INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

22nd May 2009

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. <u>However, answers to objective type questions could be written on the same sheet</u>.*

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) Consider the following zero-sum game, where the numbers represent the amount of loss of Player A.

		Player A	
		Strategy I	Strategy II
Player B	Strategy 1	2	-1
	Strategy 2	-2	3

(i) Does the game have a saddle point? Explain. (2)

(ii) Determine the *pure* strategies that are minimax for Player A and Player B, respectively. (2)

(iii) Suppose that Players A and B play this game on six consecutive days. On day 1, each player chooses the pure strategy that is minimax for him. On each following day, each player assumes that his opponent would repeat the strategy used in the previous day, and accordingly adjusts his own strategy to minimize his loss. Indicate the value of the game on these six days. (3)

(iv) If the game continues over the subsequent days in the manner described above, will the value of the game converge? (1)

(v) Explain the advantages of a randomized strategy in this game (2)

[10]

Q 2) The number of accidents in a year in a city has the Poisson distribution with parameter λ. The observed number of accidents in that city in five consecutive years happen to be 120, 135, 104, 128 and 146, respectively. On the other hand, the average annual number of accidents in comparable cities is known to have the gamma distribution with mean 110 and variance 1100.

(i) Calculate the shape and scale parameters of the gamma distribution. (2)

In the following questions, use the above gamma distribution as the prior distribution for λ.

(ii) Calculate the marginal (not conditional on λ) distribution of the annual number of accidents in a *single year*, and show that it is a negative Binomial distribution. Identify the parameters of this distribution. (4)

(iii) Obtain the posterior distribution of λ, given the five observations. (3)

(iv) Obtain the posterior mean of λ. and show that it can be written in the form of a credibility estimate. What is the credibility factor? (3)

[12]

Q 3) **(i)** State the two main types of Proportional reinsurance arrangements. How do they differ from each other? (3)

(ii) The aggregate claims process for a risk is a compound Poisson process. The expected number of claims each year is 3.65 and individual claim amounts have an exponential distribution with mean 400.

(a.) If the insurer collects risk premium continuously using a loading factor of 0.25, and starts with an initial asset of Rs. 0, show that the probability of ruin within ten days is less than or equal to

$$P(\text{Size of one claim is more than premium accrued in 10 days}) \times P(\text{One claim arises in 10 days}) + P(\text{More than one claim arises in 10 days}),$$

and evaluate this worst-case probability. (5)

(b.) The insurer decides to use a special type of reinsurance for this risk. For an individual claim of amount X, the reinsurer will pay an amount Z and the insurer will pay an amount X – Z, where

$$Z = \begin{cases} 0 & \text{if } X \le 200, \\ X - 200 & \text{if } 200 < X < 600, \\ 600 & \text{if } X \ge 600. \end{cases}$$

The reinsurer calculates premiums using a premium loading factor of 0.4. Calculate the total amount of premium that the reinsurer charges to reinsure this risk for one year. (4)

[12]

Q 4) **(i)** Explain what is meant by a 'delay triangle'. (2)

(ii) The following table gives the cumulative incurred claims data by year of accident and reporting development for a portfolio of motor insurance policies.

Cumulative incurred claim amounts (Rs '000s)

Accident year	Development year			
	0	1	2	3
1997	280	375	431	446
1998	255	343	398	
1999	248	323		
2000	260			

The corresponding cumulative number of reported claims by years of accident and reporting development are as follows:

Cumulative number of claims

Accident year	Development year			
	0	1	2	3
1997	56	74	87	91
1998	49	65	77	
1999	44	61		
2000	50			

Use an average cost per claim method to calculate an estimate of the outstanding claim amount for these policies for claims arising during these accident years. The claims paid to date are Rs 1,323,000. (8)

[10]

Q 5) **(i)** Show that if X has a lognormal distribution with parameters μ and σ^2, then $P(X > M) = 1 - \Phi((\log M - \mu)/\sigma)$ (1)

(ii) A no claims discount (NCD) system has three states 0%, 20% and 50%. The full premium for a policyholder in the system is Rs 800. Any policyholder making a claim in any year moves to the next lower discount level (or stays on 0%). Any policyholder who does not make a claim during a year moves in the next year to the next higher discount level (or remains at 50%). The probability that any policyholder in the group has an accident in any year is 0.2. However, the policyholders are reluctant to lose their discounts on the full premium and so will make an actual claim if the amount of damage incurred is greater than the difference between:

- the amount of the next two premiums assuming a claim is made,
- the amount of the next two premiums assuming no claim is made.

In each case, the policyholder assumes that no further claims will be made during the subsequent two-year period. Losses from accidents occurring to all policyholders have a lognormal distribution with $\mu = 4$ and $\sigma^2 = 4$.

(a) Find the minimum amount of accident damage for which a policyholder at each of the three levels will deem it worthwhile to make a claim. (3)

(b) Hence, find the transition probability matrix for this NCD system. (3)

(c) Determine the steady state proportions of policyholders in the three discount states, if this NCD system is allowed to run for a long time. (4)

[11]

Q 6) Each year an insurer issues a number of household insurance policies. The premium for each of these policies is Rs 125. The number of claims from each policy has a Poisson distribution with parameters 0.2 and individual claim amounts have a Lognormal distribution with parameters $\mu = 5$ and $\sigma^2 = 2$. Assuming that the distribution of the aggregate claim amount is approximately Normal and that individual claim amounts are independent, determine the minimum number of policies that must be sold each year if the employee wishes to be 99% sure of making a profit on the portfolio. Ignore expenses, interest and tax.

[7]

Q 7) Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent random variables such that

$$Y_i = \begin{cases} 0 & \text{with probability } 1 - p_i, \\ 1 & \text{with probability } p_i, \end{cases}$$

for $i = 1,2,\ldots,n$.

(i) By writing $P(Y_i = y_i) = p_i^{y_i}(1 - p_i)^{1-y_i}$, where $y_i = 0$ or 1, show that this probability function is a member of the exponential family. Identify the canonical link function. (3)

(ii) Verify that $E(Y_i) = b'(\theta_i)$ and $V(Y_i) = a(\phi_i)b''(\theta_i)$, where the functions a and b and the parameters θ_i and ϕ_i correspond to the usual notations of Generalized Linear Model. (2)

(iii) Write down the likelihood function for the parameters $p_1, p_2, \ldots, p_n$ based on observations $y_1, y_2, \ldots, y_n$. (1)

(iv) You have data of the form $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, where y_i indicates the cardiovascular health status of the i^{th} policyholder (with the value 0 for normal health and 1 for heart disease), and x_i indicates the age of that policyholder. Using the canonical link function to express p_i in terms of x_i, write down the likelihood function for the generalized linear model parameters. (3)

[9]

Q 8) Explain the terms

(i) Cointegrated time series, (2)

(ii) ARCH model. (2)

[4]

Q 9) **(i)** The following time series model is used for the monthly inflation rate $\{Y_t\}$ in India, based on wholesale price of select commodities.

$$Y_t = 0.4Y_{t-1} + 0.2Y_{t-2} + Z_t + 0.025Z_{t-1} + 0.016 .$$

Here, $\{Z_t\}$ is a sequence of uncorrelated random variables having a common variance and zero mean.

(a) Obtain the values of p, d and q for this ARIMA(p, d, q) model. (2)

(b) Assuming infinite history, calculate the expected value of the inflation rate according to this model. (1)

(c) Determine whether $\{Y_t\}$ is a stationary process. (2)

(ii) It was thought that the monthly inflation rate $\{X_t\}$ in India, based on retail price of select commodities would follow the model

$$X_t = 0.4X_{t-1} + 0.2X_{t-2} + Z_t + 0.025,$$

where $\{Z_t\}$ is a sequence of uncorrelated random variables having a common variance and zero mean.

(a) Determine the autocorrelation function of $\{X_t\}$. (5)

(b) Describe briefly two diagnostic checks that can be performed to determine whether there is any inadequacy in the fitting of such ARMA models to time series data. (2)

[12]

Q 10) **(i)** Explain the disadvantage of using random numbers, as opposed to pseudo-random numbers, for simulation. (2)

(ii) What is a linear congruential generator? (2)

(iii) State how you would use the polar method to generate pairs of uncorrelated pseudo-random numbers from the standard normal distribution. (2)

(iv) Describe how you would simulate pseudo-random samples from the following distributions.

(a) The mixture of two uniform distributions, having density

$$f(x) = \begin{cases} 0.2 & \text{if } 0 < x < 1, \\ 0.8 & \text{if } 2 < x < 3. \end{cases}$$

(3)

(b) The folded normal distribution, having density

$$f(x) = \begin{cases} \dfrac{2}{\sqrt{2\pi}} e^{-x^2/2} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

(2)

(c) The censored exponential distribution, having distribution

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - e^{-x} & \text{if } 0 \leq x < 2, \\ 1 & \text{if } x \geq 2. \end{cases}$$

(2)

[13]

****