied, where t is time.

The process starts in state A at $t = 0$.

(i) Calculate the probability that the process remains in state A until at least time s. [2]

(ii) Show that the probability that the process is in state B at time T, and that it is in the first visit to state B, is given by $T^2 \times \exp^{-T^2}$. [3]

(iii) (a) Sketch the probability function given in (ii).

(b) Give an explanation of the shape of the probability function.

(c) Calculate the time at which it is most likely that the process is in its first visit to state B.

[6]

[Total 11]

END OF PAPER

EXAMINATION

14 September 2005 (am)

Subject CT4 (104) — Models (104 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 6 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 Describe the advantages and disadvantages of graduating a set of observed mortality rates using a parametric formula. [4]

2 A lecturer at a university gives a course on Survival Models consisting of 8 lectures. 50 students initially register for the course and all attend the first lecture, but as the course proceeds the numbers attending lectures gradually fall.

Some students switch to another course. Others intend to sit the Survival Models examination but simply stop attending lectures because they are so boring. In this university, students who decide not to attend a lecture are not permitted to attend any subsequent lectures.

The table below gives the number of students switching courses and stopping attending lectures after each of the first 7 lectures of the course.

Lecture number	*Number of students switching courses*	*Number of students ceasing to attend lectures but remaining registered for Survival Models*
1	5	1
2	3	0
3	2	3
4	0	1
5	0	2
6	0	1
7	0	0

The university's Teaching Quality Monitoring Service has devised an Index of Lecture Boringness. This index is defined as the Kaplan-Meier estimate of the proportion of students remaining registered for the course who attend the final lecture. In calculating the Index, students who switch courses are to be treated as censored after the last lecture they attend.

(i) Calculate the Index of Lecture Boringness for the Survival Models course. [4]

(ii) Explain whether the censoring in this example is likely to be non-informative. [2]

[Total 6]

3 A mortality investigation has been carried out over the three calendar years, 2002, 2003 and 2004.

The deaths during the period of investigation, θ_x, have been classified by age x at the date of death, where

$$x = \text{calendar year of death} - \text{calendar year of birth.}$$

Censuses of the numbers alive on 1 January in each of the years 2002, 2003, 2004 and 2005 have been tabulated and denoted by

$$P_x(2002),\ P_x(2003),\ P_x(2004) \text{ and } P_x(2005)$$

respectively, where x is the age last birthday at the date of each census.

(i) State the rate year implied by the classification of deaths, and give the ages of the lives at the beginning of the rate year. [2]

(ii) Derive an expression for the exposed to risk in terms of the $P_x(t)$ (t = 2002, 2003, 2004, 2005) which corresponds to the deaths data and which may be used to estimate the force of mortality, μ_{x+f} at age $x + f$. [4]

(iii) Determine the value of f, stating any assumptions you make. [3]

[Total 9]

4 An investigation was carried out into the mortality of male undergraduate students at a large university. The resulting crude rates were graduated graphically. The following table shows the observed numbers of deaths at each age x, d_x, and the $\mathring{q}_x$s obtained from the graduation, together with the number of lives exposed to risk at each age.

Age x	d_x	$\mathring{q}_x$	*Exposed-to-risk*
18	6	0.0012	5,200
19	8	0.0013	5,000
20	12	0.0015	4,800
21	8	0.0017	5,000
22	9	0.0019	3,800
23	6	0.0020	3,600
24	8	0.0021	3,200

(i) Test whether the overall fit of the graduated rates to the crude data is satisfactory using a chi-squared test. [5]

(ii) Comment on your results in (i). [1]

(iii) (a) Describe three possible shortcomings in a graduation which the chi-squared test cannot detect, and

(b) State a test which can be used to detect each one. [3]

[Total 9]

5 An investigation was carried out into the effects of lifestyle factors on the mortality of people aged between 50 and 65 years. The investigation took the form of a prospective study following a sample of several hundred individuals from their 50th birthdays until their 65th birthdays and collecting data on the following covariates for each person:

X_1 Sex (a categorical variable with 0 = female, 1 = male)

X_2 Cigarette smoking (a categorical variable with 0 = non-smoker, 1 = smoker)

X_3 Alcohol consumption (a categorical variable with 0 = consumes fewer than 21 units of alcohol per week, 1 = consumes 21 or more units of alcohol per week)

In addition, data were collected on the age at death for persons who died during the period of investigation.

In order to analyse the data, it was decided to use a Gompertz hazard, $\lambda_x = Bc^x$, where x is the duration since the start of the observation.

(i) Explain why the Gompertz hazard might be appropriate for analysing the mortality of persons aged between 50 and 65 years. [2]

(ii) Show that the substitution:

$$B = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3),$$

in the Gompertz model (where β_0 ... β_3 are parameters to be estimated), leads to a proportional hazards model for this particular analysis. [3]

(iii) Using the Gompertz hazard, the parameter estimates in the proportional hazards model were as follows:

Covariate	*Parameter estimate*	*Parameter*
Sex	β_1	+0.40
Cigarette smoking	β_2	+0.75
Alcohol consumption	β_3	–0.20
	β_0	–5.00
	c	+1.10

(a) Describe the characteristics of the person to whom the baseline hazard applies in this model.

(b) Calculate the estimated hazard for a female cigarette smoker aged 55 years who does not consume alcohol.

(c) Show that, according to this model, a cigarette smoker at any age has a risk of death roughly equal to that of a non-smoker aged eight years older. [6]

[Total 11]

6 Studies of the lifetimes of a certain type of electric light bulb have shown that the probability of failure, q_0, during the first day of use is 0.05 and after the first day of use the "force of failure", μ_x, is constant at 0.01.

(i) Calculate the probability that a light bulb will fail within the first 20 days. [2]

(ii) Calculate the complete expectation of life (in days) of:

(a) a one-day old light bulb
(b) a new light bulb

[7]

(iii) Comment on the difference between the complete expectations of life calculated in (ii) (a) and (b). [2]

[Total 11]

END OF PAPER

EXAMINATION

September 2005

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

15 November 2005

EXAMINERS' COMMENTS

Comments on solutions presented to individual questions for this September 2005 paper are given below:

103 Part

Question A1 *This was not well answered.*
There was a lot of repetition in some of the solutions offered - for example several different instances of parameter error may have been mentioned.

Question A2 *This was well answered overall, even by the weaker candidates.*
Credit was not given in part (ii)(b) if the examples cited were not likely to be encountered by an actuary working in a professional capacity.

Question A3 *This was well answered overall.*
Some candidates lost marks by not explaining why the chains were not irreducible and were aperiodic. Many candidates did not correctly identify the state space of the chain C_n and most did not realise that the chain will escape to infinity as the value increases without barrier.

Question A4 *This was very well answered overall, with the majority of candidates scoring highly.*
One common mistake was the omission of the constant term from the likelihood function in part (i).

Question A5 *This was very poorly answered by all but a few candidates.*
Some candidates offered general explanations in parts (i) and (iii), which, if clear enough, were given some credit.

Question A6 *Overall this was not well answered.*
In part (i), few candidates gave the full, correct Kolmogorov equations.
Many candidates lost marks in part (ii) because of insufficient or inaccurate working.

Question A7 *Overall this was not well answered.*
However, part (i) was well answered. Some candidates reached the correct answer via a different solution and received full credit.
Many candidates struggled with part (ii), failing to identify the correct integrand required.
In part (iii), many candidates described the shape of the function, but few explained it, as required by the question.

104 Part

Question B1 *This was not well answered.*
Some candidates commented on the advantages/disadvantages of graduation in general, rather than concentrating on the parametric formula method.

Question B2 *Part (i) was well answered.*
In part (ii), many candidates clearly did not understand the meaning of non-informative censoring.

Question B3 *This was well answered overall.*
In part (ii), the question asked candidates to "derive an expression" and therefore we were looking for clearly set out steps here. Many candidates lost marks by not providing sufficient explanation of their working.

Question B4 *This was very well answered, even by the weaker candidates.*
The main areas where candidates lost marks were: not stating the null hypothesis, or not stating it clearly enough; failure to identify the correct degrees of freedom to be used in the test; and insufficient or insufficiently clear descriptions of the shortcomings.
In part (iii), the majority of candidates seemed confused between two issues in connection with bias. There are two distinct problems. Firstly, if the consistent bias is only small, the chi-squared test may fail to detect it because the resulting number (i.e. the sum of the squared deviations) is not large enough to exceed the critical value. The signs test, which ignores the magnitude of the bias and looks only at how consistent it is across the ages, can be used to identify this. The second problem is that even if the consistent bias is larger and the chi-squared test leads us to reject the null hypothesis, the test gives no indication of whether the graduated rates are too high or too low. This is because the deviations are squared and the test statistic always positive. The signs test is <u>not</u> a solution to this second problem.

Question B5 *This was well answered overall.*
Some parts of the question required candidates to "show" a result; candidates lost marks if their working was not sufficiently clear or complete.

Question B6 *This was not well answered.*
Surprisingly few candidates correctly answered part (i).
In parts (ii) and (iii), very few candidates recognised that the expectation of life was an average of the future lifetimes of those bulbs still shining. As a result, although many candidates correctly calculated the expectation of life for a one-day old bulb, few managed to do so for a new bulb. In part (iii), most candidates commented on the higher force of failure in the first day.

103 Part

A1

Items to be mentioned include:

- Models will be chosen which it is felt give a reasonable reflection of the underlying real world processes, but this may not turn out to be the case. (Model error.)

- The model may be very sensitive to parameters chosen, and the parameters are estimates because the true underlying parameters cannot be observed. (Parameter error.)

- Sampling error may result from running insufficient simulations. (It should be possible to give a confidence interval for the error that could result from this source.)

- The management actions assumed may not match what would happen in extreme circumstances.

- Policyholder behaviour, such as take-up rates for options, may differ in practice.

- There may be future events, such as legislative changes which affect the interpretation of the policy conditions, which have not been anticipated in the modelling.

- There may be errors in the coding of the model. The model is likely to be complex and difficult to verify completely.

- The model relies on input data, which may be grouped rather than being able to run every policy. Any errors in the data could cause the output to be inaccurate.

A2

(i) (a) The state space is the set of values which it is possible for each random variable X_t to take.

(b) The time set is the set J, the times at which the process contains a random variable X_t.

(c) A sample path is a joint realisation of the variables X_t for all t in J, that is a set of values for X_t (at each time in the time set) calculated using the previous values for X_t in the sample path.

(ii) Discrete State Space, Discrete Time

(a) Simple random walk, Markov chain, or any other suitable example

(b) Any reasonable example. For example: No Claims Discount systems, Credit Rating at end of each year

Discrete State Space, Continuous Time

(a) Poisson process, Markov jump process, for example

(b) Any reasonable example. For example: Claims received by an insurer, Status of pension scheme member

Continuous State Space, Discrete Time

(a) General random walk, time series, for example

(b) Any reasonable example. For example: Share prices at end of each trading day, Inflation index

Continuous State Space, Continuous Time

(a) Brownian motion, diffusion or Itô process, for example.
Compound Poisson process if the defined state space is continuous.

(b) Any reasonable example. For example: Share prices during trading period, Value of claims received by insurer

A3 (a) Given the current state (the largest outcome or the number of sixes) up to the nth roll, no additional information is required to predict the status of the chain after the next roll. Therefore both B_n and C_n have the Markov property.

(b) B_n has state space {1, 2, 3, 4, 5, 6},
the state space for C_n is the set of non-negative integers.

(c) For B_n, and $1 \le i, j \le 6$,

$$P\left(B_{n+1} = j \mid B_n = i\right) = \frac{i}{6} \quad \text{for } j = i,$$

$$P\left(B_{n+1} = j \mid B_n = i\right) = \frac{1}{6} \quad \text{for each } j > i$$

and $$P\left(B_{n+1} = j \mid B_n = i\right) = 0 \quad \text{for } i > j$$

For C_n, and for $k = 0,1,2,\ldots,$

$$P\left(C_{n+1} = k+1 \mid C_n = k\right) = \frac{1}{6},$$

$$P\left(C_{n+1} = k \mid C_n = k\right) = \frac{5}{6},$$

and $$P\left(C_{n+1} = j \mid C_n = k\right) = 0 \quad \text{for all other } j \neq k, k+1$$

(d) The chain B_n is clearly aperiodic; if currently at state i, it can remain there if the next outcome is at most i.
It is *not* irreducible, as it cannot be reached from j for $i < j$.

C_n is again aperiodic; if currently at state i, it can remain there if the next outcome is not a 6.
It is not irreducible; state k cannot be reached from m if $k < m$.

(e) In the long run, B_n will reach state 6 and will remain there; hence in equilibrium $P(B_n = 6) = 1$ for sufficiently large n.

C_n cannot decrease and has an infinite state space; therefore, it is certain that it will escape to infinity with probability one.

A4 (i) The likelihood is

$$L = K \times \exp(-652(\sigma + \mu))\exp(-44(\rho + \upsilon))\sigma^{23}\rho^{15}\mu^{3}\upsilon^{5}$$

(ii) $l = \ln L = -652\sigma + 23 \ln \sigma$ + constant with respect to σ

Differentiating with respect to σ gives

$$\frac{\partial l}{\partial \sigma} = -652 + \frac{23}{\sigma}$$

and setting equal to zero gives

$$0 = -652 + \frac{23}{\hat{\sigma}}$$

$$\Rightarrow \hat{\sigma} = \frac{23}{652} = 0.0353 \text{ p.a.}$$

Differentiating again gives

$$\frac{\partial^2 l}{\partial \sigma^2} = -\frac{23}{\sigma^2} < 0$$

therefore $\hat{\sigma}$ is the maximum likelihood estimate

(iii) The variance of $\tilde{\sigma}$ is $-\left(\dfrac{\partial^2 l}{\partial \sigma^2}\right)^{-1} = \dfrac{\sigma^2}{23}$,

which we can estimate by $\dfrac{\hat{\sigma}^2}{23}$.

Therefore the estimated standard deviation of $\tilde{\sigma}$ is $\dfrac{\hat{\sigma}}{\sqrt{23}} = 0.00736$.

A5 (i) Let N_t denote the number of claims up to time t. Since the Poisson process has stationary increments, we may take $t = 0$, so that the required conditional distribution is

$$P\left(T_0 \le y \mid N_s = 1\right) = \frac{P\left(T_0 \le y, \quad N_s = 1\right)}{P\left(N_s = 1\right)}$$

$$= \frac{P\left(\ N_y = 1, \quad N_s - N_y = 0\right)}{P\left(N_s = 1\right)}$$

But $N_s - N_y$ is independent of N_y
and has the same distribution as N_{s-y}.

Thus the right hand side above equals

$$\frac{(\lambda y e^{-\lambda y}) e^{-\lambda(s-y)}}{\lambda s e^{-\lambda s}} = \frac{y}{s},$$

which is the cdf of the uniform distribution on [0, s].

(ii) Since holding times are independent, each having an exponential distribution, their joint density is

$$\lambda^n e^{-\lambda\left(t_1 + t_2 + \ldots + t_n\right)} 1_{\left\{t_1, t_2, \ldots, t_n > 0\right\}}.$$

(iii) We have, as in part (i),

$$P\left(N_s = k \mid N_t = n\right) = \frac{P\left(N_s = k, \quad N_t = n\right)}{P\left(N_t = n\right)}$$

$$= \frac{P\left(\ N_s = k, \quad N_t - N_s = n - k\right)}{P\left(N_t = n\right)}$$

Using again that the Poisson process has stationary and independent increments, and that the number of claims in an interval [0, t] is Poisson (λt), we derive from above that

$$P\left(N_s = k \mid N_t = n\right) = \frac{\dfrac{e^{-\lambda s}(\lambda s)^k}{k!} \cdot \dfrac{e^{-\lambda(t-s)}\lambda^{n-k}(t-s)^{n-k}}{(n-k)!}}{\dfrac{e^{-\lambda t}(\lambda t)^n}{n!}}$$

$$= \frac{e^{-\lambda t}\lambda^n s^k (t-s)^{n-k}}{k!(n-k)!} \cdot \frac{n!}{e^{-\lambda t}\lambda^n t^n}$$

$$= \frac{n!}{k!(n-k)!} \cdot \frac{s^k (t-s)^{n-k}}{t^k t^{n-k}}$$

$$= \binom{n}{k}\left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}$$

which is binomial with parameters *n* and *s/t*.

A6 (i) The generator matrix is

$$A = \begin{pmatrix} -\lambda & \lambda & & & & & 0 \\ & -\lambda & \lambda & & & & \\ & & \cdot & \cdot & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ & & & & & \lambda & \\ & & & & & -\lambda & \lambda \\ 0 & & & & & & 0 \end{pmatrix},$$

all other entries being zero

The Kolmogorov equations are $P'(t) = P(t)A$.

In a component form the forward equations read

$$p'_{ii}(t) = -\lambda p_{ii}(t) \qquad \text{for } 0 \le i \le N-1$$

$$p'_{ij}(t) = -\lambda p_{ij}(t) + \lambda p_{i,j-1}(t) \qquad \text{for } i < j < N$$

$$p'_{iN}(t) = \lambda p_{i,N-1}(t).$$

(ii) Differentiating the function given in the question, we get first for $i = j$,

$$p'_{ii}(t) = -\lambda e^{-\lambda t},$$

while for $i < j \le N$,

$$p'_{ij}(t) = -\lambda e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} + \lambda e^{-\lambda t}\frac{(\lambda t)^{j-i-1}}{(j-i-1)!}$$

We can then check that the above satisfy the forward equations.

(iii) For $i = j(<N)$, the solution in (ii) implies that $p_{ii}(t) = e^{-\lambda t}$, so that the distribution of the holding times $T_0, T_1, ..., T_{N-1}$ is exponential with parameter λ.

For $i = N$, this is obviously not true; once the chain reaches state N, it stays there forever.

A7 (i) $$\frac{d}{dt}P_{\overline{AA}}(t) = -2t \times P_{\overline{AA}}(t)$$

$$\Rightarrow \frac{d}{dt}\left[\ln P_{\overline{AA}}(t)\right] = -2t$$

$$\Rightarrow \ln P_{\overline{AA}}(s) = -s^2 + \text{constant}$$

We know $P_{\overline{AA}}(0) = 1$, hence constant $= 0$

Hence, $P_{\overline{AA}}(s) = \exp^{-s^2}$

(ii) P(in first visit to B at time $T \mid$ in state A at t = 0)

$$= \int_0^T P(\text{remains in A to time } s)$$
$$\times P(\text{transition to B in time } s, s + ds)$$
$$\times P(\text{remains in B to time } T)\, ds$$

$$= \int_{s=0}^{T} P_{\overline{AA}}(s) \times 2s \times P_{\overline{BB}}(s,T)ds$$

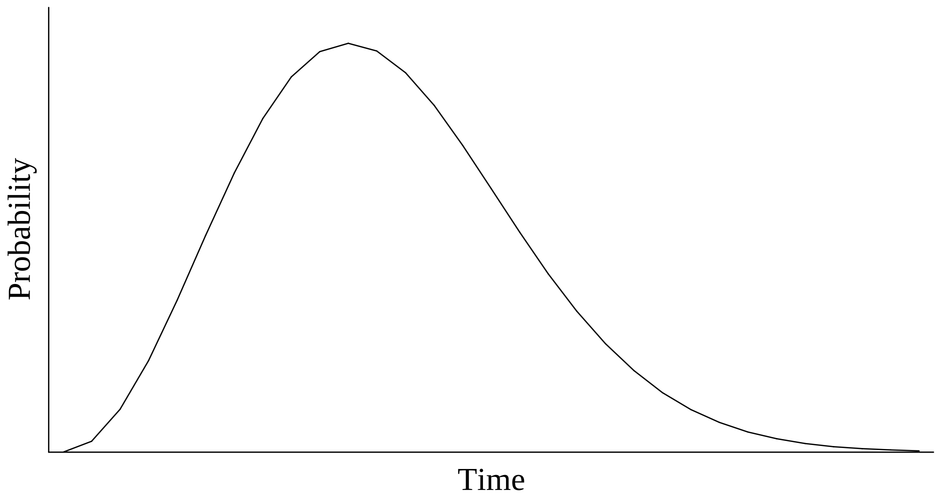
Using the result from part (i) and the similar result for P_{BB} with boundary condition $P_{BB}(s, s) = 1$, this gives us:

$$= \int_{s=0}^{T} e^{-s^2} \times 2s \times e^{-T^2+s^2} ds$$

$$= \int_{s=0}^{T} 2s \times e^{-T^2} ds$$

$$= e^{-T^2} \times T^2$$

(iii) (a) The sketch should be shaped like:



(b) Commentary:

- Initially probability increases from 0 at T = 0, and accelerates as the transition rate from A to B increases.

- However, as transitions increase, it becomes more likely that the process has already visited state B and jumped back to A. Therefore the probability of being in the first visit to B tends (exponentially) to zero.

(c) Differentiate to find turning point:

$$\frac{d}{dt}\left[e^{-t^2} \times t^2\right] = 2t \times e^{-t^2} - 2t^3 \times e^{-t^2}$$

set derivative equal to zero

$$e^{-t^2} \times 2t \times (1 - t^2) = 0$$

implies t = 1 for a positive solution
and, from above analysis, this is clearly a maximum.

104 Part

B1 *Advantages:*

The graduated rates will progress smoothly provided the number of parameters is small.

Good for producing standard tables.

Can easily be extended to more complex formulae, provided optimisation can be achieved.

Can fit the same formula to different experiences and compare parameter values to highlight differences between them.

Disadvantages:

It can be hard to find a formula to fit well at all ages without having lots of parameters.

Care is required when extrapolating: the fit is bound to be best at ages where we have lots of data, and can often be poor at extreme ages.

B2 (i) The table below gives the relevant calculations.

Lecture j	n_j	d_j	c_j	λ_j	$1-\lambda_j$	$S(j)$
1	50	1	5	1/50	49/50	0.980
2	44	0	3	0	1	0.980
3	41	3	2	3/41	38/41	0.908
4	36	1	0	1/36	35/36	0.883
5	35	2	0	2/35	33/35	0.833
6	33	1	0	1/33	32/33	0.807
7	32	0	0	0	1	0.807
8	32					

The Index of Lecture Boringness is therefore equal to 0.807.

(ii) Censoring in this case is unlikely to be non-informative.

This is because the students who switched courses were probably less interested in the subject matter of Survival Models than those who remained registered.

Therefore they would have been more likely, had they not switched courses, to cease attending lectures than those who did not switch.

B3 (i) The classification of deaths implies a calendar year rate interval.

A person who dies will be aged x on the birthday in the calendar year of death, which implies that he or she will be aged x next birthday on 1 January in the calendar year of death.

Since 1 January is the start of the rate interval, the age range at the start is $x - 1$ to x.

(ii) A census of those aged x next birthday on 1 January in each year would correspond to the classification of deaths.

But we have lives classified by age x last birthday.

However, the number alive aged x next birthday on any date is equal to the number alive aged $x - 1$ last birthday.

The number alive aged $x - 1$ last birthday on 1 January in year t is given by $P_{x-1}(t)$.

At the end of year t this cohort will be aged x last birthday.

Thus, using the trapezium rule, the correct exposed to risk at age x in year t is given by

$$\frac{1}{2}\left[P_{x-1}(t) + P_x(t+1)\right].$$

Over the three calendar years 2002, 2003 and 2004, we have, therefore, exposed to risk =

$$\frac{1}{2}\left[P_{x-1}(2002) + P_x(2003)\right]$$
$$+\frac{1}{2}\left[P_{x-1}(2003) + P_x(2004)\right]$$
$$+\frac{1}{2}\left[P_{x-1}(2004) + P_x(2005)\right].$$

(iii) Assuming birthdays are uniformly distributed over the calendar year, the average age at the start of the rate interval will be x – ½.

Therefore the average age in the middle of the rate interval is x.

Assuming a constant force of mortality between x – ½ and x + ½, therefore, $f = 0$.

B4 (i) The null hypothesis is that the observed data come from a population in which the graduated rates are the true rates.

The chi-squared statistic is given by the formula:

$$\sum_x \frac{(d_x - E_x \mathring{q}_x)^2}{E_x \mathring{q}_x}.$$

The calculations are shown in the table below.

Age	$E_x q_x$	$E_x \mathring{q}_x$	$(E_x q_x - E_x \mathring{q}_x)^2$	$\dfrac{(E_x q_x - E_x \mathring{q}_x)^2}{E_x \mathring{q}_x}$
18	6	6.24	0.0576	0.0092
19	8	6.50	2.2500	0.3461
20	12	7.20	23.0400	3.2000
21	8	8.50	0.2500	0.0294
22	9	7.22	3.1684	0.4388
23	6	7.20	1.4400	0.2000
24	8	6.72	1.6384	0.2438

Therefore the calculated chi-squared value is

0.0092 + 0.3461 + 3.2000 + 0.0294 + 0.4388 + 0.2000 + 0.2438 = 4.4673

Since we have 7 ages, we compare this with the tabulated value at the 5% level at, say, 4 degrees of freedom (since we lose 2–3 degrees for every 10 ages graduated graphically).

The tabulated value with 4 degrees of freedom is 9.488.

Since 4.4673 < 9.488 we have no evidence to reject the null hypothesis.

(ii) On the basis of the chi-squared test, the graphical graduation adheres to the data satisfactorily.

However, there is a large deviation at age 20 which requires further investigation.

(iii) Possible shortcomings, and the relevant tests are:

There may be long runs of deviations of the same sign caused by undergraduation.
These can be detected by the grouping of signs test or the serial correlations test.

There may be one or two large deviations at particular ages, balanced by lots of small deviations (as in the example in part (i))
These can be detected by the individual standardised deviations test.

The graduated rates may be too high or too low over the whole of the age range, but by an amount too small for the chi-squared test to detect.
The signs test or the cumulative deviations test will detect this.

The results of the graduation may not be smooth.
This can be detected by looking at the third order differences of the graduated rates $\mathring{q}_x$. If the rates are smooth, these should be small in magnitude compared with the quantities themselves and should progress regularly.

B5 (i) Taking logarithms of the Gompertz hazard produces

$$\log \lambda_x = \log B + x \log c$$

which indicates that the rate of increase of the hazard with age is constant.

Empirically, this is often a reasonable assumption for middle ages and older ages, which include the age range 50–65 years.

(ii) Putting $B = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)$ into the Gompertz model produces

$$\lambda_x = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3) \,.\, c^x,$$

defining x as duration since 50th birthday.

The hazard can therefore be factorised into two parts:

$\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)$, which depends only on the values of the covariates, and

c^x, which depends only on duration.

Therefore the ratio between the hazards for any two persons with different characteristics does not depend on duration, and so the model is a proportional hazards model.

(iii) (a) The baseline hazard in this model relates to

a female,
non-smoker,
who drinks less than 21 units of alcohol per week.

(b) For a female cigarette smoker who does not consume alcohol we have
$X_1 = 0$, $X_2 = 1$, $X_3 = 0$ and $x = 5$.

Therefore the hazard is given by

$$\lambda_5 = \exp(\beta_0 + \beta_1 .0 + \beta_2 .1 + \beta_3 .0) . c^5$$
$$= \exp(-5 + 0.75) \times 1.10^5$$
$$= 0.0230.$$

(c) The hazard for a non-smoker at duration u is given by the formula

$$\lambda_u = \exp(\beta_0 + \beta_1 X_1 + \beta_3 X_3) . c^u,$$

The hazard for a smoker at duration v is given by the formula

$$\lambda^*_v = \exp(\beta_0 + \beta_1 X_1 + 0.75 + \beta_3 X_3) . c^v.$$

If the smoker's and non-smoker's hazards are the same, then
$$\lambda_u = \lambda^*_v,$$

which implies that
$$\exp(\beta_0 + \beta_1 X_1 + \beta_3 X_3).c^u$$
$$= \exp(\beta_0 + \beta_1 X_1 + 0.75 + \beta_3 X_3) . c^v.$$

which simplifies to
$$c^u = \exp(0.75) . c^v,$$

so that
$$c^u/c^v = c^{u-v} = \exp(0.75) = 2.117.$$

Since $c = 1.1$, we have
$$1.1^{u-v} = 2.117.$$

Therefore

$$u - v = \log(2.117)/\log(1.1)$$
$$= 0.75/0.0953 = 7.87.$$

So when the two hazards are equal, the non-smoker is approximately eight years older than the smoker.

Alternatively this could be demonstrated by calculating λ_u and λ^*_{u-8} and showing that they are approximately the same.

B6 (i) Let the probability of failure within the first 20 days be ${}_{20}q_0$.

We have:

$$\begin{aligned}{}_{20}q_0 &= 1 - {}_{20}p_0 = 1 - {}_1p_0 \cdot {}_{19}p_1 \\ &= 1 - (1 - {}_1q_0)\exp(-19\mu) \\ &= 1 - 0.95\exp(-19 \times 0.01) \\ &= 1 - 0.95\exp(-0.19) \\ &= 1 - 0.95(0.82696)\end{aligned}$$

which is 0.21439.

(ii) (a) The complete expectation of life of a one-day old light bulb, $\mathring{e}_1$ is given by

$$\begin{aligned}\mathring{e}_1 &= \int_0^\infty {}_tp_1 dt \\ &= \int_0^\infty e^{-0.01t} dt\end{aligned}$$

Integrating, this gives

$$\mathring{e}_1 = -\frac{1}{0.01}\left[e^{-0.01t}\right]_0^\infty = -\frac{1}{0.01}[0 - 1]$$

= 100 days.

(b) The complete expectation of life of a new light bulb, $\mathring{e}_0$ is given by

$$\mathring{e}_0 = \int_0^\infty {}_tp_0 dt = \int_0^1 {}_tp_0 dt + \int_1^\infty {}_tp_0 dt. \quad (*)$$

Alternative 1

Assume a uniform distribution of failure times between exact ages 0 and 1,

the first term in (*) is equal to

$$\frac{1}{2}\left(1+{}_1p_0\right)$$

$$=\frac{1}{2}\left[1+(1-{}_1q_0)\right]$$

$$=\frac{1}{2}(1+0.95)=0.975$$

The second term is equal to

$${}_1p_0\int_0^\infty {}_tp_1 dt = 0.95(100)$$

(using the result from part (i) above).

Therefore:

$$\mathring{e}_0 = 0.975+100\times 0.95 = 95.975 \text{ days.}$$

Alternative 2

Assume a constant force of failure between exact ages 0 and 1

Let this constant force be δ.

Then

$${}_1p_0 = \exp\left[-\int_0^1 \delta ds\right] = \exp(-\delta)$$

$$=1-{}_1q_0 = 0.95.$$

So that

$$\exp(-\delta)=0.95$$

and

$$\delta = -\log(0.95) = 0.0513.$$

Thus the first term on the right-hand side of (*) is

$$\int_0^1 {}_t p_0 dt = \int_0^1 \exp(-0.0513t)dt$$

$$= \frac{1}{-0.0513}\left[\exp(-0.0513t)\right]_0^1$$

$$= \frac{1}{-0.0513}\left[\exp(-0.0513) - 1\right]$$

$$= 0.97478,$$

and the second term is equal to

$${}_1 p_0 \int_0^\infty {}_t p_1 dt = 0.95(100)$$

(using the result from part (i) above).

So that

$$\overset{\circ}{e}_0 = 0.97478 + 100 \times 0.95 = 95.97478 \text{ days.}$$

(iii) The complete expectation of life of a light bulb at any age is an average of the future lifetimes of all bulbs which have not failed before that age.

The value of $\overset{\circ}{e}_0$ is lower than $\overset{\circ}{e}_1$ because the average $\overset{\circ}{e}_0$ includes the very short lifetimes of the relatively large proportion of bulbs which fail in the first day, which deflate the average, whereas $\overset{\circ}{e}_1$ excludes these.

END OF EXAMINERS' REPORT

EXAMINATION

29 March 2006 (am)

Subject CT4 (103) — Models (103 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*
2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*
3. *Mark allocations are shown in brackets.*
4. *Attempt all 6 questions, beginning your answer to each question on a separate sheet.*
5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

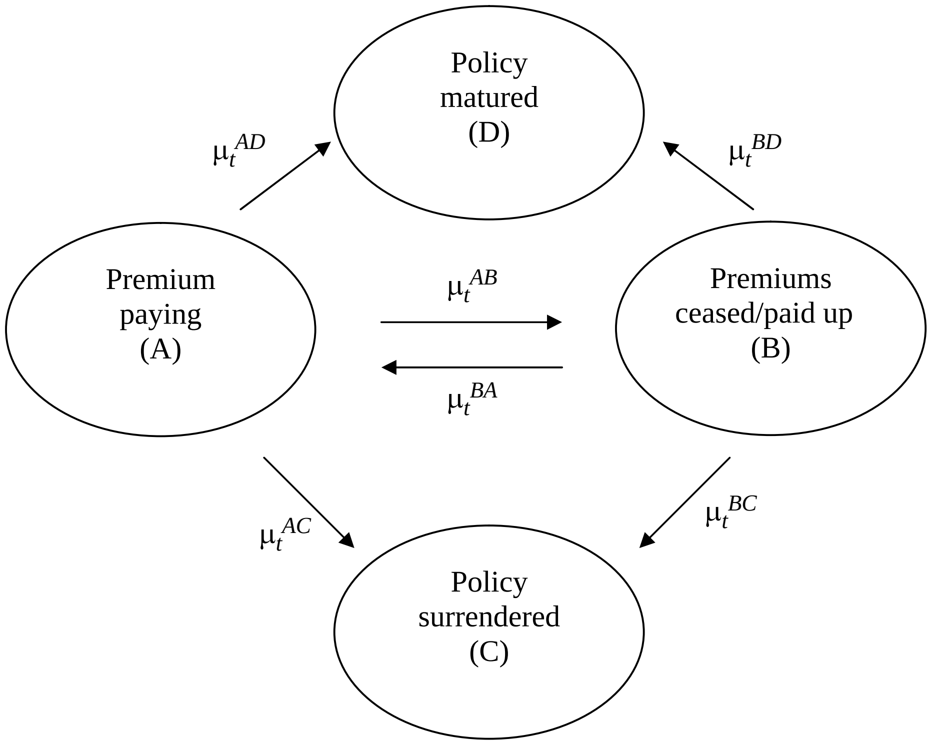
In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

A1 In the context of a stochastic process $\{X_t : t \in J\}$, explain the meaning of the following conditions:

(a) strict stationarity
(b) weak stationarity

[3]

A2 A savings provider offers a regular premium pension contract, under which the customer is able to cease paying in premiums and restart them at a later date. In order to profit test the product, the provider set up the four-state Markov model shown in the following diagram:



Show, from first principles, that under this model:

$$\frac{\partial}{\partial t}\,{}_t p_0^{AB} = {}_t p_0^{AA}.\mu_t^{AB} - {}_t p_0^{AB}.(\mu_t^{BA} + \mu_t^{BC} + \mu_t^{BD})$$ [5]

A3 A motor insurer's No Claims Discount system uses the following levels of discount {0%, 25%, 40%, 50%}. Following a claim free year a policyholder moves up one discount level (or remains on 50% discount). If the policyholder makes one (or more) claims in a year they move down one level (or remain at 0% discount).

The insurer estimates that the probability of making at least one claim in a year is 0.1 if the policyholder made no claims the previous year, and 0.25 if they made a claim the previous year.

New policyholders should be ignored.

(i) Explain why the system with state space {0%, 25%, 40%, 50%} does not form a Markov chain. [2]

(ii) (a) Show how a Markov chain can be constructed by the introduction of additional states.

(b) Write down the transition matrix for this expanded system, or draw its transition diagram.

[4]

(iii) Comment on the appropriateness of the current No Claims Discount system.

[2]

[Total 8]

A4 (i) List the benefits of modelling in actuarial work. [2]

(ii) Describe the difference between a stochastic and a deterministic model. [2]

(iii) Outline the factors you would consider in deciding whether to use a stochastic or deterministic model to study a problem. [3]

(iv) Explain how a deterministic model might be used to validate model outcomes where a stochastic approach has been selected. [2]

[Total 9]

A5 Employees of a company are given a performance appraisal each year. The appraisal results in each employee's performance being rated as High (H), Medium (M) or Low (L). From evidence using previous data it is believed that the performance rating of an employee evolves as a Markov chain with transition matrix:

$$P = \begin{array}{c} \\ H \\ M \\ L \end{array} \begin{array}{c} \begin{array}{ccc} \quad H \quad & \quad M \quad & \quad L \quad \end{array} \\ \begin{pmatrix} 1-\alpha-\alpha^2 & \alpha & \alpha^2 \\ \alpha & 1-2\alpha & \alpha \\ \alpha^2 & \alpha & 1-\alpha-\alpha^2 \end{pmatrix} \end{array}$$

for some parameter α.

(i) Draw the transition graph of the chain. [2]

(ii) Determine the range of values for α for which the matrix P is a valid transition matrix. [2]

(iii) Explain whether the chain is irreducible and/or aperiodic. [2]

(iv) For $\alpha = 0.2$, calculate the proportion of employees who, in the long run, are in state L. [3]

(v) Given that $\alpha = 0.2$, calculate the probability that an employee's rating in the third year, X_3, is L:

(a) in the case that the employee's rating in the first year, X_1, is H
(b) in the case $X_1 = M$
(c) in the case $X_1 = L$

[2]
[Total 11]

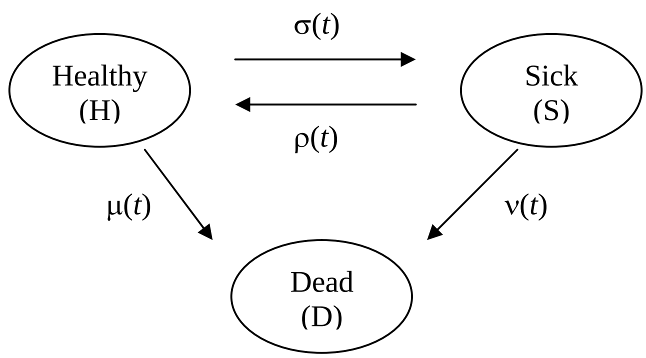
A6 (i) (a) Explain what is meant by a Markov jump process.

(b) Explain the condition needed for such a process to be time-homogeneous.

[2]

(ii) Outline the principal difficulties in fitting a Markov jump process model with time-inhomogeneous rates. [2]

A company provides sick pay for a maximum period of six months to its employees who are unable to work. The following three-state, time-inhomogeneous Markov jump process has been chosen to model future sick pay costs for an individual:



Where "Sick" means unable to work and "Healthy" means fit to work.

The time dependence of the transition rates is to reflect increased mortality and morbidity rates as an employee gets older. Time is expressed in years.

(iii) Write down Kolmorgorov's forward equations for this process, specifying the appropriate transition matrix. [1]

(iv) (a) Given an employee is sick at time $w < T$, write down an expression for the probability that he or she is sick throughout the period $w < t < T$.

(b) Given that a transition out of state H occurred at time w, state the probability that the transition was into state S.

(c) For an employee who is healthy at time τ, give an approximate expression for the probability that there is a transition out of state H in a small time interval $[w, w + dw]$, where $w > \tau$. Your expression should be in terms of the transition rates and $P_{HH}(\tau, w)$ only.

[3]

(v) Using the results of part (iv) or otherwise, derive an expression for the probability that an employee is sick at time T and has been sick for less than 6 months, given that they were healthy at time $\tau < T - ½$. Your expression should be in terms of the transition rates and $P_{HH}(\tau, w)$ only. [3]

 PLEASE TURN OVER

(vi) Comment on the suggestions that:

(a) $\rho(t)$ should also depend on the holding time in state S, and
(b) mortality rates can be ignored.

[3]
[Total 14]

END OF PAPER

EXAMINATION

29 March 2006 (am)

Subject CT4 (104) — Models (104 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*
2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*
3. *Mark allocations are shown in brackets.*
4. *Attempt all 6 questions, beginning your answer to each question on a separate sheet.*
5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

B1 A Cox proportional hazards model was estimated to assess the effect on survival of a person's sex and his or her self-esteem (measured on a three-point scale as "low", "medium" or "high"). The baseline category was males with "low" self-esteem.

Write down the equation of the model, using algebraic symbols to represent variables and parameters and defining all the symbols that you use. [4]

B2 (i) (a) Explain why it is important to sub-divide data when carrying out mortality investigations.

(b) Describe the problems that can arise with sub-dividing data.

[4]

(ii) List four factors which are often used to sub-divide life assurance data. [2]

[Total 6]

B3 (i) Assume that the force of mortality between consecutive integer ages, y and $y + 1$, is constant and takes the value μ_y.

Let T_x be the future lifetime after age x ($x \le y$) and $S_x(t)$ be the survival function of T_x.

Show that:

$$\mu_y = \log[S_x(y-x)] - \log[S_x(y+1-x)].$$ [4]

(ii) An investigation was carried out into the mortality of male life office policyholders. Each life was observed from his 50th birthday until the first of three possible events occurred: his 55th birthday, his death, or the lapsing of his policy. For those policyholders who died or allowed their policies to lapse, the exact age at exit was recorded.

Using the result from part (i) or otherwise, describe how the data arising from this investigation could be used to estimate:

(a) μ_{50}

(b) ${}_5q_{50}$

[3]

[Total 7]

B4 A company is interested in estimating policy lapse rates by age. It conducts an investigation into this, which lasts for the whole of the calendar year 2003. The investigation collects the following data for a sample of policies which are funded by annual premiums:

- the age last birthday of the policyholder when the policy was taken out;
- the number of premiums the policyholder paid before the policy lapsed.

In addition, the number of policies in-force on 1 January each year is available, classified by age x last birthday and years t elapsed since 1 January 2003, $(P_{x,t}^{\ *})$.

(i) State the rate interval in this investigation. [1]

(ii) Derive an expression for the exposed-to-risk in terms of $P_{x,t}^{\ *}$, stating any assumptions you make. [7]

(iii) Comment on the reasonableness or otherwise of the assumptions you made in your answer to part (ii). [2]

[Total 10]

B5 A life assurance company carried out an investigation of the mortality of male life assurance policyholders. The investigation followed a group of 100 policyholders from their 60th birthday until their 65th birthday, or until they died or cancelled their policy (whichever event occurred first).

The ages at which policyholders died or cancelled their policies were as follows:

Died	**Cancelled policy**
Age in years and months	*Age in years and months*
60y 5m	60y 2m
61y 1m	60y 3m
62y 6m	60y 8m
63y 0m	61y 0m
63y 0m	61y 0m
63y 8m	61y 0m
64y 3m	61y 5m
	62y 2m
	62y 9m
	63y 9m
	64y 5m

(i) Explain which types of censoring are present in the investigation. [2]

(ii) Calculate the Nelson-Aalen estimate of the integrated hazard for these policyholders. [5]

(iii) Sketch the estimated integrated hazard function. [2]

(iv) Estimate the probability that a policyholder will survive to age 65. [2]
[Total 11]

B6 An investigation was undertaken into the mortality of male term assurance policyholders for a large life insurance company. The crude mortality rates were graduated using a formula of the form:

$$\overset{\circ}{q}_x = \alpha + \beta e^{\gamma x}$$

An extract of the results is shown below.

Age	*Exposure (years)*	*Crude mortality rate*	*Graduated mortality rate*	*Standardised deviation*
x	E_x	$\hat{q}_x$	$\overset{\circ}{q}_x$	$z_x = \dfrac{E_x\left(\hat{q}_x - \overset{\circ}{q}_x\right)}{\sqrt{E_x \overset{\circ}{q}_x \left(1 - \overset{\circ}{q}_x\right)}}$
40	11,037	0.0029	0.00348	-1.035
41	12,010	0.00333	0.00358	-0.459
42	11,654	0.003	0.00368	-1.212
43	9,658	0.003	0.00379	-1.264
44	8,457	0.00319	0.00391	-1.061
45	10,541	0.00427	0.00402	0.406
46	7,410	0.00472	0.00415	0.763
47	12,042	0.00399	0.00428	-0.487
48	14,038	0.00406	0.00441	-0.626
49	11,479	0.00375	0.00455	-1.274
50	12,480	0.00409	0.00469	-0.981
51	10,567	0.00407	0.00485	-1.154
52	9,187	0.00512	0.00500	0.163
53	14,027	0.00456	0.00517	-1.007
54	11,581	0.00466	0.00534	-1.004

(i) Test the graduation for goodness of fit using the chi-squared test. [5]

(ii) (a) By inspection of the data, suggest one aspect of the graduated rates where adherence to data seems inadequate.

(b) Explain why this may not be detected by the chi-squared test.

(c) Carry out one other test that may detect this deficiency.

[5]

(iii) Suggest how the graduation could be adjusted to correct the deficiency identified. [2]

[Total 12]

END OF PAPER

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

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:

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

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:

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

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

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.

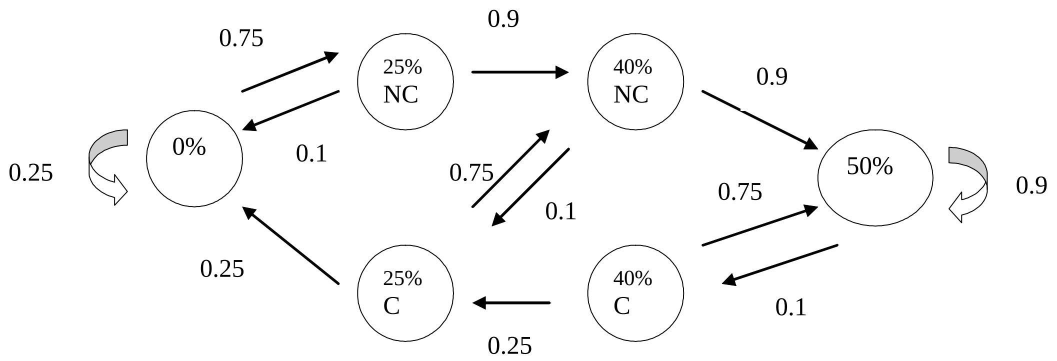
Specifically, if you are in the 25% discount level, the transition probability to state 0% is 0.25 if a claim was made last year and 0.1 if the previous year was claim free.

(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						
		0%	25%C	25%NC	40%C	40%NC	50%
Old State	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.

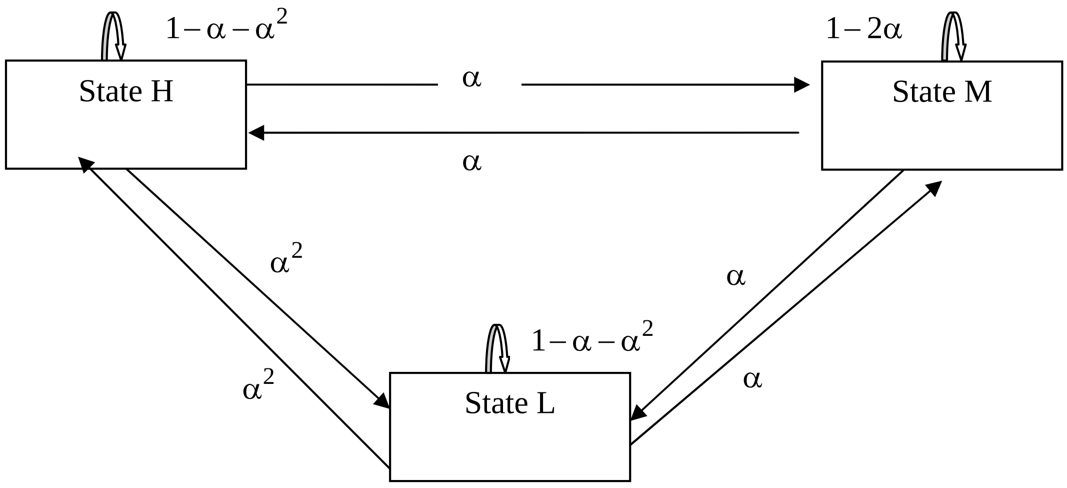
The following may favour a deterministic approach:

- 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 $P(X_t = j \mid X_s = i)$ 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,

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

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

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

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

(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.

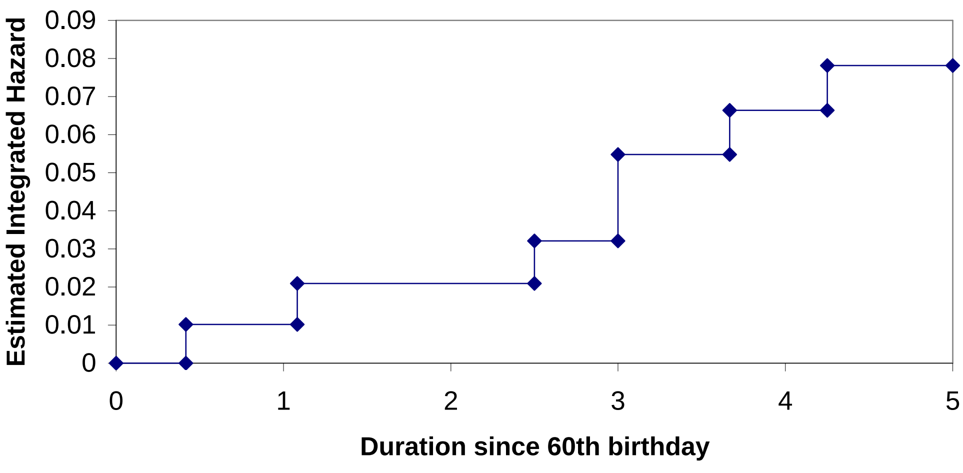
B5 (i) The following types of censoring will be present:

- 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	$\frac{d_j}{n_j}$	$\Lambda_j = \sum \frac{d_j}{n_j}$
$0 \le t < 5/12$	100	0	2	0	0
$5/12 \le t < 1\,1/12$	98	1	4	1/98	0.0102
$1\,1/12 \le t < 2\,6/12$	93	1	2	1/93	0.0210
$2\,6/12 \le t < 3$	90	1	1	1/90	0.0321
$3 \le t < 3\,8/12$	88	2	0	2/88	0.0548
$3\,8/12 \le t < 4\,3/12$	86	1	1	1/86	0.0664
$4\,3/12 \le t$	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 $\dfrac{\sum_x \left(E_x \hat{q}_x - E_x \mathring{q}_x\right)}{\sqrt{\sum_x E_x \mathring{q}_x \left(1 - \mathring{q}_x\right)}}$

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 $\left|-2.715\right|>1.96$, 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

EXAMINATION

6 September 2006 (am)

Subject CT4 (103) — Models (103 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 6 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

A1 A manufacturer uses a test rig to estimate the failure rate in a batch of electronic components. The rig holds 100 components and is designed to detect when a component fails, at which point it immediately replaces the component with another from the same batch. The following are recorded for each of the n components used in the test ($i = 1,2,\ldots,n$):

s_i = time at which component i placed on the rig
t_i = time at which component i removed from rig

$$f_i = \begin{cases} 1 & \text{Component removed due to failure} \\ 0 & \text{Component working at end of test period} \end{cases}$$

The test rig was fully loaded and was run for two years continuously.

You should assume that the force of failure, μ, of a component is constant and component failures are independent.

(i) Show that the contribution to the likelihood from component i is:

$$\exp\left(-\mu\left(t_i - s_i\right)\right)\cdot\mu^{f_i}$$ [2]

(ii) Derive the maximum likelihood estimator for μ. [4]
[Total 6]

A2 The price of a stock can either take a value above a certain point (state A), or take a value below that point (state B). Assume that the evolution of the stock price in time can be modelled by a two-state Markov jump process with homogeneous transition rates $\sigma_{AB} = \sigma$, $\sigma_{BA} = \rho$.

The process starts in state A at $t = 0$ and time is measured in weeks.

(i) Write down the generator matrix of the Markov jump process. [1]

(ii) State the distribution of the holding time in each of states A and B. [1]

(iii) If $\sigma = 3$, find the value of t such that the probability that no transition to state B has occurred until time t is 0.2. [2]

(iv) Assuming all the information about the price of the stock is available for a time interval $[0,T]$, explain how the model parameters σ and ρ can be estimated from the available data. [2]

(v) State what you would test to determine whether the data support the assumption of a two-state Markov jump process model for the stock price. [1]
[Total 7]

A3 (i) Define the following types of a stochastic process:

(a) a Poisson process
(b) a compound Poisson process; and
(c) a general random walk

[3]

(ii) For each of the processes in (i), state whether it operates in continuous or discrete time and whether it has a continuous or discrete state space. [2]

(iii) For each of the processes in (i), describe one practical situation in which an actuary could use such a process to model a real world phenomenon. [3]
[Total 8]

A4 The credit-worthiness of debt issued by companies is assessed at the end of each year by a credit rating agency. The ratings are A (the most credit-worthy), B and D (debt defaulted). Historic evidence supports the view that the credit rating of a debt can be modelled as a Markov chain with one-year transition matrix

$$X = \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

(i) Determine the probability that a company rated A will never be rated B in the future. [2]

(ii) (a) Calculate the second order transition probabilities of the Markov chain.

(b) Hence calculate the expected number of defaults within the next two years from a group of 100 companies, all initially rated A. [2]

The manager of a portfolio investing in company debt follows a "downgrade trigger" strategy. Under this strategy, any debt in a company whose rating has fallen to B at the end of a year is sold and replaced with debt in an A-rated company.

(iii) Calculate the expected number of defaults for this investment manager over the next two years, given that the portfolio initially consists of 100 A-rated bonds. [2]

(iv) Comment on the suggestion that the downgrade trigger strategy will improve the return on the portfolio. [2]
[Total 8]

 PLEASE TURN OVER

A5 A motor insurance company wishes to estimate the proportion of policyholders who make at least one claim within a year. From historical data, the company believes that the probability a policyholder makes a claim in any given year depends on the number of claims the policyholder made in the previous two years. In particular:

- the probability that a policyholder who had claims in both previous years will make a claim in the current year is 0.25
- the probability that a policyholder who had claims in one of the previous two years will make a claim in the current year is 0.15; and
- the probability that a policyholder who had no claims in the previous two years will make a claim in the current year is 0.1

(i) Construct this as a Markov chain model, identifying clearly the states of the chain. [2]

(ii) Write down the transition matrix of the chain. [1]

(iii) Explain why this Markov chain will converge to a stationary distribution. [2]

(iv) Calculate the proportion of policyholders who, in the long run, make at least one claim at a given year. [4]
[Total 9]

A6 (i) Explain the difference between a time-homogeneous and a time-inhomogeneous Poisson process. [1]

An insurance company assumes that the arrival of motor insurance claims follows an inhomogeneous Poisson process.

Data on claim arrival times are available for several consecutive years.

(ii) (a) Describe the main steps in the verification of the company's assumption.

(b) State one statistical test that can be used to test the validity of the assumption.

[3]

(iii) The company concludes that an inhomogeneous Poisson process with rate $\lambda(t) = 3 + \cos(2\pi t)$ is a suitable fit to the claim data (where t is measured in years).

(a) Comment on the suitability of this transition rate for motor insurance claims.

(b) Write down the Kolmogorov forward equations for $P_{0j}(s,t)$.

(c) Verify that these equations are satisfied by:

$$P_{0j}(s,t) = \frac{(f(s,t))^j.\exp(-f(s,t))}{j!}$$

for some $f(s,t)$ which you should identify.

[Note that $\int \cos x\,dx = \sin x$.]

(d) Comment on the form of the solution compared with the case where λ is constant.

[8]
[Total 12]

END OF PAPER

EXAMINATION

6 September 2006 (am)

Subject CT4 (104) — Models (104 Part)
Core Technical

Time allowed: One and a half hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 6 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

B1 Calculate ${}_{0.25}p_{80}$ and ${}_{0.25}p_{80.5}$, using the ELT15 (Females) mortality table and assuming a uniform distribution of deaths. [4]

B2 A national mortality investigation is carried out over the calendar years 2002, 2003 and 2004. Data are collected from a number of insurance companies.

Deaths during the period of the investigation, θ_x, are classified by age nearest at death.

Each insurance company provides details of the number of in-force policies on 1 January 2002, 2003, 2004 and 2005, where policyholders are classified by age nearest birthday, $P_x(t)$.

(i) (a) State the rate year implied by the classification of deaths.
(b) State the ages of the lives at the start of the rate interval.
[1]

(ii) Derive an expression for the exposed to risk, in terms of $P_x(t)$, which may be used to estimate the force of mortality in year t at each age. State any assumptions you make. [3]

(iii) Describe how your answer to (ii) would change if the census information provided by some companies was $P_x^*(t)$, the number of in-force policies on 1 January each year, where policyholders are classified by age last birthday.
[3]
[Total 7]

B3 An investigation was undertaken into the effect of a new treatment on the survival times of cancer patients. Two groups of patients were identified. One group was given the new treatment and the other an existing treatment.

The following model was considered:

$$h_i(t) = h_0(t) \cdot \exp\left(\underline{\beta}^T \underline{z}\right)$$

where: $h_i(t)$ is the hazard at time t, where t is the time since the start of treatment
$h_0(t)$ is the baseline hazard at time t

$\underline{z}$ is a vector of covariates such that:
z_1 = sex (a categorical variable with 0 = female, 1 = male)
z_2 = treatment (a categorical variable with 0 = existing treatment, 1 = new treatment)

and $\underline{\beta}$ is a vector of parameters, (β_1, β_2).

The results of the investigation showed that, if the model is correct:

A the risk of death for a male patient is 1.02 times that of a female patient; and

B the risk of death for a patient given the existing treatment is 1.05 times that for a patient given the new treatment

(i) Estimate the value of the parameters β_1 and β_2. [3]

(ii) Estimate the ratio by which the risk of death for a male patient who has been given the new treatment is greater or less than that for a female patient given the existing treatment. [2]

(iii) Determine, in terms of the baseline hazard only, the probability that a male patient will die within 3 years of receiving the new treatment. [2]
[Total 7]

B4 An investigation took place into the mortality of persons between exact ages 60 and 61 years. The table below gives an extract from the results. For each person it gives the age at which they were first observed, the age at which they ceased to be observed and the reason for their departure from observation.

Person	*Age at entry*		*Age at exit*		*Reason for exit*
	years	*months*	*years*	*months*	
1	60	0	60	6	withdrew
2	60	1	61	0	survived to 61
3	60	1	60	3	died
4	60	2	61	0	survived to 61
5	60	3	60	9	died
6	60	4	61	0	survived to 61
7	60	5	60	11	died
8	60	7	61	0	survived to 61
9	60	8	60	10	died
10	60	9	61	0	survived to 61

(i) Estimate q_{60} using the Binomial model. [5]

(ii) List the strengths and weaknesses of the Binomial model for the estimation of empirical mortality rates, compared with the Poisson and two-state models. [3]
[Total 8]

B5 A life insurance company has carried out a mortality investigation. It followed a sample of independent policyholders aged between 50 and 55 years. Policyholders were followed from their 50th birthday until they died, they withdrew from the investigation while still alive, or they celebrated their 55th birthday (whichever of these events occurred first).

(i) Describe the censoring that is present in this investigation. [2]

An extract from the data for 12 policyholders is shown in the table below.

Policyholder	*Last age at which policyholder was observed (years and months)*	*Outcome*
1	50 years 3 months	Died
2	50 years 6 months	Withdrew
3	51 years 0 months	Died
4	51 years 0 months	Withdrew
5	52 years 3 months	Withdrew
6	52 years 9 months	Died
7	53 years 0 months	Withdrew
8	53 years 6 months	Withdrew
9	54 years 3 months	Withdrew
10	54 years 3 months	Died
11	55 years 0 months	Still alive
12	55 years 0 months	Still alive

(ii) Calculate the Nelson-Aalen estimate of the survival function. [5]

(iii) Sketch on a suitably labelled graph the Nelson-Aalen estimate of the survival function. [2]

[Total 9]

B6 (i) (a) Describe the general form of the polynomial formula used to graduate the most recent standard tables produced for use by UK life insurance companies.

(b) Show how the Gompertz and Makeham formulae arise as special cases of this formula.

[3]

(ii) An investigation was undertaken of the mortality of persons aged between 40 and 75 years who are known to be suffering from a degenerative disease. It is suggested that the crude estimates be graduated using the formula:

$$\overset{o}{\mu}_{x+\frac{1}{2}} = \exp\left[b_0 + b_1\left(x+\frac{1}{2}\right) + b_2\left(x+\frac{1}{2}\right)^2 \right].$$

(a) Explain why this might be a sensible formula to choose for this class of lives.

(b) Suggest two techniques which can be used to perform the graduation.

[3]

(iii) The table below shows the crude and graduated mortality rates for part of the relevant age range, together with the exposed to risk at each age and the standardised deviation at each age.

Age last birthday	*Graduated force of mortality*	*Crude force of mortality*	*Exposed to risk*	*Standardised deviation*
x	$\overset{o}{\mu}_{x+1/2}$	$\hat{\mu}_{x+1/2}$	E_x^c	$z_x = \dfrac{E_x^c\left(\hat{\mu}_{x+1/2} - \overset{o}{\mu}_{x+1/2}\right)}{\sqrt{E_x^c\,\overset{o}{\mu}_{x+1/2}}}$
50	0.08127	0.07941	340	-0.12031
51	0.08770	0.08438	320	-0.20055
52	0.09439	0.09000	300	-0.24749
53	0.10133	0.10345	290	0.11341
54	0.10853	0.09200	250	-0.79336
55	0.11600	0.10000	200	-0.66436
56	0.12373	0.11176	170	-0.44369
57	0.13175	0.12222	180	-0.35225

Test this graduation for:

(a) overall goodness-of-fit

(b) bias; and

(c) the existence of individual ages at which the graduated rates depart to a substantial degree from the observed rates

[9]
[Total 15]

END OF PAPER

Faculty of Actuaries **Institute of Actuaries**

EXAMINATION

September 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 A Stocker
Chairman of the Board of Examiners

November 2006

Comments

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

103 Part

Question A1 *This was reasonably well answered, even by the weaker candidates.*
In part (ii), very few candidates used the information in the question and calculated $\sum_{i=1}^{n}(t_i - s_i)$.

Question A2 *This was reasonably well answered.*
In part (iv), many candidates wrote down a suitable estimate, but failed to provide an explanation as required.

Question A3 *This was reasonably well answered.*
In part (i), many candidates attempted to describe the simple random walk rather than the general case.
In part (ii), very few candidates identified the correct state space for the compound Poisson process or general random walk.
In part (iii), credit was not given if the examples cited were not likely to be encountered by an actuary working in a professional capacity.

Question A4 *This was not well answered overall, but many of the stronger candidates did score highly.*
In part (i), some candidates incorrectly attempted to calculate the long-run probability of being in state B.
Part (ii) was generally well answered.
In part (iv), the stronger candidates provided good answers, but overall candidates did not score well here.

Question A5 *Overall this was poorly answered, although the stronger candidates did well. Many candidates failed to split the two states labelled B and C in the solution, giving instead a 3-state chain. Some marks were still awarded for the long-run probability calculations in part (iv), but such candidates were not able to calculate the required final answer.*

Question A6 *This was poorly answered by most candidates, even though some parts of the question had been asked in previous (103) exams.*
Marks were lost in all parts of the question. Many candidates did not make a serious attempt at part (iii)(c).

104 Part

Question B1	*This was well answered.* *Some candidates assumed a constant force of mortality, for which credit was not given. Some candidates struggled with the second calculation.*
Question B2	*This was poorly answered overall, although some of the stronger candidates did manage to score highly.* *In part (ii), the question asked candidates to "derive an expression" and therefore we were looking for clearly set out steps here. Many candidates lost marks by not providing sufficient explanation of their working.*
Question B3	*This was well answered overall.* *In parts (i) and (ii), candidates were asked to "estimate" and some indication was required of how the numerical estimate was reached.*
Question B4	*This was not well answered overall.* *In part (i), many candidates did not calculate the correct exposed to risk. Marks were frequently lost because of insufficient working combined with an incorrect final answer. Candidates who wrote down the formulae they were using were given credit even if arithmetic slips were made.*
Question B5	*This was very well answered by most candidates.* *The most common errors were: inconsistency in the assumed order of death and censoring at ages 51 and 54 $^{3}/_{12}$; and continuation of the estimated survival function after age 55.*
Question B6	*This was reasonably well answered overall.* *Parts (i) and (ii) were poorly answered.* *In part (iii), the main areas where candidates lost marks were: not correctly stating the null hypothesis; failure to identify the correct degrees of freedom to be used in the chi-squared test; and a failure to state relevant and clear conclusions to the tests.*

103 Solutions

A1 (i) If the *i*th component is still working at the end of the test period its contribution to the likelihood is:

$$_{t_i - s_i}p_{s_i} = \exp(-\mu(t_i - s_i))$$

under the assumption of a constant force of failure.

If the *i*th component fails at time t_i its contribution to the likelihood is:

$$_{t_i - s_i}p_{s_i}.\mu_{t_i} = \exp(-\mu(t_i - s_i)).\mu$$

under the assumption of a constant force of failure.

In both cases the contribution equals:

$$\exp(-\mu(t_i - s_i)).\mu^{f_i}$$

(ii) Denote the total number of components used in the test by n. The likelihood for n independent components is:

$$L = \prod_{i=1}^{n} \exp(-\mu(t_i - s_i)).\mu^{f_i}$$

$$L = \exp(-\mu\sum_{i=1}^{n}(t_i - s_i)).\mu^{\sum_{i=1}^{n} f_i}$$

Now the rig contains 100 components at all times because it is fully loaded and failed components are immediately replaced, so $\sum_{i=1}^{n}(t_i - s_i) = 200(\text{years})$.

So $$L = \exp(-200\mu)\cdot\mu^{\sum_{i=1}^{n} f_i}$$

$$\ln L = -200\mu + \ln\mu.\sum_{i=1}^{n} f_i$$

$$\frac{\partial \ln L}{\partial \mu} = -200 + \frac{\sum_{i=1}^{n} f_i}{\mu}$$

Setting this to zero the MLE is:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} f_i}{200}$$

To verify this is a maximum we see that:

$$\frac{\partial^2 \ln L}{\partial \mu^2} = -\frac{\sum_{i=1}^{n} f_i}{\mu^2} < 0$$

A2 (i) The generator matrix is

$$A = \begin{pmatrix} -\sigma & \sigma \\ \rho & -\rho \end{pmatrix}$$

(ii) The distribution is exponential in both cases; with parameter σ in state A, ρ in state B.

(iii) The probability that the process stays in A throughout [0, *t*] is

$$\int_t^{\infty} \sigma e^{-\sigma s} ds = e^{-\sigma t}.$$

For $\sigma = 3$, we get $e^{-3t} = 0.2$
which gives $t = -\ln(0.2)/3 = 0.54$ weeks.

(iv) The time spent in state A before the next visit to B has mean 1/σ.

Therefore a reasonable estimate for σ is the reciprocal of the mean length of each visit:

$\hat{\sigma}$ = (Number of transitions from A to B) / (Total time spent in state A up until the last transition from A to B).

[An alternative is to use the maximum likelihood estimator for σ, which is (Number of transitions from A to B)/Total time spent in state A).]

Similarly we can estimate $\hat{\rho}$.

(v) Testing whether the successive holding times are exponential variables and independent would be best. Any procedure which does this test is acceptable.

A3 (i) (a) A Poisson process with rate λ is an 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, each having a Poisson distribution, i.e.

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

(b) Let N_t be a Poisson process, $t \geq 0$ and let $Y_1, Y_2, \ldots, Y_j, \ldots$, be a sequence of i.i.d. random variables. Then a compound Poisson process is defined by

$$X_t = \sum_{j=1}^{N_t} Y_j, \quad t \geq 0.$$

(c) Let $Y_1, Y_2, \ldots, Y_j, \ldots$, be a sequence of independent and identically distributed random variables and define

$$X_n = \sum_{j=1}^{n} Y_j$$

with initial condition $X_0 = 0$. Then $\{X_n\}_{n=0}^{\infty}$ constitutes a general random walk.

(ii) (a) A Poisson process operates in continuous time and has a discrete state space, the set of nonnegative integers.

(b) A compound Poisson process operates in continuous time.

It has a discrete or continuous state space depending on whether the variables Y_j are discrete or continuous respectively.

(c) A general random walk operates in discrete time. Again, this has a discrete or continuous state space according to whether the variables Y_j have a discrete or continuous distribution.

(iii) (a) Examples of a Poisson process:

- claims arriving to an insurance company through time
- car accidents reported over time
- arrival of customers at a service point over time

(b) A standard example of a compound Poisson process used by actuaries is for modelling the total amount of claims to an insurance company over time.

(c) Examples of a general random walk:

- modelling share prices daily
- inflation index, measured on say a monthly basis

Other reasonable examples received credit.

A4 (i) Probability that a company is never in state B is:

$$\Pr(A \to D) + \Pr(A \to A \to D) + \Pr(A \to A \to A \to D) + \ldots\ldots$$

$$= 0.03 + 0.92 \times 0.03 + 0.92^2 \times 0.03 + \ldots\ldots$$

$$= 0.03 \times \sum_{i=0}^{\infty} 0.92^i = \frac{0.03}{1-0.92} = 0.375$$

(ii) (a)

$$A^2 = \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8489 & 0.0885 & 0.0626 \\ 0.0885 & 0.725 & 0.1865 \\ 0 & 0 & 1 \end{pmatrix}$$

(b) Probability of default within 2 years for an A rated company 6.26%, so 6.26 defaults expected.

(iii) *Either*

Calculate revised transition probabilities based on the rating of bonds held by the investment manager after rebalancing:

$$A' = \begin{pmatrix} 0.97 & 0 & 0.03 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(state B is unnecessary so this can be shown as 2 × 2 or 3 × 3)

$$A'^2 = \begin{pmatrix} 0.9409 & 0 & 0.0591 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

So the expected number of defaults is $0.0591 \times 100 = 5.91$.

Or

Required probability is

$$\Pr(A \to D) + \Pr(A \to A) \times \Pr(A \to D) + \Pr(A \to B) \times \Pr(A \to D)$$

$$= 0.03 + 0.92 \times 0.03 + 0.05 \times 0.03 = 0.0591$$

So expected defaults 5.91.

(iv) The expected number of defaults has been reduced by this strategy. (The variance of the number of defaults would also reduce.)

However it is not possible to tell whether the overall return is improved as this depends on the price at which bonds were bought and sold at the end of year 1.

The price of the debt sold may have been depressed by the companies having been downgraded to rating B, and the manager loses out on any increase in price if they recover.

The "downgrade trigger" strategy will incur dealing costs, which should be considered when comparing the returns.

A5 (i) Consider the following four states that the policyholder might be at the end of a year:

- the policyholder has made at least one claim both in the year just ended and the previous one (state A)
- the policyholder has made no claims in the year just ended but s/he made at least one claim during the previous year (state B)
- the policyholder has made at least one claim in the year just ended but not in the previous one (state C)
- the policyholder has made no claim during either the year ended or the previous one (state D)

If the year ended is year n, and X_n denotes the current state of the policyholder, then X_n constitutes a Markov chain.

(ii) The transition matrix is

$$P = \begin{pmatrix} 0.25 & 0.75 & 0 & 0 \\ 0 & 0 & 0.15 & 0.85 \\ 0.15 & 0.85 & 0 & 0 \\ 0 & 0 & 0.10 & 0.90 \end{pmatrix}$$

(iii) The chain has a finite number of states (A,B,C,D). In order to show that it has a stationary distribution, it suffices to show that it is irreducible and aperiodic.

It is apparent from the transition matrix above that any state can be reached from any other; hence the chain is irreducible.

The chain is also aperiodic since for states A, D the state can remain at the same state after one step, while for states B, C the state may return to its current state after 2 or 3 steps.

Hence the chain has a stationary distribution (which is unique).

(iv) The set of equations is given (in matrix from) by $\pi P=\pi$,
where $\pi = (\pi_A, \pi_B, \pi_C, \pi_D)$ denotes the stationary distribution.

Using the transition matrix from (ii) above we obtain the equations

$$
\begin{array}{llllll}
0.25\,\pi_A + & & +0.15\,\pi_C & & = \pi_A & (1) \\
0.75\,\pi_A + & & +0.85\,\pi_C & & = \pi_B & (2) \\
 & 0.15\,\pi_B & & +0.10\,\pi_D & = \pi_C & (3) \\
 & 0.85\,\pi_B & & +0.90\,\pi_D & = \pi_D &
\end{array}
$$

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

$$\pi_A + \pi_B + \pi_C + \pi_D = 1 \qquad (4)$$

Equation (1) gives

$$0.75\,\pi_A = 0.15\,\pi_C \qquad (5)$$

Or $\quad 5\,\pi_A = \pi_C$

Substituting (5) into (2) yields immediately

$$\pi_B = \pi_C$$

and inserting this into (3) we get

$$\pi_D = \frac{17}{2}\pi_B.$$

In view of the above, we obtain now from (4) that

$$\pi_B\left(\frac{1}{5}+1+1+\frac{17}{2}\right)=1 \;\Rightarrow \pi_B = \frac{10}{107}.$$

Hence the other probabilities are

$$\pi_A = \frac{2}{107},\; \pi_C = \frac{10}{107},\; \pi_D = \frac{85}{107}.$$

The proportion of policyholders who, in the long run, make at least one claim in a given year is

$$\pi_A + \pi_B = \frac{12}{107}.$$

A6 (i) The probability that an event occurs during the short time interval between t and $t + h$ is approximately equal to $\lambda(t)\, h$ for small h where $\lambda(t)$ is called the rate of the process. For a time-inhomogeneous process, $\lambda(t)$ depends on the current time t; for a time-homogeneous process it is independent of time.

(ii) (a) Divide the time period into intervals of a suitable size, say one month. Estimate the arrival rate separately for each time period.

See if the observed data match the pattern which would be expected if the model were accurate and if the parameters had their values given by their estimates.

If not, the model should be revised.

(b) A goodness of fit test, such as the chi-squared test, should be carried out for each time period chosen.

Tests for serial correlation [e.g. portmanteau test] should use the whole data set at once.

(iii) (a) This implies that claims are seasonal with period 12 months, and that claims in the peak (presumably winter) are double those at the low point of the year.

This would be reasonable if in a climate where driving conditions are worse in winter.

(b) Kolmogorov forward equations:

$$\frac{\partial}{\partial t} P(s,t) = P(s,t).A(t) \qquad\qquad t \geq s$$

Where:

$$A(t) = \begin{pmatrix} -\lambda(t) & \lambda(t) & & \\ & -\lambda(t) & \lambda(t) & \\ & & -\lambda(t) & \ddots \\ & & & \ddots \end{pmatrix}$$

(c) Consider the case $j > 0$,

$$\frac{\partial}{\partial t} P_{0j}(s,t) = \lambda(t).P_{0,j-1}(s,t) - \lambda(t).P_{0j}(s,t) \qquad \text{(I)}$$

with $P_{0j}(s,s) = 0$

If solution is of the form

$$P_{0j}(s,t) = \frac{(f(s,t))^j.\exp(-f(s,t))}{j!}$$

LHS of I

$$(j.(f(s,t))^{j-1} - f(s,t)^j).\frac{\exp(-f(s,t))}{j!}.\frac{d}{dt}f(s,t)$$

RHS of I

$$\lambda(t).\frac{f(s,t)^{j-1}}{(j-1)!}.\exp(-f(s,t)) - \lambda(t).\frac{f(s,t)^j.\exp(-f(s,t))}{j!}$$

These are equal if

$$\frac{\partial}{\partial t}f(s,t) = \lambda(t)$$

Now

$$\int_s^t \lambda(v)dv = \int_s^t (3+\cos(2\pi v))dv$$

$$= \left[3v + \frac{1}{2\pi}\sin(2\pi v)\right]_s^t$$

$$= 3(t-s) + \frac{1}{2\pi}\left[\sin(2\pi t) - \sin(2\pi s)\right] \equiv f(s,t)$$

this satisfies the boundary condition.

Consider the case $j = 0$

$$\frac{\partial}{\partial t}P_{00}(s,t) = -\lambda(t).P_{00}(s,t) \qquad \text{(II)}$$

with boundary condition $P_{00}(s,s) = 1$

Need to verify that $P_{00}(s,t) = \exp(-f(s,t))$ satisfies II

LHS of II

$$-\exp(-f(s,t)).\frac{\partial}{\partial t}(f(s,t)) = -P_{00}(s,t).\lambda(t)$$

and $P_{00}(s,s) = 1$

(d) Solution is of the same form, except that for the homogeneous case $f(s,t) = \lambda(t - s)$.

104 Solutions

B1 $\quad {}_{0.25}p_{80} = 1 - {}_{0.25}q_{80} = 1 - 0.25 \times q_{80}$

under the assumption of a uniform distribution of deaths (UDD) between ages 80 and 81.

From ELT 15, $q_{80} = 0.05961$, so

$${}_{0.25}p_{80} = 1 - 0.25 \times 0.05961 = 0.98510$$

ALTERNATIVE 1

Under UDD we have, for $0 \leqslant s < t \leqslant 1$,

$${}_{t-s}q_{x+s} = \frac{(t-s)q_x}{1 - sq_x}.$$

Putting $t = 0.75$, $s = 0.5$ and $x = 80$, therefore,

$${}_{0.75-0.5}q_{80+0.5} = \frac{0.25q_{80}}{1 - 0.5q_{80}}, \text{ and so}$$

$${}_{0.25}p_{80.5} = 1 - \frac{0.25q_{80}}{1 - 0.5q_{80}}.$$

Using ELT15, this is evaluated as

$$1 - \frac{0.25(0.05961)}{1 - 0.5(0.05961)} = 1 - \frac{0.01490}{0.97020} = 1 - 0.01536 = 0.98464$$

ALTERNATIVE 2

Using $\quad {}_{t}p_x = {}_{s}p_x \cdot {}_{t-s}p_{x+s},$

$${}_{0.75}p_{80} = {}_{0.5}p_{80} \cdot {}_{0.25}p_{80.5}$$

Using an assumption of UDD between ages 80 and 81, we have

$${}_{0.5}p_{80} = 1 - 0.5 \times 0.05961 = 0.97020$$

$${}_{0.75}p_{80} = 1 - 0.75 \times 0.05961 = 0.95529$$

So, $\quad {}_{0.25}p_{80.5} = \dfrac{{}_{0.75}p_{80}}{{}_{0.5}p_{80}} = \dfrac{0.95529}{0.97020} = 0.98463$

B2 (i) (a) The age definition changes 6 months before/after each birthday, so this is a life year rate interval.

(b) Lives are aged x - ½ at the start of the rate interval.

(ii) Under the principle of correspondence the age definition of deaths and census should correspond, which they do here. So we do not need to adjust the census information.

The exposed to risk is given by $E_x^c = \int_0^3 P_x(t)\,dt$.

Assuming $P_x(t)$ is linear over calendar years, we can approximate this to

$$E_x^c = \sum_0^2 \frac{1}{2}\left(P_x(t) + P_x(t+1)\right)\text{, where } t \text{ is measured from 1 January 2002}$$

$$= \left(\frac{1}{2}P_x(0) + P_x(1) + P_x(2) + \frac{1}{2}P_x(3)\right)$$

(iii) The age definitions for deaths and census no longer correspond. So, we need to adjust the census information for those companies who supply details of $P_x^*(t)$.

Assuming birthdays are uniformly distributed over the calendar year, we can approximate $P_x(t) \approx \frac{1}{2}\left(P_{x-1}^*(t) + P_x^*(t)\right)$.

And the exposed to risk is then:

$$E_x^c = \sum_0^2 \frac{1}{2}\left(P_x(t) + P_x(t+1)\right)$$

$$= \sum_0^2 \frac{1}{2}\left(\frac{1}{2}\left(P_{x-1}^*(t) + P_x^*(t)\right) + \frac{1}{2}\left(P_{x-1}^*(t+1) + P_x^*(t+1)\right)\right)$$

$$= \frac{1}{4}\left(P_{x-1}^*(0) + P_x^*(0)\right) + \frac{1}{2}\left(P_{x-1}^*(1) + P_x^*(1) + P_{x-1}^*(2) + P_x^*(2)\right) + \frac{1}{4}\left(P_{x-1}^*(3) + P_x^*(3)\right)$$

B3 (i) The hazard for a female patient is:

$$h_f(t) = h_0(t) \times \exp(0 + \beta_2 z_2)$$

and the hazard for a male patient is:

$$h_m(t) = h_0(t) \times \exp(\beta_1 \times 1 + \beta_2 z_2)$$

Using $\hat{\beta}_i$ to denote our estimate of β_i, we know from A that, if the model is correct,

$$h_m(t) = 1.02 \times h_f(t), \text{ so that:}$$

$$h_0(t) \times \exp\left(\hat{\beta}_1 + \hat{\beta}_2 z_2\right) = 1.02 \times h_0(t) \times \exp\left(\hat{\beta}_2 z_2\right)$$

$$\Rightarrow \exp(\hat{\beta}_1) = 1.02$$

$$\Rightarrow \hat{\beta}_1 = \ln(1.02) = 0.0198$$

And similarly, from B, we know that:

$$h_0(t) \times \exp\left(\hat{\beta}_1 z_1 + 0\right) = 1.05 \times h_0(t) \times \exp\left(\hat{\beta}_1 z_1 + \hat{\beta}_2 z_2\right)$$

$$\Rightarrow 1 = 1.05 \times \exp\left(\hat{\beta}_2\right)$$

$$\Rightarrow \hat{\beta}_2 = \ln\left(1/1.05\right) = -0.0488$$

(ii) The hazard for a male patient who has been given the new treatment is:

$$h_{m,n}(t) \quad = h_0(t) \times \exp(\beta_1 \times 1 + \beta_2 \times 1)$$

$$= h_0(t) \times \exp(0.0198 - 0.0488)$$

$$= h_0(t) \times \exp(-0.029)$$

$$= 0.9714 \times h_0(t)$$

The hazard for a female patient given the existing treatment is the baseline hazard.

Hence, the ratio of the hazard for a male patient who has been given the new treatment to that for a female patient given the existing treatment is:

$$\frac{h_{m,n}(t)}{h_0(t)} = 0.9714$$

ALTERNATIVELY

Candidates may recognise that the proportions given in A and B can be combined to give:

$$\frac{h_{m,n}(t)}{h_{f,e}(t)} = \left[\frac{h_{m,x}(t)}{h_{f,x}(t)}\right] \times \left[\frac{h_{x,n}(t)}{h_{x,e}(t)}\right] = 1.02 \times \frac{1}{1.05} = 0.9714$$

(iii) The probability of death is given by:

$$\begin{aligned} 1 - S_{m,n}(3) &= 1 - \exp\left\{-\int_0^3 h_{m,n}(s)ds\right\} \\ &= 1 - \exp\left\{-\int_0^3 0.9714 \times h_0(s)ds\right\} \\ &= 1 - \exp\left\{0.9714 \times \left(-\int_0^3 h_0(s)ds\right)\right\} \\ &= 1 - \left(e^{-\int_0^3 h_0(s)ds}\right)^{0.9714} \end{aligned}$$

B4 (i) Let the age individual i enters observation be a_i and the age that individual i leaves observation be b_i. Define an indicator variable d_i such that $d_i = 0$ if individual i is not observed to die and $d_i = 1$ if individual i dies.

Measure all ages in years since exact age 60.

The estimate of q_{60} using the Binomial model is:

$$\hat{q}_{60} = \frac{\sum_{i=1}^{10} d_i}{\sum_{i=1}^{10}\left(1 - a_i - \left[(1-d_i)(1-b_i)\right]\right)}.$$

The denominator in this formula shows that for persons who do not die ($d_i = 0$) the exposed to risk is $b_i - a_i$ and for persons who die ($d_i = 1$) the exposed to risk is $1 - a_i$.

Thus the relevant calculations are shown in the table below (all durations are in years).

Person	a_i	b_i	d_i	$1 - a_i$	$1 - b_i$	$1 - a_i - (1 - d_i)(- b_i)$
1	0	6/12	0	1	6/12	6/12
2	1/12	1	0	11/12	0	11/12
3	1/12	3/12	1	11/12	9/12	11/12
4	2/12	1	0	10/12	0	10/12
5	3/12	9/12	1	9/12	3/12	9/12
6	4/12	1	0	8/12	0	8/12
7	5/12	11/12	1	7/12	1/12	7/12
8	7/12	1	0	5/12	0	5/12
9	8/12	10/12	1	4/12	2/12	4/12
10	9/12	1	0	3/12	0	3/12
Totals			4			74/12

Therefore $\hat{q}_{60} = \dfrac{4}{74/12} = 0.6486$.

ALTERNATIVELY

Take the central exposed to risk, $\sum_{1}^{10}(b_i - a_i)$ (in years) and add ½d_{60} to give the initial exposed to risk.

This involves estimating q_{60} using the formula

$$\hat{q}_{60} = \frac{d_{60}}{E_{60}^c + 0.5d_{60}} = \frac{4}{(59/12)+2} = \frac{4}{83/12} = 0.5783.$$

[This approach is inferior to the first, as it does not use all the information available in the data, and involves the assumption that the deaths take place, on average, half way through the year.]

(ii) Strengths of Binomial model

- avoids numerical solution of equations
- can be generalised to give the Kaplan-Meier estimate

Weaknesses of Binomial model

- need to compute an initial exposed-to-risk is a pointless complication if census-type data are available
- not so easily generalised as two-state or Poisson models to processes with more than one decrement, and not so easily generalised as two-state model to increments
- estimate of q_x has a higher variance than that of the two-state Poisson models (though the difference is very small unless mortality is very high)

B5 (i) There will be Type I censoring of lives that survive to age 55 years.

There will be random censoring of lives that withdraw before age 55 years.

(ii) The calculations are shown in the table below, where durations are measured in years since the 50th birthday.

Using the convention that, when deaths and withdrawals are observed at the same duration, deaths occur first:

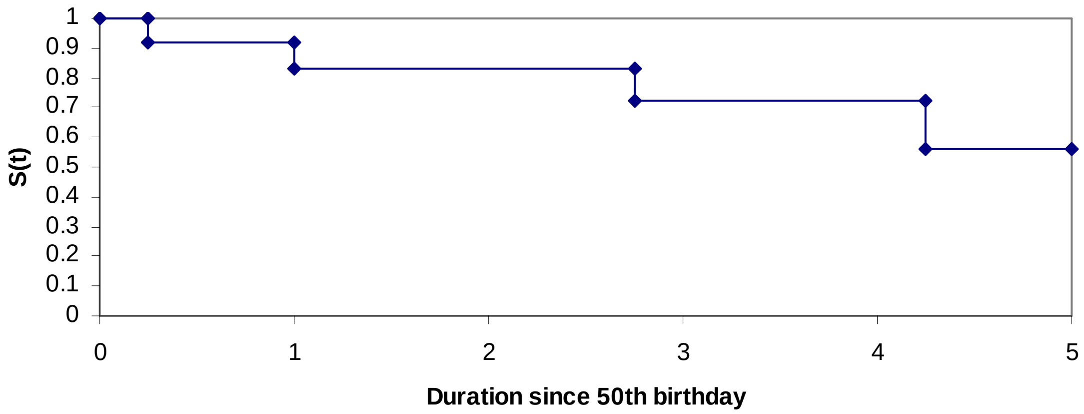
t_j	N_j	d_j	c_j	d_j / N_j	$\hat{\Lambda}_t = \sum_{t_j \le t} (d_j / N_j)$
0	12				
0.25	12	1	1	0.0833	0.0833
1.00	10	1	2	0.1000	0.1833
2.75	7	1	2	0.1429	0.3262
4.25	4	1	3	0.25	0.5762

Since $\hat{S}(t) = \exp(-\hat{\Lambda}_t)$

the estimated survival function is

t	$\hat{S}(t)$
$0 \le t < 0.25$	1.0000
$0.25 \le t < 1.00$	0.9201
$1.00 \le t < 2.75$	0.8325
$2.75 \le t < 4.25$	0.7217
$4.25 \le t < 5.00$	0.5620

(iii)



B6 (i) (a) The general form is

$$\mu_x = (\text{polynomial}(1)) + \exp(\text{polynomial}(2)),$$

where polynomial (1) takes the form

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots$$

and polynomial (2) takes the form

$$\beta_0 + \beta_1 x + \beta_2 x^2 + \ldots.$$

(b) In the case of the Gompertz formula $\mu_x = Bc^x$, then putting

$$B = \exp(\beta_0) \text{ and } c = \exp(\beta_1),$$

we can re-write the formula as

$$\mu_x = \exp(\beta_0)\exp(\beta_1 x) = \exp(\beta_0 + \beta_1 x),$$

which is of the required form if

$$\alpha_i = 0 \text{ for all } i$$

and

$$\beta_i = 0 \text{ for } i = 2, 3, \ldots.$$

Similarly the Makeham formula $\mu_x = A + Bc^x$
can be expressed in the required form by putting

$$A = \alpha_0,\ B = \exp(\beta_0) \text{ and } c = \exp(\beta_1).$$

(ii) (a) The Gompertz formula written

$$\mu_x = \exp(\beta_0 + \beta_1 x)$$

is an exponential function which implies that the rate of increase of mortality with age is constant.

This is often a reasonable assumption for ordinary lives at middle ages and older ages.

In the special case of the impaired lives known to be suffering from a degenerative disease, it is plausible to suppose that the rate of increase of mortality might increase with age.

The term $b_2\left(x+\frac{1}{2}\right)^2$ in the formula can allow for this possibility.

(b) The graduation can be achieved by

maximum likelihood estimation of the parameters

or by ordinary least squares regression

of $\log\left[\hat{\mu}_{x+\frac{1}{2}}\right]$ on $x+\frac{1}{2}$ and $\left(x+\frac{1}{2}\right)^2$.

(iii) (a) The null hypothesis is that there is no difference between the graduated rates and the underlying rates in the population from which the crude rates are derived.

To test overall goodness-of-fit we use the chi-squared test.

$$\sum_x z_x^{\ 2} \sim \chi^2_{\ m},$$

where m is the number of degrees of freedom.

In this case, we have 8 ages, but 3 parameters were estimated when performing the graduation, so $m = 5$.

The calculations are shown in the table below.

Age x last birthday	z_x	$z_x^{\ 2}$
50	-0.12031	0.01447
51	-0.20055	0.04022
52	-0.24749	0.06125
53	0.11341	0.01286
54	-0.79336	0.62942
55	-0.66436	0.44137
56	-0.44369	0.19686
57	-0.35225	0.12408
Sum		1.52053

The critical value of the chi-squared distribution with 5 degrees of freedom at the 5 per cent level is 11.07.

Since $1.52052 \ll 11.07$, we do not reject the null hypothesis and conclude that the graduation adheres satisfactorily to the data.

(b) To test for bias we use EITHER the Signs Test or the Cumulative Deviations test.

Signs Test

The test statistic, P, is the number of signs that is positive.

Under the null hypothesis, $P \sim \text{Binomial}(8,0.5)$

In this case $P = 1$, and $\text{Prob}[P \leq 1] = 0.0352$.

Since this probability > 0.025 (two-tailed test) we do not reject the null hypothesis.

We conclude that the graduated rates are not biased above or below the crude rates.

Cumulative deviations test

The test statistic

$$\frac{\sum_x (\hat{\mu}_{x+\frac{1}{2}} E_x - \overset{o}{\mu}_{x+\frac{1}{2}} E_x)}{\sqrt{\sum_x \overset{o}{\mu}_{x+\frac{1}{2}} E_x}} \sim \text{Normal}(0,1).$$

The calculations are shown in the table below.

Age x last birthday	$\hat{\mu}_{x+\frac{1}{2}} E_x - \overset{o}{\mu}_{x+\frac{1}{2}} E_x$	$\overset{o}{\mu}_{x+\frac{1}{2}} E_x$
50	-0.63	27.63
51	-1.06	28.06
52	-1.32	28.32
53	0.61	29.39
54	-4.13	27.13
55	-3.20	23.20
56	-2.03	21.03
57	-1.72	23.72
Sum	-13.48	208.48

The value of the test statistic is therefore

$(-13.48/\sqrt{208.48}) = -0.9335.$

using a two-tailed test, the absolute value of the test statistics is less than 1.96, so we do not reject the null hypothesis.

We conclude that the graduated rates are not biased above or below the crude rates.

(c) To test for the existence of individual ages at which the graduated rates depart greatly from the observed rates we can use the Individual Standardised Deviations Test.

There are no ages at which the absolute value of z_x exceeds 1.96.

Therefore we do not reject the null hypothesis and conclude that there are no outliers.

END OF EXAMINERS' REPORT

EXAMINATION

20 April 2007 (am)

Subject CT4 — Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*
2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*
3. *Mark allocations are shown in brackets.*
4. *Attempt all 11 questions, beginning your answer to each question on a separate sheet.*
5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 (a) Define, in the context of stochastic processes, a:

1. mixed process
2. counting process

(b) Give an example application of each type of process.

[4]

2 An insurance company is investigating the mortality of its annuity policyholders. It is proposed that the crude mortality rates be graduated for use in future premium calculations.

(i) (a) Suggest, with reasons, a suitable method of graduation in this case.

(b) Describe how you would graduate the crude rates.

[3]

(ii) Comment on any further considerations that the company should take into account before using the graduated rates for premium calculations. [2]

[Total 5]

3 The government of a small country has asked you to construct a model for forecasting future mortality.

Outline the stages you would go through in identifying an appropriate model. [6]

4 The actuary to a large pension scheme carried out an investigation of the mortality of the scheme's pensioners over the two years from 1 January 2005 to 1 January 2007.

(i) List the data required by the actuary for an exact calculation of the central exposed to risk for lives aged x. [2]

The following is an extract from the data collected by the actuary.

Age x nearest birthday	*Number of pensioners at:*			*Deaths during:*	
	1 January 2005	*1 January 2006*	*1 January 2007*	*2005*	*2006*
63	1,248	1,312	1,290	10	6
64	1,465	1,386	1,405	13	15
65	1,678	1,720	1,622	16	23
66	1,719	1,642	1,667	22	19
67	1,686	1,695	1,601	19	25

(ii) (a) Derive an expression that could be used to estimate the central exposed to risk using the available data. State any assumptions you make.

(b) Use the data to estimate μ_{65}. State any further assumptions that you make. [4]

[Total 6]

5 (i) Define the *hazard rate*, $h(t)$, of a random variable T denoting lifetime. [1]

(ii) An investigation is undertaken into the mortality of men aged between exact ages 50 and 55 years. A sample of n men is followed from their 50th birthdays until either they die or they reach their 55th birthdays.

The hazard of death (or force of mortality) between these ages, $h(t)$, is assumed to have the following form:

$$h(t) = \alpha + \beta t$$

where α and β are parameters to be estimated and t is measured in years since the 50th birthday.

(a) Derive an expression for the survival function between ages 50 and 55 years.

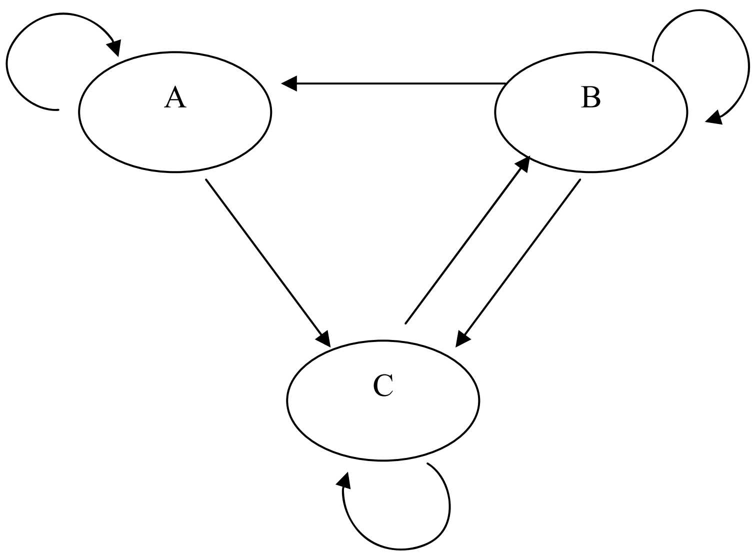
(b) Sketch this on a graph.

(c) Comment on the appropriateness of the assumed form of the hazard for modelling mortality over this age range.

[6]

[Total 7]

6 A three state process with state space {A, B, C} is believed to follow a Markov chain with the following possible transitions:



An instrument was used to monitor this process, but it was set up incorrectly and only recorded the state occupied after every two time periods. From these observations the following two-step transition probabilities have been estimated:

$P_{AA}^{2} = 0.5625$

$P_{AB}^{2} = 0.125$

$P_{BA}^{2} = 0.475$

$P_{CC}^{2} = 0.4$

Calculate the one-step transition matrix consistent with these estimates. [8]

7 Every person has two chromosomes, each being a copy of one of the chromosomes from one of their parents. There are two types of chromosomes labelled X and Y. A child born with an X and a Y chromosome is male and a child with two X chromosomes is female.

The blood-clotting disorder haemophilia is caused by a defective X chromosome (X*). A female with the defective chromosome (X*X) will not usually exhibit symptoms of the disease but may pass the defective gene to her children and so is known as a *carrier*. A male with the defective chromosome (X*Y) suffers from the disease and is known as a *haemophiliac*.

A medical researcher wishes to study the progress of the disease through the first born child in each generation, starting with a female carrier.

You may assume:

- every parent has a equal chance of passing either of their chromosomes to their children
- the partner of each person in the study does not carry a defective X chromosome; and
- no new genetic defects occur

(i) Show that the expected progress of the disease through the generations may be modelled as a Markov chain and specify carefully:

(a) the state space; and
(b) the transition diagram

[5]

(ii) State, with reasons, whether the chain is:

(a) irreducible; and
(b) aperiodic

[2]

(iii) Calculate the stationary distribution of the Markov chain. [3]
[Total 10]

8 A medical study was carried out between 1 January 2001 and 1 January 2006, to assess the survival rates of cancer patients. The patients all underwent surgery during 2001 and then attended 3-monthly check-ups throughout the study.

The following data were collected:

For those patients who died during the study exact dates of death were recorded as follows:

Patient	*Date of surgery*	*Date of death*
A	1 April 2001	1 August 2005
B	1 April 2001	1 October 2001
C	1 May 2001	1 March 2002
D	1 September 2001	1 August 2003
E	1 October 2001	1 August 2002

For those patients who survived to the end of the study:

Patient	*Date of surgery*
F	1 February 2001
G	1 March 2001
H	1 April 2001
I	1 June 2001
J	1 September 2001
K	1 September 2001
L	1 November 2001

For those patients with whom the hospital lost contact before the end of the investigation:

Patient	*Date of surgery*	*Date of last check-up*
M	1 February 2001	1 August 2003
N	1 June 2001	1 March 2002
O	1 September 2001	1 September 2005

(i) Explain whether and where each of the following types of censoring is present in this investigation:

(a) type I censoring
(b) interval censoring; and
(c) informative censoring [3]

(ii) Calculate the Kaplan-Meier estimate of the survival function for these patients. State any assumptions that you make. [7]

(iii) Hence estimate the probability that a patient will die within 4 years of surgery. [1]

[Total 11]

9 An insurance company is concerned that the ratio between the mortality of its female and male pensioners is unlike the corresponding ratio among insured pensioners in general. It conducts an investigation and estimates the mortality of male and female pensioners, $\hat{\mu}^m_{x+1/2}$ and $\hat{\mu}^f_{x+1/2}$. It then uses the $\hat{\mu}^m_{x+1/2}$ to calculate what the expected mortality of its female pensioners would be if the ratio between male and female mortality rates reflected the corresponding ratio in the PMA92 and PFA92 tables, $S_{x+1/2}$, using the formula

$$\tilde{\mu}^f_{x+1/2} = \hat{\mu}^m_{x+1/2} S_{x+1/2}.$$

The table below shows, for a range of ages, the numbers of female deaths actually observed in the investigation and the number which would be expected from the $\tilde{\mu}^f_{x+1/2}$.

Age x	*Actual deaths* $E^c_x \hat{\mu}^f_{x+1/2}$	*Expected deaths* $E^c_x \tilde{\mu}^f_{x+1/2}$
65	30	28.4
66	20	30.1
67	25	31.2
68	40	33.5
69	45	34.1
70	50	41.8
71	50	46.5
72	45	44.5

(i) Describe and carry out an overall test of the hypothesis that the ratios between male and female death rates among the company's pensioners are the same as those of insured pensioners in general. Clearly state your conclusion. [5]

(ii) Investigate further the possible existence of unusual ratios between male and female death rates among the company's pensioners, using two other appropriate statistical tests. [6]

[Total 11]

10 The members of a particular profession work exclusively in partnerships.

A certain partnership is concerned that it is losing trained technical staff to its competitors. Informal debriefing interviews with individuals leaving the partnership suggest that one reason for this is that the duration elapsing between becoming fully qualified and being made a partner is longer in this partnership than in the profession as a whole.

The partnership decides to investigate whether this claim is true using a multiple-state model with three states: (1) fully qualified but not yet a partner, (2) fully qualified and a partner, (3) working for another partnership. The period of the investigation is to be 1 January 1997 to 31 December 2006.

(i) (a) Draw and label a state-space diagram depicting the chosen model, showing possible transitions between the three states.

(b) State any assumptions implied by the diagram you have drawn and comment on their appropriateness. [3]

(ii) (a) State what data would be required in order to estimate the transition intensity of moving from state (1) to state (2) for employees aged 30 years last birthday.

(b) Write down the likelihood of these data.

(c) Derive an expression for the maximum likelihood estimate of this transition intensity.

The investigation assumes that all transition intensities are constant within each year of age. [7]

In order to estimate the corresponding transition intensity for competitors, the partnership is compelled to rely on data kept by the relevant professional institute, of which all fully qualified individuals must be members. The institute keeps data on the numbers of members actively working on 1 January each year, classified by year of birth, according to whether or not they are partners. It also keeps data on the number of members who become partners each year, classified by age in completed years upon election to partnership.

(iii) Derive, using these data, an estimate for the profession as a whole of the corresponding transition intensity of becoming a partner among persons aged 30 years last birthday during the period of the investigation. State any assumptions you make. [5]

[Total 15]

11 (i) Consider two Poisson processes, one with rate λ and the other with rate μ. Prove that the sum of events arising from either of these processes is also a Poisson process with rate $(\lambda + \mu)$. [2]

(ii) (a) Explain what is meant by a Markov jump chain.

(b) Describe the circumstances in which the outcome of the Markov jump chain differs from the standard Markov chain with the same transition matrix. [4]

An airline has N adjacent check-in desks at a particular airport, each of which can handle any customer from that airline. Arrivals of passengers at the check-in area are assumed to follow a Poisson process with rate q. The time taken to check-in a passenger is assumed to follow an exponential distribution with mean $1/a$.

(iii) Show that the number of desks occupied, together with the number of passengers waiting for a desk to become available, can be formulated as a Markov jump process and specify:

(a) the state space; and
(b) the transition diagram

[3]

(iv) State the Kolmogorov forward equations for the process, in component form. [2]

(v) Comment on the appropriateness of the assumptions made regarding passenger arrival and the check-in process. [2]

(vi) (a) Set out the transition matrix of the jump chain associated with the airline check-in process.

(b) Determine the probability that all desks are in use before any passenger has completed the check-in process, given that no passengers have arrived at check-in at the outset. [4]

[Total 17]

END OF PAPER

EXAMINATION

April 2007

Subject CT4 — Models
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 A Stocker
Chairman of the Board of Examiners

June 2007

Comments

Comments on solutions presented to individual questions for this April 2007 paper are given below and further comments, where appropriate, are given in the solutions that follow.

Question 1 *This was poorly answered by most candidates.*

Question 2 *This was reasonably well answered.*
In part (iii), many candidates did not take into account that the question related to annuities.

Question 3 *This was reasonably well answered, although many candidates took no account of the particular circumstances referred to in the question.*

Question 4 *Again, this was reasonably well answered overall.*
Many candidates failed to state the correct assumptions.

Question 5 *Overall this was poorly answered,*
Many candidates did not provide a correct definition for the hazard function.
In part (ii), marks were lost by candidates who evaluated the survival function at $t = 5$*, rather than providing the expression for* $0 \le t \le 5$*, and by those who provided graphs which were incorrectly or incompletely labelled.*

Question 6 *This was well answered by most candidates.*

Question 7 *Overall this was reasonably well answered, with the stronger candidates scoring highly.*

Question 8 *This was well answered overall.*
In part (ii), a relatively common error was to ignore the date of surgery, effectively assuming that all lives entered into the study on 1 January 2001.

Question 9 *This was reasonably well answered overall.*
As for similar questions in previous years, the main areas where candidates lost marks were: failing to provide sufficient and sufficiently clear working; failing to identify the correct degrees of freedom to be used in the chi-squared test; and failing to state relevant and clear conclusions to the tests.
Many candidates who carried out the test for individual standardised deviations failed to address the issue of outliers.
Many candidates carried out the Grouping of Signs test, which was not appropriate with so few age groups.

Question 10 *Parts (i) and (ii) were fairly well answered overall, but few candidates scored well in part (iii).*

Question 11 *This was very poorly answered by most candidates.*
The most common error in part (iii) was to give the state space as $\{0, 1, 2, \ldots, N - 1, N\}$*. Few candidates attempted part (vi).*

1 **Mixed process**

(a) Is a stochastic process that operates in continuous time, which can also change value at predetermined discrete instants.

(b) The number of contributors to a pension scheme can be modelled as a mixed process with state space $S = \{1, 2, 3, ...\}$ and time interval $J = [0, \infty]$.

Counting process

(a) Is a process, X, in discrete or continuous time, whose state space is the natural numbers {0, 1, 2, …}.

$X(t)$ is a non-decreasing function of t.

(b) Number of claims reported to an insurer by time t.

2 (i) (a) Graduation by reference to a standard table would be appropriate.

There are likely to be existing standard tables which are suitable and this method is suitable for relatively small data sets.

Alternatively, graduation by parametric formula would be suitable if the volume of data was large enough. But that is unlikely to be the case here.

Graphical graduation would not be appropriate for rates for premium calculations.

(b) Assuming graduation by reference to a standard table:

- Select a suitable table, based on a similar group of lives.
- Plot the crude rates against q_x^s from the standard table to identify a simple relationship.
- Find the best-fit parameters, using maximum likelihood or least squares estimates.
- Test the graduation for goodness of fit. If the fit is not adequate, the process should be repeated.

(ii) Considerations include:

- As the premiums are for annuity policies, it is important not to overestimate the mortality rates, as the premiums would be too low.

- The rates will be based on current mortality; the company should also take into account expected future changes, especially any reductions in mortality rates.

- Premiums charged by other insurer: if rates are too high the company will fail to attract business; if too low, it may attract too much, unprofitable business.

3 Clarify the purpose of the exercise. Why does the government want forecasts of mortality? What is the period for which the forecast is wanted? Is it short (e.g. 5–10 years) or long (e.g. 50–70 years).

Consult the existing literature on models for forecasting mortality, and speak to experts in this field of application. Consider using or adapting existing models which are employed in other countries.

Establish what data are available (e.g. on past mortality trends in the country, preferably with deaths classified by age and cause of death).

On the basis of what data are available, define the model you propose to use. If the data are simple and not detailed, then a complex model is not justified. Will a deterministic or a stochastic model be appropriate in this case?

Identify suitable computer software to implement the model, or, if none exists, write a bespoke program.

Debug the program or, if existing software is used, check that it performs the operations you intend it to do.

Run the model and test the reasonableness of the output. Consider, for example, the forecast values of quantities such as the expectation of life at birth.

Test the sensitivity of the results to changes in the input parameters.

Analyse the output.

Write a report documenting the results and the model and communicate the results and the output to the government of the small country.

4 (i) For each pensioner in the investigation, the actuary would need:

Date of entry into the investigation
(the latest of date of retirement, date of *x*th birthday and 1 January 2005)

Date of exit from the investigation
(the earliest of date of death, date of $(x+1)$th birthday and 1 January 2007)

(ii) (a) The central exposed to risk of pensioners aged *x* nearest birthday is given by

$$E_x^c = \int_0^2 P_{x,t}$$

$$\approx \sum_0^1 \tfrac{1}{2}\left(P_{x,t} + P_{x,t+1}\right) = \tfrac{1}{2} P_{x,0} + P_{x,1} + \tfrac{1}{2} P_{x,2}$$

Where $P_{x,t}$ is the number of pensioners aged *x* nearest birthday at time *t*, measured from 1 January 2005.

This assumes that $P_{x,t}$ is linear over the calendar year.

(b) This is a life year rate interval, from age $x-½$ to $x+½$. The age in the middle of the rate interval is *x*, so $\hat{\mu}$ estimates μ_x, assuming a constant force of mortality over the life year.

The estimate of μ_x is therefore given by:

$$\hat{\mu}_{65} = \frac{d_{65,2005} + d_{65,2006}}{E_{65}^c}$$

$$= \frac{16+23}{\left(\tfrac{1}{2}\times 1678 + 1720 + \tfrac{1}{2}\times 1622\right)} = \frac{39}{3370}$$

$$= 0.01157$$

5 (i) The hazard function is defined as

$$h(t) = \lim_{dt \to 0^+} \frac{1}{dt}\left(\Pr[T \le t + dt \mid T > t]\right).$$

(ii) (a) Since the survival function $S(t)$ is given by

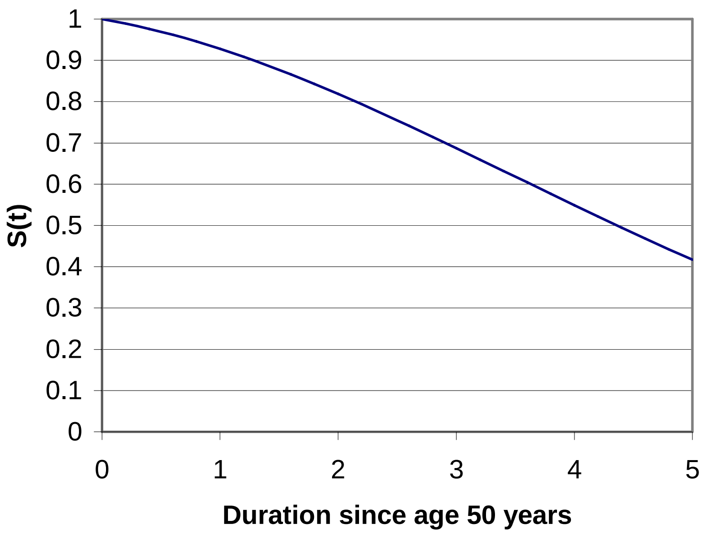
$$S(t) = \exp\left(-\int_0^t h(s)ds\right),$$

then

$$S(t) = \exp\left(-\int_0^t (\alpha + \beta s)\, ds\right) = \exp\left[-\alpha s - \frac{\beta s^2}{2}\right]_0^t = \exp\left[-\alpha t - \frac{\beta t^2}{2}\right]$$

where $0 \le t \le 5$.

(b) A suitable plot is shown below.



Both concave and convex plots were acceptable as this depends on parameters, α *and* β.

(c) If both α and β are positive, then the formula implies a force of mortality which increases with age, which is sensible for this age range.

The parameter α measures the 'level' of mortality and the parameter β measures the rate of increase with age. Varying these permits quite a wide range of forms for $S(t)$.

So the formula seems appropriate.

6 Based on the given transition diagram, the one-step transition matrix must be of the form:

$$\begin{pmatrix} a & 0 & c \\ d & e & f \\ 0 & h & i \end{pmatrix}$$

The two-step transition matrix is given by:

$$\begin{pmatrix} a & 0 & c \\ d & e & f \\ 0 & h & i \end{pmatrix} * \begin{pmatrix} a & 0 & c \\ d & e & f \\ 0 & h & i \end{pmatrix} = \begin{pmatrix} a^2 & ch & c(a+i) \\ d(a+e) & e^2+fh & cd+ef+fi \\ dh & h(e+i) & fh+i^2 \end{pmatrix}$$

$$P_{AA}^2 = 0.5625 \Rightarrow a^2 = 0.5625 \Rightarrow a = 0.75$$

Rows of transition matrix must sum to 1.

So, $a + c = 1$
and $c = 0.25$

$$P_{AB}^2 = 0.125 \Rightarrow ch = 0.125 \Rightarrow h = 0.5$$

$h + i = 1$
so i=0.5

$$P_{CC}^2 = 0.4 \Rightarrow f \times 0.5 + 0.5^2 = 0.4 \Rightarrow f = 0.3$$

$$P_{BA}^2 = 0.475 \Rightarrow d(0.75+e) = 0.475$$

Rows sum to 1 so, $d + e$ =0.7

Substitute for e:

$$d(1.45-d) = 0.475 \Rightarrow d^2 - 1.45d + 0.475 = 0$$

Solving using standard quadratic formula:

$$d = \frac{1.45 \pm \sqrt{1.45^2 - 4\times 0.475}}{2} = \frac{1.45 \pm 0.45}{2} = 0.95 \text{ or } 0.5$$

0.95 is not possible because e would need to be negative

So $d = 0.5$ and $e = 0.2$

Transition matrix is:

$$\begin{pmatrix} 0.75 & 0 & 0.25 \\ 0.5 & 0.2 & 0.3 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

7 (i) Consider the sequence of the status of the first born child in each generation.

The state space consists of the four possible combinations of chromosomes:

Female non-carrier	(FN)	*or* XX
Female carrier	(FC)	*or* X*X
Male non-sufferer	(MN)	*or* XY
Male haemophiliac	(MH)	*or* X*Y

Using the assumption that there is an equal chance of either chromosome being inherited:

- A female non-carrier will lead to a female non-carrier or male non-carrier.

- A female carrier may produce:

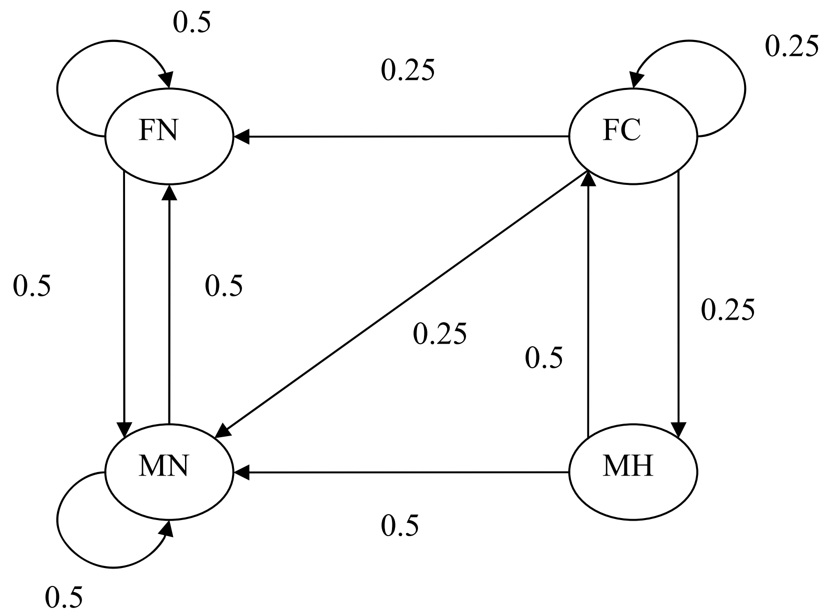
  X*X, XX, X*Y, XY all with equal probability.

- A male non-sufferer will lead to female non-carrier or male non-carrier.

- A male haemophiliac may produce:

  X*X or XY (because his partner must provide an X) with equal probability.

The transition diagram is therefore:



Each of the transition probabilities depends only on state currently occupied, so the process possesses the Markov property.

(ii) (a) The chain is reducible because once it enters states FN or MN it cannot access FC or MH.

(b) The chain is aperiodic.
As it is reducible we need to consider each group of states. FN/MN clearly have no period, and MH/FC do not either because a loop is possible in state FC.

(iii) The transition matrix is

$$A = \begin{array}{r|cccc} & FN & FC & MN & MH \\ FN(0) & 0.5 & 0 & 0.5 & 0 \\ FC(1) & 0.25 & 0.25 & 0.25 & 0.25 \\ MN(2) & 0.5 & 0 & 0.5 & 0 \\ MH(3) & 0 & 0.5 & 0.5 & 0 \end{array}$$

The stationary distribution π must satisfy:

$$\pi_0 = 0.5\pi_0 + 0.25\pi_1 + 0.5\pi_2$$
$$\pi_1 = 0.25\pi_1 + 0.5\pi_3$$
$$\pi_2 = 0.5\pi_0 + 0.25\pi_1 + 0.5\pi_2 + 0.5\pi_3$$
$$\pi_3 = 0.25\pi_1$$

So,

$$\pi_1 = 0.25\pi_1 + 0.5 \times 0.25\pi_1$$
$$\Rightarrow \pi_1 = \pi_3 = 0$$
$$\Rightarrow \pi_0 = \pi_2 = 0.5$$

An alternative solution combines the states FN and MN to give a 3-state model. This was given credit.

8 (i) (a) Type I censoring is present for those lives still under observation at 31 December 2005 as the censoring times are known in advance.

(b) Interval censoring would be present if we only knew death occurred between check-ups. However, actual dates of death are known, so interval censoring is not present.

Right censoring can be seen as a special case of interval censoring (for those censored before death, we know death occurs in the interval (c_i, ∞) where c_i is the censoring time for person i).

(c) Informative censoring is not likely to be present. The censoring of lives gives us no information about future lifetimes.

(ii) The durations at which lives died or were censored are shown below. Duration is measured in years and months from the date of surgery.

Patient	*Death or censored*	*Duration*
A	death	4 years 4 months
B	death	6 months
C	death	10 months
D	death	1 year 11 months
E	death	10 months
F	censored	4 years 11 months
G	censored	4 years 10 months
H	censored	4 years 9 months
I	censored	4 years 7 months
J	censored	4 years 4 months
K	censored	4 years 4 months
L	censored	4 years 2 months
M	censored	2 years 6 months
N	censored	9 months
O	censored	4 years

The calculation of the survival function is shown in the table below. We assume that at duration 4 years 4 months, the death occurred before lives were censored.

t_j	n_j	d_j	c_j	$\hat{\lambda}_j = d_j / n_j$
0	15	0	0	0
0.5	15	1	1	1/15
0.833	13	2	0	2/13
1.917	11	1	3	1/11
4.333	7	1	6	1/7

The estimated survival function is given by, $\hat{S}(t) = \prod_{t_j \le t} \left(1 - \lambda_j\right)$. So,

t	$\hat{S}(t)$
$0.000 \le t < 0.500$	1.0000
$0.500 \le t < 0.833$	0.9333
$0.833 \le t < 1.917$	0.7897
$1.917 \le t < 4.333$	0.7179
$4.333 \le t < 5.0$	0.6154

Solutions using different assumptions (for example assuming the death at 4 years 4 months occurred after lives were censored, or assuming lives M, N and O were censored sometime within 3 months of their last check-up) were acceptable and received credit.

(iii) The probability that a patient will die within 4 years of surgery is estimated by:

$$1-\hat{S}(4) = 1 - 0.7179 = 0.2821$$

9 (i) The chi-squared test is a suitable overall test.

The test statistic is $\sum_x z_x^{\,2}$, where

$$z_x = \frac{E_x^c \hat{\mu}_{x+1/2}^f - E_x^c \tilde{\mu}_{x+1/2}^f}{\sqrt{E_x^c \tilde{\mu}_{x+1/2}^f}}.$$

$\sum_x z_x^{\,2}$ has the χ_8^2 distribution.

The calculations are shown in the table below

Age x	*Actual deaths* $E_x^c \hat{\mu}_{x+1/2}^f$	*Expected deaths* $E_x^c \tilde{\mu}_{x+1/2}^f$	z_x	$z_x^{\,2}$
65	30	28.4	0.3002	0.0901
66	20	30.1	-1.8409	3.3890
67	25	31.2	-1.1100	1.2321
68	40	33.5	1.1230	1.2612
69	45	34.1	1.8666	3.4842
70	50	41.8	1.2683	1.6086
71	50	46.5	0.5133	0.2634
72	45	44.5	0.0750	0.0056

$\sum_x z_x^{\,2} = 11.3343.$

The critical value of the χ_8^2 distribution at the 5% level of statistical significance is 15.51.

Since 11.3343 < 15.51, we have no reason to reject the null hypothesis that the sex ratios of death rates among the company's pensioners are the same as those prevailing in the PMA92 and PFA92 tables.

(ii) **Standardised deviations test**

Using the individual standardised deviations test, we note that none of the z_xs exceeds 1.96 in absolute value, so there is no evidence that the sex ratios among the company's pensioners are unusual at any specific ages

Signs test

Under the null hypothesis of no difference between the company's pensioners and insured pensioners in general, the number of positive signs should have a Binomial (8, 0.5) distribution.

There are 2 negative and 6 positive signs.

The probability of obtaining 6 positive signs if the null hypothesis is true is

$$\binom{8}{6}0.5^8 = 0.1094$$

Since this is greater than 0.025 (two-tailed test), the sex ratios of death rates among the company's pensioners are not systematically higher or lower than those derived from the PMA92 and PFA92 tables.

Cumulative deviations test

The cumulative deviation

$$\sum_x (E_x^c \hat{\mu}_{x+1/2}^f - E_x^c \tilde{\mu}_{x+1/2}^f) \sim \text{Normal}\,(0, E_x^c \tilde{\mu}_{x+1/2}^f)\,,$$

so that under the null hypothesis

$$\frac{\sum_x (E_x^c \hat{\mu}_{x+1/2}^f - E_x^c \tilde{\mu}_{x+1/2}^f)}{\sqrt{\sum_x E_x^c \tilde{\mu}_{x+1/2}^f}} \sim \text{Normal (0,1)}.$$

Using the figures in the table above we have

$$\frac{\sum_x (E_x^c \hat{\mu}_{x+1/2}^f - E_x^c \tilde{\mu}_{x+1/2}^f)}{\sqrt{\sum_x E_x^c \tilde{\mu}_{x+1/2}^f}} = \frac{14.9}{\sqrt{290}} = 0.875$$

and since $|0.875| < 1.96$ using a two-tailed test, the sex ratios of death rates among the company's pensioners are not systematically higher or lower than those derived from the PMA92 and PFA92 tables.

Credit was only given for one of the Signs test and the Cumulative Deviations test as they both test for bias.

Serial correlations test (lag 1)

The calculations are shown in the tables below

$\overline{z}^{(1)} = \frac{1}{7}\sum_{1}^{7} z_x = 0.3029$, and $\overline{z}^{(2)} = \frac{1}{7}\sum_{2}^{8} z_x = 0.2707$

Age x	$z_x - \overline{z}^{(1)}$	$z_{x+1} - \overline{z}^{(2)}$	$(z_x - \overline{z}^{(1)})(z_{x+1} - \overline{z}^{(2)})$
65	-0.0027	-2.1117	0.0057
66	-2.1439	-1.3807	2.9601
67	-1.4129	0.8523	-1.2042
68	0.8201	1.5958	1.3087
69	1.5637	0.9976	1.5598
70	0.9654	0.2425	0.2341
71	0.2103	-0.1958	-0.0412
Sum			4.8231

Age	$\left[z_x - \overline{z}^{(1)}\right]^2$	$\left[z_{x+1} - \overline{z}^{(2)}\right]^2$
65	0.0000	4.4592
66	4.5962	1.9064
67	1.9963	0.7264
68	0.6726	2.5467
69	2.4450	0.9951
70	0.9320	0.0588
71	0.0442	0.0383
Sum	10.6863	10.7310

The correlation coefficient is therefore

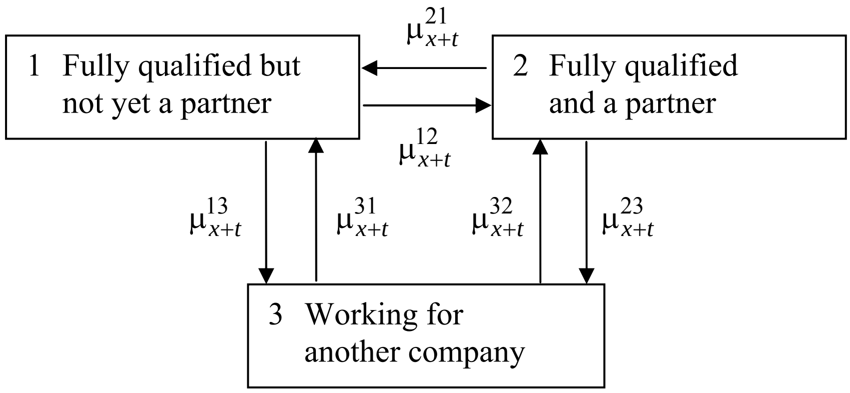
$$r_1 = \frac{4.8231}{\sqrt{(10.6863)(10.7310)}} = 0.4503$$

We test $r_1\sqrt{8}$ = 1.27 against the Normal (0,1) distribution using a one-tailed test.

Since 1.27 < 1.645, we conclude that there is no evidence that the sex ratios of death rates among the company's pensioners vary with age in a way different from the ratios derived from PMA92 and PFA92.

*Note that the **Grouping of Signs test** is not appropriate with 8 ages, 6 positive and 2 negative signs.*

10 (i) (a) A suitable diagram is shown below.



(b) The chosen model ignores death among persons in the relevant age groups. Since mortality in this age group among professional people is likely to be low, this seems reasonable.

This diagram assumes that demotion is possible, i.e. some-one who has become a partner can return to non-partnership status without leaving the company.

The assumption is also made that a new employee joining from another company can do so as a partner.

Credit was given for models based on alternative assumptions, provided these were reasonable.

(ii) (a) Assume we have data on N individuals ($i = 1, ..., N$).

We should need to know for each individual:

- the total waiting time during the calendar years 1997–2006 in state (1) when aged 30 last birthday

- whether or not the individual was made a partner between exact ages 30 and 31 years during the calendar years 1997–2006 while remaining in the company.

(b) The likelihood of the data is:

$$L = \prod_{i=1}^{N} K \exp[-(\mu^{13} + \mu^{12})v_i](\mu^{12})^{d_i}$$

where

v_i is the waiting time at age 30 last birthday in state (1) for individual *i*.

d_i is an indicator variable such that $d_i = 1$ if individual *i* was made a partner while aged 30 last birthday during the period of the investigation and $d_i = 0$ otherwise.

K is a constant denoting terms that do not depend on μ^{12}.

(c) The logarithm of the likelihood is

$$\log_e L = \sum_{i=1}^{N} \log_e K - (\mu^{12} + \mu^{13})v_i + d_i \log_e \mu^{12}$$

Differentiating this with respect to μ^{12} we obtain

$$\frac{\partial \log_e L}{\partial \mu^{12}} = -\sum_{i=1}^{N} v_i + \frac{\sum_{i=1}^{N} d_i}{\mu^{12}},$$

and setting this equal to zero and solving for μ^{12} gives

$$\hat{\mu}^{12} = \frac{\sum_{i=1}^{N} d_i}{\sum_{i=1}^{N} v_i}.$$

This is the maximum likelihood estimate, as can be seen by noting that

$$\frac{\partial^2 \log_e L}{(\partial \mu^{12})^2} = -\frac{\sum_{i=1}^{N} d_i}{(\mu^{12})^2}$$ which must be negative.

(iii) The data on becoming a partner are classified by age last birthday, which is the same classification as used in the company's own investigation, therefore the relevant intensities will relate to the same age range.

For the correct exposed to risk we only consider those who are members of the institute but not yet partners.

Let the number of such members in the census in year t who were born in year s be $P_{t,s}$.

All persons born in year s would be aged x last birthday on 1 January in year $s+x+1$.

Therefore, assuming that the $P_{t,s}$ change linearly during each calendar year the correct exposed to risk for the year 1997 is

$$\frac{1}{2}(P_{1997,1956} + P_{1998,1957})$$

and the exposed to risk for the entire 10-year period of the investigation is

$$\sum_{t=1997}^{t=2006} \frac{1}{2}(P_{t,t-31} + P_{t+1,t-30}).$$

If the number of persons becoming partners aged 30 last birthday in year t is θ_t, then an estimate of the relevant transition intensity is

$$\frac{\sum_{t=1997}^{t=2006} \theta_t}{\sum_{t=1997}^{t=2006} \frac{1}{2}(P_{t,t-31} + P_{t+1,t-30})}.$$

11 (i) Consider a small time interval dt

The probability of an arrival from the first process in time dt is $\lambda.dt + o(dt)$ and the probability of a arrival from the second process in time dt is $\mu.dt + o(dt)$.

The arrival probability for the sum of the processes in dt is therefore $(\lambda + \mu).dt + o(dt)$

This is by definition a Poisson process with rate $(\lambda + \mu)$.

Alternative solutions, based on the Moment Generating Function or the Probability Generating Function of a Poisson distribution were acceptable.

(ii) (a) A jump chain is formed by recording the state of a Markov jump process only at the instant when a transition has just been made.

The jump chain is in itself a Markov chain.

(b) The outcome of the jump chain can only differ from that of the standard Markov chain if the jump process enters an absorbing state.

As the jump process will make no further transitions once it enters an absorbing state, the jump chain "stops".

It is possible to model the jump chain as though transitions continue to occur but the chain continues to occupy the same state.

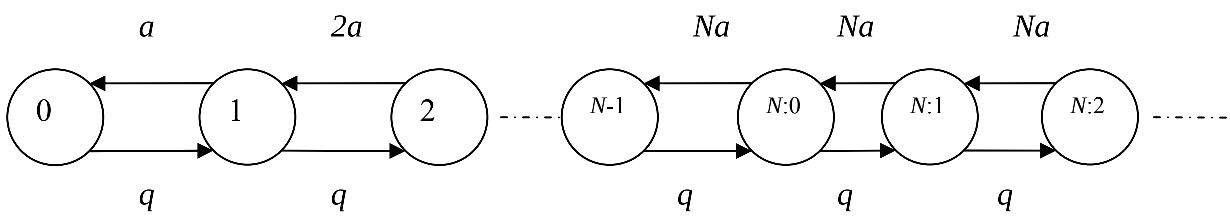
(iii) The possible states are 0 to N desks in use with no passengers queuing, and N desks in use with 0, 1, 2, ….. passengers in the queue.

When all desks are occupied and there are M passengers in the queue denote the state as N:M.

State space is:

{0, 1, 2, …., N - 1, N : 0, N : 1, N : 2, …..}

Transition diagram:



(iv) Kolmogorov forward equations in component form are:

$$\frac{d}{dt}P_0(t) = aP_1(t) - qP_0(t)$$

$$\frac{d}{dt}P_r(t) = a(r+1)P_{r+1}(t) + qP_{r-1}(t) - (ar+q)P_r(t) \qquad r + 1 \leq N$$

$$\frac{d}{dt}P_{N:0}(t) = aNP_{N:1}(t) + qP_{N-1}(t) - (aN+q)P_{N:0}(t)$$

$$\frac{d}{dt}P_{N:m}(t) = aNP_{N:m+1}(t) + qP_{N:m-1}(t) - (aN+q)P_{N:m}(t) \qquad m \geq 1$$

(v) Poisson process is usually suitable for arrivals at a service point.

Rate may be time inhomogeneous because passengers may aim to arrive a couple of hours before the flight — so a time-inhomogeneous Poisson process may be better.

However if the airline operates many flights this may not be an issue.

Passengers may be checked-in in family groups rather than individually.

There is likely to be a minimum time for processing a check-in due to standard security questions etc, so exponential distribution may not hold.

(vi) (a) The transition matrix is:

$$\begin{pmatrix} 0 & 1 & & & & & & \\ \frac{a}{a+q} & 0 & \frac{q}{a+q} & & & & & \\ & \frac{2a}{2a+q} & 0 & \frac{q}{2a+q} & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & \frac{Na}{Na+q} & 0 & \frac{q}{Na+q} & & \\ & & & & \frac{Na}{Na+q} & 0 & \frac{q}{Na+q} & \\ & & & & & \ddots & \ddots & \end{pmatrix}$$

(b) This is the probability that all the first N transitions are to the right in the transition diagram.

The probability of each transition is given by the elements in the upper half of the jump chain transition matrix in (vi)(a).

Required probability is therefore $q^{N-1}.\prod_{i=1}^{N-1}\frac{1}{ia+q}$

END OF EXAMINERS' REPORT

EXAMINATION

3 October 2007 (am)

Subject CT4 — Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*
2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*
3. *Mark allocations are shown in brackets.*
4. *Attempt all 11 questions, beginning your answer to each question on a separate sheet.*
5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 List the factors you would consider when assessing the suitability of an actuarial model for its purpose. [4]

2 A particular baker's shop in a small town sells only one product: currant buns. These currant buns are delicious and customers travel many miles to buy them. Unfortunately, the buns do not keep fresh and cannot be stored overnight.

The baker's practice is to bake a certain number of buns, K, before the shop opens each morning, and then during the day to continue baking c buns per hour. He is concerned that:

- he does not run out of buns during the day; and
- the number of buns left over at the end of each day is as few as possible

(i) Describe a model which would allow you to estimate the probability that the baker will run out of buns. State any assumptions you make. [3]

(ii) Determine the relevant expression for the probability that the baker will run out of buns, in terms of K, c, and B_j, the number of buns bought by the day's jth customer. [1]

[Total 4]

3 A no-claims discount system has 3 levels of discount: 0%, 25% and 50%. The rules for moving between discount levels are:

- After a claim-free year, move up to the next higher level or remain at the 50% discount level.

- After a year with one or more claims, move down to the next lower level or remain at the 0% discount level.

The long-run probability that a policyholder is in the maximum discount level is 0.75.

Calculate the probability that a given policyholder has a claim-free year, assuming that this probability is constant.

[5]

4 A national mortality investigation was carried out. It was suggested that the mortality of the male population could be represented by the following graduated rates:

$$\overset{o}{\mu}_{x+\frac{1}{2}} = \mu^{s}_{x+2\frac{1}{2}}$$

where μ^s_x is from the standard tables, ELT15(males).

The table below shows the graduated rates for part of the age range, together with the exposed to risk, expected and actual deaths at each age. The squared standardised deviations that were calculated are also shown.

The standardised deviations were calculated as $z_x = \dfrac{\left(\theta_x - E^c_x \cdot \overset{o}{\mu}_{x+\frac{1}{2}}\right)}{\sqrt{E^c_x \cdot \overset{o}{\mu}_{x+\frac{1}{2}}}}$

Age	*Graduated rates*	*Exposed to risk*	*Expected deaths*	*Deaths*	*Squared standardised deviations*
x	$\overset{o}{\mu}_{x+\frac{1}{2}}$	E^c_x	$E^c_x \cdot \overset{o}{\mu}_{x+\frac{1}{2}}$	θ_x	z^2_x
50	0.00549	10,850	59.57	52	0.9611
51	0.00610	9,812	59.85	54	0.5724
52	0.00679	10,054	68.27	60	1.0010
53	0.00757	9,650	73.05	65	0.8872
54	0.00845	8,563	72.36	64	0.9653
55	0.00945	10,656	100.70	87	1.8637
56	0.01057	9,667	102.18	88	1.9679
57	0.01182	9,560	113.00	97	2.2653
58	0.01323	8,968	118.65	103	2.0634
59	0.01483	8,455	125.39	105	3.3150

(i) Test this graduation for overall goodness-of-fit. [5]

(ii) Comment on your findings in (i). [2]
[Total 7]

5 (i) Explain why crude mortality rates are graduated before being used for financial calculations. [3]

(ii) List two methods of graduating a set of crude mortality rates and state, for each method:

(a) under what circumstances it should be used; and
(b) how smoothness is ensured

[4]
[Total 7]

6 Below is an extract from English Life Table 15 (Males)

Age x	l_x
58	88,792
62	84,173

(i) Estimate l_{60} under each of the following assumptions:

(a) a uniform distribution of deaths between exact ages 58 and 62 years; and

(b) a constant force of mortality between exact ages 58 and 62 years

[5]

(ii) Find the actual value of l_{60} in the tables and hence comment on the relative validity of the two assumptions you used in part (i). [3]
[Total 8]

7 In order to boost sales, a national newspaper in a European country wishes to compile a "fair play league table" for the country's leading football clubs. On 1 December it undertakes a survey of all the players who play for these clubs, in which it collects the following data:

- number of games played by each player since the beginning of the season (the football season in this country begins in September); and

- for each player who had been dismissed from the field of play between the beginning of the season and 1 December (inclusive), the number of games he had played before the game in which he was first dismissed

No games were played on 1 December.

The statistic the newspaper proposes to use in order to construct its "fair play league table" is the probability that a player will not have been dismissed in any of his first 10 games. It plans to calculate this statistic for each of the 20 leading clubs.

The following table shows the data collected for the players of the club which was top of the league on 1 December.

Player	*Total number of games played*	*Number of times dismissed*	*Games played before first dismissal*
1	12	0	
2	12	0	
3	12	1	5
4	12	0	
5	12	1	7
6	12	0	
7	10	0	
8	9	1	0
9	9	1	5
10	8	0	
11	6	2	2
12	5	0	
13	5	0	
14	4	1	0
15	4	0	

(i) (a) Explain how the Kaplan-Meier estimator can be used to estimate the newspaper's statistic from these data.

(b) Comment on the way in which censoring arises and on the type of censoring produced. [4]

(ii) Calculate the newspaper's statistic using the data above. [4]

[Total 8]

8 (i) Describe the difference between the central exposed to risk and the initial exposed to risk. [2]

The following data come from an investigation of the mortality of participants in a dangerous sport during the calendar year 2005.

Age x	*Number of lives aged x last birthday on:* 1 January 2005	1 January 2006	*Number of deaths during 2005 to persons aged x last birthday at death*
22	150	160	20
23	160	155	25

(ii) (a) Estimate the initial exposed to risk at ages 22 and 23.
(b) Hence estimate q_{22} and q_{23}.
[4]

Suppose that in this investigation, instead of aggregate data we had individual-level data on each person's date of birth, date of death, and date of exit from observation (if exit was for reasons other than death).

(iii) Explain how you would calculate the initial exposed-to-risk for lives aged 22 years last birthday. [4]
[Total 10]

9 In a game of tennis, when the score is at "Deuce" the player winning the next point holds "Advantage". If a player holding "Advantage" wins the following point that player wins the game, but if that point is won by the other player the score returns to "Deuce".

When Andrew plays tennis against Ben, the probability of Andrew winning any point is 0.6. Consider a particular game when the score is at "Deuce".

(i) Show that the subsequent score in the game can be modelled as a Markov Chain, specifying both:

(a) the state space; and
(b) the transition matrix [3]

(ii) State, with reasons, whether the chain is:

(a) irreducible; and
(b) aperiodic [2]

(iii) Calculate the number of points which must be played before there is more than a 90% chance of the game having been completed. [3]

(iv) (a) Calculate the probability that Andrew wins the game.
(b) Comment on your answer. [4]
[Total 12]

10 (i) Compare the advantages and disadvantages of fully parametric models and the Cox regression model for assessing the impact of covariates on survival. [3]

You have been asked to investigate the impact of a set of covariates, including age, sex, smoking, region of residence, educational attainment and amount of exercise undertaken, on the risk of heart attack. Data are available from a prospective study which followed a set of several thousand persons from an initial interview until their first heart attack, or until their death from a cause other than a heart attack, or until 10 years had elapsed since the initial interview (whichever of these occurred first).

(ii) State the types of censoring present in this study, and explain how each arises. [2]

(iii) Describe a criterion which would allow you to select those covariates which have a statistically significant effect on the risk of heart attack, when controlling the other covariates of the model. [4]

Suppose your final model is a Cox model which has three covariates: age (measured in age last birthday minus 50 at the initial interview), sex (male = 0, female = 1) and smoking (non-smoker = 0, smoker = 1), and that the estimated parameters are:

Age	0.01
Sex	−0.4
Smoking	0.5
Sex *x* smoking	−0.25

where “sex *x* smoking” is an additional covariate formed by multiplying the two covariates “sex” and “smoking”.

(iv) Describe the final model’s estimate of the effect of sex and of smoking behaviour on the risk of heart attack. [3]

(v) Use the results of the model to determine how old a female smoker must be at the initial interview to have the same risk of heart attack as a male non-smoker aged 50 years at the initial interview. [3]

[Total 15]

11 The following data have been collected from observation of a three-state process in continuous time:

State occupied	*Total time spent in state (hours)*	*Total transitions to:* *State A*	*State B*	*State C*
A	50	Not applicable	110	90
B	25	80	Not applicable	45
C	90	120	15	Not applicable

It is proposed to fit a Markov jump model to this data set.

(i) (a) List all the parameters of the model.
(b) Describe the assumptions underlying the model. [4]

(ii) (a) Estimate the parameters of the model.
(b) Give the estimated generator matrix. [4]

The following additional data in respect of secondary transitions were collected from observation of the same process.

Triplet of successive transitions	*Observed number of triplets* n_{ijk}	*Triplet of successive transitions*	*Observed number of triplets* n_{ijk}
ABC	42	BCA	38
ABA	68	BCB	7
ACA	85	CAB	64
ACB	4	CAC	56
BAB	50	CBA	8
BAC	30	CBC	7

(iii) State the distribution of the number of transitions from state i to state j, given the number of transitions out of state i. [1]

(iv) Test the goodness-of-fit of the model by considering whether triplets of successive transitions adhere to the distribution given in (iii). [5]

[Hint: Use the test statistic $\chi^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - E)^2}{E}$ where E is the expected number of triplets under the distribution in (iii)]

(v) Identify two other aspects of the appropriateness of the fitted model that could be tested, stating suitable tests in each case. [2]

(vi) Outline two methods for simulating the Markov jump process, without performing any calculations. [4]

[Total 20]

END OF PAPER

EXAMINATION

September 2007

Subject CT4 — Models
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 A Stocker
Chairman of the Board of Examiners

December 2007

Comments

Comments on solutions presented to individual questions for this September 2007 paper are given below and further comments may be written within in the solutions that follow.

Q1 *This straightforward bookwork question was not especially well answered.*

Q2 *This was the most poorly answered question on the examination paper. Very few candidates recognised that the baker's problem could be modelled using the compound Poisson process described in Unit 2 section 3.4 of the Core Reading.*

Q3 *This was well answered, with many candidates scoring full marks.*

Q4 *Although most candidates performed the chi-squared test correctly, few realised that when using this to test a graduation some degrees of freedom are lost, a fact which is clearly stated in the Core Reading in Unit 12, section 7.3. In part (ii) comments tended not to be related to the data in the question; rather they focused rather mechanically on the shortcomings of the chi-squared test.*

Q5 *This straightforward bookwork question was well answered by many candidates.*

Q6 *Most candidates obtained the correct numerical answers in part (i) of this question, but answers to part (ii) were rather sketchy and vague.*

Q7 *This was more demanding than some previous questions on the Kaplan-Meier or Nelson-Aalen estimators, and the standard of the answers was lower than expected.*

Q8 *This exposed-to-risk question was easier than many questions on the same topic in previous papers. Most candidates scored well on parts (i) and (ii), although few explained that the method relied on the assumption of a uniform distribution of deaths. Answers to part (iii) were less impressive and tended to lack detail. Some candidates couched their answers to this part in aggregate terms, despite the question clearly referring to individual-level data.*

Q9 *Most candidates scored well on parts (i) and (ii). Common errors included the use of a three-state model (Deuce, Advantage and Game) which is inappropriate as the transition out of the state "Advantage" is ill-defined. Few candidates made attempts at parts (iii) and (iv) and several of these wrongly thought that part (iv) could be solved by finding the stationary distribution of the chain.*

Q10 *Parts (i) and (iv) of this question tested knowledge of Unit 7, sections 2, 3 and 5 of the Core Reading, which has not been tested in previous CT4 examination papers. Perhaps because of this, many candidates gave very sketchy and vague answers. In part (ii), while most candidates spotted that Type I censoring was present, only a small minority also registered the existence of random censoring. In part (iii) few candidates correctly interpreted the sex x smoking interaction. Part (v) was well answered by most candidates.*

Q11 *Many candidates only attempted parts (i) and (ii) of this question. The remainder was very poorly answered, with few candidates making serious attempts at part (vi), despite this being bookwork based on Core Reading, Unit 4, section 5.4.*

1 Factors to be considered include:

- the objectives of the modelling exercise,
- the validity of the model for the purpose to which it is to be put,
- the validity of the data to be used,
- the possible errors associated with the model or parameters used not being a perfect fit,
- representation of the real world situation being modelled,
- the impact of correlations between the random variables that drive the model,
- the extent of correlations between the various results produced from the model,
- the current relevance of models written and used in the past,
- the credibility of the data input,
- the credibility of the results output,
- the dangers of spurious accuracy,
- the ease with which the model and its results can be communicated.

Not all these factors needed to be mentioned for full marks to be awarded.

2 (a) Assume that, during each day, customers arrive at the shop according to a Poisson process.

Assume that the numbers of buns bought by each customer, the B_j, are independent and identically distributed random variables.

Then if X_t is the total number of buns sold between the beginning of the day and time t, (where t is measured in hours since the shop opens), X_t is a compound Poisson process defined by

$$X_t = \sum_{j=1}^{N_t} B_j ,$$

where the number of customers arriving between the shop opening and time t is N_t .

(b) The probability that the baker will run out of buns is

$$\Pr[K + ct - \sum_{j=1}^{N_t} B_j < 0]$$

for some t.

3 The transition matrix for the chain is:

$$\begin{pmatrix} 1-\alpha & \alpha & \\ 1-\alpha & & \alpha \\ & 1-\alpha & \alpha \end{pmatrix}.$$

To determine the long-run probability, we need to solve the equation $\pi P = \pi$, which reads:

(I) $\pi_1 = (1-\alpha)\pi_1 + (1-\alpha)\pi_2$

(II) $\pi_2 = \alpha\pi_1 + (1-\alpha)\pi_3$

(III) $\pi_3 = \alpha\pi_2 + \alpha\pi_3$.

The probabilities must also satisfy:

(IV) $\pi_1 + \pi_2 + \pi_3 = 1$.

(III) gives $\pi_2 = \left(\dfrac{1-\alpha}{\alpha}\right)\pi_3$.

Substituting in (I) gives $\pi_1 = \left(\dfrac{1-\alpha}{\alpha}\right)^2 \pi_3$,

and so (IV) leads to $\left(\left(\dfrac{1-\alpha}{\alpha}\right)^2 + \left(\dfrac{1-\alpha}{\alpha}\right) + 1\right)\pi_3 = 1$.

We know that $\pi_3 = 0.75$, which leads to:

$$\left(\frac{(1-\alpha)^2 + \alpha(1-\alpha) + \alpha^2}{\alpha^2}\right) \times 0.75 = 1,$$

$$\Rightarrow 0.75\left(\left(1-2\alpha+\alpha^2\right) + \left(\alpha+\alpha^2\right) + \alpha^2\right) = \alpha^2,$$

$$\Rightarrow 0.25\alpha^2 + 0.75\alpha - 0.75 = 0 \ .$$

Using the quadratic equation formula, this leads to

$$\alpha = \frac{-0.75 \pm \sqrt{0.75^2 + 4 \times 0.25 \times 0.75}}{2 \times 0.25}.$$

As $\alpha > 0$, we must have $\alpha = 0.7913$.

4 (i) The null hypothesis is that graduated rates are the same as the true underlying rates in the population.

To test overall goodness-of-fit we use the chi-squared test.

$\sum_x z_x^2 \sim \chi^2_m$, where m is the number of degrees of freedom.

In this case, we have 10 ages.

The graduation was carried out by reference to a standard table, so we lose a number of degrees of freedom because of the choice of standard table.

So, $m < 10$, and let us say $m = 8$.

The observed value of the test statistic is $\sum_x z_x^2 = 15.8623$

The critical value of the chi-squared distribution with 8 degrees of freedom at the 5 per cent level is 15.51.

Since 15.8623 > 15.51,

we reject the null hypothesis and conclude that the graduated rates do not adhere to the data.

[Credit was given for using other values of m, say m = 7 or m = 9, provided candidates recognized that some degrees of freedom should be lost for the choice of standard table. Note that if m = 9, the null hypothesis will not be rejected.]

(ii) From the data we can see that the actual deaths are lower than those expected at all ages.

The graduated rates are too high; the graduation should be revisited.

At these ages the force of mortality increases with age,
so a suitable adjustment may be to reduce the age shift relative to the standard table from 2 years.

The standardised deviations also appear to show a systematic increase with age, showing that departure of the graduated rates from the actual rates increases with age.

There appear to be no outliers (all the z_xs have absolute values below 1.96).

5 (i) We assume that mortality rates progress smoothly with age.

Therefore a crude estimate at age x carries information about the rates at adjacent ages, and graduation allows us to use this fact to "improve" the estimate at age x by smoothing.

This reduces the sampling errors at each age.

It is desirable that financial quantities progress smoothly with age, as irregularities are hard to justify to clients.

(ii) Any two of the following three methods are acceptable:

By parametric formula:

Should be used for large experiences, especially if the aim is to produce a standard table;

Depends on a suitable formula being found which fits the data well.

Provided the number of parameters is small, the resulting curve should be smooth.

With reference to a standard table

Should be used if a standard table for a class of lives similar to the experience is available, and the experience we are interested in does not provide much data.

The standard table will be smooth,

and provided the function linking the graduated rates to the rates in the standard table is simple, this smoothness will be "transferred to the graduated rates".

Graphical

if a quick check is needed, or data are very scanty.

The graduation should be tested for smoothness using the third differences of the graduated rates, which should be small in magnitude and progress regularly with age.

If the smoothness is unsatisfactory, the curve can be adjusted ("hand-polishing") and the smoothness tested again.

6 (i) (a) Assuming a uniform distribution of deaths between ages 58 and 62 implies that half of those who die between those ages die between ages 58 and 60.

Therefore

$$l_{60} = l_{58} - 0.5(l_{58} - l_{62})$$

$$= 88{,}792 - 0.5(88{,}792 - 84{,}173)$$

$$= 86{,}482.5.$$

(b) ALTERNATIVE 1

Let the constant force of mortality be μ.

Then we have ${}_4p_{58} = \exp\left(-\int_0^4 \mu dx\right) = e^{-4\mu}$.

But ${}_4p_{58} = \dfrac{l_{62}}{l_{58}} = \dfrac{84{,}173}{88{,}792} = 0.94798$.

Therefore $e^{-4\mu} = 0.94798$,

so that $-4\mu = \log_e(0.94798) = -0.05342$,

whence $\mu = 0.01336$.

Therefore with a constant force of mortality,

$$l_{60} = l_{58} \exp[-2(0.01336)] = 88{,}792(0.97363)$$

so $l_{60} = 86{,}452$.

ALTERNATIVE 2

Let the constant force of mortality be μ.

Then we have ${}_4p_{58} = \exp\left(-\int_0^4 \mu dx\right) = e^{-4\mu}$.

But ${}_4p_{58} = \dfrac{l_{62}}{l_{58}}$.

Now $l_{60} = l_{58} \cdot {}_2p_{58}$.

and, since ${}_2p_{58} = e^{-2\mu} = \sqrt{e^{-4\mu}} = \sqrt{\frac{l_{62}}{l_{58}}}$,

$$l_{60} = l_{58}\sqrt{\frac{l_{62}}{l_{58}}} = \sqrt{l_{58}l_{62}} = \sqrt{(88,792)(84,173)}$$

so l_{60} = 86,452

(ii) The actual value of l_{60} from the tables is 86,714.

This shows that neither assumption is very accurate, but that the uniform distribution of deaths (UDD) is closer than the constant force of mortality.

The UDD assumption is better than the constant force of mortality assumption because UDD implies an increasing force of mortality over this age range, which is biologically more plausible than the assumption of a constant force.

The fact that the actual value of l_{60} is considerably greater than that implied by the UDD assumption suggests that the true rate of increase of the force of mortality over this age range in English Life Table 15 (males) is even greater than that implied by UDD.

7 (i) (a) If, for player i, T_i is the number of games played before he is dismissed, and C_i is the total number of games played before 1 December, and d_i = 1 if the player had been dismissed before 1 December and 0 otherwise.

then

EITHER

from the data given we can create the two variables

min(T_i,C_i)
and d_i,
e.g. for player 1, min(T_i,C_i) = 12 and d_i = 0

OR

The required data for the Kaplan-Meier estimator are therefore

Player	*min*(T_i, C_i)	d_i
1	12	0
2	12	0
3	5	1
4	12	0
5	7	1
6	12	0
7	10	0
8	0	1
9	5	1
10	8	0
11	2	1
12	5	0
13	5	0
14	0	1
15	4	0

(b) Censoring in these data arises because not all players have been dismissed before 1 December. Those players who have yet to be dismissed on that data are right-censored.

This censoring is random [NOT Type I], because the metric of "duration" is the number of games played since the start of the season, and this may vary from player to player.

(ii) ALTERNATIVE 1 (where censorings are assumed to occur immediately before events)

t_j	N_j	D_j	C_j	$\frac{D_j}{N_j}$	$1-\frac{D_j}{N_j}$
0	15	2	0	2/15	13/15
2	13	1	3	1/13	12/13
5	9	2	0	2/9	7/9
7	7	1	6	1/7	6/7

Then the Kaplan-Meier estimate of the survival function is

t	$\hat{S}(t)$
$0 \le t < 2$	0.8667
$2 \le t < 5$	0.8000
$5 \le t < 7$	0.6222
$7 \le t < 12$	0.5333

Therefore the value of the chosen statistic, $\hat{S}(10)$ is 0.5333.

ALTERNATIVE 2 (where censorings are assumed to occur immediately after events)

t_j	N_j	D_j	C_j	$\frac{D_j}{N_j}$	$1-\frac{D_j}{N_j}$
0	15	2	0	2/15	13/15
2	13	1	1	1/13	12/13
5	11	2	2	2/11	9/11
7	7	1	6	1/7	6/7

Then the Kaplan-Meier estimate of the survival function is

t	$\hat{S}(t)$
$0 \le t < 2$	0.8667
$2 \le t < 5$	0.8000
$5 \le t < 7$	0.6545
$7 \le t < 12$	0.5610

Therefore the value of the chosen statistic, $\hat{S}(10)$ is 0.5610.

8 (i) The central exposed to risk at age x, E_x^c, is the observed waiting time in a multiple-state or a Poisson model. It is the sum of the times spent under observation by each life at age x.

In aggregate data, the central exposed to risk is an estimate of the number of lives exposed to risk at the mid-point of the rate interval.

The initial exposed to risk requires adjustments for those lives who die, whom we continue observing until the end of the rate interval.

It may be approximated as $E_x^c + 0.5d_x$, where d_x is the number of deaths to persons aged x.

(ii) The age definition used for both deaths and exposed to risk is the same, so no adjustment is necessary.

Using the census formula, and assuming that the population aged 22 and 23 years changes linearly over the year, we have, for the central exposed to risk:

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

so that

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

The initial exposed to risk, E_x , is then obtained using the approximation $E_x^c + 0.5d_x$.

This assumes that deaths are uniformly distributed across each year of age.

Therefore, at age 22 we have

$$E_{22} = \frac{1}{2}(150+160) + \frac{20}{2} = 165 ,$$

and

$$E_{23} = \frac{1}{2}(160+155) + \frac{25}{2} = 170 .$$

Hence $q_{22} = \frac{20}{165} = 0.1212$ and $q_{23} = \frac{25}{170} = 0.1471$.

[The complete derivation was not required for full marks.]

(iii) ALTERNATIVE 1

The central exposed to risk is calculated as $\sum_i (b_i - a_i)$, for all lives *i* for whom $b_i - a_i > 0$,

where a_i and b_i are measured in years since the person's 22nd birthday, and

where b_i is the earliest of

the date of person *i*'s death
the date of person *i*'s 23rd birthday
the end of the calendar year 2005
the date of person *i*'s exit from observation for reasons other than death

and a_i is the latest of

the date of person *i*'s 22nd birthday
the start of the calendar year 2005
the date of person *i*'s entry into observation.

The initial exposed to risk is then calculated by adding on to the central exposed to risk a quantity equal to $1 - b_i$ for all lives who died aged 22 last birthday during the calendar year 2005.

ALTERNATIVE 2

The initial exposed to risk is calculated as $\sum_i (b_i - a_i)$,

where a_i and b_i are measured in years since the person's 22nd birthday, and

where b_i is the earliest of

the date of person i's 23rd birthday
the date of person i's exit from observation for reasons other than death

and a_i is the latest of

the date of person i's 22nd birthday
the start of the calendar year 2005
the date of person i's entry into observation.

for all lives i for whom $b_i - a_i > 0$.

9 (i) State space:

{Deuce, Advantage A(ndrew), Advantage B(en), Game A(ndrew), Game B(en)}.

Transition matrix:

	Deuce	*Adv A*	*Adv B*	*Game A*	*Game B*
Deuce	0	0.6	0.4	0	0
Adv A	0.4	0	0	0.6	0
Adv B	0.6	0	0	0	0.4
Game A	0	0	0	1	0
Game B	0	0	0	0	1

The chain is Markov because the probability of moving to the next state does not depend on history prior to entering that state (because the probability of each player winning a point is constant)

(ii) The chain is reducible because it has two absorbing states Game A and Game B.

States Game A and Game B are absorbing so have no period. The other three states each have a period of 2 so the chain is not aperiodic.

(iii) The game either ends after 2 points or it returns to Deuce.

The probability of it returning to Deuce after two points is:

Prob A wins 1st point × Prob B wins 2nd point
+ Prob B wins 1st point × Prob A wins 2nd point

$= 0.6 \times 0.4 + 0.4 \times 0.6 = 0.48.$

[This can also be obtained by calculating the square of the transition matrix.]

Need to find number of such cycles N such that:

$$0.48^N < 1 - 0.9,$$

so that

$$N > \frac{\ln 0.1}{\ln(0.48)} > 3.14.$$

But the game can only finish every two points so we require 4 cycles, that is 8 points.

(iv) (a) Define A_X to be the probability that A ultimately wins the game when the current state is X.

We require A_{Deuce}.

By definition $A_{\text{Game A}} = 1$ and $A_{\text{Game B}} = 0$.

Conditioning on the first move out of state Adv A:

$$A_{\text{Adv A}} = 0.6 \times A_{\text{Game A}} + 0.4 \times A_{\text{Deuce}} = 0.6 + 0.4 \times A_{\text{Deuce}}.$$

Similarly:

$$A_{\text{Adv B}} = 0.6 \times A_{\text{Deuce}},$$

and

$$A_{\text{Deuce}} = 0.6 \times A_{\text{Adv A}} + 0.4 \times A_{\text{Adv B}} = 0.6 \times A_{\text{Adv A}} + 0.24 \times A_{\text{Deuce}}.$$

So,

$$A_{\text{Deuce}} = \frac{0.6}{0.76} A_{\text{Adv A}},$$

$$A_{\text{Adv A}} = 0.6 + 0.4 \times \frac{0.6}{0.76} A_{\text{Adv A}},$$

and

$$A_{\text{Adv A}} = 0.8769,$$

and

$$A_{\text{Deuce}} = 0.6923.$$

ALTERNATIVELY

Probability A wins after 2 points = 0.6*0.6 =0.36

Probability that A wins from Deuce
= $\sum_{i=1}^{\infty}$ Probability A wins after i points have been played

= Probability A wins after 2 points
+ Probability A wins after 4 points +…..
(as period 2)

= 0.36 + 0.48 * 0.36 + 0.48^2 * 0.36 +…….

= 0.36/(1-0.48) as a geometric progression

= 0.6923

(b) This is higher than 0.6 because Ben has to win at least two points in a row to win the game.

10 (i) Fully parametric models are good for comparing homogenous groups, as confidence intervals for the fitted parameters give a test of difference between the groups which should be better than non-parametric procedures, or semi-parametric procedures such as the Cox model.

But parametric methods need foreknowledge of the form of the hazard function, which might be the object of the study.
The Cox model is semi-parametric so such knowledge is not required.

The Cox model is a standard feature of many statistical packages for estimating survival model, but many parametric distributions are not, and numerical methods may be required, entailing additional programming.

(ii) Type I censoring, since the investigation ends after a period which is fixed in advance.

Random censoring, since death from a cause other than a heart attack is a random variable and may occur at any time.

(iii) The likelihood ratio statistic is a common criterion.

Suppose we fit a model with p covariates and another model with $p+q$ covariates which include all the p covariates of the first model.

Then if the maximised log-likelihoods of the two models are L_p and L_{p+q}, then the statistic

$$-2(L_p - L_{p+q})$$

has a chi-squared distribution with q degrees of freedom, under the hypothesis that the extra q covariates have no effect in the presence of the original p covariates.

This statistic can be used either will full likelihoods or with partial likelihoods in the Cox model

This statistic can be used to test the statistical significance of any set of q covariates in the presence of any other disjoint set of p covariates.

(iv) Holding other factors constant,

females have a lower risk of heart attack than males,

and smokers have a higher risk than non-smokers,

but the effect of smoking varies for men and women.

The relative risks, compared with the baseline category of male non-smokers are as follows.

female non-smokers	$\exp(-0.4)$	$= 0.67$
male smokers	$\exp(0.5)$	$= 1.65$
female smokers	$\exp(-0.4+0.5-0.25)$	$= 0.86$

(or any other numerical example to illustrate the previous points)

(v) Let the required age for the woman smoker be $50+x$.

The hazard for this woman is

$h(t,x) = h_0(t)\exp(0.01x - 0.4 + 0.5 - 0.25)$,

The hazard for a male non-smoker aged 50 at the initial interview is simply $h_0(x)$, since this is the baseline category.

Thus we have

$h_0(t)\exp(0.01x - 0.4 + 0.5 - 0.25) = h_0(t)$

so that

$\exp(0.01x - 0.4 + 0.5 - 0.25) = 1$

or

$\exp(0.01x - 0.15) = 1$

so that

$0.01x = 0.15$

Therefore $x = 15$, and the woman's age at interview must be 65 years.

11 (i) (a) The parameters are:

- the rate of leaving state i, λ_i, for each i,
- the jump-chain transition probabilities, r_{ij}, for $j \neq i$, where r_{ij} is the conditional probability that the next transition is to state j given the current state is i.

[Alternatively the parameters may be expressed as σ_{ij}, where $\sigma_{ii} = -\lambda_i$ and (for $j \neq i$), $\sigma_{ij} = \lambda_i r_{ij}$.]

(b) The assumptions are as follows.

- The holding time in each state is exponentially distributed. The parameter of this distribution varies only by state i. The distribution is independent of anything that happened prior to the current arrival in state i.

- The destination of the jump on leaving state i is independent of holding time, and of anything that happened prior to the current arrival in state i.

ALTERNATIVELY

The holding time in each state is exponentially distributed and the destination of the jump on leaving state i is independent of holding time

Both holding time distribution and destination of jump on leaving state i are independent of anything that happened prior to arrival in state i

(ii) (a) The estimator *[it is the MLE but this need not be stated]* of λ_i, $\hat{\lambda}$, is the inverse of the average duration of each visit to state i.

so $\hat{\lambda}_A = 4$ per hour, $\hat{\lambda}_B = 5$ per hour, $\hat{\lambda}_C = 1.5$ per hour

The estimator *[it is the MLE but this need not be stated]* of r_{ij}, $\hat{r}_{ij}$, is the proportion of observed jumps out of state i to state j.

$\hat{r}_{AB} = 11/20$
$\hat{r}_{AC} = 9/20$
$\hat{r}_{BA} = 80/125 = 16/25$
$\hat{r}_{BC} = 9/25$
$\hat{r}_{CA} = 24/27 = 8/9$
$\hat{r}_{CB} = 1/9$

(b) The estimated generator matrix (in hr^{-1}) is:

$$\begin{pmatrix} -4 & 11/5 & 9/5 \\ 16/5 & -5 & 9/5 \\ 4/3 & 1/6 & -3/2 \end{pmatrix}$$

(iii) Distribution is binomial with mean $n.r_{ij}$ and variance $n.r_{ij}(1 - r_{ij})$, where n is the given number of transitions.

(iv) Null hypothesis is that the Markov property applies to successive transitions, or that the observed triplets are from a Binomial distribution with the estimated parameters (given the number of transitions to the middle state).

Using test statistic given in the hint, we can draw up the table below.

Triplet	n_{ijk}	$E = n_{ij}\hat{r}_{jk}$	$\dfrac{(n_{ijk} - E)^2}{E}$
ABC	42	39.6	0.1455
ABA	68	70.4	0.08182
ACA	85	80	0.3125
ACB	4	10	3.6
BAB	50	44	0.8182
BAC	30	36	1
BCA	38	40	0.1
BCB	7	5	0.8
CAB	64	66	0.0606
CAC	56	54	0.07407
CBA	8	9.6	0.2667
CBC	7	5.4	0.4741

Test statistic 7.7335

Under the null hypothesis, the test statistic follows a χ^2 distribution with the following number of degrees of freedom:

	Number of triplets	12
Minus	Number of pairs	6
Plus	Number of states	3
Minus	One	1
		8 degrees of freedom

The critical value of χ_8^2 at the 5% significance level is 15.51

As 7.7335 < 15.51 there is no evidence to reject the null hypothesis.

[Alternative approaches could be taken which resulted in a slightly different result for the test statistic. These were given full credit where appropriate.]

(v) *[Refer back to part (i) — the test in (iv) has only tested that there is no evidence that the destination that the next jump depends on the previous state occupied. Need to test the other assumptions].*

Holding times — are these exponentially distributed?

A chi-squared goodness of fit test would be appropriate

Is destination of jump independent of the holding time?

There is no obvious test statistic for doing this. A suitable test would be to classify jumps as being from short, medium and long holding times and investigating these graphically.

(vi) APPROXIMATE METHOD

Divide time into very short intervals, h, such that $\sigma_{ij}h$ is much less than 1.

Simulate a discrete-time Markov chain $\{Y_n : n \geq 0\}$, with transition probabilities $p^*_{ij}(h) = \delta_{ij} + h\sigma_{ij}$.

The jump process, X_t is given by $X_t = Y_{[t/h]}$.

EXACT METHOD

Simulate the jump chain as a Markov chain, with transition probabilities $p_{ij} = \sigma_{ij} / \lambda_i$.

Once the path $\{\hat{X}_n : n = 0, 1, ...\}$ has been generated, the holding times $\{T_n : n = 0, 1, ...\}$ are a sequence of independent exponential random variables, having parameter $\lambda_{\hat{X}_n}$.

END OF EXAMINERS' REPORT

EXAMINATION

9 April 2008 (am)

Subject CT4 — Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 11 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator from the approved list.

1 List four factors in respect of which life insurance mortality statistics are often subdivided. [2]

2 Describe how smoothness is ensured when mortality rates are graduated using each of the following methods:

(a) fitting a parametric formula
(b) graphical graduation

[3]

3 (i) Define the following stochastic processes:

(a) Poisson process
(b) compound Poisson process

[4]

(ii) Identify the circumstances in which a compound Poisson process is also a Poisson process. [1]
[Total 5]

4 Describe the benefits and limitations of modelling in actuarial work. [6]

5 A survey of first marriage patterns among women in a remote population in central Asia collected the following data for a sample of women:

- calendar year of birth
- calendar year of first marriage

Data are also available about the population of never-married women on 1 January each year, classified by age last birthday.

You have been asked to estimate the intensity, λ_x, of first marriage for women aged x.

(i) State the rate interval implied by the first marriages data. [1]

(ii) Derive an appropriate exposed to risk which corresponds to the first marriages data. State any assumptions that you make. [4]

(iii) Explain to what age x your estimate of λ_x applies. State any assumptions that you make. [2]
[Total 7]

6 An investigation was carried out into mortality rates among a certain class of female pensioners. Crude mortality rates were estimated by single years of age from ages 65–89 years last birthday inclusive. The investigators decided to ask an actuary to compare the crude rates with a standard table. They calculated the relevant standardised deviations, printed them out and sent them to the actuary.

Unfortunately, because of a printing error, the right-hand edge of the document containing the standardised deviations failed to print properly. The actuary was unable to read the magnitude of the standardised deviations. However, the sign of each deviation was clear. This revealed that the crude mortality rates were higher than the standard table rates at ages 65–72 years and 75–84 years inclusive, but that the crude mortality rates were lower than the standard table rates at ages 73–74 years and 85–89 years inclusive.

The null hypothesis to be tested is that the crude mortality rates come from a population with underlying mortality consistent with that in the standard table.

(i) List two statistical tests of the null hypothesis which the actuary could carry out on the basis of the information received. [1]

(ii) Carry out both tests. For each test, state what feature of the experience it is specifically testing, and give your conclusion. [10]
[Total 11]

7 In a certain small country all listed companies are required to have their accounts audited on an annual basis by one of the three authorised audit firms (A, B and C). The terms of engagement of each of the audit firms require that a minimum of two annual audits must be conducted by the newly appointed firm. Whenever a company is able to choose to change auditors, the likelihood that it will retain its auditors for a further year is (80%, 70%, 90%) where the current auditor is (A,B,C) respectively. If changing auditors a company is equally likely to choose either of the alternative firms.

(i) A company has just changed auditors to firm A. Calculate the expected number of audits which will be undertaken before the company changes auditors again. [2]

(ii) Formulate a Markov chain which can be used to model the audit firm used by a company, specifying:

(a) the state space
(b) the transition matrix

[4]

(iii) Calculate the expected proportion of companies using each audit firm in the long term. [5]
[Total 11]

8 An education authority provides children with musical instrument tuition. The authority is concerned about the number of children giving up playing their instrument and is testing a new tuition method with a proportion of the children which it hopes will improve persistency rates. Data have been collected and a Cox proportional hazards model has been fitted for the hazard of giving up playing the instrument. Symmetric 95% confidence intervals (based upon standard errors) for the regression parameters are shown below.

Covariate		*Confidence Interval*
Instrument		
	Piano	0
	Violin	[−0.05,0.19]
	Trumpet	[0.07,0.21]
Tuition method		
	Traditional	0
	New	[−0.15,0.05]
Sex		
	Male	[−0.08,0.12]
	Female	0

(i) Write down a general expression for the Cox proportional hazards model, defining all terms that you use. [3]

(ii) State the regression parameters for the fitted model. [2]

(iii) Describe the class of children to which the baseline hazard applies. [1]

(iv) Discuss the suggestion that the new tuition method has improved the chances of children continuing to play their instrument. [3]

(v) Calculate, using the results from the model, the probability that a boy will still be playing the piano after 4 years if provided with the new tuition method, given that the probability that a girl will still be playing the trumpet after 4 years following the traditional method is 0.7. [3]

[Total 12]

9 An investigation into the mortality of patients following a specific type of major operation was undertaken. A sample of 10 patients was followed from the date of the operation until either they died, or they left the hospital where the operation was carried out, or a period of 30 days had elapsed (whichever of these events occurred first). The data on the 10 patients are given in the table below.

Patient number	*Duration of observation (days)*	*Reason for observation ceasing*
1	2	Died
2	6	Died
3	12	Died
4	20	Left hospital
5	24	Left hospital
6	27	Died
7	30	Study ended
8	30	Study ended
9	30	Study ended
10	30	Study ended

(i) State whether the following types of censoring are present in this investigation. In each case give a reason for your answer.

(a) Type I
(b) Type II
(c) Random [3]

(ii) State, with a reason, whether the censoring in this investigation is likely to be informative. [1]

(iii) Calculate the value of the Kaplan-Meier estimate of the survival function at duration 28 days. [5]

(iv) Write down the Kaplan-Meier estimate of the hazard of death at duration 8 days. [1]

(v) Sketch the Kaplan-Meier estimate of the survival function. [2]
[Total 12]

10 An internet service provider (ISP) is modelling the capacity requirements for its network. It assumes that if a customer is not currently connected to the internet ("offline") the probability of connecting in the short time interval $[t,dt]$ is $0.2dt + o(dt)$. If the customer is connected to the internet ("online") then it assumes the probability of disconnecting in the time interval is given by $0.8dt + o(dt)$.

The probabilities that the customer is online and offline at time t are $P_{ON}(t)$ and $P_{OFF}(t)$ respectively.

(i) Explain why the status of an individual customer can be considered as a Markov Jump Process. [2]

(ii) Write down Kolmogorov's forward equation for $P'_{OFF}(t)$. [2]

(iii) Solve the equation in part (ii) to obtain a formula for the probability that a customer is offline at time t, given that they were offline at time 0. [3]

(iv) Calculate the expected proportion of time spent online over the period $[0,t]$. [HINT: Consider the expected value of an indicator function which takes the value 1 if offline and 0 otherwise.]

[4]

(v) (a) Sketch a graph of your answer to (iv) above.
(b) Explain its shape. [3]

[Total 14]

11 An investigation was carried out into the relationship between sickness and mortality in an historical population of working class men. The investigation used a three-state model with the states:

1 Healthy
2 Sick
3 Dead

Let the probability that a person in state i at time x will be in state j at time $x+t$ be ${}_tp_x^{ij}$. Let the transition intensity at time $x+t$ between any two states i and j be μ_{x+t}^{ij}.

(i) Draw a diagram showing the three states and the possible transitions between them. [2]

(ii) Show from first principles that

$$\frac{\partial}{\partial t}\,{}_tp_x^{23} = {}_tp_x^{21}\mu_{x+t}^{13} + {}_tp_x^{22}\mu_{x+t}^{23}\,. \qquad [5]$$

(iii) Write down the likelihood of the data in the investigation in terms of the transition rates and the waiting times in the Healthy and Sick states, under the assumption that the transition rates are constant. [3]

The investigation collected the following data:

• man-years in Healthy state	265
• man-years in Sick state	140
• number of transitions from Healthy to Sick	20
• number of transitions from Sick to Dead	40

(iv) Derive the maximum likelihood estimator of the transition rate from Sick to Dead. [3]

(v) Hence estimate:

(a) the value of the constant transition rate from Sick to Dead
(b) 95 per cent confidence intervals around this transition rate

[4]
[Total 17]

END OF PAPER

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

$P_{x,t} = P_{x\text{-}1,t}{}^*$

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

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

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

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

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

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

$= \sum_{i=1}^{\infty}$ probability ith 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:

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

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

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

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

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

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

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

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

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

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

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.

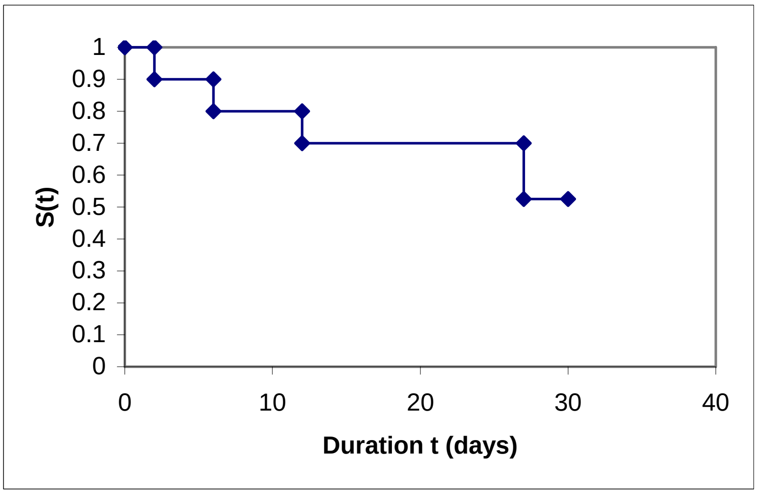
(iii) The Kaplan-Meier estimate of the survival function is estimated as follows

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

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

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

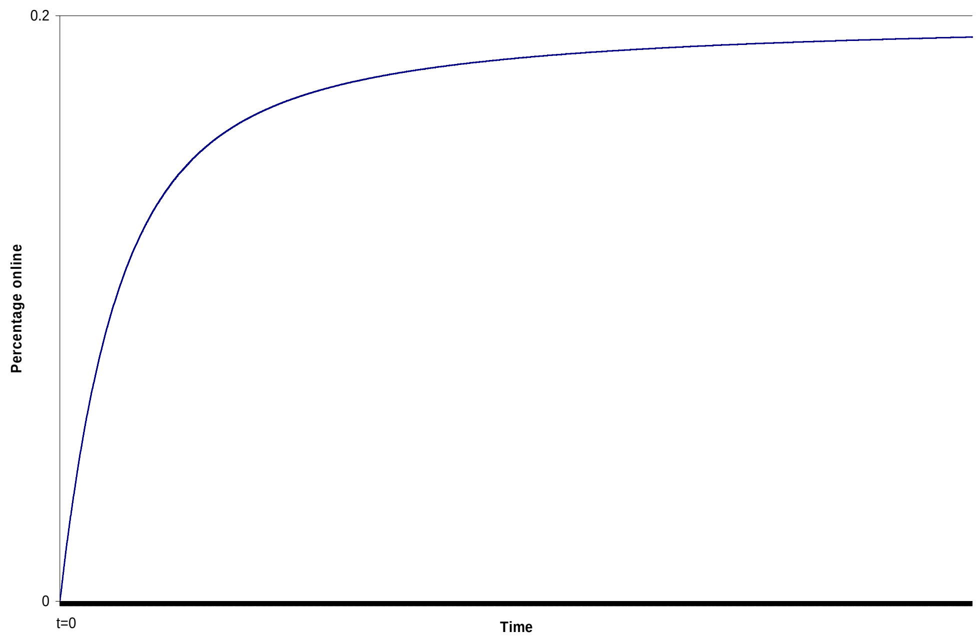
Either online or offline at any time so time spent online is:

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

So proportion spent online is:

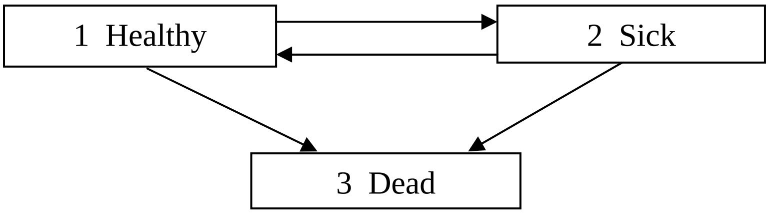
$$\frac{0.2t - 0.2(1 - e^{-t})}{t} = 0.2 - 0.2(\frac{1 - e^{-t}}{t})$$

(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 $\frac{0.2857}{140} = 0.00204$.

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

EXAMINATION

17 September 2008 (am)

Subject CT4 — Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 12 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator from the approved list.

1 You work for a consultancy which has created an actuarial model and is now preparing documentation for the client.

List the key items you would include in the documentation on the model. [4]

2 The classification of stochastic models according to:

- discrete or continuous time variable
- discrete or continuous state space

gives rise to a four-way classification.

Give four examples, one of each type, of stochastic models which may be used to model observed processes, and suggest a practical problem to which each model may be applied. [4]

3 Compare the advantages and disadvantages of the Binomial and the multiple-state models in the following situations:

(a) analysing human mortality without distinguishing between causes of death
(b) analysing human mortality when distinguishing between causes of death

[5]

4 In the village of Selborne in southern England in the year 1637 the number of babies born each month was as follows

January	2	July	5
February	1	August	1
March	1	September	0
April	2	October	2
May	1	November	0
June	2	December	3

Data show that over the 20 years before 1637 there was an average of 1.5 births per month. You may assume that births in the village historically follow a Poisson process.

An historian has suggested that the large number of births in July 1637 is unusual.

(i) Carry out a test of the historian's suggestion, stating your conclusion. [4]

(ii) Comment on the assumption that births follow a Poisson process. [1]
[Total 5]

5 An investigation into the mortality experience of a sample of the male student population of a large university has been carried out. The university authorities wish to know whether the mortality of male students at the university is the same as that of males in the country as a whole. They have drawn up the following table.

Age x	*Number of deaths*	*Expected number of deaths assuming national mortality*
18	13	10
19	15	12
20	14	14
21	20	12
22	12	8
23	8	5

Carry out an overall test of the university authorities' hypothesis, stating your conclusion. [5]

6 A portfolio of term assurance policies was transferred from insurer A to insurer B on 1 January 2001. Each policy in the portfolio was written with premiums payable annually in advance. Insurer B wishes to investigate the mortality experience of its acquired portfolio and has collected the following data over the period 1 January 2001 to 1 January 2005:

d_x numbers of deaths aged x

$P_{x,t}$ number of policies in force aged x at time t (t = 0, 1, 2, 3, 4 years measured from 1 January 2001)

Where x is defined as:

age last birthday at the most recent policy anniversary prior to the portfolio transfer + number of premiums received by insurer B.

(i) (a) State the rate interval implied by the above data.
(b) Write down the range of ages at the start of the rate interval. [2]

(ii) Give an expression which can be used to estimate the initial exposed to risk at age x, E_x, stating any assumptions made. [2]

The following is an extract from the data collected in the investigation:

x	d_x	$\sum P_{x,t}$	$\sum P_{x,t+1}$
39	28	10,536	11,005
40	36	10,965	10,745
41	33	10,421	10,577

where the summations are from t = 0 to t = 3.

(iii) Estimate q_{40}, stating any further assumptions made. [3]
[Total 7]

7 (i) Explain why, under Continuous Mortality Investigation investigations, the data analysed are usually based upon the number of policies in force and number of policies giving rise to claims, rather than the number of lives exposed and number of lives who die during the period of study. [2]

Suppose N identical and independent lives are observed from age x exact for one year or until death if earlier.

Define:

π_i to be the proportion of the N lives exposed who hold i policies (i = 1,2,3,….);

$\mathbf{D}_i$ to be a random variable denoting the number of deaths amongst lives with i policies

$\mathbf{C}_i$ to be a random variable denoting the number of claims arising from lives with i policies.

(ii) Derive an expression for the ratio of the variance of the number of claims arising compared with that if each policy covered an independent life. [4]

(iii) Explain how the expression derived in (ii) could be used in practice. [2]
[Total 8]

8 A No-Claims Discount system operated by a motor insurer has the following four levels:

Level 1: 0% discount
Level 2: 25% discount
Level 3: 40% discount
Level 4: 60% discount

The rules for moving between these levels are as follows:

- Following a year with no claims, move to the next higher level, or remain at level 4.
- Following a year with one claim, move to the next lower level, or remain at level 1.
- Following a year with two or more claims, move down two levels, or move to level 1 (from level 2) or remain at level 1.

For a given policyholder in a given year the probability of no claims is 0.85 and the probability of making one claim is 0.12.

(i) Write down the transition matrix of this No-Claims Discount process. [1]

(ii) Calculate the probability that a policyholder who is currently at level 2 will be at level 2 after:

(a) one year.
(b) two years. [3]

(iii) Calculate the long-run probability that a policyholder is in discount level 2.
[5]
[Total 9]

9 A company pension scheme, with a compulsory scheme retirement age of 65, is modelled using a multiple state model with the following categories:

1 currently employed by the company
2 no longer employed by the company, but not yet receiving a pension
3 pension in payment, pension commenced early due to ill health retirement
4 pension in payment, pension commenced at scheme retirement age
5 dead

(i) Describe the nature of the state space and time space for this process. [2]

(ii) Draw and label a transition diagram indicating appropriate transitions between the states. [2]

For *i*,*j* in {1,2,3,4,5}, let:

${}_t p_x^{1i}$ the probability that a life is in state *i* at age *x*+*t*, given they are in state 1 at age *x*

μ_{x+t}^{ij} the transition intensity from state *i* to state *j* at age x+t

(iii) Write down equations which could be used to determine the evolution of ${}_t p_x^{1i}$ (for each *i*) appropriate for:

(a) $x + t < 65$.
(b) $x + t = 65$.
(c) $x + t > 65$.

[6]
[Total 10]

10 In an investigation of reconviction rates among those who have served prison sentences, let X be a random variable which measures the duration from the date of release from prison until the ex-prisoner is convicted of a subsequent offence. The investigation monitored a sample of 100 ex-prisoners (who were all released on the same date) at one-monthly intervals from their date of release for a period of 6 months. Those who could not be traced in any month were removed from the sample at that point and not traced in subsequent months. Reconviction was assumed to take place at the duration that a prisoner was first known to have been reconvicted.

(i) Express the hazard rate at duration x months in terms of probabilities. [1]

The investigation produced the following data for a sample of 100 ex-prisoners.

Months since release	*Number of prisoners contacted*	*Number who had been reconvicted since last contact*
1	100	0
2	97	0
3	95	4
4	90	3
5	85	5
6	80	0

(ii) Calculate the Nelson-Aalen estimate of the survival function. [5]

A previous investigation found that the probability that a prisoner would be reconvicted within 6 months of release was 0.2.

(iii) Estimate confidence intervals around the integrated hazard using the results from part (ii) to test the hypothesis that the rate of reconviction has declined since the previous investigation. [6]

[Total 12]

 PLEASE TURN OVER

11 Consider the random variable defined by $X_n = \sum_{i=1}^{n} Y_i$ with each Y_i mutually independent with probability:

$$P[Y_i = 1] = p, \; P[Y_i = -1] = 1 - p \qquad 0 < p < 1$$

(i) Write down the state space and transition graph of the sequence X_n. [2]

(ii) State, with reasons, whether the process:

(a) is aperiodic.
(b) is reducible.
(c) admits a stationary distribution. [3]

Consider $j > i > 0$.

(iii) Derive an expression for the number of upward movements in the sequence X_n between t and $(t + m)$ if $X_t = i$ and $X_{t+m} = j$. [2]

(iv) Derive expressions for the m-step transition probabilities $p_{ij}^{(m)}$. [3]

(v) Show how the one-step transition probabilities would alter if X_n was restricted to non-negative numbers by introducing:

(a) a reflecting boundary at zero.
(b) an absorbing boundary at zero.

[2]

(vi) For each of the examples in part (v), explain whether the transition probabilities $p_{ij}^{(m)}$ would increase, decrease or stay the same.
(Calculation of the transition probabilities is not required.) [3]
[Total 15]

12 (i) Explain the meaning of the rates of mortality usually denoted q_x and m_x , and the relationship between them. [3]

(ii) Write down a formula for ${}_t q_x$, $0 \le t \le 1$, under each of the following assumptions about the distribution of deaths in the age range [x, x+1]:

(a) uniform distribution of deaths
(b) constant force of mortality
(c) the Balducci assumption

[2]

A group of animals experiences a mortality rate $q_x = 0.1$.

(iii) Calculate m_x under each of the assumptions (a) to (c) above. [8]

(iv) Comment on your results in part (iii). [3]
[Total 16]

END OF PAPER

Subject CT4 — Models
Core Technical

EXAMINERS' REPORT

September 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.

R D Muckart
Chairman of the Board of Examiners

November 2008

Comments

Comments on solutions presented to individual questions for the September 2008 paper are given below.

Q1 *This standard bookwork question was fairly well answered. Some candidates simply wrote down a list of steps in the development of the model, rather than answering the question that was set.*

Q2 *This straightforward question was well answered. Some candidates were vague about emphasising that continuous time models are applied to problems which require continuous monitoring.*

Q3 *Answers to this question were very poor. Many candidates did not go beyond making the point that the Binomial model is hard to extend to multiple decrements, whereas the multiple state model extends quite naturally.*

Q4 *Only a minority of candidates answered this question using the approach intended. Many tried to do a chi-squared test comparing the observed and expected numbers of births. This received some credit, especially when candidates combined the months into half-years, or thirds of a year, before performing the chi-squared test, so that the expected values in each cell were greater than 5.*

Q5 *This straightforward question was very well answered. The most common error was in reducing the number of degrees of freedom below 6. This is incorrect in this case, because the comparison is between an observed experience and a pre-existing experience, not between crude rates and graduated rates.*

Q6 *As with many exposed-to-risk questions, answers to this question were disappointing. In part (iii), few candidates realised that* $\hat{q}_{39\frac{1}{2}}$ *was required to estimate* q_{40}.

Q7 *Part (ii) of this question was standard bookwork, but was nevertheless answered in a brief or cursory fashion by many candidates. On the other hand, a good number of candidates were able to make the points required in part (i).*

Q8 *This question was very well answered, with many candidates scoring full marks. Some candidates were penalised in part (iii) for simply calculating the stationary distribution and not stating explicitly which of the numbers represented the long-run probability of being in discount level 2.*

Q9 *Answers to this question were disappointing, especially to part (iii). In part (ii), candidates who included additional transitions between states 2 and 3, and between states 2 and 1, were not penalised. However, such candidates were expected to produce answers to part (iii) which were consistent with the transition diagram they had drawn in part (ii).*

Q10 *Part (i) of this question was very disappointingly answered, as the required definition is in the Core Reading. Most candidates were able to compute the estimated survival function in part (ii). Some candidates interpreted the question as meaning that the* *numbers contacted at any duration include those reconvicted prior to that duration, so that those reconvicted must be subtracted from those contacted to obtain the relevant n_j. These candidates were given credit for part (ii). In part (iii) many candidates correctly calculated the variance of the integrated hazard but then incorrectly used this variance to compute a confidence interval around the survival function, rather than first computing the confidence interval around the integrated hazard and then using the formula $S(x) = exp(-\Lambda_x)$ to convert this into a confidence interval around S(6).*

Q11 *Answers to this question were disappointing. Many candidates were able to answer parts (i) and (ii) reasonably well, but made little or no attempt at the remaining sections.*

Q12 *Answers to this question varied widely, but overall were disappointing. There was a large variation by centre, with average scores for some centres being several marks higher than for other centres. Perhaps this is the result of different training and education materials being used in different locations? While most candidates could write down the formula $m_x = \dfrac{q_x}{\int_0^1 {}_t p_x dt}$ and the formulae required to answer part (ii), it was clear from the answers to parts (iii) and (iv) that understanding of what these formulae mean was very shaky.*

1 **Instructions on how to run the model**

Tests performed to validate the output of the model.
Definition of input data.
Any limitations of the model identified (e.g. potential unreliability).
Basis on which the form of the model chosen (e.g. deterministic or stochastic)
References to any research papers or discussions with appropriate experts.
Summary of model results.
Name and professional qualification.
Purpose or objectives of the model.
Assumptions underlying the model.
How the model might be adapted or extended.

2 **Discrete time, discrete state space**

Counting process, random walk, Markov chain
No claims bonus in motor insurance.

Continuous time, discrete state space

Counting process, Poisson process, Markov jump process
Healthy-sick-dead model in sickness insurance

Discrete time, continuous state space

General random walk, ARIMA time series model, moving average model
Share price at end of each day

Continuous time, continuous state space

Compound Poisson process, Brownian motion, Ito process, white noise
Value of claims reaching an insurance company monitored
continuously

3 (a) Both models produce consistent and unbiased estimators.

The estimate of q_x made using the Binomial model will have a higher variance than that made using the multiple-state model, though the difference is tiny if the forces of mortality are small.

If data on exact ages at entry into and exit from observation are available, the multiple state model is simpler to apply. The Binomial model requires further assumptions (e.g. uniform distribution of deaths).

The Binomial model also does not use all the information available if exact ages at entry into and exit from observation are available.

However, if the forces of mortality are small, both models will give very similar results.

(b) The multiple state model can simply be extended

The estimators have the same form and the same statistical properties as in the classic life table.

The Binomial model is hard to extend to several causes of death. Although the life table as a computational tool can be extended, the calculations are more complex and awkward than those in the multiple-state model.

4 (i) Suppose that the number of births each month, B, is the outcome of a Poisson process with a rate $\lambda = 1.5$.

The probability of obtaining b births per month is given by the formula $\Pr[B=b] = \dfrac{\exp(-1.5).1.5^b}{b!}$

Therefore we have

b	$\Pr[B=b]$
0	0.223
1	0.335
2	0.251
3	0.126
4	0.047
5	0.014
6+	0.004

Therefore, if the number of births per month is the outcome of a Poisson process with a rate of 1.5 per month the probability of obtaining 5 or more births in a single month is 0.014 + 0.004 = 0.018.
EITHER This is very small OR this is < 0.05

which suggests that the historian may be correct to suspect something unusual about July 1637.

But only July has a number of births more than 5, and at the 5% level of statistical significance we expect 1 month in 20 to have such a large number, then unless we have a prior expectation that July is unusual, we should be cautious before accepting the historian's suggestion.

(ii) The assumption that births follow a Poisson process is unlikely to be entirely realistic

EITHER because of the occurrence of multiple births (twins and triplets)
OR because births tend to occur seasonally
OR because the process might be time inhomogeneous.

5 Using the chi-squared test (a suitable overall test).

If $z_x = \dfrac{\text{actual deaths - expected deaths}}{\sqrt{\text{expected deaths}}}$, then the test statistic is $\sum_x z_x{}^2 \sim \chi^2{}_m$,
where m is the number of ages, which in this case is 6.

The calculations are shown below.

Age x	z_x	$z_x{}^2$
18	0.9487	0.9
19	0.8660	0.75
20	0	0
21	2.3094	5.3333
22	1.4142	2
23	1.3416	1.8

Therefore the value of the test statistic is 10.783.

The critical value of the chi-squared distribution at the 5% level of significance with 6 degrees of freedom is 12.59.

Since 10.783 < 12.59 there is insufficient evidence to reject the hypothesis that the mortality rate of men in the University is the same as that of the national population.

6 (i) Age label changes on the receipt of the premium on the policy anniversary so this is a policy year rate interval.

Policyholders' ages range from x to $x+1$ at start of the rate interval.

(ii) Central exposed to risk $E_x^c = \int_{t=0}^{4} P_{x,t}dt \approx \frac{1}{2}\sum_{t=0}^{3}(P_{x,t} + P_{x,t+1})$

Approximation assumes population changes linearly over each year during the period of investigation.

Initial exposed to risk $E_x \approx \frac{1}{2}\sum_{t=0}^{3}(P_{x,t} + P_{x,t+1}) + \frac{1}{2}d_x$,

assuming deaths are uniform over the rate interval OR deaths occur on average half way through the rate interval.

(but NOT deaths are uniform over the "year", or occur on average half way through the "year")

(iii) $\hat{q}_x = \frac{d_x}{E_x}$ estimates q_x for the average age at the start of the rate interval.

Assuming birthdays are uniformly distributed across policy years,

the average age at the start of the rate interval is $x+½$, so we require $\hat{q}_{39\frac{1}{2}}$ to estimate q_{40}.

Assuming $\hat{q}_{39\frac{1}{2}} = \frac{1}{2}[\hat{q}_{39} + \hat{q}_{40}]$ we have

$$\hat{q}_{39} = \frac{28}{\frac{1}{2}(10536 + 11005) + \frac{1}{2} * 28} = 0.002596$$

$$\hat{q}_{40} = \frac{36}{\frac{1}{2}(10965 + 10745) + \frac{1}{2} * 36} = 0.003311$$

and hence our estimate of q_{40} is 0.5[0.002596 + 0.003311) = 0.002954.

7 (i) Individual life offices are likely to have their systems set up to provide information on a "by policy" basis.

When data from different offices is pooled, it would not be practicable to establish whether an individual held policies with other companies.

(ii) If the mortality rate is q_x then since the lives are independent the number of deaths $\mathbf{D}_i$ will be distributed Binomial$\left(q_x, \pi_i N\right)$

So $\sum_i \mathbf{C}_i = \sum_i i\mathbf{D}_i$.

Hence $\text{Var}[\mathbf{C}] = \text{Var}\left[\sum_i \mathbf{C}_i\right] = \text{Var}\left[\sum_{\text{i}} i\mathbf{D}_i\right] = \sum_i i^2 \text{Var}\left[\mathbf{D}_i\right]$

by independence of deaths

$$= \sum_i i^2 \pi_i N q_x (1 - q_x)$$

If instead there were $\sum_i i\pi_i N$ independent policies/lives the variance would be additive so:

$$\text{Var}\left[\mathbf{C}'\right] = \sum_i i\pi_i N q_x (1 - q_x)$$

So the variance is increased by the ratio $\dfrac{\sum_i i^2 \pi_i}{\sum_i i\pi_i}$

(iii) If the proportions of lives holding *i* policies were known, the variance ratio could be allowed for in statistical tests

by using the ratio to adjust the variance upwards.

However, the variance ratio is unlikely to be known exactly.

Special investigations may be performed from time to time to estimate the variance ratios by matching up policyholders, which could then be applied to subsequent mortality investigations.

8 (i) The transition matrix of the process is

$$P = \begin{pmatrix} 0.15 & 0.85 & 0 & 0 \\ 0.15 & 0 & 0.85 & 0 \\ 0.03 & 0.12 & 0 & 0.85 \\ 0 & 0.03 & 0.12 & 0.85 \end{pmatrix}$$

(ii) (a) For the one year transition, $p_{22} = 0,$ as can be seen from above (or is obvious from the statement).

(b) The second order transition matrix is

$$\begin{pmatrix} 0.15^2+0.85\times0.15 & 0.85\times0.15 & 0.85^2 & 0 \\ 0.15^2+0.85\times0.03 & 0.85\times0.15+0.85\times0.12 & 0 & 0.85^2 \\ 0.03\times0.15+0.12\times0.15 & 0.85\times0.03\times2 & 0.85\times0.12\times2 & 0.85^2 \\ 0.03\times0.15+0.12\times0.03 & 0.12^2+0.85\times0.03 & 0.85\times0.03+0.85\times0.12 & 0.12\times0.85+0.85^2 \end{pmatrix}$$

$$= \begin{pmatrix} 0.15 & 0.1275 & 0.7225 & 0 \\ 0.048 & 0.2295 & 0 & 0.7225 \\ 0.0225 & 0.051 & 0.204 & 0.7225 \\ 0.0081 & 0.0399 & 0.1275 & 0.8245 \end{pmatrix}$$

hence the required probability is 0.2295.

(iii) In matrix form, the equation we need to solve is $\pi P = \pi$, where π is the vector of equilibrium probabilities.

This reads

$$\begin{aligned} 0.15\pi_1 + 0.15\pi_2 + 0.03\pi_3 \qquad\qquad &= \pi_1 \quad (1) \\ 0.85\pi_1 + \qquad +0.12\pi_3 + 0.03\pi_4 &= \pi_2 \quad (2) \\ +0.85\pi_2 \qquad\qquad +0.12\pi_4 &= \pi_3 \quad (3) \\ 0.85\pi_3 + 0.85\pi_4 &= \pi_4 \quad (4) \end{aligned}$$

Discard the first of these equations and use also the fact that $\sum_{i=1}^{4} \pi_i = 1$.

Then, we obtain first from (4) that $0.85\pi_3 = 0.15\pi_4$
or, that $\pi_4 = 17\pi_3 / 3$

Substituting in (3) this gives

$$0.85\pi_2 \; + 0.12 \times \frac{17}{3}\pi_3 = \pi_3 \Rightarrow \pi_3 = 2.65625\pi_2$$

(2) now yields that

$$0.85\pi_1 = \pi_2 - 0.12\pi_3 - 0.03\pi_4 = \frac{1}{2.65625}\pi_3 - 0.12\pi_3 - 0.17\pi_3 = 0.0865\pi_3,$$

so that finally we get $\pi_1 = 0.10173\pi_3$.

Using now that the probabilities must add up to one, we obtain

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = (0.10173 + 0.3765 + 1 + 5.666)\pi_3 = 1,$$

or that $\pi_3 = 0.13996$.

Solving back for the other variables we get that

$$\pi_1 = 0.01424, \quad \pi_2 = 0.05269, \quad \pi_4 = 0.79311$$

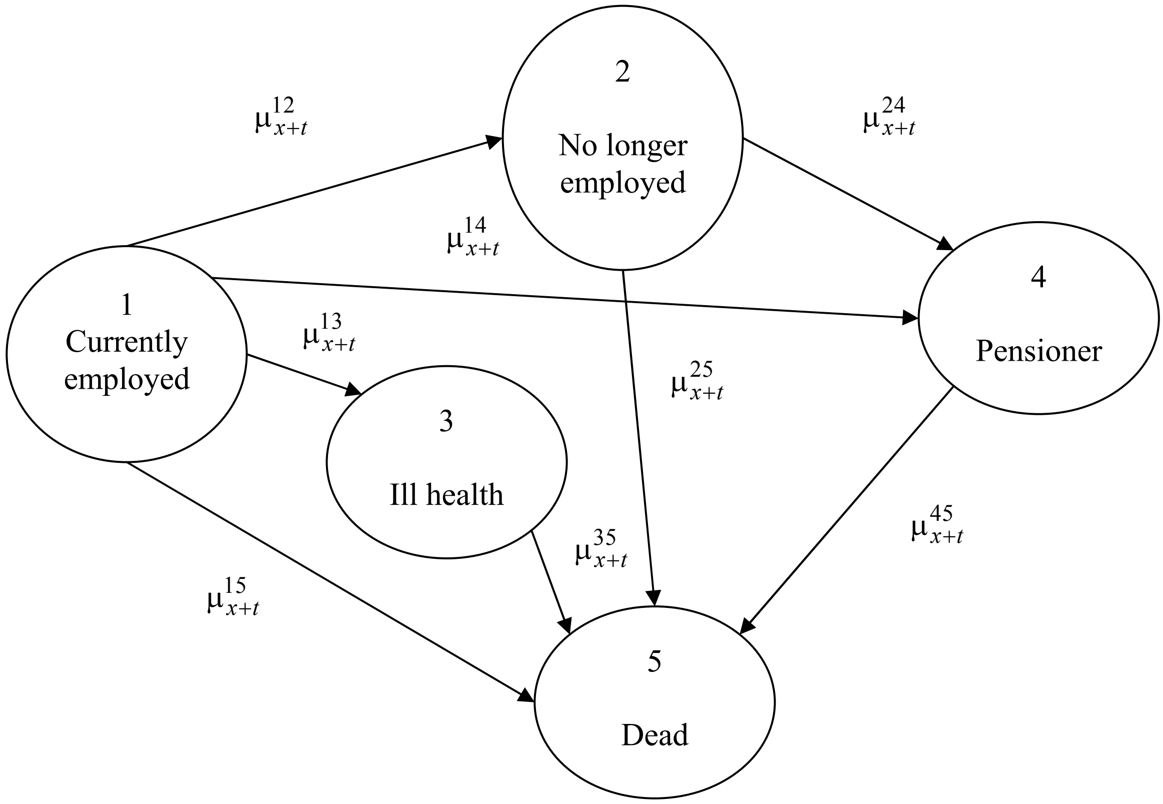
The long-run probability that the motorist is in discount level 2 is therefore 0.05269.

9 (i) The state space is discrete with states as given in the question.

The process operates in continuous time.
However, at the compulsory scheme retirement
age of 65 there is a discrete step change.

This is sometimes described as a mixed process.

(ii)



(iii) (a) For $x + t < 65$

$$\partial/\partial t \; {}_t p_x^{11} = -(\mu^{12}_{x+t} + \mu^{13}_{x+t} + \mu^{15}_{x+t})\, {}_t p_x^{11}$$

$$\partial/\partial t \; {}_t p_x^{12} = \mu^{12}_{x+t} \cdot {}_t p_x^{11} - \mu^{25}_{x+t} \cdot {}_t p_x^{12}$$

$$\partial/\partial t \; {}_t p_x^{13} = \mu^{13}_{x+t} \cdot {}_t p_x^{11} - \mu^{35}_{x+t} \cdot {}_t p_x^{13}$$

$$\partial/\partial t \; {}_t p_x^{15} = \mu^{15}_{x+t} \cdot {}_t p_x^{11} + \mu^{25}_{x+t} \cdot {}_t p_x^{12} + \mu^{35}_{x+t} \cdot {}_t p_x^{13}$$

and ${}_t p_x^{14}$ is zero.

(b) For $x + t = 65$

${}_t p_x^{11}$ and ${}_t p_x^{12}$ become 0 at $x + t = 65 + \delta$

$${}_{t+\delta} p_x^{14} = {}_{t-\delta} p_x^{11} + {}_{t-\delta} p_x^{12}$$

(c) For $x + t > 65$

$${}_t p_x^{11} = {}_t p_x^{12} = 0$$

$$\partial/\partial t \; {}_t p_x^{13} = -\mu_{x+t}^{35} \cdot {}_t p_x^{13}$$

$$\partial/\partial t \; {}_t p_x^{14} = -\mu_{x+t}^{45} \cdot {}_t p_x^{14}$$

$$\partial/\partial t \; {}_t p_x^{15} = \mu_{x+t}^{35} \cdot {}_t p_x^{13} + \mu_{x+t}^{45} \cdot {}_t p_x^{14}$$

10 (i) EITHER

The hazard rate at duration x is given by

$$\lim_{dt \to 0} \frac{\Pr[X \le x + dt \mid X > x]}{dt}.$$

OR

In discrete time, the hazard rate at duration x is given by, $\Pr[X = x \mid X \ge x]$.

OR

The hazard rate at duration x is given by $h(x) = -\dfrac{1}{S(x)} \dfrac{d}{dx}[S(x)]$,

where $S(x)$ is the survival function defined as $\Pr[X > x]$.

(ii) The integrated hazard, Λ_x, is estimated as follows:

x_j	n_j	d_j	c_j	$\dfrac{d_j}{n_j}$	$\Lambda_x = \sum_{x_j \le x} \dfrac{d_j}{n_j}$
0	100	0	0	0	0
1	100	0	3	0	0
2	97	0	2	0	0
3	95	4	1	4/95 = 0.0421	0.0421
4	90	3	2	3/90 = 0.0333	0.0754
5	85	5	0	5/85 = 0.0588	0.1343
6	80	0	80	0	0.1343

The survival function $S(x)$ is given by $\exp(-\Lambda_x)$, so that we have

x	$S(x)$
$0 \le x < 3$	1.0000
$3 \le x < 4$	0.9588
$4 \le x < 5$	0.9274
$5 \le x$	0.8744

(iii) Confidence intervals around the integrated hazard may be estimated using the formula

$$Var[\tilde{\Lambda}_x] = \sum_{x_j \le x} \frac{d_j(n_j - d_j)}{n_j^3}$$

Applying this to the data gives

x_j	n_j	d_j	$\frac{d_j(n_j - d_j)}{n_j^3}$	$\sum_{x_j \le x} \frac{d_j(n_j - d_j)}{n_j^3}$
0	100	0	0	0
1	100	0	0	0
2	97	0	0	0
3	95	4	0.000425	0.000425
4	90	3	0.000358	0.000783
5	85	5	0.000651	0.001434
6	80	0	0	0.001434

95 per cent confidence intervals around the integrated hazard at duration 6 can therefore be computed as

$$\hat{\Lambda}_6 \pm 1.96\sqrt{\text{var}\,\hat{\Lambda}_6}$$

$$= 0.1343 \pm 1.96\sqrt{0.001434}$$

$$= (0.1343 - 0.0742,\ 0.1343 + 0.0742)$$

$$= (0.0601,\ 0.2085).$$

THEN EITHER

The estimated survival function, $\hat{S}(x)$ is given

by $\exp(-\hat{\Lambda}_x)$,

so that the 95 per cent confidence interval for $\hat{S}(x)$ is

[exp(–0.0601), exp(–0.2085)]

which is (0.9417, 0.8118).
In the previous investigation the probability that a prisoner would not be reconvicted within 6 months of release was 1 – 0.2 = 0.8.

Since the 95 per cent confidence interval around $\hat{S}(x)$ in the current investigation does not include the value 0.8, and our estimate of $\hat{S}(x) > 0.8$ we conclude that the rate of reconviction has declined since the previous investigation.

OR

In the previous investigation the probability that a prisoner would not be reconvicted within 6 months of release was 1 – 0.2 = 0.8 – i.e. $S(6) = 0.8$

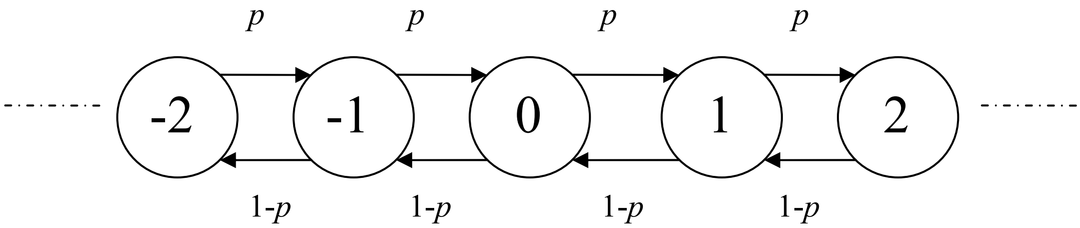
Since $S(x) = \exp(-\Lambda_x)$, the value of Λ_6 corresponding to $S(6) = 0.8$ is

$\Lambda_6 = -\log_e(0.8) = 0.2231$.

Since this is higher than the upper limit in the range (0.0601, 0.2085) we conclude that the rate of reconviction has declined since the previous investigation.

11 (i) State space is the set of integers $\mathbb{Z}$.

Transition graph:



(ii) (a) The process is not aperiodic

because it has period 2:
for example, starting from an even number the process is only even after an even number of steps

(b) The process is irreducible

as the probabilities of X_n increasing and decreasing by 1 are both non-zero so any state can be reached.

(c) No stationary distribution will exist because the state space is infinite.

(iii) Suppose there are u upward movements.

Then there must be $m - u$ downward movements,

and $u - (m - u) = j - i$

So $u = \dfrac{m + j - i}{2}$.

(iv) The maximum number of upward steps is m so the transition probability is zero if $j - i > m$.

As the chain is periodic with period 2, it can only occupy state j after m steps if $m + j - i$ is even.

If $m + j - i$ is even and $j - i \leq m$ then there must be u upward jumps and $(m - u)$ downward jumps.

These can be ordered in $\begin{pmatrix} m \\ u \end{pmatrix}$ ways.

So the transition probabilities are:

$$p_{ij}^{(m)} = \begin{cases} \binom{m}{u} p^u (1-p)^{m-u} & \text{if } j-i \le m \text{ and } m+j-i \text{ even} \\ 0 & \text{otherwise} \end{cases}$$

(v) EITHER

In both cases the transition probabilities are unaltered unless $X_i = 0$.

(a) Reflecting boundary implies $P[X_{i+1} = 1 \mid X_i = 0] = 1$ (or $p_{01}{}^{(1)} = 1$)

(b) Absorbing boundary implies $P[X_{i+1} = 0 \mid X_i = 0] = 1$ (or $p_{00}{}^{(1)} = 1$)

OR

A matrix solution for the transition probabilities is acceptable

Reflecting:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ 1-p & 0 & p & 0 & 0 & \dots \\ 0 & 1-p & 0 & p & 0 & \dots \\ 0 & 0 & 1-p & 0 & p & \dots \\ 0 & 0 & 0 & 1-p & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

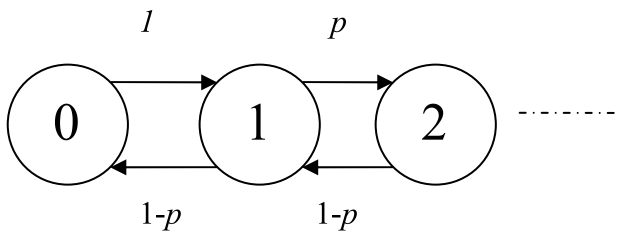
Absorbing:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 1-p & 0 & p & 0 & 0 & \dots \\ 0 & 1-p & 0 & p & 0 & \dots \\ 0 & 0 & 1-p & 0 & p & \dots \\ 0 & 0 & 0 & 1-p & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

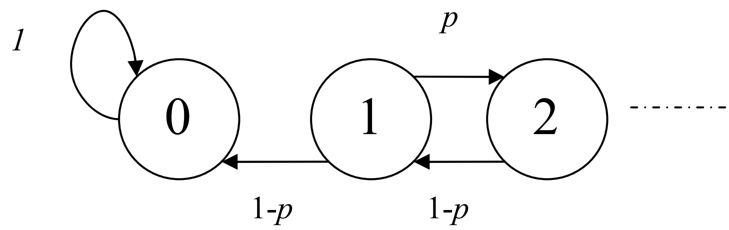
OR

A diagrammatic solution is also acceptable:

Reflecting



Absorbing:



(vi) In both cases the zero transition probabilities remain zero as the period is still 2 where relevant.

If i is sufficiently above 0 then conditions at zero will not be relevant and all the m-step transition probabilities will remain the same. (This applies if $m < i$.)

Otherwise

In (a) some sample paths which would have taken X below zero will be reflected, increasing the probability of reaching j at step m.

So the m-step transition probabilities would increase.

In (b) any sample path which reaches zero would no longer be able to access state j

so the transition probabilities would decrease.

12 (i) q_x is the probability that a life aged exactly x will die before reaching exact age x+1, and is called the initial rate of mortality.

m_x is called the central rate of mortality and represents the probability that a life alive between the ages of x and x+1 dies

They are related by:

$$m_x = \frac{q_x}{\int_0^1 {}_t p_x dt}$$

(ii) (a) Uniform distribution of deaths (UDD)

$${}_t q_x = t * q_x$$

(b) Constant force of mortality (CFM)

$${}_t q_x = 1 - e^{-\mu * t}$$

(c) Balducci assumption

$${}_{1-t} q_{x+t} = (1-t) * q_x$$

(iii) (a) **UDD**

$$\int_0^1 {}_t p_x dt = \int_0^1 (1 - 0.1t) dt = 1 - 0.1 \left[\frac{t^2}{2} \right]_0^1 = 0.95$$

(or other reasoning why exposure is 0.95 under UDD)

$m_x = 0.1/0.95 = 0.105263$

(b) **CFM**

μ given by:

$$1 - e^{-\mu} = 0.1$$

$$\mu = -\ln 0.9 = 0.1053605$$

EITHER

If force of mortality constant over [x, x+1] then central rate must be equal to the force μ

so $m_x = 0.1053605$

OR

$$\int_0^1 {}_t p_x dt = \int_0^1 (1-(1-e^{-\mu t}))dt = -\frac{1}{\mu}\left[e^{-\mu t}\right]_0^1 = \frac{1}{\mu}(1-e^{-\mu}) = 0.949122$$

$m_x = 0.1/0.949122 = 0.1053605$

(c) **Balducci**

For consistency, observe that ${}_1 p_x = {}_t p_x \cdot {}_{1-t} p_{x+t}$

So

$${}_t p_x = \frac{{}_1 p_x}{{}_{1-t} p_{x+t}} = \frac{0.9}{1-{}_{1-t} q_{x+t}} = \frac{0.9}{0.9+0.1t}$$

$$\int_0^1 {}_t p_x dt = \int_0^1 \frac{0.9}{0.9+0.1t} dt = \frac{0.9}{0.1}\left[\ln(0.9+0.1t)\right]_0^1 = -9\ln 0.9 = 0.9482446$$

So $m_x = 0.1/0.9482446 = 0.1054580$

(iv) The Balducci assumption implies a decreasing mortality rate over [x, x+1] and UDD an increasing mortality rate.

CFM is obviously constant

For a given number of deaths over the period, the estimated exposure would be highest if we assumed an increasing mortality rate.

We would expect the central rate to be highest for that with the lowest estimate exposure, hence Balducci > CFM > UDD is the expected order.

END OF EXAMINERS' REPORT

EXAMINATION

29 April 2009 (am)

Subject CT4 — Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*
2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*
3. *Mark allocations are shown in brackets.*
4. *Attempt all 12 questions, beginning your answer to each question on a separate sheet.*
5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator from the approved list.

1 A life insurance company has a small group of policies written on impaired lives and has conducted an investigation into the mortality of these policyholders. It is proposed that the crude mortality rates be graduated for use in future premium calculations.

Discuss the suitability of two methods of graduation that the insurance company could use. [3]

2 (i) Explain what is meant by a time-homogeneous Markov chain. [2]

Consider the time-homogeneous two-state Markov chain with transition matrix:

$$\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

(ii) Explain the range of values that a and b can take which result in this being a valid Markov chain which is:

(a) irreducible
(b) periodic [3]

[Total 5]

3 List the benefits and limitations of modelling in actuarial work. [5]

4 Below is an extract from English Life Table 15 (females).

Age x (years)	*Number of survivors to exact age x out of 100,000 births*
30	98,617
40	97,952

(i) Calculate ${}_5q_{30}$ under each of the two following alternative assumptions:

(a) a uniform distribution of deaths (UDD) between ages 30 and 40 years
(b) a constant force of mortality between ages 30 and 40 years [3]

(ii) Calculate the number of survivors to exact age 35 years out of 100,000 births under each of the assumptions in (i) above. [1]

English Life Table 15 (females) was originally calculated using data classified by single years of age. The number of survivors to exact age 35 years was 98,359.

(iii) Comment on the appropriateness of the assumptions of UDD and a constant force of mortality between ages 30 and 40 years in this example. [3]

[Total 7]

5 Explain the basis underlying the grouping of signs test, and derive the formula for the probability of exactly t positive groups by considering the possible arrangements of a set of positive and negative signs. [5]

6 An investigation by a hospital into rates of recovery after a specific type of operation collected the following data for each month of the calendar year 2008:

- number of persons who recovered from the operation during the month (defined as being discharged from the hospital) classified by the month of their operation.

You may assume that there were no deaths.

On the first day of each month from January 2008 to January 2009, the hospital listed all in-patients who were yet to recover from this operation, classified according to the length of time elapsing since their operation, to the nearest month.

(i) (a) Write down an expression which will enable the hospital to calculate rates of recovery, r_x, during 2008 at various durations x since the operation using the available data.

(b) Derive a formula for the exposed to risk based on the information in the hospital's monthly lists of in-patients which corresponds to the data on recovery from the operation.

[5]

(ii) Determine the value of f such that the expression in (i)(a) applies to an actual duration $x + f$ since the operation. [2]

[Total 7]

7 (i) Explain how the classification of stochastic processes according to the nature of their state space and time space leads to a four way classification. [2]

(ii) For each of the four types of process:

(a) give an example of a statistical model

(b) write down a problem of relevance to the operation of:

- a food retailer
- a general insurance company

[6]

[Total 8]

8 There is a population of ten cats in a certain neighbourhood. Whenever a cat which has fleas meets a cat without fleas, there is a 50% probability that some of the fleas transfer to the other cat such that both cats harbour fleas thereafter. Contacts between two of the neighbourhood cats occur according to a Poisson process with rate μ, and these meetings are equally likely to involve any of the possible pairs of individuals. Assume that once infected a cat continues to have fleas, and that none of the cats' owners has taken any preventative measures.

(i) If the number of cats currently infected is x, explain why the number of possible pairings of cats which could result in a new flea infection is $x(10 - x)$. [1]

(ii) Show how the number of infected cats at any time, $X(t)$, can be formulated as a Markov jump process, specifying:

(a) the state space
(b) the Kolmogorov differential equations in matrix form

[4]

(iii) State the distribution of the holding times of the Markov jump process. [2]

(iv) Calculate the expected time until all the cats have fleas, starting from a single flea-infected cat. [2]
[Total 9]

9 (i) Prove that, under Gompertz's Law, the probability of survival from age x to age $x + t$, ${}_t p_x$, is given by:

$$ {}_t p_x = \left[\exp\left(\frac{-B}{\ln c} \right) \right]^{c^x (c^t - 1)} . \quad [3]$$

For a certain population, estimates of survival probabilities are available as follows:

$${}_1 p_{50} = 0.995$$
$${}_2 p_{50} = 0.989 .$$

(ii) Calculate values of B and c consistent with these observations. [3]

(iii) Comment on the calculation performed in (ii) compared with the usual process for estimating the parameters from a set of crude mortality rates. [3]
[Total 9]

10 Let T_x be a random variable denoting future lifetime after age x, and let T be another random variable denoting the lifetime of a new-born person.

(i) (a) Define, in terms of probabilities, $S_x(t)$, which represents the survival function of T_x.

(b) Derive an expression relating $S_x(t)$ to $S(t)$, the survival function of T.

[2]

(ii) Define, in terms of probabilities involving T_x, the force of mortality, μ_{x+t}.

[1]

The Weibull distribution has a survival function given by

$$S_x(t) = \exp\left(-(\lambda t)^{\beta}\right),$$

where λ and β are parameters (λ, β > 0).

(iii) Derive an expression for the Weibull force of mortality in terms of λ and β.

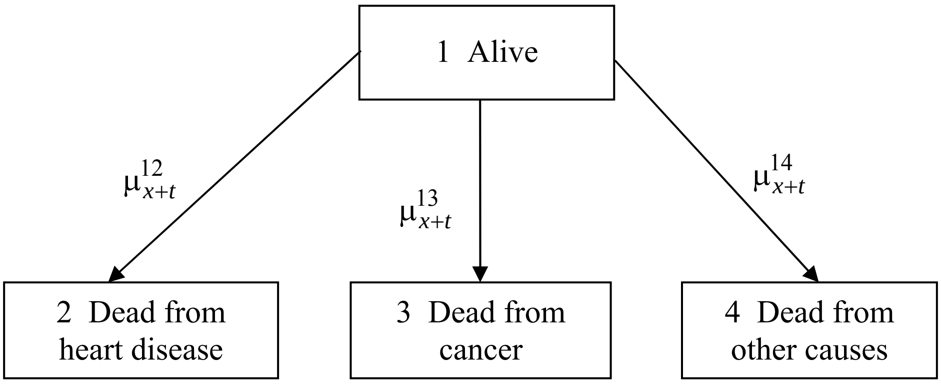
[3]

(iv) Sketch, on the same graph, the Weibull force of mortality for $0 \le t \le 5$ for the following pairs of values of λ and β:

$\lambda = 1,\ \beta = 0.5$
$\lambda = 1,\ \beta = 1.0$
$\lambda = 1,\ \beta = 1.5$

[4]
[Total 10]

11 An investigation into mortality by cause of death used the four-state Markov model shown below.



(i) Show from first principles that

$$\frac{\partial}{\partial t}\, {}_t p_x^{12} = \mu_{x+t}^{12}\, {}_t p_x^{11}\,.$$ [5]

The investigation was carried out separately for each year of age, and the transition intensities were assumed to be constant within each single year of age.

(ii) (a) Write down, defining all the terms you use, the likelihood for the transition intensities.

(b) Derive the maximum likelihood estimator of the force of mortality from heart disease for any single year of age. [5]

The investigation produced the following data for persons aged 64 last birthday:

Total waiting time in the state Alive	1,065 person-years
Number of deaths from heart disease	34
Number of deaths from cancer	36
Number of deaths from other causes	42

(iii) (a) Calculate the maximum likelihood estimate (MLE) of the force of mortality from heart disease at age 64 last birthday.

(b) Estimate an approximate 95% confidence interval for the MLE of the force of mortality from heart disease at age 64 last birthday. [3]

(iv) Discuss how you might use this model to analyse the impact of risk factors on the death rate from heart disease and suggest, giving reasons, a suitable alternative model. [3]

[Total 16]

12 A motor insurer operates a no claims discount system with the following levels of discount {0%, 25%, 50%, 60%}.

The rules governing a policyholder's discount level, based upon the number of claims made in the previous year, are as follows:

- Following a year with no claims, the policyholder moves up one discount level, or remains at the 60% level.

- Following a year with one claim, the policyholder moves down one discount level, or remains at 0% level.

- Following a year with two or more claims, the policyholder moves down two discount levels (subject to a limit of the 0% discount level).

The number of claims made by a policyholder in a year is assumed to follow a Poisson distribution with mean 0.30.

(i) Determine the transition matrix for the no claims discount system. [3]

(ii) Calculate the stationary distribution of the system, π. [5]

(iii) Calculate the expected average long term level of discount. [1]

The following data shows the number of the insurer's 130,200 policyholders in the portfolio classified by the number of claims each policyholder made in the last year. This information was used to estimate the mean of 0.30.

No claims	96,632
One claim	28,648
Two claims	4,400
Three claims	476
Four claims	36
Five claims	8

(iv) Test the goodness of fit of these data to a Poisson distribution with mean 0.30. [5]

(v) Comment on the implications of your conclusion in (iv) for the average level of discount applied. [2]
[Total 16]

END OF PAPER

Subject CT4 — Models
Core Technical

EXAMINERS' REPORT

April 2009

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.

R D Muckart
Chairman of the Board of Examiners

June 2009

Comments

Comments on solutions presented to individual questions for the April 2009 paper are given below.

Q1 Answers to this question were satisfactory. Most candidates realised that graduation by reference to a standard table was potentially appropriate, and that graphical graduation might have to be used as a last resort. Credit was given for sensible points other than those mentioned in the specimen solution below.

Q2 In part (ii) some explanation of the correct possible values of a and b was required for full credit. A common error in part (ii) (a) was to write 0 < a < 1 and 0 < b < 1, ignoring the possibility that a and b could equal 1.

Q3 This bookwork question was well answered by many candidates. Credit was given for sensible points other than those mentioned in the specimen solution below.

Q4 Answers to parts (i) and (ii) were generally good, with a substantial proportion of candidates scoring full marks. Part (iii) was much less convincingly answered. Although not all the points mentioned in the specimen solutions below were required for full credit, many candidates only included the briefest of comments, and consequently scored few marks.

Q5 Most candidates simply wrote down the formula for Pr[G=t] (which is given in the book of Formulae and Tables) and then explained what each bracketed expression in the formula meant. Few candidates gave more than the briefest explanation of why the test is useful, and what it is designed to achieve, and still fewer gave any indication of how the test was to be performed.

Q6 Answers to this question were very disappointing. Although this was slightly more demanding than some exposed-to-risk questions in the past, many candidates seemed to have little notion of how to approximate the central exposed to risk.

Q7 This question was generally well answered, although part (ii) was less well answered than similar questions on previous papers in which examples relevant to actuarial work were asked for. Marks were deducted in part (ii) for problems which seemed trivial, or where essentially the same examples were given for more than one class of models.

Q8 Few candidates made a serious attempt at this question. Many answers consisted of an attempt at part (i) followed by a description of the state space in part (ii)(a), the general expression for the Kolmogorov equations, and a statement in part (iii) that the distribution of holding times was exponential. Few candidates attempted to write down the matrix in part (ii). Note that credit was given in part (iv) for errors carried forward from incorrect matrices in part (ii).

Q9 Part (i) of this question was well answered by a good proportion of candidates. Fewer managed to calculate the values of B and c in part (ii), partly due to algebraic errors. Credit was given for the calculation of B to candidates who calculated an incorrect value for c but then correctly computed the value of B which corresponded to their value of c. Part (iii) was poorly answered, with many candidates offering no comments at all.

Q10 Answers to this question were very disappointing. Parts (i) and (ii) were bookwork based on the Core Reading, yet many candidates seemed not to understand what was required. Part (iii) was rather better answered. Candidates who derived an incorrect hazard function in part (iii) could score full credit in part (iv) for correct sketches of these incorrect hazards. Indeed, of the relatively small number of candidates who scored highly for the sketches in part (iv), some did indeed produce correct plots of incorrect (and sometimes much more complicated) hazard functions.

Q11 This question was well answered by many candidates. The only general weaknesses were steps missing in part (i) and the lack of explanation of where the approximate variance came from in part (iii)(b). In part (iv), an encouraging number of candidates realised that the Cox model was an obvious alternative model, though few made any further comments on how it might be applied to the problem mentioned in the question.

Q12 This question was also well answered by the majority of candidates. Many scored full marks on parts (i), (ii) and (iii), and made a good attempt at part (iv). The comments asked for in part (v) were, however, much less convincingly made. In part (iv), several candidates combined the two categories "4 claims" and "5 claims" because the expected value was small. Full credit was given for this if the chi-squared statistic was computed correctly, and the number of degrees of freedom was correct for this alternative. However, candidates who performed the test on the reduced number of categories "0 claims", "1 claim" and "2 or more claims" were penalised.

1 Graduation by reference to a standard table might be appropriate, if a suitable standard table could be found.

However the fact that the company insures non-standard lives makes it unlikely that a suitable standard table would exist.

Graphical graduation might be used if no suitable standard table can be found.

However it is a last resort as it is difficult to obtain results which are smooth and which adhere to the data.

Graduation using a parametric formula is unlikely to be appropriate as the amount of data in this investigation is likely to be small and it is unlikely that the company will want to produce a standard table.

2 (i) A Markov chain is a stochastic process with discrete states operating in discrete time in which the probabilities of moving from one state to another are dependent only on the present state of the process.

EITHER

If the transition probabilities are also independent of time.

OR

If the *l*-step transition probabilities are dependent only on the time lag, the chain is said to be time-homogeneous.

(ii) (a) In this case the chain is irreducible if the transition probability out of each state is non-zero (or, equivalently, if it is possible to reach the other state from both states)

So requires $0 < a \le 1$ and $0 < b \le 1$

(b) The chain is only periodic if the chain must alternate between the states.

So $a = 1$ and $b = 1$.

3 Benefits

Complex systems with stochastic elements can be studied.

Different future policies or possible actions can be compared.

In models of complex systems we can control the experimental conditions and thus reduce the variance of the results without upsetting their mean values.

Can calibrate to observed data and hence model interdependencies between outcomes.

Often models are the only practicable means of answering actuarial questions.

Systems with a long time-frame can be studied and results obtainedrelatively quickly.

Limitations

Time or cost or resources required for model development.

In a stochastic model, many independent runs of the model are needed to obtain results for a given set of inputs.

Models can look impressive and there is a danger this results in false sense of confidence.

Poor or incredible data input or assumptions will lead to flawed output.

Users need to understand the model and the uses to which it can safely be put — the model is not a "black box".

It is not possible to include all future events in a model (e.g. change in legislation).

Interpreting the results can be a challenge.

Any model will be an approximation.

Models are better for comparing the impact of input variations than for optimising outputs.

4 (i) (a) Under UDD the number of deaths between exact ages 30 and 35 years is half the number of deaths between exact ages 30 and 40 years.

So the number of deaths between exact ages 30 and 35 years is

½(98,617 – 97,952) = 332.5

and ${}_5q_{30} = \dfrac{332.5}{98,617} = 0.0033716$.

(b) Let the constant force of mortality be μ.

Then, since ${}_tp_x = \exp\left(-\int_0^t \mu_{x+s}\,ds\right)$,

${}_{10}p_{30} = \exp(-10\mu)$

so

$$\mu = \frac{-\log_e({}_{10}p_{30})}{10} = \frac{-\log_e(97,952/98,617)}{10} = 0.0006766 .$$

${}_5q_{30} = 1 - {}_5p_{30} = 1 - \exp(-5\mu)$

$= 1 - \exp[(-5)(0.0006766)] = 0.0033773$.

(ii) EITHER

The number of survivors to exact age 35 years is

$98,617\,{}_5p_{30} = 98,617(1 - {}_5q_{30})$,

so for UDD this is

$98,617(1 - 0.0033716) = 98,284.5$,

and under a constant force of mortality this is

$98,617(1 - 0.0033773) = 98,283.9$.

OR

Under UDD the number of survivors to exact age 35 years is (98,617 + 97,952)/2 = 98,284.5.

Under a constant force of mortality the number of survivors to exact age 35 years is given by

$$\sqrt{98,617*97,952} = 98,283.9$$

(iii) The actual number of survivors to exact age 35 years is higher (or, equivalently, mortality is lighter) than that under either the UDD or the constant force assumptions.

The actual number of survivors implies that there were 258 deaths between ages 30 and 35 years and 407 deaths between ages 35 and 40 years.

The actual data reveal that the force of mortality is higher between ages 35 and 40 years than it is between ages 30 and 35 years for females in English Life Table 15, which suggests that the force of mortality is increasing over this age range.

The assumption of UDD implies an increasing force of mortality.

The actual force of mortality seems to be increasing even faster than is implied by UDD.

A constant force of mortality is unlikely to be realistic for this age range.

Used over a 10-year age span the assumption of UDD is unlikely to be appropriate, whereas used over single years of age it is acceptable.

5 Suppose we have a set of n crude mortality rates for a given age range x to $x + n - 1$, and we wish to compare them to a standard set of n mortality rates for the same age range.

If the mortality underlying the crude rates is the same as that of the standard set of rates (the null hypothesis), then we should expect the difference between the two sets of rates to be due only to sampling variability.

The grouping of signs test tests the null hypothesis by examining the number of groups of consecutive positive deviations among the n ages, where a positive deviation occurs when the crude rate exceeds the corresponding rate in the standard set.

Suppose there are a total of m positive deviations, $n - m$ negative deviations and G positive groups.

Then the number of possible ways to arrange t positive groups among $n - m$ negative deviations is $\binom{n-m+1}{t}$.

There are $\binom{m-1}{t-1}$ ways to arrange *m* positive signs into *t* positive groups.

There are $\binom{n}{m}$ ways to arrange *m* positive and *n* – *m* negative signs.

Therefore the probability of exactly *t* positive groups is

$$\Pr[G = t] = \frac{\binom{n-m+1}{t}\binom{m-1}{t-1}}{\binom{n}{m}}$$

The grouping of signs test then evaluates $\Pr[t \le G]$ under the null hypothesis.

If this is less than 0.05 we reject the null hypothesis at the 5% level.

6 (i) (a) The relevant recovery rates can be estimated as

$$r_x = \frac{d_x}{E_x^c}, \quad x = 0, 1, 2, \ldots \text{ months}$$

where d_x is the number of persons recovering in the calendar month that was *x* months after the calendar month of their operation, and E_x^c is the central exposed to risk.

(b) We need to ensure that the E_x^c correspond to the data on persons recovering

The hospital's data imply a calendar month rate interval for the recoveries, running from the first day of each month until the last day of each month.

Using the monthly "census" data, a definition of E_x^c which corresponds to the deaths data can be obtained as follows.

We observe $P_{x,t}$ = number of lives under observation for whom the time elapsing since the operation was between $x - ½$ and $x + ½$ months, where *t* is the time in months since 1 January 2008.

Therefore, using the census formula:

$$E_x^c = \int_0^{12} P^*_{x,t}\, dt = \sum_0^{11} \tfrac{1}{2}\left(P^*_{x,t} + P^*_{x+1,t+1}\right),$$

where $P^*_{x,t} = \tfrac{1}{2}(P_{x-1,t} + P_{x,t})$.

We assume all months are the same length, and that the numbers in the hospital vary linearly across each month.

(ii) At the start of the rate interval, durations since the operation range from $x - 1$ to x months, so the average duration is $x - ½$, assuming operations take place evenly across the month.

r_x estimates the recovery rate at the mid-point of the rate interval.

This is exactly x months since the operation, so $f = 0$.

7 (i) Processes can be classified, first, according to whether their state space (i.e. the range of states they can possibly occupy) is discrete or continuous

For processes operating in both discrete and continuous state space the time domain can either be discrete or continuous

Therefore we have four possible types of process

EITHER

2 types of state space × 2 types of time domain

OR

State space	Time domain
Discrete	Discrete
Discrete	Continuous
Continuous	Discrete
Continuous	Continuous

(ii)

Type of process	*Statistical model*	*Problem of relevance to food retailer*	*Problem of relevance to a general insurer*
SS Discrete/ T Discrete	Markov chain Markov jump chain Counting process Random walk	Whether or not particular product out of stock at the end of each day	No claims bonus
SS Discrete/ T Continuous	Counting process Poisson process Markov jump process Compound Poisson process	Rate of arrival of customers in shop	Number of claims received monitored continuously
SS Continuous/ T Discrete	ARIMA time series model General random walk White noise	Value of goods in stock at the end of each day	Total amount insured on a certain type of policy valued at the end of each month
SS Continuous/ T Continuous	Compound Poisson process Brownian motion Ito process	Volume (or value) of trade in shop over a continuous period of time	Value of claims arriving monitored continuously

8 (i) There are x infected cats and hence $10 - x$ uninfected cats.

Flea transmission requires one of the x infected cats to meet one of the $(10 - x)$ uninfected cats.

(ii) The total number of pairings of cats is $\binom{10}{2} = 45$.

So the probability of a meeting resulting in an increase in the number of cats with fleas is $0.5x(10 - x)/45$.

As this depends only on the number of cats currently infected, and meetings occur according to a Poisson process, the number of infected cats over time follows a Markov jump process.

(a) The state space is the number of cats infected {0,1,2,,…..10}

(b) The generator matrix is

$$A = \frac{\mu}{90}\begin{pmatrix} 0 & 0 & & & & & & & & & \\ & -9 & 9 & & & & & & & & \\ & & -16 & 16 & & & & & & & \\ & & & -21 & 21 & & & & & & \\ & & & & -24 & 24 & & & & & \\ & & & & & -25 & 25 & & & & \\ & & & & & & -24 & 24 & & & \\ & & & & & & & -21 & 21 & & \\ & & & & & & & & -16 & 16 & \\ & & & & & & & & & -9 & 9 \\ & & & & & & & & & & 0 \end{pmatrix}$$

Kolmogorov's equations:

EITHER

forward form $\frac{d}{dt}P(t) = P(t)A$

OR

backward form $\frac{d}{dt}P(t) = AP(t)$

(iii) Holding times are exponentially distributed.

With mean $\frac{90}{\mu x(10-x)}$ OR parameter $\frac{\mu x(10-x)}{90}$.

(iv) Total expected time is the sum of the mean holding times.

$$= \frac{90}{\mu}\sum_{x=1}^{9}\frac{1}{x(10-x)} = \frac{90}{\mu}\left(\frac{1}{9}+\frac{1}{16}+\frac{1}{21}+\frac{1}{24}+\frac{1}{25}+\frac{1}{24}+\frac{1}{21}+\frac{1}{16}+\frac{1}{9}\right)$$

$= 50.92/\mu$

9 (i) Under Gompertz's Law

$\mu_x = Bc^x$.

Since

$${}_t p_x = \exp\left(-\int_0^t \mu_{x+w}\,dw\right),$$

we have ${}_t p_x = \exp\left(-\int_0^t Bc^{x+w}\,dw\right) = \exp\left(-\left|\dfrac{Bc^x c^w}{\ln c}\right|_0^t\right)$,

which is $\exp\left(-\dfrac{\left[Bc^x c^t - Bc^x\right]}{\ln c}\right) = \left[\exp\left(\dfrac{-B}{\ln c}\right)\right]^{c^x(c^t-1)}$.

(ii) Define $Q = \left[\exp\left(\dfrac{-B}{\ln c}\right)\right]^{c^{50}}$

ln 0.995 = (c – 1) ln Q

ln 0.989 = (c^2 – 1) ln Q

$$\frac{(c^2-1)}{(c-1)} = \frac{(c-1)(c+1)}{(c-1)} = 2.20665$$

c = 1.20665

Therefore Q = 0.976036128

$$\left[\exp\left(\frac{-B}{\ln 1.20665}\right)\right]^{1.20665^{50}} = 0.976036128$$

$B = 3.797*10^{-7}$.

(iii) In this example, only two observations are provided so there is an analytical solution to the Gompertz model.

This is unrealistic as in general a graduation process would be used to provide a fit to a set of crude rates.

This could be done by weighted least squares or maximum likelihood.

The more general graduation process allows the fitting of more complex models from the Gompertz-Makeham family which have the form

μ_x = polynomial(1) + exp(polynomial(2))

the parameters of which cannot always so easily be estimated by the method used in part (ii).

10 (i) (a) $S_x(t) = \Pr[T_x > t]$

(b) EITHER

Since $\Pr[T_x > t] = \Pr[T > x+t \mid T > x] = \dfrac{\Pr[T > x+t]}{\Pr[T > x]}$

and $S(t) = \Pr[T > t]$,

then $S_x(t) = \dfrac{S(x+t)}{S(x)}$.

OR

Since $S_x(t) = {}_t p_x$, then using the consistency principle ${}_{x+t}p_0 = {}_t p_x \cdot {}_x p_0$

Therefore ${}_t p_x = S_x(t) = \dfrac{{}_{x+t}p_0}{{}_x p_0} = \dfrac{S(x+t)}{S(x)}$.

(ii) EITHER

$$\mu_{x+t} = -\frac{1}{\Pr[T_x > t]}\frac{d}{dt}[\Pr(T_x > t)]$$

OR

$$\mu_{x+t} = \lim_{h \to 0^+} \frac{1}{h}\left(\Pr[T_x \le t+h \mid T_x > t\right)$$

(iii) EITHER

If the density function of T_x is $f_x(t)$, then we can write

$$f_x(t) = S_x(t)\mu_{x+t} = -\frac{d}{dt}S_x(t)$$

Therefore $\mu_{x+t} = -\frac{1}{S_x(t)}\frac{d}{dt}S_x(t)$

If $S_x(t) = \exp\left(-(\lambda t)^\beta\right)$, therefore, we have

$$\mu_{x+t} = -\frac{1}{\exp\left(-(\lambda t)^\beta\right)}\frac{d}{dt}\exp\left(-(\lambda t)^\beta\right)$$

$$\mu_{x+t} = -\frac{1}{\exp\left(-(\lambda t)^\beta\right)}\left(\exp\left(-(\lambda t)^\beta\right)\right)\left(-\lambda^\beta \beta t^{\beta-1}\right) = \lambda^\beta \beta t^{\beta-1}$$

OR

$$S_x(t) = \exp\left[-\int_0^t \mu_{x+s}ds\right] = \exp\left[-(\lambda t)^\beta\right].$$

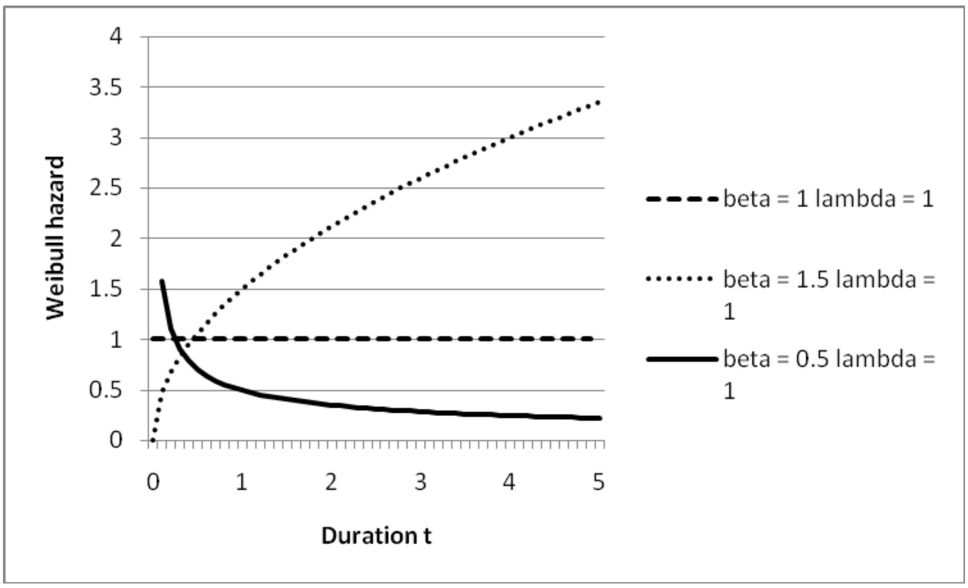
So

$$\frac{d}{dt}\left[\int_0^t \mu_{x+s}ds\right] = \mu_{x+t} = \frac{d}{dt}\left[(\lambda t)^\beta\right],$$

and hence

$$\mu_{x+t} = \beta\lambda^\beta t^{\beta-1}.$$

(iv)



11 (i) Condition on the state occupied at t.

We have

$${}_{t+dt}p_x^{12} = {}_t p_x^{11}\,{}_{dt}p_{x+t}^{12} + {}_t p_x^{12}\,{}_{dt}p_{x+t}^{22}.$$

since it is impossible to leave states 3 and 4 once entered.

Also, ${}_{dt}p_{x+t}^{22} = 1$,

since state 2 is an absorbing state.

We now assume that, for small dt,

$${}_{dt}p_{x+t}^{12} = \mu_{x+t}^{12}dt + o(dt)$$

where $o(dt)$ is the probability that a life makes two or more transitions in the time interval dt, and

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

Substituting for ${}_{dt}p_{x+t}^{12}$ gives

$${}_{t+dt}p_x^{12} = \mu_{x+t}^{12}\,{}_t p_x^{11}dt + {}_t p_x^{12} + o(dt)$$

Thus

$${}_{t+dt}p_x^{12} - {}_t p_x^{12} = \mu_{x+t}^{12}\,{}_t p_x^{11}dt + o(dt)$$

and

$$\frac{\partial}{\partial t}\,{}_t p_x^{12} = \lim_{dt \to 0^+} \frac{{}_{t+dt}p_x^{12} - {}_t p_x^{12}}{dt} = \mu_{x+t}^{12}\,{}_t p_x^{11}$$

(ii) (a) Suppose we observe d^{12} deaths from heart disease, d^{13} deaths from cancer and d^{14} deaths from other causes.

Suppose also that we observe the waiting time for each life, and that the total observed waiting time is V, being the sum of the waiting times for each life.

Then the likelihood of the data is given by

$$L \propto \exp\left[-\left(\mu^{12}+\mu^{13}+\mu^{14}\right)V\right](\mu^{12})^{d^{12}}(\mu^{13})^{d^{13}}(\mu^{14})^{d^{14}}.$$

(b) The maximum likelihood estimator of μ^{12} isobtained by differentiating this expression (or its logarithm) with respect to μ^{12} and setting the derivative equal to zero.

Taking logarithms produces

$$\log L = -(\mu^{12}+\mu^{13}+\mu^{14})V + d^{12}\log\mu^{12} + d^{13}\log\mu^{13} + d^{14}\log\mu^{14} + K$$

(where K is a constant)

Partially differentiating this with respect to μ^{12} leads to

$$\frac{\partial \log L}{\partial\mu^{12}} = -V + \frac{d^{12}}{\mu^{12}},$$

and setting the partial derivative equal to zero leads to the solution

$$\hat{\mu}^{12} = \frac{d^{12}}{V}.$$

Since $\dfrac{\partial^2 \log L}{(\partial\mu^{12})^2} = -\dfrac{d^{12}}{(\mu^{12})^2}$, the second derivative is always negative and so we have a maximum.

(iii) (a) The maximum likelihood estimate of the force of mortality from heart disease is 34/1,065 = 0.0319249

(b) The variance of the maximum likelihood estimator of μ^{12} is asymptotically $\dfrac{\mu^{12}}{E[V]}$, where $E[V]$ is the expected waiting time in the state "alive" and μ^{12} is the "true" population value of the force of mortality from heart disease.

This may be approximated by using the observed force of mortality and the observed waiting time, so that an estimate of the variance is

$$\frac{0.0319249}{1,065} = 0.000029976.$$

The estimated standard error is therefore

$\sqrt{0.000029976} = 0.00547507$.

The 95% confidence interval is therefore

$$0.0319249 \pm (1.96)0.00547507 = 0.0319249 \pm 0.0107311$$
$$= (0.0212, 0.0427).$$

(iv) Using the four state model, the lives in the investigation would have to be stratified according to the risk factors and the transition intensities estimated separately for each stratum.

This is likely to run into problems of small numbers.

Using a Cox regression model with death from heart disease as the event of interest and the risk factors as covariates would avoid this problem.

Lives who died from other causes could be treated as censored at the durations when they died.

12 (i) The probability of making the relevant number of claims is:

P[0 claims] = exp(–0.3) = 0.740818
P[1 claim] = 0.3exp(–0.3) = 0.222245

So *P*[2 or more claims] = 1 – 0.740818 – 0.222245 = 0.036936

Therefore the transition matrix *P* is given by:

$$\begin{pmatrix} 0.259182 & 0.740818 & 0 & 0 \\ 0.259182 & 0 & 0.740818 & 0 \\ 0.036936 & 0.222245 & 0 & 0.740818 \\ 0 & 0.036936 & 0.222245 & 0.740818 \end{pmatrix}$$

(ii) $\pi = \pi P$

$$\pi_1 = 0.259182\pi_1 + 0.259182\pi_2 + 0.036936\pi_3 \qquad (1)$$
$$\pi_2 = 0.740818\pi_1 + 0.222245\pi_3 + 0.036936\pi_4 \qquad (2)$$
$$\pi_3 = 0.740818\pi_2 + 0.222245\pi_4 \qquad (3)$$
$$\pi_4 = 0.740818\pi_3 + 0.740818\pi_4 \qquad (4)$$
$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$$

Using (4)

$\pi_3 = [(1-0.740818)/0.740818]*\pi_4 = 0.349859\pi_4$.

In (3)

$\pi_2 = [(0.349859-0.222245)/0.740818]*\pi_4 = 0.17226\pi_4$.

Then in (2)

$\pi_1 = [(0.17226-0.036936-0.222245*0.349859)/0.740818]*\pi_4 = 0.07771\pi_4$

So

$\pi_4 = 1/(1+0.349859+0.17226+0.07771)=0.625067$
$\pi_3 = 0.218685$
$\pi_2 = 0.107674$
$\pi_1 = 0.048574$

(iii) Average discount =

60%*0.625067+50%*0.218685+25%*0.107674 = 51.13%

(iv) The total number of policyholders shown is 130,200.

Number of claims	*Probability*	*Expected Number*	*Observed*	*$(O-E)^2/E$*
0	0.740818221	96454.53	96632	0.327
1	0.222245466	28936.35	28648	2.873
2	0.03333682	4340.45	4400	0.817
3	0.003333682	434.05	476	4.054
4	0.000250026	32.55	36	0.366
5	1.50016*E*–05	1.95	8	18.771

Null hypothesis: the data come from a source where the underlying distribution of number of claims follows a Poisson distribution with mean 0.30.

The test statistic $z = \sum_i (O_i - E_i)^2 \big/ E_i$ is distributed as chi-square with (6 – 1(parameter) – 5 degrees of freedom under the null hypothesis.

This is a one-tailed test, and the upper 5% point of the chi-squared distribution with 5 degrees of freedom is 11.07.

The observed value of the test statistic is 27.2.

As 27.2 > 11.07 we reject the null hypothesis.

(v) As the goodness of test fails, the discount level calculated assuming the Poisson distribution may be incorrect.

The goodness-of-fit test fails due to a larger number of multiple claims than expected.

Conversely a higher number of policyholders make no claims than expected (within the mean of 0.30), so the average discount level may be understated.

The average discount level calculated from the data could usefully be compared with that estimated using the Poisson distribution.

END OF EXAMINERS' REPORT

EXAMINATION

8 October 2009 (am)

Subject CT4 — Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 11 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator from the approved list.

1 Describe the difference between the following assumptions about mortality between any two ages, x and y ($y > x$):

- uniform distribution of deaths
- constant force of mortality

In your answer, explain the shape of the survival function between ages x and y under each of the two assumptions. [2]

2 (i) List the key steps in constructing a new actuarial model. [4]

You work for an actuarial consultancy which is taking over responsibility for a modelling process which has previously been conducted in house by a client.

(ii) Discuss the extent to which the steps required for this task differ from those listed in your answer to (i). [2]

[Total 6]

3 (i) List the data needed for the exact calculation of a central exposed to risk depending on age. [2]

An investigation studied the mortality of persons aged between exact ages 40 and 41 years. The investigation began on 1 January 2008 and ended on 31 December 2008. The following table gives details of 10 lives involved in the investigation.

Life	*Date of 40th birthday*	*Date of death*
1	1 March 2007	–
2	1 May 2007	1 October 2008
3	1 July 2007	–
4	1 October 2007	–
5	1 December 2007	1 February 2008
6	1 February 2008	–
7	1 April 2008	–
8	1 June 2008	1 November 2008
9	1 August 2008	–
10	1 December 2008	–

Persons with no date of death given were still alive when the investigation ended.

(ii) Calculate a central exposed to risk using the data for the 10 lives in the sample. [3]

(iii) (a) Calculate the maximum likelihood estimate of the hazard of death at age 40 last birthday.

(b) Hence, or otherwise, estimate q_{40}. [2]

[Total 7]

4 (i) In the context of mortality investigations describe the principle of correspondence and give an example of a situation in which it may be hard to adhere to this principle. [2]

On 1 January 2005 a country introduced a comprehensive system of death registration, which classified deaths by age last birthday on the date of death.

The government of the country wishes to obtain estimates of the force of mortality, μ_x, by single years of age x for the period between 1 January 2005 and 1 January 2008. Annual population censuses have been taken on 30 June each year since 2004, which classify the population by age last birthday. However the only copy of the data from the population census of 30 June 2006 was lost when the computer disc on which it was stored was being transferred between government departments.

Let the population aged x last birthday on 30 June in year t be denoted by the symbol $P_{x,t}$, and the number of deaths during the period of investigation of persons aged x be denoted by the symbol d_x.

(ii) Derive an expression in terms of $P_{x,t}$ and d_x which may be used to estimate μ_x. [6]

[Total 8]

5 (i) State the Markov property. [1]

A stochastic process $X(t)$ operates with state space S.

(ii) Prove that if the process has independent increments it satisfies the Markov property. [3]

(iii) (a) Describe the difference between a Markov chain and a Markov jump process.

(b) Explain what is meant by a Markov chain being irreducible.

[2]

An actuarial student can see the office lift (elevator) from his desk. The lift has an indicator which displays on which of the office's five floors it is at any point in time. For light relief the student decides to construct a model to predict the movements of the lift.

(iv) Explain whether it would be appropriate to select a model which is:

(a) irreducible
(b) has the Markov property

[3]

[Total 9]

 PLEASE TURN OVER

6 The complaints department of a company has two employees, both of whom work five days per week.

The company models the arrival of complaints using a Poisson process with rate 1.25 per working day.

(i) List the assumptions underlying the Poisson process model. [2]

On receipt of a complaint, it is immediately assessed as being straightforward, of medium difficulty or complicated. 60% of cases are assessed as straightforward and 10% are assessed as complicated. The time taken in person-days' effort to prepare responses is assumed to follow an exponential distribution, with parameters 2 for straightforward complaints, 1 for medium difficulty complaints and 0.25 for complicated complaints.

(ii) Calculate the average number of person-days' work expected to be generated by complaints arriving during a five-day working week. [2]

(iii) Define a state space under which the number of outstanding complaints can be modelled as a Markov jump process. [2]

The company has a service standard of responding to complaints within a fixed number of days of receipt. It is considering using this Markov jump process to model the probability of failing to meet this service standard.

(iv) Discuss the appropriateness of using the model for this purpose, with reference to the assumptions being made. [3]
[Total 9]

7 A firm rents cars and operates from three locations — the Airport, the Beach and the City. Customers may return vehicles to any of the three locations.

The company estimates that the probability of a car being returned to each location is as follows:

	Car returned to		
Car hired from	*Airport*	*Beach*	*City*
Airport	0.5	0.25	0.25
Beach	0.25	0.75	0
City	0.25	0.25	0.5

(i) Calculate the 2-step transition matrix. [2]

(ii) Calculate the stationary distribution π. [3]

It is suggested that the cars should be based at each location in proportion to the stationary distribution.

(iii) Comment on this suggestion. [2]

(iv) Sketch, using your answers to parts (i) and (ii), a graph showing the probability that a car currently located at the Airport is subsequently at the Airport, Beach or City against the number of times the car has been rented. [3]
[Total 10]

8 A researcher is studying a certain incurable disease. The disease can be fatal, but often sufferers survive with the condition for a number of years. The researcher wishes to project the number of deaths caused by the disease by using a multiple state model with state space:

{H – Healthy, I – Infected, $D_{\text{(from disease)}}$ – Dead (caused by the disease), $D_{\text{(not from disease)}}$ – Dead (not caused by the disease)}.

The transition rates, dependent on age x, are as follows:

- a mortality rate from the Healthy state of $\mu(x)$
- a rate of infection with the disease $\sigma(x)$
- a mortality rate from the Infected state of $\upsilon(x)$ of which $\rho(x)$ relates to Deaths caused by the disease

(i) Draw a transition diagram for the multiple state model. [2]

(ii) Write down Kolmogorov's forward equations governing the transitions by specifying the transition matrix. [3]

(iii) Determine integral expressions, in terms of the transition rates and any expressions previously determined, for:

(a) $P_{HH}(x, x+t)$

(b) $P_{HI}(x, x+t)$

(c) $P_{HD\text{(from disease)}}(x, x+t)$

[5]
[Total 10]

9 An electronics company developed a revolutionary new battery which it believed would make it enormous profits. It commissioned a sub-contractor to estimate the survival function of battery life for the first 12 prototypes. The sub-contractor inserted each prototype battery into an identical electrical device at the same time and measured the duration elapsing between the time each device was switched on and the time its battery ran out. The sub-contractor was instructed to terminate the test immediately after the failure of the 8th battery, and to return all 12 batteries to the company.

When the test was complete, the sub-contractor reported that he had terminated the test after 150 days. He further reported that:

- two batteries had failed after 97 days
- three further batteries had failed after 120 days
- two further batteries had failed after 141 days
- one further battery had failed after 150 days

However, he reported that he was only able to return 11 batteries, as one had exploded after 110 days, and he had treated this battery as censored at that duration when working out the Kaplan-Meier estimate of the survival function.

(i) State, with reasons, the forms of censoring present in this study. [2]

(ii) Calculate the Kaplan-Meier estimate of the survival function based on the information supplied by the sub-contractor. [5]

In his report, the sub-contractor claimed that the Kaplan-Meier estimate of the survival function at the duration when the investigation was terminated was 0.2727.

(iii) Explain why the sub-contractor's Kaplan-Meier estimate would be consistent with him having stolen the battery he claimed had exploded. [4]

[Total 11]

10 An investigation into the mortality of men engaged in a hazardous occupation was carried out. The following is an extract from the results.

Age x	*Initial exposed-to-risk* E_x	*Observed deaths* θ_x	$\hat{q}_x$
30	950	12	0.0126
31	1,200	14	0.0117
32	1,200	16	0.0133
33	900	9	0.0100
34	1,000	11	0.0110
35	1,100	15	0.0136
36	800	10	0.0125
37	1,250	16	0.0128
38	1,400	17	0.0121

It was decided to graduate the results with reference to English Life Table 15 (males). The formula used for the graduation was $\overset{o}{q}_x = 10q_x^s$.

(i) Using a test of the overall fit of the graduated rates to the data, test the hypothesis that the underlying mortality of men in the hazardous occupation is in accordance with the graduation formula given above. [6]

(ii) Test the graduation using two other tests which detect different features of the graduation. For each test you apply:

(a) State the feature of the graduation it is designed to detect.
(b) Carry out the test.
(c) State your conclusion.

[7]
[Total 13]

11 A study was undertaken into the length of spells of unemployment among young people in a certain city. A sample of young people was monitored from the time they started to claim unemployment benefit until either they resumed work, or they moved away from the city. None of the members of the sample died during the study.

The study investigated the impact of age, sex and educational qualifications on the hazard of returning to work using the following covariates:

A a young person's age when he or she started claiming benefit (measured in exact years since his or her 16th birthday)

S a dummy variable taking the value 1 if the person was male and 0 if the person was female

E a dummy variable taking the value 1 if the person had passed a school leaving examination in mathematics, and 0 otherwise

with associated parameters β_A, β_S and β_E.

The investigators decided to use a Cox proportional hazards regression model for the study.

(i) Explain what is meant by a proportional hazards model. [3]

(ii) Explain why the Cox model is a popular model for the analysis of survival data. [3]

(iii) (a) Write down the equation of the model that was estimated, defining the terms you use (other than those defined above).

(b) List the characteristics of the young person to whom the baseline hazard applies. [3]

The results showed:

- The hazard of resuming work for males who started claiming benefit aged 17 years exact and who had passed the mathematics examination was 1.5 times the hazard for males who started claiming benefit aged 16 years exact but who had not passed the mathematics examination.

- Females who had passed the mathematics examination were twice as likely to take up a new job as were males of the same age who had failed the mathematics examination.

- Females who started claiming benefit aged 20 years exact and who had passed the mathematics examination were twice as likely to resume work as were males who started claiming benefit aged 16 years exact and who had also passed the mathematics examination.

(iv) Calculate the estimated values of the parameters β_A, β_S and β_E. [6]

[Total 15]

END OF PAPER

Subject CT4 — Models.
Core Technical

September 2009 Examinations

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.

R D Muckart
Chairman of the Board of Examiners

December 2009

Comments for individual questions are given with the solutions that follow.

Examiners' Comments

Comments on solutions presented to individual questions for the September 2009 paper are given below. In general, those using this report should be aware that in the case of non-numerical answers full credit could often be obtained for rather less than is given in the solutions which follow. The solutions are meant as a guide to the various points which could have been made and considered relevant.

1

A uniform distribution of deaths means

EITHER

that deaths are evenly spaced between the ages *x* and *y*.

OR

that ${}_tq_x = tq_x \quad (t \leq y - x)$

OR

that ${}_tp_x\mu_{x+t}$ is constant for $t \leq y - x$.

It also means that the survival function decreases linearly between ages *x* and *y*. The assumption of a constant force of mortality between any two ages means

EITHER

that the hazard does not change with age over this age range.

OR

that ${}_tp_x = (p_x)^t$.

This implies that the survival function decreases exponentially between *ages* x *and* y.

Answers to this straightforward bookwork question were disappointing. Although most candidates could describe the difference between a constant force of mortality and the increasing force implied by a uniform distribution of deaths, few made correct reference to the form of the survival function. An alarming number of candidates referred to survival functions which increased with age! Credit was given for graphs which correctly depicted the shape of the survival function under the two assumptions.

2

(i) Define objectives of modelling process.

Plan the modelling process and how it will be validated.

Collect and validate the data required.

Define the form of the model.

Involve experts on the real world system/get feedback on validity.

Decide on software to be used, choose random number generator etc.
Write the computer program.

Debug the program.

Analyse the output

Test the reasonableness of the output.

Consider appropriateness of response of the model to small changes in input parameters.

Communicate and document results.

[*½ mark was awarded for each point up to a maximum of 4 marks]*

(ii) Whilst in theory all steps are still required, some may take the form of reviewing the appropriateness of existing decisions made, such as how the form of the model was determined.

Extent of work will depend on whether the existing model is to be used, adapted or superseded.

An understanding of how results compare with those previously used by the company will be required.

Process maps for the existing approach, or discussions with the people running the process about what they do, may be helpful.

The scope needs to be tightly defined up front to ensure it is clear what is expected of the consultancy.

Data sources may already be established.

[½ mark was awarded for each point up to a maximum of 2 marks]

Part (i) of this question was basic bookwork and was extremely well answered. Part (ii) required more thought, but many candidates were able to write down some relevant points.

3

(i) For each life we need

EITHER date of birth OR exact age at entry into observation OR exact age at exit from observation

Date of entry into observation

Date of exit from observation

[Alternatives were given full credit, provided the information given allowed the calculation of the date of entry into and exit from observation and the life's age]

(ii) The contribution of each life to the central exposed to risk is the number of months between STARTDATE and ENDDATE, where STARTDATE is the latest of date of 40th birthday 1 January 2008 and ENDDATE is the earliest of date of 41st birthday date of death 31 December 2008

Life	*STARTDATE*	*ENDDATE*	*number of months between STARTDATE and ENDDATE*
1	1 January 2008	1 March 2008	2
2	1 January 2008	1 May 2008	4
3	1 January 2008	1 July 2008	6
4	1 January 2008	1 October 2008	9
5	1 January 2008	1 February 2008	1
6	1 February 2008	31 December 2008	11
7	1 April 2008	31 December 2008	9
8	1 June 2008	1 November 2008	5
9	1 August 2008	31 December 2008	5
10	1 December 2008	31 December 2008	1

Summing the number of months over the 10 lives gives a total of 53 months, which is 4.42 years, which is the central exposed to risk.

(iii)

a. The total number of deaths during the period of observation is 2. So the maximum likelihood estimate of the hazard of death is 2/4.42 = 0.4528.

b. ALTERNATIVE 1

If the hazard of death at age 40 years is μ_{40}, then

$$q_{40} = 1 - p_{40} = 1 - \exp(-\mu_{40})$$

$$= 1 - \exp(-0.4528) = 1 - 0.6358 = 0.3642.$$

ALTERNATIVE 2

If the central exposed to risk is E_{40}^c, then if we work in years

$$q_{40} \approx \frac{d_{40}}{E_{40}^c + 0.5 d_{40}}$$

$$= \frac{2}{4.42+1} = \frac{2}{5.42} = 0.3690.$$

This was well answered. A common error was to count 3 deaths rather than 2. Although 3 deaths are mentioned in the data given in the question, one of these occurred after the life's 41st birthday and so should not be included in the estimation of μ_{40}. Another common error was to forget that exposure ends at exact age 41 years. Each of these errors was only penalised once, so that calculations which followed through correctly in (iii) were awarded full marks for part (iii). Note also that candidates who made BOTH the above errors were only penalised for one, as if exposure is assumed to continue past exact age 41 years, it is consistent to count 3 deaths!

4

(i) The principle of correspondence states that a life alive at time t should be included in the exposure at age x at time t if and only if, were that life to die immediately, he or she would be counted in the deaths data at age x. Problems in adhering to this can arise when the deaths data and the exposed-to-risk data come from two different sources. These may classify lives differently.

(ii) Since deaths are classified by age last birthday at date of death, a central exposed to risk which corresponds to the deaths data is given by

$$E_x^c = \int_{t=0}^{t=3} P_{x,t}$$

where $P_{x,t}$ is the population aged x last birthday at time t, and t is measured in years since 1 January 2005. We have censuses on 30 June 2004, 30 June 2005, 30 June 2007 and 30 June 2008.

Assuming that the population varies linearly across the period between each successive census for which we have data the population aged x last birthday on 1 January 2005 is equal to

$$\tfrac{1}{2}(P_{x,30/6/2004} + P_{x,30/6/2005})$$

and the population aged x last birthday on 1 January 2008 is equal to

$$\tfrac{1}{2}(P_{x,30/6/2007} + P_{x,30/6/2008}).$$

Dividing the period of the investigation into three sub-periods

from 1 January 2005 to 30 June 2005

from 30 June 2005 to 30 June 2007

from 30 June 2007 to 1 January 2008

and applying the trapezium rule to each sub-period produces the following exposed to risk for persons aged x last birthday

For the sub-period between 1 January 2005 and 30 June 2005

$$\tfrac{1}{2}\left[\tfrac{1}{2}(P_{x,1/1/2005} + P_{x,30/6/2005})\right]$$
$$= \tfrac{1}{2}\left[\tfrac{1}{2}(\tfrac{1}{2}(P_{x,30/6/2004} + P_{x,30/6/2005}) + P_{x,30/6/2005})\right]$$

For the sub-period between 30 June 2005 and 30 June 2007

$$2\left[\tfrac{1}{2}(P_{x,30/6/2005} + P_{x,30/6/2007})\right]$$

For the sub-period between 30 June 2007 and 1 January 2008

$$\tfrac{1}{2}\left[\tfrac{1}{2}(P_{x,30/6/2007} + P_{x,1/1/2008})\right]$$
$$= \tfrac{1}{2}\left[\tfrac{1}{2}(P_{x,30/6/2007} + \tfrac{1}{2}(P_{x,30/6/2007} + P_{x,30/6/2008}))\right]$$

Summing these gives

$$E_x^c = \tfrac{1}{8}P_{x,30/6/2004} + \tfrac{1}{8}P_{x,30/6/2005} + \tfrac{1}{4}P_{x,30/6/2005} + P_{x,30/6/2005}$$
$$+P_{x,30/6/2007} + \tfrac{1}{4}P_{x,30/6/2007} + \tfrac{1}{8}P_{x,30/6/2007} + \tfrac{1}{8}P_{x,30/6/2008}$$

which simplifies to

$$E_x^c = \tfrac{1}{8}P_{x,30/6/2004} + 1\tfrac{1}{8}P_{x,30/6/2005} + 1\tfrac{1}{8}P_{x,30/6/2007} + \tfrac{1}{8}P_{x,30/6/2008}.$$

The force of mortality may be estimated using the formula

$$\mu_x = \frac{d_x}{E_x^c},$$

where d_x denotes deaths to persons aged x last birthday when they died.

This was very poorly answered. It was perhaps rather more difficult than some exposed-to-risk questions in previous examination papers, but nevertheless the standard of most attempts was disappointing. In part (ii) credit was given for various alternative approximations provided that they were explained clearly.

5

(i) The Markov property states that the future development of a process can be predicted from its present state alone without reference to its past history.

(ii) Formally, for times $s_1 < s_2 < ... < s_n < s < t$ and for states $x_1, x_2, ..., x_n, x$ in the state space S and all subsets A of S, the Markov property can be written

$$\Pr[X(t) \in A \mid X(s_1) = x_1, X(s_2) = x_2, ..., X(s_n) = x_n, X(s) = x] = \Pr[X_t \in A \mid X(s) = x]$$

For independent increments we can write

$$\Pr[X(t) \in A \mid X(s_1) = x_1, X(s_2) = x_2, ..., X(s_n) = x_n, X(s) = x]$$
$$= \Pr[X(t) - X(s) + x \in A \mid X(s_1) = x_1, X(s_2) = x_2, ..., X(s_n) = x_n, X(s) = x]$$
$$= \Pr[X(t) - X(s) + x \in A \mid X(s) = x]$$
$$= \Pr[X(t) \in A \mid X(s) = x]$$

(iii)

a. A Markov chain is a stochastic process with the Markov property which has a discrete time set with a discrete state space. A Markov jump process is a stochastic process with the Markov property which has a continuous time set with a discrete state space.

b. A Markov chain is irreducible if any state can be reached from any other state.

(iv)

a. A lift could not serve its purpose unless it could return to each of the floors which it serves. This means an irreducible model would be appropriate.

b. Suppose, for example, the lift is currently at the third floor, with its last two states being the fourth floor and the fifth floor. In such a case the lift is more likely to be heading downwards than upwards. So the past history is likely to provide information on the likely future movement of the lift, unless the state space is very complicated (involving a number of past floors as well as the current floor). Therefore a Markov model is unlikely to be appropriate.

This question was generally well answered, apart from section (iv)(b) in which few candidates spotted the point that the direction of travel of the lift as well as its current floor will influence its next location.

6

(i) A Poisson process is a continuous-time integer valued process

$N_t,\ t \geq 0$ with

$N_0 = 0$

independent increments

EITHER

increments follow a Poisson distribution

OR

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

(ii) Average work created by a complaint is

60%* ½+ 30%* 1 + 10%*4 = 1 day.

Complaints arrive at a rate 1.25 per working day

So, work expected to be generated is 1.25*1*5 = 6.25 person-days.

(iii) As the time to handle complaints follows an exponential (memoryless) distribution, only need to know how many unanswered complaints there are –

but do need to know how many of each type. If cases are allocated randomly rather than in order, then the state space consists of (in terms of complaints not resolved):

r – straightforward,

s – medium,

t – complicated.

where $r = 0,1,2,3,4,5,\ldots.$

$s = 0,1,2,3,4,5,\ldots\ldots$

$t = 0,1,2,3,4,5,\ldots..$

(iv) EITHER The model will only give an approximation.

OR The model is not suitable for this purpose.

The model could not be used to do this without extending the state space to consider the time the complaint has been in the queue. There are only two employees, so holidays and sickness are important factors not taken into account.

The model assumes complaints are time-homogeneous. We do not know the nature of the business, but for some industries complaints would be seasonal e.g. holiday companies.

The model assumes that complaint arrivals are independent, but more complaints might be expected if the company has had a quality control problem at a particular time. If struggling to meet the service standard, action would be. Taken, such as overtime, or prioritising easy cases. Staff may be able to deal with complaints which are similar to other recent complaints very quickly, using standard 'template' responses.

The memoryless property is unlikely to be realistic as the work required to complete the case could be assessed and then worked through to a schedule.

The Markov jump process could be used to estimate the probability that a complaint is responded to within a given number of days of receipt.

So the model could be used to estimate the probability of a complaint not being responded to in the stated time, that is the failure to meet the service standard.

[½ mark was awarded for each point up to a maximum of 3 marks]

Answers to this question were disappointing. Most candidates were able to tackle the calculation in part (ii) but few correctly identified the state space in part (iii), and most only made a cursory attempt at part (iv).

7

(i) Two step transition matrix

$$= \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.75 & 0 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} . \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.75 & 0 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.375 & 0.375 & 0.25 \\ 0.3125 & 0.625 & 0.0625 \\ 0.3125 & 0.375 & 0.3125 \end{pmatrix}$$

(ii) $\pi = \pi \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.75 & 0 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$

$\pi_1 = 0.5\pi_1 + 0.25\pi_2 + 0.25\pi_3$

$\pi_2 = 0.25\pi_1 + 0.75\pi_2 + 0.25\pi_3$

$\pi_3 = 0.25\pi_1 + 0.5\pi_3$

and $\pi_1 + \pi_2 + \pi_3 = 1$

$\pi_1 = 2\pi_3$

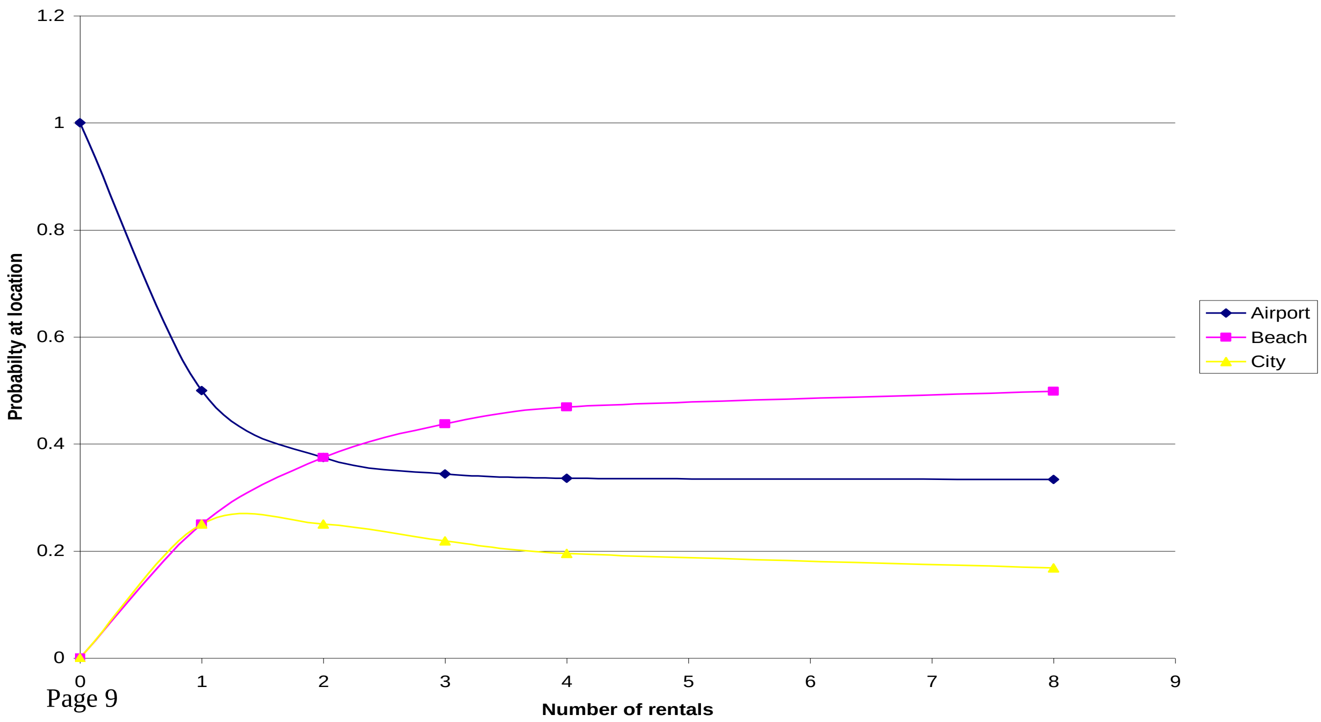
$\pi_2 = 3\pi_3$

$\pi_1 = 1/3$

$\pi_2 = 1/2$

$\pi_3 = 1/6$

(iii) The stationary distribution gives the long run probability that a particular car will be at each location. However this does not take into account the demand for hiring vehicles at each location, or the amount of space available at each location. These factors are likely to be more important in determining how many cars to base at each site.



(iv) A starts at 1, B and C at zero

Asymptote to the stationary distribution probs.

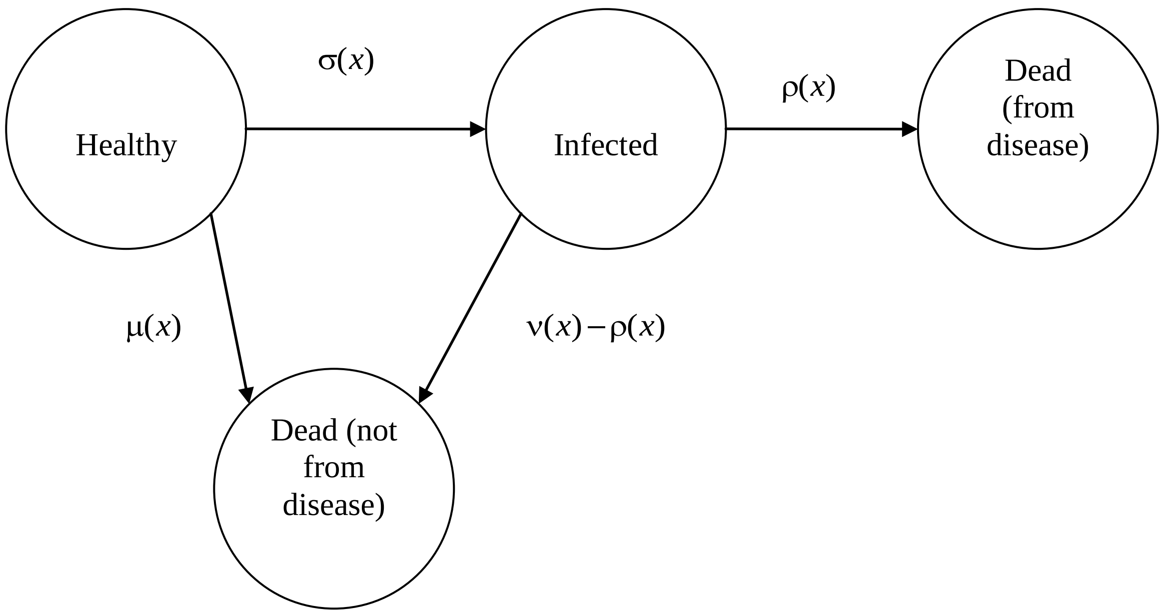
B and C same after 1 period

A and B same after 2 periods.

The calculations in parts (i) and (ii) were, as is usually the case in CT4 examinations, successfully completed by the vast majority of candidates. However only a minority made the point that, whereas the stationary distribution gives the long run probability that cars will be returned to each location, the company would be better advised to position cars at the three locations to reflect the demand for rentals. In part (iv), some candidates drew a set of histograms. Credit was given for this, provided that histograms were presented for 1 rental, 2 rentals, and the long run distribution, together with a statement that at 0 rentals the car must be at the Airport.

8

(i)



(ii) $\frac{d}{dt}P(x) = P(x)A(x)$ where with order of state space

{Healthy, Infected, Dead (not disease), Dead(from disease)}

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

(iii)

a. $P_{HH}(x, x+t) = \exp[-\int_{w=0}^{t} (\sigma(x+w) + \mu(x+w))dw]$

b. $P_{HI}(x, x+t) = \int_{w=0}^{t} P_{HH}(x, x+w).\sigma(x+w).\exp[-\int_{u=w}^{t} \nu(x+u)du].dw$

c. EITHER

$$P_{HD(\text{from disease})}(x, x+t) = \int_{w=0}^{t} P_{HI}(x, x+w).\rho(x+w).dw$$

OR (backwards alternative)

$P_{HD(\text{from disease})}(x, x+t)$

$$= \int_{w=0}^{t} P_{HH}(x+w).\sigma(x+w).P_{ID(fromdisease)}(x+w, x+t).dw.$$

Now $P_{ID(fromdisease)}(x+w, x+t) = \int_{s=w}^{t} P_{II}(x+w, x+s).\rho(x+s).1.ds$

and $P_{II}(x+w, x+s) = \exp\left[-\int_{u=w}^{s} \nu(x+u)du\right].$

So $P_{HD(\text{from disease})}(x, x+t)$

$$= \int_{w=0}^{t} P_{HH}(x+w).\sigma(x+w).\int_{s=w}^{t} \exp\left[-\int_{u=w}^{s} \nu(x+u)du\right].\rho(x+s).ds.dw$$

This question was considerably better answered than were similar questions in previous examinations. In particular, the proportion of candidates making serious attempts at part (iii) was greater than has been the case for similar questions in the past.

9

(i) Type II censoring as the study was terminated after a pre-determined number of failures. Random censoring of the device which exploded.

(ii) According to the information supplied by the sub-contractor, the Kaplan-Meier estimate of the survival function should be calculated as follows:

j	t_j	N_j	d_j	c_j	d_j/N_j	$1 - d_j/N_j$
0	0	12				
1	97	12	2	1	2/12	10/12

2	120	9	3	0	3/9	6/9
3	141	6	2	0	2/6	4/6
4	150	4	1	3	1/4	3/4

The Kaplan-Meier estimate is then

$$\hat{S}(t) = \prod_{t_j \le t} \left(1 - \frac{d_j}{N_j}\right)$$

so we have

t	$\hat{S}(t)$
$0 \le t < 97$	1
$97 \le t < 120$	5/6
$120 \le t < 141$	5/9
$141 \le t < 150$	10/27
$150 \le t$	5/18 = 0.2778

(iii) Since 5/18 is not equal to 0.2727, the sub-contractor's story is internally inconsistent. The Kaplan-Meier estimate of the survival function after the failure of the 8th battery of 0.2727 would be obtained had only 11 batteries been tested at the start, and no battery being censored, as shown in the following table.

j	t_j	N_j	d_j	c_j	d_j/N_j	$1 - d_j/N_j$
0	0	11				
1	97	11	2	0	2/11	9/11
2	120	9	3	0	3/9	6/9
3	141	6	2	0	2/6	4/6
4	150	4	1	0	1/4	3/4
					+½	+½

The Kaplan-Meier estimate is then

$$\hat{S}(t) = \prod_{t_j \le t} \left(1 - \frac{d_j}{N_j}\right)$$

so we have

t $\hat{S}(t)$

$0 \le t < 97$	1
$97 \le t < 120$	9/11
$120 \le t < 141$	6/11
$141 \le t < 150$	4/11
$150 \le t$	3/11 = 0.2727

Therefore the value of $\hat{S}(150)$ reported by the sub-contractor is consistent with him having stolen the last battery.

Many candidates scored highly on this question. Credit was given in part (i) for other types of censoring provided that a sensible reason was given. In part (iii), for full credit some kind of calculation of an alternative survival function was needed, together with an explanation of why this provided evidence to support the suggestion that the sub-contractor has stolen the battery.

10

(i) The chi-squared test is for the overall fit of the graduated rates to the data

The test statistic is $\sum z_x^2$, where

$$z_x = \frac{\theta_x - E_x \overset{o}{q}_x}{\sqrt{E_x \overset{o}{q}_x (1 - \overset{o}{q}_x)}}.$$

The calculations are shown in the table below (since $\overset{o}{q}_x$ is

small we use the approximation $z_x \approx \dfrac{\theta_x - E_x \overset{o}{q}_x}{\sqrt{E_x \overset{o}{q}_x}}$.

Age x	θ_x	$\overset{o}{q}_x$	$E_x \overset{o}{q}_x$	z_x	z_x^2
30	12	0.0091	8.645	1.141	1.302
31	14	0.0094	11.28	0.810	0.656
32	16	0.0097	11.64	1.278	1.633
33	9	0.0099	8.91	0.030	0.001
34	11	0.0106	10.60	0.123	0.015
35	15	0.0116	12.76	0.627	0.393
36	10	0.0127	10.16	-0.050	0.003
37	16	0.0138	17.25	-0.301	0.091

38	17	0.0149	20.86	-0.845	0.714
				$\sum$	4.808

The test statistic has a chi-squared distribution with degrees of freedom (d.f.) given by number of ages

– 1 (for parameter of function linking $\overset{o}{q}_x$ and q_x^s)

– some d.f. for constraints imposed by choice of standard table

The critical value of the chi-squared distribution is

11.07 with 5 d.f.

12.59 with 6 d.f.

14.07 with 7 d.f.

15.51 with 8 d.f.

16.92 with 9 d.f. at the 5% level (from tables)

Since 4.808 < 11.07 (or 12.59 etc.) there is no evidence to reject the null hypothesis that the graduated rates are the true rates underlying the crude rates.

(ii) EITHER

Signs test

a. The Signs test looks for overall bias.

b. The number of positive signs among the z_x s

is distributed Binomial (9, 0.5).

We observe 6 positive signs.

The probability of obtaining 6 or more positive signs is

(from tables)

1 – 0.7461 = 0.2539.

[Alternatively, candidates could calculate the probability of obtaining exactly 6 positive signs, which is 0.1641]

Since this is greater than 0.025 (two-tailed test)

c. we cannot reject the null hypothesis and we conclude that the graduated rates are not systematically higher or lower than the crude rates.

OR

Cumulative Deviations test

a. When applied over the whole age range, the Cumulative Deviations test looks for overall bias

b. The test statistic is

$$\frac{\sum_x \left(\theta_x - E_x \overset{o}{q}_x\right)}{\sqrt{\sum_x E_x \overset{o}{q}_x}} \sim \text{Normal}(0,1)$$

Age x	θ_x	$E_x \overset{o}{q}_x$	$\theta_x - E_x \overset{o}{q}_x$
30	12	8.645	3.355
31	14	11.28	2.72
32	16	11.64	4.36
33	9	8.91	0.09
34	11	10.60	0.40
35	15	12.76	2.24
36	10	10.16	-0.16
37	16	17.25	-1.25
38	17	20.86	-3.86
	$\sum$	112.105	7.895

So the value of the test statistic is $\dfrac{7.895}{\sqrt{112.105}} = 0.7457$

Using a 5% level of significance, we see that

$-1.96 < 0.7457 < 1.96$

c. We accept the null hypothesis at the 5% level of significance and conclude there is no overall bias in the graduation.

Grouping of Signs test

a. The Grouping of Signs test looks for runs or clumps of deviations of the same sign OR the grouping of signs test tests for overgraduation.

b. We have:

9 ages in total

6 positive deviations

3 negative deviations

We have 1 positive run

Pr[1 positive run] is therefore equal to

$$\frac{\binom{5}{0}\binom{4}{1}}{\binom{9}{6}} = \frac{4}{\left(\frac{9.8.7}{3.2}\right)} = \frac{4}{84} = 0.0476$$

Since this is less than 0.05 (using a one-tailed

test)

c. We reject the null hypothesis that the graduated rates are the true rates underlying the crude rates (OR we conclude that the graduation is unsatisfactory OR there is evidence of over-graduation).

Individual Standardised Deviations test

a. The Individual Standardised Deviations tests looks for individual large deviations at particular ages.

b. If the graduated rates were the true rates underlying the observed rates we would expect the individual deviations to be distributed Normal (0,1) and therefore only 1 in 20 z_x s should have absolute magnitudes greater than 1.96. Looking at the z_x s we see that the largest individual deviation is 1.278. Since this is less in absolute magnitude than 1.96

c. we cannot reject the null hypothesis that the graduated rates are the true rates underlying the crude rates.

Answers to this question were disappointing compared with previous years. A common error was for candidates to misread the question and to try to compare the observed number of deaths with an 'expected' number computed on the basis of the $\hat{q}_x$ *given in the question. These candidates were, in effect, examining deviations based solely on rounding! Candidates who made this error were penalised in part (i), but could gain credit for some of the alternative tests in part (ii) provided that they performed the tests correctly.*

11

(i) A proportional hazards (PH) model is a model which allows investigators to assess the impact of risk factors, or covariates, on the hazard of experiencing an event.

In a PH model the hazard is assumed to be the product of two terms, one which depends only on duration, and the other which depends only on the values of the covariates.

Under a PH model, the hazards of different lives with covariate vectors z_1 and z_2 are in the same proportion at all times:

for example in the Cox model

$$\frac{\lambda(t; z_1)}{\lambda(t; z_2)} = \frac{\exp(\beta z_1^T)}{\exp(\beta z_2^T)}.$$

(ii) Cox's model ensures that the hazard is always positive. Standard software packages often include Cox's model.

Cox's model allows the general "shape" of the hazard function for all individuals to be determined by the data, giving a high degree of flexibility while an exponential term accounts for differences between individuals.

This means that if we are not primarily concerned with the precise form of the hazard, we can ignore the shape of the baseline hazard and estimate the effects of the covariates from the data directly.

(iii)

a. $\lambda(t) = \lambda_0(t)\exp(\beta_A A + \beta_E E + \beta_S S)$, where $\lambda(t)$ is the estimated hazard and $\lambda_0(t)$ is the baseline hazard.

b. A female aged exactly 16 years when she first claimed benefit who had not passed the school mathematics examination.

(iv) "The hazard of resuming work for males aged 17 years who had passed the mathematics examination was 1.5 times the hazard for males aged 16 years who had not passed the mathematics examination" implies that

$$\frac{\exp[(\beta_A * 1) + \beta_S + \beta_E]}{\exp(\beta_S)} = \exp(\beta_A + \beta_E)$$

$$= \exp(\beta_A)\exp(\beta_E) = 1.5$$

"Females who had passed the examination were twice as likely to take up a new job as were males of the same age who had failed" implies that

$$\frac{\exp(\beta_E)}{\exp(\beta_S)} = 2$$

since the age terms cancel out.

"Females aged 20 years who had passed the examination were twice as likely to resume work as were males aged 16 years who had also passed the examination" implies that

$$\frac{\exp(\beta_A * 4)}{\exp(\beta_S)} = 2.$$

Substituting from (2) into (1) gives

$$2\exp(\beta_A)\exp(\beta_S) = 1.5$$

so

$$\exp(\beta_S) = 0.75\exp(-\beta_A).$$

Substituting into (3) gives

$$\frac{\exp[\beta_A * 4)}{0.75\exp(-\beta_A)} = 2,$$

$$\exp(5\beta_A) = 1.5$$

$$\beta_A = \frac{\log_e 1.5}{5} = 0.0811$$

From (1) then, we obtain

$$\exp(\beta_E)\exp(0.0811) = 1.5$$

$$\beta_E + 0.0811 = 0.4055$$

$$\beta_E = 0.3244.$$

Finally, from (2) we obtain

$$\frac{\exp(0.3244)}{\exp(\beta_S)} = 2$$

$$0.3244 - \beta_S = \log_e 2 = 0.6931$$

$$\beta_S = -0.3688$$

This was satisfactorily answered by many candidates. Although it is still the case than only a minority of candidates seem to understand the essential feature of a proportional hazards model that the hazard can be factorised into one part depending on duration and another part depending on the values of covariates, many candidates could list some advantages of the Cox model in part (ii). In part (iii)(b) very few candidates spotted that the baseline person was aged 16 years <u>when first claiming benefit</u>. In part (iv) candidates who failed to write down the correct equations implied by the three statements in the question were given some credit for correctly solving the equations they did produce.

END OF EXAMINERS' REPORT