STATISTICS

3RD SEM SEC, 2025

PAPER CODE: SEC0326403N

PAPER NAME: STATISTICS FOR DATA ANALYSIS USING SOFTWARE PACKAGE- R PROGRAMMING

Full marks: 30
Time: 1 hour 30 minutes

1. Answer the following questions: 1x5=5

   a. Define R language.
   b. Define null hypothesis and alternative hypothesis.
   c. Range of Karl Pearson coefficient is between ___ to ____.
   d. Taking a=1, b=2, c=3. Write a R code to evaluate the following expression:

$$\frac{a}{b^{(a+c)}}$$

   e. Give the concept of degrees of freedom.

2. Answer any five questions: 2x5=10
   a. Define goodness of fit test.
   b. What is Type I and Type II error?
   c. Write a short note on the basic principles of sample survey.
   d. Generate first 10 natural numbers using R code and find their sum.
   e. Write a short note on the types of graphs that can be used for drawing continuous data.
   f. Write two applications of Chi-square distribution.
   g. X and Y take the following values

X	10	8	7	5	1
y	5	3	10	9	1

Evaluate mean of x and mean of y using R coding.

3. Answer any **three** questions: 5x3=15

a. Discuss about the two methods by which simple random sample can be drawn.

b. The means of two large samples of sizes 500 and 600 members are 55.5 inches and 56.0 inches respectively. Can the samples be regarded as drawn from the same population of standard deviation 2.5 inches at 5% level of significance.

c. Prepare a blank table

Main Characteristics	Sub Characteristics
Place	Rural, Urban
Religion	Hindu, Non-Hindu
Gender	Male, Female, Transgender
Family Income	0-1,1-5, 5-10, above 10 (in lakhs)
Age	0-20,20-40,40-60, 60 above (in years)

d. Given the values of X and Y

X	1	2	3	4	5	6	7	8	9
Y	2	5	7	9	4	2	1	3	6

Write a R code to find

i) Mean of x and mean of y

ii) Correlation between x and y

e. Give reasons why sample survey is better than complete census.

XXXXXXXX