

## Statistics I

**Course Title:** Statistics I

**Course No:** STA169

**Nature of the Course:** Theory + Lab

**Semester:** II

**Full Marks:** 60 + 20 + 20

**Pass Marks:** 24 + 8 + 8

**Credit Hrs:** 3

**Course Description:** This course contains basics of statistics, descriptive statistics, probability, sampling, random variables and mathematical expectations, probability distribution, correlation and regression.

**Course Objectives:** The main objective of this course is to impart the knowledge of descriptive statistics, correlation, regression, sampling, theoretical as well as applied knowledge of probability and some probability distributions.

### **Course Contents:**

#### **Unit 1: Introduction (4 Hrs.)**

Basic concept of statistics; Application of Statistics in the field of Computer Science & Information technology; Scales of measurement; Variables; Types of Data; Notion of a statistical population

#### **Unit 2: Descriptive Statistics (6 Hrs.)**

Measures of central tendency; Measures of dispersion; Measures of skewness; Measures of kurtosis; Moments; Stem and leaf display; five number summary; box plot

Problems and illustrative examples related to computer Science and IT

#### **Unit 3: Introduction to Probability (8 Hrs.)**

Concepts of probability; Definitions of probability; Laws of probability; Bayes theorem; prior and posterior probabilities

Problems and illustrative examples related to computer Science and IT

#### **Unit 4: Sampling (3 Hrs.)**

Definitions of population; sample survey vs. census survey; sampling error and non sampling error; Types of sampling

#### **5. Random Variables and Mathematical Expectation (5 Hrs.)**

Concept of a random variable; Types of random variables; Probability distribution of a random variable; Mathematical expectation of a random variable; Addition and multiplicative theorems of expectation

Problems and illustrative examples related to computer Science and IT

**Unit 6: Probability Distributions (12 Hrs.)**

Probability distribution function, Joint probability distribution of two random variables; Discrete distributions: Bernoulli trial, Binomial and Poisson distributions; Continuous distribution: Normal distributions; Standardization of normal distribution; Normal distribution as an approximation of Binomial and Poisson distribution; Exponential, Gamma distribution

Problems and illustrative examples related to computer Science and IT

**Unit 7: Correlation and Linear Regression (7 Hrs.)**

Bivariate data; Bivariate frequency distribution; Correlation between two variables; Karl Pearson's coefficient of correlation( $r$ ); Spearman's rank correlation; Regression Analysis: Fitting of lines of regression by the least squares method; coefficient of determination

Problems and illustrative examples related to computer Science and IT

**Laboratory Works:**

The laboratory work includes using any statistical software such as Microsoft Excel, SPSS, STATA etc. whichever convenient using Practical problems to be covered in the Computerized Statistics laboratory

**Practical problems**

| S. No. | Title of the practical problems   | No. of practical problems |
|--------|---|---------------------------|
| 1      | Computation of measures of central tendency (ungrouped and grouped data)<br>Use of an appropriate measure and interpretation of results and computation of partition Values | 1                         |
| 2      | Computation measures of dispersion (ungrouped and grouped data) and computation of coefficient of variation.  | 1                         |
| 3      | Measures of skewness and kurtosis using method of moments, Measures of Skewness using Box and whisker plot.   | 2                         |
| 4      | Scatter diagram, correlation coefficient (ungrouped data) and interpretation. Compute manually and check with computer output.  | 1                         |
| 5      | Fitting of lines of regression (Results to be verified with computer output)  | 1                         |
| 6      | Fitting of lines of regression and computation of correlation coefficient, Mean residual sum of squares, residual plot.   | 1                         |
| 7      | Conditional probability and Bayes theorem   | 3                         |
| 8      | Obtaining descriptive statistics of probability distributions   | 2                         |
| 9      | Fitting probability distributions in real data (Binomial, Poisson and Normal)   | 3                         |
|        | <b>Total number of practical problems</b>   | <b>15</b>                 |

**Text Books:**

1. Michael Baron (2013). Probability and Statistics for Computer Scientists. 2<sup>nd</sup> Ed., CRC Press, Taylor & Francis Group, A Chapman & Hall Book.
2. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, & Keying Ye (2012). Probability & Statistics for Engineers & Scientists. <sup>th</sup> Ed., Printice Hall.