INSTITUTE AND FACULTY OF ACTUARIES

EXAMINATION

22 April 2015 (pm)

Subject CT4 – Models
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 12 questions, beginning your answer to each question on a new page.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is NOT required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator from the approved list.

1 For a simple random walk:

(i) Define the process. [2]

(ii) Write down the nature of the state space and time space in which it operates. [1]

(iii) Describe an example of a practical application of the process. [1]
[Total 4]

2 The mortality of a rare form of flying beetle is being studied. It has been discovered that beetles kept in a protected environment have a constant force of mortality, μ, but that those in the wild have a force of mortality which is 50% higher. It has been proven that the beetles revert immediately to the higher rate of mortality if they are released from the protected environment.

A beetle born and always living in the wild has a 58% chance of living for eight days.

Calculate the probability of living the same length of time for:

(a) a beetle born and reared in the protected environment.

(b) a beetle born in the protected environment which is scheduled to be released into the wild after six days.

[4]

3 (i) Explain what is meant by a proportional hazards model. [3]

(ii) Outline three reasons why the Cox proportional hazards model is widely used in empirical work. [3]
[Total 6]

4 (i) List the stages you would go through in creating a model. [4]

(ii) Discuss, for three of these stages, the specific issues that could arise when creating a model to price a new sickness benefit product. [3]
[Total 7]

5 (i) State the principle of correspondence as it applies to death rates. [1]

A nightclub opens at 10.00 p.m. and closes at 2.00 a.m. It admits only people aged over 21 years on the production of an identity card giving date of birth.

The table below shows the number of people entering in various intervals between 10.00 p.m. and 2.00 a.m. on 30 June 2013. No-one was admitted after 1.00 a.m., and you may assume that all those who enter the premises stay until 2.00 a.m.

Year of birth	*10.00–11.30 p.m.*	*11.30–12.00 p.m.*	*12.00 p.m.–1.00 a.m.*
1989	100	300	200
1990	200	400	350
1991	150	400	300
1992	100	250	200

During the period of opening, 40 people aged 22 last birthday required medical attention for heat exhaustion.

(ii) Calculate the rate per person-hour at which those attending the night club aged 22 last birthday required medical attention for heat exhaustion, stating any assumptions you make. [6]

[Total 7]

6 A health insurance company has collected data on sickness rates during the calendar year 2013 among a sample of its policyholders aged 40–64 years inclusive. It compares these to the rates among its policyholders of the same age in 2012. It finds that at ages 40–50 years inclusive, and at ages 56–61 years inclusive, the sickness rates in 2013 are higher than those in 2012. At other ages, the sickness rates in 2013 were lower than those in 2012.

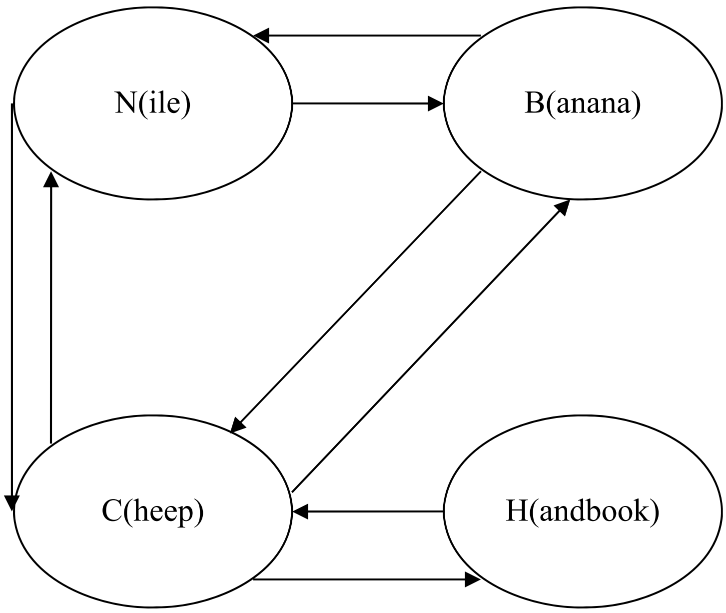
(i) Carry out two tests of the null hypothesis that the underlying sickness rates in 2013 are the same as those in 2012. [6]

(ii) Comment on the implications of the results of your test for the company's sickness insurance business. [2]

[Total 8]

7 (i) Describe what is meant by a Markov chain. [2]

A simplified model of the internet consists of the following websites with links between the websites as shown in the diagram below.



An internet user is assumed to browse by randomly clicking any of the links on the website he is on with equal probability.

(ii) Calculate the transition matrix for the Markov chain representing which website the internet user is on. [2]

(iii) Calculate, of the total number of visits, what proportion are made to each website in the long term. [4]

[Total 8]

8 (i) State why it is important to divide data into homogeneous classes when undertaking mortality investigations. [2]

(ii) List four factors, apart from smoking behaviour, by which mortality data are often classified by life insurance companies. [2]

In a particular life insurance market, it has for many years been the practice for all companies to charge smokers higher premiums than non-smokers for the same term assurance policy. Suppose one company decides to switch to charging smokers and non-smokers the same premiums for term assurance policies. The other companies retain differential pricing for smokers and non-smokers.

(iii) Discuss the likely implications for the company making the switch. [4]

[Total 8]

9 (i) Describe an example of a situation when graduation by parametric formula would be used. [1]

(ii) State two advantages and two disadvantages of graduation by parametric formula. [4]

(iii) (a) Explain why the χ^2 test is different when considering the goodness of fit of graduated data compared with when considering the similarity of two sets of data.

(b) Describe how this is dealt with when the graduation has been carried out by parametric formula. [4]

[Total 9]

10 In a computer game a player starts with three lives. Events in the game which cause the player to lose a life occur with a probability $\mu dt + o(dt)$ in a small time interval dt. However the player can also find extra lives. The probability of finding an extra life in a small time interval dt is $\lambda dt + o(dt)$. The game ends when a player runs out of lives.

(i) Outline the state space for the process which describes the number of lives a player has. [1]

(ii) Draw a transition graph for the process, including the relevant transition rates. [3]

(iii) Determine the generator matrix for the process. [2]

(iv) Explain what is meant by a Markov jump chain. [1]

(v) Determine the transition matrix for the jump chain associated with the process. [2]

(vi) Determine the probability that a game ends without the player finding an extra life. [1]

[Total 10]

11 A new disease has been discovered which is transmitted by an airborne virus. Anyone who contracts the disease suffers a high fever and then in 60% of cases dies within an hour and in 40% of cases recovers. Having suffered from the disease once, a person builds up antibodies to the disease and thereafter is immune.

(i) Draw a multiple state diagram illustrating the process, labelling the states and possible transitions between states. [2]

(ii) Express the likelihood of the process in terms of the transition intensities and other observable quantities, defining all the terms you use [4]

(iii) Derive the maximum likelihood estimator of the rate of first time sickness. [2]

Three years ago medical students visited the island where the disease was first discovered and found that of the population of 2,500 people, 860 had suffered from the disease but recovered. They asked the leaders of the island to keep records of the occurrence and the outcome of each incidence of the disease. The students intended to return exactly three years later to collect the information.

(iv) Derive an expression (in terms of the transition intensities) for the probability that an islander who has never suffered from the disease will still be alive in three years' time. [4]

(v) Set out the information which the students would need when they returned three years later in order to calculate the rate of sickness from the disease. [2]
[Total 14]

12 A study was made of a group of people seeking jobs. 700 people who were just starting to look for work were followed for a period of eight months in a series of interviews after exactly one month, two months, etc. If the job seeker found a job during a month, the job was assumed to have started at the end of the month. Unfortunately, the study was unable to maintain contact with all the job seekers.

The data from the study are shown in the table below:

Months since start of study	*Found employment*	*Contact lost*
1	100	50
2	70	0
3	50	20
4	40	20
5	20	30
6	20	60
7	12	38
8	6	0

(i) (a) Describe two types of censoring present in the investigation.

(b) Describe an example of a person to whom each type applies. [3]

(ii) Calculate the Kaplan-Meier estimate of the function for "remaining without employment". [6]

A Weibull distribution with a rate $h(t)$ given by the formula $h(t) = \lambda^{\beta}\beta\, t^{\beta-1}$ was fitted to these data. The estimated value of λ was 0.18 and the estimated value of β was 0.3.

(iii) Test the goodness-of-fit of the data to this Weibull distribution. [6]
[Total 15]

END OF PAPER