INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

September 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.

Mike Hammer
Chair of the Board of Examiners
December 2018

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. Candidates' performance was weakest on questions that tested higher order skills or were demanding applications, and much better on knowledge based questions and straightforward application questions. The one exception was a knowledge based question which was on a part of the syllabus that had not been tested recently.

2. The Pass Mark was adjusted to 58 on the basis that, as marking proceeded, it became clear that the Marking Schedule included marks for two small part questions that almost no candidates were scoring.

3. There was 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 58

Solutions

Q1

(a) Not a valid transition matrix, [+½]
because the first row sums to 0 not 1. [+½]

(b) Not a valid transition matrix, [+½]
because it is not square. [+½]

(c) This is a valid transition matrix. [+½]

Each row sums to one, and each entry is between 0 and 1 inclusive, and the matrix is square. [+½]

(d) Not a valid transition matrix, [+½]
because there is an entry which is less than 0. [+½]

[Total 4]

Most candidates answered this question well. The most common error was not to recognise that transition matrices for a Markov chain must be square.

Q2

The starting point is the number of age groups used. [+1]

If the age groups have been chosen with reference to the data, an unknown number of degrees of freedom should be deducted. [+½]

Then you deduct a number of degrees of freedom depending upon the method of graduation used. [+½]

If a standard table is used, deduct, say, 2 degrees of freedom for the choice of standard table (though the exact number to deduct is not determined easily). [+1]

If a link function is used to a standard table or a parametric formula is used, deduct one degree of freedom per parameter estimated. [+1]

If graphical graduation is used, deduct 2 or 3 degrees of freedom for every 10 or so ages. [+1]

[Total max. 4]

This question was well answered by most candidates.

Q3

Process	*State space*	*Time set*	
Simple random walk	Discrete	Discrete	[+1]
Markov jump process	Discrete	Continuous	[+1]
Compound Poisson process	Either	Continuous	[+1]
Markov chain	Discrete	Discrete	[+1]
Counting process	Discrete	Either	[+1]

[Total 5]

This question was well answered. Common errors were not to realise that a Compound Poisson process can have either a continuous or a discrete state space, and that a Counting process can have either a continuous or a discrete time set.

Q4

(i) Internal data for an old model performing the same/similar function. [+1]

Internal data for a model performing a different function [+1]

Market observable yields or rates. [+1]

Expert opinion. [+1]

Industry data, for example a standard table or surveys. [+1]

Regulations set out by regulatory authorities. [+1]

Government statistical data. [+1]

[max. 3]

(ii) **Internal data for an old model performing the same/similar function**

Past experience may not be representative of future experience. [+1]

Internal data on a model performing a different function

Subjective adjustments may be needed. [+1]

Market observable yields or rates

May be a time delay before they become available,
OR
different sources may give slightly different rates for the same item. [+1]

Expert opinion

May be hard or expensive to find a relevant expert,
OR
the expert's advice may be theoretical and hard to adapt
into a pragmatic model. [+1]

Industry data, for example a standard table or surveys

May not be directly relevant to the situation to be modelled,
OR
survey may be expensive
OR
the experience likely to differ by firm due to distribution
approaches/target markets, etc.,
OR
Firms for whom the assumption is insignificant may take a high-level
approach and this may not be readily apparent. [+1]

Regulations set out by regulatory authorities.

Current regulations may change in the future. [+1]

Government statistical data

Will tend to apply to the population as a whole, and the
model may apply to a non-representative subset of the population. [+1]

[max. 3]

[Total 6]

> *This higher skills question proved the most demanding of any question on the examination paper. Many candidates did not attempt it or made only token attempts. Credit was given for sensible answers other than those listed above, but vague answers, such as "past experience" did not receive full credit. Some candidates offered alternative data sources such as "reinsurers' data" or "data from other countries" and these were given credit. The answers to part (ii) were expected to relate to the sources listed in part (i).*

Q5

(i) Assume that the hazard of death (or force of mortality) is constant between ages x and x+1 and takes the unknown value μ.

Probability of observing all the data we actually observe – both censored lives and those which died is the likelihood L, which is

$$L = \prod_{\text{all censored lives}} S(t_i) \prod_{\text{all lives which died}} f(t_i)\,,$$ [+½]

where t_i is the duration for which life i is observed, and $S(t_i)$ and $f(t_i)$ are the survival and probability density functions of the chosen survival distribution. [+½]

To obtain $\hat{\mu}$, define a variable δ_i such that

$\delta_i = 1$ if life i died

$\delta_i = 0$ if life i was censored. [+½]

Then the likelihood can be written

$$L = \prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i} = \prod_{i=1}^{n} \mu^{\delta_i} \exp(-\mu t_i)\,.$$ [+½]

Thus $\log L = \sum_{i=1}^{n} \delta_i \log\mu - \sum_{i=1}^{n} \mu t_i$ [+½]

We differentiate this with respect to μ to give

$$\frac{\partial \log L}{\partial \mu} = \frac{\sum_{i=1}^{n} \delta_i}{\mu} - \sum_{i=1}^{n} t_i\,.$$ [+½]

Setting this equal to zero produces

$$\frac{\sum_{i=1}^{n}\delta_i}{\mu} = \sum_{i=1}^{n} t_i \,, \qquad [+½]$$

so that

$$\hat{\mu} = \frac{\sum_{i=1}^{n}\delta_i}{\sum_{i=1}^{n} t_i}. \qquad [+½]$$

We can check that this is a maximum by noting that $\dfrac{\partial^2 \log L}{\partial \mu^2} = -\dfrac{\sum_{i=1}^{n}\delta_i}{\mu^2}$, [+½]

which is negative. [+½]

[max. 4]

(ii) Obtain a series of separate estimates for the different hazards in each year of age for the calendar year 2017. [+1]

Suppose that the maximum likelihood estimate of the constant force during the single year of age from x to x+1 is μ_x . [+1]

Then the probability that a person alive at exact age x will still be alive at exact age x+1 is just $S_x(1)$. Given the constant force, then

$$S_x(1) = \exp(-\hat{\mu}_x)\,. \qquad [+1]$$

In general, therefore

$$S_x(m) = {}_m p_x = \exp\left(-\sum_{j=0}^{m-1}\hat{\mu}_{x+j}\right). \qquad [+½]$$

By 'chaining' together the probabilities in this way, we can create a life table from our estimates and evaluate probabilities over any relevant age range. [+½]

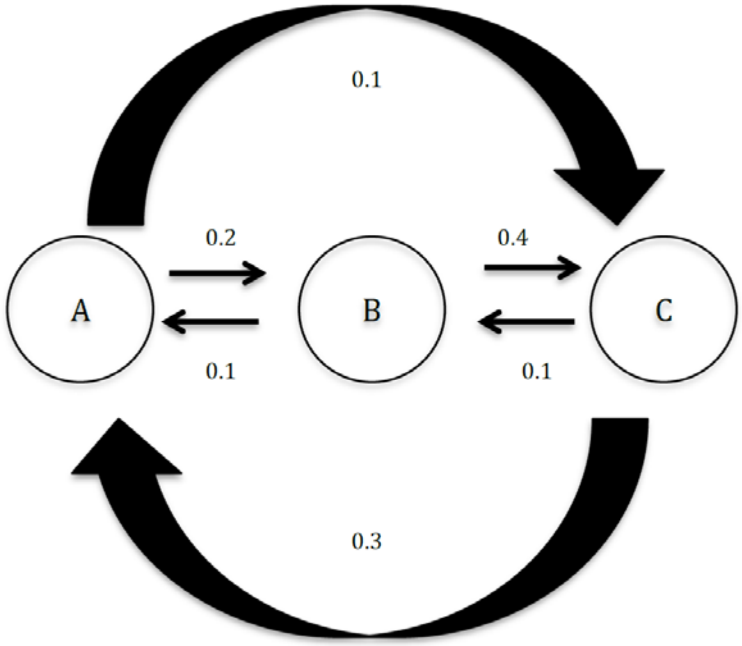
[max. 3]

[Total 7]

Part (i) of this question was from the Core Reading, Unit 6, pages 8-11. A large number of candidates based their answers on the Poisson likelihood. This is not quite correct as, strictly speaking, it requires observation of all lives for the same fixed period, and the question states clearly that this is not the case. Nevertheless, candidates could score credit for correctly deriving the correct maximum likelihood estimator from the Poisson likelihood. On the other hand, Binomial likelihoods were given little credit, as they do not furnish a maximum likelihood estimator of the hazard of death (they allow one to construct an estimator of q_x). The answer to part (ii) given above follows that in the Core Reading, Unit 6, pages 11-12. Some candidates framed their answers to part (ii) around the need for graduation and smoothing. Some credit was given for such answers.

Q6

(i)



[+2]
[2]

(ii) $$\frac{d}{dt}P_{AA}(t) = -0.3P_{AA}(t) + 0.1P_{AB}(t) + 0.3P_{AC}(t)$$

$$\frac{d}{dt}P_{AB}(t) = 0.2P_{AA}(t) - 0.5P_{AB}(t) + 0.1P_{AC}(t)$$

$$\frac{d}{dt}P_{AC}(t) = 0.1P_{AA}(t) + 0.4P_{AB}(t) - 0.4P_{AC}(t)$$

[+2]
[2]

(iii) EITHER

To stay in state *A* the equation reduces to:

$$\frac{d}{dt}P_{\overline{AA}}(t) = -0.3P_{\overline{AA}}(t)$$ [+½]

which has solution

$P_{\overline{AA}}(t) = \exp(-0.3t)$ [+½]

So for t = 2 we have exp(-0.6) = 0.5488. [+1]

OR

We can model this as Poisson with parameter (0.1 + 0.2)*2 = 0.6 [+1]

$P(Poi(0.6) = 0) = \dfrac{e^{-0.6}\,0.6^0}{0!}$ [+½]

$= e^{-0.6} = 0.5488$ [+½]

[2]

(iv) The only paths under which the third jump is into state *C* are *BAC, CAC* and *CBC*. [+1]

The probabilities of each jump are given by the ratio of the transition rates.

So the probabilities for each path are:

BAC = $\frac{2}{3}.\frac{1}{5}.\frac{1}{3} = \frac{2}{45}$ [+½]

CAC = $\frac{1}{3}.\frac{3}{4}.\frac{1}{3} = \frac{1}{12}$ [+½]

CBC = $\frac{1}{3}.\frac{1}{4}.\frac{4}{5} = \frac{1}{15}$ [+½]

Sum = 7/36 = 0.194. [+½]

[3]

[Total 9]

Parts (i)-(iii) of this question were very well answered, with many candidates scoring full marks. Part (iv) was only answered correctly by a minority of candidates. An alternative solution involving writing down the transition matrix, P, of the process, and then pointing out that the correct probability would be found in cell {1,3} of the matrix P3 was awarded credit.

Q7

(i) **Random**

When a subject is removed from the investigation for a reason other than by death, and the timing of the removal can be considered a random variable. [+1]

Right

When a subject is removed from the investigation for a reason other than by death, so that
EITHER
we do not know exactly when death will occur, just that it occurs after the time of removal.
OR
censoring cuts short the observation in progress. [+1]

Informative

When the future mortality of a subject censored from the investigation is likely to be different from those remaining in the investigation. [+1]

[3]

(ii) **Random**

The three eaten by the goat on day 3 and the one stolen by the boy down the road on day 8, [+1]

as the times of these events could not have been known in advance. [+1]

Arguably, those remaining fresh on day 9, if we did not know at what remaining number it would be difficult to make them look good (otherwise it would be type II). [+1]

Right

Day 3, Day 8 and Day 9 as listed above, [+1]

as we do not know when they would have wilted if they had not been removed from observation, just that it would have been after the day they were removed. [+1]

Informative.

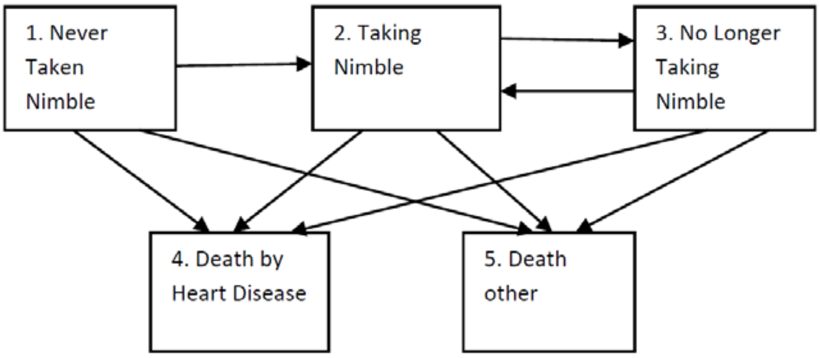
Arguably the one stolen on day 8 [+½]
as it is likely he took the freshest-looking one to present to his girlfriend [+1]
so it may have had a longer future fresh life than the others. [+½]

[max. 6]
[Total 9]

In part (i) many candidates did not include in their answers a definition of censoring. This omission was penalised. The understanding of many candidates about informative censoring is still rather shaky. Many answers were vague, such as "censoring gives information about the remaining lives" without specifying what that information might be. In part (ii) most candidates identified that the roses censored on Days 3 and 8 were examples of random censoring, and the roses left on Day 9 were examples of right censoring. Fewer candidates spotted that the roses censored on Days 3 and 8 were also right censoring. Fewer still gave persuasive examples of informative censoring.

Q8

(i)



[+2]
[2]

(ii) Using the Markov assumption

OR
the Chapman Kolmogorov equation is [+½]

$$_{dt+t}p_x^{34} = {_t}p_x^{31}\,{_{dt}}p_{x+t}^{14} + {_t}p_x^{32}\,{_{dt}}p_{x+t}^{24} + {_t}p_x^{33}\,{_{dt}}p_{x+t}^{34} + {_t}p_x^{34}\,{_{dt}}p_{x+t}^{44} + {_t}p_x^{35}\,{_{dt}}p_{x+t}^{54}.$$

Since $_{dt}p_{x+t}^{54} = {_t}p_x^{31} = 0$ [+½]

$$_{dt+t}p_x^{34} = {_t}p_x^{32}\,{_{dt}}p_{x+t}^{24} + {_t}p_x^{33}\,{_{dt}}p_{x+t}^{34} + {_t}p_x^{34}\,{_{dt}}p_{x+t}^{44}.$$ [+1]

Given that $_{dt}p_{x+t}^{44} = 1$ [+½]

And assuming that, for small dt

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

where $\lim_{dt\to 0}\frac{o(dt)}{dt} = 0$, [+½]

then substituting, we have

$$_{dt+t}p_x^{34} = {_t}p_x^{32}\mu_{x+t}^{24}dt + {_t}p_x^{33}\mu_{x+t}^{34}dt + {_t}p_x^{34} + o(dt)$$ [+½]

so that $_{dt+t}p_x^{34} - {_t}p_x^{34} = {_t}p_x^{32}\mu_{x+t}^{24}dt + {_t}p_x^{33}\mu_{x+t}^{34}dt + o(dt)$ [+½] and hence $\frac{d}{dt}({}_t p_x^{34}) = \lim_{dt \to 0} \frac{{}_{t+dt} p_x^{34} - {}_t p_x^{34}}{dt} = {}_t p_x^{32} \mu_{x+t}^{24} + {}_t p_x^{33} \mu_{x+t}^{34}$. [+1]

[max. 5]

(iii) EITHER INDIVIDUAL-LEVEL DATA

If data are held at an individual level we would need, for the period of the investigation [+½]

dates the individual moved into and out of the local health authority (if such movement took place) [+½]

date the individual attained exact age 50 or 51 years (or date of birth) [+½]

dates the individual started taking Nimble [+½]

dates the individual stopped taking Nimble [+½]

date the individual died [+½]

cause of death. [+½]

whether the individual had taken Nimble before (or date the individual first took Nimble). [+½]

OR AGGREGATE-LEVEL DATA

If data is held at an aggregate level we would need the amount of time spent within the investigation period [+½]

by lives aged between ages 50 and 51 years exact [+½]

for each of the states "Never taken Nimble", "Taking Nimble" and "No longer taking Nimble" . [+½]

The number of deaths [+½]

from each of the three states "Never taken Nimble", "Taking Nimble" and "No longer taking Nimble" [+½]

split by cause of death (heart disease or not) [+½]

[max. 3]

[Total 10]

In part (i), most candidates wrote down a transition graph which was either correct or close to being correct. Answers to part (ii) were better than answers to similar questions in previous sessions. In part (iii) most candidates provided answers assuming aggregate-level data. Candidates could assume either individual-level or aggregate-level data, but credit was not awarded from both alternatives given in the solutions. Candidates were expected to refer to the specific scenario in the question, so vague answers of the form 'number of transitions from i to j scored little credit.

Q9

(i) We believe that mortality varies smoothly with age
OR
evidence from large experiences suggests mortality varies smoothly with age . [+½]

Therefore the crude estimate of mortality at any age carries information about mortality at adjacent ages. [+½]

By smoothing the experience, we can make use of data at adjacent ages to improve the estimates at each age. [+½]

This reduces sampling (or random) errors. [+½]

The mortality experience may be used in financial calculations. [+½]

Irregularities, jumps and anomalies in financial quantities (such as premiums for life insurance contracts) are hard to justify to customers
OR
jumps and anomalies in financial quantities may be taken advantage of by customers. [+½]
[3]

(ii) **Parametric formula**

Rates are automatically smooth provided that a formula with sufficiently few parameters is used. [+1]

Graphical

Reliance is placed on the skill of the practitioner to draw a sufficiently smooth line through the crude rates. [+½]

The third differences test for smoothness is useful here. [+½]

It is usually necessary to make several attempts,
and to adjust the results by hand (rather than re-drawing the curve),
a process called hand-polishing. [+½]

With reference to a standard table

A standard table will already be smooth. [+½]

Provided a link function is selected with few parameters,
this smoothness should be preserved in the graduated rates. [+½]

[max. 3]

(iii) The null hypothesis is that the graduated rates are the same as the true underlying mortality rates for this block of business. [+½]

The test statistic $\sum_x z_x^2 \approx \chi_m^2$ where m is the degrees of freedom.

Age	*Exposed to risk*	*Observed deaths*	*Graduated rates*	*Expected deaths*	z_x	z_x^2
55	1550	15	0.00673	10.432	1.41449	2.00079
56	2100	18	0.00689	14.469	0.92828	0.86170
57	2300	15	0.00709	16.307	-0.32366	0.10476
58	2450	21	0.00736	18.032	0.69894	0.48852
59	2700	18	0.00770	20.790	-0.61190	0.37442
60	3250	29	0.00820	26.650	0.45522	0.20722
61	3100	25	0.00891	27.621	-0.49871	0.24871
62	3450	30	0.00978	33.741	-0.64403	0.41478
63	3600	45	0.01084	39.024	0.95663	0.91514
64	3750	41	0.01210	45.375	-0.64949	0.42183
				Total		6.0378

[+1½]

The observed test statistic is 6.0378 [+½]

The degrees of freedom are 10 minus an unknown number for the choice of standard table (say 2) and a further one for the parameter in the link function. [+½]

So m = 7 say (but could also use 6 or 8 degrees of freedom) [+½]

The critical value of the χ^2 distribution with 7 degrees of freedom at the 95% significance level is 14.07 (6 d.f.12.59, 8 d.f. 15.51) [+½]

Since 6.0378 < 14.07 (or 12.59, or 15.51 [+½]

We have insufficient evidence to reject the null hypothesis. [+½]

[5]

[Total 11]

This question was reasonably well answered by most candidates. The weakest section was part (ii). Few candidates realised that a parametric formula would automatically furnish smoothness if the number of parameters was small. Similarly, few mentioned that a standard table will already be smooth, so the requirement is to find a link function capable of transferring that smoothness to the graduated rates. In part (iii), many candidates did not apply the approach outlined in the solution to Q2 to the determination of the number of degrees of freedom.

Q10

(i) As the sick list is updated weekly, we can assume that events (falling sick, recovering, dying) take place mid-week, so that

duration of sickness = last week – first week + 1 [+1]

The exposed to risk is therefore calculated as shown below

Member number	*Duration*	*Outcome (needed for part (iii))*
1	1	Assumed recovered
2	3	Died
3	2	Died
4	63	Still alive
5	12	Died
6	8	Assumed recovered
7	3	Assumed recovered
8	51	Died
9	1	Died
10	54	Still alive
Total	198	

[+2]

There were 5 deaths, [+½]

so the death rate is 5/198 = 0.02525 per week. [+½]

[4]

(ii) $$p = \exp\left(-\int_0^{52} \mu dt\right) = e^{-0.02525(52)} = 0.26898.$$ [+1]

[1]

(iii) The Nelson-Aalen estimate calculations are shown in the table below. We adopt the convention that censoring happened immediately after death where censored observations and deaths have the same duration.

t_j	n_j	d_j	c_j	d_j/n_j	$\sum(d_j/n_j)$	$\exp[-\sum(d_j/n_j)]$
0	10					
1	10	1	1	1/10	0.1000	0.9048

2	8	1	0	1/8	0.2250	0.7985	
3	7	1	2	1/7	0.3679	0.6922	
12	4	1	0	1/4	0.6179	0.5391	
51	3	1	2	1/3	0.9512	0.3863	[+4]

Since we have information up to week 65, [+½]

the Nelson-Aalen estimate of $S(52)$ is 0.3863. [+½]

[5]

(iv) The Nelson-Aalen estimate of the one-year survival probability is higher than that obtained using the exact exposed to risk. [+½]

The exact exposed to risk approach constrains the death rate to be constant over the 52 weeks at the "average" rate implied by the number of deaths and exposed to risk. [+1]

The Nelson-Aalen estimate allows the death rate to vary with time according to the data. [+½]

The sample size is very small so the results are not likely to be reliable. [+½]

The group of lives being considered is very varied, so we do not have a homogeneous group. [+½]

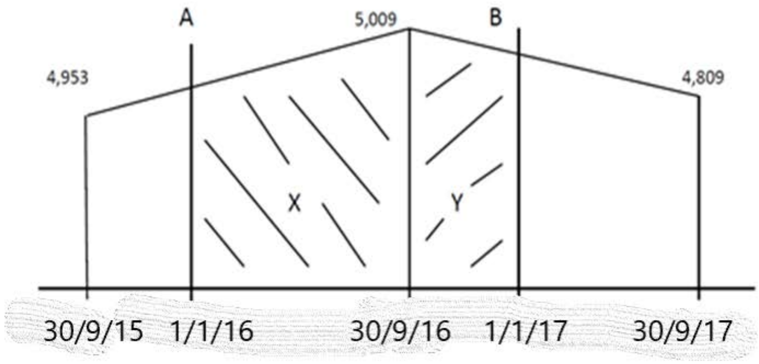
The lives under observation are by definition sick, so the rates we are coming out with are very high. [+½]

[max. 2]
[Total 12]

In part (i) a common error was to fail to add +1 to the difference between the last week and the first week. This produced an exposed to risk of 188 weeks and a death rate of 0.0266 per week. This was penalised by the loss of 1 mark. Candidates who made errors in part (i) which were carried forward into parts (ii) and (iii) were not penalised again in those later parts. Many candidates made only token efforts at part (iv). This was one of the part questions which led to the Pass Mark being reduced to 58, as almost no candidates scored more than +1 for this part.

Q11

(i) **Man Life**



See the diagram above. The required exposed to risk is represented by Area X + Area Y

Assuming that the population varies linearly over inter-census periods, [+½]

and that the data for 31 December in a year can be taken to represent the data for 1 January the following year

Number of policies in force on 1 January 2016 (A)

= (¾ * 4,953) + (¼ * 5,009) = 4,967 [+½]

Number of policies in force on 1 January 2017 (B)

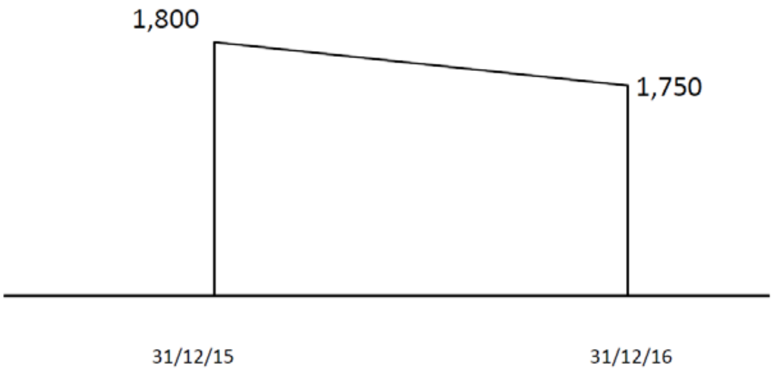
= (¾ * 5,009) + (¼ *4,809) = 4,959 [+½]

Area X = 9/24 * (4,967 + 5,009) = 3,741 [+½]

Area Y = 3/24 * (5,009 + 4,959) = 1,246 [+½]

Exposed to risk = 3,741 + 1,246 = 4,987 [+½]

Mixed Life



Assuming that the population varies linearly over inter-census periods, and that the data for 31 December in a year can be taken to represent the data for 1 January the following year

Exposed to risk = ½ (1,800 + 1,750) = 1,775 [+1]

Total male exposed to risk = 4,987 + 1,775 = 6,762 [+1]
[5]

(ii) We need to adjust the age definition for the female lives.

Assuming birthdays are spread evenly over calendar years, [+½]

and that the data for 31 December in a year can be taken to represent the data for 1 January the following year,

the number of policies in force aged 50 last birthday is equal to

0.5 * number of policies in force aged 50 nearest birthday
+ 0.5 * number of policies in force aged 51 nearest birthday [+½]

on 31 December 2015 this is ½ (1,506 + 1610) = 1,558 [+½]

on 31 December 2016 this is ½ (1,497 + 1,587) = 1,542 [+½]

so the exposed to risk for the female lives at age 50 last birthday is

½ (1,558 + 1,542) = 1,550 [+½]

Total exposed to risk of the combined portfolio is therefore

6,762 + 1,550 = 8,312 [+½]

[3]

(iii) This approach will only work if the mix of males and females remains the same. [+½]

It is not clear whether this will happen in the future. [+½]

Need to know what competitors are doing. [+½]

Other companies may base their rates on a different mix of in force business, or some estimate of future mix. [+½]

Consider mortality improvements going forward, and in particular the future development of the ratio between male and female death rates. [+1]

What demographic does the company want to target, e.g. only males? [+½]

Some selection effects are nullified by the fact that all companies are required to charge unisex rates. [+½]

The overall mix of business by gender may alter temporarily as those who are likely to lose out by the introduction of the new legislation may make a dash to get cover before the legislation comes into force. [+½]

[max. 3]

[Total 11]

Answers to parts (i) and (ii) were, overall, better than answers to similar questions on this part of the syllabus in previous sessions. Answers to part (iii) were very poor. This was the other part question which led to the Pass Mark being reduced to 58, as almost no candidates scored more than +2 for this part. However, it was disappointing that most candidates seemed not to have read the question. They wrote answers arguing that customers would switch from the company to other companies who charged different premiums to males and females, without realising that all companies were required by law to charge the same premiums to males and females. Assuming this applies to new business, not to existing business, then the common premium will be determined by the sex ratio applied to the pricing basis and profitability by how this compares with the mix of sales in the future.

Q12

(i) EITHER

The number of corpses in the refrigerator one morning is the number the previous morning, plus the number of deaths that day less the one the embalmer embalmed [+1]

OR

If no corpses are in the refrigerator, then
$p(0,0) = 0.497 + 0.348$ (since if there is 1 death the embalmer will embalm the one who dies)
$p(0,1) = 0.122$
$p(0,2) = 0.028$
$p(0,3) = 0.005$
$p(0,4) = 0$.

If one corpse is in the refrigerator, then
$p(1,0) = 0.497$
$p(1,1) = 0.348$
$p(1,2) = 0.122$
$p(1,3) = 0.028$
$p(1,4) = 0.005$ [+1]

Using similar calculations for 2, 3 and 4 corpses, we obtain the transition matrix, P, of the number of corpses in the refrigerator:

$$\begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{bmatrix} 0.845 & 0.122 & 0.028 & 0.005 & 0 \\ 0.497 & 0.348 & 0.122 & 0.028 & 0.005 \\ 0 & 0.497 & 0.348 & 0.122 & 0.033 \\ 0 & 0 & 0.497 & 0.348 & 0.155 \\ 0 & 0 & 0 & 0.497 & 0.503 \end{bmatrix}.$$ [+2]

[max. 3]

(ii) If x_i is the probability that the refrigerator contains i corpses at the start of a day, then

Using the transition matrix in part (i) and $\pi = \pi P$ we get [+½]

$x_0 = 0.845x_0 + 0.497x_1$
$x_1 = 0.122x_0 + 0.348x_1 + 0.497x_2$
$x_2 = 0.028x_0 + 0.122x_1 + 0.348x_2 + 0.497x_3$ [+2]
$x_3 = 0.005x_0 + 0.028x_1 + 0.122x_2 + 0.348x_3 + 0.497x_4$
$x_4 = 0.005x_1 + 0.033x_2 + 0.155x_3 + 0.503x_4$

Proceeding recursively we obtain

$x_1 = (0.155/0.497)x_0 = 0.312x_0$

$0.312x_0 = 0.122x_0 + 0.348(0.312x_0) + 0.497x_2$

$x_2 = (0.081/0.497)x_0 = 0.164x_0$

$0.164x_0 = 0.028x_0 + 0.122(0.312x_0) + 0.348(0.164x_0) + 0.497x_3$

$x_3 = (0.041/0.497)x_0 = 0.082x_0$

$0.082x_0 = 0.005x_0 + 0.028(0.312x_0) + 0.122(0.164x_0) + 0.348(0.082x_0)$
$+ 0.497x_4$ [+1½]

$x_4 = (0.020/0.497)x_0 = 0.040x_0$

So we have

$x_0 + 0.312x_0 + 0.164x_0 + 0.082x_0 + 0.040x_0 = 1$ [+½]

$x_0 = 0.626$
$x_1 = 0.195$
$x_2 = 0.103$
$x_3 = 0.051$
$x_4 = 0.025$ [+½]
[5]

(iii) The probability the funeral director has to contact the hospital is:

x_2Pr[4 deaths] + x_3Pr[3 or 4 deaths] + x_4[Pr 2 or more deaths] [+1]

$= 0.005x_2 + 0.033x_3 + 0.155x_4$

$= 0.005(0.103) + 0.033(0.051) + 0.155(0.025) = 0.006.$ [+1]
[2]

(iv) Probability the funeral director has to contact the hospital on Christmas Day is

x_1Pr[4 deaths] + x_2Pr[3 or 4 deaths] + x_3[Pr 2 or more deaths]
+ x_4[Pr 1 or more deaths] [+1]

= 0.005(0.195) + 0.033(0.103) + 0.155(0.051) +0.503(0.025) = 0.025. [+1]

[2]
[Total 12]

This question was based on a practical problem faced by a funeral director known to the Principal Examiner. It was poorly answered by most candidates. In part (i) there was +2 marks for the correct matrix and +1 for some explanation of how the matrix was determined. Few candidates could correctly formulate the matrix. A common error was to ignore the fact that the embalmer can embalm one corpse per day and will always do so provided a corpse is available. Some candidates assumed that once the fridge was full the funeral director would offload the entire contents to the hospital morgue, rather than just those corpses in excess of 4. Candidates whose matrices in part (i) were incorrect could score full credit for part (ii) if they correctly calculated the stationary distribution for the matrix they had produced in past (i). Few candidates attempted parts (iii) and (iv) and most attempts were incorrect. Full credit could be obtained in parts (iii) and (iv) for answers which applied the correct method to incorrect matrices.

END OF EXAMINERS' REPORT