INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

April 2015 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 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 at the date the examination was set. Candidates should take into account the possibility that circumstances may have changed if using these reports for revision.

F Layton
Chairman of the Board of Examiners

June 2015

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 2015 paper

The average performance was similar to that of recent April sessions. Well-prepared candidates scored highly across most of the paper, with one in ten candidates scoring 70% or more, and a highest mark of 86%. There were one or two sections of the paper where very few candidates scored full credit, even though these dealt with bookwork which was in the Core Reading.

In general, there was a tendency for candidates to fail to score marks by missing out the more "wordy" sections of questions even when these were straightforward bookwork.

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 include these areas in their revision.

1 (i) This is defined as $X_n = Y_1 + Y_2 + \dots Y_n$

where the random variables Y_j (the steps of the walk) are mutually independent with the common probability distribution:

$\Pr[Y_j = 1] = p$,

$\Pr[Y_j = -1] = 1 - p$.

(ii) It operates in discrete time with a discrete state space.

(iii) Any reasonable practical application
e.g. cumulative results of the Oxford vs Cambridge boat race (net lead of Cambridge over Oxford) measured annually.
OR how much a gambler has won or lost if he wins or loses £1 on every bet.
OR number of cars in a car park controlled by a single entry/exit barrier measured after each time the barrier goes up.

Most candidates answered parts (i) and (ii) well, though many missed the point about the Y_t being independent. Some of the examples in part (iii) were rather contrived.

2 (a) A beetle in the wild has a force of mortality equal to $3\mu/2$.

So for a beetle in the wild

we have $S(8) = \exp[-8(3\mu/2)] = \exp(-12\mu) = 0.58$

Hence

$-12\mu = -0.5447$

$\mu = 0.0454$.

Therefore a beetle reared in the protected environment will have an 8 day survival probability of

$\exp(-8*0.0454) = 0.6955$.

(b) A beetle in the protected environment has a probability of surviving 6 days equal to

$\exp(-6*0.0454) = 0.7616$,

and a probability of surviving 2 days in the wild of

$\exp(-2*1.5*0.0454) = 0.8727$

Therefore this beetle's probability of surviving 8 days is

$0.7615 \times 0.8727 = 0.6646$.

This was the best answered question on the entire paper, with many candidates scoring full marks. The most common error was to assume the rate in the wild was twice, rather than 1.5 times, the rate in the protected environment in part (a). If this was carried through correctly into the rest of the answer then credit was given for subsequent calculations.

3 (i) In a proportional hazards model the hazard of experiencing an event may be factorised into two components:

one depending only on duration since some start event, which is known as the baseline hazard, and the other depending only on a set of covariates and associated parameters.

Thus the ratio between the hazards for any two individuals with different values of the covariates is constant across all durations.

The baseline hazard applies to an individual with the value zero on all covariates.

(ii) The proportionality of the hazards makes estimating the impact of covariates on the hazard straightforward (through partial likelihood).

Widely available statistical software packages have built-in routines for the Cox model.

The Cox model is semi-parametric so the baseline hazard does not need to be specified, and can be determined by the data (as with a Kaplan-Meier hazard).

It ensures that the hazard is always positive.

It is easy to communicate.

There were some good attempts at this question. However, many candidates seemed to think that the Cox model and the proportional hazards (PH) model were the same thing. In fact, the Cox model is just one of a class of PH models. In part (i) we looked for knowledge of the attractive characteristics of PH models <u>in general</u>, whereas in part (ii) we gave credit for advantages of the Cox model <u>in particular</u>, as well as for general attributes of PH models *which are useful in practice. In part (i) some candidates simply wrote down a formula for a PH model. If all the terms were defined, partial credit was given for this.*

4 (i) Develop a well-defined set of objectives which need to be met by the modelling process.

Plan the modelling process and how the model will be validated.

Collect and analyse the necessary data.

Define the parameters for the model and consider appropriate parameter values.

Define the model initially by capturing the essence of the real world system (refining the level of detail in the model can come at a later stage).

Involve experts on the real world system you are trying to imitate so as to get feedback on the validity of the conceptual model.

Decide on whether a simulation package or a general purpose language is appropriate for the implementation of the model.

Choose a statistically reliable random number generator that will perform adequately in the context of the complexity of the model.

Write the computer program for the model.

Debug the program to make sure it performs the intended operations in the model definition.

Test the reasonableness of the output of the model.

Review and carefully consider the appropriateness of the model in the light of small changes in input parameters.

Analyse the output from the model.

Ensure that any relevant professional guidance has been complied with.

Communicate and document the results of the model.

(ii) Objectives.
Is a single pricing table needed for a defined set of cover/deferred period/definition of sickness?
Is a range of prices needed for different cover levels?
Is a simple price needed, or a confidence interval around profitability etc.?

Data.
Look at national industry data and possibly international data on sickness and recovery rates.
Take care as to the definitions of sickness.
Look at trends.
Is recovery rate affected by reduction in benefits?
Might future trends change due to economic influences, medical developments, change in government policy to state benefits?
What might the likely level of expenses be? This will depend on expected sales volumes (and commission levels).
It is a new product so perhaps the company has no experience of claim monitoring. Look at industry data and in house data for new product launches.

Modelling process.
Start with a simple model reflecting sickness and recovery rates.
Build in expenses and reserving later.

Define the model.
Type of model will be determined by the objectives, deterministic for a simple set of premium rates, stochastic if confidence intervals required.

Reasonableness of output.
Look at the premium rates generated and compare them with competitors in the market.
Are they roughly where you would expect them to be?

Sensitivity to changes in input parameters.
Change each important parameter slightly; sickness rates, recovery rates, expenses and ensure that the impact on the premium rates is not huge.
If it is, review the product design.
Can risk be reduced by introducing, say, annually reviewable premium rates.

Compliance with professional guidance and regulatory environment.
Look at professional guidance and legislative requirements, might Solvency II impact the cost of capital to make the product uncompetitive or uneconomic?

Part (i) was well answered by most candidates. In part (ii) other suggestions scored credit provided they related to one of the stages identified in the answer to part (i) and dealt with a model to price a new sickness benefit. No credit was given for comments which were not specifically related to the pricing of a new sickness benefit product (many candidates made general points which were applicable to any model). Note that the level of detail given above is well in excess of that the Examiners required for full credit. Nevertheless, many candidates gave responses which were too brief and vague to gain much credit for this part.

5 (i) A life alive at age x at time t should be included in the exposed-to-risk if and only if, were that life to die immediately, his or her death would be included in the deaths at age x, d_x.

(ii) Those aged 22 last birthday on 30 June 2013 were born between 1 July 1990 and 30 June 1991, so half of them were born in 1990 and half in 1991.

Assuming that birthdays are evenly distributed across calendar years,

The number of persons aged 22 last birthday entering during each period is

10.00 – 11.30 p.m.	0.5(200 + 150) = 175
11.30 p.m. – 12.00 midnight	0.5(400 + 400) = 400
12.00 midnight – 1.00 a.m.	0.5(350 + 300) = 325

THEN EITHER

The number of persons aged 22 last birthday in the nightclub at 10.00 p.m., 11.30 p.m., 12.00 midnight, 1.00 a.m. and 2.00 a.m. is therefore

10.00 p.m.	0
11.30 p.m.	175
12 midnight	575
1.00 a.m.	900
2.00 a.m.	900

Using the census approximation and assuming that arrivals are evenly distributed across time,

the exposed to risk in person-hours is

$$1.5\left(\frac{0+175}{2}\right)+0.5\left(\frac{175+575}{2}\right)+1\left(\frac{575+900}{2}\right)+900$$
$$=131.25+187.5+737.5+900$$

$$= 1{,}956.25$$

OR

Using the census approximation and assuming that arrivals are evenly distributed across time,

the exposed-to-risk in person-hours is

175(3.25) + 400(2.25) + 325(1.5) = 1,956.25.

AND HENCE

the rate of requiring medical attention is

$$\frac{40}{1{,}956.25} = 0.02045 \text{ per person hour.}$$

Common errors were to use the wrong year or years of birth, to fail to cumulate the arrivals (i.e. to realise that once inside the building, customers remained until 2.00 a.m.), and to forget the final hour, during which the club was full and no more customers entered. In general, answers to this question were rather better than answers to similar questions on other recent examination papers. In part (ii) candidates were expected to relate the assumptions to the specific stage of the derivation to which they applied. Candidates who wrote down lists of assumptions – some relevant to the answer, others not – scored little credit.

6 (i) **Signs Test**

Under the null hypothesis, the number of positive deviations (2013 higher than 2012) is distributed Binomial (25, 0.5).

We have 17 positive deviations

ALTERNATIVE 1: NORMAL APPROXIMATION

As the number of ages is large enough, we can use the normal approximation, in which the number of positive deviations is distributed Normal$\left(\frac{25}{2}, \frac{25}{4}\right)$.

THEN EITHER

A z-score for 17 positive deviations is $\left(\frac{17-12.5}{\sqrt{6.25}}\right) = 1.8$

OR

with a continuity correction a z-score for 17 positive deviations is $\left(\frac{16.5-12.5}{\sqrt{6.25}}\right) = 1.6$.

AND HENCE

Since 1.8 (or 1.6) < 1.96 (2-tailed test)

we do not have sufficient evidence to reject the null hypothesis.

ALTERNATIVE 2: EXACT TEST

Pr[exactly 17 positive signs] = $\binom{25}{17}0.5^{25} = 0.0322$.

Since 0.0322 > 0.025 (2-tailed test)

we do not have sufficient evidence to reject the null hypothesis.

Grouping of Signs Test

We have 25 age groups, 17 positive signs, and 2 positive groups

ALTERNATIVE 1: NORMAL APPROXIMATION

Using the Normal approximation (as we have more than 20 ages),
the number of positive groups is distributed Normal $\left(\frac{17(8+1)}{25}, \frac{(17*8)^2}{(25)^3}\right)$,

which is Normal(6.12, 1.18).

We therefore compute a z-score for 2 runs as $\left(\frac{2-6.12}{\sqrt{1.18}}\right) = -3.79$.

Since Pr ($z < -3.79$) << 0.05 (one-tailed test) (or -3.79 < -1.645),

we reject the null hypothesis.

ALTERNATIVE 2: EXACT CALCULATION

Probability of getting 2 or fewer positive groups is

$$\frac{\binom{16}{0}\binom{9}{1}}{\binom{25}{17}} + \frac{\binom{16}{1}\binom{9}{2}}{\binom{25}{17}} = \frac{9}{1{,}081{,}575} + \frac{576}{1{,}081{,}575} = 0.000541$$

Since this is less than 0.05

we reject the null hypothesis.

ALTERNATIVE 3: USING THE TABLE IN THE "GOLD BOOK"

Using the table on p. 189 of the "Gold Book", with $n_1 = 17$, $n_2 = 8$

the table shows that we reject the null hypothesis with 3 or fewer runs of positive signs.

Since we only have 2 positive runs and 2 < 3 we reject the null hypothesis.

(ii) The results of the Signs Test suggest that the underlying rates in 2013 are not systematically higher or lower than those in 2012.

The null hypothesis was rejected by the Grouping of Signs Test implying that the shape of the distribution of sickness rates in 2013 is different from that in 2012.
However this is only one year's data and the company might wait to see if a trend develops, or investigate whether there was a specific factor operating in 2012 or 2013 which caused the change.

If the shape of sickness rates makes them markedly different in 2013 from 2012 at ages where much business is sold, this will have implications for profitability and pricing.

In part (i) most candidates used the Normal approximation for the Signs Test. If the exact version was used, it is not necessary to compute Pr[17 or more positive signs] as Pr[exactly 17 positive signs] > 0.025. The correct value for Pr[17 or more positive signs] is 0.0538. Answers to part (ii) were poor. Full credit was given for summarising the immediate implications of the results obtained in part (i) for the comparison of the rates in 2013 and 2012 and for making some comment about the potential financial implications for the company. Thus full credit could be obtained for less than is given in the model solution above. Comments in part (ii) that were consistent with the actual results obtained by the candidate in part (i) were given credit, even if the tests in part (i) were performed incorrectly and reached conclusions different from those above.

7 (i) A Markov chain is a stochastic process with discrete states operating in discrete time in which

EITHER

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

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

OR

the probabilities of moving from one state to another depend only on the present state of the process: the history of the process before the current state is irrelevant.

(ii) This is based on the number of links to the site and where they go.

$$P = \begin{array}{ccccc} From\backslash To & N & B & C & H \\ N & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ B & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ C & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ H & 0 & 0 & 1 & 0 \end{array}$$

(iii) Stationary distribution satisfies $\pi = \pi P$

$$\frac{1}{2}\pi_B + \frac{1}{3}\pi_C = \pi_N \qquad (1)$$

$$\frac{1}{2}\pi_N + \frac{1}{3}\pi_C = \pi_B \qquad (2)$$

$$\frac{1}{2}\pi_N + \frac{1}{2}\pi_B + \pi_H = \pi_C \qquad (3)$$

$$\frac{1}{3}\pi_C = \pi_H \qquad (4)$$

Also $\pi_N + \pi_B + \pi_H + \pi_C = 1$

From (1) and (2)

$\pi_N = \pi_B$

From (3) and (4)

$\pi_N = \pi_B = \frac{2}{3}\pi_C$

So $\left(\frac{2}{3} + \frac{2}{3} + 1 + \frac{1}{3}\right)\pi_C = 1$

Hence

$$\pi_N = \frac{1}{4}, \pi_B = \frac{1}{4}, \pi_C = \frac{3}{8}, \pi_H = \frac{1}{8}$$

Both parts of this question were well answered by most candidates. In the final answer to part (ii), it was important for candidates to indicate which probability applied to which state, rather than just listing four numbers.

8 (i) All our models and analyses are based on the assumption that we can observe groups of identical lives (or at least, lives whose mortality characteristics are the same).

Although in practice, this is never possible.

We can at least subdivide our data according to characteristics known, from experience, to have a significant effect on mortality.

This ought to reduce the heterogeneity of each class so formed.

(ii) Sex
Age
Type of policy
Level of underwriting
Duration in force
Sales channel
Policy size
Occupation or socio-economic group
Known impairments/medical history
Postcode/geographic location
Marital status

(iii) EITHER

If the company changing its policy charges both smokers and non-smokers a premium equal to the rate typically charged to smokers, then, relative to other companies, it will become poor value for non-smokers.

The company changing its policy will therefore lose business from non-smokers (whom it will charge more than an actuarially fair premium).

The portfolio will (eventually) be made up mostly of smokers (whom it will charge an actuarially fair premium).

The volume of business sold is likely to decrease, possibly to the extent that it does not cover the expenses estimated in the pricing basis.

OR

If the company changing its policy charges both smokers and non-smokers a premium equal to the rate typically charged to non-smokers, then relative to other companies, it will become good value for smokers (and acceptable value for non-smokers).

The company changing its policy will therefore attract more business from smokers (whom it will charge less than an actuarially fair premium). This is a form of anti-selection.

The smoker business is likely to be unprofitable, although the increase in business will reduce the overheads per policy

This is likely to lead to losses for the company changing its policy.

OR

If the company changing its policy charges both smokers and non-smokers a premium somewhere between the rate typically charged to smokers and the rate typically charged to non-smokers, then relative to other companies, it becomes good value for smokers and poor value for non-smokers.

The company changing its policy will therefore attract business from smokers and lose business from non-smokers (whom it will charge more than an actuarially fair premium). This is a form of anti-selection.

The smoker business is likely to be unprofitable, but any remaining non-smoker business will be profitable.

This may eventually lead to losses of the company changing its policy.

This question was generally well answered, and part (iii) was very well answered. In part (i) the question asked about the reasons why data are subdivided when undertaking investigations, so the answers were expected to reflect the assumptions underlying our models, rather than the convenience of users (e.g. pricing issues). In part (iii) most candidates implicitly supposed that the new single premium was between the previous smoker and non-smoker premiums, though few explicitly stated this. For full credit the Examiners were looking for consideration of at what level the single premium might be set and the consequences of this: hence the range of alternative approaches given above.

9 (i) When using a large experience

EITHER to produce a standard table.
OR where a suitable formula can be found to fit at all ages.

(ii) **Advantages**

It is straightforward to extend the statistical theory of estimation from one parameter to several.

Provided a reasonably small number of parameters is used, the resulting graduation will be acceptably smooth.

When comparing several experiences, the same parametric formula can be fitted to all of them. Differences between the parameters, given their standard errors, give insight into the differences between the experiences.

Disadvantages

It can be difficult to find a single formula to fit at all ages.

Care is required when extrapolating. The fit of the curve will probably be best where there is most data, but results where data are scanty (e.g. at extreme ages) may be poor and require adjustment.

(iii) (a) The χ^2 test compares an "observed" experience with an "expected" experience.

It is essential when making this comparison that the two sets of experiences be independent.

This is normally the case when considering the similarity of two sets of data.

However when comparing the difference between an "observed" experience and that "expected"from graduated data, there is a problem because the graduated data have been derived from the observed experience.

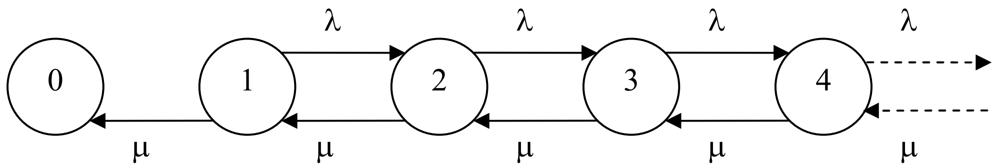
Because of this, we need to make it easier to reject the null hypothesis, and we achieve this by reducing the number of degrees of freedom used in the chi-squared test.

(b) When the graduation has been carried out by parametric formula, we reduce the degrees of freedom by one for each parameter estimated from the observed data.

This was a bookwork question, to which answers were very disappointing. In part (ii) many candidates were unable to reproduce the points made in Unit 12, page 7 of the Core Reading. In part (iii) most candidates knew that the number of degrees of freedom should be reduced, *and many wrote that the reduction was equal to the number of parameters in the formula used for the graduation. But very few candidates explained why the reduction was needed (para. 5.4 in Unit 12, page 9 of the Core Reading).*

10 (i) $\{0,1,2,3,4\ldots.\}$

(ii)



(iii) Generator matrix

$$
\begin{array}{c}
\text{Lives} \quad 0 \qquad 1 \qquad\qquad 2 \qquad\qquad 3 \qquad\qquad 4 \qquad \ldots
\end{array}
$$

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & . \\
\mu & -(\mu+\lambda) & \lambda & 0 & 0 & \\
0 & \mu & -(\mu+\lambda) & \lambda & 0 & \\
0 & 0 & \mu & -(\mu+\lambda) & \lambda & \\
0 & 0 & 0 & \mu & -(\mu+\lambda) & \\
. & & & & & .
\end{pmatrix}
$$

(iv) EITHER

If a Markov jump process X_t is examined only at the times of transition, the resulting process is called the jump chain associated with X_t.

OR

A jump chain is each distinct state visited in the order visited where the time set is the times when states are moved between.

(v) Lives 0 1 2 3 4 ...

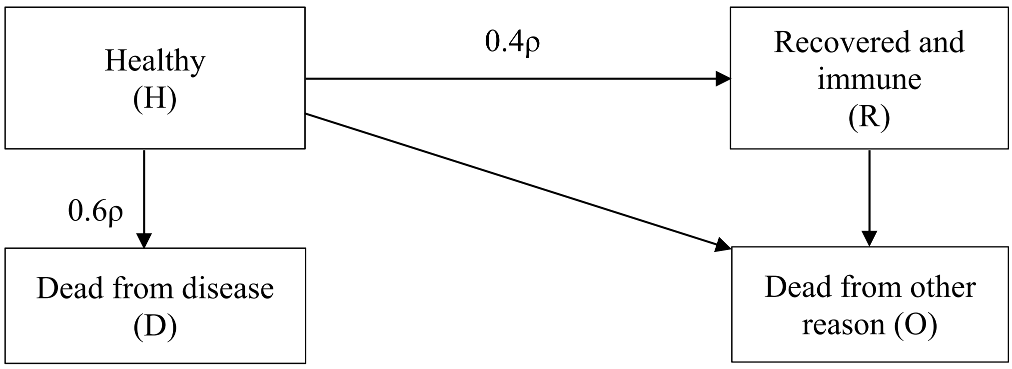
$$
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & \text{etc.} \\
\mu/(\mu+\lambda) & 0 & \lambda/(\mu+\lambda) & 0 & 0 & \\
0 & \mu/(\mu+\lambda) & 0 & \lambda/(\mu+\lambda) & 0 & \\
0 & 0 & \mu/(\mu+\lambda) & 0 & \lambda/(\mu+\lambda) & \\
0 & 0 & 0 & \mu/(\mu+\lambda) & 0 & \\
\text{etc.} & & & & &
\end{pmatrix}
$$

(vi) $\left(\dfrac{\mu}{\mu+\lambda}\right)^3$

Many candidates scored respectably on this question. The most common error was to define the state space as {0, 1, 2, 3}, ignoring the possibility that the player could, at any time, have found more extra lives than (s)he has lost. Candidates who used this state space were penalised in parts (i) and (ii) but could score full credit for later parts by carrying through their answer correctly. The other commonly occurring errors were to allow a transition out of state 0 (this is not possible, as the game ends when the player has no lives left), or to ignore the absorbing state 0 completely. Again, these errors were penalised in parts (i) and (ii) but credit was given for correctly following through into later parts. Candidates who ignored state 0 completely were unable to give a sensible answer to part (vi). In part (iv) a disappointing number of candidates simply provided a general definition of a Markov chain (i.e. repeating the answer to Question 7, part (i)), rather than relating the Markov jump chain to a Markov jump process.

11 ALTERNATIVE 1

(i)



(ii) The likelihood is

$$L \propto \exp\left\{\left(-0.6\rho\ -0.4\rho\ -\mu^{HO}\right)\nu^{H}\right\}\exp\left\{\left(-\mu^{RO}\right)\nu^{R}\right\}\left(0.6\rho\ \right)^{d^{HD}}$$

$$\left(0.4\rho\right)^{d^{HR}}\left(\mu^{HO}\right)^{d^{HO}}\left(\mu^{RO}\right)^{d^{RO}}$$

where

the four states are H – healthy, R – recovered, D – dead from the disease and O – dead for some other reason, and

ν^{I} is the waiting time in state I ($I = H, R$)

d^{IJ} is the number of transitions from state I to state J ($J = R, D, O$) and μ^{IJ} is the intensity of the transition from state I to state J and ρ is the rate of first time sickness

(iii) Taking logarithms of the likelihood we have:

$\ln L = (-0.6\rho - 0.4\rho)\nu^H + d^{HS}\ln(0.6\rho) + d^{HR}\ln(0.4\rho)$ plus terms not dependent on ρ

Differentiating with respect to ρ gives:

$$\frac{d(\ln L)}{d\rho} = -\nu^H + \frac{0.6d^{HS}}{0.6\rho} + \frac{0.4d^{HR}}{0.4\rho}$$

and setting this to zero gives a maximum likelihood estimate of ρ

$$\hat{\rho} = \frac{d^{HD} + d^{HR}}{\nu^H}$$

This is a maximum as the second derivative $\frac{d^2(\ln L)}{(d\rho)^2} = -\frac{d^{HD} + d^{HR}}{(\rho)^2}$ must be negative.

(iv) We want a person in H at $t = 0$ not dead at $t = 3$.

So they can either be healthy throughout:

$${}_t p_0^{HH} = \exp\left\{-\int_0^3 \left(0.4\rho + 0.6\rho + \mu^{HO}\right)du\right\} = \exp\left\{-\int_0^3 \left(\rho + \mu^{HO}\right)du\right\}$$

or go H to R and stay there which is the integral between
0 and 3 of the product of

surviving healthy for a period u i.e. $\exp-\left\{\left(\rho + \mu^{HO}\right)u\right\}$

getting sick and recovering at time u i.e. $0.4\rho du$

and staying recovered i.e. $\exp\left[-\left\{\left(\mu^{RO}\right)(3-u)\right\}\right]$,

where we ignore the short time spent in the sick state.

So altogether $\int_0^3 \exp\left[-\left\{\left(\rho + \mu^{HO}\right)u\right\}\right] 0.4\rho \exp\left[-\left\{\left(\mu^{RO}\right)(3-u)\right\}\right] du$.

(v) We need enough information to calculate the number of transitions from Healthy to Sick and the waiting time in the "Healthy" state. Hence we will need:

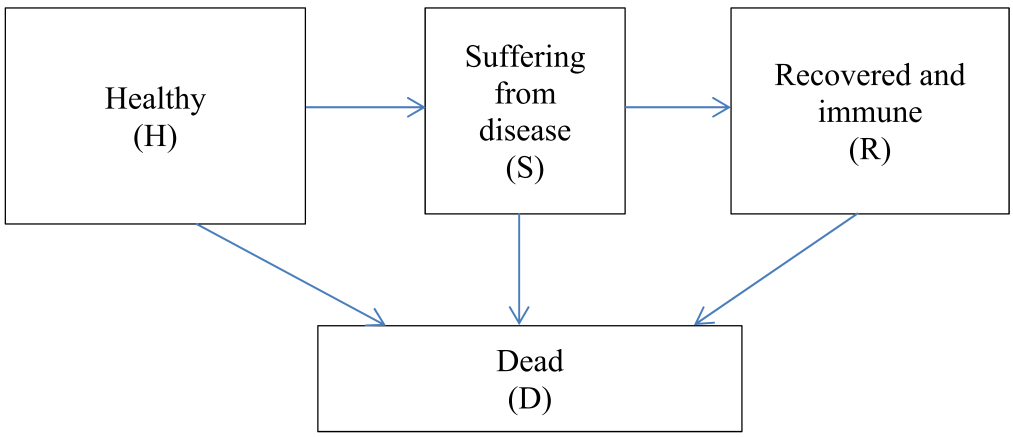
Date of birth of all births in last three years.
Date of death of all "Healthy" deaths in last three years.

Date of any other immigration or emigration of Healthy people in last three years.

Date at which each person who contracted the disease fell ill

ALTERNATIVE 2

(i)



(ii) The likelihood is

$$L \propto \exp\left\{\left(-\mu^{HS} - \mu^{HD}\right)v^{H}\right\}\exp\left\{\left(-\mu^{RD}\right)v^{R}\right\}\exp\left\{\left(-\mu^{SR} - \mu^{SD}\right)v^{S}\right\}\left(\mu^{HS}\right)^{d^{HS}}$$
$$\left(\mu^{HD}\right)^{d^{HD}}\left(\mu^{RD}\right)^{d^{RD}}\left(\mu^{SR}\right)^{d^{SR}}\left(\mu^{SD}\right)^{d^{SD}}$$

where

v^{I} is the waiting time in state I,
d^{IJ} is the number of transitions from state I to state J,
and μ^{IJ} is the intensity of the transition from state I to state J.

The four states are H – healthy, S – suffering from disease,
R – recovered, D – dead.

(iii) Taking logarithms of the likelihood we have:

$$\ln L = -\mu^{HS}\nu^{H} + d^{HS}\ln\left(\mu^{HS}\right) \text{ plus terms not dependant on } \mu^{HS}$$

Differentiating with respect to μ^{HS} gives:

$$\frac{d\left(\ln L\right)}{d\mu^{HS}} = -\nu^{H} + \frac{d^{HS}}{\mu^{HS}}$$

and setting this to zero gives a maximum likelihood estimate of μ^{HS}

$$\hat{\mu}^{HS} = \frac{d^{HS}}{\nu^{H}}$$

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

(iv) We want a person in H at $t = 0$ neither suffering from the disease nor dead at $t = 3$.

So they can either be healthy throughout:

$${}_{t}p_0^{HH} = \exp\left\{-\int_0^3\left(\mu^{HS} + \mu^{HD}\right)du\right\}$$

or to go H to S to R and stay there which is the integral between 0 and 3 of the product of:

surviving healthy for a period u i.e. $\exp-\left\{\left(\mu^{HS} + \mu^{HD}\right)u\right\}$

getting sick at time u i.e. $\mu^{HS}du$ and recovering, i.e. 0.4

and staying recovered i.e. $\exp\left[-\left\{\left(\mu^{RD}\right)\left(3-u\right)\right\}\right]$,

where we ignore the short time spent in the sick state.

So altogether $\int_0^3 \exp\left[-\left\{\left(\mu^{HS} + \mu^{HD}\right)u\right\}\right]\mu^{HS}\,0.4\exp\left[-\left\{\left(\mu^{RD}\right)\left(3-u\right)\right\}\right]du$.

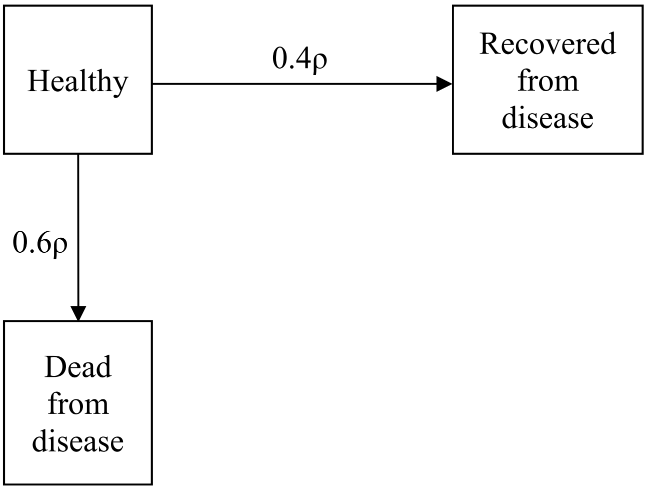
(v) We need enough information to calculate the number of transitions from Healthy to Sick and the waiting time in the "Healthy" state. Hence we will need:

Date of birth of all births in last three years.

Date of death of all "Healthy" deaths in last three years.
Date of any other immigration or emigration of Healthy people in last three years.

Date at which each person who contracted the disease fell ill.

Various alternative (usually simpler) models were suggested by some candidates. For example a much simpler model was proposed with only three states:



This was treated sympathetically for parts (i) to (iii) as the rate of first-time sickness can be derived from this model. However in part (iv) the probability of remaining alive for three years depends on the rate of death from causes other than the disease, even among islanders who have never suffered from the disease, so this needs to be introduced. Moreover, the rate of death from causes other than the disease may vary for persons who have never had the disease and persons who have recovered from the disease.

Few candidates attempted part (iv). Candidates were expected to calculate the probability of being alive in three years' time for a person who had never had the disease at the start of that three-year period, not the probability that a person who had not had the disease would be alive and still not have had the disease in three years' time. In part (iv) some candidates assumed constant transition intensities and evaluate the integral, which was given full credit.

In part (v) many candidates stated that what is required is the number of healthy persons at the beginning of the period (three years ago) and the current number of healthy persons, together with the number of persons who had fallen sick. This would allow an approximate sickness rate to be calculate using various assumptions and was given partial credit.

12 (i) Right censoring

The exact duration of the event is not known, but only that it exceeds some duration.

Example: job seekers with whom contact was lost during the investigation (or those still seeking jobs at the end of the investigation)

Random censoring

The time at which contact was lost may be regarded as a random variable.

Example: a job seeker with whom contact was lost during the investigation.

Type I censoring

The censoring times were known in advance (as they were determined by the fixed period of the investigation).

Example: a person still without work after 8 months.

Interval censoring

The censoring mechanism prevents us from knowing exactly when the event of interest took place, only that it fell within a certain period.

Example: EITHER a person who actually found a job after 5.5 months (say) is recorded as having found a job after 6 months;
OR a person who was still seeking work at the end of the investigation found a job within the interval $[8,\infty)$

Informative censoring

Censoring gives information about the lifetimes of those who remain (survival function for each censored observation for t greater than the time of censoring is the same as that for the non-censored observations).

Example: a person lost to the investigation will have a greater or lesser chance of finding a job than those who remained. [3]

(ii) The calculations are shown in the table below.

t_j	N_j	d_j	c_j	d_j/N_j	$1 - d_j/N_j$
0	700				
1	700	100	50	0.1429	0.8571
2	550	70	0	0.1273	0.8727
3	480	50	20	0.1042	0.8958
4	410	40	20	0.0976	0.9024
5	350	20	30	0.0571	0.9429
6	300	20	60	0.0667	0.9333
7	220	12	38	0.0545	0.9455
8	170	6	164	0.0353	0.9647

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

t	*Kaplan-Meier estimate of S(t)*
$0 \leq t < 1$	1.0000
$1 \leq t < 2$	0.8571
$2 \leq t < 3$	0.7480
$3 \leq t < 4$	0.6701
$4 \leq t < 5$	0.6047
$5 \leq t < 6$	0.5702
$6 \leq t < 7$	0.5321
$7 \leq t < 8$	0.5031
$t = 8$	0.4854

(iii) The null hypothesis is that the durations at which job seekers find work follow a Weibull distribution with parameters $\lambda = 0.18$ and $\beta = 0.3$.

Using the chi-squared test we have the following calculations:

t	*h(t)*	N_j	*expected*	*observed*	z_x	z_x^2
1	0.1794	700	125.55	100	−2.28	5.20
2	0.1104	550	60.72	70	1.19	1.42
3	0.0831	480	39.90	50	1.60	2.56
4	0.0680	410	27.86	40	2.30	5.29
5	0.0581	350	20.35	20	−0.08	0.01
6	0.0512	300	15.35	20	1.19	1.41
7	0.0459	220	10.11	12	0.60	0.36
8	0.0418	170	7.11	6	−0.42	0.17

The calculated value of the chi-squared statistic is 16.40.

This should be compared with the critical value at the 5% level
with 6 degrees of freedom (because we have eight ages and two parameters have been fitted, and 8 – 6 = 2)

which is 12.59.

Since 16.40 > 12.59

we reject the null hypothesis that the time to employment follows the Weibull distribution.

In part (i) credit was given for up to two different forms of censoring. For informative/non-informative censoring, a candidate could decide that censoring is EITHER informative OR non-informative and gain credit for a sensible explanation and example which are consistent with this decision. In part (ii) a common error was to suggest that the estimate of S(t) extended to values of t above 8. This is not the case as there is no information in the data about what might happen after 8 months. In part (ii) some candidates decided that censoring precedes the event. Provided they explained this, full credit was given.

A substantial proportion of candidates did not attempt part (iii). Of those who did, the approach given above was the most common. Little credit was given to candidates who tried to compare hazards or survival functions directly using the chi-squared test, as the assumptions of the test are not met. A common error was to use an exposed-to-risk of 700 to compute the expected deaths at all durations.

END OF EXAMINERS' REPORT