INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

April 2012 examinations

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 who are 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. Although Examiners have access to the Core Reading, which is designed to interpret the syllabus, the Examiners are not required to examine the content of Core Reading. Notwithstanding that, the questions set, and the following comments, will generally be based on Core Reading.

For numerical questions the Examiners' preferred approach to the solution is reproduced in this report. Other valid approaches are always given appropriate credit; where there is a commonly used alternative approach, this is also noted in the report. For essay-style questions, and particularly the open-ended questions in the later subjects, this report contains all the points for which the Examiners awarded marks. This is much more than a model solution – it would be impossible to write down all the points in the report in the time allowed for the question.

T J Birse
Chairman of the Board of Examiners

July 2012

General comments on Subject CT4

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.

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 be penalised where excessive rounding has been used or where insufficient working is shown.

Comments on the April 2012 paper

The general performance was better than in any session since April 2008. Well-prepared candidates scored highly across the whole paper. A feature of this diet was that parts of questions which required an element of explanation or interpretation (such as Q7(ii) and Q12(vii)) were better answered than in previous diets, and this largely accounted for the increased pass rate. The comments that follow the questions concentrate on areas where candidates could have improved their performance. Candidates approaching the subject for the first time are advised to concentrate their revision in these areas.

1

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

where the Y_j are i.i.d. random variables and $X_0 = 0$

(ii) Is simple random walk when Y_j can have values $+1$ and -1 only

In part (i) few candidates gave the initial condition that $X_0=0$. Many candidates were confused as to the definitions of general, simple, and symmetric random walks (for example defining a simple random walk in part (i) and then stating for part (ii) that the probabilities of Y_j being +1 and -1 were both equal to 0.5).

2

(i) Users of data require rates subdivided by age and other criteria.

Models are based on the assumption that we can observe groups of identical lives.

Therefore it is important that we analyse groups of lives which are homogenous (or have the same mortality).

This can, for example, help avoid anti-selection.

(ii) Small numbers in some sub-groups leading to scanty data and non-credible rates or a large variance.

Sometimes relevant factors cannot be used because the relevant information cannot be collected on the proposal form because questions are unlikely to be answered honestly,

or because the key questions are intrusive or impractical for marketing or administrative reasons or make the questionnaire too long, or cannot be asked by law.

Can be difficult to ensure that events data and exposed-to-risk data are subdivided in the same way, leading to the principle of correspondence being violated.

Answers to this question were disappointing, even though not all the points listed above were required for full credit. In part (ii), many candidates made only the first point, about sparse data. Some candidates approached this question as practitioners or users of data rather than giving the general principles for which the question was asking. Nevertheless, if good points were made, this approach could earn full credit.

3

To work out the number of degrees of freedom (d.f.) we start with the number of age groups.

We reduce the d.f. because of the constraints imposed by the graduation process.

The reduction varies according to the graduation method:

parametric formula – one d.f. lost for each parameter estimated;

standard table – one d.f. lost for each parameter fitted and a further reduction due to the constraints imposed by the choice of standard table;

graphical – two or three d.f. lost for about every 10 ages graduated.

This question was generally well answered. Common errors were to suppose that only one d.f. is lost for the choice of standard table, and that for graphical graduation, two or three d.f. were lost in total, regardless of the number of ages being graduated.

4

(i)

Month 1	5/200 = 0.025
Month 2	8/190 = 0.042
Month 3	15/175 = 0.086
Month 4	10/150 = 0.067
Month 5	6/135 = 0.044
Month 6	3/125 = 0.024

(ii) To assess the impact of risk factors, a proportional hazards model would be useful because of its simple interpretation or because it allows the effect of each individual risk factor to be assessed.

The Gompertz model can be framed as a proportional hazards model, as can a semi-parametric model (such as the Cox model).

The Gompertz model would not be appropriate here, as it has a monotonically increasing or decreasing hazard,

whereas it is clear from part (i) that the hazard of symptoms returning first rises and then falls with duration.

A semi-parametric model allows the shape of the hazard to be determined by the data.

The semi-parametric model would be better than the Gompertz in this case.

In part (i) a minority of candidates subtracted half the deaths from the exposed-to-risk. Partial credit was given for this. Part (ii) was a higher skills question, and was poorly attempted by many candidates. Only a small proportion related their answers to the data *given and spotted that the empirical hazard calculated in part (i) was non-monotonic and so the Gompertz model would be a poor fit. Hardly any candidates pointed out that the Gompertz model can be framed as a proportional hazards model.*

5

(i) The likelihood of the data is given by:

$$L = \prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i},$$

where $f(t_i)$ is the probability density function and $S(t_i)$ is the survivor function.

Since $f(t_i)$ is related to the hazard function by

$f(t_i) = h(t_i)\, S(t_i)$

the likelihood can be rewritten:

$$L = \prod_{i=1}^{n} h(t_i)^{\delta_i} S(t_i).$$

Since

$$S(t_i) = \exp\left[-\int_{r=o}^{t_i} h(r)dr\right] = \exp[-At_i - \frac{1}{2}Bt_i^{\,2}],$$

$$L = \prod_{i=1}^{n} (A+Bt_i)^{\delta_i} \exp[-At_i - \frac{1}{2}Bt_i^2] \text{ as required.}$$

(ii) The log likelihood is given by:

$$\log L = \sum_{i=1}^{n}\left[\delta_i \log(A+Bt_i) - At_i - \frac{1}{2}Bt_i^{\,2}\right].$$

We are trying to maximise likelihood with respect to two parameters, so need partial differentials with respect to A and B:

$$\frac{\partial}{\partial A}\log L = \sum_{i=1}^{n}\left[\frac{\delta_i}{A+Bt_i} - t_i\right],$$

$$\frac{\partial}{\partial B}\log L = \sum_{i=1}^{n}\left[\frac{\delta_i t_i}{A+Bt_i} - \frac{1}{2}t_i^2\right].$$

The simultaneous equations satisfied by the MLEs are obtained by setting these to zero:

$$\sum_{i=1}^{n}\left[\frac{\delta_i}{A+Bt_i} - t_i\right] = 0,$$

$$\sum_{i=1}^{n}\left[\frac{\delta_i t_i}{A+Bt_i} - \frac{1}{2}t_i^2\right] = 0.$$

In part (i) many candidates failed to explain where the components of the likelihood came from by explaining the different contributions of the lives who were observed to die and those who were not. In part (ii) credit was given for knowing the correct method even if this was not executed. Credit was also given for differentiating a second time and showing that the second derivatives were negative (and hence that we do have maxima), even though this was not required for full marks.

6

(i) **Benefits**

Systems with long time frames can be studied in compressed time

Complex systems with stochastic elements can be studied (especially by simulation modelling).

Different future policies or possible actions can be compared either to see which best suits the requirements of a user or to examine different scenarios without carrying them out in practice, or to avoid potential costs associated with trialing in real life.

Models allow control over experimental conditions, so that we can reduce the variance of the results output without upsetting the mean values.

Parameters can be sensitivity tested using a model.

Limitations

Model development requires a lot of time and expertise, and hence can be /costly.

May need to run model lots of times (essential if it is a stochastic model).

Models more useful for comparing the results of input variations than for optimising outputs.

Models can look impressive, but can lull the user into a false sense of security. Impressive output is not a substitute for validity and close imitation of the real world.

This is more true the further into the future you project

Models rely heavily on the data input. If this is poor or lacking in credibility the output is likely to be flawed.

Users need to understand the model sufficiently well to be able to know when it is appropriate to apply it.

Interpretation of models can be difficult, and often outputs need to be seen in relative rather than absolute terms.

Models cannot take into account all possible future events (e.g. changes in legislation).

(ii) The model should be simple to apply.

The data specified are likely to be available from reliable sources.

Although it is possible that the starting point for the planned population may be wrong

Unforeseen events may take place such as a national epidemic which change the rates.

The model is relatively straightforward to explain to the planners/developers.

Should consider whether there are trends in fertility rates, rather than simply using current rates.

Mortality rates unlikely to be significant relative to the uncertainty in the projection,

because rates at ages with non-zero fertility rates should be small and child mortality rates should be low.

Current age distribution for the area may not be representative of that for the new town as, for example, rural areas may have different distributions to urban areas

Consider the type of houses being built and how they are marketing e.g. are they family houses?

May wish to consider experience of similar new towns.

May wish to consider whether national fertility rates are appropriate for a new town, where many young families may live.

Migration may affect the profile of the population, for example older families moving away and younger families buying their houses may mean the age structure remains relatively constant over time regardless of mortality and fertility rates.

The approach does not take account of non-state schooling or the possibility of children going to boarding school.

Part (i) of the question was standard bookwork and was well answered. The quality of answers to part (ii) varied: some candidates wrote lengthy and well-argued discussions; others made only cursory attempts. In part (ii), not all the points listed above were needed for full credit, and other sensible comments could also score marks.

7

(i) The sequence of events described may be summarised in the table below

Duration t_j	*Pies in shop n_j*	*Pies bought d_j*	*Pies destroyed or stolen, c_j*
1	12	2	0
2	10	3	0
3	7	0	1
4	6	2	0
5	4	0	2
6	2	1	0

The hazard of pies being bought is thus

2/12 at duration	1 hour
3/10 at duration	2 hours
2/6 at duration	4 hours
1/2 at duration	6 hours

The Nelson-Aalen estimate of the survival function, $S(t)$, is then

Duration	*Nelson-Aalen estimate of S(t)*
$0 \le t < 1$	1
$1 \le t < 2$	$\exp[-2/12] = 0.8465$
$2 \le t < 4$	$\exp[-(2/12 + 3/10)] = 0.6271$
$4 \le t < 6$	$\exp[-(2/12 + 3/10 + 2/6)] = 0.4493$
$6 \le t < 8$	$\exp[-(2/12 + 3/10 + 2/6 + 1/2)] = 0.2725$

The Nelson-Aalen estimate is a step function.

We need t for which $S(t) = 0.6$.

Therefore it will be 4 hours until Mr Bunn has sold 40% of his pies.

(ii) The estimate would not be a good basis on which to plan future production.

And how long it takes to sell 40% of your goods is not very relevant for future production.

It is based on only one day's experience, and a good basis for future production should be based on several days, probably involving different days of the week.

Sales of pies may vary seasonally: data from a winter's day may tell Mr Bunn little about the demand for pies in summer.

Mr Bunn might be more careful in future not to sit on his pies, and might take steps to avoid the dog from across the street stealing pies.

The proportion of pies sold will depend on the number of pies Mr Bunn stocks. He should not assume if he had twice as many pies he would still sell 40% of them in 4 hours.

Mr Bunn may vary his sales strategy, by, for example, reducing his prices

The method does, however, take account to of censored data.

In part (i) the question said "estimate", so some indication of how the answer was arrived at was necessary, although not every detail was required. As a bare minimum full credit could be obtained for first three hazards at times 1, 2 and 4, some statement of what the Nelson-Aalen estimate of S(t) is, the fact that we are looking for S(t) = 0.6, and some numbers to demonstrate that S(t) = 0.6 happens at duration 4 hours. Answers that used the logarithm of S(t) were acceptable. Answers to part (ii) were encouraging. A substantial proportion of candidates made sensible points.

8

(i) Chi-squared test (for overall goodness of fit)

(Modified) individual standardised deviations test (for outliers)

(ii) **Chi-squared test**

The null hypothesis is that the mortality among the members of the company's pension scheme is represented by the standard table.

The test statistic is $\sum_x z_x^{\,2}$, where the z_x are the standardised deviations.

Under the null hypothesis, this statistic has a chi-squared distribution with 8 degrees of freedom.

$$\sum_x z_x^{\ 2} = 0.052^2 + 0.967^2 + 2.528^2 + 0.328^2 + 1.234^2 + 0.250^2$$
$$+ 1.023^2 + 0.756^2 = 10.64.$$

The critical value of the chi-squared distribution with 8 degrees of freedom at the 5% significance level is 15.51.

Since 10.64 < 15.51

we do not reject the null hypothesis.

(Modified) individual standardised deviations test

Under the null hypothesis (same as for the chi-squared test)

we would expect individual deviations to be distributed Normal (0,1)

Only 1 in 20 of the z_x should lie above 1.96 in absolute value
OR
none should lie above 3 in absolute value
OR
about two thirds of the z_x should lie between −1 and +1
OR

Interval	(0,1)	(1,2)	(2,∞)
Actual deaths	5	2	1
Expected deaths	5.44	2.24	0.32

The largest deviation we have here is 2.528 in absolute value,

which is well outside the range −1.96 to +1.96,

therefore we have reason to reject the null hypothesis.

but, since we have 8 ages we cannot say definitively whether the null hypothesis should be rejected, but the large deviation of 2.528 suggests there may be a problem.

Many candidates scored highly on this question, though the chi-squared test was generally better done than the individual standardised deviations test. A surprising proportion of candidates thought that it was possible to perform the serial correlations test with the data given. The most common errors were to reduce the number of degrees of freedom in the chi-squared test (incorrect here as we are not testing a graduation) and a failure to spot the large deviation of 2.528, and state that this is a source of concern.

9

(i) Gender

Type of policy

Level of underwriting

Duration in force

Sales channel

Policy size

Occupation

Known impairments

Postcode/geographical area

Education

Socio-economic class / income

Marital status

(ii) For Gasperton we have, using the census formula central death rate

$$= \frac{25}{\frac{1}{2}(2{,}000 + 2{,}100)} = 0.0122 .$$

For Great Hawking we have central death rate

$$= \frac{21}{\frac{1}{2}(1{,}770 + 1{,}674)} = 0.0122 .$$

(iii) Let the death rate for smokers in Gasperton be γ_s, and that for non-smokers be γ_n.

We therefore have

$$0.5\gamma_s + 0.5\gamma_n = 0.0122$$

$$\gamma_s = 1.4\gamma_n$$

and hence

$$0.5(1.4)\gamma_n + 0.5\gamma_n = 0.0122$$

$$\gamma_n = \frac{0.0122}{1.2} = 0.0102$$

$$\gamma_s = \frac{0.0122(1.4)}{1.2} = 0.0142$$

Let the death rate for smokers in Great Hawking be ζ_s, and that for non-smokers be ζ_n.

We therefore have

$$0.2\zeta_s + 0.8\zeta_n = 0.0122$$

$$\zeta_s = 1.4\zeta_n$$

and hence

$$0.2(1.4)\zeta_n + 0.8\zeta_n = 0.0122$$

$$\zeta_n = \frac{0.0122}{1.08} = 0.0113$$

$$\zeta_s = \frac{0.0122(1.4)}{1.08} = 0.0158$$

(iv) The company would do better to vary the premiums on the basis of geographical area,

as it is clear that death rates in Great Hawking for both smokers and non-smokers are higher than those in Gasperton.

If the company does not differentiate its prices on the basis of geographical area, it may lose business in Gasperton to a rival company which does differentiate; conversely in Great Hawking it may attract new business from rival companies, but will underprice the product and hence risk its life assurance fund becoming insolvent.

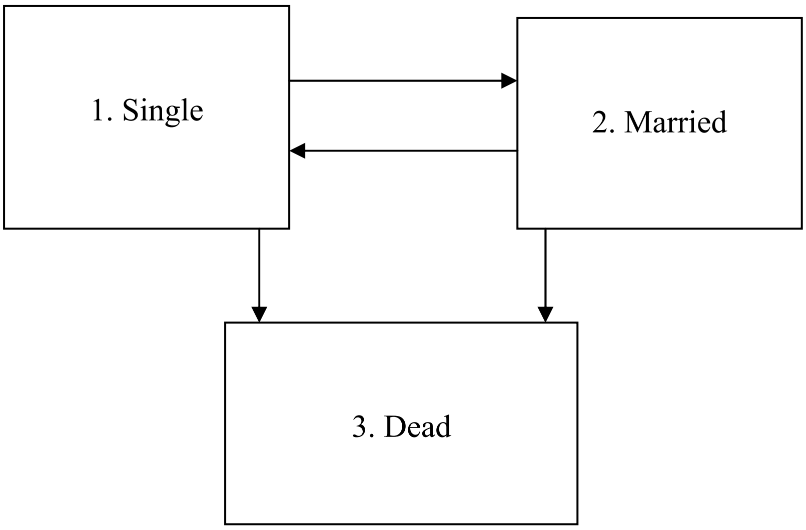
There are relatively little data, so it might be worth adopting a "wait and see" approach.

1.4 times the death rate will not translate as 1.4 times the premium. The difference may me relatively small, (although it is a 25 year term assurance so it probably is pretty significant).

Most candidates scored highly on parts (i) and (ii). Part (iii) was very poorly answered. A large number of candidates misinterpreted the question as meaning that the ratio of the <u>*numbers of deaths*</u> *to smokers and non-smokers was 1.4. This works for Gasperton because there are equal numbers of smokers and non-smokers in the exposed-to-risk, but for Great Hawking it produces incorrect results. Only a minority of candidates made a serious attempt at part (iv). Credit was given for any sensible comments in part (iv) which were consistent with the answers given to parts (ii) and (iii).*

10

(i)



(ii) $$L \propto \exp\left\{\left(-\mu^{12}-\mu^{13}\right)\nu^{1}\right\}\exp\left\{\left(-\mu^{23}-\mu^{21}\right)\nu^{2}\right\}\left(\mu^{12}\right)^{d^{12}}\left(\mu^{21}\right)^{d^{21}}\left(\mu^{13}\right)^{d^{13}}\left(\mu^{23}\right)^{d^{23}}$$

where

μ^{ij} is the transition intensity from state i to state j

ν^{i} is the total observed waiting time in state i

d^{ij} is the number of transitions from state i to state j

(iii) Taking the logarithm of the likelihood we get

$$\log_e(L) = -\mu^{13}\nu^1 + d^{13}\log_e(\mu^{13}) + \text{terms not involving } \mu^{13}.$$

Differentiate with respect to μ^{13}

$$\frac{d\ln(L)}{d\mu^{13}} = -\nu^1 + \frac{d^{13}}{\mu^{13}}.$$

Setting this to zero we obtain

$$\hat{\mu}^{13} = \frac{d^{13}}{\nu^1}.$$

To check it is a maximum differentiate again giving

$$\frac{d^2\log_e(L)}{(d\mu^{13})^2} = -\frac{d^{13}}{(\mu^{13})^2} \quad \text{which is always negative.}$$

(iv) The maximum likelihood estimate of μ^{13} is 12/10,298 = 0.001165.

The variance is $-1 \Big/ \dfrac{d^2\ln(L)}{(d\mu^{13})^2} = 12 \Big/ 10{,}298^2 = 1.13 \times 10^{-7}$.

This was the best answered question on the paper, with most candidates scoring at least 9 of the 11 marks available. In part (ii) many candidates omitted the constant of proportionality. In part (iv) the question says "estimate", so we needed some indication of where the answers came from for full marks.

11

(i) $X_{n+1} = X_n + Y_{n+1} = X_n + f(X_n)$

so the series X_i depends only on the current state and hence satisfies the Markov property.

$$Y_{n+1} = \frac{1}{4}\left(1 + \frac{X_n}{n}\right) = \frac{1}{4}\left(1 + \frac{\sum_{j=1}^{n} Y_j}{n}\right)$$

and hence depends on all the previous values of Y_i.

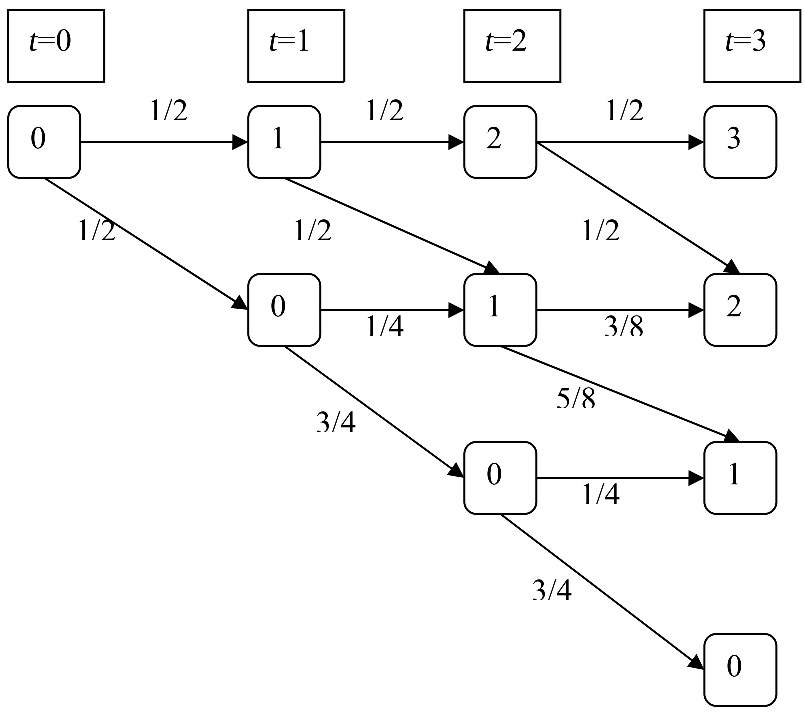
(ii) (a) It is not possible for the cumulative number of accidents to reduce (OR the cumulative number of accidents is an increasing/ non-decreasing function)

and so the process is not irreducible.

(b) The probabilities depend on the number of time periods n

so the process is not time homogeneous

(iii)



(iv) From the diagram above (or otherwise) it can be seen that there are three paths to the 2 accidents by time 3 box.

Required probability = Pr(0–0–1–2)+Pr(0–1–1–2)+Pr(0–1–2–2)

$$=\frac{1}{2}.\frac{1}{4}.\frac{3}{8}+\frac{1}{2}.\frac{1}{2}.\frac{3}{8}+\frac{1}{2}.\frac{1}{2}.\frac{1}{2}=\frac{3+6+8}{64}=\frac{17}{64}$$

(v) It is reasonable to assume that probability of having an accident depends on the number of previous accidents.

It is also reasonable that the effect of a previous accident should wear off over time.

There are likely to be other factors which have a significant effect on the probability of an accident, such as the fact that people who have recently had an accident might drive more carefully.

May want to give more weighting to recent years.

This was a demanding question, and only a minority of candidates scored highly. In part (i) few answers were sufficiently rigorous, many either re-stating the question or simply stating the Markov property. Part (ii)(a) was well answered, but in part (ii)(b) many candidates did not understand the term "time homogeneous". Most candidates made only sketchy attempts at part (iii) and (iv). Credit was given for calculations in part (iv) which demonstrated that candidates knew the correct method, even though the numbers were incorrect.

12

(i) A process with a discrete state space and discrete time space

where the future development is only dependent on the current state occupied.

OR

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

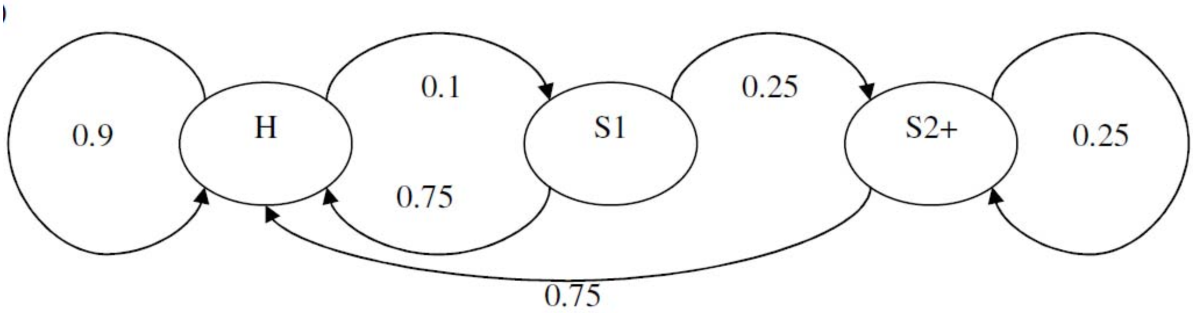
THEN EITHER THE THREE STATE SOLUTION

(ii) The sick pay depends on the duration of sickness, so to model with a time homogeneous Markov chain needs as a minimum the states:

Healthy (H)
Sick month 1 (S1)
Sick month 2 or more (S2+)

So the minimum number of states is 3.

(iii)



(iv) If using H, S1, S2+ then the stationary distribution π is given by:

$$\pi\begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.75 & 0 & 0.25 \\ 0.75 & 0 & 0.25 \end{pmatrix} = \pi$$

$$\pi_H = 0.9\pi_H + 0.75\pi_{S_1} + 0.75\pi_{S_{2+}}$$
$$\pi_{S_1} = 0.1\pi_H$$
$$\pi_{S_{2+}} = 0.25\pi_{S_1} + 0.25\pi_{S_{2+}}$$
$$\pi_H + \pi_{S_1} + \pi_{S_{2+}} = 1$$

$$3\pi_{S_{2+}} = \pi_{S_1}$$
$$30\pi_{S_{2+}} = \pi_H$$

Implies

$$\pi_{S_{2+}} = \frac{1}{34},\ \pi_{S_1} = \frac{3}{34},\ \pi_H = \frac{15}{17}$$

(v) Let percentage of salary when healthy be *a*%

Then in the stationary state looking at payments for the next month we need

Expected income = Expected outgo.

Probability healthy **a*% of salary = Probability of 100% sick pay*100% of salary + Probability on 50% sick pay*50% of salary:

$$\left(\frac{15}{17}*0.9+\frac{2}{17}*0.75\right)a = \left(\frac{3}{34}*0.25+\frac{1}{34}*0.25*50\%+\frac{15}{17}*0.1\right)$$

$0.88235a = 0.113971$
a =12.917% of salary.

(vi) Now just need a two state version {H,S}

$$\pi_H = \frac{15}{17},\ \pi_S = \frac{2}{17}$$

and need contribution rate > 2/15 = 13.333% of salary.

OR THE FOUR STATE SOLUTION

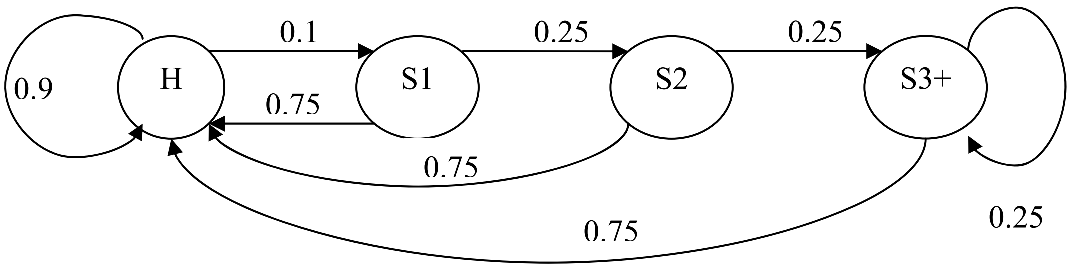
(ii) The sick pay depends on the duration of sickness, so to model with a time homogeneous Markov chain needs as a minimum the states:

Healthy (H)
Sick month 1 (S1)
Sick month 2 (S2)
Sick month 3 or more (S3+)

So the minimum number of states is 4.

[2]

(iii)



(iv) If using H, S1, S2, S3+ then the stationary distribution π is given by:

$$\pi\begin{pmatrix} 0.9 & 0.1 & 0 & 0 \\ 0.75 & 0 & 0.25 & 0 \\ 0.75 & 0 & 0 & 0.25 \\ 0.75 & 0 & 0 & 0.25 \end{pmatrix} = \pi$$

$$\pi_H = 0.9\pi_H + 0.75\pi_{S_1} + 0.75\pi_{S_2} + 0.75\pi_{S_{3+}}$$

$$\pi_{S_1} = 0.1\pi_H$$
$$\pi_{S_2} = 0.25\pi_{S_1}$$
$$\pi_{S_{3+}} = 0.25\pi_{S_2} + 0.25\pi_{S_{3+}}$$
$$\pi_H + \pi_{S_1} + \pi_{S_2} + \pi_{S_{3+}} = 1$$

$$\pi_H = 10\pi_{S_1}$$

$$\pi_{S_2} = \frac{1}{4}\pi_{S_1}$$

$$0.75\pi_{S_{3+}} = 0.25\pi_{S_2}$$

$$\pi_{S_{3+}} = \frac{1}{3}.\frac{1}{4}\pi_{S_1}$$

$$\pi_{S_1}\left\{10+1+\frac{1}{4}+\frac{1}{12}\right\}=1$$

Implies

$$\pi_{S_1} = \frac{12}{136} = \frac{3}{34},\ \pi_H = \frac{120}{136} = \frac{15}{17},\ \pi_{S_2} = \frac{3}{136}\ \pi_{S_{3+}} = \frac{1}{136}$$

(v) Let percentage of salary when healthy be *a*%

Expected income = Expected outgo

Probability healthy **a*% of salary = Probability of 100% sick pay*100% of salary
Probability on 50% sick pay*50% of salary

$$\frac{15}{17}a = \left(\frac{3}{34}+\frac{3}{136}\right)+\frac{1}{2}\left(\frac{1}{136}\right)$$

0.88235*a* = 0.113971

a =12.917% of salary.

(vi) Now all those not healthy get 100% so

$$\frac{15}{17}a = \frac{2}{17}$$

and need contribution rate > 2/15 = 13.333% of salary

(vii) The reduction in cost is calculated as 3.23%.

This is not particularly significant either relative to the likely uncertainty in the assumptions or because recovery rates are so high.

The reduction in sick pay is likely to encourage employees to try to get back into work.

This question was well answered despite its complexity. Most candidates went for the four state solution, and there were many correct answers to parts (i)-(iv). In parts (v) and (vi) a *common mistake was to fail to divide by the proportion of healthy employees, as only healthy employees (i.e. those not receiving sick pay) contribute to the scheme. Answers to part (vii) often included sensible comments that gained credit, even if some candidates answered as if the scheme had an unlimited supply of funds!*

END OF EXAMINERS' REPORT