Stat-315 Data Processing Using Software II

50 Marks: 02 Credits

Number of Class: 20-26

Introduction: R versus STATA, R and the window system, downloading R, installing R, using R: a simple R session with some basic R commands, case-sensitivity, recall and correction of previous commands, assignments and expressions, simple manipulations of numbers, getting help with existing R functions and features, executing commands from a file, diverting output to a file.

STATA: Concept about the different windows in STATA, Converting data into STATA format from other formats by StatTransfer, importing and exporting data, data entry, data cleaning, data management: imputing, editing data, creating and changing variables, saving and reusing data, data reorganization, data merging and appending, basic STATA Commands and comparing them with R, STATA command for different probability distributions.

Vectors: Generating regular sequences, creating vectors, vector arithmetic, logical vectors character vectors, missing values, selecting and modifying subsets of vector, combining elements of two vectors.

Matrices: Creating matrices, matrix operations, selecting and modifying subsets of matrix, linear equations and inversion of matrices, eigen values and eigen vectors, combining rows (case) and columns (variables) of two matrices.

List and data frames: Making list and data frames, attaching and detaching data frames.

Probability Distributions: R names for different distributions, obtaining densities, cumulative probabilities, quantiles and random samples from different distributions.

Control Statements in R: Conditional execution with if statement, repetitive execution with for, repeat and while.

Writing new R functions: Simple examples, arguments and defaults, assignments within functions, returning multiple objects as output.

Importing data in R: Reading text files with read, table and scan functions, importing data from other systems like SAS, SPSS, S-Plus, Excel, Stata, editing data.

Application in Statistics: Basic statistical techniques, graphs, correlation and regression, estimation of parameters of multiple regression model , inference in multiple regression, partial correlation, multiple correlation and related tests, model selection, fitting polynomial regression , examination of residuals, outliers influential points in both STATA and R.

References

1. R Reference Manuals.
2. Stata Manuals, Stata Press.
3. Venables, W.N. and Smith, D. M. , An Introduction to R