

Faculty of Science

Bachelor of Computer Application (B.C.A.)

Semester: V

Paper: DSE-I

Subject: Statistics and R Programming

Course Outcomes

CO. No.	Course Outcomes	Cognitive Level
CO 1	Develop a comprehensive understanding of descriptive statistics, including data classification, tabulation, and graphic representation.	U, R
CO 2	Gain proficiency in correlation and regression analysis techniques and their applications in statistical modelling.	U, R, Ap
CO 3	Acquire a solid foundation in probability theory and probability distributions, and apply these concepts to real-world scenarios.	U, Ap, C
CO 4	Enhance skills in data visualization and statistical analysis using R, including hypothesis testing and linear regression.	An, Ap, C
CO 5	Understand the significance of eigenvalues and eigenvectors in data analysis and machine learning, and apply eigenvalue decomposition techniques using R.	U, Ap, C

Credit and Marking Scheme

	Credits	Marks		Total Marks
		Internal	External	
Theory	4	40	60	100
Practical	2	40	60	100
Total	6	200		

Evaluation Scheme

	Marks	
	Internal	External
Theory	3 Internal Exams of 20 Marks (During the Semester) (Best 2 will be taken)	1 External Exams (At the End of the Semester)
Practical	3 Internal Exams (During the Semester) (Best 2 will be taken)	1 External Exams (At the End of the Semester)



Content of the Course Theory

No. of Lectures (in hours per week): 2 Hrs. per week

Total No. of Lectures: 60 Hrs.

Maximum Marks: 60

Units	Topics	No. of Lectures
Unit-I	Descriptive Statistics Sampling Techniques – Data Classification – Tabulation – Frequency and graphic Representation – Measures of Central Tendency – Measures of Variation – Quartiles and Percentiles – Moments - Skewness and Kurtosis.	10
Unit-II	Correlation and Regression Scatter Diagram – Karl Pearson's Correlation Coefficient – Rank Correlation – Correlation Coefficient for Bivariate Frequency Distribution – Regression Coefficients – Fitting of Regression Lines.	10
Unit-III	Probability Theory and Statistical Inference Random Experiment – Sample Space – Events – Axiomatic Definition of Probability – Addition Theorem – Multiplication Theorem – Baye's Theorem - Applications. Statistical Inference- Hypothesis Testing, Estimation, Test of Significance for large samples (Standard Error of Mean), Test of Significance for small samples (t and Z test), Chi-square test and Goodness of Fit, Variance Ratio test (F-test)	10
Unit-IV	Data Visualization and Statistical Analysis in R Basic Plotting: Creating histograms, bar plots, box plots, and scatter plots using base R. Statistical Analysis in R: Implementing descriptive statistics, hypothesis testing (ANOVA) and linear regression analysis using R.	15
Unit-V	Eigenvalues and Eigenvectors: Calculating eigenvalues and eigenvectors, their significance in data analysis and machine learning, application of eigenvalue decomposition in R.	15

References

Text Books:

- Statistical Methods by S.P. Gupta
- Introductory Statistics with R by Peter Dalgaard
- The R Book by Michael J. Crawley

Reference Books:

- Statistical Methods" by N.G. Das
- Statistical Analysis Handbook" by Dr. Michael J. De Smith
- R for Data Science" by Hadley Wickham and Garrett Grolemund
- An Introduction to Statistical Learning" by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani

List of Practical

1. WAP in R to perform t-test.
2. WAP in R to perform Chi-square test.
3. WAP in R to perform One Way ANOVA.
4. WAP in R to perform Two Way ANOVA.
5. WAP in R to create basic plots (histogram, bar plot, box plot, scatter plot).
6. WAP in R to perform correlation analysis.
7. WAP in R to perform linear regression analysis.
8. WAP in R to manipulate and clean a dataset (subsetting, filtering, merging).
9. WAP in R to calculate eigenvalues and eigenvectors of a matrix.
10. WAP in R to perform principal component analysis (PCA) using eigenvalue decomposition.



Four handwritten signatures in black ink, arranged horizontally. The signatures are stylized and cursive, with varying lengths and orientations.