INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

19th May 2009

Subject CT3 – Probability & Mathematical Statistics

Time allowed: Three Hours (10.00 – 13.00 Hrs)

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. <u>However, answers to objective type questions could be written on the same sheet</u>.***
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) A civil engineer monitors water quality by measuring the amount of suspended solids in a sample of river water. Over 11 days, she observed

14 12 21 28 30 63 29 63 55 19 20

suspended solids (parts per million)

a) Draw a dot diagram. (1)
b) Find the median and the mean. Locate both on the dot diagram. (3)
c) Find the sample standard deviation. (2)
[6]

Q 2) The mean weight of 150 students in a class is 60 *kgs.* The mean weight of boys in the class is 70 *kgs* and that of girls is 55 *kgs*.

Find the number of boys and girls. **[3]**

Q 3) A pair of events A and B cannot be simultaneously mutually exclusive and independent. Prove that if $P(A) > 0$ and $P(B) > 0$, then

a) If A and B are mutually exclusive, they cannot be independent.

b) If A and B are independent, they cannot be mutually exclusive. **[2]**

Q 4) Generate three random samples from Binomial B(3, 0.4) using the following values from U(0,1)

0.196 0.351 0.975 **[3]**

Q 5) **a)** State the law of total probability for exclusive events and Bayes theorem. (2)

b) A shop keeper buys a particular kind of light bulbs from three manufacturers. A,B and C she buys 30% of her stock from A, 45% from B and 25% from C. In the past, she has found that 2% of $C^{'s}$ bulbs are faulty whereas only 1% of $A^{'s}$ and $B^{'s}$ are. Suppose that she chooses a bulb at random and it is faulty, what is the probability that it was one of $C^{'s}$ bulbs. (3)
[5]

Q 6) A random variable X has pdf

$$f(x) = \frac{x}{a^2} e^{-\frac{x^2}{2a^2}} ; a > 0,\ 0 < x < \infty$$

a) Find the Inter Quartile Range (IQR). (2)

b) Show that the ratio of IQR to standard deviation of X is free from a. (4)
[6]

Q 7) Suppose that for a given population with $\sigma = 8.4$ square inches. One wants to test the null hypothesis $H_0 : \mu = 80.0$ square inches against the alternative $H_1 : \mu \neq 80$ square inches on the basis of a random sample of size $n = 100$.

If the null hypothesis is rejected for $\overline{X} < 78.0$ square inches, what is the probability of type I error? Find the power at $\mu = 82$. **[3]**

Q 8) Define compound distribution. (1)

The number of claims N, which arises in a year from a group of policies, has negative binomial distribution

$$P(N = n) = \binom{n+2}{n}(0.9)^3(0.1)^n \quad n = 0,1,2,...$$

The claim amounts (in Rs.1000s) are independent and identically distributed as gamma (6,2) and also independent of N. Let Y be the total claim amount arising from these policies.

a) Obtain an expression for the *mgf* of Y (You may assume the *mgf* of negative binomial and gamma). (3)

b) Compute the standard deviation of Y. (2)

[6]

Q 9) Let X and Y have joint *pmf*

$$p(x, y) = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} \quad \text{if} \quad y = 0,1,...,x; \quad x = 0,1,...; \quad \lambda > 0; \quad 0 < p < 1$$

a) Find marginal *pmfs* of X and Y. (4)

b) Find conditional distribution of Y for a given X, conditional distribution of X for a given Y and comment on these results. (4)

c) Find $E[Y/X=x]$ and $E[X/Y=y]$ (2)

[10]

Q 10) Let $X_1, X2, ..., X_n$ be a random sample from $U(-\theta, \theta)$, $\theta > 0$.

a) Find the maximum likelihood estimator of θ. (4)

b) Find the moment estimator of θ. (2)

c) State *CRLB*? Explain why *CRLB* cannot be applied in this case. (2)

d) Using the data below compute *ML* estimate moments estimate of θ.

4.0 4.5 -3.8 -1.4 4.3 2.8 1.2 2.7 3.1 -2.2 (2)

[10]

Q 11) The yields of tomato plants grown using different types of fertilizers are given below.

Fertilizer	Yield (in kgs)				
X	3.5	4.0	3.8	4.1	4.4
Y	4.7	5.0	4.5	5.3	4.6
Z	3.6	3.9	4.2	4.1	4.0

a) Is there evidence that the fertilizers produce different yields? Test at 5% level. State the hypothesis and assumptions you make. (4)

b) Irrespective of the conclusions in (a) above, if X and Z are chemical fertilizers and Y is a natural fertilizer, examine whether there is evidence to suggest that the average effect of the two chemical fertilizers are inferior to natural fertilizers regarding the yield. (3)

[7]

Q 12) State the postulates of Poisson process. If the number of accidents in a town follows a Poisson process with a mean of 2 per day and the number X_i of people involved in the accident has the distribution

$$P(X_i = k) = \frac{1}{2^k}, k = 1,2,...$$

Obtain the mean and variance of the number of people involved in accidents per week. **[5]**

Q 13) An experiment was conducted on 20 identical square metal sheets, in order to study the increases in (a) their length (X in cms), (b) their breadth (Y in cms) and (c) the difference between their length and breadth ((X-Y) in cms), when they were subjected to a heat process. The following summary was observed.

Sample variance of X (s_X^2) = 1.2011 ; Sample variance of Y (s_Y^2) = 2.2958

Sample variance of (X-Y) ($s_{(X-Y)}^2$) = 1.8673

a) Calculate the sample correlation coefficient between X and Y. (1)

b) Construct the 95% CI for the population correlation coefficient ρ, on assuming bivariate normality. (3)

c) Without carrying out a statistical test, conclude the acceptance or rejection of the null hypothesis $H_0 : \rho = 0$ against the alternative $H_1 : \rho \neq 0$ at 5% level of significance. (1)

[5]

Q 14) In an examination in Mathematics 20 randomly selected students at a government school had obtained mean mark of 52 and the sum of squares of deviations from this mean mark was 1312. From a private school, 17 randomly selected students, taking the same examination, obtained a mean mark of 36, and the sum of squares of deviations from this mean was 1401.

Assume that the marks obtained in government schools follow a normal distribution (μ_1, σ_1^2) and that in private schools follows a normal distribution (μ_2, σ_2^2), and that the two populations are independent.

a) Test the null hypothesis $H_0 : \mu_1 = 2\mu_2$ vs $H_1 : \mu_1 \neq 2\mu_2$, when $\sigma_1^2 = 70$ and $\sigma_2^2 = 90$. (3)

b) Test the hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ (2)

c) If your conclusions is to accept the hypothesis in (b), test the null hypothesis $H_0 : \mu_1/\mu_2 = 1$ vs $H_1 : \mu_1/\mu_2 \neq 1$ when the common population variance is unknown. Comment on testing H_0, if the hypothesis in (b) is rejected. (3)

d) Assuming the population variances to be equal, obtain *95%* confidence interval for $(\mu_1 - \mu_2)$ (2)

[10]

Q 15) In a spring balance, when weights are added to the scale pan, the spring stretches. The following table shows the results obtained when different load X (in *Newtons*) were applied in a random order and the length of the spring Y (*in cms*).

X :	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Y :	10.7	11.3	12.0	12.4	13.0	13.7	14.5	15.1	15.6	16.0

a) Representing the relationship between X and Y as $Y_i = \alpha + \beta X_i + e_i$ where e_i's are *iid* $N(0,\sigma^2)$ random variables, estimate α and β by the method of least squares. (5)

b) If the problem were to be to estimate α and β by maximum likelihood (ML) method, what will be your answer? (1)

c) Estimate the value of σ^2. (2)

d) Obtain the residuals using the fitted regression and plot the same against the fitted responses $\left(\hat{Y}\right)$.Comment on the adequacy of the model (4)

[12]

Q 16) A company employs drivers to work round the clock. The driver's union is concerned that some periods are more dangerous than to others. The management contests this, claiming that no one period is more dangerous than any other. The following data gives the details of the incidence of traffic accidents by time of day for this group over the previous two years.

Time of Day	No. of Accidents
00.01 – 04.00	14
04.01 – 08.00	16
08.01 – 12.00	24
12.01 – 16.00	22
16.01 – 20.00	24
20.01 – 24.00	20

a) Are these results consistent with the management claim? (4)

b) The company wishes to examine whether there is an association between accidents proneness and colour blindness. The results for a group of 80 drivers (with a minimum of 5 years employment) are as in the following table.

	Colour Blindness	
Accidents driving last 5 years	Yes	No
None	22	5
One or more	38	15

Is there sufficient evidence to conclude that there is an association between colour blindness and accident proneness at 5% level? (3)

[7]

****