INSTITUTE AND FACULTY OF ACTUARIES

EXAMINATION

28 April 2014 (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 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 (i) List six of the characteristics that insurable risks usually have. [3]

(ii) List two key characteristics of a short term insurance contract. [1]
[Total 4]

2 Ruth takes the bus to school every morning. The bus company's ticket machine is unreliable and the amount Ruth is charged every morning can be regarded as a random variable with mean 2 and non-zero standard deviation. The bus company does offer a "value ticket" which gives a 50% discount in return for a weekly payment of 5 in advance. There are 5 days in a week and Ruth walks home each day.

Ruth's mother is worried about Ruth not having enough money to pay for her ticket and is considering three approaches to paying for bus fares:

A Give Ruth 10 at the start of each week.

B Give Ruth 2 at the start of each day.

C Buy the 50% discount card at the start of the week and then give Ruth 1 at the start of each day.

Determine the approach that will give the lowest probability of Ruth running out of money. [4]

3 The table below shows the payoff to a player from a decision problem with three uncertain states of nature θ_1, θ_2 and θ_3 and four possible decisions D_1, D_2, D_3 and D_4.

	D_1	D_2	D_3	D_4
θ_1	10	3	−7	9
θ_2	−5	12	6	−7
θ_3	−8	−3	13	−10

(i) Determine whether any of the decisions are dominated. [2]

(ii) Determine the optimal decision using the minimax criteria. [2]

Now suppose $P(\theta_1) = 0.5$ and $P(\theta_2) = 0.3$ and $P(\theta_3) = 0.2$.

(iii) Determine the optimal decision under the Bayes criterion. [2]
[Total 6]

4 Individual claim amounts on a portfolio of motor insurance policies follow a Gamma distribution with parameters α and λ. It is known that $\lambda = 3$ for all drivers, but the parameter α varies across the population. 70% of drivers have $\alpha = 300$ and the remaining 30% have $\alpha = 600$.

Claims on the portfolio follow a Poisson process with annual rate 500 and the likelihood of a claim arising is independent of the parameter α.

Calculate the mean and variance of aggregate annual claims on the portfolio. [6]

5 A particular portfolio of insurance policies gives rise to aggregate claims which follow a Poisson process with parameter $\lambda = 25$. The distribution of individual claim amounts is as follows:

Claim	50	100	200
Probability	30%	50%	20%

The insurer initially has a surplus of 240. Premiums are paid annually in advance.

Calculate approximately the smallest premium loading such that the probability of ruin in the first year is less than 10%. [7]

6 Claim amounts arising from a certain type of insurance policy are believed to follow a Lognormal distribution. One thousand claims are observed and the following summary statistics are prepared:

mean claim amount	230
standard deviation	110
lower quartile	80
upper quartile	510

(i) Fit a Lognormal distribution to these claims using:

(a) the method of moments.
(b) the method of percentiles.

[6]

(ii) Compare the fitted distributions from part (i). [2]
[Total 8]

7 The heights of adult males in a certain population are Normally distributed with unknown mean μ and standard deviation $\sigma = 15$.

Prior beliefs about μ are described by a Normal distribution with mean 187 and standard deviation 10.

(i) Calculate the prior probability that μ is greater than 180. [2]

A sample of 80 men is taken and the mean height is found to be 182.

(ii) Calculate the posterior probability that μ is greater than 180. [4]

(iii) Comment on your results from parts (i) and (ii). [2]
[Total 8]

8 (i) (a) Write down the Box-Muller algorithm for generating samples from a standard Normal distribution.

(b) Give an advantage and a disadvantage of the Box-Muller algorithm relative to the Polar method.
[3]

(ii) Extend the algorithm in part (i) to generate samples from a Lognormal distribution with parameters μ and σ^2. [1]

A portfolio of insurance policies contains n independent policies. The probability of a claim on a policy in a given year is p and the probability of more than one claim is zero. Claim amounts follow a Lognormal distribution with parameters μ and σ^2. The insurance company is interested in estimating the probability θ that aggregate claims exceed a certain fixed level M.

(iii) Construct an algorithm to simulate aggregate annual claims from this portfolio. [2]

The insurance company estimates that θ is around 10%.

(iv) Calculate the smallest number of simulations the insurance company should undertake to be able to estimate θ to within 1% with 95% confidence. [2]

The insurance company is considering the impact on θ of entering into a reinsurance arrangement.

(v) Explain whether the insurance company should use the same pseudo random numbers when simulating the impact of reinsurance. [1]
[Total 9]

9 The table below sets out incremental claims data for a portfolio of insurance policies.

Accident year	*Development year* 0	1	2
2011	1,403	535	142
2012	1,718	811	
2013	1,912		

Past and projected future inflation is given by the following index (measured to the mid point of the relevant year).

Year	*Index*
2011	100
2012	107
2013	110
2014	113
2015	117

Estimate the outstanding claims using the inflation adjusted chain ladder technique. [9]

10 For a certain portfolio of insurance policies the number of claims on the i^{th} policy in the j^{th} year of cover is denoted by Y_{ij}. The distribution of Y_{ij} is given by

$$P(Y_{ij} = y) = \theta_{ij}(1 - \theta_{ij})^y \qquad y = 0, 1, 2, \ldots$$

where $0 \le \theta_{ij} \le 1$ are unknown parameters with $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, l$.

(i) Derive the maximum likelihood estimate of θ_{ij} given the single observed data point y_{ij}. [4]

(ii) Write $P(Y_{ij} = y)$ in exponential family form and specify the parameters. [4]

(iii) Describe the different characteristics of Pearson and deviance residuals. [2]
[Total 10]

11 Let θ denote the proportion of insurance policies in a certain portfolio on which a claim is made. Prior beliefs about θ are described by a Beta distribution with parameters α and β.

Underwriters are able to estimate the mean μ and variance σ^2 of θ.

(i) Express α and β in terms of μ and σ. [3]

A random sample of n policies is taken and it is observed that claims had arisen on d of them.

(ii) (a) Determine the posterior distribution of θ.

(b) Show that the mean of the posterior distribution can be written in the form of a credibility estimate. [5]

(iii) Show that the credibility factor increases as σ increases. [3]

(iv) Comment on the result in part (iii). [1]

[Total 12]

12 A sequence of 100 observations was made from a time series and the following values of the sample auto-covariance function (SACF) were observed:

Lag	*SACF*
1	0.68
2	0.55
3	0.30
4	0.06

The sample mean and variance of the same observations are 1.35 and 0.9 respectively.

(i) Calculate the first two values of the partial correlation function $\hat{\phi}_1$ and $\hat{\phi}_2$. [1]

(ii) Estimate the parameters (including σ^2) of the following models which are to be fitted to the observed data and can be assumed to be stationary.

(a) $Y_t = a_0 + a_1 Y_{t-1} + e_t$

(b) $Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + e_t$

In each case e_t is a white noise process with variance σ^2. [12]

(iii) Explain whether the assumption of stationarity is necessary for the estimation for each of the models in part (ii). [2]

(iv) Explain whether each of the models in part (ii) satisfies the Markov property. [2]

[Total 17]

END OF PAPER