INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

30th October 2007

Subject CT3 – Probability and Mathematical Statistics

Time allowed: Three Hours (10.00 – 13.00 Hrs)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. Do not write your name anywhere on the answer sheet/s. You have only to write your Candidate's Number on each answer sheet/s.
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
Please return your answer scripts and this question paper to the supervisor separately.

Q. 1) Show that a Binomial *(n, p)* random variable is a sum of *n iid* Bernoulli *(p)* random variables. **[2]**

Q. 2) If *A* and *B* are independent and P [*A*] = P [*B*] =1/2, Find P [$A\bar{B} \cup \bar{A}B$] where $\bar{B}$ denotes complement of *B*. **[3]**

Q. 3) A truck can hold 25 containers (each of them is identical in shape and size). The weight of the containers follows a normal distribution with mean 150 *Kg* and standard deviation 15 *Kg*. What is the probability that 25 containers will overload the truck if the maximum load limit is 4000 *Kg*? You can use approximation. **[3]**

Q. 4) For three candidates *A*, *B* and *C*, the chances of becoming an actuary of a certain insurance company have the ratio 4:2:3. The probabilities that a new insurance scheme will be introduced if *A*, *B* or *C* becomes the actuary are 0.3, 0.5 and 0.8 respectively. Find the probability that the insurance scheme will be introduced by *A* **[4]**

Q. 5) The Head of the personal department of an insurance company maintains records of yearly medical leave taken by 30 employees. The data collected by him is as given below:

13	47	10	3	16
7	25	8	21	19
12	45	1	8	4
6	2	14	13	7
34	13	41	28	50
14	26	10	24	36

a) Construct a *stem-leaf* display of the table. The entries on the *leaves* should be in an increasing order. (2)

b) Construct a frequency distribution by taking class intervals of the form 0-10, 10-20 and so on. (2)

c) Draw a histogram of the data and suggest whether the distribution is symmetrical, positively skewed or negatively skewed. (2)

[6]

Q. 6) A continuous random variable *X* has the following *pdf*

$f(x) = k\, x\, e^{-x/2}$; *k* is a constant. $x > 0$

a) Find the value of *k* for *f(x)* to be a valid probability density function. (2)

b) Find the cumulant generating function of *X*. (3)

c) Using the cumulant generating function or otherwise, find the mean and variance of *X*. (2)

[7]

Q. 7) The random variable $Y = Log\, X$ has N (10, 4) distribution. Find

a) The *pdf* of *X* (2)

b) Mean and variance of *X* (3)

c) $P(X \leq 1000)$ (2)

[7]

Q. 8) The manager of a nationalized bank is examining the mortgage payments made by customers of the bank. A payment is classified as:

'Good' if it arrives on or before time; 'Delinquent' if it arrives late or is not paid.

In addition the customers' income is classified as 'Low', 'Medium' or 'High'. The distribution of a group of randomly selected 200 customers is as under:

Payment	Income Level		
	Low	Medium	High
Good	45	50	65
Delinquent	5	20	15

a) Test the hypothesis that payment being 'Good' or 'Delinquent' is independent of income level at significance level 0.05. (4)

b) What would you say about the dependence between payment type and income level when the income levels are classified as 'High' or 'Not high'? (3)

[7]

Q. 9) **(a)** State the conditions under which a counting process is called a Poisson process with parameter λ. (3)

(b) Consumers arrive in an insurance company branch office for premium payment according to a Poisson process with rate $\lambda = 4$ per hour. The office opens at 9:00 AM. Find the probability that "exactly one customer visits by 9:30 AM and four more visit till 11:30 AM for premium payment". (4)

[7]

Q. 10) A computer manufacturing company has three plants at X, Y and Z. To measure how much employees at these plants know about total quality management, a random sample of six employees were selected and a quality awareness test was administered. The test scores are as given below:

	Test Scores		
Observation No.	**Plant *X***	**Plant *Y***	**Plant *Z***
1	85	71	59
2	75	75	64
3	82	73	62
4	76	74	69
5	71	69	75
6	85	82	67

a) Test the null hypothesis that the average test scores are same for all the three plants. Assume that the variances of the score in the plants are same. (5)

b) Obtain 95% confidence interval of the population mean for the plant at X. Assume that the variances of the score in the plants are same. (3)

[8]

Q. 11) Suppose that the joint *pdf* of X_1 and X_2 is given by

$$f(x_1, x_2) = x_1^2 + \frac{x_1 x_2}{3} \quad ; \; 0 \le x_1 \le 1 \; , \; 0 \le x_2 \le 2$$

$$= 0 \qquad \textit{otherwise.}$$

Find **a)** The conditional density of X_2 given $X_1 = x_1$. (3)

b) $E(X_2 \mid X_1 = x_1)$ (3)

c) Verify that $E[E(X_2 \mid X_1 = x_1)] = E(X_2)$ (2)

[8]

Q. 12) **a)** Let X be a continuous random variable having *pdf*

$$f(x,\theta)=\frac{\theta^m x^{m-1}e^{-\theta x}}{(m-1)!} \quad ; x \geq 0 \text{ , } \theta > 0,$$
$$= 0 \qquad \textit{otherwise}$$

where m is a known integer ≥ 2 .Show that $\frac{m-1}{x}$ is an unbiased estimator of θ. (4)

b) Let $X_1, X_2,\ldots, X_n$ be a random sample from $f_\theta(x)$ where

$$f_\theta(x) = (1+\theta)x^\theta \quad ; \quad 0 < x < 1$$
$$= 0 \qquad \textit{otherwise}$$

Obtain the m.l.e. of θ. (3)

c) Let $X_1, X_2,\ldots, X_n$ be a random sample from $f_\theta(x)$ where

$$f_\theta(x) = \theta\ e^{-\theta x} \ ; \quad x > 0$$
$$= 0 \qquad \textit{otherwise}$$

Obtain Cramer-Rao Lower Bound (CRLB) for the variance of the unbiased estimator of θ, assuming that the regularity conditions are satisfied. (3)

[10]

Q. 13) The diameter of steel rods manufactured on two different machines A and B is studied. Two random samples of sizes n_1 = 12 and n_2 = 15 are selected and the sample means and sample standard deviations respectively are

$\overline{x_1}$ =24.6, s_1 = 0.85

$\overline{x_2}$ =22.1, s_2 = 0.98

Assuming that the diameters of the rods follow $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ and

a) Test for the equality of the variances. (2)
b) Test the equality of means of the diameters of the rods manufactured by the two machines assuming $\sigma_1^2 = \sigma_2^2$ (3)
c) Construct 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 = \sigma_2^2$ (3)
d) Construct 95% confidence interval for σ_1^2 / σ_2^2 (3)

[11]

Q. 14) The following data refers to the number of claims *(X)* received by a motor insurance company in a week and the number of settlements *(Y)* of these claims in the following week during 10 randomly selected weeks in a year.

X:	100	110	120	130	140	150	160	170	180	190
Y:	45	51	54	61	66	70	74	78	85	89

A regression model $Y = \alpha + \beta X + \varepsilon;\ \varepsilon \sim N(\mu, \sigma^2)$ is to be fitted on the above data.

a) Display the data in a scatter diagram and comment on the selection of a linear model for regression. (2)
b) Compute $\overline{X}$, $\overline{Y}$, S_{XX}, S_{YY}, S_{XY}. Hence, find the estimates of α and β . (3)
c) Obtain the estimate of σ^2 (3)
d) Test the hypothesis $\beta = 0$ against $\beta \neq 0$ (3)
e) Obtain 95% confidence interval for β (3)
f) Let the population correlation coefficient between X and Y be ρ. Compute the sample correlation coefficient and test whether $\rho = 0.75$ against $\rho \neq 0$. (3)

[17]

****