

GUJARAT TECHNOLOGICAL UNIVERSITY

Master in Computer Application

Year II – (Semester-IV) (W.E.F. January 2017)

Subject Name: Basic Statistics

Subject Code: 3640001

Objectives: To understand and apply various concepts, techniques and methods used in Descriptive Statistics and Inferential Statistics. The knowledge and skills gained will equip students in carrying out preliminary Data Analytics tasks, and to prepare foundation to understand and apply the statistical techniques in various fields such as Total Quality Management, Simulation, Game Theory, Operations Research, etc in addition to Computer Science topics such as Machine Learning, Cryptography, Artificial Intelligence, Operating Systems, Data Structures and Algorithms, etc.

Prerequisites: Preliminary mathematical concepts

Contents:

- 1. Introduction to Statistics and Descriptive Statistics** [10 Lectures]
Introduction, Broad areas (classification) of Statistics; **Describing Data Visually:** Frequency Distributions and Histograms; Line Charts; Bar Charts: Pareto Chart, Scatter Plots (Degree of Association); Pie Charts;
Descriptive Statistics: Central Tendency: Mean and its Characteristics, Median and its Characteristics, Quartiles and Percentiles, Mode; **Dispersion:** Range, Variance, Standard Deviation and its Characteristics, Coefficient of Variation; Standardized Data: Chebyshev's Theorem, Outliers; Mean Absolute Deviation, Interquartile Range (IQR); **Box Plots:** Fences and Unusual Data Values;
Grouped Data: Nature, Mean and Standard Deviation, Accuracy Issues;
Skewness: Coefficient of Skewness; **Kurtosis:** Leptokurtic, Platykurtic, Mesokurtic;
Measures of Association: Correlation, Coefficient of Correlation; Correlation and Causation
- 2. Probability and Probability Distributions** [12 Lectures]
Introduction: Common Framework: Experiment, Event, Elementary Events, Sample Space; Definition of Probability; Marginal Probability; Probability of Union of Events (Addition Laws), Probability Matrix; Probability of Complement of a Union; Probability of Joint Events (General Laws of Multiplication); Conditional Probability; Mutually Exclusive Events, Independent Events; Revision of Probabilities: Bayes' Rule
Discrete Probability Distributions: Introduction, Binomial Distribution, Poisson Distribution, Applications; Overview of other Discrete Probability Distribution
Probability Distributions: Introduction, Normal Distribution, Exponential Distribution, Applications; Overview of other Continuous Probability Distribution
- 3. Sampling, Sampling Distributions and Estimation** [12 Lectures]
Types of Sampling: Random, Nonrandom; Sampling Distribution of \bar{x} ; Central Limit Theorem; z Formula for Sample Mean; Standard Error of Mean; Sampling from a Finite Population; Sampling Distribution of a Proportion, Standard Error of Proportion
Estimation for Single Population: Estimating the Population Mean using z Statistic (σ Known); Estimating the Population Mean using the z Statistic when the Sample Size is Small; Estimating the Population Mean using t Statistic (σ Unknown); Estimating the

Population Proportion; Estimating the Population Variance; Estimating Sample Size

4. One Sample Hypothesis Tests

[10 Lectures]

Introduction; Null Hypothesis, Alternate Hypothesis; Type I & Type II Errors, Testing Hypotheses about a Population Mean using z Statistic (σ Known); Using Critical Value Method to test Hypotheses, Examples; Population Mean Testing Hypotheses about a Population Mean using t Statistic (σ Unknown); Testing Hypotheses about a Proportion; Testing Hypotheses about a Variance

Overview: Statistical Inferences about Two Populations; Analysis of Variance

5. Regression

[08 Lectures]

Introduction, Simple Regression Analysis, Least Square Analysis to Determine the Equation of Regression Line; Residual Analysis, Using Residual to Test the Assumptions of the Regression Model; Standard Error of the Estimate; Coefficient of Determination; Hypothesis Testing for the Slope of the Regression Model; Testing the Overall Model; Using Regression to Develop a Forecasting Trend Line

Overview: Multiple Regression Model; Mathematical Transformation of Nonlinear Models

Text Books:

1. Ken Black, "Business Statistics for Contemporary Decision Making", Wiley Student Edition, 2010

Reference Books:

1. David P. Doane, Lori E. Seward, "Applied Statistics in Business and Economics" Tata McGraw-Hill, 2010
2. Anderson, Sweeney, Williams, "Statistics for business and economics", 9 th edition, Thompson Publication
3. Bharat Jhunjhunwala, "Business Statistics", first edition, S Chand, 2008
4. Richard Levin, David Rubin, "Statistics for Management", 7 th edition, PHI
5. Nabendu Pal, Sahadeb Sarkar, "Statistics-Concepts and Applications", 2 nd edition, PHI
6. J. Susan Milton & Jesse Arnold, "Introduction to Probability & Statistics: Principles & Applications for Engineering & Computing Sciences", McGraw-Hill Education
7. S P Gupta, "Statistical Methods", 30th edition, S Chand

Chapter wise coverage from the Text Books:

Unit-1: Chapters 1, 2, 3

Unit-2: Chapters 4, 5, 6

Unit-3: Chapters 7, 8

Unit-4: Chapters 9, 10, 11

Unit-5: Chapters 14, 15, 16

Accomplishment of the student after completing the course:

The student will be able to understand various algorithms and implement them in C language. More specifically, she will be able to understand and apply the concepts of sets, cross product of sets and relation, recursion, functions, hash functions, matrices, and basic algorithms related with binary tree and graphs.