INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

April 2017

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
July 2017

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 not penalised 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. The performance of candidates in this examination was below the average over the past five years, but somewhat better than that in April or September 2016. The examination paper was considered to be of slightly greater difficulty than the September 2016 paper, and a slightly lower Pass Mark was therefore used.

2. This examination paper included questions on parts of the syllabus that have not been examined for many sessions. A substantial number of candidates scored poorly on these questions. This suggests that candidates are relying too much on previous examination papers in their preparation, training themselves to pass the examination by making sure they can do certain commonly-asked questions. If questions are asked on parts of the syllabus which are not often examined, or questions are asked on commonly examined

parts of the syllabus in a slightly unusual form, these candidates struggle.

3. A number of candidates 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 exam was 56.

Solutions

Q1 We require ${}_5q_{100}\cdot{}_7p_{93} = (1 - {}_5p_{100})\cdot{}_7p_{93}$ +½

$${}_7p_{93} = \exp\left(-\int_0^2 0.10dt - \int_2^7 0.15dt\right)$$ +½

$${}_7p_{93} = \exp(-0.95) = 0.386741$$ +½

$${}_5p_{100} = \exp\left(-\int_0^5 0.20dt\right)$$ +½

$${}_5p_{100} = \exp(-1.00) = 0.367879$$ +½

So ${}_5q_{100}\cdot{}_7p_{93} = 0.632121 * 0.386741 = 0.244467$ +½

[Total 3]

Full credit could be obtained for the correct numerical answer and some indication of the method used. A common error was to use an incorrect age range when evaluating survival probabilities. Most candidates, however, scored well on this question.

Q2 (i) An increment of a process is the amount by which its value increases over a period of time, for example $X(u + t) - X(u)$ where $t > 0$. +1

[1]

(ii) EITHER

$$\log[h(30+u+t)]-\log[h(30+u)]=\log[B(1+\gamma)^{u+t}]-\log[B(1+\gamma)^{u}]$$ +1

$$=\log B+(u+t)\log(1+\gamma)-(\log B+u\log(1+\gamma)).$$ +½

$$=t\log(1+\gamma)$$ +½

The increment thus depends on t, the duration of the process but not on u, hence the process is stationary. +1

OR

$$\log h(30+u)=\log B+u\log(1+\gamma)$$ +1

$$\frac{d}{du}\log h(30+u)=\log(1+\gamma)$$ +1

which is constant and does not depend on u, so the process is stationary. +1

[3]

[Total 4]

Most candidates (though by no means all) could define an *increment*. Many candidates did not make a serious attempt at part (ii), which dealt with a topic that had not been examined for several sessions. The better-prepared candidates saw that the increment for each time unit was constant and hence the process was stationary. A few candidates failed to spot this and tried to demonstrate weak stationarity. Credit was given for this, although such candidates often encountered difficulty in calculating the covariance of the process.

Q3 (i) A Markov Chain is a process operating in discrete time with a discrete state space +1

EITHER

It obeys the Markov property: +½

that the future state of the process can be predicted from its present state alone, without any reference to its past history +½

OR

$P[X_t \in A \mid X_{s_1} = x_1, X_{s_2} = x_2, ..., X_{s_n} = x_n, X_s = x] = P[X_t \in A \mid X_s = x]$ +½

for all times $s_1 < s_2 < ... < s_n < s < t$, all states $x_1, x_2, ..., x_n, x$ in S and all subsets A of S. +½

[2]

(ii) In a time-homogeneous Markov chain the transition probabilities are time-independent. +½

In a time-inhomogeneous Markov chain the transition probabilities depend on the absolute values of time, not just the time difference. +½

Examples: time-homogeneous – no claims discount system in which the probability of a claim in each year is constant. +½

time-inhomogeneous – no claims discount system in which accident probabilities reflect changing traffic conditions from one year to the next. +½

[2]

[Total 4]

Most candidates could define a Markov Chain. There was less sure-footedness about describing the difference between a time-homogeneous and a time-inhomogeneous Markov chain. Many candidates referred to "transition rates" when they meant one-step transition probabilities. Others made vague statements about the "Chain depending [or not] on time", for which only limited credit was given. In part (ii) credit was given only for examples which could sensibly be analysed using a Markov Chain.

Q4 (i) The state space may be either discrete or continuous. +½

The time set may be either discrete or continuous. +½

Hence we have $2 \times 2 = 4$ possibilities:

State space	*Time set*
Discrete	Discrete
Discrete	Continuous
Continuous	Discrete
Continuous	Continuous

+1
[2]

(ii)

State space	*Time set*	*Examples*
Discrete	Discrete	Simple random walk Counting process Markov chain Markov jump chain
Discrete	Continuous	Poisson process Counting process Markov jump process Compound Poisson process
Continuous	Discrete	White noise General random walk Time series
Continuous	Continuous	Compound Poisson process White noise Brownian motion Itô process

+2
[2]
[Total 4]

This question was well answered, with many candidates scoring full marks. As indicated in the model solutions above, credit was given in part (ii) for examples which are not on the CT4 syllabus, but which were correctly classified (e.g. Itô process).

Q5 (i)

$$\begin{array}{c}0\\1\\2\\3\\4\end{array}\begin{pmatrix}0.7 & 0.3 & 0 & 0 & 0\\ 0.2 & 0.5 & 0.3 & 0 & 0\\ 0 & 0.2 & 0.5 & 0.3 & 0\\ 0 & 0 & 0.2 & 0.5 & 0.3\\ 0 & 0 & 0 & 0.2 & 0.8\end{pmatrix}.$$ +2

[2]

(ii) For the long run proportion of time that the rack is either full or empty, we need to find the stationary distribution.
OR
Let the long-run probability of there being *i* bicycles in the rack be π_i. +½

Then if *P* is the transition matrix we need to solve

$$(\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3 \quad \pi_4) = (\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3 \quad \pi_4)P.$$ +½

The equations are

$$\pi_0 = 0.7\pi_0 + 0.2\pi_1$$
$$\pi_1 = 0.3\pi_0 + 0.5\pi_1 + 0.2\pi_2$$
$$\pi_2 = 0.3\pi_1 + 0.5\pi_2 + 0.2\pi_3$$ +1½
$$\pi_3 = 0.3\pi_2 + 0.5\pi_3 + 0.2\pi_4$$
$$\pi_4 = 0.3\pi_3 + 0.8\pi_4$$

Starting with the equation for π_0 we have

$$0.3\pi_0 = 0.2\pi_1$$
$$\pi_1 = 1.5\pi_0$$

Then in the equation for π_1 we have

$$1.5\pi_0 = 0.3\pi_0 + 0.5(1.5\pi_0) + 0.2\pi_2$$
$$0.45\pi_0 = 0.2\pi_2$$
$$\pi_2 = 2.25\pi_0$$

and in the equation for π_2 we have

$$2.25\pi_0 = 0.3(1.5\pi_0) + 0.5(2.25\pi_0) + 0.2\pi_3$$
$$0.675\pi_0 = 0.2\pi_3$$
$$\pi_3 = 3.375\pi_0$$

and, finally, in the equation for π_4 we have

$$\pi_4 = 0.3(3.375\pi_0) + 0.8\pi_4$$
$$0.2\pi_4 = 1.0125\pi_0$$
$$\pi_4 = 5.0625\pi_0$$

+1½

Since

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1,$$ +½

we have

EITHER

$$\pi_0 + 1.5\pi_0 + 2.25\pi_0 + 3.375\pi_0 + 5.0625\pi_0 = 13.1875\pi_0 = 1,$$

OR

$$\pi_0\left(\frac{16}{24} + \frac{24}{24} + \frac{36}{24} + \frac{54}{24} + \frac{81}{24}\right) = 1$$

and hence

$$\pi_0 = \frac{16}{211} = 0.0758 \quad \pi_4 = \frac{81}{211} = 0.3839$$ +½

Thus the rack is full or empty 45.97 per cent (97/211) of the time, which is more than 35 per cent, +½

and the city will increase the size of the rack. +½

[6]

(iii) The increase in the size of the rack is likely to reduce the proportion of time for which the rack is empty or full, as the fact that there are more states will "dilute" the probabilities. +1

This assumes that behaviour of the users of the scheme remains the same. +½

However, if the behaviour/characteristics of those using the bicycle scheme change in response to the increased size of the rack (for example new people join the scheme) +½

then the transition matrix may change so that the proportion of time the rack is empty or full may shift. +½

[Max 2]
[Total 10]

Many candidates scored highly on parts (i) and (ii) of this question. Candidates who produced an incorrect matrix in part (i) could potentially score full credit in part (ii) if they followed through correctly. In part (iii) a substantial number of candidates spotted that, with more states, the dilution of the probability should mean that the proportion of time the rack is empty or full will decrease. Some, however, thought that that chance of it being full would be less, but that the chance of it being empty would remain the same. Only a few candidates went further to comment on the possible effect on the increase in the size of the rack on the behaviour of those using the bicycle scheme.

Q6 (i) A Poisson process is a counting process in continuous time $\{N_t, t \geq 0\}$, where N_t records the number of occurrences of a type of event within the time interval from 0 to *t*. +1

Events occur singly and may occur at any time; +½

the probability that an event occurs during the short time interval from time *t* to time *t* + *h* is approximately equal to λ*h* for small *h*, where the parameter λ is the rate of the Poisson process. +½

OR

A Poisson process is an integer valued process in continuous time $\{N_t, t \geq 0\}$, where +½

$N_0 = 0$ +½

$\Pr[N_{t+h} - N_t = 1 \mid F_t] = \lambda h + o(h)$

$\Pr[N_{t+h} - N_t = 0 \mid F_t] = 1 - \lambda h + o(h)$

$\Pr[N_{t+h} - N_t \neq 0,1 \mid F_t] = o(h)$ +1

OR

A Poisson process with rate λ is a continuous-time integer-valued process $N_t, t \geq 0$, with the following properties: +½

$N_0 = 0$ +½

N_t has independent increments +½

N_t has Poisson distributed stationary increments

$$P\left[N_t - N_s = n\right] = \frac{\left[\lambda\left(t-s\right)\right]^n e^{-\lambda(t-s)}}{n!}, \qquad s < t,\ n = 0,1,...$$ +½

[2]

(ii) Because it is total claims divided by number of weeks (to the nearest whole week):

$$3.846 = \frac{200}{52},$$ +½

and this can be shown by formal procedures (e.g. maximum likelihood) to be a good estimate. +½

[1]

(iii) The null hypothesis is that the number of claims per week follows a Poisson distribution with parameter 3.846. +½

Using the Poisson formula, the probability of getting exactly *d* claims in a week is given by

$$\Pr[D = d] = \frac{e^{-(200/52)}(200/52)^d}{d!}$$

This produces an expected distribution of claims per week as follows:

Claims	*Probability*	*Expected number of weeks*
0	0.02136	1.11
1	0.08216	4.27
2	0.15800	8.22
3	0.20257	10.53
4	0.19478	10.13
5	0.14983	7.79
6	0.09604	4.99
7	0.05278	2.74
8	0.02537	1.32
9 or more	0.01711	0.89

+2

Combining the categories 0 and 1, and the categories 6–9 so that each of the expected values exceeds 5, we can test the fit using a chi-squared test as follows: +½

Claims	*Actual number of weeks, A*	*Expected number of weeks, E*	*$(A - E)^2/E$*
0–1	6	5.38	0.071
2	8	8.22	0.006
3	10	10.53	0.027
4	12	10.13	0.345
5	4	7.79	1.844
6–9	12	9.94	0.427

+1

The chi-squared statistic is 2.72. +½

The number of degrees of freedom is the number of categories less one because of the constraint imposed by the number of weeks in the year. +½

So there are 5 degrees of freedom.. +½

The critical value at the 5% level is 11.07. +½

Since 2.72 < 11.07 +½

we do not reject the null hypothesis, and conclude that the Poisson distribution with parameter 3.846 fits the data well. +½

ACCEPTABLE ALTERNATIVE GROUPING OF CLAIMS

Claims	*Actual number of weeks, A*	*Expected number of weeks, E*	*(A – E)^2/E*
0–1	6	5.38	0.071
2	8	8.22	0.006
3	10	10.53	0.027
4	12	10.13	0.345
5	4	7.79	1.844
6	6	4.99	0.203
7–9	6	4.95	0.223

+1

The chi-squared statistic is 2.72. +½

The number of degrees of freedom is the number of categories less one because of the constraint imposed by the number of weeks in the year. +½

So there are 6 degrees of freedom. +½

The critical value at the 5% level is 12.59. +½

Since 2.72 < 12.59 +½

we do not reject the null hypothesis, and conclude that the Poisson distribution with parameter 3.846 fits the data well. +½

[7]

(iv) In order to test for independence of the number of claims in successive weeks. +1

[1]

[Total 11]

Parts (i) and (ii) of this question were well answered. In part (ii) many candidates correctly calculated the expected number of weeks having 0, 1, 2, … claims using a Poisson distribution with parameter 3.846. Fewer knew how to test the hypothesis that this Poisson distribution fit the data. Of those who did a chi-squared test, only a minority grouped the claims categories so that the expected number of claims in each category was 5 or more. A significant minority of candidates calculated a table based on the actual and expected number of claims occurring in weeks with 0, 1, 2, … claims, and tried to do a chi-squared test on this. This test cannot include the weeks in which no claims take place (when both actual and expected numbers are zero whatever the value of the parameter), so limited credit was given for this approach. A common error was to state that the number of degrees of freedom was the number of groups minus 1 because of the estimation of the Poisson parameter. This is incorrect; it is the number of groups minus one because of the constraint that there are 52 weeks in the year. In part (iv) only a small number of candidates correctly identified the reason for performing a serial correlations test. Many candidates made vague reference to "clumps" or "runs", suggesting that they were thinking in a general way about statistical tests of a graduation, rather than trying to understand the scenario in the question.

Q7 (i) EITHER

In a proportional hazards model the hazard factorises +½

into a component which depends only on duration and a component which depends only on the covariates. +½

OR

In a proportional hazards model the hazard, $h(t)$, may be represented as $h(t) = h_0(t)g(z)$, +½

where $h_0(t)$ depends only on duration and $g(z)$ depends only on the covariates +½

As a consequence, the effect of a covariate is to shift the hazard up or down by a the same proportion at all durations OR the ratio between the hazards for two individuals, A and B, with different covariate vectors, $h_A(t)/h_B(t)$, is constant/does not depend on duration t. +1

[2]

(ii) μ is the hazard for a man who does not drink beer. +1

β measures the impact on the hazard of a one glass increase in the daily amount of beer drunk +1

[2]

(iii) The hazard for a man who drinks two glasses of beer a day is

$0.03\exp(0.2*2) = 0.0448.$ +1

[1]

(iv) (a) The probability that a man aged 60 years who drinks three glasses of beer a day will survive to his 70th birthday is

$$\exp\left(-\int_0^{10} 0.03\exp(0.2*3)dt\right) = \exp(-0.547) = 0.579\,.$$ +1

(b) Because the hazard of death is constant, the expectation of life at age 60 years is given by

$$\frac{1}{h(t)} = \frac{1}{\mu \exp(\beta x)} = \frac{1}{0.03 \exp(0.2*3)} = 18.3 \text{ years.}$$ +1

[2]

(v) The owner's total revenue R is proportional to the average number of glasses of beer drunk per day multiplied by the man's expectation of life:

$$R \propto \frac{x}{0.03 \exp(0.2x)}.$$ +1

THEN EITHER

We maximise R with respect to x.

$$\frac{dR}{dx} = \frac{0.03 \exp(0.2x) - x(0.03*0.2)\exp(0.2x)}{[0.03 \exp(0.2x)]^2} = \frac{1-0.2x}{0.03 \exp(0.2x)}.$$ +½

This is zero when $1 – 0.2x = 0$, or when $x = 5$. +½

$$\frac{d^2R}{dx^2} = \frac{-0.2[0.03\exp(0.2x)] - (1-0.2x)[0.03*0.2\exp(0.2x)]}{[0.03\exp(0.2x)]^2}$$

+½

$$= \frac{-0.2-0.2(1-0.2x)}{[0.03\exp(0.2x)]} = \frac{0.2(0.2x-2)}{[0.03\exp(0.2x)]}$$

which is negative when $x = 5$, +½

so we have a maximum, +½

and the owner should sell the man five glasses of beer per day. +½

OR

Taking the logarithm of R we have

$$\log R = \log x - \log 0.03 - 0.2x$$

$$\frac{d \log R}{dx} = \frac{1}{x} - 0.2 .$$ +½

This is zero when $x = 5$. +½

Since $\dfrac{d^2 \log R}{dx^2} = -\dfrac{1}{x^2}$ +½

which is always negative, +½

we have a maximum. +½

So the owner should sell the man five glasses of beer per day. +½

[4]

[Total 11]

> The better-prepared candidates answered part (i) well. Part (ii) was answered poorly by many candidates. The question asked for the interpretation of the parameters μ and β "in the context of this model", but most candidates simply gave general interpretation which were not given credit. Few candidates made serious attempts at part (v). Most seemed to have little idea of how to express the total amount of beer the man buys over his remaining lifetime as a function of the number of glasses of beer to be sold to him per day. Credit was given for an alternative approach in which candidates computed the expected revenue for 0, 1, 2, … glasses per day and showed that this reached a maximum at 5 glasses.

Q8 (i) Right censoring +½

Of those packets of cheese which are sold or discarded after 10 days. We do not know when these packets would have gone mouldy, but we know that it was after they ceased to be observed. +½

Random censoring +½

The time at which a packet of cheese is sold may be considered as a random variable. +½

Type 1 censoring +½

For those packets thrown out at day 10, the time of censoring is known in advance. +½

Interval censoring +½

Because data were only collected once per day: we only know that events happened between 8 a.m. one day and 8 a.m. the next day; we do not know exactly when within this period they happened. +½

EITHER

Informative censoring +½

If packets that were sold were those which looked the freshest then those packets censored for this reason might be less likely to go mouldy at each duration in excess of the censoring time than those which remain on the shelves. +½

OR

Non-informative censoring +½

If shoppers select packets of cheese at random there is no reason to believe that those packets bought are any more or less likely to go mouldy than those remaining on the shelves. +½

[Max 3]

(ii) Right censoring

Hard to change this, as would need to prevent customers buying cheese. +1

Random censoring

Hard to change as would need to prevent customers buying cheese. +1

Type 1 censoring

Could be removed by leaving all the cheese in the shelves until the last packet went mouldy or was sold. +1

OR

Could be removed by throwing the remaining cheese away once a certain number had already gone mouldy, in which case we would have introduced Type II censoring. +1

Interval censoring

Could be reduced by more frequent checks. Hard to remove, other than by continuous monitoring of the cheese and the removal of cheeses at the first sign of mould. Video surveillance might be a solution. +1

Informative censoring

Could be removed by giving customers no choice about which packet of cheese they bought. However, customers might object to being treated this way. +1

Non-informative censoring

This feature is desirable so we would not want to remove it. +1

[Max 3]

(iii)

t_j	n_j	d_j	c_j	d_j/n_j	$1 - d_j/n_j$
0	20		3		
2	17		1		
2	16	1		1/16	15/16
3	15	2		2/15	13/15
4	13		4		
6	9	2		2/9	7/9
9	7		2		
10	5	3	2	3/5	2/5
+½	*+1*	*+½*	*+½*	*+½*	*+½*

+3½

The Kaplan-Meier estimate is $S(t) = \prod_{t_j \le t} \left(1 - \frac{d_j}{n_j}\right)$. +½

t	*Kaplan-Meier estimate of S(t)*
$0 \le t < 3$	1
$3 \le t < 4$	15/16=0.9375
$4 \le t < 8$	13/16=0.8125
$8 \le t < 10$	91/144=0.6319
$t = 10$	91/360=0.2528
+1	*+1*

+2

[6]

[Total 12]

The most common error with this question was to assume the wrong decrement (i.e. that the decrement was cheese being sold, whereas the question stated that interest is in "the hazard of cheese going mouldy"). Candidates who assumed the wrong decrement scored little for part (i) but could gain credit for part (ii) if their suggestions were sensible given what they had written in part (i), and for part (iii) if they applied the correct method.

Q9 (i) The matrix is

$$\begin{array}{c} A \\ B \\ C \end{array}\begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.6 & 0.15 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$ +1

[1]

(ii) We need second order transition probabilities. +½

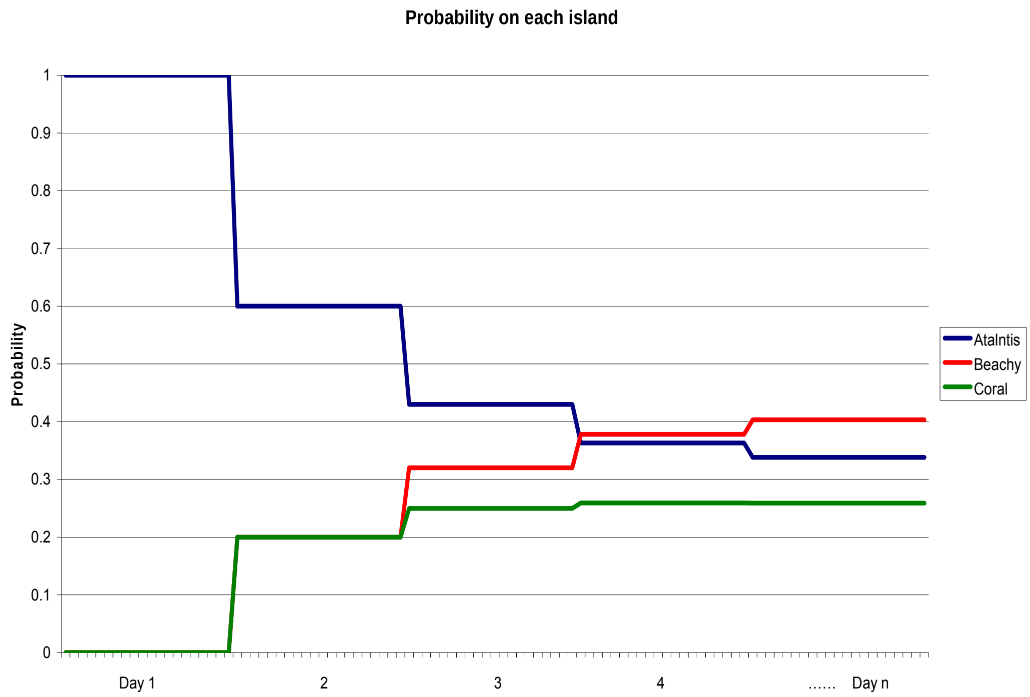
The second order transition matrix is:

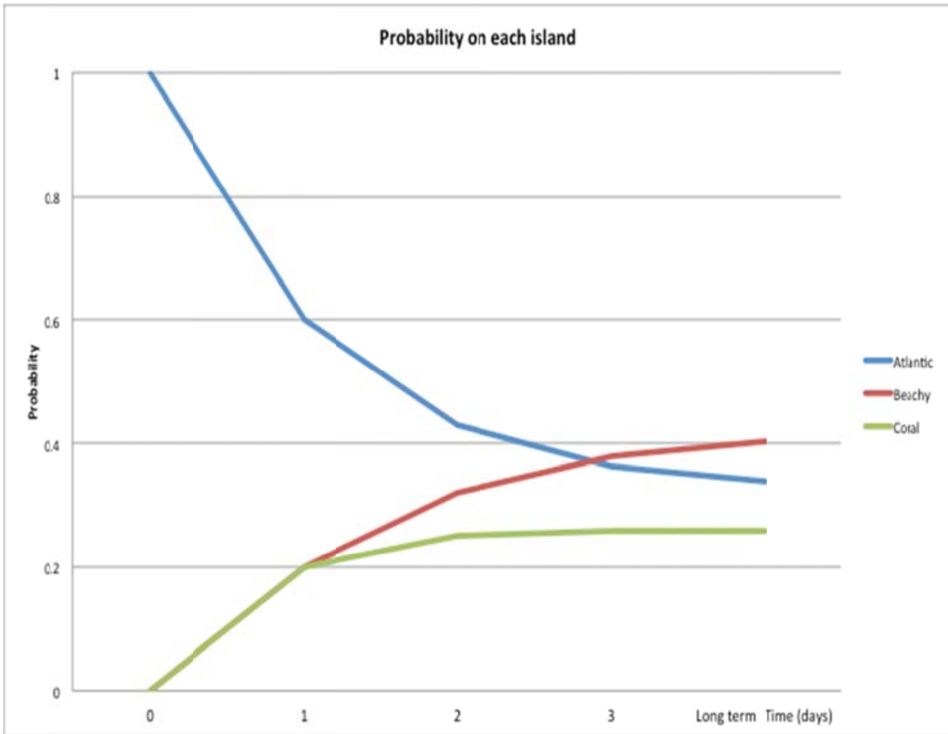
$$\begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.6 & 0.15 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} * \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.6 & 0.15 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.43 & 0.32 & 0.25 \\ 0.315 & 0.47 & 0.215 \\ 0.21 & 0.46 & 0.33 \end{pmatrix}$$ +1

So the probabilities are 0.43 in Atlantis, 0.32 in Beachy and 0.25 in Coral. +½

[2]

(iii) *Either of the following two graphs was acceptable.*





+3
[3]

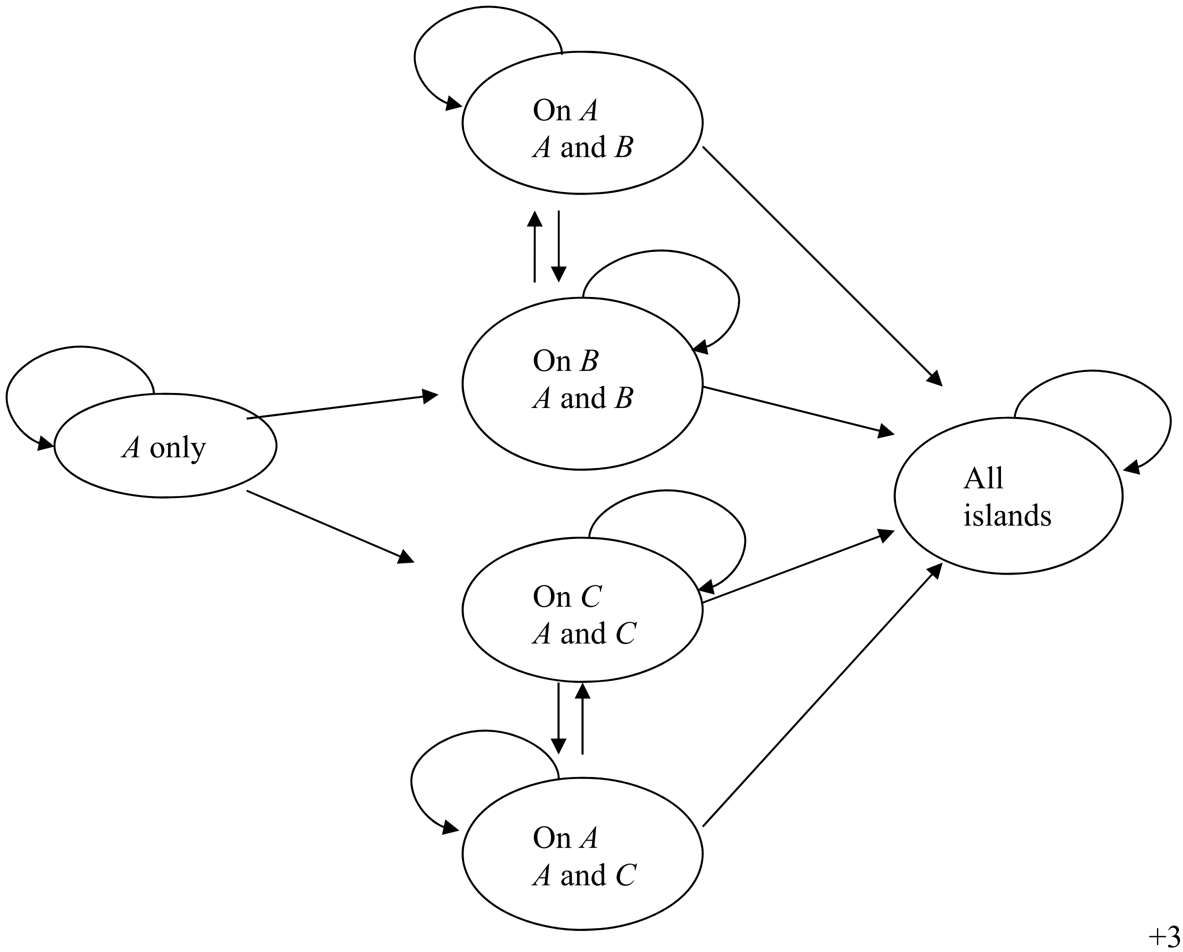
(iv) This is the probability he is in each state multiplied by the chance of taking a flight if he is in that state. +½

$(0.326 \times 0.4) + (0.419 \times 0.4) + (0.256 \times 0.5) = 0.426.$ +½
[1]

(v) This needs six states: +½

{*A* visited only, On *B* (*A* and *B* visited), On *A* (*A* and *B* visited), On *C* (*A* and *C* visited), On *A* (*A* and *C* visited), All islands visited} +1½
[2]

(vi)



+3

[3]

[Total 12]

This was a demanding question, and many candidates only attempted parts (i) and (ii). In part (iii) a graph with smooth lines was acceptable. Many candidates simply computed the long-run probabilities of being in each country and plotted these on a bar chart. This did not answer the question, which was about the "probabilities … over time". Of those who did try parts (v) and (vi), many posited incorrect state spaces, the most common which considered where the foreign correspondent was and had states {A for the first time, A for a second or subsequent time, B for the first time, B for a second or subsequent time, C for the first time, C for a second or subsequent time}. This is incorrect because the purpose of the exercise is to model "which countries he has visited so far" and the state B for the first time, say, does not distinguish individuals who have visited country A only, country C only, and both countries A and C. Partial credit was given in parts (v) and (vi) for various alternative state spaces, notably {A, AB, AC, ABC} where the letters refer to the countries the foreign correspondent has already visited. This is not Markov, since the probability of moving from state AB to state ABC depends on whether he is in country A or in country B.

Q10 (i) The crude rates are estimated independently at each age. +½

Therefore they are subject to sampling error. +½

This is a fairly small country, so there may be scanty data at certain ages. +½

This means that the crude rates may exhibit “roughness” in the progression from age to age. +½

We believe that underlying mortality actually progresses smoothly from age to age. +½

Graduation “irons out” the roughness in the crude rates which is due to sampling error while preserving the underlying level and shape of the mortality curve. +½

It uses information from adjacent ages to adjust the rate at any particular age. +½

We believe that the graduated rates are close to the true mortality rates underlying the crude rates. +½

The government of the country may wish to use the new table for financial calculations and economic planning (e.g. for a state pension scheme), so it is important that sampling errors are removed. +½

[Max 3]

(ii) Graduation using a parametric formula. +½

As the results are being used to create a standard table. +½

As the data reflect adult ages, a formula from the Gompertz or Makeham family would seem appropriate. +½

The parameters can be estimated by maximum likelihood or least squares, using weights which are proportional to the amount of data at each age. +½

[Max 1]

(iii) Might not find a suitable formula with a small number of parameters to fit the data at all ages. +1

At extreme ages data might be scanty and so less weight should be given to those ages when estimating the parameters of the formula. +1

[2]

(iv) Under the null hypothesis that the graduated rates are equal to the mortality underlying the crude rates, +½

the individual standardised deviations should be distributed $N(0,1)$ +½

There are 80 deviations. +½

The number of positive deviations P is distributed Binomial (80, 0.5). +½

We have a large number of deviations so
we can use the Normal approximation +½

$P \sim$ Normal(40, 20)

We have 31 positive deviations. +½

EITHER WITHOUT CONTINUITY CORRECTION

We compute the z-score which is $\frac{31-40}{\sqrt{20}} = -2.01$. +½

Since –2.01 < –1.96 +½

we have sufficient evidence to reject the null hypothesis at the
5% level (two-tailed test). +½

OR WITH CONTINUITY CORRECTION

We compute the z-score which is $\frac{31.5-40}{\sqrt{20}} = -1.9007$. +½

Since –1.9007 > –1.96 +½

we have insufficient evidence to reject the null hypothesis at the 5% level
(two-tailed test). +½

EITHER

We expect 1 in 20 of the standardised deviations
to be greater than 2 in absolute magnitude. Here we have
12 out of 80 which is 3 in 20, which is more than we should expect. +1

OR

We also expect fewer than 1 in 100 of the standardised
deviations to be greater than 3 in absolute magnitude, and here
we have 3 out of 80, which is again a larger proportion than we
should expect. +1

We can compute a chi-squared test. The calculations are shown in the table below.

Range	*Actual number of z_xs*	*Expected number of z_xs*	*$(A - E)^2/E$*
$z_x < -1$	23	12.8	8.2
$0 > z_x \geq -1$	26	27.2	0.1
$1 > z_x \geq 0$	16	27.2	4.6
$z_x \geq 1$	15	12.8	0.4

+1

The chi-squared statistic is 13.3. +½

There are 3 degrees of freedom (the number of groups minus 1) +½

because the total number of ages is fixed, so once the z_xs have been allocated to three groups, the number to go in the final group is determined. +½

The critical value at the 5% level is 7.82. +½

Since 13.3 > 7.82 we reject the null hypothesis. +½

[Max 6]

(v) The tests seem to indicate that the graduated rates are statistically different from the "true" mortality rates underlying the crude rates. +½

The government would be unwise to use these graduated rates for financial calculations or economic planning. +½

They will tend to overestimate mortality and therefore underestimate the amount of money required to fund future pensions. +½

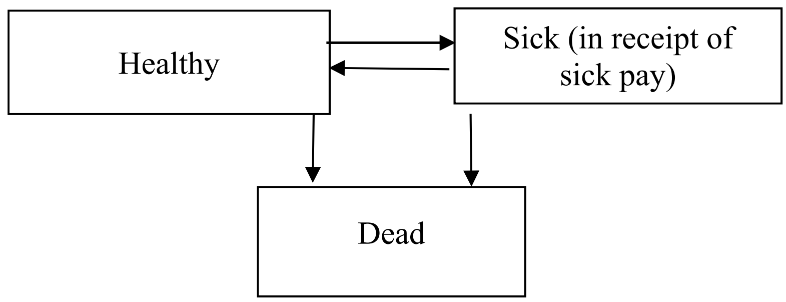
This could lead to fiscal problems for future governments. +½

[2]

[Total 14]

Answers to this question were often relatively weak, especially part (iv). The data were provided to allow the chi-squared test described on Unit 11, page 10 of the Core Reading, but only a small minority of candidates attempted such a test. Most were content with a non-rigorous check for outliers. Several candidates, however, made sensible comments in part (v). Credit was given in part (v) for any thoughtful comments which were consistent with candidates' answer to part (iv). So, if a candidate used the continuity correction in part (iv) and did not reject the null hypothesis, credit was given in part (v) for stating that the graduated rates were not biased up or down.

Q11 (i)



+2
[2]

(ii) Let the number of transitions observed from state i to state j be d^{ij}. +½

Let the total waiting time in state i be W^i. +½

Let the intensity of the transition from state i to state j be μ^{ij}. +½

Then the likelihood, L, is

$$L \propto \exp[-W^H(\mu^{HS}+\mu^{HD})]\exp[-W^S(\mu^{SH}+\mu^{SD})](\mu^{HS})^{d^{HS}}(\mu^{HD})^{d^{HD}}(\mu^{SH})^{d^{SH}}(\mu^{SD})^{d^{SD}}$$

+2½
[4]

(iii) Taking logarithms of the likelihood we have

$$\log_e L = -W^H(\mu^{HS}) + d^{HS}\log_e \mu^{HS} + \text{terms not involving } \mu^{HS}$$ +½

To maximize this with respect to μ^{HS} we proceed as follows:

$$\frac{d\log_e L}{d\mu^{HS}} = -W^H + \frac{d^{HS}}{\mu^{HS}} \ .$$ +½

Setting this to zero produces the estimate +½

$$\hat{\mu}^{HS} = \frac{d^{HS}}{W^H}$$ +½

and since

$$\frac{d^2\log_e L}{(d\mu^{HS})^2} = -\frac{d^{HS}}{(\mu^{HS})^2}$$ +½

which is negative, we have a maximum. +½
[3]

(iv) Let the number of healthy members aged x last birthday on 1 January in year t be $H_{x,t}$, and the number of sick members be $S_{x,t}$.

We need to adjust the exposed-to-risk to age nearest birthday. +½

Assuming birthdays are evenly distributed across the calendar year, +½

those aged 52 nearest birthday consist of half of those aged 51 last birthday and half of those aged 52 last birthday. +½

Further, assuming the population varies linearly over each calendar year we can apply the trapezium rule. +½

The required central exposed to risk for sick members, $E^c_{52,S}$ is given by

$$E^c_{52,S} = \frac{1}{2}\left[\frac{1}{2}\left(S_{51,2014}+S_{52,2014}\right)+\frac{1}{2}\left(S_{51,2015}+S_{52,2015}\right)+\frac{1}{2}\left(S_{51,2015}+S_{52,2015}\right)+\frac{1}{2}\left(S_{51,2016}+S_{52,2016}\right)\right]$$

+1

which is

$$E^c_{52,S} = \frac{1}{2}\left[\frac{1}{2}(12+10)+\frac{1}{2}(20+18)+\frac{1}{2}(20+18)+\frac{1}{2}(8+7)\right] = 28.25\,.$$ +½

Similarly, we have

$$E^c_{52,H} = \frac{1}{2}\left[\frac{1}{2}\left(H_{51,2014}+H_{52,2014}\right)+\frac{1}{2}\left(H_{51,2015}+H_{52,2015}\right)+\frac{1}{2}\left(H_{51,2015}+H_{52,2015}\right)+\frac{1}{2}\left(H_{51,2016}+H_{52,2016}\right)\right]$$

+1

and, since $H_{x,t}$ = total members – $S_{x,t}$, we have +1

$$E^c_{52,H} = \frac{1}{2}\left[\frac{1}{2}(136+136)+\frac{1}{2}(142+130)+\frac{1}{2}(142+130)+\frac{1}{2}(172+153)\right] = 285.25$$ +½

[6]
[Total 15]

This question was well answered by most candidates. A very common error in part (iii) was to forget that the exposed to risk for falling sick should not include those persons who are already sick. The total exposed to risk was 313.5, of whom 285.25 were Healthy (and hence at risk of falling sick) and 28.25 were Sick (and hence at risk of recovery).

END OF EXAMINERS' REPORT