



St. Xavier's College – Autonomous
Mumbai
TYBSc

Syllabus for Applied Component in Statistics

5th Semester

(June 2021 onwards)

**THIS COURSE IS OFFERED TO STUDENTS WHO HAVE TAKEN
STATISTICS UPTO THE 4TH SEMESTER**

Contents:

Syllabus (Theory and Practical) for Courses:

SSTA05AC - Statistical computing using R

SSTA05ACPR

Academic/field/industrial visits and seminars may be organized by the Department, at other venues, as part of the curriculum.

T.Y.B.Sc. Statistics (Applied Component)

Course: SSTA05AC

Title: Statistical computing using R

Learning objectives

To enable students to:

1. Install, update & learn the preliminaries of R - software
2. Input, Access & Index data.
3. Save, Store & Retrieve files.
4. Import, Review, Manipulate & Summarize a variety of data formats into R using RStudio
5. Use R-programming for statistical analysis.
6. Explore different ways of Data Visualization.

Number of lectures: 60 Lectures

Unit 1

(15 Lectures)

Introduction to R Programming and Data Management

R Programming- History, programming rules, packages in R and using RStudio as IDE

Importing data in R, checking data features using functions - dim, str, head, tail etc.

Sorting of data, merging two data sets, data aggregation, creating subsets.

Numeric and String Functions

Data Cleaning- Removing duplicate records, making data consistent and treating missing values using simple methods

Unit 2

(15 Lectures)

Descriptive Statistics and Data Visualization

Measures of Central Tendency & Dispersion

Measures of Skewness & Kurtosis

Impact of missing data on R functions to calculate above measures

Box-Whisker Plot and Histogram

Bar Charts and Pie chart

Heat Map

Unit 3

(15 Lectures)

Standard Distributions and Statistical Inference using R

Probability calculations for Binomial, Poisson, Normal and Lognormal distributions

Generating random observations from specific distribution

Independent samples t test and paired t test

F test for two variances

Analysis of Variance-One way and two way

Unit 4

(15 Lectures)

Bivariate Data Analysis in R

Scatter Plot

Analysis of Correlation, Simple Linear Regression-Estimation of parameters, hypothesis testing, R squared, Residual Analysis and predictions for out of sample data.

Fitting of Curves

Cross Table for two categorical variables
Chi Square test for independence of attributes

List Of Recommended Reference Books:

1. Maria. L. Rizzo: Statistical Computing with R (2007), (Chapman & Hall/CRC), 2nd edition.
2. S.G. Purohit, S.D. Gore, S.R. Deshmukh: Statistics using R (Narosa Publishing), 2nd edition.
3. Garrett Golemund, Hadley Wickham: R for Data Science, 1st edition
4. Hadley Wickham: Advanced R, 2nd edition
5. College Statistics in R: <http://www.r-tutor.com/elementary-statistics>
6. Google R Style Guide: <https://google.github.io/styleguide/Rguide.xml>

Topics for practicals:

- Importing data in R & Data pre-processing
- Descriptive Statistics
- Graphs & Diagrams
- Standard probability distributions
- Testing of Hypothesis
- Bivariate data analysis

Course Outcomes:

At the end of course the students will be able to:

1. Be confident of using R and R-Studio.
2. Develop their own simple programs in R.
3. Create and edit visualizations with R.
4. Appreciate and apply R programming from a statistical perspective.

EVALUATION

Template of Theory Question paper

Theory: 100 Marks

CIA I – 20 marks, 45 mins.

CIA II - 20 marks, 45 mins.

End Semester exam – 60 marks, 2 hours.

Template of Practical Question paper

Practical: 50 Marks

Practical Examination - 45 marks

Journal: 05 marks