Total number of printed pages–15

3 (Sem–4/CBCS) STA SE 1/2/3

2023

STATISTICS

(Skill Enhancement Course)

Answer the Questions from any one Option.

OPTION – A

(Statistical Data Analysis using R)

Paper : STA–SE–4014

OPTION – B

(Database Management System)

Paper : STA–SE–4024

OPTION – C

(Statistical Techniques for Research Methods)

Paper : STA–SE–4034

Full Marks : 50

Time : Two hours

The figures in the margin indicate full marks for the questions.

Answer ***either*** *in English* ***or*** *in Assamese.*

Contd.

OPTION – A

(*Statistical Data Analysis using R*)

Paper : STA–SE–4014

1. Answer the following questions : 1×4=4

(a) What is R ?

(i) An object-oriented programming language

(ii) An open source project from CRAN

(iii) A programming language for statistical computing

(iv) All of the above

(Choose the correct option)

(b) What is primary file type of R ?

(i) Vector

(ii) Text file

(iii) R scripts

(iv) Statistical file

(Choose the correct option)

(c) _____ function is used for reading the .csv file in R language.

(i) write.csv( )

(ii) read.csv( )

(iii) let.csv( )

(iv) table.csv( )

(Choose the correct option)

(d) If a command is not complete at the end of a line, R will give a different prompt, by default it is

(i) *

(ii) –

(iii) +

(iv) 1

(Choose the correct option)

2. Answer the following questions : 2×3=6

(a) Discuss the steps required for saving data in R.

(b) Write the summary and table function used in R.

(c) Write R program for addition of two 2×2 matrices.

3. Answer the following questions : ***(any two)*** 5×2=10

(a) What are the data frames? Write its significance in R language.

(b) Write a note on import data into R program from various data sources.

(c) Write a note on different data types in R.

(d) Write R program to find the correlation matrix of a multivariate dataset.

4. Answer the following questions : 10×3=30

(a) The frequency distribution of number of heads obtained in an experiment of tossing 5 coins 110 times is given below :

Number of heads	:	0	1	2	3	4	5
Frequency	:	6	15	25	42	18	4

Fit a binomial distribution to the above data and find the expected frequency. Plot observed and expected frequencies and comment on the adequency of model.

Or

Write R code to

(i) draw histogram for a given data. 5

(ii) draw ogives with graphical summary of a given data. 5

(b) Write R code to calculate arithmetic mean, geometric mean, harmonic mean, mode, median, sixth decile for given $x_i / f_i \ (i = 1, 2, ..., 5)$.

Or

(i) Write an R program for a two-sample *t*-test. 5

(ii) Write R program for confidence interval of population mean of a normal distribution. 5

(c) Write R program to fit a normal distribution for $x_i / f_i \ (i = 1, 2, 3, 4, 5)$ and also test the goodness of fit.

Or

Fit a second degree parabola $y = a + bx + cx^2$ to the following data :

Year *(X)* :	1999	2000	2001	2002	2003	2004	2005
Profit *(Y)* : (in'000)	24	27	32	38	45	20	17

Estimate profit for the year 2006. Find estimated values of *Y* for each *X*. Plot estimated and observed values and represent parabola by graph.
