INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

11th May 2010

Subject CT6 – Statistical Methods

Time allowed: Three Hours (10.00 – 13.00 Hrs)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. ***Please read the instructions on the front page of answer booklet and instructions to examinees sent along with hall ticket carefully and follow without exception***

2. ***Mark allocations are shown in brackets.***

3. ***Attempt all questions, beginning your answer to each question on a separate sheet. However, answers to objective type questions could be written on the same sheet.***

4. ***In addition to this paper you will be provided with graph paper, if required.***

AT THE END OF THE EXAMINATION

Please return your answer book and this question paper to the supervisor separately.

Q. 1) The claims paid to on a motor insurance policy are as follows (Figures in Rs. '000s).

Year of Accident	Claim payment in the year of development			
	1	2	3	4
2006	141	94	54	17
2007	137	97	58	
2008	139	101		
2009	143			

Number of claims in the same period are as follows.

Year of Accident	No. of claims settled in the year of development			
	1	2	3	4
2006	240	139	67	11
2007	248	145	71	
2008	251	141		
2009	246			

It is assumed that the accident year 2006 is fully developed.

Inflation for the 12 months to the middle of each year was as follows:

Year	2007	2008	2009
Inflation rate %	6.2	7	7.6

Beyond the year 2009, the rate of inflation is assumed to be uniform at 7% per annum. Calculate the outstanding claim reserve as at 31st December 2009 using each of the following methods, stating the assumptions implied in each of the methods:

(i) Inflation adjusted chain-ladder method, and (8)

(ii) Inflation adjusted average cost per claim method. (9)

[17]

Q. 2) **(i)** What are the disadvantages of using truly random numbers, as opposed to pseudo-random numbers, in Monte Carlo simulations? (3)

(ii) It is necessary to generate a set of simulated samples from the density function

$$g(x) = \frac{ke^{-3x}}{(1+x)^2}, \quad x > 0,$$

where k is an appropriate constant.

a) Give a procedure that applies the Acceptance-Rejection method to draw samples from the density g, while making sure that you

1. describe the steps clearly and specifically for the given density, and
2. clarify how you would generate samples from any other distribution that may be needed for this purpose. (4)

b) Derive an expression, in terms of k, for the expected number of pseudo-random variables required to generate a single observation from the density g using this method. (2)

[9]

Q. 3) **(i)** In the context of generalized linear models, explain what you understand by the following terms:

a) Covariate,

b) Linear predictor, and

c) Link function (3)

(ii) Consider a Generalised Linear Model (GLM) with independent Poisson responses with structure

$g(\mu) = \eta$

comprising a linear predictor, η, linked to the mean response, μ, through a link function, g.

a) Prove that the model belongs to an exponential family, and identify the different aspects of the exponential family for this model. (3)

b) Identify the canonical link function. (1)

c) Find maximum likelihood estimates of the model parameters from the following data on the response (y) and a single covariate (x). (5)

Serial no. (i)	covariate (x_i)	response (y_i)
1	2	3
2	4	3
3	6	9

[12]

Q. 4) The Marketing Director of an insurance company has the option for one day of promoting Term Plan (d_1), Unit-linked Plan (d_2) or Whole Life Endowment Plan (d_3) at a road show. He believes that the prospective customer is equally likely to be single (θ_1), married with at most one child (θ_2) or married with children (θ_3) and estimates his profits under each customer profile to be as follows.

	θ_1	θ_2	θ_3
d_1	25	19	7
d_2	10	30	8
d_3	0	2	34

The Marketing Director has to decide which plan to promote at the road show, with a view to maximizing profit.

(i) Determine the minimax solution to this problem. (2)

(ii) The Sales Director is very optimistic and believes that the criterion to adopt in deciding which product to sell should be to maximize the maximum profit. What decision would the Sales Director make based on these predicted Profits? (1)

(iii) Determine the Bayes criterion solution to this problem. (2)

(iv) The Sales Director thinks that there is more than two-thirds chance that the target customer is married with at most one child (θ_2). Show that, according to this assumption, the Bayes criterion solution is d_2. (4)

[9]

Q. 5) Consider the AR(2) process

$$Y_t = -\alpha Y_{t-1} + \alpha^2 Y_{t-2} + Z_t,$$

where $\{Z_t\}$ is a zero-mean white noise process with $\text{Var}(Z_t) = \sigma^2$.

(i) Determine the range of values of α for which the above process can be stationary. (3)

(ii) Derive the auto-covariances γ_1 and γ_2 of Y in terms of α and σ. (6)

[9]

Q. 6) An insurer currently has an individual excess of loss reinsurance arrangement with retention limit of Rs. 1,000. The total expected claims this year from the portfolio before reinsurance are Rs. 1,000,000. Individual claim sizes are independent and these have a lognormal distribution with parameters $\mu = 4$ and $\sigma^2 = 4$. The number of claims is independent of the claim sizes, and it has an unspecified distribution. The premium for the reinsurance is 115% of the expected losses covered by the reinsurance.

(i) Calculate the expected number of claims for this year. (2)

(ii) Calculate the reinsurance premium for this year. (3)

(iii) It is known that individual claim sizes will increase next year by 10% due to inflation, but the distribution of the number of claims will remain the same. Will the reinsurance premium increase by 10%, more than 10%, or less than 10%? Explain. (4)

[9]

Q. 7) The claim sizes for two groups of drivers covered in a portfolio have the exponential distribution with mean $1/\lambda$ and $2/\lambda$, respectively. The following claim sizes are observed in a particular month in respect of the two groups of drivers.

Group 1 : 250, 100, 90, 410.

Group 2 : 70, 430, 320, 640.

It is presumed that the prior distribution of the parameter λ is exponential with mean 0.01. Derive the posterior distribution of λ, and calculate the posterior mean.

[5]

Q. 8) The number of claims arising from a particular portfolio in a year has the binomial distribution with parameters 100 and p. The prior distribution of the parameter p is the beta distribution, having density $\frac{(k+l-1)!}{(k-1)!(l-1)!}p^{k-1}(1-p)^{l-1}$, $0 \le p \le 1$. In the last year, N claims were observed.

(i) Determine the posterior distribution of p, given N. (2)

(ii) Calculate the prior mean of p. (2)

(iii) Calculate the posterior mean of p, given N, show that it has the form of a credibility estimate, and identify the credibility factor. (3)

[7]

Q. 9) The observed values of aggregate claims from a portfolio over 5 years are $X_1, X_2, \ldots, X_5$. An actuary wishes to calculate the premium for this risk using these data, as well as collateral data on 3 similar risks over 8 years, denoted by

$Y_{11}, Y_{21}, \ldots, Y_{81},$ $\quad Y_{12}, Y_{22}, \ldots, Y_{82},$ $\quad Y_{13}, Y_{23}, \ldots, Y_{83}.$

The Empirical Bayes Credibility Theory (EBCT) Model 1, with underlying parameter θ, is to be used to calculate the credibility premium.

(i) State whether the following statements in respect of EBCT model 1 are correct.

(a) The observations $X_1, X_2, \ldots, X_5$, given θ, are assumed to be independent and identically distributed.

(b) In order to calculate the credibility premium, one needs to know how $E\left[X_1 | \theta\right]$ depends on θ.

(c) In order to calculate the credibility premium, one needs to know the prior distribution of θ.

(d) The parameter θ is assumed to take the same value in respect of the data sets $\{Y_{11}, Y_{21}, \ldots, Y_{81}\}$, $\{Y_{12}, Y_{22}, \ldots, Y_{82}\}$, and $\{Y_{13}, Y_{23}, \ldots, Y_{83}\}$. (4)

(ii) Give a specific expression for the credibility premium under EBCT Model 1, in terms of the above data only. (The answer should not involve any quantity that has not been specified in the question.) (4)

[8]

Q. 10) The aggregate claims process for a particular risk is a compound Poisson process with $\lambda = 20$. Individual claim amounts are 100 with probability ¼, 200 with probability ½, or 250 with probability ¼. The initial surplus is 1,000. Using a Normal approximation, calculate the smallest premium loading factor θ such that the probability of ruin at time 3 is at most 0.05. **[7]**

Q. 11) The number of claims occurring in a year has the Poisson distribution with mean λ, while the claims are independent and have the exponential distribution with mean θ. There is an excess of loss reinsurance arrangement, with retention limit 1000. The claim amounts paid by the reinsurer in a particular year are observed to be 1500, 200, 600 and 500. The reinsurer wishes to estimate the parameters λ and θ from these data.

(i) Obtain the distribution of the claim amount paid by the reinsurer (i.e., claim amount in excess of the retention limit, disregarding any claim that is below the retention limit). (1)

(ii) Obtain the distribution of the claim frequency experienced by the reinsurer. (2)

(iii) Obtain the maximum likelihood estimate of λ and θ from the above data. (2)

(iv) Obtain the maximum likelihood estimate of the mean aggregate claims paid together by the direct insurer and the reinsurer in a year. (1)

[6]

Q. 12) Give two examples of non-linear time series models, while clearly specifying their functional forms.

[2]

****