INSTITUTE AND FACULTY OF ACTUARIES

EXAMINATION

23 April 2013 (pm)

Subject CT6 – Statistical Methods
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 11 questions, beginning your answer to each question on a separate sheet.*
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 Use the $U(0,1)$ random numbers 0.238 and 0.655 to generate two observations from the Weibull distribution with parameters $c = 0.002$ and $\gamma = 1.1$. [4]

2 Claims on a certain type of insurance policy are believed to follow an exponential distribution. The upper quartile claim size is 240.

Calculate the mean claim size. [4]

3 An actuary has a tendency to be late for work. If he gets up late then he arrives at work X minutes late where X is exponentially distributed with mean 15. If he gets up on time then he arrives at work Y minutes late where Y is uniformly distributed on [0,25]. The office manager believes that the actuary gets up late one third of the time.

Calculate the posterior probability that the actuary did in fact get up late given that he arrives more than 20 minutes late at work. [5]

4 (i) Explain what is meant by a two player zero-sum game. [2]

Sally and Fiona agree to play a game. The rules of the game are as follows:

- Each player chooses either the number 10 or the number 40.
- Neither player knows the other player's choice before selecting her number.
- If both players choose the same number, Fiona pays Sally the sum of the numbers.
- If the players choose differently, Sally pays Fiona the sum of the numbers.

Sally decides to adopt a randomised strategy where she chooses 10 with probability p and 40 with probability $1-p$.

(ii) (a) Determine the value of p for which Sally's expected payoff is the same regardless of what Fiona chooses.

(b) Explain why this strategy is optimal for Sally.

(c) Calculate Sally's expected payout each time the game is played, assuming that she follows this strategy.

[4]
[Total 6]

5 The following table shows incremental claims data from a portfolio of insurance policies for the accident years 2010, 2011 and 2012. Claims from this type of policy are fully run off after the end of development year two.

Incremental Claims		*Development year* 0	1	2
Accident year	2010	2,328	1,484	384
	2011	1,749	1,188	
	2012	2,117		

Estimate the total claims outstanding using the basic chain ladder technique. [7]

6 Claim numbers on a portfolio of insurance policies follow a Poisson process with parameter λ. Individual claim amounts X follow a distribution with moments $m_i = E(X^i)$ for $i = 1, 2, 3, \ldots$. Let S denote the aggregate claims for the portfolio. You may assume that the mean of S is λm_1 and the variance of S is λm_2.

(i) Derive the third central moment of S and show that the coefficient of skewness of S is $\dfrac{\lambda m_3}{(\lambda m_2)^{3/2}}$. [4]

(ii) Show that S is positively skewed regardless of the distribution of X. [2]

(iii) Show that the distribution of S tends to symmetry as $\lambda \to \infty$. [2]

[Total 8]

7 An insurance company believes that individual claim amounts from house insurance policies follow a gamma distribution with distribution function given by:

$$f(y) = \frac{\alpha^{\alpha}}{\mu^{\alpha}\Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\alpha}{\mu}y} \text{ for } y > 0$$

where α and μ are positive parameters.

(i) Show that the gamma distribution can be written in exponential family form, giving the natural parameter and the canonical link function. [5]

The insurance company has data for claim amounts from previous claims. It believes that the claim amount is primarily influenced by two variables:

x_i the type of geographical area in which the house is situated. This can take one of 4 values.

y_i the category of the age of the house where the three categories are 0–29 years, 30-59 years and 60 years +.

It wishes to model claim amounts using this data and the generalised linear model from part (i) with canonical link function. The insurance company is investigating models which take into account these variables and has the following table of values:

Model	*Choice of predictor*	*Scaled Deviance*
A	1	900
B	Age	789
C	Age +location	544
D	Age * location	541

(ii) Explain, by analysing the scaled deviances, which model the insurance company should use. [6]
[Total 11]

8 An insurance company has a portfolio of 1,000 car insurance policies. Claims arise on individual policies according to a Poisson process with annual rate μ. The insurance company believes that μ follows a gamma distribution with parameters $\alpha = 2$ and $\lambda = 8$.

(i) (a) Show that the average annual number of claims per policy is 0.25.

(b) Show that the variance of the number of annual claims per policy is 0.28125.

[5]

Individual claim amounts follow a gamma distribution with density

$$f(x) = \frac{x}{1{,}000{,}000} e^{\frac{-x}{1000}} \quad \text{for } x > 0.$$

(ii) Calculate the mean and variance of the annual aggregate claims for the whole portfolio. [3]

The insurance company has agreed an aggregate excess of loss reinsurance contract with a retention of £0.55m (this means that the reinsurance company will pay the excess above £0.55m if the aggregate claims on the portfolio in a given year exceed £0.55m).

(iii) Calculate, using a Normal approximation, the probability of aggregate claims exceeding the retention in any year. [2]

For each of the last three years, the total claim amount has in fact exceeded the retention.

(iv) Comment on this outcome in light of the calculation in part (iii). [2]

[Total 12]

9 Claims on a portfolio of insurance policies arise as a Poisson process with rate λ. The mean claim amount is μ. The insurance company calculates premiums using a loading of θ and has an initial surplus of U.

(i) Explain how the parameters λ, μ, θ and U affect $\psi(U)$, the probability of ultimate ruin. [4]

Now suppose that $\lambda=50$, $\mu=200$ and $\theta=30\%$. There are three models under consideration for the distribution of individual claim amounts:

A fixed claims of 200
B exponential with mean 200
C gamma with mean 200 and variance 800

Let the corresponding adjustment coefficients be R_A, R_B and R_C.

(ii) Find the numerical value of R_B and show that R_B is less than both R_A and R_C. [7]

(iii) Use the fact that $\left(1+\dfrac{x}{n}\right)^n \approx e^x$ for large n to show that R_A and R_C are approximately equal. [2]
[Total 13]

10 An insurance company has a portfolio of building insurance policies. The company classifies buildings into three types and believes that the number of claims on buildings of each type follows a Poisson distribution with parameters as shown:

Type	*Parameter*
1	λ
2	2λ
3	5λ

where λ is an unknown positive constant.

Actual claim numbers over the last five years have been as follows. Here X_{ij} represents the number of claims from the *i*th type in the *j*th year:

Number of claims X_{ij}

Year (j)

Type(i)	*5*	*4*	*3*	*2*	*1*	$\sum_{j=1}^{5}(X_{ij}-\bar{X}_i)^2$
1	23	17	9	21	12	139.2
2	56	39	44	29	35	417.2
3	87	115	141	92	84	2322.8

(i) Derive the maximum likelihood estimate of λ. [5]

(ii) Estimate the average number of claims per year for each type of building using EBCT Model 1. [7]

(iii) Comment on the results of parts (i) and (ii). [2]

(iv) Explain the main weakness of the model in part (ii). [1]
[Total 15]

11 An actuary is considering the time series model defined by

$$X_t = \alpha X_{t-1} + e_t$$

where e_t is a sequence of independent Normally distributed random variables with mean 0 variance σ^2. The series begins with the fixed value $X_0 = 0$.

(i) Show that the conditional distribution of X_t given X_{t-1} is Normal and hence show that the likelihood of making observations $x_1, x_2, \ldots, x_n$ from this model is:

$$L \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \alpha x_{i-1})^2}{2\sigma^2}}.$$ [3]

(ii) Show that the maximum likelihood estimate of α can also be regarded as a least squares estimate. [2]

(iii) Find the maximum likelihood estimates of α and σ^2. [4]

(iv) Derive the Yule-Walker equations for the model and hence derive estimates of α and σ^2 based on observed values of the autocovariance function. [5]

(v) Comment on the difference between the estimates of α in parts (iii) and (iv). [1]

[Total 15]

END OF PAPER