[This question paper contains 6 printed pages.]

Sr. No. of Question Paper : 863   E   Your Roll No................

Unique Paper Code   : 234615

Name of the Course   : **B.Sc. (H) Computer science**

Name of the Paper   : Statistical Methodology (STC-402)

Semester   : VI

Duration : 3 Hours   Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.

2. Attempt **15** questions in all, Selecting **five** questions from each section.

3. All questions carry equal marks.

4. Use of simple calculators is allowed.

SECTION - I

1. Define Karl Pearson's coefficient of correlation $r(X, Y)$ for a bivariate frequency distribution (X_i, y_i), $i = 1, 2, ..., n$. Prove that $-1 \leq r(X, Y) \leq 1$.

2. Two random variables X and Y have zero means, the same variance σ^2 and zero correlation. Obtain the coefficient of correlation between $u = x \cos \alpha + y \sin \alpha$ and $v = x \sin\alpha + y \cos\alpha$.

3. The coefficient of rank correlation between marks in Statistics and marks in Mathematics obtained by certain group of students is 0.8. If the sum of squares of the difference in ranks is given to be 33, find the number of students in the group.

P.T.O.

4. The following information is available about advertisement expenditure and sales:

	Advertisement Exp.	Sales
	(Rs. Cr.)	(Rs. Cr.)
Mean	20	120
Standard Deviation	5	25

Correlation coefficient = 0.8

(a) Determine two regression lines.

(b) What will be likely sales when advertisement expenditure is Rs. 25 Cr?

(c) What should be the advertisement budget if the company wants to attain the sales target of Rs. 150 Cr.?

5. For certain X and Y series which are correlated, the two lines of regression are:

$5X - 6Y + 90 = 0$

$15X - 8Y - 130 = 0$

(a) Find the means of two series;

(b) Find the correlation coefficient.

6. Differentiate between multiple and partial correlation coefficients. Show that

$1 - R_{1.23}^2 = (1 - r_{12}^2) \cdot (1 - r_{13.2}^2),$

where the symbols have their usual meaning.

SECTION-II

7. If X is a random variable with mean μ and variance σ^2, then show that:

$$P\{|X-\mu|\le k\}\ge 1-\frac{\sigma^2}{k^2}$$

where $k > 0$ is any constant. If E(X)=10 and var(X)=15, then find a bound on the probability P(5 < X < 15).

8. State Central Limit Theorem (CLT). Let X_i, $i= 1, 2, \dots ,10$ be independent random variables, each being uniformly distributed over (0,1) . Apply CLT to estimate $P\left\{\sum_{i=1}^{10} X_i > 7\right\}$.

9. For standard exponential distribution

$$f(x) = \begin{cases} e^{-x}, & x \ge 0 \\ 0, & \text{otherwise,} \end{cases}$$

determine the CDF of the largest order statistic $X_{(n)}$ in a random sample of size n.

10. The random variable X. representing the number of cherries in a cherry puff, has the following probability distribution:

X:	4	5	6	7
P(X=x):	0.2	0.4	0.3	0.1

(a) Find the mean and variance of the mean $\bar{x}$ for random sample of 36 cherry puffs.

(b) Find the probability that the average number of cherries in 36 cherry puffs will be less than 5.5.

11. The average life of a bread making machine is 7 years, with the standard deviation of 1 year. Assuming that the lives of these machines follow approximately a normal distribution, find

   (a) The probability that the mean life of a random sample of nine such machines falls between 6.4 and 7.2 years.

   (b) The value of x to the right of which, 15% of the means computed from random samples of size 9 would fall.

12. Write short notes on the following,

   (a) Sampling distribution of mean

   (b) Type I and Type II error.

SECTION - III

13. Suppose that it is known from experiences that the standard deviation of the weight of 8-ounce packages of cookies made by a certain bakery is 0.16–ounce. To check whether its production is under control on a given day, that is, to check whether the true average weight of the package is 8-ounces, employees select a random sample of 25 packages and find that there mean weight is 8.091-ounces. Since the bakery stands to loose money when $\mu > 8$ and the customer looses out when $\mu < 8$ test the hypothesis that $\mu = 8$ against the hypothesis $\mu \neq 8$ at the 0.05 level of significance.

14. Test the hypothesis that the average content of containers of a particular lubricant is 10 liters if the contents of a random sample of 10 containers are 10.2, 9.7, 10.1, 10.3, 10.1, 9.8, 9.9, 10.4, 10.3 and 9.8 liters. Use a 0.01 level of significance and assume that the distribution of contents is normal.

15. Two new methods for producing a tire have been proposed. The manufacturer believes there will be no appreciable difference in the lifetimes of the tire produced by these methods. To test the plausibility of such a hypothesis, a sample of 9 tires is produced by method 1 and a sample of 7 tires by method 2. The first sample of tires is to be road tested at location A and the second at location B. It is known from previous experience, that the lifetime of a tire tested at either of these locations is a normal random variable with a mean life due to the tire but with a variance that is due to the location. Specifically, it is known that the lifetimes of tires tested at location A are normal with mean 62.24 and the standard deviation equal to 3000 kilometers, whereas those tested at location B have lifetimes that are normal with mean 58.27 and the standard deviation of 4000 kilometers. Test the hypothesis that the mean lifetime is same for both types of tires? Use a 5% level of significance.

16. In a study to estimate the proportion of residents in a certain city and its suburbs who favour the construction of a nuclear power plant, it is found that 63 of 100 urban residents favour the construction while only 59 of 125 suburban residents are in favour. Is there a significant difference between the proportion of urban and suburban residents who favour construction of the nuclear plant?

17. A die is tossed 180 times and the number shown (X) and the corresponding frequencies (f) are recorded with the following results:

X:	1	2	3	4	5	6
f:	28	36	36	30	27	23

Is this a balanced die? Use an appropriate level of significance.

18. A study is conducted to compare the length of time between men and women to assemble a certain product. Past experience indicates that the distribution of times for both men and women is approximately normal but the variance of the times for woman is less than that for man. A random sample of times for eleven man and fourteen woman produced standard deviations of 6.1 and 5.3 respectively. Test the hypothesis that $\sigma_1^2 = \sigma_1^2$ against the alternative $\sigma_1^2 > \sigma_1^2$. Use a 5% level of significance.