

Actuarial Society of India

EXAMINATIONS

17th June 2005

Subject CT6 – Statistical Models

Time allowed: Three Hours (10.30 am – 13.30 pm)

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.*
- 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."

AT THE END OF THE EXAMINATION

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

Q.1 An insurance company operates a No Claims Discount system for its motor insurance business, with discount levels 0%, 15%, 30% and 50%. The full annual premium is Rs100. The rules for moving between discount levels are:

- If no claims are made during a year, the policyholder moves to the next higher level of discount or remains at the maximum discount level.
- If one or more claims are made during a year, a policyholder at the 30% or 50% discount level moves to the 15% discount level and a policyholder at the 15% or 0% discount level moves to, or remains at, the 0% discount level.

When an accident occurs, the distribution of the loss is exponential with mean Rs500. In the event of an accident, a policyholder will claim only if the loss is greater than the total extra premiums that would have to be paid over the next three years, assuming that no further accidents occur.

For each discount level, calculate:

- i) the smallest loss for which a policyholder will make a claim. [2]
- ii) the probability of a claim being made in the event of an accident occurring. [3]

[Total 5]

Q.2 The total amount claimed for a particular risk in a portfolio is observed for each of 3 consecutive years. From past knowledge of similar portfolios, an insurer believes that the claims are normally distributed with mean q and variance 16, and that the prior distribution of the mean is normal with mean 100 and variance 49.

- i) Derive the Bayesian estimate under quadratic loss, and show that it can be written in the form of a credibility estimate combining the mean observed claim size for this risk with the prior mean for q . [4]
- ii) State the credibility factor, and calculate the credibility premium if the mean claim size over the 3 years is 110. [1]
- iii) Comment on how the credibility factor and the credibility estimate change if the variance of 16 is decreased. [1]

Total [6]

Q.3 A portfolio of crop insurance has claims distributed as Pareto with mean 500 and standard deviation of 1500. In the following Kharif season, 200 claims are expected, with claims being assumed to be occurring according to a Poisson process. To limit its losses, the insurer decides to introduce a policy excess of Rs100. Calculate the percentage reduction in the mean of aggregate claims to the insurer following the introduction of the policy excess. [5]

Q.4 The table below shows the cumulative claims paid (before adjustment for inflation) from a portfolio of insurance policies:

Accident year	Development year			
	0	1	2	3
2001	2,047	2,862	3,217	3,485
2002	2,471	3,728	3,918	
2003	2,388	3,826		
2004	2,580			

It may be assumed that payments are made in the middle of a calendar year. Inflation rate on these policies has been estimated to be 10% per annum over the relevant period.

- i) What are the assumptions for the inflation adjusted chain ladder method? [1]
 ii) Use the inflation adjusted chain ladder method to estimate the total outstanding payments, up to the end of development year 3, for accident year 2004 in mid-2004 prices, ignoring future inflation. [7]
Total [8]

Q.5 The number of claims, N , arising from a particular group of policies has a negative binomial distribution with parameters $k=3$ and $p=0.9$. Individual claim amounts, X , have the following distribution:

$$P(X = 500) = 0.5$$

$$P(X = 1,000) = 0.25$$

$$P(X = 2,000) = 0.25$$

- i) State the assumptions underlying the individual risk model for modelling the aggregate claim amount for a portfolio. [2]
 ii) Show that the relationship

$$P(N = n) = \left(a + \frac{b}{n} \right) P(N = n - 1)$$

- holds for $n = 1, 2, \dots$ and determine the values of a and b . [2]
 iii) Calculate the probability of the aggregate claim being at most 2,000 using the recursive method. [6]

Total [10]

Q.6 An insurer writes policies with individual excesses of Rs.1,000 per claim. The insurer has taken out a reinsurance policy whereby the insurer pays out a maximum of Rs. 20,000 in respect of each individual claim, the rest being paid by the reinsurer. The individual claims, gross of reinsurance and the excess, are believed to follow an exponential distribution with parameter \ddot{e} . Over the last year, the insurer has gathered the following data:

- There were 8 claims which were not processed by the insurer because the loss was less than the excess.
- There were 13 claims where the insurer paid out Rs. 20,000 for each claim and the reinsurer paid the remainder.
- There were 19 other claims in respect of which the insurer paid out a total of Rs. 66,666.

Derive the loglikelihood function of \ddot{e} . [6]

Q.7 Claims occur on a portfolio of insurance policies according to a Poisson process with rate 0.4 per year. The insurer's initial surplus is Rs. 10,000. The claim size is equal to Rs. 5,000 or Rs. 15,000 with equal probability and the different claim sizes are independent. Premium is paid continuously with a loading factor 0.25. Calculate

- i) The probability that ruin occurs at first claim. [4]
- ii) The probability that surplus is negative immediately after second claim. [4]
- iii) The probability that ruin does not occur in the first year. [4]

[You can use the fact that inter-occurrence times of a Poisson process with rate \ddot{e} are independent and have the exponential distribution with mean $1/\ddot{e}$.]

Total [12]

Q.8

A company is considering setting up a new type of insurance. All risks will be at one of four possible levels of intensity, I_1, I_2, I_3 and I_4 . The company has to decide at what level to set the premium, which will attract differing amounts of business as follows:

<i>Annual premium</i>	Rs. 850	Rs. 810	Rs. 790
<i>Number of policies ('000s)</i>	100	150	200

The company has annual fixed costs of Rs. 1.5 crores plus annual per policy expenses of Rs. 180. Under each level of intensity the company expects to pay out an average in claims per policy of :

<i>Level of intensity</i>	I_1	I_2	I_3	I_4
<i>Average cost (Rs.)</i>	Rs. 400	Rs. 450	Rs. 570	Rs. 600

- i) Determine the minimax solution based on annual profits. [7]
- ii) Given the probabilities 0.1, 0.4, 0.3 and 0.2 of I_1, I_2, I_3 and I_4 , respectively, determine the Bayes criterion solution based on annual profits. [2]

Total [9]

Q.9

A claim amount distribution is normal with unknown mean m and known standard deviation Rs. 100. Based on past experience a suitable prior distribution for m is normal with mean Rs. 500 and standard deviation Rs. 20.

- i) Calculate the prior probability that m , the mean of the claim amount distribution, exceeds Rs. 535. [1]
- ii) A random sample of 10 current claims have a mean of Rs. 535. Determine the posterior distribution of m . [2]
- iii) Calculate the posterior probability that m exceeds Rs. 535 and comment briefly on your answer. [3]

Total [6]

Q.10

There are m male drivers in each of three age groups, and data on the number of claims made during the last year are available. Assume that the numbers of claims are independent Poisson random variables. If Y_{ij} is the number of claims for the j th male driver in group i ($i = 1, 2, 3; j = 1, \dots, m$), let $E(Y_{ij}) = m_j$ and suppose $\log(m_j) = a_i$.

- i) Show that this is a generalized linear model, identifying the link function and the linear predictor. [3]
- ii) Determine the log-likelihood, and the maximum-likelihood estimators of a_1, a_2 and a_3 . [4]

For a particular data set with 20 observations in each group, several models are fitted, with deviances as shown:

		Deviance
Model 1	$\log(\mathbf{m}_j) = \mathbf{a}_i$	60.40
Model 2	$\log(\mathbf{m}_j) = \mathbf{a}$ if $i = 1, 2$ $= \mathbf{b}$ if $i = 3$	61.64
Model 3	$\log(\mathbf{m}_j) = \mathbf{a}$	72.53

- iii) Determine whether or not model 2 is a significant improvement over model 3, and whether or not model 1 is a significant improvement over model 2. [5]
 - iv) Interpret these three models. [3]
- Total [15]**

Q.11 Consider the second-order autoregressive process

$$Y_t = -2\mathbf{a}Y_{t-1} + \mathbf{a}^2Y_{t-2} + Z_t$$

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

- i) Determine the range of values of \mathbf{a} for which the process Y can be stationary. [3]
- ii) Derive the autocovariances \mathbf{g}_1 and \mathbf{g}_2 of Y in terms of \mathbf{a} and \mathbf{s}^2 . [6]

Total [9]

Q.12 A client wishes to model the behaviour of a stochastic process $\{X_t : t = 0\}$ which represents the average annual return for a particular class of asset. After a number of observations the client has determined that $\text{Corr}(X_t, X_{t-1}) = 0.7$ and $\text{Corr}(X_t, X_{t-2}) = 0.5$. He thinks that one of the two models

I: $X_t = \mathbf{m} + 0.7(X_{t-1} - \mathbf{m}) + 0.5(X_{t-2} - \mathbf{m}) + e_t$

II: $X_t = \mathbf{m} + e_t + 0.7e_{t-1} + 0.5e_{t-2}$

will be best, but cannot decide which. He has simulated both processes from time $t = 1$ to time $t = 200$, but has not obtained the results he expected, so is seeking your advice.

- i) (a) Outline a suitable method of simulating a second-order autoregression, assuming you have access to a reliable stream $\{u_k : k = 0\}$ of pseudo-random numbers uniformly distributed over the range $[0, 1]$.
(b) Explain why it might be desirable to ensure that the stream $\{u_k\}$ can be re-used if necessary. [4]
- ii) State why neither of the suggested models is suitable. [1]
- iii) (a) Derive the lag-1 and lag-2 autocorrelations, \mathbf{r}_1 and \mathbf{r}_2 , of a second order autoregressive process $X_t = \mathbf{m} + \mathbf{a}_1(X_{t-1} - \mathbf{m}) + \mathbf{a}_2(X_{t-2} - \mathbf{m}) + e_t$
(b) Find values of the parameters \mathbf{a}_1 and \mathbf{a}_2 which would provide a suitable AR(2) model for $\{X_t : t = 0, 1, 2, \dots\}$. [4]

Total [9]
