No. of Printed Pages : 5 **MSTL–013(Set-II)**

# M. SC. (APPLIED STATISTICS)

## (MSCAST)

## Term-End Practical Examination

## December, 2024

### MSTL–013(Set-II) : STATISTICAL COMPUTING USING R—III

*Time : $2\frac{1}{2}$ Hours* *Maximum Marks : 50*

---

***Note*** **:** *(i) Attempt any **two** questions.*

*(ii) Solve the questions using R-software and create a script file.*

*(iii) Mention necessary steps, hypotheses interpretation, etc.*

*(iv) Symbols have their usual meanings.*

---

1. A population consisting of 6 clusters, each containing 6 units. The value of the study variable Y, observed for each unit within each cluster, are given as follows : 7+6+6+6

| Cluster | Values of Y |
|---|---|
| 1 | 2, 4, 6, 1, 3, 5 |
| 2 | 2, 5, 3, 4, 7, 4 |
| 3 | 4, 3, 6, 2, 1, 5 |
| 4 | 3, 2, 5, 1, 6, 4 |
| 5 | 2, 4, 6, 8, 3, 5 |
| 6 | 4, 1, 2, 7, 5, 3 |

A random sample of three clusters was selected from the population (first-stage sampling) and 3 units were randomly chosen from each of the selected clusters (second-stage sampling). The selected clusters and corresponding units are as follows :

(i) Cluster 2 : The Y-values 2, 7 and 3 were selected.

(ii) Cluster 4 : The Y-values 6, 3 and 1 were selected.

(iii) Cluster 6 : The Y-values 7, 4 and 3 were selected.

Answer the following questions :

(a) Write R-code to define the population data as a list of clusters.

(b) Using R, extract the selected clusters (2, 4 and 6) from the population.

(c) Extract the second-stage samples (specified Y-values) from the selected clusters.

(d) Calculate the sample mean of Y for the second-stage sample and print the results.

2. (a) An economist believes that the median starting salary for a computer programmer in a certain city is ₹ 12,000. To varify this claim, 12 computer programmers with similar backgrounds who recently got the jobs were randomly selected. Their starting salaries (in ₹) were as follows :

| **Salaries (in ₹)** | | | | |
|---|---|---|---|---|
| 18,000 | 15,000 | 10,000 | 13,000 | 12,000 |
| 10,000 | 16,000 | 9,000 | 10,000 | 9,000 |

Write an R-script to test the economist belief using the sign test at 1% level of significance.

(b) The following are the grades in Mathematics of High School students

taken randomly from two schools : School-I and School-II :

| Grades | No. of Students in School-I | No. of Students in School-II |
|---|---|---|
| A | 2 | 1 |
| $A^+$ | 2 | 1 |
| B | 4 | 2 |
| $B^+$ | 6 | 3 |
| C | 3 | 4 |
| $C^+$ | 3 | 2 |
| D | 0 | 4 |
| $D^+$ | 0 | 3 |

Write an R-code to apply Kolmogorov-Smirnov test to examine the performance in Mathematics of High School students in School-I and School-II is same at 5% level of significance.

(c) Generate 50 random numbers from a Poisson distribution with $\lambda = 3.5$ using R. Match the result using the direct function of R. 10+7+8

3. (a) Let :

$$x = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \text{ and } \mathrm{A} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}$$

Write R-script to compute the quadratic form $\mathrm{Q}(x) = x'\mathrm{A}\,x$.

(b) Given matrix :

$$\mathrm{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Write a R-script to apply the Gram-Schmidt process to make the vectors orthogonal, then normalize each resulting vector to create an orthonormal matrix.

10+15

× × × × × × ×