

SEMESTER-II

COURSE 4: DESCRIPTIVE STATISTICS

Theory

Credits: 3

3 hrs/week

Course Learning Outcomes: Students will acquire:

- knowledge of Statistics and its implementation through practical understanding for various domains related to data science.
- knowledge of various types of data, their organization and evaluation of summary measures such as measures of central tendency and dispersion etc.
- knowledge of other types of data reflecting quality characteristics including concepts of independence and association between two attributes,
- insights into preliminary exploration of different types of data.
- Knowledge of correlation, regression analysis, regression diagnostics, partial and multiple correlations.

UNIT I:

Introduction to Statistics: Importance of Statistics. Scope of Statistics in different fields. Concepts of primary and secondary data. Diagrammatic and graphical representation of data: Histogram, frequency polygon, Pie. Measures of Central Tendency: Mean, Median, Mode, Geometric Mean and Harmonic Mean. Median and Mode through graph.

UNIT II:

Measures of Dispersion: Range, Quartile Deviation, Mean Deviation and Standard Deviation, Variance. Central and Non-Central moments and their interrelationship. Skewness and kurtosis.

UNIT III:

Curve fitting: Bi-variate data, Principle of least squares, fitting of degree polynomial. Fitting of straight line, Fitting of Second degree polynomial or parabola, Fitting of power curve and exponential curves.

Correlation: Meaning, Types of Correlation, Measures of Correlation: Scatter diagram, Karl Pearson's Coefficient of Correlation, Rank Correlation Coefficient (with and without ties), Bi-variate frequency distribution, correlation coefficient for bi-variate data and simple problems. Concept of multiple and partial correlation coefficients (three variables only) and properties

UNIT IV:

Regression : Concept of Regression, Linear Regression: Regression lines, Regression coefficients and its properties, Regressions lines for bi-variate data and simple problems. Correlation vs regression, sigmoid curve, derivation from linear regression to logistic regression.

UNIT-V

Attributes : Notations, Class, Order of class frequencies, Ultimate class frequencies, Consistency of data, Conditions for consistency of data for 2 and 3 attributes only, Independence of attributes, Association of attributes and its measures, Relationship between association and colligation of attributes, Contingency table: Square contingency, Mean square contingency, Coefficient of mean square contingency,

TEXT BOOKS:

1. V.K.Kapoor and S.C.Gupta: Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.
2. BA/BSc I year statistics - descriptive statistics, probability distribution - Telugu Academy - Dr M. Jaganmohan Rao, Dr N. Srinivasa Rao, Dr P. Tirupathi Rao, Smt. D. Vijayalakshmi.
3. K.V.S. Sarma: Statistics Made Simple: Do it yourself on PC. PHI

REFERENCE BOOKS:

1. Willam Feller: Introduction to Probability theory and its applications. Volume –I, Wiley
2. Goon AM, Gupta MK, Das Gupta B : Fundamentals of Statistics , Vol-I, the World Press Pvt.Ltd., Kolakota.
3. Hoel P.G: Introduction to mathematical statistics, Asia Publishinghouse.
4. M. Jaganmohan Rao and Papa Rao: A Text book of Statistics Paper-I.
5. Sanjay Arora and Bansi Lal: New Mathematical Statistics: Satya Prakashan , New Delhi

SEMESTER-II

COURSE 4: DESCRIPTIVE STATISTICS

Practical

Credits: 1

2 hrs/week

List of the experiments:

1. Graphical presentation of data (Histogram, frequency polygon).
2. Diagrammatic presentation of data (Bar and Pie).
3. Computation of measures of central tendency (Mean, Median and Mode)
4. Computation of measures of dispersion (Q.D, M.D and S.D)
5. Computation of non-central, central moments, μ_1 and μ_2 for ungrouped data.
6. Computation of Karl Pearson's coefficients of Skewness and Bowley's coefficients of Skewness.
7. Fitting of straight line by the method of least squares
8. Fitting of parabola by the method of least squares
9. Fitting of power curve of the type by the method of least squares.
10. Fitting of exponential curve of the type and by the method of least squares.
11. Computation of correlation coefficient and regression lines for ungrouped data.
12. Computation of correlation coefficient, forming regression lines for grouped data