2020

STATISTICS — GENERAL

Paper : SEC-A-1

(Statistical Data Analysis using R)

Full Marks : 80

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Answer ***question nos.*** **1** & **2** and ***any two*** questions from the rest.

1. Answer ***any fifteen*** questions : 2×15

(a) Write the output for z <– (1 : 10); z <– z[2 * 1 : 5]; z.

(b) Write a one line program to make use of the function which().

(c) What is the use of typeof() function?

(d) Create a vector using the seq() function.

(e) Explain the statement "R is an open source software".

(f) Why is it safer to use "<–" instead of "=" as an assignment operator?

(g) How to install any additional *R* package from the internet?

(h) Mention the use of lines() function.

(i) Write a program for filtering a vector using the subset() function.

(j) How can you access files on remote machines via URLs?

(k) x <– matrix (1:9, ncol = 3); y <– matrix(c (1, 2, 3, 4), nrow = 2); x[2:3, 2:3] <– y. Write the output for x.

(l) Differentiate between NA and NaN.

(m) Explain briefly the use of the argument "type =" in the function plot().

(n) Write down the use of read.csv().

(o) What is the use of defining rm(list = ls (all = TRUE)) at the start of any new *R* program?

(p) Write the output for a <– c(3, 4, 5, 6, 7, 8); mean (a[2 : 5]).

(q) What function would you use to find the 1st quartile of an *R* object?

(r) Write the output for a <– c(1, 2); b <– matrix(1:4, nrow = 2); a*b.

(s) Differentiate between rbind() and cbind().

(t) Give an example to find the variance of a vector in *R*.

2. Answer ***any six*** questions :

(a) What are the hierarchy of modes of an *R* object? Write a code to change the mode of an *R* object from numeric to character. 2+3

(b) What do you mean by indexing in *R*? Give examples of three ways you can do indexing in *R*. 2+3

(c) Let $z <- c(1, 1, 2, 3, 4, 5, 5, 6, 7, 1, 1)$. Write a program to draw a boxplot on z and mark its mean with a horizontal line. 5

(d) Write programs to demonstrate the use of all(), any() and identical() functions. 5

(e) What do you mean by replication of a vector? Write down the output for $a <- c(1, 2, 3, 4)$; $b <- c(1, 2)$; $a + b$. 2+3

(f) Explain with examples the different kinds of matrix operations. Let $a <-$ matrix(c(1,2,3,4,5,6,7,8,9), nrow = 3, ncol = 3); $z <- c(5,12,13)$; a[z%%2= =1,]. Write the output. 3+2

(g) What are the problems of dealing with "NA" in *R*? How can you overcome them? 2+3

(h) How to take care of table headers while doing any numerical operations on an entire row or column of a table? 5

3. What is an *R* object? What are the different types of *R* objects? What do you mean by the mode and attribute of an *R* object? 2+4+4

4. Write a program to draw a scatter plot in *R*. Make sure to label the x and y axes and name the plot "Scatter plot". Next carry out a linear regression analysis based on the same data as used in the scatter plot. Interpret the tabular results in brief. 10

5. (a) Discuss, in brief, the different kinds of input and output operations in *R*.

(b) Define a 3×3 matrix and find out its transpose without using the funtion t(). 6+4

———————