

Total number of printed pages-20

3 (Sem-4/CBCS) STA SE 1/2/3

2022

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

Paper : STA-SE-4014

(Statistical Data Analysis Using R)

1. Answer the following questions : *(any four)*

$$1 \times 4 = 4$$

(a) A command used for 'logarithm of x with base n ' is _____.

(Fill in the blank)

(b) In R, missing values are represented by _____ which should be in capital letters. *(Fill in the blank)*

(c) For a given vector $x = c(4, 7, 2, 9, 5, 10)$, the values obtained by using `sort(x)` are _____. *(Fill in the blank)*

(d) Which of the following is used for reading in saved workspaces ?

- (i) Unserialize
- (ii) Load
- (iii) Get
- (iv) Set

(Choose the correct option)

(e) What plot(s) are used to view the linear regression ?

- (i) Scatter plot
- (ii) Box plot
- (iii) Density plot
- (iv) All of the above

(f) Which of the following codes creates n samples of size 'size' with probability 'prob' from the binomial ?

- (i) `z <- rnom(n, size, prob)`
- (ii) `z <- rbinom(n, size, prob)`
- (iii) `z <- binom(n, size, prob)`
- (iv) `z > - nom(n, size, prob)`

(Choose the correct option)

(g) Which of the following codes creates an 'n' item vector of random normal deviates ?

- (i) `X1 <- C(snorm(n))`
- (ii) `X1 <- C(pnorm(n))`
- (iii) `X1 <- C(rnorm(n))`
- (iv) `X1 <- C(norm(n))`

(Choose the correct option)

(h) _____ lists the variables in the workspaces.

- (i) `rm(x)`
- (ii) `rm(list = ls())`
- (iii) `ls()`
- (iv) `attach(mat)`

(Fill in the blank)

2. Answer the following questions : **(any three)**
2×3=6

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

(b) Write down the R command to construct the following matrix :

$$\begin{pmatrix} 1 & 3 & 9 \\ 4 & 2 & 6 \end{pmatrix}$$

(c) Write down a short note on the relational operators used in R programming language.

(d) Write down the R command to draw a histogram for the data set —
12, 14, 5, 7, 14, 17, 4, 9, 9, 5.

(e) Write down a short note on the logical operators used in R programming language.

(f) Write down the R command to find out the sequence of odd numbers less than or equal to 21.

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

(a) Write a note on different data structures in R.

(b) Write down different commands available in R to load text data.

(c) Write the procedure for t-test of one sample in R.

(d) Write down the R commands of the two samples t-test.

(e) From a population of size 15000, draw a sample of size 8 (without replacement) by using appropriate R-programming)

(f) Write R codes to construct box plot and qq plot to check the normality of the parent population.

4. Answer the following : **(any three)**

10×3=30

(a) Given,

Speed	Distance
4	2
4	10
7	4
7	22
8	16
9	10

Write down the commands to carry out the linear regression in R programming language.

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

(c) Write R codes to —

(i) draw bar chart for a given data vector r .

(ii) draw pie chart for a given data vector r .

(d) The test scores of 8 students of class XI are as follows :

Test 1 : 90 90 80 90 92 88 90 63

Test 2 : 84 84 82 94 90 85 89 62

Using R programming language determine whether these students can perform equally in two different tests.

(e) Write R codes to fit a binomial distribution for given x_i/f_i ($i = 1, 2, \dots, 5$) and also test the goodness of fit.

(f) Using R, fit a Poisson distribution to the following data :

Number of mistakes per page	Number of pages
0	109
1	65
2	22
3	3
4	1
Total	200

(g) A random sample of size 10 is taken from a normal population with mean μ and variance σ^2 , where σ^2 is unknown.

$$\text{Given } \sum_{i=1}^{10} x_i = 60.5 \quad \text{and} \quad \sum_{i=1}^{10} x_i^2 = 390.$$

Write R codes to determine a 90% confidence interval for the population mean.

(h) For the given vectors x and y , write R codes to fit a line of regression and plot a graph.

OPTION - B

Paper : STA-SE-4024

(Database Management System)

1. Answer **any four** of the following : $1 \times 4 = 4$

(i) Which of the following is not a type of database ?

- (a) Hierarchical
- (b) Network
- (c) Distributed
- (d) Decentralized

(Choose the correct option)

(ii) Which of the following is a feature of DBMS ?

- (a) Minimum Duplication and Redundancy of Data
- (b) High level of security
- (c) Single-user access only
- (d) Support ACID property

(Choose the correct option)

(iii) Which of the following is a function of the DBMS ?

- (a) Storing data
- (b) Providing multi-users access control

(d) Data integrity

(e) All of the above

(Choose the correct option)

(iv) The DBMS acts as an interface between _____ and _____ of an enterprise-class system.

- (a) data and the DBMS
- (b) application and SQL
- (c) database application and the database
- (d) the user and the software

(Choose the correct option)

(v) Who created the first DBMS ?

- (a) Edgar Frank Codd
- (b) Charles Bachman
- (c) Charles Babbage
- (d) Sheron B Codd

(Choose the correct option)

(vi) In which of the following formats data is stored in DBMS ?

(a) Image

(b) Text

(c) Table

(d) Graph

(Choose the correct option)

(vii) What command is used to create a new table in SQL ?

(a) CREATE TABLE

(b) BUILD TABLE

(c) GENERATE TABLE

(d) None of the above

(Choose the correct option)

(viii) Which of the following is the full form of DDL ?

(a) Data Definition Language

(b) Data derivation Language

(c) Dynamic Data Language

(d) Detailed Data Language

(Choose the correct option)

2. Answer **any three** of the following : $2 \times 3 = 6$

(a) What is meant by ACID properties in DBMS ?

(b) What are the disadvantages of file processing system ?

(c) Define tuple and attribute.

(d) What is meant by a database ?

(e) What are the main differences between primary key and unique key ?

(f) What is meant by RDBMS ?

3. Answer **any two** of the following : $5 \times 2 = 10$

(a) What are different languages present in DBMS ? Explain them.

(b) What is the use of DROP command ? What are the differences among DROP, TRUNCATE and DELETE commands ? $2+3=5$

(c) Write a short note on relational model.

- (d) Write a SQL query for the following table where it returns only the top five selling items in the table :

iD	Name	Amount Sold
1	Cup	11
2	Saucer	22
3	Plate	46
4	Fork	34
5	Spoon	45
6	Knife	78
7	Mug	23
8	Glass	64
9	Tumbler	24

- (e) Create a table of students, taking the attributes given below :
Roll No, Student name, Address, Dept., Contact no.

- (f) Write the difference between primary, candidate and key.

4. Answer **any three** of the following : $10 \times 3 = 30$

- (a) What is SQL ? Why do we use SQL ?
Explain different types of joins in SQL. $2+2+6=10$

- (b) Explain different types of keys in DBMS with example. 10

- (c) Explain different types of relationships in DBMS. What are the major advantages of DBMS ? Write the differences between DBMS and RDBMS. $4+3+3=10$

- (d) What do you mean by data model ? Explain all the types of data model used in DBMS. What is an entity-relationship model ? $2+6+2=10$

- (e) Explain 2-tier and 3-tier architecture in DBMS with example.

- (f) What is an index in SQL ? Describe different types of indexes in SQL with example. What are the set operators in SQL ? $2+6+2=10$

- (g) What is data abstraction in DBMS ? Explain different levels of data abstractions in DBMS. What is query in DBMS ? $2+6+2=10$

(h)

B

id	Name	Age
1	Bob	21
2	Sam	19
3	Jill	18
4	Jim	21
5	Sally	19
6	Jess	20
7	Will	21

- (i) Write an SQL query that will group all the people by their age, along with a count of people that are the same age in the table. 5
- (ii) Write an SQL query that will return a sum of all the ages in the table. 5

OPTION - C

Paper : STA-SE-4034

(Statistical Techniques for Research Methods)

1. Answer the following questions as directed :
(any four) 1×4=4

নিৰ্দেশনা অনুসৰি নিম্নলিখিত প্ৰশ্নবোৰৰ উত্তৰ দিয়া :
(যিকোনো চাৰিটা)

- (a) The research problem undertaken for study must be carefully _____.
(Fill in the blank)

অধ্যয়নৰ বাবে কৰা গৱেষণাৰ সমস্যাটো সাৱধানে
_____ হ'ব লাগিব। (খালী ঠাই পূৰ কৰা)

- (b) Data collected for survey from warranty cards are secondary data.
(True or False)

ৱাৰেণ্টি কাৰ্ডৰ পৰা জৰীপৰ বাবে সংগ্ৰহ কৰা তথ্য
হৈছে গৌণ তথ্য। (সঁচা বা মিছা)

- (c) Collection of data through questionnaires is free from the bias of the interviewer. (True or False)

প্ৰশ্নাৱলীৰ জৰিয়তে তথ্য সংগ্ৰহ পদ্ধতিত সাক্ষাৎকাৰ লণ্ডতাৰ পক্ষপাতমুক্ত।
(সঁচা/মিছা লিখা)

- (d) What is editing in processing and analysis of data ?

ডাটা প্ৰক্ৰিয়াকৰণ আৰু বিশ্লেষণত সম্পাদনা কি ?

- (e) What is coding ?

কোডিং কি ?

- (f) Define unemployment rates.

নিবনুৱাৰ হাৰ নিৰ্ধাৰণ কৰা।

- (g) What is the observational method of research ?

গৱেষণাৰ পৰ্যবেক্ষণ পদ্ধতি কি ?

- (h) What is the interview method of research ?

গৱেষণাৰ সাক্ষাৎকাৰ পদ্ধতি কি ?

2. Answer the following questions : (any three) $2 \times 3 = 6$

নিম্নলিখিত প্ৰশ্নবোৰৰ উত্তৰ দিয়া : (যিকোনো তিনিটা)

- (a) What is research ?

গৱেষণা কি ?

- (b) What do you mean by research problem ?

গৱেষণাৰ সমস্যাৰ অৰ্থ কি ?

- (c) Write a note on questionnaire.

প্ৰশ্নাৱলীৰ ওপৰত এটা টোকা লিখক।

- (d) What is data processing ?

ডাটা প্ৰচেছিং কি ?

- (e) What do you mean by removal of subsidy ?

ৰাজসাহায্য আঁতৰোৱাৰ অৰ্থ কি ?

- (f) When are both quantitative and qualitative methods beneficial ?

পৰিমাণগত আৰু গুণগত দুয়োটা পদ্ধতি কেতিয়া লাভজনক ?

- (f) Write different steps of data processing. Also write a note on different types of data processing tools.

ডাটা প্রকিয়াকৰণৰ বিভিন্ন পদক্ষেপ লিখা। লগতে বিভিন্ন প্ৰকাৰৰ ডাটা প্রকিয়াকৰণৰ সঁজুলিৰ এটা টোকা লিখা।

- (g) Develop a questionnaire to collect survey data pertaining to a research problem. (Gender discriminations in private v/s government sector).

গৱেষণা সম্পৰ্কীয় জৰীপৰ তথ্য সংগ্ৰহ কৰিবলৈ প্ৰশ্নাৱলী এখন প্ৰস্তুত কৰা। (ব্যক্তিগত বনাম চৰকাৰী খণ্ডত লিংগ বৈষম্য)

- (h) What do I need to consider when doing interviews ? What are the issues, or concerns in conducting interviews ?

সাক্ষাৎকাৰ লওঁতে মই কি বিবেচনা কৰিব লাগিব ?
সাক্ষাৎকাৰ লোৱাৰ সমস্যা বা উদ্বেগবোৰ কি ?