No. of Printed Pages : 7 **MST–015**

M. SC. (APPLIED STATISTICS)

(MSCAST)

Term-End Examination

June, 2025

MST-015 : INTRODUCTION TO R SOFTWARE

Time : 2 Hours *Maximum Marks : 25*

Note : *(i) Question No.* ***1*** *is compulsory.*

(ii) Attempt any ***two*** *questions out of the remaining Question Nos.* ***2*** *to* ***4****.*

(iii) Use of scientific calculator (non-programmable) is allowed.

(iv) Symbols have their usual meanings.

---

1. Answer the following : 5×1=5

(a) Write a command to get help on the **for** reserved word.

(b) Differentiate between the use of the **sep** and **collapse** arguments of the **paste( )** function.

(c) Write the output of the following statement :

is.na (c(9, –2, 0, NA, NaN, 10))

(d) Which of the following user defined function names is inappropriate and why ?

(i) **matrix**

(ii) **abs**

(e) Write an assignment statement equivalent to the following equation in R :

$$\text{Area} = 2\pi r^2 + 2\pi r h$$

2\. (a) Write R code to create the graph of the following function : 3

$$f(x) = |x|, \qquad -5 \le x \le 5$$

(b) Consider the following data : 3

37, NA, 30, 49, 110, 96, NA,

23, 21, 7, 21, 9, NA, 37, 30, 30, 21

Write R commands to perform the following tasks :

(i) Remove NA's from this data and save it under the name **Air**.

(ii) Create a histogram by specifying the colors for filling bars.

(iii) Extract the frequency distribution corresponding to the histogram.

(c) Write an **if-else** statement to compute the value of y, where : 4

$$y = \begin{cases} 4x^3 \text{ and increase } x \text{ by } \dfrac{1}{2}, & \text{if } x \le 4 \\ 4(x-1)^2 \text{ and decrease } x \text{ by } \dfrac{1}{2}, & \text{if } x > 4 \end{cases}.$$

3. (a) Write R commands to create a function to compute the variance of the following discrete frequency distribution : 3

$$x_i \mid f_i, \quad i = 1, 2, 3, 4$$

(b) Consider the following height and weight data : 3

Height (in inches)	Weight (lbs)
57	113
58	116
59	118
60	121
61	124
62	127
63	130
64	132
65	135
66	138

Create a data frame named **DF** of the data and write R commands to do the following tasks :

(i) Extract the first 9 rows of the data frame and assign it under the name **women**.

(ii) Compute the Karl Pearson's correlation coefficient between the columns of **women**.

(iii) Create scatter plot of the variables of **DF** in a single plot.

(c) Write step-by-step execution of the following code : 4

```
m <- 6
while (m > 3) {
n ← m %/% 2
if (n==0) m ← m*5 else m < - m – 1
}
Cat ("n=", n, "m=", m)
```

4. (a) Consider the following three matrices : 4

```
A ← matrix (1 : 4, nrow = 2)
B ← matrix (rep(1, 4), ncol = 2)
C ← matrix (5 : 8, nrow = 2,
            ncol = 2)
```

Create the matrices and write the outputs of the following commands :

(i) **A – C + B % * % C**

(ii) **cbind (A, B)**

(iii) **diag(C)**

(iv) **A[1,]**

(b) Write R commands to create a .csv file consisting of the following data : 4

Air. Flow	**Water. Temp**	**Acid. Conc.**	**Stack. loss**
80	26	90	42
80	24	89	37
75	22	91	37
62	21	88	28
62	23	94	18

If the data is saved in SL.csv file, then read the data and name it as **SLOSS**. Also, write R commands to :

(i) Name the rows as **Row1, Row2, Row3, Row4** and **Row5** and columns as **Col1, Col2, Col3** and **Col4.**

(ii) Compute the row means and column sums of the data frame.

(iii) Sort **SLOSS** according to the **Col2** of it.

(c) Write R commands to get the following outputs : 2

(i) **a##1$, b##3$, c##5$**

(ii) **A:1, B:2, C:3, D:4**

× × × × ×