INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

April 2018

Subject CT4 – Models
Core Technical

Introduction

The Examiners' Report is written by the Principal Examiner with the aim of helping candidates, both those who are sitting the examination for the first time and using past papers as a revision aid and also those who have previously failed the subject.

The Examiners are charged by Council with examining the published syllabus. The Examiners have access to the Core Reading, which is designed to interpret the syllabus, and will generally base questions around it but are not required to examine the content of Core Reading specifically or exclusively.

For numerical questions the Examiners' preferred approach to the solution is reproduced in this report; other valid approaches are given appropriate credit. For essay-style questions, particularly the open-ended questions in the later subjects, the report may contain more points than the Examiners will expect from a solution that scores full marks.

The report is written based on the legislative and regulatory context pertaining to the date that the examination was set. Candidates should take into account the possibility that circumstances may have changed if using these reports for revision.

Luke Hatter
Chair of the Board of Examiners
June 2018

© Institute and Faculty of Actuaries

A. General comments on the *aims of this subject and how it is marked*

1. The aim of the Models subject is to provide a grounding in stochastic processes and survival models and their application.

2. Subject CT4 comprises five main sections:

   (1) a study of the properties of models in general, and their uses for actuaries, including advantages and disadvantages (and a comparison of alternative models of the same processes);

   (2) stochastic processes, especially Markov chains and Markov jump processes;

   (3) models of a random variable measuring future lifetime;

   (4) the calculation of exposed to risk and the application of the principle of correspondence;

   (5) the reasons why mortality (or other decremental) rates are graduated, and a range of statistical tests used both to compare a set of rates with a previous experience and to test the adherence of a graduated set of rates to the original data.

   Throughout the subject the emphasis is on estimation and the practical application of models. Theory is kept to the minimum required in order usefully to apply the models to real problems.

3. Different numerical answers may be obtained to those shown in these solutions depending on whether figures obtained from tables or from calculators are used in the calculations but candidates are still awarded points for this. However, candidates may lose marks where excessive rounding has been used or where insufficient working is shown. Credit is given for valid solutions different from those shown below. Partial credit is also given to candidates submitting incomplete solutions with valid intermediate workings.

B. General comments on *student performance in this diet of the examination*

1. One or two questions in the April 2018 examination were on areas of the syllabus that have not been examined for some time. Performance on these questions was poor. This suggests that, in preparing for the examination, candidates are concentrating too much on making sure they can do previous examination questions, rather than learning the syllabus in general.

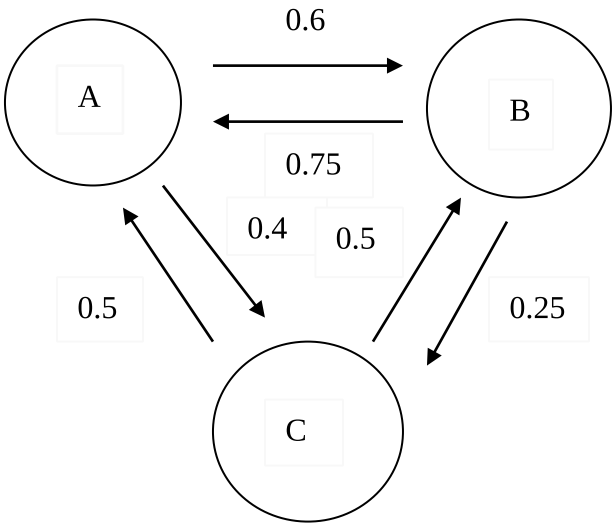
2. There was still a large number of candidates who did not read the wording of the questions closely enough, and so lost marks on straightforward sections of the paper because they did not answer the question asked.

C. Pass Mark

The Pass Mark for this examination was 60.

Solutions

Q1 (i)



[Total for part (i): 2]

(ii) The chain is irreducible [½]

because every state can eventually be reached from every other state. [½]

The chain is aperiodic [½]

because it can return to each state in multiples of 2 or 3, giving no overall period. [½]

[Total for part (ii): 2]
[Total 4]

> *This straightforward question was well answered. In part (ii), a minority of candidates thought that the Markov chain was periodic with period 2. Some candidates wrote that the Markov chain was aperiodic but gave vague or incorrect explanations.*

Q2 (i) Team *A* goes first, so at i = 1 the process can have the values 1 (if Team *A* scores) or 0 (if Team *A* misses). [½]

Team B then has a go. If Team B scores, then $X_2 = X_1 - 1$.
If Team B misses, then $X_2 = X_1$. [½]

Team A then has another go. If Team A scores, then $X_3 = X_2 + 1$.
If Team A misses, then $X_3 = X_2$. [½]

Hence possible sample paths for X_i ($i = 1, 2, 3, 4$) are:

0, 0, 0, 0
0, 0, 0, –1
0, 0, 1, 0
0, 0, 1, 1
0, –1, 0, 0
0, –1, 0, –1
0, –1, –1, –1
0, –1, –1, –2
1, 0, 0, –1
1, 0, 0, 0
1, 0, 1, 0
1, 0, 1, 1
1, 1, 1, 0
1, 1, 1, 1
1, 1, 2, process ends at $i = 3$ [2]

[Total for part (i): max. 3]

(ii) Taking the paths in (i) and considering only the first three penalties we can compute the probabilities as follows:

0, 0, 0	0.5 x 0.6 x 0.5 = 0.15
0, 0, 1	0.5 x 0.6 x 0.5 = 0.15
0, -1, 0	0.5 x 0.4 x 0.5 = 0.10
0, -1, -1	0.5 x 0.4 x 0.5 = 0.10
1, 0, 0	0.5 x 0.4 x 0.5 = 0.10
1, 0, 1	0.5 x 0.4 x 0.5 = 0.10
1, 1, 1	0.5 x 0.6 x 0.5 = 0.15
1, 1, 2	0.5 x 0.6 x 0.5 = 0.15

[2]

so the distributions are as follows

x	$Pr[X_i = x \mid i = 2]$	$Pr[X_i = x \mid i = 3]$
–1	0.2	0.10
0	0.5	0.35
1	0.3	0.40
2		0.15

[1]

[Total for part (ii): 3]

[Total 6]

This question was on a part of the syllabus which had not been tested for several sessions, but which is an important part of the Core Reading. Performance was poor, with only a minority of candidates managing to list the possible sample paths in part (i), and fewer being able to compute the probabilities in part (ii). A common error was to fail to read the question closely and to consider pairs of penalties. Some credit was given for answers which were correct on the basis of pairs of penalties.

Q3 (i) **Parametric formula**

The resultant graduation will be sufficiently smooth provided few parameters are used. [1]

It is a suitable method to produce standard tables. [½]

It can be useful to fit the same formula to several experiences to give insight into the differences between experiences. [½]

Reference to a standard table

It can be used to fit relatively small data sets in cases where a suitable standard table exists. [1]

The graduated rates should be smooth provided that a simple function is used. [1]

The standard table can provide information at extreme ages where data may be scanty
OR

The shape of the table can be used to "fill in gaps" in the data [½]

It can be useful to fit the same table to several experiences with the same link function to give insight into how the experience differs over time. [½]

Graphical graduation

It can be used for scanty data sets where no suitable standard table exists
OR
no more sophisticated method is justifiable. [1]

It enables an experienced analyst to allow for known (or likely) features of the data. [½]

It can give a quick initial feel for the rates. [½]
[Total for part (i): max. 4]

(ii) Select a suitable standard table. [½]

In making this selection, consider the nature of lives of involved, and compare their characteristics with the description of data used in a range of standard tables; and [½]

the date range for information used in preparation of the standard tables, with a general preference for using data closer in date to the period for the crude rates if possible. [½]

Select one or more link functions to try. [½]

Exploratory graphical or regression analysis may help with the selection of the link function. [½]

Estimate the parameters [½]

using a method such as maximum likelihood or least squares. [½]

Compute the graduated rates. [½]

Perform statistical tests on the graduation to check adherence to the data. [½]

If necessary, repeat some or all steps until satisfied with graduation. [½]
[Total for part (ii): max. 3]
[Total 7]

This question was generally well answered. In both parts, full credit could be gained for rather less than is written in this Examiners' Report.

Q4 (i) (a) The set of possible states that the process can take in a case where the process can only take a countable number of different values,
OR
the set of states that the process can take where it can take only distinct states. [1]

(b) A model in which at least one of the components is random in nature,
OR
a collection of random variables, one for each time point. [1]

(c) A model in which changes in state may take place at any point in time (between the start and end times). [1]

(d) Processes which operate in continuous time but which can also change value at predetermined discrete instants. [1]

[Total for part (i): 4]

(ii) Whether outputs from the model are only required at discrete points in time. [½]

The objectives of the modelling
OR
the accuracy required. [½]

The nature of the input data (which may override the nature of the process). [½]

The expertise of the analyst. [½]

Time, cost, IT resources. [½]

The nature of previous models. [½]

If simulation is required it may be easier to make the time step discrete. [½]

Continuous time models are ultimately more flexible than discrete time models. [½]

Some results for continuous time models cannot be obtained by discrete simulation at all. [½]

Regulatory requirements. [½]

The need to explain the model to a non-technical audience. [½]

[Total for part (ii): max. 3]

[Total 7]

Answers to this question were generally poor. In part (i)(a) and (c), a substantial minority of candidates provided unnecessary answers rather than offering descriptions of what the terms mean: thus "a continuous time model is a model in which time is continuous". No credit was awarded for such answers. In part (i)(d) many candidates incorrectly wrote that a stochastic process of mixed type is one in which the state space is discrete and the time domain is continuous, or vice versa. In part (ii), full credit could be obtained for less than is written in this Examiners' Report, though candidates generally did not offer factors beyond the first three points listed in the solution above.

Q5 (i) Writing the state space in the order {Bid (B), Offer (O)}, [½]

the generator matrix is:

$$\begin{array}{c} B \\ O \end{array}\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

[½]

[Total for part (i): 1]

(ii) The holding times are exponentially distributed with [1]

parameter λ in state B, [½]

and μ in state O. [½]

[Total for part (ii): 2]

(iii) $$\frac{\partial}{\partial t}\,{}_tP_s^{BB} = -\lambda.{}_tP_s^{BB} + \mu.{}_tP_s^{BO}\ .$$ [1]

$$\frac{\partial}{\partial t}\,{}_tP_s^{BO} = \lambda.{}_tP_s^{BB} - \mu.{}_tP_s^{BO}\ .$$ [1]

[Total for part (iii): 3]

(iv) We have a two-state model so:

$${}_tP_s^{BB} + {}_tP_s^{BO} = 1.$$ [½]

Substituting:

$$\frac{\partial}{\partial t}\,{}_tP_s^{BB} = -\lambda.{}_tP_s^{BB} + \mu.(1 - {}_tP_s^{BB});$$ [½]

$$\frac{\partial}{\partial t}\left[\exp((\lambda+\mu)t).{}_tP_s^{BB}\right] = \mu.\exp((\lambda+\mu)t);$$ [1]

and hence

$$\exp((\lambda+\mu)t).{}_tP_s^{BB} = \frac{\mu}{\lambda+\mu}.\exp((\lambda+\mu)t) + \text{constant.}$$ [1]

Since the process is in state Bid at time s (i.e. t = 0),
the constant is $\dfrac{\lambda}{\mu+\lambda}$, [½]

and thus ${}_tP_s^{BB} = \dfrac{\mu}{\lambda+\mu} + \dfrac{\lambda}{\lambda+\mu}.\exp(-(\lambda+\mu)t)$. [½]

[Total for part (iv): 4]
[Total 9]

This question was well answered, with many candidates scoring full marks, or close to full marks. In part (iv) a common error was to determine $\exp((\lambda+\mu)t).{}_tP_s^{BB}$ *to be 0 for t = 0, which led to the constant being evaluated as* $\dfrac{-\mu}{\lambda+\mu}$. *Where candidates wrote down incorrect Kolmogorov forward equations in part (iii), credit was given in part (iv) for sensible attempts to solve the equations that had been written.*

Q6 (i) To ensure that they follow the principle of correspondence, [1]

which states that a life alive 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 death data at age x. [½]

The deaths data "carry most information" when mortality rates are small, so we adjust the census data not the deaths data. [1]

[Total for part (i): max. 2]

(ii) Let $P^*_{x,t}$ be the population in the census aged *x* nearest birthday at time *t*. [½]

Then

$$P^*_{x,t} = \frac{1}{2}(P_{x,t} + P_{x-1,t}).$$ [½]

The required exposed-to-risk is then given by

$$\text{exposed-to-risk} = \int_{t_1}^{t_2} P^*_{x,u}du \ \text{ OR } \int_0^{10} P^*_{x,u}du$$ [½]

Where t_1 and t_2 are the start and end times.

We require the exposed to risk for a period in which t_1 = 2005 and t_2 = 2015. [½]

Using the trapezium rule [½]

we can approximate these exposed-to-risks as

$$\frac{10}{2}(P^*_{x,2005} + P^*_{x,2015}) \ .$$ [½]

Applying the same rule for the inter-censal period 2002–2015 we have

$$P^*_{x,2005} = \frac{10}{13}P^*_{x,2002} + \frac{3}{13}P^*_{x,2015} \ .$$ [½]

Hence the exposed-to-risk for *x* last birthday or the period 2005-2014 is:

$$\frac{10}{2}\left(\frac{10}{13}P^*_{x,2002} + \frac{3}{13}P^*_{x,2015} + P^*_{x,2015}\right) = \frac{10}{2}\left(\frac{10}{13}P^*_{x,2002} + \frac{16}{13}P^*_{x,2015}\right).$$ [1]

So the required exposed-to-risk for *x* nearest birthday for the period 2005–2014 is:

$$\frac{5}{2}\left(\frac{10}{13}\left(P_{x,2002} + P_{x-1,2002}\right) + \frac{16}{13}\left(P_{x,2015} + P_{x-1,2015}\right)\right).$$ [½]

[Total for part (ii): 5]

(iii) When adjusting the age data [½]

we need to assume that births are uniformly distributed across the calendar year. [½]

To use the trapezium rule [½]

we must assume that the population varies linearly between census dates. [½]

We assume that the population enumerated in the census of 1 January 2015 can be taken to be the population at the end of the calendar year 2014. [½]

[Total for part (iii): max. 2]

(iv) The deaths data are available for the years 2015 and 2016 and will not need adjusting. [½]

The exposed-to-risk will still need adjusting from an age last birthday basis to an age nearest birthday basis. [½]

To compute the population aged x last birthday in each of the calendar years 2015 and 2016 some kind of forecasting/modelling will be required. [½]

Extrapolation of the linear change at each age x between 1 January 2002 and 1 January 2015 is one option. [½]

Better might be to use the deaths in 2015 to estimate an adjusted population aged x last birthday for 1 January 2016 and to use this to estimate the exposed-to-risk at age nearest birthday for the calendar year 2015. [½]

The procedure could then be iterated to produce an exposed-to-risk for the calendar year 2016. [½]

This should be fairly accurate for the first two years immediately following a census. [½]

Data could be gathered on births or migration in and out to improve the estimate of the exposed to risk. [½]

[Total for part (iv): max. 3]

[Total 12]

Answers to part (i) of this question were generally satisfactory, though only a minority of candidates pointed out that the deaths data "carry most information" when mortality rates are small. Most candidates had a general idea of how to approach part (ii) but only a minority had the details of the estimation of the population in 2005 correct. Credit was given for part (iii) where candidates had included the assumptions in their answers to part

> *(ii) at the appropriate point in their argument. Few candidates offered extensive answers to part (iv). Most were content to say that some kind of modelling or extrapolation of the population would be required. In parts (iii) and (iv) full credit could be obtained for somewhat less than is written in this Examiners; Report.*

Q7 (i) A Markov Chain is a stochastic process [½]

with discrete states operating in discrete time [½]

in which the probabilities of moving from one state to another are dependent only on the present state of the process. [1]

OR

A Markov chain is a sequence of random variables $X_0, X_1, \ldots, X_n, \ldots$ [½]

with the following property:

$$P\left[X_n = j \middle| X_0 = i_0, X_1 = i_1, \ldots, X_{m-1} = i_{m-1}, X_m = i\right] = P\left[X_n = j \mid X_m = i\right] \quad [1]$$

for all integer times $n > m$ and states $i_0, i_1, \ldots, i_{m-1}, i, j$ in S. [½]

[Total for part (i): 2]

(ii) This needs the second order transition matrix:

$$\begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{pmatrix} . \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.31 & 0.36 & 0.33 \\ 0.31 & 0.4 & 0.29 \\ 0.27 & 0.36 & 0.37 \end{pmatrix}. \quad [1]$$

So the required probabilities are:

Andy 0.31
Brenda 0.40
Carol 0.29. [1]

OR

Calculate directly, for example:

Carol 0.2*0.5 + 0.3*0.3 + 0.5*0.2 = 0.29 +2

[Total for part (ii): 2]

(iii) The long term probabilities satisfy $\pi P = \pi$ [½]

$0.4\pi_A + 0.3\pi_B + 0.2\pi_C = \pi_A$ (1)

$0.3\pi_A + 0.5\pi_B + 0.3\pi_C = \pi_B$ (2)

$0.3\pi_A + 0.2\pi_B + 0.5\pi_C = \pi_C$ (3) [1]

and $\pi_A + \pi_B + \pi_C = 1$ (4) [½]

(2) minus (3) gives

$0.3\pi_B - 0.2\pi_C = \pi_B - \pi_C$

so

$\pi_B = 8/7\ \pi_C = 1.1429\ \pi_C$ [½]

substitute in (1)

$0.6\pi_A = 0.5429\pi_C$

$\pi_A = 0.9048\pi_C$

$(0.9048 + 1.1429$ [1]$)\ \pi_C = 1$ [½]

$\pi_C = 0.3281 = 21/64$
$\pi_B = 0.375 = 3/8$
$\pi_A = 0.2969 = 19/64$

So the required probabilities are:

Andy	19/64 or 0.2969
Brenda	3/8 or 0.375
Carol	21/64 or 0.3281.

[1]

[Total for part (iii): 4]

(iv) In this case, to know who will be "Employee of the week", we need to know who was "Employee of the week" last week as well as who made most sales this week. [½]

Suppose Andy made most sales this week. If he was "Employee of the week" last week his probability of being "Employee of the week" this week is 0, but if Brenda was "Employee of the week" last week, Andy will be "Employee of the week" this week. [1]

So additional states are needed to model "Employee of the week" as a Markov Chain. [½]

[Total for part (iv): 2]

(v) This needs nine states i.e. 3 by 3, [1]

defined by

Most Sales : "Employee of the Week" last week

Andy : Andy
Andy : Brenda
Andy : Carol
Brenda : Andy
Brenda : Brenda
Brenda : Carol
Carol : Andy
Carol : Brenda
Carol : Carol. [1]

[Total for part (v): 2]

[Total 12]

Part (i) was well answered. In part (ii), common errors were to use the third order transition matrix, or to select the wrong elements from the second order transition matrix. Part (iii) was well answered, though a substantial number of candidates lost a small amount of credit by failing to identify which probability applied to which member of staff. In parts (iv) and (v), most candidates understood why the "Employee of the week" could not be modelled as a Markov Chain with three states; however, only a small number of candidates realised than nine states were required to model the process as a Markov Chain. The most common answer was "six states". These were then typically defined as follows on the basis of Most Sales: "Employee of the week" last week {Andy: Andy, Andy: not Andy; Brenda: Brenda, Brenda: not Brenda, Carol: Carol, Carol: not Carol}. This can be shown to be insufficient to model the process as a Markov Chain by writing down the nine-state transition matrix.

Q8 (i) The Gompertz model has been shown to approximate human mortality closely in the middle and/or older ages in human populations. [1]

The Gompertz model is simple to understand, [½] and is easy to fit. [½]

[Total for part (i): 2]

(ii) EITHER

The Gompertz model is

$$\mu_x = Bc^x .$$

Hence

$$\frac{\mu_x}{\mu_y} = \frac{Bc^x}{Bc^y} = c^{x-y} \quad . \qquad [½]$$

$$\frac{\mu_{70}}{\mu_{60}} = c^{10} \qquad [½]$$

Using the values for ages 60 and 70 years we have

$$c = \sqrt[10]{\frac{0.04749}{0.02029}}$$

So c = 1.08876. [1]

Hence

$$\log_e B = \log_e 0.02029 - 60(0.08504)$$

and B = 0.0001234. [1]

OR

The Gompertz model is

$$\mu_x = \exp(\alpha_0 + \alpha_1 x)$$

Hence

$$\log_e \mu_x = \alpha_0 + \alpha_1 x \qquad [½]$$

Using the values for ages 60 and 70 years we have

$$\log_e 0.02029 = \alpha_0 + 60\alpha_1$$

$$\log_e 0.04749 = \alpha_0 + 70\alpha_1 \qquad [½]$$

Hence

$$\log_e 0.04749 = \log_e 0.02029 + 10\alpha_1$$

$$\alpha_1 = \frac{\log_e(0.04749/0.02029)}{10} = 0.085039.$$ [1]

So that

$$\alpha_0 = \log_e 0.02029 - 60(0.08504) = -9.0000$$ [1]

[Total for part (ii): 3]

(iii) We compute the actual and expected deaths in the table below.

EITHER using μ_x

Age x	*Actual deaths*	*Gompertz* μ_x	*Expected deaths*	z_x	$z_x^{\ 2}$
61	51	0.02209	50.52	0.0673	0.0045
62	55	0.02405	53.64	0.1863	0.0347
63	68	0.02619	65.44	0.3164	0.1001
64	70	0.02851	67.94	0.2498	0.0624
65	67	0.03104	67.05	–0.0061	0.0000
66	69	0.03380	73.00	–0.4683	0.2193
67	66	0.03680	79.48	–1.5121	2.2864
68	84	0.04006	78.32	0.6416	0.4116
69	88	0.04362	87.15	0.0911	0.0083

OR using $\mu_{x+1/2}$

Age x	*Actual deaths*	*Gompertz* $\mu_{x+1/2}$	*Expected deaths*	z_x	$z_x^{\ 2}$
61	51	0.02305	52.72	-0.2364	0.0559
62	55	0.02510	55.97	-0.1290	0.0166
63	68	0.02732	68.28	-0.0342	0.0012
64	70	0.02975	70.89	-0.1060	0.0112
65	67	0.03239	69.96	–0.3541	0.1254
66	69	0.03526	76.17	–0.8217	0.6752
67	66	0.03839	82.93	–1.8594	3.4572
68	84	0.04180	81.72	0.2517	0.0634
69	88	0.04551	90.94	-0.3078	0.0947

[+2]

The test is the chi-squared test, and the test statistic is $\sum_x {z_x}^2$ [½]

The null hypothesis is that the observed deaths come from a population in which the underlying mortality is described by the fitted model in (ii). [½]

The calculated value of the test statistic is 3.1267.
OR,
using $\mu_{x+1/2}$ it is 4.5009 [½]

We compare this with the critical value of the chi-squared distribution with 9 degrees of freedom at the 95% level, [½]

because we have not used the data involved in the test to estimate the expected deaths. [½]

The critical value is 16.92. [½]

Since 3.1267(4.5009) < 16.92, [½]

we have no reason to reject the null hypothesis. [½]

[Total for part (iii): 6]

(iv) The Gompertz model seems to fit the data well. [½]

However there is a relatively large negative deviations at age 67 years [½]

The exposed-to-risk is exactly the same at ages 65 to 67 years. Could there be a transcription error in the data? [½]

The model might be a better fit if MLE or weighted least squares had been used to fit the parameters. [½]

[Total for part (iv): max. 1]

[Total 12]

Parts (i) and (ii) of this question were well answered, with the majority of candidates correctly determining the parameters of the Gompertz model. Answers to part (iii) varied. Common errors included using the data given in the question, rather than the fitted model, to estimate the expected deaths, and hence effectively carrying out a chi-squared test on the rounding errors (unsurprisingly, the fit was found to be extremely good). Few candidates realised that, as the data for ages 61-69 years had not been used to estimate the parameters of the Gompertz distribution, and we can suppose the estimated μ_xs for each age to have been obtained independently, it was not necessary to deduct two degrees of freedom when carrying out the chi-squared test (it would

> *have been necessary had data for all ages been used to estimate the parameters).*

Q9 (i) Using the Markov assumption or from the Chapman-Kolmogorov equations we can write [½]

$${}_{dt+t}p_x^{13} = {}_t p_x^{11}\, {}_{dt}p_{x+t}^{13} + {}_t p_x^{12}\, {}_{dt}p_{x+t}^{23} + {}_t p_x^{13}\, {}_{dt}p_{x+t}^{33}.$$ [½]

But ${}_{dt}p_{x+t}^{33} = 1$. [½]

Assuming that, for small dt,

$${}_{dt}p_{x+t}^{ij} = \mu_{x+t}^{ij}dt + o(dt),$$ [½]

where $\lim\limits_{dt\to 0}\dfrac{o(dt)}{dt} = 0$;

then substituting, we have

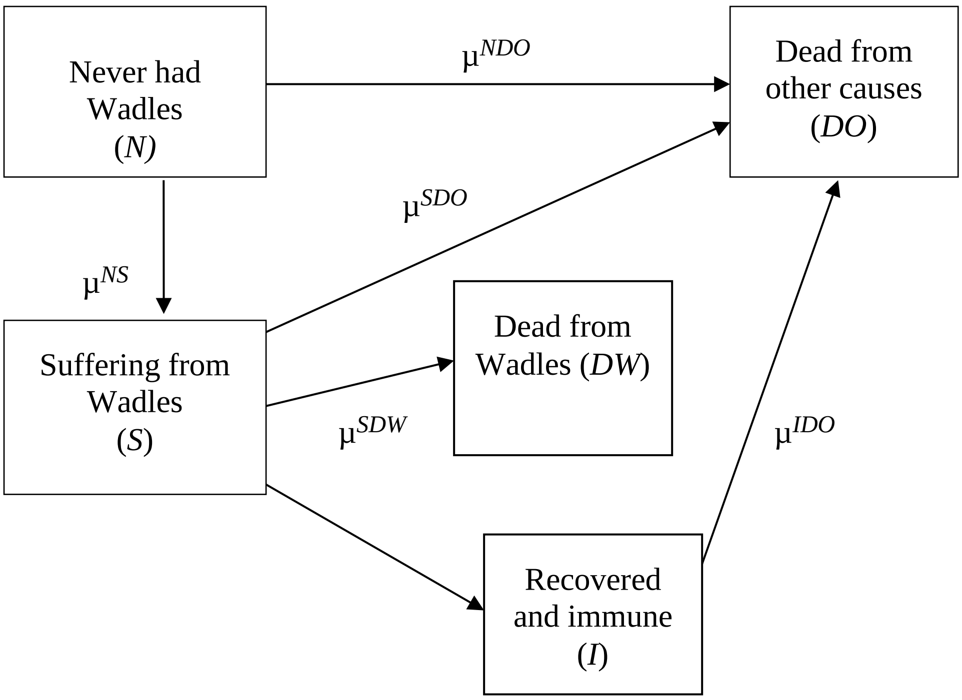
$${}_{dt+t}p_x^{13} = {}_t p_x^{11}\mu_{x+t}^{13}dt + {}_t p_x^{12}\mu_{x+t}^{23}dt + {}_t p_x^{13} + o(dt) \;,$$ [½]

so that ${}_{dt+t}p_x^{13} - {}_t p_x^{13} = {}_t p_x^{11}\mu_{x+t}^{13}dt + {}_t p_x^{12}\mu_{x+t}^{23}dt + o(dt)$, [½]

and hence $\dfrac{d}{dt}({}_t p_x^{13}) = \lim\limits_{dt\to 0}\dfrac{{}_{t+dt}p_x^{13} - {}_t p_x^{13}}{dt} = {}_t p_x^{11}\mu_{x+t}^{13} + {}_t p_x^{12}\mu_{x+t}^{23}$. [1]

[Total for part (i): 4]

(ii)



μ^{SI}

[+2]

OR

[Total for part (ii): 2]

(iii) The likelihood is

$$L \propto \exp\left\{\left(-\mu^{NS}-\mu^{NDO}\right)\nu^{N}\right\}\exp\left\{\left(-\mu^{SDW}-\mu^{SI}-\mu^{SDO}\right)\nu^{S}\right\}\exp\left\{-\mu^{IDO}\nu^{I}\right\}$$
$$\left(\mu^{NS}\right)^{d^{NS}}\left(\mu^{NDO}\right)^{d^{NDO}}\left(\mu^{SDW}\right)^{d^{SDW}}\left(\mu^{SDO}\right)^{d^{SDO}}\left(\mu^{SI}\right)^{d^{SI}}\left(\mu^{IDO}\right)^{d^{IDO}}.$$

[+2]

Here

ν^{I} is the waiting time in state *I*, [½]

d^{IJ} is the number of transitions from state *I* to state *J*, and [½]

and μ^{IJ} is the intensity of the transition from state *I* to state *J* . [½]

[Total for part (iii): max. 3]

(iv) Taking logarithms of the likelihood we have:

$\ln L = \left(-\mu^{SDW}\right)\nu^{S} + d^{SDW}\ln\left(\mu^{SDW}\right)$ plus terms not dependent on μ^{SDW}.

[½]

Differentiating with respect to μ^{SDW} gives:

$$\frac{d\left(\ln L\right)}{d\mu^{SDW}} = -\nu^{S} + \frac{d^{SDW}}{\mu^{SDW}},$$ [½]

and setting this to zero gives a maximum likelihood estimate of μ^{SDW} [½]

$$\hat{\mu}^{SDW} = \frac{d^{SDW}}{\nu^{S}}.$$ [½]

This is a maximum as the second derivative $\dfrac{d^{2}\left(\ln L\right)}{\left(d\mu^{SDW}\right)^{2}} = -\dfrac{d^{SDW}}{\left(\mu^{SDW}\right)^{2}}$ [½] must be negative. [½]

[Total for part (iv): 3]

[Total 12]

Part (i) was reasonably well answered by most candidates. Identifying the state space in part (ii) proved challenging for many candidates. Credit was given for an alternative four-state solution {Never had Wadles, Sick with Wadles, Recovered and Immune, Dead}. This is not ideal for estimating the death rate from Wadles, as a person who has Wadles may die from a cause other than Wadles, but is a reasonable answer to part (ii) as asked in the Examination Paper. A common error was to label the first state "Healthy" rather than "Never had Wadles". This is ambiguous, as a person who has recovered from Wadles is also "Healthy". In part (iii) most candidates successfully wrote down a likelihood consistent with the state space and transitions they had sketched in part (ii); they also managed to derive the maximum likelihood estimator of the death rate from Wadles in part (iv). Where the diagram in part (ii) did not include the death rate from Wadles (for example, it might include the death rate of any sick person, regardless of whether that person had Wadles or not), credit was given in part (iv) for a derivation of the maximum likelihood estimator of the transition rate which was closest to the death rate from Wadles.

Q10 (i) Random censoring when the time at which the life is censored is a random variable [½]

of the lawns damaged by sheep. [½]

Right censoring when the censoring mechanism cuts short the observations in progress [½]

such as when the sheep damage the lawns or any lawns still mossy when the study ends after 16 days. [½]

Type 1 censoring when the censoring times are known in advance [½]

of the lawns still mossy after 16 days. [½]

Interval censoring, [½]

because lawns are only checked once per day, so we do not know when in the day the moss disappeared. [½]

Non-informative censoring,

if the destruction of lawns by sheep gives no information about when they might have been free of moss.

Informative censoring,

if sheep are attracted to lawns where the moss has started to die.

[Total for part (i): max. 3]

(ii)

t_j	N_j	d_j	c_j	λ_j	$(1 - \lambda_j)$
ForLawn					
5	7	1	0	1/7	6/7
6	6	1	0	1/6	5/6
8	5	1	2	1/5	4/5
11	2	1	1	1/2	1/2
[½]	*[½]*	*[½]*		*[½]*	

[+2]

Alternative treatment

t_j	N_j	d_j	c_j	λ_j	$(1 - \lambda_j)$
3	7	1	0	1/7	6/7
4	6	1	0	1/6	5/6
5	5	1	3	1/5	4/5
10	1	1	0	1	0
[½]	*[½]*	*[½]*		*[½]*	

[+2]

From which we obtain the survival functions as follows:

ForLawn

t	$S(t)$
$0 \le t < 5$	1
$5 \le t < 6$	6/7
$6 \le t < 8$	5/7
$8 \le t < 11$	4/7
$11 \le t < 16$	2/7
[½]	*[½]*

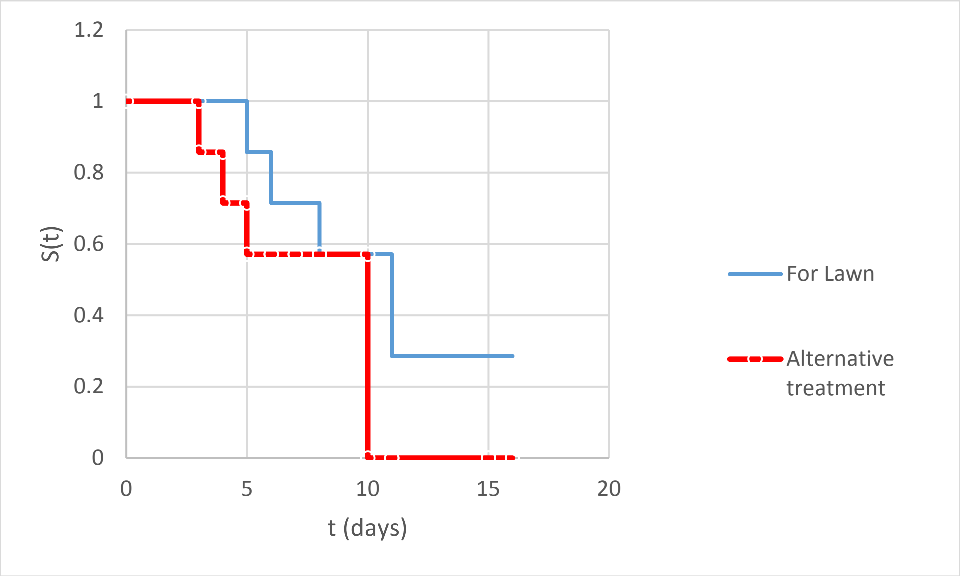
[1]

Alternative treatment

t	$S(t)$
$0 \le t < 3$	1
$3 \le t < 4$	6/7
$4 \le t < 5$	5/7
$5 \le t < 10$	4/7
$10 \le t < 16$	0

[½] *[½]* [1]
[Total for part (ii): 6]

(iii)



[+2]
[Total for part (iii): 2]

(iv) ForLawn does not seem to be better than the alternative treatment. [½]

Indeed, the survival function for the alternative treatment is below that of ForLawn at most durations, indicating that moss lasts less long under the alternative treatment than it does when treated with For Lawn. [1]

However the study is small, so the difference may not be statistically significant at conventional levels. [½]
[Total for part (iv): 2]

(v) For ForLawn we have

$$h(t \mid X = 1) = h_0(t) \exp \beta.$$ [½]

For the alternative treatment we have

$$h(t \mid X = 0) = h_0(t).$$ [½]

Where $h(t)$ is the hazard, $h_0(t)$ is the baseline hazard, β is the coefficient measuring the impact of the different treatments and X is a dummy variable taking the value 1 for ForLawn and 0 for the alternative treatment. [1]
[Total for part (v): 2]

(vi) The contributions to the partial likelihood, L, at durations where events occur are:

$t = 3$ $\qquad \dfrac{1}{7+7e^{\beta}}$ [½]

$t = 4$ $\qquad \dfrac{1}{6+7e^{\beta}}$ [½]

$t = 5$ $\qquad \dfrac{e^{\beta}}{(5+7e^{\beta})^2}$ (using Breslow correction for ties) [½]

$t = 6$ $\qquad \dfrac{e^{\beta}}{1+6e^{\beta}}$ [½]

$t = 8$ $\qquad \dfrac{e^{\beta}}{1+5e^{\beta}}$ [½]

$t = 10$ $\qquad \dfrac{1}{1+2e^{\beta}}$ [½]

$t = 11$ $\qquad \dfrac{e^{\beta}}{2e^{\beta}} = \dfrac{1}{2}$ [½]

Multiplying these elements together gives

$$L = \frac{1}{7+7e^{\beta}} \cdot \frac{1}{6+7e^{\beta}} \cdot \frac{e^{\beta}}{(5+7e^{\beta})^2} \cdot \frac{e^{\beta}}{1+6e^{\beta}} \cdot \frac{e^{\beta}}{1+5e^{\beta}} \cdot \frac{1}{1+2e^{\beta}} \cdot \frac{1}{2}.$$ [1]

[Total for part (vi): max. 4]
[Total 19]

This question was generally well answered, with many candidates scoring 12 marks or more. Answers to parts (i) – (iii) were especially good: many candidates scored full credit for these parts. In part (ii), some candidates did not realise that we have no information after t = 16, so the ranges for the last interval of the survival function for ForLawn should be 11 ≤ t < 16, not just 11 ≤ t. In part (v) many candidates simply wrote down the general form of the Cox model, rather than determining the equations of the model for the two groups in this specific study. Answers to part (vi) fell into two groups: candidates who made little or no attempt, and candidates whose derivations were largely correct. Some of those who had the correct terms did not state to which time interval each term applied, and a small number of marks were lost for this.

END OF EXAMINERS' REPORT