INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

20th November 2012

Subject CT6 – Statistical Models

Time allowed: Three Hours (10.00 – 13.00)

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.***

AT THE END OF THE EXAMINATION

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

Q. 1) The tables below show the cumulative cost of incurred claims in thousands of rupees and the number of claims reported each year for a small cohort of general insurance policies. The claims are assumed to be fully run-off at the end of development year 2.

The cumulative costs (in thousands of rupees) of claims incurred are as follows.

Accident Year	**Development Year**		
	0	**1**	**2**
2009	320	628	895
2010	490	980	
2011	750		

The number of claims reported in each year are as follows.

Accident Year	**Development Year**		
	0	**1**	**2**
2009	105	90	45
2010	152	128	
2011	285		

Given that total amount paid towards claims to date, relating to accident years 2009, 2010 and 2011 is Rs. 2,500,000, calculate the outstanding claims reserve using the average cost per claim method, on the basis of simple average form of grossing up factors.

[9]

Q. 2) The number of claims on a portfolio of insurance policies follows a Poisson distribution with parameter 25. Individual claims may be regarded as realizations of a random variable $200X$, where X has the distribution with probability density function

$$f(x) = \begin{cases} \left(\frac{2}{25}\right)(5-x), & 0 < x < 5, \\ 0, & \text{otherwise.} \end{cases}$$

In addition, for each claim, there is a 25% chance that an additional fixed expense of 500 will be incurred.

[i] Calculate the mean and variance of the total individual claim amounts. (7)

[ii] Calculate the mean and variance of the aggregate claims on the portfolio. (2)

[9]

Q. 3) The aggregate claim process from a portfolio of 1000 one year health insurance policies, each with monthly premium of Rs. 200 (payable in advance), is a compound Poison process. The expected number of claims per month is 1. The policy covers three types of health risk events viz. Minor, moderate and major, for which the claim amounts are Rs. 100000, Rs. 250000 and Rs 500000 respectively. The probabilities that a claim is for a Minor, Moderate and Major risk event are 70%, 25% and 5%, respectively. Claims are settled at the end of each month. The insurer holds initial surplus of Rs. 90000.

Calculate the probability of ruin at the end of the first month. **[8]**

Q. 4) A health insurance company has option of three products to sell in the market and must decide which product to sell in the coming year. There are three possible choices: basic, lean or rich product, each with different related costs based on the complexity of the product. The manufacturer has fixed overheads of Rs. 1,500,000.

The revenue and cost for each product are as follows.

Policy	Cost	Revenue per policy
Basic	500,000	1500
Lean	300,000	1000
Rich	1,000,000	2000

The insurer had sold 2,100 policies last year, and is preparing forecasts of profitability for the coming year based on three scenarios: Low sales (80% of last year's level), Medium sales (same as last year's level) and High sales (20% higher than last year's level).

[i] Determine the annual profit in rupees for each product under each scenario. (3)

[ii] Determine the minimax choice of product. (2)

[iii] Determine the Bayes solution, if there is 20% chance of Low sales, 60% chance of Medium sales and 20% chance of High sales. (2)

[7]

Q. 5) For a general insurance policy, the settled amount of a particular claim (denoted by X) has uniform distribution over the interval $(0, \theta)$, where θ is the maximum claim size permissible for that risk. An analyst, who has knowledge of X but does not know θ, wants to guess the value of θ from a single observed value of X. The analyst prefers to use the absolute error loss function for this purpose.

[i] If the prior distribution of θ is $f(\theta) = \theta e^{-\theta}$ for $\theta > 0$, derive the Bayes estimator of θ with respect to the absolute error loss function. (5)

[ii] Suppose no prior is used, and that the analyst uses the estimator kX, where k is an appropriate constant. Determine the value of k he should use so that the estimator has the smallest mean absolute error for any fixed value of θ. (3)

[iii] Compare the two quantities that are minimized in order to produce the estimators of parts (i) and (ii). (2)

[10]

Q. 6) The yearly claim amounts arising from n independent motor insurance policies are $Y_1, Y_2, \ldots, Y_n$. The ages of the insured vehicles are $X_1, X_2, \ldots, X_n$. The i^{th} claim amount, Y_n, is modeled as a random variable, whose mean μ_i depends on X_i. Exploratory analysis indicates that the claim amounts may be assumed to have the exponential distribution, and the mean could depend on the age through the model $\mu_i = e^{a+bX_i}$.

[i] Identify the link function corresponding to the suggested model. Is it the canonical link function for the exponential distribution? (2)

[ii] Use the suggested model to derive equations that are to be solved in order to obtain the maximum likelihood estimators of a and b. (4)

[iii] Define scaled deviance, and derive an expression for it in terms of the fitted values of $\mu_1, \mu_2, \dots, \mu_n$. (4)

[10]

Q. 7) **[i]** Explain the concept of cointegrated time series. (2)

[ii] Give two examples of circumstances when it is reasonable to expect that two processes may be cointegrated. (2)

[iii] Define a weakly stationary time series. (1)

[iv] Let Y_n be a time series, which satisfies the relation

$$Y_n = 0.65Y_{n-1} + Z_n + 0.35Z_{n-1},$$

where Z_n is a sequence of independent zero-mean variables with common variance σ^2. Derive the autocorrelation sequence of Y_n for all lags. (5)

[v] Let X_n be a time series, which satisfies the relation

$$X_n = 1.2X_{n-1} + 0.7X_{n-2} - 0.1X_{n-3} + Z_n,$$

where Z_n is a sequence of independent zero-mean variables with common variance σ^2.

a. Explain why this series does not have the Markov property. (1)

b. Describe how one can construct a vector-valued process, using lagged samples of this series, which would have the Markov property. (1)

c. Justify the construction. (3)

[15]

Q. 8) The table below shows aggregate annual claim statistics for four different products over a period of five years. Annual aggregate claims for product *i* in year *j* are denoted by X_{ij}.

Product (i)	$\bar{X}_i = \frac{1}{5}\sum_{j=1}^{5} X_{ij}$	$\frac{1}{4}\sum_{j=1}^{5}\left(X_{ij} - \bar{X}_i\right)^2$
1	125	300
2	85	60
3	140	35
4	175	100

[i] Calculate the credibility premium of each product under the assumptions of EBCT Model 1. (7)

[ii] Explain why the credibility factor is relatively high in this case. (2)

[9]

Q. 9) You have to generate pseudo-random samples from a gamma distribution with probability density function $f(x)$ given by

$$f(x) = \frac{x^{1.5}e^{-x}}{\Gamma(2.5)}, \qquad x > 0,$$

by using the acceptance/rejection method. As for the density to draw samples from, there are three candidates:

$h_1(x) = e^{-x}$, $h_2(x) = \frac{1}{2}e^{-x/2}$ and $h_3(x) = \frac{1}{3}e^{-x/3}$.

Explain which one would be most appropriate, and why.

[6]

Q. 10) State six conditions that are either necessary or desirable for a risk to be insurable.

[3]

Q. 11) A reinsurer believes that claims from a catastrophic event insurance policy follow a Pareto distribution with parameters $\alpha = 3$ and $\lambda = 500$. The reinsurer wishes to draft an excess of loss treaty such that 50% of the losses result in no claim on the reinsurer.

[i] Calculate the size of the deductible. (2)

[ii] Calculate the average claim amount net of the deductible, in respect of those losses that result in some amount of claim on the reinsurer. (4)

[iii] In case a direct insurer agrees to the above reinsurance arrangement, calculate the average claim amount for that insurer. (4)

[iv] The reinsurer wishes not to use the assumed values of the parameters of the Pareto distribution, and estimate them from real data observed during a year of the above reinsurance arrangement. Let $x_1, x_2, \ldots, x_n$ be the claims recorded by the reinsurer. Write down an expression for the likelihood. (2)

[v] The direct insurer also wishes not to use the assumed values of the parameters of the Pareto distribution, and estimate them from real data observed during the same period. His claims records includes claim amount $y_1, y_2, \ldots, y_m$, and the fact that a total of n claims have involved the reinsurer. The insurer has no record of the actual values of these n claims. Write down an expression for the likelihood on the basis of the direct insurer's data. (2)

[14]

****