Aliah University

Department of Statistics & Informatics

Detailed Syllabus
for
5-year Integrated B. Sc. - M. Sc. in Statistics & Informatics
with effect from 2014-15

Aliah University, IIA/27, Newtown, Kolkata-156

Contents

5-year Integrated B. Sc. - M.Sc. in Statistics & Informatics

1.1 Semester-I

In semester -I there are 3 departmental courses, 2 subsidiary courses and 1 compulsory course.

1.1.1 ST-101 Descriptive Statistics-I

1.1.2 Course Objectives

In this course students will learn

- Different types of data, collection and scrutiny of data, tabulation of data etc.
- Data visualization techniques with practical applications
- Various measures of central tendency and different measures of dispersion, their use in comparing variability of different distribution.
- the concept of skewness and kurtosis and their significance.

Learning Outcomes

Upon completion of this course, students will be able to:

- Have visual perception of different type of data, with help of Tables, Histogram, Charts, Graphs, etc. using Excel.
- Apply correctly a variety of descriptive statistical techniques.
- Compute various summary statistics from data
- Make decision based on measure of central tendencies and measures of dispersion

Detailed syllabus

Introduction: Nature of Statistics, Uses of Statistics, Statistics in relation to other disciplines, Abuses of Statistics. Types of Data: Concepts and population and Sample, quantitative qualitative data, discrete and continuous data. Collection of Scrutiny of Data: Primary data, Secondary data - its major sources. Complete enumeration and Sample Surveys. Scrutiny of data for internal consistency and detection of errors in recording. Ideas of cross-validation. Presentation of data: Constructions of Tables with one or more factors of classification, diagrammatic representations, frequency distributions and cumulative frequency distributions and their graphical representations, stem and leaf displays. Univariate data - Measures - Arithmetic mean, Geometric mean, Harmonic mean; Standard deviation.

Univariate data - different measures of location, dispersion, relative dispersion, skewness and kurtosis, Moments inequalities of moments, Quantiles and measures based on them - comparison with moment measures. Box Plot. Outlier Detection.

[Credit-4 (Th-3, P-1)]

1.1.3 ST-103 Probability Theory-I

Course objectives

In this course students will learn

- Various concepts of probability: definition of probability, independence of two or more events, finding probabilities
- Theorem on total probability, Bayes' theorem, Boole's inequality, Kolmogorov's inequality and related problems
- Random variables and its distributions
- Expectations and moments of a random variable
- Probability inequalities and their applications

Learning Outcomes

At the end of the course a student will be able to

- Find probability using set theory, permutations and combinations
- Solve problems related to the distribution of a random variable, cumulative distribution function etc.
- Find various moments of a distribution and their interpretations
- Solve problems related to probability inequalities

Detailed syllabus

Random Experiment: Trial, Sample Point, Sample Space, Different types of events. Definition of Probability: Classical and relative-frequency approach to probability, Kolmogorov's Axiomatic definition (detailed discussion on discrete space only), limitations of classical definition. Probability of union and intersection of events, Probability of occurrence of exactly m and atleast m events out of n events. Conditional Probability and Independence of events, Bayes' Theorem and its applications. Examples based on classical approach and repeated trials.

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (with proof), probability mass function (p.m.f.) and probability density function (p.d.f.), Expectation and Moments, Dispersion, Skewness, Kurtosis, Quantiles.

[Credit-4 (Th-4)]

1.1.4 ST-105 Linear Algebra and Numerical Analysis

Detailed syllabus

Course objectives

In this course students will learn

- Various aspects of linear algebra which will be required in various courses in this academic programme.

- Linear dependence and independence of vectors and how to check
- Interpolation using various methods and applications.
- Numerical differentiation and integration
- How to find solution of an equation numerically using various method

Learning Outcomes

At the end of the course a student will be able to

- Find rank and inverse of a matrix, determinant of square matrix
- Find solution of a system of linear equations
- Check definiteness of a quadratic form.
- Fit a polynomial using Newton's forward, backward and Lagrange's interpolation
- Compute numerical integral using Trapezoidal, Simpson's 1/3 & Weddle's rule
- Solve equation numerically.

Detailed syllabus

- **ST-105 A Linear Algebra**

  Vector Algebra: Vector spaces with real field, Basis and dimension of a vector space, Orthogonal vectors, Gram-Schmidt Orthogonalization. Matrix Algebra: Linear transformation and Matrices, Matrix operations, Elementary matrices and their uses, Rank of a matrix and related results, Inverse of a matrix, Determinants, the Sweep-out and the Pivotal Condensation methods, Characteristic roots and vectors, Quadratic forms - classification and canonical reduction.

  Systems of Linear Equations: Homogeneous and Non-homogeneous systems - conditions for solvability.

  Characteristic roots and vectors, Quadratic forms - classification and canonical reduction. Systems of Linear Equations: Homogeneous and Non-homogeneous systems - conditions for solvability.

- **ST-105 B Numerical Analysis**

  Approximation of numbers and functions, Absolute and Relative errors.

  Interpolation : Polynomial approximation, Difference Table, Newton's Forward and Backward interpolation formulae and Lagrange's general interpolation formula, Error terms.

  Numerical Differentiation and its applications.

  Numerical Integration : Trapezoidal, Simpson's 1/3 & Weddle's rule, Euler - Maclaurin's Formula

  Numerical solution of equations : method of fixed point iteration and Newton-Raphson method in one unknown, Conditions of convergence, rates of convergence. Extension of the iteration method to two unknowns.

  Stirling's approximation to factorial n.

[Credit-4 (Th-3, P-1)]

1.1.5 MA131 Mathematics-I*

This is a subsidiary course offered by Department of Mathematics. [Credit-4 (Th-4)]

1.1.6 PH131 Physics-I / EC131 Economics-I*

This is a subsidiary course offered by Department of Physics / Economics. [Credit-4 (Th-4)]

1.1.7 EN131 Common English**

This is a compulsory course offered by English Department. [Credit-4 (Th-4)]

1.2 Semester-II

In semester -II there are 3 departmental courses, 2 subsidiary courses and 1 compulsory course.

1.2.1 ST-102 Descriptive Statistics-II

1.2.2 Course Objectives

In this course students will learn

- Relationship between two variables and visualization using scatter diagram
- Usefulness of correlation and significance of its study and the interpretations
- To explain some real phenomena using some linear model and model fitting using linear regression
- Differentiate between variables and attributes.

Learning Outcomes

Upon completion of this course, students will be able to:

- Calculate correlation coefficient and interpret the index.
- Calculate correlation ratio, correlation index and interpret the result.
- Fit linear model using linear regression and interpret the model parameters.
- Develop the analytical understanding of cause and effect of regression models.
- Fit linear model when the response variable is a character variable.

Detailed syllabus

Bivariate data - scatter diagram, correlation coefficient and its properties, Concept of Regression, Principle of Least Squares, Fitting of polynomial and exponential curves.

Correlation ratio, Correlation index, Intra-class correlation, Rank correlation - Spearman's and Kendall's measures. Analysis of Categorical Data: Consistency of data, independence and association of attributes, measures of association - Pearson's and Yule's measures, Goodman-Kruskal's γ.

Odds Ratio. Fitting of logit model through least squares.

[Credit-4 (Th-4)]

1.2.3 ST-104 Probability Theory - II

Course objectives

In this course students will learn

- Various univariate discrete and continuous distributions and their examples in real life and their statistical properties
- Two dimensional random variables: discrete as well as continuous type, joint, marginal and conditional distributions

Learning Outcomes

At the end of the course a student will be able to

- Solve problems related to various discrete as well as continuous distributions
- Fit a data set to a distribution such as binomial, Poisson etc.
- Solve problems related to bivariate distributions

Detailed syllabus

The c.d.f., p.m.f. and p.d.f. in bivariate case. Marginal and Conditional distributions, Independence, Conditional Expectation, Correlation and Regression.Theorems on sum and product of expectations of random variables. Probability Inequalities: Markov's & Chebyshev's inequalities.

Univariate Discrete Distributions: Uniform, Bernoulli, Hypergeometric, Binomial, Poisson, Negative Binomial, Geometric distributions and their properties.

Univariate Continuous Distributions: Rectangular, Normal (Normal approximation to the Poisson distribution), Cauchy, Gamma, Beta, Exponential, Laplace, Logistic, Pareto, Log-normal distributions and their properties. Truncated distributions.

[Credit-4 (Th-4)]

1.2.4 ST-192 Computer Application and Statistical Lab-I

Course Objectives

In this course students will learn

- Basics of computer.
- uses of Microsoft office and related hands on.
- Practical on descriptive statistics.

Learning Outcomes

At the end of this course students will be able to

- Write in word and make various formatting in word.
- Calculation of various statistics in excel and graphical representation in excel
- Make presentation in power point

Computer Application-I

Basics of Computer: Operations of a computer, Different units of a computer system like central processing unit, memory unit, arithmetic and logical unit, input unit, output unit etc., Hardware including different types of input, output and peripheral devices, Software, system and application software, number systems.

Understanding Word Processing: Word Processing Basics; Opening and Closing of documents; Text creation and Manipulation; Formatting of text; Table handling; Spell check, language setting and thesaurus; Printing of word document.

Using Spread Sheet: Basics of Spreadsheet; Manipulation of cells; Formulas and Functions; Editing of Spread Sheet, printing of Spread Sheet.

Making Small Presentation: Basics of presentation software; Creating Presentation; Preparation and Presentation of Slides; Slide Show; Taking printouts of presentation / handouts. [Credit-2 (P-2)]

Statistical Lab-I

Practical based on ST-102 & ST-104.

[Credit-2 (P-2)]

1.2.5 MA132 Mathematics-II*

This is a subsidiary course offered by Department of Mathematics. [Credit-4 (Th-4)]

1.2.6 PH132 Physics-II / EC132 Economics-II*

This is a subsidiary course offered by Department of Physics / Economics. [Credit-4 (Th-4)]

1.2.7 AI131 Elementary Arabic and Islamic Studies**

This is a compulsory course offered by Arabic department. [Credit-4 (Th-4)]

1.3 Semester-III

In semester -III there are 2 departmental courses, 4 subsidiary courses.

1.3.1 ST-201 Introduction to Sample Survey

Course objectives

In this course a student will learn

- Population, Sample and their definitions and examples
- Need for Sampling
- Planning and execution of sample surveys.
- Judgment and probability sampling schemes.
- Simple Random Sampling with and without replacement
- Stratified Random Sampling.

Learning Outcomes

At the end of the course a student will be able to

- Draw random numbers
- Organize a sample survey
- Determination of sample size
- Prepare questionnaire
- Do the analysis of the data and write report

Detailed syllabus

Introduction: Concepts of finite Population and Sample, Need for Sampling, Complete Enumeration and Sample Surveys. General Ideas: Planning and execution of sample surveys, analysis of data and reporting, Biases and Errors. Judgement and probability sampling schemes. Tables of Random numbers and their uses.

Simple Random Sampling with and without replacement, Determination of sample size in simple random sampling. Stratified Random Sampling.

[Credit-4 (Th-3, P-1)]

1.3.2 ST-203 Sampling Distributions

Course Objectives

In this course students will learn

- Distribution of sample statistic (such as, sample mean, sample variance and sample proportion).
- The pattern of variability of sample statistic will be explored in this course.
- Different properties of chi-square distribution, t-distribution and F-distribution.

Learning Outcomes

At the end of this course students will be able to

- Find the sampling distribution of various statistics.
- Explore different properties of sample statistics and its role in hypothesis testing.
- Solve problems related to bivariate distributions including bi-variate normal distribution.

Detailed syllabus

Introduction: Concepts of Random Sampling, Statistics and Sampling Distributions of Statistics. Illustrations using different distributions, reproductive properties of the distributions.

Bivariate distribution- marginal & conditional distributions, properties of Bivariate normal distribution.

Some Standard Sampling Distributions: χ^2 distribution, distribution of the mean and variance of a random sample from a normal population, t and F distributions, distributions of means, variances and correlation coefficient (null case) of a random sample from a bivariate normal population, distribution of the simple regression coefficient (for both stochastic and non- stochastic independent variable cases).

Distributions of Order Statistics and Sample Range.

Non-central-χ^2, t and F distributions.

[Credit-4 (Th-3, P-1)]

1.3.3 MA231 Mathematics-III*

This is a subsidiary course offered by Department of Mathematics.

[Credit-4 (Th-4)]

1.3.4 MA235 Mathematics-IV*

This is a subsidiary course offered by Department of Mathematics.

[Credit-4 (Th-4)]

1.3.5 PH231 Physics-III / EC231 Economics-III*

This is a subsidiary course offered by Department of Physics / Economics.

[Credit-4 (Th-4)]

1.3.6 PH261 Physics Practical -IV / EC233 Economics-IV*

This is a subsidiary course offered by Department of Physics / Economics.

[Credit-4 (P-4)]

1.4 Semester-IV

In semester -IV there are 2 departmental courses, 4 subsidiary courses.

1.4.1 ST-202 Statistical Inference-I

Course Objectives

Upon completion of this course, students will be able to

- Establish mathematical and conceptual foundation in the methods of statistical inference.
- Have idea of Point and Interval estimations and Testing of Hypothesis.
- Understand the requirements and the criterion of good estimators.
- Know the most common distributions and the exponential family.
- Know the different notions of convergence in statistics like convergence in probability, almost sure convergence and convergence in distribution.
- Be familiar with the concept of sufficiency and the likelihood principle.

Learning Outcomes

Upon completion of this course, students will be able to:

- Explain the concept of estimation of parameters.
- Calculate the problems related to point estimation and interval estimation.
- Explain the concepts of Testing of Hypotheses, (Large Sample Tests small sample test).
- Know the most important estimation methods like maximum likelihood, least square and the method of moments.
- Hypothesize various advanced statistical techniques for modelling and exploring practical situations.
- Interpret, in plain language, the application and outcomes of statistical techniques.

Detailed syllabus

Idea of Inference - Point & Interval Estimations and Testing of Hypothesis. Point estimation: Requirements of a good estimator - notions of Mean Square Error, Unbiasedness: Minimum Variance Unbiasedness and Best Linear Unbiasedness, Sufficiency, Factorization Theorem (Discrete case only), Properties of minimum variance unbiased estimators, consistent estimators and asymptotic efficiency, Cramer- Rao lower bound, Rao- Blackwell Theorem. Methods of Estimation - Moment, Least-square, Maximum Likelihood & Minimum χ^2 methods and their properties (excluding proofs of large sample properties).

[Credit-4 (Th-4)]

1.4.2 ST-292 Computer Application and Statistical Lab-II

Course Objectives

In this course students will learn

- basics of computer programming
- how to write programs in C
- practical on Estimation of parameters

Learning Outcomes

At the end of this course students will be able to

- Write algorithm using flowchart
- Write and execute simple programs in C
- estimation of model parameters using method of moments and MLE

Detailed syllabus

Computer Application-II

Operating systems, packages and utilities, Low and High level languages, Compiler, Assembler, Memory- RAM, ROM, unit of computer memory (bits, bytes etc.), Network - LAN, WAN, internet, intranet, basics of computer security, virus, antivirus, firewall, spyware, malware etc. Basics of Programming: Algorithm, Flowchart, Data, Information, Database, overview of different programming languages, frontend and backend of a project, variables, control structures, arrays and their usages, functions, modules, loops, conditional statements, exceptions, debugging and related concepts. Introduction to C programming

[Credit-2 (P-2)]

Statistical Lab-II

Practical based on ST202.

[Credit-2 (P-2)]

1.4.3 MA232 Mathematics - V*

This is a subsidiary course offered by Department of Mathematics.

[Credit-4 (Th-4)]

1.4.4 MA236 Mathematics - VI*

This is a subsidiary course offered by Department of Mathematics.

[Credit-4 (Th-4)]

1.4.5 PH232 Physics - V / EC232 Economics-V*

This is a subsidiary course offered by Department of Physics / Economics.

[Credit-4 (Th-4)]

1.4.6 PH262 Physics Practical - VI / EC234 Economics-VI*

This is a subsidiary course offered by Department of Physics / Economics.

[Credit-4 (P-4)]

1.5 Semester-V

In semester -V there are 5 departmental courses and 1 compulsory course.

1.5.1 ST-301 Probability Theory - III

Course objectives

In this course students will learn

- Generating Functions: Probability generating function and moment generating function in the univariate and bivariate cases.
- Characteristic function- Inversion theorem and Continuity theorem.
- Convergence of sequence of random variables: Almost sure convergence, Convergence in probability, Convergence in distribution, mean-square convergence.
- Weak law and Strong law of large numbers.
- Central limit theorem for iid random variables and applications.

Learning Outcomes

At the end of the course a student will be able to

- Understand the role of generating functions in probability and to solve related problems.
- Derive the distribution of sum and average of independent random variables.
- Identify a distribution using generating functions.
- Solve problem related to probability inequalities.
- To check convergence of a sequence of random variables.
- Derive the asymptotic distributions.

Detailed syllabus

Generating Functions: Probability generating function and moment generating function in the univariate and bivariate cases. Characteristic function- Inversion theorem and Continuity theorem (statement only)

Sequences of random variables, Almost sure convergence. Convergence in probability. Convergence in distribution.

Borel-Cantelli lemma, Independence. Weak law and Strong law of large numbers. Kolmogorov inequality. Central limit theorem for iid random variables and applications.

[Credit-4 (Th-3, P-1)]

1.5.2 ST-303 Statistical Inference II

Course Objective

This course is a continuation of statistical Inference I, that was being taught in semester IV. Students who take this course have the idea of estimation and through this course the training on statistical testing are aimed to be build.

Learning Outcomes

Upon taking the course, the students will learn the basics of testing of hypothesis, for example: test, errors, level and size of a test, concept of critical region, confidence interval and most importantly how to build a test. From this, one can make themselves ready to learn ANOVA which is offered during Semester VI, under the course Statistical Inference-III (ST-304). Through this statistical Inference course, one would be able to build proper understanding regarding the choice of the project topic, which is mandatory in Semester VI (ST-372).

Detailed syllabus

Elements of Hypothesis Testing: Null and Alternative hypotheses, Simple and Composite hypotheses, Critical Region, type I and Type II Errors, Level of Significance and Size, p-value, power.

Theory of hypothesis Testing: Most powerful (MP), Uniformly Most powerful (UMP), Randomized and Nonrandomized tests, Neyman-Pearson Fundamental Lemma (sufficiency part only), and its use in the construction of MP and UMP tests (single parameter), Uniformly Most Powerful Unbiased (UMPU) tests (definition only). [Credit-4 (Th-3, P-1)]

1.5.3 ST-305 SQC and Population Statistics

Course Objectives

1. The main objective of statistical quality control (SQC) is to achieve quality in production and service organizations, through the use of adequate statistical techniques. Different techniques of process control and product control will be explored through lecture, tutorial and assignment.

2. The objective of the Population Statistics is to focus on different measures of fertility and mortality rate. Measurement of population growth and population estimation will also be discussed through lecture, tutorial and assignment.

Learning Outcomes

1. Given sampled process data over time, students can establish control charts for monitoring processes. They also understand the concept of Operating Characteristic (OC) curve and average outgoing quality.

2. Understand the core social demographic variables, and how these variables influence population growth.

Detailed syllabus

Statistical Quality Control (SQC)

Introduction: Concept of Quality and Quality Control, Process Control and Product Control. Process Control: Control Charts and their uses, Choice of Subgroup sizes, Construction of control charts by attributes (p, c, np) (including unequal subgroup size) and variables (x, R).

Interpretation of non-random patterns of points. Product control: Producer-s risk, Consumer's risk, Acceptance Sampling plan, Single and double Sampling plan by attributes, their OC, ASN (and ATI), LTPD and AOQL. Single sampling plan for inspection by variables (one sided specification, known and unknown cases), use of IS plans and tables.

Population Statistics

Introduction : Sources of Population Data - Census data, Regression data and the errors in such data. Rates and ratios of vital events.

Measurements of Mortality: Crude Death rate, Specific Death Rate, Standardized death rate, Case fatality rate and Cause of Death Rate, Infant Mortality Rate, Neonatal and Prenatal Mortality Rates.

Life Tables: Descriptions of Complete and Abridged Life Tables and their uses, Cohort vs. Current Life Tables, Stable population and Stationary population, Construction of complete life table from population and death statistics.

Measurements of Fertility: Crude Birth Rate, General Fertility Rate, Age Specific Fertility Rate, Total Fertility Rate. Measurement of Population Growth: Crude Rate of Natural Increase and Vital Index, Gross and Net Reproduction Rates. Population Estimation, Projection and Forecasting: Use of A.P. and G.P. methods for population estimates, Fitting of Logistic curve for population forecasting using Rhode's method.

[Credit-4 (Th-3, P-1)]

1.5.4 ST-307 Data Structures and Algorithm using C

Course objectives

In this course students will learn

1. The fundamental concept of data structures and to emphasize the importance of data structures in developing and implementing efficient algorithms.
2. Construct and analysis various data structures and abstract data types including lists, stacks, queues, trees, and graphs using C.
3. Various sorting and searching algorithms.

Learning Outcomes

In completion of this course, students will be able to

1. Describe how arrays, records, linked structures, stacks, queues, trees, and graphs are represented in memory and used by algorithms.
2. Write programs that use arrays, records, linked structures, stacks, queues, trees, and graphs.
3. Discuss the computational efficiency of the principal algorithms for sorting and searching.
4. Write programs on different shortest path algorithm.

Detailed syllabus

This course introduces the fundamental concept of data structures and to emphasize the importance of data structures in developing and implementing efficient algorithms. Using C/C++, students will construct and analysis various data structures and abstract data types including lists, stacks, queues, trees, and graphs. Students will implement various sorting and searching algorithms. Students will build a substantial, complex data structure.

Tools for Analysis of Algorithms (Asymptotics, Recurrence Relations). Algorithms on arrays and matrices.

Data Structures (Linked Lists and their variants, Stacks, Queues, Trees, Heaps and some variants) and applications. Sorting, Searching and Selection (Binary Search, Insertion Sort, Merge Sort, Quick Sort, Radix Sort, Counting Sort, Heap Sort etc.. Median finding using Quick-Select, Median of Medians). Basic Graph Algorithms (BFS, DFS, strong components etc.).

Dijkstra's Shortest Paths algorithm, Bellman Ford algorithm, All pairs shortest path problem - Floyd Warshall's algorithm.

[Credit-4 (Th-4)]

1.5.5 ST-309 Economic Statistics and Official Statistics

Course Objectives

Upon completion of this course, students will be able to:

- Understand institutional, legal and organizational aspects of official statistics in India.
- Understand the functioning of official statistics.
- Interpret and use a range of index numbers commonly used in the business sector .
- Understand the basic structure of the consumer price index (CPI) and perform calculations involving its use.
- Informally define and explain terminology used to describe time series, including trend, seasonal effects, cyclical effects and outlier .

Learning Outcomes

Upon completion of this course, students will be able to:

- Describe the institutions in the India that shape monetary and fiscal policy and the processes these institutions use to achieve their goals in addressing issues of unemployment, inflation, exchange rates, balance of payments, or economic growth in a global context.
- Collect and integrate information from a variety of sources, assessing its meaning, accuracy, and timeliness; discuss the bias and variance of possible measurement and estimation procedures and recognize the agendas of points of view of various purveyors of data and analysis.
- Understand the origins and basic features of axiomatic, economic and stochastic approaches to price index.
- Apply quantitative reasoning skills to interpret and explain economic information contained in data, statistical relationships, and graphs using R, SAS, SPSS, STATA, etc. statistical software's.
- Recognize when curve fitting may be an appropriate method for modelling a time series data, identifying linear, quadratic, Gompertz and Logistic models where appropriate.

Official Statistics:

The Statistical System in India: The central and State Government organizations, the functions of the Central Statistical Organization (CSO), the National Sample Survey Organization (NSSO).

National Income Statistics: Income, expenditure and production approaches. Their applications in various sectors in India.

Economic Statistics:

Index Numbers: Price, Quantity and Value indices.

Price Index Numbers: Construction, Uses, Limitations, Tests for index numbers, Various formulae and their comparisons, Chain Index Number.

Some Important Indices: Consumer Price Index, Wholesale Price Index and Index of Industrial Production - methods of construction and uses.

Measurement of income inequality: Gini's coefficient, Lorenz curves, Application of Pareto and Log-normal as income distributions.

Time series : Components of time series, Determination of trend by fitting a mathematical curve, moving average, determination of seasonal index. Stationary time series. Auto correlation.

[Credit-4 (Th-3, P-1)]

1.5.6 ES-331 Environmental Studies**

This is a compulsory course.

[Credit-4 (Th-4)]

1.6 Semester-VI

In semester -VI there are 5 departmental courses and 1 compulsory course.

1.6.1 ST-302 Multivariate Analysis

Course Objectives

This course is the fundamental of multivariate data analysis. The exposure provided to multivariate data structure, multinomial and multivariate normal distribution, multiple correlation and regression.

Learning Outcomes

Acquire knowledge on multivariate data, multiple correlation, multiple regression and different multivariate probability distributions.

Detailed syllabus

Multivariate Data, multiple regression, multiple correlation and partial correlation-their properties and related results. Multivariate Distribution: Multinomial, Multivariate Normal Distribution and their properties.

Random Vector: Probability mass and density functions, Distribution Function, Mean vector and Dispersion matrix, Marginal and Conditional Distribution, Ellipsoid of Concentration.

Transformation of random variables.

[Credit-4 (Th-4)]

1.6.2 ST-304 Statistical Inference-III

Course Objectives

1. This course introduces students to the basic theory behind the development and assessment of statistical analysis techniques in the areas of hypothesis testing and interval estimation, as well as analysis of variance (ANOVA).

2. Test the significance related to various population parameter and compare the different population parameter.

3. Learn the basic theory of analysis of variance and apply it to one way and two way classified data.

Learning Outcomes

1. Acquire knowledge on hypothesis testing and applying it on large sample population.

2. Analyze one way and two way classified data.

3. testing simple regression coefficients, correlation ratio, linearity of simple regression, multiple correlation and partial correlation coefficients.

Detailed syllabus

Test of Significance related to a single binomial proportion and Poisson parameter; two binomial proportion and Poisson parameters; the mean(s) and variance(s) of a single univariate normal Distribution, two independent normal distribution and a single bivariate normal Distribution; regression and correlation coefficients of a single bivariate normal distribution, Combination of probabilities in tests of significance.

Analysis of Variance (ANOVA)

Introduction: Heterogeneity and Analysis of Variance and Covariance, Linear Hypothesis, orthogonal splitting of total variation, selection of Valid Error.

Applications of the ANOVA technique to: one-way classified data, two way classified data with equal number of observation per cell, testing simple regression coefficients, test for parallelism and identity, correlation ratio, linearity of simple regression, multiple correlation and partial correlation coefficients.

[Credit-4 (Th-4)]

1.6.3 ST-306 Design of Experiments - I

Course Objectives

Upon completion of this course, students will be able to:

- Discuss the fundamental principles of Design of Experiment.

- List the different types of formal experimental designs (e.g. completely randomised, randomised block, repeated measures, Latin square and factorial experimental designs).

- Understand total and partial confounding and its analysis.

- Estimate and analysis the design with missing observation.

Learning Outcomes

Upon completion of this course, students will be able to:

- Design and conduct experiments efficiently and effectively, and analyze the resulting data to obtain objective conclusions.
- Describe some of the factors affecting reproducibility and external validity.
- Describe possible causes of bias and ways of alleviating it.
- Explain the fundamental principles behind the output of an ANOVA, including "blocking" and "interactions".
- Interpret, in plain language, the application and outcomes of statistical techniques.

Detailed syllabus

Principle of Experimental Design: Randomization, Replication and Local Control, Uniformity trials, Shape and size of Plots and Blocks.

Standard Design and their Analysis: Completely Randomized Design (CRD), Randomized Block Design(RBD), Latin Square Design(LSD), comparison of efficiencies. Applications of the technique ANOVA to the analysis of the above designs. Missing Plot Technique: Analysis with one missing plot in a RBD and in a LSD.

Factorial Experiments: 2^n experiment, Advantages, Total and Partial Confounding, Analysis.

[Credit-4 (Th-4)]

1.6.4 ST-372 Computer Application-III and Project

Course objective

In this course students will learn

- Writing programs in SAS.
- Import and export datasets
- Calculate various statics including mean, median, mode, variance, standard deviation etc.
- Plot graphs in SAS
- Analyze data set using SAS
- Writing report in SAS

Learning Outcomes

At the end of the course a student will be able to

- Write and compile SAS codes
- Compute various statistics using SAS
- Analyze data using SAS
- Writing statistical report

Detailed syllabus

Statistical techniques using SAS / SPSS: introduction about SAS/SPSS, data import, export, creation, editting, printing output, arranging dataset, subsetting dataset calculation of various summary statistics, calculating frequency table, graphical representations: line diagram , pie chart, bar diagram, box plot, correlation, regression, testing of hypothesis

[Credit-4 (P-4)]

1.6.5 ST-392 Statistical Lab-III

Practical based on ST-302, ST-304 and ST-306. [Credit-4 (P-4)]

1.6.6 PS-332 Presentation Skill**

This is a compulsory course.

[Credit-4 (Th-4)]

1.7 Semester-VII

1.7.1 ST-401 Mathematical Analysis and Linear Algebra

Course objectives

In this course students will learn

- Real numbers and their convergence, limit, Continuity and uniform Continuity, Differentiability and related problems
- Riemann integral, Improper integrals, Line integral, Multiple integrals.
- Several aspects of linear algebra which are necessary to develop various statical theories
- Real quadratic forms, reduction of pair of real symmetric matrices, Singular value decomposition.
- Vector and matrix differentiation, Optimum values of a quadratic forms.

Learning Outcomes

At the end of the course a student will be able to

- Check convergence of sequence and series of real numbers
- Solve problems related to limit, continuity, differentiability of univariate and multivariate real valued functions
- Find rank of various matrices and solve system of linear equations.
- Find generalized inverses
- Find optimum values of linear form, quadratic forms, bilinear form and their examples in Statistics

Detailed syllabus

Review of Real and complex numbers; Open, closed and compact sets in R^n; Sequences and Series, Convergence, Real valued functions, Limit, Continuity and uniform Continuity, Differentiability of Univariate and Multivariate functions, Sequences and series of functions, improper integrals, Mean value theorems etc.

Uniform convergence, Power series, Taylor series. Riemann integral, Line integral, Multiple integrals.

Review of finite dimensional vector spaces (Null space and nullity), Linear dependence and independence, Matrix algebra, Rank of a Matrix, Inverse of a non-singular matrix. Hermite canonical forms, Generalized inverses, solution of linear equations, Projection and orthogonal projection matrices, Idempotent matrices. Real quadratic forms, reduction of pair of real symmetric matrices, Singular value decomposition. Optimum values of a quadratic forms, Vector and matrix differentiation.

1.7.2 ST-403 Probability Theory-IV

Course objectives

In this course students will learn

- probability as a measure.
- probability inequalities
- convergence of random variables

Learning Outcomes

At the end of the course a student will be able to

- Check field, sigma field, and related problems.
- Check whether a measurable or not and study its properties.
- Solve problems involving probability inequalities
- Check convergence of random variables.
- Find asymptotic distributions.

Detailed syllabus

Review of Probability concepts. Fields, sigma-fields and generators, semi-fields, Borel sigma-field on R. Monotone classes, monotone class theorem, pi-lambda theorem.

Measures, finite, sigma-finite measures. Measurable functions and properties. Generated sigma-fields. Induced measures. Compositions. Examples. Probability measures, properties. Independence of events, Borel-Cantelli lemmas. Fubini's theorem. Independence of random variables. Sums, variances, covariances. Kolmogorov's 0-1 law. Weak and strong laws of large numbers. Kolmogorov's inequality.

Convergence in distribution. Scheffe's theorem, Slutsky's theorem. Law of Types. Asymptotic normality. Integration of complex-valued functions, characteristic functions. Inversion and Continuity theorems. Central Limit Theorems.

1.7.3 ST-405 Regression Analysis

Course Objectives

This is one of the most important courses in the post-graduation curriculum. The course is designed to make students aware of the linear models, usual linear model assumptions, deviations from the usual linear model assumptions and regression diagnostics. Among several other components, this course aims to give students practical exposures through assignments and problem sets.

Learning Outcomes

The course makes students industry ready. It exposes them towards the deviations from Gauss Marcov set up. It will teach them the techniques for model parameter assumptions under the presence of heteoscadasticity, autocorrelation and multi-colinearity. It helps students to handle data and have a strong building block on statistical modeling. Also upon taking the course, students might consider taking the elective paper, STAT206B (Advanced Regression Analysis), which gives them huge exposure on discrete data analysis.

Detailed syllabus

General theory of regression, fitting of polynomial regression by orthogonal methods, multiple regression, examination of regression equation. Analysis of variance for fixed, random and mixed models.

Gauss-Markov set-up, estimation function BLUE and Gauss-Markov Theorem estimation and error spaces, estimation with correlated observations. Least squares estimations with restriction on parameters.

Model selection problems. Concept of best subset regression, stepwise regression, Ridge regression, Logistic regression etc.

Departures from the Gauss-Markov set-up: Heteroscedasticity, Autocorrelation and Multicolinearity. Checking for normality: Q-Q plots, Normal Probability plot, Shapiro-Wilks test. Residuals and their plots. Detection of outliers.

1.7.4 ST-407 Statistical Methodology

Course Objectives

Upon completion of this course, students will be able to:

- Know the most common distributions and the exponential family.
- Be familiar with transformation of univariate and multivariate densities.
- Know the concept of covariance and conditional probability.
- Have complete idea of truncated distribution, compound distribution and mixture of distribution.
- Know the different notions of convergence in statistics like convergence in probability, almost sure convergence and convergence in distribution.
- Be familiar with the concept of sufficiency and the likelihood principle.
- Understand Order Statistics.
- Have some knowledge of asymptotic statistics.

Learning Outcomes

Upon completion of this course, students will be able to:

- Explain the concept of estimation of parameters.
- Calculate the problems related to point estimation and interval estimation.
- Explain the concepts of Testing of Hypotheses, (Large Sample Tests small sample test).
- Know the most important estimation methods like maximum likelihood, least square and the method of moments.

- Handle a parametric hypothesis testing problem and to use the likelihood ratio method.
- Hypothesize various advanced statistical techniques for modelling and exploring practical situations.
- Interpret, in plain language, the application and outcomes of statistical techniques.

Detailed Syllabus

Brief review of basic distribution theory. Symmetric Distributions, truncated and compound distributions, mixture of distributions, Power series distribution, exponential family of distributions, Characterization of distributions (Geometric, negative exponential, normal, gamma), non-central χ^2, t and F distributions and their properties.

Parametric models, parameters, random sample and its likelihood, statistic and its sampling distributions, problems of inference. Examples from standard discrete and continuous models such as Bernoulli, Binomial, Poisson, Negative Binomial, Normal, Exponential, Gamma, Weibull, Pareto etc. Concept of sufficiency, minimal sufficiency, Neyman factorization criterion, Fisher information, exponential families. Maximum likelihood estimators, method of moment estimators. Rao Blackwell theorem, Cramer-Rao lower bound, different examples.

Approximating distributions, Delta method and its applications. Approximating distributions of sample moments, limiting moment generating function, Poisson approximation to negative binomial distribution. Order statistics-their distributions and properties. Joint and marginal distributions of order statistics. Extreme values and their asymptotic distributions (statement only) with applications. Tolerance intervals, covariance of $(X_{(r)}, X_{(s)})$.

1.7.5 ST-491 C and R Programming

Course Objectives

In this course students will learn and practice

- Various statistical computation techniques using C
- Writing and compiling programs in R, exporting and importing data sets, writing outputs in word.
- Loops and controls in R, functions in R
- Statistical computations in R

Learning Outcomes

At the end of the course a student will be able to

- Write program for statistical computations in C.
- Write and execute programs in R.
- Do statistical computations using R.

Detailed syllabus

Introduction to C language. Simple Syntax, loops, pointers, arrays, functions, files. Introduction to data structures in C Programming Language. Algorithms and corresponding C-programs for Sorting and Searching, Generation of Random Numbers, generation of samples from different theoretical distributions etc.

Introduction to R; R help; help.search(), R mailing list, contributed documentation on CRAN. Data types in R : numeric/character/logical; real/integer/complex, strings and the paste command, matrices, dataframes, lists, Creation of new variables, Creation of patterned variables, Saving workspace/history.

Writing programs in R markdown.

Graphs in R: the plot command, histogram, bar plot, box plot, points, lines, segments, arrows, inserting mathematical symbols in a plot,pie diagram, Customization of plot setting graphical parameters, adding text, saving to a file; Adding a legend.

Functions & loops in R, Programming in R.

Basic statistics using R: one and two sample t tests, Bartletts test for variance,F test for equality of variances, multi sample means, Nonparametric tests, Chi squared tests, Exact tests and confidence intervals.

Vector matrix operations: Matrix operations such as addition, subtraction, multiplication,; Linear equations and eigenvalues, matrix decomposition - LU, QR and SVD; matrix inverse.

Linear models: the lm function; ANOVA/ANCOVA/regression, models, the summary function, goodness of fit measures, predicted values and residuals.

Random no. generation & Simulations: runif, rnorm, rchisq, rt, rbinom, sample etc.; set.seed, Monte Carlo techniques.

1.7.6 ST-493 Statistical Lab-IV

Course Objectives

In this course students will do hands on

- Finding inverse, determinant, g-inverse, and rank using R
- Fitting linear model using linear regression
- Fitting best polynomial model
- Model section techniques
- Outlier detection and multicoliearity problem
- Shrinkage Methods in regression
- Estimation of model parameters using method of moments and maximum likelihood method.

Learning Outcomes

At the end of the course a student will be able to

- Find the inverse, determinant, g-inverse, and rank using R
- Fit best linear model and do the interpretation of model
- Non-linear model using polynomial regression
- Fit regression model using Ridge regression and Lasso
- Estimate model parameters using method of moments and maximum likelihood method

Practical based on ST-401, ST-403, ST-405 & ST-407.

1.8 Semester-VIII

1.8.1 ST-402 Design of Experiments - II

Course Objectives

Upon completion of this course, students will be able to:

- Discuss the fundamental principles of Design of Experiment.
- List the different types of formal experimental designs (e.g. completely randomised, randomised block, repeated measures, Latin square and factorial experimental designs).
- Understand total and partial confounding and its analysis.
- Explain Multiple comparison techniques due to Scheffe and Tukey.
- Explain various types of incomplete block design and its analysis.
- Frame mutual orthogonal latin square design and its analysis.

Learning Outcomes

Upon completion of this course, students will be able to:

1. Design and conduct experiments efficiently and effectively, and analyze the resulting data to obtain objective conclusions.
2. Make use of the basics of the Design of Experiments such as randomization and blocking
3. Describe some of the factors affecting reproducibility and external validity.
4. Identify common and important types of experimental designs with respective advantages and disadvantages (power, cost/workload).
5. Explain the fundamental principles behind the output of an ANOVA, including "blocking" and "interactions".
6. Interpret the model and report the findings scientifically.

Detailed syllabus

Review of factorial designs, fractional, factorial design. Ideas of CRD, RBD and LSD.

Split Plot Design and Strip Arrangements.

General linear hypothesis - related sampling distribution, Multiple comparison techniques due to Scheffe and Tukey. Block designs: - concepts of connectedness, orthogonality and balance; intra-block analysis - BIB: BIB design ,Youden Square design.

Construction of complete classes of mutual orthogonal Latin squares (MOLS): construction of BIBD through MOLS finite geometry and Bose's fundamental method of differences, other orthogonal designs.

1.8.2 ST-404 Database Management Systems

Course Objectives

1. The objective of the course is to present an introduction to database management systems, with an emphasis on how to organize, maintain and retrieve - efficiently, and effectively - information from a DBMS.
2. Learn the basic concepts of relational data model, entity-relationship model, relational database design, relational algebra and SQL.
3. Design ER-models to represent simple database application scenarios and convert the ER-model to relational tables, populate relational database and formulate SQL queries on data.

Learning Outcomes

1. Know the basic data models.
2. Draw ER-Diagram and convert into a RDBMS.
3. Write SQL queries from basic queries in DBMS by using standard relational operators.

Detailed Syllabus

Introduction to Databases and Transactions. Data Models: The importance of data models, Basic building blocks, Business rules, The evolution of data models, Degrees of data abstraction. Database Design, ER-Diagram and Unified Modeling Language: Database design and ER Model: overview, ER-Model, Constraints, ER-Diagrams, ERD Issues, weak entity sets, Codd's rules, Relational Schemas, Relational algebra: introduction, Selection and projection, set operations, renaming, Joins, Division, syntax, semantics, computational capabilities Constraints, Views and SQL What is constraints, types of constrains, Integrity constraints, Views: Introduction to views, data independence, security, updates on views, comparison between tables and views SQL: data definition, aggregate function, Null Values, nested sub queries, Joined relations.

1.8.3 ST-406 Parametric Statistical Inference

Course objectives

In this course students will learn

- Statistical hypothesis testing and construction of test statistics under different distributional assumptions with examples.
- construction of Likelihood ratio test with examples.
- Optimum parametric confidence intervals for large samples

Learning Outcomes

At the end of the course a student will be able to

- perform statical hypothesis testing and interpretation of results.
- do likelihood ratio test.
- Test for significance of correlation coefficient.
- perform Large sample tests

Detailed syllabus

Statistical Hypotheses-simple and composite, statistical tests, critical regions, Type-I and Type-II errors, size and power of a test, Neyman Pearson lemma and its different applications. Most powerful test, uniformly most powerful test, unbiased test and uniformly most unbiased test. Likelihood ratio test.

Likelihood Ratio tests and its applications to tests for the equality of means and variances of several normal populations. Generalized Neyman - Pearson Lemma. UMPU test or one parameter families, Locally best tests, Similar tests, Neyman structure. UMPU tests for composite hypotheses.

Confidence sets: relation with hypothesis testing. Optimum parametric confidence intervals Large Sample Theory

Delta method, Derivation of large sample standard error of sample moments, standard deviation, coefficient of variation, b_1 and b_2 measures, and correlation coefficient and their uses in large sample tests under, normality assumption, large sample distribution of sample quartile.

Transformation of statistics to stabilize variance: derivation and use of Sin^{-1}, square root, logarithmic and z transformations. Large sample tests for binomial proportions, Poisson means (single and two independent samples cases).

1.8.4 ST-408 Simulation and Data Analysis

Course Objectives

The purpose of this course is to provide students with an opportunity to develop skills in modeling and simulating data from different probability distribution. After learning the simulation techniques, the students are expected to be able to solve real world problems which cannot be solved strictly by mathematical approaches.

Learning Outcomes

Students will have ability to generate sample from different probability distributions. They can also acquire knowledge on computer intensive inference methods like Jack-Knife, Bootstrap, cross validation, Monte Carlo methods.

Detailed syllabus

Simulation of random variables from discrete, continuous, multivariate distributions and stochastic processes, Monte-Carlo methods. Regression analysis, scatter plot, residual analysis. Computer Intensive Inference Methods - Jack-Knife, Bootstrap, cross validation, Monte Carlo methods and permutation tests, Importance sampling, Metropolis Hastings Algorithm.

1.8.5 ST-4010 Applied Stochastic Processes

Course objective

In this course students will learn

- Stochastic process and deterministic processes and their distinction and examples
- Markov chains with finite and countable state space, classification of states, Chapman - Kolmogorov equation
- Calculation of n-step transition probability and its limit, stationery distribution for discrete time Markov processes
- Continuous time discrete state space stochastic process including Birth and death process, Branching process

- Distribution of population size for Continuous time Markov chain with examples

Learning Outcomes

At the end of the course a student will be able to

- Classify a stochastic process with respect to state space and time
- Derive the m-step transition probability distribution
- Will be able to estimate the transition probability matrix when the realization of the process is available.
- Classify a state and its behavior
- Explain when stationary distribution exists and able to derive the stationary distribution if exists
- Model discrete state space stochastic process and study their behavior

Detailed syllabus

Introduction to stochastic process. Markov chains with finite and countable state space, classification of states, Chapman - Kolmogorov equations. Calculation of n-step transition probability and its limit. Stationery distribution. Random walk. Discrete state space continuous time Markov chains. Poisson process, gambler's ruin problem.

Birth and death process, Galton-Watson branching process, estimation of probability of extinction. Brownian motion.

1.8.6 ST-492 Statistical Lab-V

Practical based on ST-402, ST-404, ST-406, ST-408 & ST-4010.

1.9 Semester-IX

1.9.1 ST-501 Multivariate Data Analysis

Course Objectives

Objective of this course is to focus on the statistical inference of the multivariate model. Different techniques of point estimation, testing of hypothesis and multivariate analysis of variance will be discussed through lecture.

Learning Outcomes

Acquire knowledge on Hotelling T2 and Mahalanobis's D2 and its application on multivariate testing of hypothesis. Students will be able to transfer any practical problem to a multivariate model and will be able to provide inference of the model parameters.

Detailed syllabus

Multivariate normal distribution and its properties. Sampling from Multivariate normal distribution - independence of sample mean vector and variance-covariance matrix. Wishart distribution and their properties. Distribution of quadratic forms.

Hotelling T^2 and Mahalanobis's D^2 application in testing and confidence set construction. Multivariate linear model: estimation of parameters, tests of linear hypotheses, Multivariate Analysis of variance of one and two way classified data (only LR test).

1.9.2 ST-503 Survey Sampling

Course Objective

1. The aim of this course is to cover sampling design and analysis methods that would be useful for real life decision making problem. A well designed sampling procedure ensures that we can summarize and analyze data with a minimum of assumptions and complications.
2. Learn the basics of sampling from finite populations and analyze various sampling design such as simple random sampling, stratified sampling, unequal probability sampling, and other sampling design.

Learning Outcome

1. Construct different estimators for population parameters by knowing the sampling design.
2. Learn the the differences of equal and unequal probability sampling methods.
3. Explicit uses of auxiliary variables in different sampling techniques.
4. Derive optimum sample size in a given sampling design with predefined constraints and different estimator from a given sample.

Detailed Syllabus

Probability sampling from a finite population - Notions of sampling design, sampling scheme, inclusion probabilities, Horvitz-Thompson estimator of a population total. Basic sampling schemes - Simple random sampling with and without replacement, Unequal probability sampling with and without replacement, Systematic sampling. Related estimators of population total/mean, their variances and variance estimators, Hansen-Hurwitz estimator in unequal probability sampling with replacement, Stratified sampling - Allocation problem and construction of strata.

1.9.3 ST-505 Non parametric and Bayesian Inference

Course objectives

In this course students will learn

- Statistical hypothesis testing under non-parametric setup.
- Bayesian inference and its comparison with parametric and non-parametric inference
- How to find Bays estimator under different loss function.

Learning Outcomes

At the end of the course a student will be able to

- Perform various non-parametric tests.
- Find Bayes estimator.

Detailed syllabus

Nonparametric Methods: Sign test, Mann-Whitney test, Run test, Test of randomness, Confidence limits for Quantiles based on Sign test statistic.

Bayesian Analysis: Overview and comparison of the three paradigms, classical statistics, data analysis and Bayesian analysis. Relative advantages and disadvantages.

Choice of subjective priors conjugate priors. Loss functions - squared error, absolute error and 0 - 1; reach function. Bayesian estimation of parameters.

1.9.4 ST-507 Information and Coding Theory

Course Objectives

1. To equip students with the basic understanding of the fundamental concept of entropy and information as they are used in communications.

2. To enhance knowledge of probabilities, entropy, measures of information.

3. To teach study how information is measured in terms of probability and entropy, and the relationships among conditional and joint entropies.

4. To guide the student through the implications and consequences of fundamental theories and laws of information theory and coding theory with reference to the application in modern communication and computer systems.

Learning Outcome

1. Calculate the information content of a random variable from its probability distribution.

2. Relate the joint, conditional, and marginal entropies of variables in terms of their coupled probabilities.

3. Define channel capacities and properties using Shannon's Theorems.

4. Construct efficient codes for data on imperfect communication channels.

5. Generalize the discrete concepts to continuous signals on continuous channels.

Detailed Syllabus

Structured, unstructured and Semi structured information. Capture, validation, classification, storage, retrieval, analysis, dissemination and archival of Information. Description of communication system, Mathematical definition of information, Axiomatic approach to information. Entropy and its property. Mutual information. Data compression. Channel capacity, efficiency and redundancy. Asymptotic equi-partition property. General concept of coding theory, noise and error. Universal source coding. Shannon-Fano encoding procedures. Huffman coding .Block codes and convolutional codes.

1.9.5 ST-509 Elective Paper-I

Any one from the following:

1. **ST-509A Reliability Theory**

   Course Objectives

   This course is a part of statistical methods for industrial engineering. This course aims students to make students aware of the tools and techniques that are used to cater customer satisfaction. Together with the course on statistical quality control (SQC) from under-graduation, this course has been built to make students fully aware of how manufacturing industry operates.

Learning Outcomes

Upon completion of the course, students become suitable for jobs in manufacturing industry. This course helps them to understand several manufacturing systems like coherent system, series and parallel system and many more. It will expose them to bivariate life distributions and several censoring techniques. The inferential methods for the reliability theory will also be discussed. Those who take this course, will acquire the capability of collaborating with people from industrial engineering background. This course gives the students the knowledge of systems engineering too and thereby the opportunity to work as a statistical consultant of several software companies open up.

Detailed syllabus

Reliability: concepts and measures, components and systems, coherent systems, reliability of coherent systems.

Life-distributions, reliability function, hazard rate, Mean residual life, common univariate life distributions - exponential, weibull, gamma, etc..Bivariate exponential.

Notions of ageing - IFR, IFRA, NBU, DMRL and NBUE classes and their duals, preservation of such classes under reliability operations, Loss of memory property,Partial ordering of life distributions.

Reliability estimation based on failure times from variously censored life-tests data for parametric families.

Kaplan - Meier estimation of reliability curve, Greenwood formula, Non - parametric methods for comparison of several reliability curves, Log rank tests.

Regression models in reliability, Cox PH and Accelerated failure time models; Estimation of parameters and diagnostics.

2. **ST-**509**B Statistical Techniques in data Mining**

   Introduction to Data Mining and its Virtuous Cycle.

   Cluster Analysis: Hierarchical and Non hierarchical techniques. Classification and Discriminant Analysis Tools: CART, Random forests, Fisher's discriminant functions and other related rules, Bayesian classification and learning rules.

   Dimension Reduction and Visualization Techniques: Multidimensional scaling, Principal Component Analysis, Chernoff faces, Sun-ray charts.

   Algorithms for data-mining using multiple nonlinear and nonparametric regression. Neural Networks: Multi-layer perception, predictive ANN model building using back propagation algorithm. Exploratory data analysis using Neural Networks - self organizing maps. Genetic Algorithms, Neuro-genetic model building.

3. **ST-**509**C Linear Programming**

   Formation of Linear programming problems. Convexity, Extreme points, Supporting hyper planes etc., Simplex Algorithm- Algebraic and Geometrical approaches, Artificial variable technique, Duality Theory: Fundamental theorem, Dual simplex method, Primal-dual method, Sensitivity Analysis, Bounded Variable L.P.P. Transportation Problems: Models and Algorithms, examples.

1.9.6 ST-591 Statistical Lab-VI

Practical based on ST-501, ST-503, ST-505, ST-507 & ST-509.

1.10 Semester-X

1.10.1 ST-502 Applied Multivariate Analysis

Course Objectives

Applied multivariate data analysis techniques like cluster analysis, factor analysis, principal component analysis will be explored in this course. Multivariate classification and discrimination techniques will also be discussed.

Learning Outcomes

Students will be able to reduce the dimension of multivariate data and they can summarize the data principal component analysis technique. They also can classify or cluster the data coming from different populations.

Detailed syllabus

Clustering: Hierarchical clustering for continuous and categorical data - different choices of proximity measures, Agglomerative and Divisive algorithms, K-means clustering - optimum choice of the number of clusters.

Classification and discrimination procedures: Discrimination between two known populations - Bayes, Minimax and Likelihood Ratio procedures. Discrimination between two multivariate normal populations. Sample discriminant function. Likelihood ratio rule. Probabilities of misclassification and their estimation. Principal Component Analysis: Population and sample Principal components and their uses. Plotting techniques.

Factor Analysis: The orthogonal factor model, Estimation of factor loading, Estimation of Factor scores, Interpretation of Factor Analysis.

1.10.2 ST-504 Survival Analysis

Course Objectives

Course is a part of the courses in Bio Statistics, where the aim is to expose students to several tools and techniques that are used for complete/incomplete biological experimentations. This course describes the various methods used for modelling and evaluating survival data, also called time-to-event data. It aims to expose the students to handle real life data from clinical trials through several assignments and small projects.

Learning Outcomes

Upon taking this course, Students will gain the ability to recognize the difference between parametric and non-parametric (Kaplan Meier method) survival models, estimate survival probabilities both with and without the presence of covariates (Cox PH model). They will acquire the practical exposures to handle time to event data. This helps them to make themselves suitable for pharmaceutical industries, collaborations with the clinicians and pursuing higher studies in biostatistics.

Detailed syllabus

Introduction. Basic functions and Models. Censoring and Truncation.

Parametric univariate estimation : Standard models - exponential, Weibull, log-logistic, log-normal and Gamma.

Nonparametric univariate estimation : Actuarial, Kaplan-Meier and Nelson-Aalen estimators. Semi-parametric regression models: Cox proportional hazard model - estimation, tests, diagnostics. Some ideas of Additive Models, Accelerated Models and Frailty Models.

1.10.3 ST-506 Elective Paper-II

Any one from the following:

1. **ST-506A Operations research**

   Review of LPP, Inventory models: Objective, Nature, Definitions and Scope of O.R., O.R. Models.

   Inventory Control: General Inventory model; Deterministic models - Economic Lots size models (single item) with uniform rate of demand, different rates of demand, finite rate of replenishment, Quantity discounts, with shortages. Single period EOQ model, Probabilistic EOQ model. Single period with uniform demand, single period with discontinuous demand & time independent costs. ABC analysis.

   Queuing Theory : Queuing problems. $(M/M/1)$: $FIFO$, $(M/M/C)$.

   Replacement Models and their solutions. Sequencing.

2. **STAT-506B Advanced Regression Analysis**

   Course Objective

   (a) This course covers techniques and applications of generalized linear model for different data type.

   (b) Learn association/relationship between different data type and will construct suitable GLM.

   (c) Check for over-dispersion of the model through suitable techniques.

   Learning Outcome

   (a) Learn association between nominal-ordinal, nominal-nominal and ordinal-ordinal type table.

   (b) Communicate the role of generalized linear modelling techniques (GLMs) in modern applied statistics and implement methodology.

   (c) Explain the underlying assumptions for GLMs and perform diagnostic checks whilst identifying potential problems.

   (d) Explicitly fit the model for binary and count type data.

   (e) Perform statistical analysis using statistical software, incorporating underlying theory and methodologies.

   Detailed Syllabus

   Measures of association for classified nominal and ordinal categorical data. Generalized Linear Models: Introduction, Components of a GLM, Maximum Likelihood estimation, Deviance.

   Binary data and Count data: un-grouped and grouped. Models with constant coefficient of variation. Polytomous data. over dispersion and fitting by quasi-likelihood.

3. **ST-506C Risk Analysis**

Course Objectives

- Understand the concept of decision problem and two person game.
- Basics of Non-randomized and randomized rules, Risk function, Admissibility of decision rules.
- Know the details of class of non-randomized rules for convex loss.
- Introduction to Bayes' rule and admissibility of Bayes rule.

Learning Outcomes

- Understand the main tools from probability and statistics that are used in modeling and analyzing risk in a business context.
- Become familiar with the relevant tests and tools for carrying out statistical computations.
- Become familiar with the basic concepts and methods of risk analysis.
- Become acquainted with the applications or risk analysis to several domains such as valuation with risk, financial risk minimization, supply chain management, and program management.

Detailed syllabus

Decision Problem and two - person game, Non-randomized and randomized rules, Risk function, Admissibility of decision rules, Complete, essentially complete, minimal complete and minimal essentially complete classes. Essential completeness and completeness of class of rules based on sufficient statistic and the class of non-randomized rules for convex loss. Bayes rules, Extended Bayes, Generalized Bayes and Limit of Bayes rules, Admissibility of Bayes rule. Minimax rules, Method for finding minimax rules.

1.10.4 ST-508 Elective Paper-III

Any one from the following:

1. **ST-508A Time Series & Forecasting**

   Linear stationary processes, AR, MA, ARMA and ARIMA; identification, estimation of the models; forecasting time series regression; Fourier analysis, spectral representation of a stochastic process, properties of ARMA processes in the frequency domain; estimation of the spectrum, Kalman filter.

2. **ST-508B Advanced Data Analytic Techniques**

   Course Objectives

   The main goal of this course is to help students learn, understand and practice advanced data analytic techniques, which include the study and analysis of different types of data: cross sectional and longitudinal data. It aims to expose students to several resampling techniques which are used for handling small samples.

   Learning Outcomes

   After taking this course, students are expected to be highly skilled in handling varied types of datasets and also the datasets where a substantial number of missing points are observed. They will learn Jackknife, bootstrap and other resampling techniques. Several restricted likelihood methods of estimation will also let them aware of handling critical situations. By learning the techniques, the students are expected to excel as an applied statistician (both if they choose a career for pursuing research as well as for joining industry).

Detailed syllabus

Resampling Techniques : Introduction to Jackknife and Bootstrap - methods for estimating bias, standard error and distribution function based on iid random variables, standard examples, Bootstrap confidence intervals.

Missing data analysis : Informative or non-informative missingness, complete case / available case estimation, Imputation, EM & MCEM algorithms and data augmentation techniques. Standard error estimation.

Longitudinal data analysis : Longitudinal regression, Cohort vs longitudinal effect, Weighted least-squares, ML and REML techniques.

3. **ST-**508**C Development Statistics**

Course Objectives

- Understand the role of statistics in National and international statistical systems of economic development.
- Basics of Distribution of income for measurement of poverty and inequality
- Know the details measures of unemployment.
- Basic concepts of development indices.
- Know the different indicators of development.

Learning Outcomes

- Understand the national account system and estimate national and state incomes.
- Project the future populations and the income distribution to eradicate poverty and inequality.
- Calculate different development indices which includes crop-forecasting, crop insurance, procurement, buffer stock management, foreign trade, balance of payments, planning, allocation of resources and evaluation of family welfare programmes.

Detailed syllabus

Concept of economic development - role of statistics. National and international statistical systems. National accounts - estimation of national and state incomes and their components. Projection of populations. Distribution of income - measurement of poverty and inequality. Measures of unemployment.

Development indices. Other indicators of development (includes agriculture - crop-forecasting and estimation, crop insurance, procurement and buffer stock management; foreign trade and balance of payments; planning and allocation of resources; evaluation of family welfare programmes)

1.10.5 ST-572 Dissertation

1.10.6 ST-592 Statistical Lab-VII

Practical based on ST-502, ST-504, ST-506 & ST-508.

References

Descriptive Statistics

1. Goon A.M.,Gupta M. K., Dasgupta B.(1998): Fundamentals of Statistics (V-1),World Press
2. Yule G.U & Kendall M.G. (1950): An Introduction to the Theory of Statistics, C.Griffin
3. Snedecor & Cochran (1967): Statistical Methods (6thed), Iowa State Univ. Press
4. Croxton F.E., Cowden D.J. & Klein (1969): Applied General Statistics, Prentice Hall
5. Wallis F.E. & Roberts H.V. (1957): Statistics- a new approach, Methuen
6. Tukey J.W. (1977): Exploratory Data Analysis, Addison-Wesley Publishing Co.
7. Lewis-Beck M.S. (edt.) (1993) : Regression Analysis, Sage Publications
8. A. Agresti (1984): Analysis of Ordinal Categorical Data
9. Kendall M.G. & Stuart A. (1966): Advanced Theory of Statistics (Vols 1 & 2)

Probability

1. Chung K. L. (1983): Elementary Probability Theory with Stochastic Process, Springer / Narosa
2. Feller W. (1968): An Introduction to Probability Theory & its Applications, John Wiley
3. Goon A.M., Gupta M.K. & Dasgupta B. (1994): An Outline of Statistical Theory (Vol-1), World Press
4. Rohatgi V. K. (1984): An Intro. to Probability Theory & Math. Statistics, John Wiley
5. Hoel P. J., Port S. C. & Stone C.J. ( ): Introduction to Probability Theory (Vol-1), Mifflin & UBS
6. Cramer H. (1954): The Elements of Probability Theory, John Wiley
7. Parzen E. (1972): Modern Probability Theory and its Applications, John Wiley
8. Uspensky J. V. (1937): Introduction to Mathematical Probability, McGraw Hill
9. Cacoullos T. (1973): Exercises in Probability. Narosa
10. Rahman N. A. (1983): Practical Exercises in Probability and Statistics, Griffen
11. Pitman J. (1993): Probability, Narosa
12. Stirzaker D. (1994): Elementary Probability, Cambridge University Press
13. Chandra T. K. & Chatterjee D. (2001): A First Course in Probability, Narosa
14. Bhat B. R. (1999): Modern Probability Theory, New Age International

Population Statistics

1. Goon AM,Gupta MK,Dasgupta B(2001): Fundamentals of Statistics (V-2), World Press
2. Spiegelman M. (1980): Introduction to Demography, Harvard University Press
3. Cox P.R. (1976): Demography
4. Biswas S. (1988): Stochastic Processes in Demography and Applications
5. Mishra B.D. (1980): An Introduction to the Study of Population, South Asian Pub.
6. Keyfitz. N and Caswell. H (2005): Applied Mathematical Demography (3rd edition), Springer

Linear Algebra

1. Hadley G. (1995): Linear Algebra, Addison Wesley/ Narosa
2. Rao A.R. & Bhimasankaran P. (1996): Linear Algebra
3. Searle S.R. (1982): Matrix Algebra- useful for Statistics, John Wiley
4. Rao C.R. (1974): Linear Statistical Inference & its Applications, Wiley Eastern
5. Hoffman K. & Kunze R. (2001): Linear Algebra
6. Goon A.M. (1988): Vectors and Matrices

Numerical and Mathematical methods

1. Scarborough J.B. (1958): Numerical Mathematical Analysis, Oxford Univ. Press
2. Atkinson K. (1985): Elementary Numerical Analysis
3. Sastry S.S. (1998): Introductory Methods of Numerical Analysis
4. Hildebrand F.B. (1974): Introduction to Numerical Analysis, Tata McGraw-Hill
5. Apostol T.M. (1985): Mathematical Analysis, Narosa
6. Apostol T.M. (1968): Calculus (Vols 1 & 2)
7. Goldberg R.R. (1953): Methods of Real Analysis, Oxford & IBH Pub. Co.
8. Widder D.V. (1994): Advanced Calculus
9. Piskunov N. (1977): Calculus (Vols 1 & 2)
10. Malik S.C. & Arora S. (1991): Mathematical Analysis

Probability Models and Stochastic Processes

1. Chung K.L. (1983): Elementary Probability Theory with Stochastic Process, Springer / Narosa
2. Feller W. (1968): An Introduction to Probability Theory & its Applications, John Wiley
3. Goon A.M., Gupta M.K. & Dasgupta B. (1994): An Outline of Statistical Theory (Vol-1), World Press
4. Rohatgi V.K. (1984): An Intro. to Probability Theory & Math. Statistics, John Wiley
5. Hoel P.J., Port S.C. & Stone C.J. ( ): Introduction to Probability Theory (Vol-1), Mifflin & UBS
6. Cramer H. (1954): The Elements of Probability Theory, John Wiley
7. Parzen E. (1972): Modern Probability Theory and its Applications, John Wiley
8. Uspesky J.V. (1937): Introduction to Mathematical Probability,McGraw Hill
9. Cacoullos T. (1973): Exercises in Probability. Narosa
10. Rahman N.A. (1983): Practical Exercises in Probability and Statistics, Griffen
11. Pitman J. (1993): Probability, Narosa
12. Stirzaker D. (1994): Elementary Probability, Cambridge University Press
13. Chandra T.K. & Chatterjee D. (2001): A First Course in Probability, Narosa
14. Bhat B.R. (1999): Modern Probability Theory, New Age International
15. Goon A.M., Gupta M.K. & Dasgupta B. (1994): Fundamentals of Statistics (Vol-II), World Press
16. Hoel P.J., Port S.C. & Stone C.J. ( ): Introduction to Stochastic Processes, Mifflin & UBS
17. Allen, Linda J. S.: An Introduction to Stochastic Processes with Applications to Biology
18. Ross, S. M.: Introduction to Probability Models
19. Anderson, D. F.: Introduction to Stochastic Processes with Application in Biosciences
20. Ross, S. M.: Stochastic Processes
21. Karlin, S. and Taylor, H. M.: A First Course in Stochastic Processes

Official, Economic Statistics and SQC

1. C.S.O. (1984) : Statistical System in India
2. Goon A. M.,Gupta M. K,, and Dasgupta. B. (2001): Fundamentals of Statistics (V-2),World Press
3. Yule G.U. & Kendall M.G. (1953): An Introduction to the Theory of Statistics, C. Griffin
4. Kendall M.G. & Stuart A. (1966): Advanced Theory of Statistics (Vol 3), C.Griffin
5. Croxton F.E., Cowden D.J. & Klein (1969): Applied General Statistics, Prentice Hall
6. Mudgett B.D. (1951): Index Numbers, John Wiley
7. Allen R.G.D. (1975): Index Numbers in Theory and Practice, Macmillan

8. Mukhopadhyay P. (1999): Applied Statistics
9. Johnston J. & Dinardo J. (1997):Econometric Methods, McGraw Hill
10. Nagar A.L. & Das R.K. (1976): Basic Statistics
11. Duncan A.J. (1953): Quality Control and Industrial Statistics, Richard DIrwin
12. Cowden D.J. (1957): Statistical Methods in Quality Control, Prentice Hall
13. Grant E.L. & Leavenworth (1964): Statistical Quality Control, McGraw Hill
14. Bowley A.H. & Goode H.P. (1952): Sampling Inspection by Variables, McGraw Hill
15. Ekambaram S. K. (1960): The Statistical Basis of Quality Cont. Charts, Asia Publishing House
16. Montgomery D.C. (1985): Introduction to Statistical Quality control, John Wiley
17. IS2500 Part I and Part II
18. Bureau of Indian Standards (1994): Handbook on Statistical quality Control
19. Indian Standards Institution (1982): Manual on Basic Principles of Lot Sampling

Sampling Distributions

1. Goon A.M., Gupta M.K. & Dasgupta B. (1994): An Outline of Statistical Theory (Vol-1), World Press
2. Johnson, N.I. & Kotz S. (1970): Distributions in Statistics, John Wiley
3. Ross S.M. (1972): Introduction to Probability Models, Academic Press
4. Mood A.M., Graybill F. & Boes D.C. (1974): An Introduction to the Theory of Statistics (3rd ed), McGraw Hill
5. Rao C.R. (1952): Advanced Statistical Methods in Biometric Research, John Wiley
6. Hogg R.V. & Craig A.T. (1978): Introduction to Mathematical Statistics
7. Rohatgi V.K. (1984): An Introduction to Probability Theory & Mathematical Statistics, John Wiley
8. Stuart G & Ord J.K. (1991): Advanced Theoryof Statistics (Vol 2), Charles Griffin
9. Goon A. M., Gupta M. K.and Dasgupta B. (1997): Fundamentals of Statistics (V-1), World Press
10. Bhattacharya GK & Johnson R. A. (1977): Concepts & Methods of Statistics, John Wiley

Time Series

1. Kendall M.G. (1976): Time Series, Charles Griffin
2. Chatfield C. (1980): The Analysis of Time Series -An Introduction, Chapman & Hall
3. Mukhopadhyay P. (1999): Applied Statistics
4. Johnston J. & Dinardo J. (1997): Econometric Methods, McGraw Hill

Inference and ANOVA

1. Goon A. M., Gupta M. K. & Dasgupta B. (1994): An Outline of Statistical Theory(Vol-2), World Press
2. Mood A. M., Graybill F. & Boes D. C. (1974): An Introduction to the Theory of Statistics (3rd ed), McGraw Hill
3. Rao C. R. (1952): Advanced Statistical Methods in Biometric Research, John Wiley
4. Hogg R. V. & Craig A.T. (1978): Introduction to Mathematical Statistics
5. Rohatgi V. K. (1984): An Introduction to Probability Theory & Mathematical Statistics, John Wiley
6. Stuart G & Ord J. K. (1991): Advanced Theory of Statistics (Vol 2), Charles Griffin
7. Goon A. M., Gupta M. K. and Dasgupta B. (1997): Fundamentals of Statistics (V-1 and 2), World Press
8. Bhattacharya G. K. & Johnson R. A. (1977): Concepts & Methods of Statistics, John Wiley
9. Scheffe H. (1959): The Analysis of Variance, John Wiley
10. Ross S.M. (1972): Introduction to Probability Models, Academic Press

Sample Survey

1. Goon A. M., Gupta M. K., Dasgupta B.(2001):Fundamentals of Statistics (V-2), World Press
2. Murthy M.N. (1977): Sampling Theory and Methods, Statistical Pub. Soc., Calcutta
3. Des Raj & Chandhok P.(1998): Sample Survey Theory, Narosa Publishing House
4. Cochran W.G. (1984): Sampling Techniques (3rd edition), Wiley Eastern
5. Mukhopadhyay P. (1998): Theory and Methods of Survey Sampling, Prentice Hall
6. Sukhatme P. V. & Sukhatme B. V. (1970): Sampling Theory of Surveys with, Asia Publishing House
7. Sampathy S. (2001): Sampling Theory and Methods, Narosa
8. NSSO Publications

Design of Experiments

1. Kempthorne O. (1965): The Design and Analysis of Experiments, Wiley Eastern
2. Das M. N. & Giri N. C. (1986) : Design and Analysis of Experiments. (2nd edition), Wiley Eastern
3. Montgomery D. C. (1976): Design and Analysis of Experiments, John Wiley
4. Cochran W. G. & Cox G.M. (1957):Experimental Designs, John Wiley
5. Federer W. T. (1975): Experimental Designs- Theory and Application, Oxford & IBH
6. Mukhopadhyay P. (1999): Applied Statistics

Multivariate Analysis

1. Kendall M.G. & Stuart A. (1966): Advanced Theory of Statistics (Vol 3), C.Griffin
2. Anderson T.W. (1958): An Introduction to Multivariate Statistical Analysis, 3rd edition, Wiley inter science
3. Goon A.M., Gupta M.K. & Dasgupta B. (1994): An Outline of Statistical Theory (Volumes 1 & 2), World Press
4. Rohatgi V.K. (1984): An Introduction to Probability Theory & Math. Statistics, John Wiley
5. Johnson, N.L. & Kotz S. (1970): Distributions in Statistics, John Wiley
6. Hogg R.V. & Craig A.T. (1978): Introduction to Mathematical Statistics
7. Rao C. R. (1974): Linear Statistical Inference and its Applications, John Wiley
8. Mukhopadhyay P. (1996): Mathematical Statistics
9. Johnson R. A. and Wichern, W (2001): Applied Multivariate Statistical Analysis, 5th edition, Prentice Hall

Computer Application & Programming

1. Sinha, P. K. & Sinha, P.: Computer Fundamentals
2. Kanetkar Y. P.: Let us C
3. Introductory Statistics with R, Peter Dalgaard, Springer, 2008, 2nd Edition
4. An Introduction to Statistical Learning with applications in R, Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani; Springer, 2013.
5. Data Analysis & Graphics using R, An example based approach, John Maindonald and W. John Braun; 3rd Edition, 2010, Cambridge University Press.
6. Matcher, L.: Step-by-Step Basic Statistics Using SAS: Students' Guide
7. Matcher, L.: Step-by-Step Basic Statistics Using SAS: Exercises
8. Wicklin, R.: Statistical Programming with SAS/IML Software
9. **Online Tutorial:**

   https://www.tutorialspoint.com/computer_fundamentals/

   https://www.tutorialspoint.com/cprogramming/

   https://www.tutorialspoint.com/data_structures_algorithms/

   https://www.tutorialspoint.com/r/