

R FOR PROBABILISTIC MODELING AND REASONING LABORATORY

| III Semester: CSE(AI & ML) | | | | | | | | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|------------------------------|---|------------------------------|---------|---------------|-------------------------|-------|
| Course Code | Category | Hours / Week | | | Credits | Maximum Marks | | |
| ACAC02 | Core | L | T | P | C | CIA | SEE | Total |
| | | 1 | 0 | 2 | 2 | 30 | 70 | 100 |
| Contact Classes: 12 | | Tutorial Classes: Nil | | Practical Classes: 24 | | | Total Classes:36 | |
| Prerequisite: Probability and Statistics | | | | | | | | |
| <p>I. COURSE OVERVIEW: This course will expose the students to R programming environment, introduces to sampling and exploring data. It also provides a foundation in both probability theory and mathematical statistics and provides an indication of the relevance and importance of the theory in solving practical problems in the real world.</p> <p>II. COURSE OBJECTIVES: The students will try to learn:</p> <ol style="list-style-type: none"> I. How to manipulate data within R and to create simple graphs and charts used in introductory statistics. II. The given data using different distribution functions in R. III. The hypothesis testing and calculate confidence intervals; perform linear regression models for data analysis. IV. The relevance and importance of the theory in solving practical problems in the real world. <p>III. COURSE SYLLABUS:</p> <p>Week – 1: INTRODUCTION TO COMPUTING</p> <ol style="list-style-type: none"> a. Installation of R b. The basics of R syntax, workspace c. Matrices and lists d. Subsetting e. System-defined functions; the help system f. Errors and warnings; coherence of the workspace <p>Week – 2: GETTING USED TO R: DESCRIBING DATA</p> <ol style="list-style-type: none"> a. Viewing and manipulating Data b. Plotting data c. Reading the data from console, file (.csv) local disk and web d. Working with larger datasets <p>Week – 3: SHAPE OF DATA AND DESCRIBING RELATIONSHIPS</p> <ol style="list-style-type: none"> a. Tables, charts and plots. b. Univariate data, measures of central tendency, frequency distributions, variation, and Shape. c. Multivariate data, relationships between a categorical and a continuous variable, d. Relationship between two continuous variables – covariance, correlation coefficients, comparing multiple correlations. e. Visualization methods – categorical and continuous variables, two categorical variables, two continuous variables. <p>Week – 4: PROBABILITY DISTRIBUTIONS</p> <ol style="list-style-type: none"> a. Sampling from distributions – Binomial distribution, normal distribution b. tTest, zTest, Chi Square test c. Density functions d. Data Visualization using ggplot – Box plot, histograms, scatter plotter, line chart, bar chart, heat maps | | | | | | | | |

Week – 5: EXPLORATORY DATA ANALYSIS

Demonstrate the range, summary, mean, variance, median, standard deviation, histogram, box plot, scatter plot using population dataset.

Week – 6: TESTING HYPOTHESES

- a. Null hypothesis significance testing
- b. Testing the mean of one sample
- c. Testing two means

Week – 7: PREDICTING CONTINUOUS VARIABLES

- a. Linear models
- b. Simple linear regression
- c. Multiple regression
- d. Bias-variance trade-off – cross-validation

Week – 8: CORRELATION

- a. How to calculate the correlation between two variables.
- b. How to make scatter plots.
- c. Use the scatter plot to investigate the relationship between two variables

Week – 9: TESTS OF HYPOTHESES

- a. Perform tests of hypotheses about the mean when the variance is known.
- b. Compute the p-value.
- c. Explore the connection between the critical region, the test statistic, and the p-value

Week – 10: ESTIMATING A LINEAR RELATIONSHIP

Demonstration on a Statistical Model for a Linear Relationship

- a. Least Squares Estimates
- b. The R Function lm
- c. Scrutinizing the Residuals

Week – 11: APPLY-TYPE FUNCTIONS

- a. Defining user defined classes and operations, Models and methods in R
- b. Customizing the user's environment
- c. Conditional statements
- d. Loops and iterations

Week – 12: STATISTICAL FUNCTIONS IN R

- a. Write Demonstrate Statistical functions in R
- b. Statistical inference, contingency tables, chi-square goodness of fit, regression, generalized linear models, advanced modeling methods.

IV. REFERENCE BOOKS:

1. Sandip Rakshit, “Statistics with R Programming”, McGraw Hill Education, 2018.
2. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, “AN Introduction to Statistical Learning: with Applications in R”, Springer Texts in Statistics, 2017.
3. Joseph Schmuller, “Statistical Analysis with R for Dummies”, Wiley, 2017.
4. K G Srinivasa, G M Siddesh, Chetan Shetty, Sowmya B J, “Statistical Programming in R”, Oxford Higher Education, 2017.

V. WEB REFERENCES

1. www.oikostat.ch
2. <https://learningstatisticswithr.com/>
3. <https://www.coursera.org/learn/probability-intro#syllabus>
4. <https://www.isibang.ac.in/~athreya/psweur/>