Actuarial Society of India

EXAMINATIONS

19^{th} May 2006

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.

Q1) A zero sum game involving two players A and B requires choosing numbers x and y over the set {1,2,3} by A and B, respectively. B's gain resulting from the choice (x,y) is 6 min(x,y)/max(x,y)

(i) Draw the pay-off matrix (2)

(ii) Determine the optimal choices for A and B under the maximini and minimax criteria (4)

[6]

Q2) An insurer arranges excess of loss reinsurance with retention limit of Rs. 10,000. The reinsurer suspects that the original claims (including the claims settled entirely by the direct insurer) are independent and have a Pareto distribution. Recent claim amounts paid by the reinsurer in respect of this risk happen to be Rs. 4,253, Rs. 22,320, Rs. 9,724, Rs. 3,692 and Rs. 85,035. The reinsurer wants to estimate the proportion of claims that are settled directly by the insurer. Use the method of moments to obtain an estimate of this proportion. **[5]**

Q3) The annual number of medical claims of a policyholder having annual income q has the Poisson distribution with mean lq where l is a fixed parameter. The annual income q has the exponential distribution with mean m.

(i) Compute the unconditional variance of the annual number of claims of a policy holder and show that it is larger than the mean number of claims (2)

(ii) Is the conditional variance of the annual number of claims for a policy holder of *given annual income* larger than the corresponding conditional mean? Compare with the result of part (i) and comment. (2)

(iii) If the unconditional variance of the annual number of claims is 20, and average annual income of a policy holder is Rs. 16,000/- what must be the value of l ? (1)

(iv) Assume that claims sizes are independent of one another and of the number of claims, and has the exponential distribution with mean d . Give expressions for the unconditional mean and variance of the total annual claims size of a policy holder, in terms of l , m and d . (2)

(v) Assume that, *given annual income* q , the claims sizes are independent of one another and of the number of claims, that the claim size distribution is exponential with mean aq and the annual number of claims distribution is Poisson with mean lq . The annual income q has the exponential distribution with mean m . Calculate the mean and the variance of the total claim size. (3)

You may use the fact, $\mathrm{var}(Y) = E(\mathrm{var}(Y \mid X)) + \mathrm{var}(E(Y \mid X))$. **[10]**

Q4) An insurer provides disability coverage for a group of 100 employees for a year. During this period, there can be at most one claim from each employee. The claims occur independently of one another with probability 0.05. The claim size distribution is log-normal with parameters m =10 and s =0.2. The event of ruin is relevant for the end of the year only. Determine the premium loading needed to ensure that the probability of ruin is at most 0.05. You can use normal approximation for the total claim amount. **[5]**

Q5) Number of claims $N_1, N_2, \ldots\ldots, N_n$ for a group insurance in successive years is assumed to follow the Poisson distribution. There had been a change in government policy in the *m*th year. It is postulated that the number of claims in the first *m* years have a different mean than the number of claims in the *n-m* subsequent years. In order to examine this postulate, define the covariate.

$$x_i = \begin{cases} 0 & if\ 1 \le i \le m, \\ 1 & if\ m < i \le n. \end{cases}$$

There is no other covariate

(i) Let *g* be an *unspecified* link function which is invertible. Using the generalized linear model for the claim numbers $N_1, N_2, \ldots\ldots, N_n$, the covariates $x_1, x_2, \ldots\ldots\ldots, x_n$, and this link function, give an expression for the log-likelihood for the model parameters (4)

(ii) Derive explicit expressions for the maximum likelihood estimators of the model parameters (4)

(iii) Show that the fitted value of the mean number of claims of the *i*th year does not depend on the function *g*. (2)

(iv) Can the choice of *g* be irrelevant in the sense of part (iii) when x_i has *three* possible levels? (2)

(v) Which choice of *g* would make it the canonical link function? (1)

(vi) Calculate the scaled deviance for this model. (2)

(vii) Show that the scaled deviance for the model under the constraint “Covariate has no effect” is

$$2\sum_{i=1}^{n}[N_i \log N_i - N_i - \log(N_i!)] - 2\sum_{i=1}^{n}[N_i \log(\sum_{i=1}^{n} N_i / n) - \sum_{i=1}^{n} N_i / n - \log(N_i!)]$$

(2)

(viii) Explain how you will test the hypothesis that the mean number of claims did not change after the first *m* years. (3)

[20]

Q6) Consider the time series model

$$Y_t = Y_{t-1} + 0.5Y_{t-2} - 0.5Y_{t-3} + Z_t + 0.3Z_{t-1},$$

where $\{Z_t\}$ is a sequence of uncorrelated random variables each having the normal distribution with mean zero and variance s^2.

(i) Show that the above model is a special case of the ARIMA (*p*,*d*,*q*) model, and determine *p*, *d* and *q*. (2)

(ii) Let $X = (1-B)^d Y$. Determine whether $\{X_t\}$ is a stationary time series. (2)

(iii) Calculate the autocorrelation function of $\{X_t\}$. (6)

[10]

Q7) The number of employees of an organization who die in a year has the Binomial (n, q) distribution, where n is the number of employees alive at the beginning of that year. The prior distribution of q is Beta with mean 0.015 and standard deviation 0.005. At the start of 2005, the organization had 250 employees in its payroll. Given that 5 employees died in 2005, calculate

(i) the posterior distribution of q, (6)

(ii) the Bayes estimator of q under the quadratic loss function, (2)

(iii) the Bayes estimator of q under the all-or-nothing loss function. (2)

[10]

Q8) Given a parameter m, losses arising from a risk, $X_1, X_2, \ldots\ldots,$ are independent and have the lognormal distribution with parameters m and s^2, where s^2 is known. From collateral data, it is surmised that the prior distribution of the parameter m is normal mean q and variance t^2, where q and t are known numbers. Let $a = e^{m+\frac{s^2}{2}}$, the conditional mean of X_1 given m.

(i) Using the prior distribution of m, show that the prior mean a is $e^{q+\frac{s^2}{2}+\frac{t^2}{2}}$. (3)

(ii) Using the posterior distribution of m, show that posterior mean of a is (2)

$$\exp\left(\frac{n\sum_{i=1}^{n} \log X_i / s^2 + q/t^2}{n/s^2 + 1/t^2} + \frac{s^2}{2} + \frac{1}{2(n/s^2 + 1/t^2)} \right)$$

[5]

Q9) The no claims discount system operated by a motor insurer has seven classes. The class at the start of a year depends only on the class at the start of the previous year and the number of claims in the previous year . The rules for moving between classes are given by the following table.

Class	Premium (% of full premium)	Class in the following year after			
		0	1	2	≥ 3
		Claims in current year			
7	100	6	7	7	7
6	75	5	7	7	7
5	65	4	6	7	7
4	55	3	5	7	7
3	45	2	5	7	7
2	40	1	4	6	7
1	33	1	4	6	7

On 1st January 2004 a group of 10,000 policyholders entered the system in class 6. For each policyholder the number of claims in a year has a Poisson distribution with mean 0.2. Calculate the expected number of policyholders in each class on 1st January 2006, assuming no policyholder leaves and no more join the group. **[8]**

Q10) The following table of cumulative claims incurred is derived from a small property account. It gives the amounts for each of the last nine accident years, and the related earned premium. The cumulative paid claims for all accident years since, and including 2001 is Rs. 20,485.

Accident year	Development year						Ultimate	Earned Premium
	0	1	2	3	4	5		
1997	2323	2713	2902	3009	3081	3065	3065	3606
1998	2489	2907	3109	3224	3301	3287	3287	3864
1999	2709	3165	3385	3509	3393	3572	3572	4206
2000	2966	3464	3705	3842	3934	3914	3914	4604
2001	3512	4042	4205	4394	4458			5305
2002	4054	4610	4938	5101				5896
2003	4614	5421	5690					6578
2004	5354	6180						7546
2005	5700							8304

Estimate the reserve required at the end of 2005 using the Bornhuetter-Ferguson method, stating the assumptions you make. **[12]**

Q11) If $X_1, X_2, \ldots\ldots,$ are independent random variables having the exponential distribution with mean m, t is a fixed time and N is the number such that

$$\sum_{i=1}^{N} X_i \le t < \sum_{i=1}^{N+1} X_i,$$

then show that N has the Poisson distribution with mean t/m. Hence, suggest a way of generating samples from the Poisson distribution with specified mean l. **[6]**

Q12) Consider the stationary autoregressive process of order 1 given by

$Y_t = aY_{t-1} + Z_t,\qquad |a|<1$

Where $\{Z_t\}$ denotes white noise with mean zero and variance s^2.

Express Y_t in the form of $Y_t = \sum_{j=0}^{\infty} a_j Z_{t-j}$ and hence or otherwise, find an expression for $V(Y_t)$, the process variance, in terms of a and s. **[3]**

****