INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

22nd May 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.***
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) Based on past experience, a football team classifies prospective players as good (decision d_1), medium (decision d_2) and bad (decision d_3), at the beginning of a season. Pre-season offers are made by using this classification. At the end of season, the actual performance of the players can be good (outcome θ_1), medium (outcome θ_2) and bad (outcome θ_3).

The losses associated with the decisions and the outcomes are given below.

	d_1	d_2	d_3
θ_1	0	12	16
θ_2	6	0	10
θ_3	18	20	0

i) Determine the minimax solution when classifying a player at the beginning of a season. (1)

ii) The performance of a player in the past season has been medium. It is known that for players of this category, 25% turn out to be good performers at the end of the new season and 30% of them turn out to be bad performers. What would be the Bayes decision for this player? (4)

[5]

Q. 2) An insurance company uses the basic chain ladder method to compute the reserves for their pets insurance business. Some of the data are lost due to system error. The following is the table of the cumulative claim amounts, with the lost values represented by English letters.

Accident Year	**Payment Year**			
	2007	**2008**	**2009**	**2010**
2007	200	248	287	*A*
2008		215	*C*	290
2009			234	260

Also, some of the projected ultimate claim amounts are available, as follows.

Accident Year	**Ultimate claims**
2007	*B*
2008	380
2009	405
2010	450

All the claims are paid by the end of the development year 3.

i) Calculate the values of *A*, *B* and *C*. (5)

ii) For the accident year 2010, the earned premium is Rs. 400 and the loss ratio is 95%. By using the Bornhuetter-Ferguson method, calculate the claims paid in the year 2010 itself. (3)

iii) For a particular accident year, the company has loss ratio 125%, whereas the usual loss ratio for this business is 95%. List four possible reasons for the discrepancy. (2)

[10]

Q. 3) **i)** Give four different purposes, for which one carries out time series analysis. (2)

ii) State whether each of the following processes is Markov or not.

a. An MA(1) process;
b. An AR(1) process;
c. An AR(2) process;
d. A random walk. (2)

iii) Suppose that $Y(t)$ is a stationary time series and $X(t) = a + bt + Y(t)$, where a and b are fixed constants. Indicate, with reasons, whether

a. $X(t)$ is stationary,
b. $X(t)$ is $I(d)$ for any value of d. (4)

iv) Let $X_1(t)$ and $X_2(t)$ be time series, which are $I(1)$, and satisfy the relations

$$X_1(t) = aX_1(t-1) + bX_2(t-1) + e_1(t),$$
$$X_2(t) = bX_1(t-1) + aX_2(t-1) + e_2(t),$$

where a and b be non-zero constants, and $e_1(t)$ and $e_2(t)$ are white noise processes. Can the two time series be said to be cointegrated? Explain. (3)

[11]

Q. 4) An insurance company has two portfolios of policies, motor and household insurance, on each of which claims occur according to a Poisson process. For the household portfolio, all claims are for a fixed amount of Rs. 10,000 and 20 claims are expected per annum. For the motor portfolio, claim amounts are exponentially distributed with mean Rs. 5,000 and 30 claims are expected per annum.

The insurer has a loading of 15% for both the portfolios. Assume that the two portfolios are independent and the event of ruin is checked only at the end of the financial year. Also assume that normal approximation can be used for the annual aggregate claims.

i) Calculate the mean and variance of the annual aggregate claims S. (2)

ii) Calculate the initial capital required so that the probability of ruin at the end of year 1 is 0.05. (4)

iii) The insurer is planning to purchase a proportional reinsurance, with 80% retention by the insurer. The reinsurance premium has a loading of ε. Show that, for any specified probability of ruin, the initial capital required with reinsurance (u') and the initial capital required without reinsurance (u) are related by the equation

$$u' = 0.8u + (0.2\varepsilon - 0.83)E(S).$$
(4)

iv) By setting the target 0.05 for the probability of ruin (with or without reinsurance), determine the range of loadings for the reinsurance premium, so that reinsurance would reduce the requirement of initial capital. (2)

[12]

Q. 5) Identify any three features of the individual risk model, which separate it from the collective risk model.

[3]

Q. 6) An insurer is interested in estimating the probability (θ) of a particular policy resulting in a legal dispute. It is presumed that the number of disputes, arising out of a total of *n* policies in force, has the binomial distribution with parameters *n* and θ.

The experience of disputes over the past five years has been as under.

Serial no. (i)	***Year***	***Number of policies in force (n_i)***	***Number of disputes (x_i)***
1	2011-12	1043	98
2	2010-11	1057	107
3	2009-10	1005	97
4	2008-09	1021	96
5	2007-08	991	96

i) Write down the likelihood of θ computed from the five years of data. (1)

ii) It appears from collateral data that θ has the beta distribution with parameters $\alpha = 10$ and $\beta = 20$. Treating this distribution as the prior for θ, and the expression obtained in part (i) as the likelihood, determine the posterior distribution. (1)

iii) Can the prior said to be conjugate? (1)

iv) Give an explicit expression for the Bayes estimator of θ with respect to the zero/one loss function, and evaluate it for the given data. (2)

v) Can the estimator obtained in part (iv) be expressed as a credibility estimate? Explain, and give the credibility factor, if any. (2)

[7]

Q. 7) You have to generate pseudo-random samples from a distribution with probability density function $f(x)$ given by

$$f(x) = \begin{cases} kx^{-\frac{1}{2}} \ , & \text{if } 0 < x \leq 1, \\ ke^{-x} \ , & \text{if } x > 1, \end{cases}$$

where *k* is an appropriate constant. You have to make use of a pseudo-random number generator from the uniform distribution over [0, 1].

i) Determine the value of *k* so that $f(x)$ is a valid probability density function. (1)

ii) Describe *specific* steps to generate pseudo-random samples from $f(x)$, by using the inverse transform method. (5)

iii) If 0.777, 0.203, 0.905 and 0.999 are four samples from the uniform distribution over [0,1], determine the corresponding samples from $f(x)$, as per the above method. (2)

[8]

Q. 8) An insurer studies aggregate annual claims from its households, shopkeeper and auto insurance portfolios. The table below shows the sample average and sample standard deviation of each risk over the past five years (in Rs.'000s).

	Household	*Shopkeeper*	*Auto*
Sample mean over last five years	587	743	929
Sample standard deviation over last five years	163	211	231

Calculate the credibility premium for each risk using EBCT Model 1.

[6]

Q. 9) A motor insurance company models the number of claims (N), filed by a customer of age x over a five year period, as a random variable having the Poisson distribution.

i) Show that the Poisson distribution with mean parameter μ can be written in the form of an exponential family. (2)

ii) By using properties of an exponential family, find the mean and the variance of the Poisson distribution. (1)

iii) The insurer believes that the mean μ depends on the age of the customer through a linear predictor and the canonical link function. Write this relation in the form an explicit equation, while clearly specifying the parameters that have to be estimated from data. (1)

iv) The insurer has past data for n customers, i.e., the number of claims (N_i) and age (x_i) of the i^{th} customer, for $i = 1, 2, \ldots n$, are available. Derive explicit equations that must be satisfied maximum likelihood estimators of the parameters specified in part (iii). [There is no need to derive the estimators or to show that the likelihood is indeed maximized at these estimators.] (3)

v) Show how the equation obtained in part (iii) will change if the gender of the insured is also included as a covariate (2)

a. with no interaction between age and gender,

b. with interaction between age and gender.

[9]

Q. 10) A health insurer offers one year "mediclaim" policies covering some specific events with yearly premium Rs.180. The total annual claim amount arising from an individual policy has a compound Poisson distribution with Poison parameter ¼. Individual claim amounts have a normal distribution with mean Rs. 600 and standard deviation Rs. 50. The claim related expenses (incurred at the time of settlement of claim) is a random variable, uniformly distributed over the interval Rs. 40 to Rs. 80, and is independent of the claim amount. Let S be the total aggregate claim amounts and expenses arising over one year from the portfolio.

i) Derive the moment generating function of S. (5)

ii) If S is assumed to have an approximately normal distribution, estimate the minimum number policies that the insurer must sell in order to be at least 99% sure of making a profit from the portfolio in that year. (5)

[10]

Q. 11) The claim amount (X) arising from policies of a general insurance portfolio is assumed to have probability density function $f(x)$ given by

$$f(x) = 2cxe^{-cx^2} \text{ for } x > 0,$$

c being an unspecified parameter. In a year, there are 1000 claims of amounts $X_1, X_2, \ldots, X_{1000}$. The median of these claim amounts is 5000, the mean is 5120.5, and it is also known that

$$\sum_{i=1}^{1000} X_i^2 = 97{,}644{,}400{,}000.$$

i) Derive an explicit expression for the maximum likelihood estimator (MLE) of the parameter c on the basis of the data $X_1, X_2, \ldots, X_{1000}$. (3)

ii) Compute the MLE of c from the given data summary. (1)

ii) Compute the method of moments estimate of c from the given data summary. (3)

iv) Compute the method of percentiles estimate of c from the given data summary. (2)

[9]

Q. 12) The individual claim amounts for the current year, from a stable portfolio of a large insurer, has the probability density function

$$f(x) = \frac{2(500)^2}{(x+500)^3}$$

for x > 0. The portfolio is reinsured by an excess of loss reinsurance arrangement with a fixed retention limit Rs 600 lakhs. The claim amount is expected to inflate at a constant rate of 10% per annum from now.

i) Calculate the probability density function of the individual claim amounts after *n* years. (3)

ii) Calculate the expected size of the individual claim amounts after *n* years. (1)

iii) Calculate the expected claim amount paid by the insurer in respect of an individual claim, after *n* years. (4)

iv) What happens to the expected claim amount paid by the insurer after *n* years, as *n* tends to infinity? Explain either by general reasoning or by analyzing the result of part (iii). (1)

v) What happens to the insurer's share of the expected claim amount paid, as *n* tends to infinity? Explain. (1)

[10]

****