


RAN-2411001003040001
S.Y.B.Sc. (Sem. - III) Examination November - 2025
Computer Science : Statistical Methods And Data Analysis (Paper - MDC _ 302)
Time: 2 Hours]
[Total Marks: 50
સૂચના : / Instructions

(1) ઉપરોક્ત દર્શાવેલ વિગતો ઉત્તરવહી પર અવશ્ય લખવી.

Fill up strictly the above details on your answer book.

Seat No.:

--	--	--	--	--	--

Student's Signature

Q. 1. A. Choose correct option (any 5 out of 7)
(05)

- Which of the following is an example of a categorical variable?
 - Age
 - Gender
 - Weight
 - Height
- What is the median of the numbers 1, 2, 3, 4, 5 and 6?
 - 2
 - 3
 - 4
 - 3.5
- What is the formula for calculating the range of a data set?
 - Maximum value - Minimum value
 - (Maximum value + Minimum value) ÷ 2
 - Sum of all values ÷ number of values
 - Maximum value × Minimum value
- Which of the following is used for comments in R?
 - '//'
 - '#'
 - '/* */'
 - '<!-->'
- Which of the following syntax is used for selecting 3rd coloum elements?
 - mat[2,]
 - mat[3,]
 - mat[,3]
 - mat[,2]
- If two dice are thrown simultaneously what is the probability of getting total sum is 9?
 - $\frac{1}{6}$
 - $\frac{1}{36}$
 - $\frac{1}{9}$
 - $\frac{1}{3}$
- Which function is used to merge two data frames by a common column?
 - combine()
 - merge()
 - cbind()
 - join()

(B) Answer the following (5 out of 7)
(05)

- In a dataset, if the numbers are 2,4, 6, 8 and 10 then mean is _____ of the set.
- How can you access a specific column of a data frame?
- If mean and median of a group are 5 and 4 respectively then find the value of mode.

4. Write mean formula for ungrouped observations.
5. Write probability function for Binomial distribution.
6. What is the data type of a value like 'TRUE' in R?
7. What is the subset() function used for?

Q. 2. (A) 1. Find the value of p if mean is 23.4. (05)

class	0-8	8-16	16-24	24-32	32-40	40-48
Frequency	5	3	10	P	4	2

2. For a Binomial variate $n = 10$ and $P(x = 5) = 2P(x = 4)$ find the value of p . (05)

OR

(A) 1. Find median for following data. (05)

class	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
f	35	45	70	105	90	74	51	30

2. In the normal distribution mean $\mu = 21.5$ and s.d. $\sigma = 2.5$. Find the following values: (05)

- | | |
|------------------------------|-------------------------------|
| (i) $P\{18 \leq x \leq 25\}$ | (ii) $P\{22 \leq x \leq 28\}$ |
| (iii) $P\{x \geq 28\}$ | (iv) $P\{x \leq 18\}$ |

Q. 2. [B] 1. Find range and coefficient of range of following data. (03)

Value	3	7	10	15	23	27
Frequency	2	4	6	8	6	4

2. The mean and variance of a Binomial distribution are 15 and 6 respectively. Find the value of n, p and q. (03)

OR

Q. 2. [B] 1. Find mean, standard deviation and variance of 10 observations if, $\sum x_i^2 = 900$ and $\sum x_i = 60$ (03)

2. In a normal distribution mean $\mu = 21.5$ and $\sigma = 2.5$ then find the value of $P\{18 \leq x\}$ (03)

Q. 3. [A] 1. Explain Data Cleaning giving appropriate example (05)

2. Explain Slice Function with syntax and example. (05)

OR

[A] 1. Using "df.csv", Explain following topics. (05)

- (i) How to import data into data frame from csv and excel file formats in R?
- (ii) How to add, remove and rename variables and attributes in R?

2. How to use cut() function? Explain with example. (05)

[B] 1. How to remove duplicate data using R programming? Give an example. (03)

2. State 6 uses of R programming in Data Analysis. (03)

OR

[B] 1. List characteristics of clean data. (03)

2..Explain Select() function of R giving syntax and example? (03)

Q. 4. [A] Discuss the difference between combining vectors using `cbind()` and combining data frames using `rbind()` in R. **(05)**

OR

Explain how to calculate summary statistics such as mean, median, Q_1 , and Q_3 in R. **(05)**

[B] What is the purpose of merging data frames in R? **(03)**

OR

What is the purpose of reordering dataframes in R? **(03)**
