Basic Statistics

Course Title: Basic Statistics　　　　　　　　　　**Full Marks:** 60 + 20 + 20
Course No: STA154　　　　　　　　　　　　　　**Pass Marks:** 24 + 8 + 8
Nature of the Course: Theory + Lab　　　　　　　**Credit Hrs:** 3
Semester: II

Course Description:
The course familiarizes students with the basic concepts of statistics including introduction, diagrammatical and graphical representation, descriptive statistics, probability, random variables, sampling, and correlation and regression.

Course Objective:
To impart the knowledge of descriptive statistics, correlation, regression, concept of sampling and sampling distribution, theoretical as well as applied knowledge of probability and some probability distributions.

Course Contents:
Unit 1: Introduction (5 Hrs.)
Basic concept of statistics; Application of Statistics in different fields including information technology; Scales of measurement; Variables; Types of Data and data source; Data preparation-editing, coding, and transcribing.

Unit 2: Diagrammatical and Graphical Presentation of Data (3 Hrs.)
Bar diagrams; Pie diagrams; Pareto chart; Graph of frequency distribution

Unit 3: Descriptive Statistics (7 Hrs)
Measures of central tendency; Measures of dispersion; Measures of skewness; Measures of kurtosis; Moments; Steam and leaf display; five number summary; box plot.
Problems and illustrative examples related to IT

Unit 4: Introduction to Probability (7 Hrs.)
Concepts of probability; Definitions of probability; Laws of probability; Bayes theorem; prior and posterior probabilities
Problems and illustrative examples related to IT

Unit 5: Random Variables and Mathematical Expectation (3 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(without proof).
Problems and illustrative examples related to IT

Unit 6: Probability Distributions (6 Hrs.)

Probability distribution function; Binomial distribution; Poisson distribution; Normal distribution and their characteristic features. Applications of these distributions in IT related data problems.

Problems and illustrative examples related to computer Science and IT

Unit 7: Sampling and Sampling Distribution (7 Hrs.)

Definitions of population; sample survey vs. census survey; sampling error and non sampling error; Types of sampling; Standard error of mean; standard error of proportion; sampling distribution of mean and proportion; Need of inferential Statistics; Concept of estimation; confidence interval estimation for mean and proportion.
Problems and illustrative examples related to IT

Unit 8: 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 IT

Laboratory Works:
Practical (Computational Statistics):
Practical problems to be covered in the Computerized Statistics laboratory

Practical problems

S. No.	Title of the practical problems (Using any statistical software such as Microsoft Excel, SPSS, STATA etc. whichever convenient).	No. of practical problems
1	Diagrammatical and graphical presentation of data	1
2	Computation of measures of central tendency (ungrouped and grouped data) Use of an appropriate measure and interpretation of results and computation of partition Values	1
3	Computation measures of dispersion (ungrouped and grouped data) and computation of coefficient of variation.	1
4	Measures of skewness and kurtosis using method of moments, Measures of Skewness using Box and whisker plot.	2
5	Scatter diagram, correlation coefficient (ungrouped data) and interpretation. Compute manually and check with computer output.	1
6	Fitting of simple linear regression model (Results to be verified with computer output), Mean residual sum of squares, residual plot	1
7	Conditional probability and Bayes theorem	3
8	Problems related to Binomial, Poisson and Normal probability distributions	2
9	Problems related sampling and sampling distribution of mean and proportion, confidence interval estimation for mean and proportion	3
	Total number of practical problems	**15**

Text Books:
1. Michael Baron (2013). Probability and Statistics for Computer Scientists. 2nd 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. 9th Ed., Printice Hall.

Reference Books:
1. Douglas C. Montgomery & George C. Ranger (2003). Applied Statistics and Probability for Engineers. 3rd Ed., John Willey and Sons, Inc.
2. Richard A. Johnson (2001). Probability and Statistics for Engineers. 6th Ed., Pearson Education, India