Actuarial Society of India

EXAMINATIONS

18th November 2005

Subject CT6 – Statistical Models

Time allowed: Three Hours (10.30 – 13.30 pm)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. ***Do not write your name anywhere on the answer scripts. You have only to write your Candidate's Number on each answer script.***

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. ***Fasten your answer sheets together in numerical order of questions. This, you may complete immediately after expiry of the examination time.***

5. ***In addition to this paper you should have available graph paper, Actuarial Tables and an electronic calculator.***

Professional Conduct:

"It is brought to your notice that in accordance with provisions contained in the Professional Conduct Standards, If any candidate is found copying or involved in any other form of malpractice, during or in connection with the examination, Disciplinary action will be taken against the candidate which may include expulsion or suspension from the membership of ASI."

Candidates are advised that a reasonable standard of handwriting legibility is expected by the examiners and that candidates may be penalized if undue effort is required by the examiners to interpret scripts.

AT THE END OF THE EXAMINATION

Hand in both your answer scripts and this question paper to the supervisor.

Q.1)

General insurance business is typically considered to be short-term business. Describe what is meant by short-tailed and long-tailed business within the context of general insurance and give an example of each type of business.

[2]

Q.2)

If Z_1, Z_2…is an *I(1)* time series, derive an appropriate operation to perform on the time series X_1, X_2…defined by

$$X_t = \exp(a + bt + Z_t)$$

so that a stationary series may be obtained.

[2]

Q.3)

Claims arrive in a Poisson process at rate λ, and *N(t)* is the number of claims arriving by time *t*. The claim sizes are independent random variables $X_1, X_2 ...$, with mean μ, independent of the arrivals process. The initial surplus is *u* and the premium loading factor is θ.

(i) Give an expression for the surplus process *U(t)* at time *t*.

(1)

(ii) Define the probability of ruin with initial surplus u, $\psi(u)$. State the value of $\psi(u)$ when $\theta = 0$.

(1)

(iii) Your colleague comments, "As the value of λ increases, the probability of ruin must also increase." Comment on your colleague's statement.

(2)

(iv) Describe the simplifications adopted in the surplus process as given in (i) above.

(2)

Total [6]

Q.4)

You are given a stream of independent uniform [0,1] random variables U_1, $U_{2,}$, U_3…Describe how to use these random variables to generate random variates with the following distributions:

(i) The Weibull distribution with probability density function

$$f(x) = \begin{cases} 30x^2 \exp(-10x^3) & if \quad x > 0 \\ 0 & otherwise \end{cases}$$

(2)

(ii) The continuous distribution with probability density function

$$f(x) = \begin{cases} \frac{2}{p} \sin^2 x & if \quad 0 < x < p \\ 0 & otherwise \end{cases}$$

(3)

(iii) The empirical distribution function obtained from the data set $\{X_1, X_2 ....X_n\}$, (that is, a discrete distribution with equal probability of occurrence of the values $X_1, X_2 \ldots X_n$).

(2)

Total [7]

Q.5)

(i) Write down the general form of a statistical model for a claims run-off triangle, defining all terms used.

(3)

(ii) You are an actuarial student working for *InsureSecure*, a leading general insurance company. The table below shows the incurred claims on a portfolio of home contents insurance policies of *InsureSecure*.

Accident Year	***Delay Year***		
	0	*1*	*2*
2002	4,253	955	235
2003	3,142	1,945	
2004	4,002		

The ultimate loss ratio assumption is 90% and *InsureSecure* uses the Bornheutter-Ferguson method to calculate reserves. Calculate the reserves for accident year 2004 if the earned premium is 4,500 and the paid claims are 1,885.

(5)

Total [8]

Q.6)

The no claims discount system operated by an insurance company for their annual motor insurance business has four levels of discount:

- Level 1: 0%
- Level 2: 25%
- Level 3: 45%
- Level 4: 60%

If a policyholder does not make a claim under the policy in a particular year then he or she will go up one level (or stay at level 4), whereas if any claims are made he or she will go down by two levels (or remain at, or move to, level 1). The full premium payable at the 0% discount level is 900.

The probability of an accident occurring is assumed to be 0.2 each year for all policyholders and losses are assumed to follow a lognormal distribution with mean 1,188 and standard deviation 495. However, policyholders claim only if the loss is greater than the total additional premiums that would have to be paid over the next three years, assuming that no further accidents occur.

(i) Calculate the smallest loss amount for which a claim will be made for a policyholder at the 0% discount level.

(1)

(ii) Complete the transition matrix below by calculating the missing values.

$$\begin{pmatrix} * & * & * & * \\ 0.147 & 0 & 0.853 & 0 \\ 0.120 & 0 & 0 & 0.880 \\ 0 & 0.197 & 0 & 0.803 \end{pmatrix}$$

(5)

(iii) Calculate the proportion of policyholders at each discount level when the system reaches a stable state.

(3)
Total [9]

Q.7)

Let X be a sample from the uniform distribution with probability density function

$$f(x) = \begin{cases} \dfrac{1}{q} & if \quad 0 < x < q \\ 0 & otherwise \end{cases}$$

The prior distribution of *è* is exponential with probability density function

$$f(x) = \begin{cases} qe^{-q} & q > 0 \\ 0 & otherwise \end{cases}$$

(i) What is the Bayes estimator of *è* with respect to the squared error loss function, based on the single observation X?

(4)

(ii) What is the Bayes estimator of *è* with respect to the absolute error loss function, based on the single observation X?

(2)

(iii) Derive an expression for the Bayes estimator of *è* with respect to the squared error loss function, based on two independent samples X_1 and X_2 from the stated uniform distribution.

(3)
Total [9]

Q.8)

Consider the probability density function of the gamma distribution
$f(y) = y^{a-1} b^{-a} e^{-y/b} / \Gamma(a)$ for $y>0$ where the parameters *á* and *â* are both positive.

(i) Show that the gamma distribution, with suitably redefined parameters belongs to the exponential family with density of the type

$$\exp\left[\frac{yq - b(q)}{a(f)} + c(y, f)\right].$$

Identify the natural parameter *è* and the functions *b*(*è*) and $a(f)$. **(3)**

(ii) Using the functions *b*(*è*) and $a(f)$, derive the mean and variance for the gamma distribution in terms of the original parameters *á* and *â*.

(2)

(iii) Size of claims arising from an automobile liability insurance for a given customer profile is sought to be modelled by the gamma distribution. The profile consists of the age of the policyholder (continuous variable x_1), gender of the policyholder (binary variable x_2), age of the car (continuous variable x_3) and the price of the car when purchased (continuous variable x_4).

Describe precisely a generalized linear regression model which can be used in this context, and explain how one can test whether claim size depends on the gender of the policyholder.

(4)
Total [9]

Q.9)

An insurance company has a portfolio of policies with a per-risk excess of loss reinsurance arrangement with a retention level of *M* (>0). Claims made to the direct insurer, denoted by *X*, have a Pareto distribution with cumulative distribution function:

$$F(x, a) = 1 - \left(\frac{100}{100 + x}\right)^{a}$$

There were a total of *n* claims from the portfolio. Of these, *h* claims were for amounts less than *M*. The claims less than *M* are $\{x_i : i = 1,\ 2 \ldots h\}$

The value of the statistic $\sum_{i=1}^{h} \log(100 + x_i) = y$ is given.

(i) Derive the maximum likelihood estimate of a in terms of *n*, *h*, *M* and *y*.

(5)

(ii) Assuming that α = 1.5 and M = 500, estimate the average amounts paid by the insurer and the reinsurer on a claim made during the year.

(5)
Total [10]

Q.10)

Y_t, $t = 1, 2, \ldots$ is a time series defined by

$$Y_t \quad \acute{a} Y_{t-1} = Z_t + (1 \quad \acute{a}) Z_{t-1},\ |\acute{a}| < 1,$$

where Z_t, $t = 1, 2, \ldots$, is a sequence of independent zero-mean variables with common variance $ó^2$.

(i) State, giving reasons, the values of p, d and q for which Y_1, Y_2, . . .is an ARIMA(p, d, q) process. **(2)**

(ii) Derive the autocorrelation function $\tilde{n}_k$, k = 0, 1, 2, . . . of the time series, after it is suitably differenced to achieve stationarity. **(6)**

(iii) Suppose that you have observed Y_1, Y_2 . . . , Y_{100}. Suggest an estimator of *á*. (Hint: You may use the result of part (ii) and the sample autocorrelation of lag 1, assuming that the latter is smaller than 1 in magnitude.)

(2)
Total [10]

Q.11)

(i) Claims occur on a portfolio of general insurance policies independently and at random. Claims are classified as "Type A claims" or "Type B claims". It is estimated that 25% of all claims are type B claims. Type A claims are distributed according to a Pareto distribution with parameters α = 3.5 and λ = 200. Type B claims are distributed according to a Pareto distribution with parameters α = 4 and λ = 1,200. Let X denote a claim chosen at random from this portfolio.

(a) Calculate $P[X > 2000]$

(3)

(b) Calculate $E[X]$ and $Var[X]$.

(3)

(ii) The random variable Y has a Pareto distribution with the same mean and variance as X in part (i).
Calculate $P[Y > 2000]$.

(3)

(iii) Comment on the differences in your answers to parts (i) and (ii)

(2)
Total [11]

Q.12)

An insurance company has arranged a policy with a large organisation to cover a particular type of disastrous event. Under the terms of the policy the insurer will cover the total cost of the claim for up to three disasters in any year. The insurer is considering taking out reinsurance for the policy whereby the reinsurer will cover the excess, if any, of the total cost over Rs. 2 million per disaster. The reinsurer has quoted the premiums for two different types of cover as follows:

	Cover provided by reinsurer	*Premium*
D_1	Claims over Rs. 2 million on the first disaster Only	Rs. 500,000
D2	Claims over Rs. 2 million on each of the first two disasters	Rs. 1,000,000

The insurer needs to decide whether to opt for cover levels D_1, D_2 or not to reinsure (D_0).
The insurer models the total loss from a single disaster as having a Pareto distribution with probability density function . $\frac{a\, l^{a}}{(l + x)^{a+1}}$ with parameters á = 3.125 and ë = Rs. 3,187,500.

(i) Determine the expected total loss per individual disaster, with and without reinsurance. **(5)**

(ii) Complete the decision matrix below based on total outgoings (you may ignore expenses):

	No disaster è$_0$	*1 disaster* è$_1$	*2 disaster* è$_2$	*3 or more disasters* è$_3$
D_0				
D_1				
D_2				

(6)

(iii) (a) Determine the minimax solution. **(1)**
(b) Explain your answer using general reasoning. **(1)**

(iv) The insurer believes that the number of disasters each year follows a Poisson distribution with parameter 0.9. Determine the Bayes Criterion solution.

(4)
Total [17]

****