UNIVERSITY OF CALCUTTA

Notification No. CSR/69/2026

It is notified for information of all concerned that in terms of the provisions of Section 54 of the Calcutta University Act, 1979, (as amended), and, in the exercise of his powers under 9(6) of the said Act, the Vice-Chancellor has, by an order dated 14.05.2026, approved the new revised Syllabus including Question Patterns / Examination Modalities and PO/CO/PSO of 4-year Honours and Honours with Research and 3-year MDC courses of studies in Statistics under CCF, 2022.

The above shall take effect from the Odd semester examinations, 2026 and onwards.

19/06/2026

Prof.(Dr.) Debasis Das
Registrar

SENATE HOUSE
Kolkata-700073
19.06.2026

UNIVERSITY OF CALCUTTA

Syllabus

for

Four-year / Three-year B.Sc. Courses of Studies

in

STATISTICS

(Under Curriculum & Credit Framework, 2022)

University of Calcutta

Programme Outcomes (PO), Programme Specific Outcomes (PSO) and Course Outcomes (CO)
STATISTICS

Four-Year B.Sc. (Honours and Honours with Research) Course
(Under Curriculum and Credit Framework, 2022)

Programme Outcomes (PO):

- **PO1: Disciplinary knowledge** – Capability to demonstrate comprehensive knowledge, understanding, and application of core concepts within the field of study. It ensures students possess foundation to advanced proficiency, enabling them to analyze, evaluate, and apply theoretical knowledge to practical, real-world or research contexts.

- **PO2: Communication skills** – Focuses on enabling students to articulate ideas clearly, confidently, and ethically across written, verbal, and digital platforms. Graduates will possess the ability to produce professional documents, deliver effective presentations, utilize digital media, and engage in active listening and empathetic dialogue in diverse workplaces.

- **PO3: Critical thinking** – Enables graduates to analyze, evaluate, and synthesize information to make informed decisions, solve complex problems, and challenge assumptions. It involves applying reasoning, logical, and evidence-based approaches to understand different perspectives, fostering intellectual independence in both professional and personal contexts.

- **PO4: Problem solving** – Measures a graduate's ability to identify, analyze, and solve complex real-life or domain-specific problems using critical thinking, scientific inquiry, and modern tools. It involves defining problems, designing sustainable solutions, and interpreting data.

- **PO5: Analytical reasoning** – Focuses on developing a student's capability to objectively evaluate, interpret, and analyze evidence, arguments, and data to make informed, logical decisions. It equips graduates to identify patterns, solve complex problems, and formulate evidence-based conclusions across various fields.

- **PO6: Research related skills** – Abilities designed to enable students to gather, analyze, interpret, and apply information effectively to address complex problems and generate new knowledge. These skills are crucial for transitioning students from passive knowledge consumers to active contributors, particularly in professional and advanced academic environments.

- **PO7: Scientific reasoning** – Refers to a student's ability to analyze, interpret, and evaluate quantitative or qualitative data, experimental evidence, and scientific claims using logical, hypothetical-deductive reasoning and evidence-based thinking. It involves formulating testable hypotheses, designing and conducting experiments, controlling variables, interpreting data, distinguishing between correlation and causation, and constructing scientific knowledge to solve complex problems.

- **PO8: Reflective thinking** – Signifies a student's ability to critically analyze their own experiences, actions, and decisions to foster deep learning, self-awareness, and continuous improvement. It bridges theory and practice, ensuring learners can adapt, solve complex problems, and engage in lifelong professional growth.

- **PO9: Information and digital literacy** – Aim to empower learners to efficiently find, evaluate, manage, create, and communicate information using digital tools responsibly and ethically. Key outcomes include critical thinking, data management, and the ability to navigate digital environments for academic, professional, and civic participation.

- **PO10: Self-directed learning** – Focus on cultivating autonomous, lifelong learners capable of independently identifying learning needs, setting goals, searching for resources, and evaluating their own progress. Key outcomes include enhanced critical thinking, increased motivation, better time management, and the ability to apply learning to new situations.

- **PO11: Moral / ethical awareness** – Aim to prepare graduates to act as responsible, principled, and engaged citizens. These outcomes ensure that students can identify, analyze, and resolve ethical dilemmas while adhering to professional codes of conduct and societal values.

- **PO12: Lifelong Learning** – Ensures graduates possess the curiosity, mindset, and skills for continuous, self-directed learning. It focuses on adaptability to evolving technologies, professional growth, and the ability to independently acquire knowledge throughout a career.

Programme Specific Outcomes (PSO):

- **PSO1: Understanding of Theories and Concepts** – Aim to equip students with a solid foundational knowledge of statistical methodology, enabling them to move from theoretical comprehension to practical application. Key outcomes include the ability to recall and apply statistical principles, identify appropriate distributions, understand inferential techniques, and utilize software for modelling.

  - ***Foundation in Probability Theory*:** Students will understand fundamental probability axioms, conditional probability, random variables (discrete and continuous), and their probability distributions (e.g., normal, binomial, Poisson, gamma, exponential).

  - ***Knowledge of Statistical Inference*:** Students will gain a deep understanding of point and interval estimation (consistency, unbiasedness, efficiency), Fisher's information function, and Cramer-Rao lower bound.

  - ***Hypothesis Testing Framework*:** Students will understand the concepts of null and alternative hypotheses, Type I and Type II errors, confidence intervals, and P-values, as well as the distinction between parametric and non-parametric methods.

  - ***Modelling & Multivariate Analysis*:** Understanding the theory behind linear models, simple and multiple regression, residual analysis, and data reduction techniques like Principal Component Analysis (PCA) and Factor Analysis.

  - ***Sampling and Design Theory*:** Knowledge of sampling techniques (random, stratified, systematic) and the design of experiments (CRD, RBD, LSD, ANOVA) to ensure efficient data collection.

  - ***Theoretical Basis of Data Analytics*:** Understanding foundational concepts in modern statistical learning, such as bias-variance trade-off, over-fitting, and supervised learning algorithms.

  - ***Application-Oriented Theory*:** Ability to map real-world scenarios in finance, biostatistics, and industry (e.g., reliability, quality control, insurance) to theoretical statistical models.

- **PSO2: Analytical and Quantitative Skills** – Designed to ensure graduates can effectively manage, analyze, and interpret data to solve complex problems. These outcomes emphasize

the transition from theoretical knowledge to practical application using statistical software and mathematical modelling.

- *Quantitative Modelling Ability*: Graduates will be able to translate real-world problems into testable mathematical/statistical models, applying probability and inferential techniques to make informed decisions.
- *Data Analytics & Interpretation*: Ability to perform comprehensive Exploratory Data Analysis (EDA), applying appropriate statistical techniques (such as regression, time series analysis, or sampling) to identify patterns, trends, and relationships.
- *Statistical Software Proficiency*: Mastery in using software tools to extract, manipulate, and analyze large datasets, fostering proficiency in statistical programming.
- *Critical Evaluation & Decision Making*: Capacity to critically evaluate statistical findings, considering limitations and ethical implications, to draw valid, evidence-based conclusions.
- *Research & Experimental Design*: Skills in planning and executing statistical experiments or survey designs to collect high-quality data while minimizing bias, often applied in medical research, finance, or social sciences.
- *Machine Learning Application*: Familiarity with fundamental machine learning algorithms and their applications to enhance predictive modelling and data analytics.

➢ **PSO3: Critical Thinking and Problem-solving Ability** – Aim to equip students to analyze complex data, apply appropriate statistical models to real-world problems, interpret findings logically, and use software for data-driven decision-making. These outcomes emphasize developing an analytical approach to research and evaluating results objectively.

- ***Statistical Modelling and Application*:** Graduates will be able to translate real-world information into statistical models, selecting and applying appropriate techniques—such as regression, experimental design, or multivariate analysis—to solve problems across various fields.
- ***Critical Data Interpretation*:** Ability to analyze, evaluate, and interpret data evidence and arguments, ensuring conclusions are supported by logical reasoning rather than mere description.

- ***Problem Diagnosis and Research Design*:** Graduates can formulate testable research hypotheses, identify potential biases, and design experiments to effectively investigate open-ended problems.

- ***Technological Proficiency in Solving Problems*:** Proficiency in utilizing statistical software and programming languages to compute, analyze, and visualize data for decision-making.

- ***Analytical Reasoning & Evaluation*:** Ability to assess the validity of statistical practices, identify potential ethical issues (e.g., data fabrication), and use mathematical tools to solve complex numerical problems.

- ***Systemic Evaluation*:** Capability to evaluate the impact of decisions and interpretations on data, considering the broader context of sustainability and societal implications.

➢ **PSO4: Communication of Ideas** – Focus on the ability to translate complex numerical data into actionable insights for diverse audiences. Graduates are expected to articulate statistical findings, methodology, and interpretations accurately in both oral and written forms. It ensures that graduates can, not only analyze data, but also convey its significance efficiently.

- *Effective Data Interpretation*: Ability to communicate statistical analysis results, interpretation, and findings clearly to non-statistical audiences and stakeholders.

- *Technical Reporting*: Capability to write comprehensive, systematic, and technically accurate statistical project reports, ensuring they are well-structured.

- *Oral Presentation Skills*: Ability to present statistical findings confidently and accurately using visual aids like graphs, posters, and PowerPoint presentations.

- *Contextualizing Results*: Ability to translate numerical outcomes into logical, actionable business or research insights.

- *Communication with Stakeholders*: Capacity to give and receive clear instructions on data-related inquiries and collaborate in multidisciplinary environments.

- *Data Visualization*: Effectively organizing and presenting data using appropriate diagrams and graphs.

- *Technical Writing*: Presenting research findings and methodological details, such as experimental designs or sampling techniques, in a written format.

- *Interdisciplinary Collaboration*: Communicating technical statistical solutions to problems in areas like data science, finance, and social sciences.

➢ **PSO5: Moral / ethical awareness and social responsibility** – Focus on ensuring that graduates apply quantitative methods with integrity, transparency, and concern for the impact of their data analyses on individuals and society. Graduates will be able to evaluate the ethical implications of using specific datasets and justify the choice of statistical methods based on both technical efficiency and moral integrity. Graduates will demonstrate an understanding of how statistical tools (e.g., in biostatistics or economics) contribute to social welfare, public policy, and environmental sustainability.

- *Ethical Data Practice & Integrity*: Ability to apply statistical methodologies honestly, avoiding data manipulation or fabrication of results to fit predetermined outcomes or stakeholder pressure.
- *Confidentiality and Data Privacy*: Competence in protecting the identity and sensitive information of research subjects, ensuring compliance with legal standards and confidentiality agreements.
- *Social Responsibility in Analytics*: Awareness of the societal impact of statistical modelling, ensuring that analyses do not perpetuate bias, discriminate against marginalized populations, or cause harm.
- *Transparency in Reporting*: Commitment to documenting and communicating the limitations of data, assumptions made, potential sources of error, and algorithmic biases to ensure stakeholders understand the context.
- *Informed Consent and Ethical Stewardship*: Ensuring that data is collected, handled, and used only within the scope of informed consent and that ethical considerations are prioritized during the entire data lifecycle.
- *Professionalism*: Adherence to established ethical guidelines and fostering a culture of accountability in professional practice.

➢ **PSO6: Research skill development and Lifelong Learning** – Aim to equip students with the ability to analyze complex data, utilize modern computational tools, and adapt to evolving statistical methodologies. These outcomes prepare students for careers in academia, industrial research, and continued professional growth.

- ***Formulate Research Hypotheses*:** Students will be able to translate real-world problems into testable research hypotheses and select appropriate statistical tools for investigation.
- ***Data Handling and Modelling*:** Students will acquire skills to collect, manage, and analyze complex datasets using modern software and develop statistical models for forecasting and prediction.
- ***Experimental Design & Validation*:** Students will be able to design, implement, and validate experimental studies, including ANOVA techniques, to draw accurate inferences.
- ***Technical Writing and Interpretation*:** Students will be capable of producing systematic project reports, interpreting data outputs, and communicating findings, preparing them for research roles.
- ***Advanced Analysis*:** Mastery of advanced techniques such as regression and time series modelling to address industrial and scientific research questions.
- ***Adaptation to New Technologies*:** Students will develop the ability to independently learn and adapt to new programming languages, software updates, and advanced methodologies in data analysis.
- ***Higher Education Readiness*:** Students will be stimulated with self-learning skills that assist them for higher studies.
- ***Professional Certification*:** Students will be equipped to appear for professional examinations such as NET, SLET, GATE, and ISS (Indian Statistical Service).
- ***Continuous Professional Development*:** Students will develop a positive attitude toward statistics, encouraging ongoing engagement with the subject for industrial research & development and consultancy purposes.
- ***Ethical Practice*:** Students will recognize the importance of ethical decision-making and professional ethics in data analysis and research.

➢ **PSO7: Interdisciplinary Perspective** – Designed to bridge rigorous mathematical theory with practical applications across various fields, including data science, finance, economics, and biology. These outcomes ensure graduates can not only understand statistical methodologies but also apply them effectively in diverse, real-world contexts.

- *Statistical Modelling and Analysis*: Ability to apply advanced statistical models (e.g., regression analysis, time series analysis, design of experiments) to real-world data across disciplines such as business analytics, pharmaceuticals, and social sciences.

- *Data Science and Technical Proficiencies*: Proficiency in using statistical computing software and programming languages for data manipulation, visualization, and mining in high-dimensional settings.

- *Interdisciplinary Problem Solving*: Capacity to act as a statistical consultant, using quantitative data to inform decisions in fields like economics, public health, and industrial production.

- *Research and Specialized Skills*: Capability to perform research-related tasks, including literature review, data collection, and developing new statistical techniques, specifically in areas like actuarial science and survival analysis.

- *Ethical Data Practice and Communication*: Ability to interpret data within ethical, legal, and social contexts, ensuring findings are communicated effectively to stakeholders.

- *Industry Application*: Graduates are equipped to analyze data for industrial improvement, such as quality control, process monitoring, and reliability studies.

- *Data-Driven Decision Making*: A strong focus is placed on using data to make informed predictions, evaluate policies, and minimize uncertainty in economic or environmental scenarios.

- *Simulation and Data Mining*: Graduates gain skills in simulation, pattern recognition, and predictive modelling (machine learning/AI).

➢ **PSO8: Professional and Career Readiness** – Ensure graduates can translate statistical theory into actionable business insights, utilize industry-standard software, and demonstrate the technical, ethical, and soft skills required for employment.

- *Technical Proficiency and Data Handling*: Students will gain the ability to clean, process, manage, and analyze large, complex, and real-world datasets using programming languages.

- *Statistical Modelling and Analytics*: Graduates can apply advanced statistical methods – including predictive modelling, machine learning algorithms (random forests, k-nearest

neighbor), time-series analysis, and regression – to solve industry-specific problems and forecast trends.

- *Professional Software Expertise*: Competence in utilizing analytical software packages to perform high-speed, accurate, and reproducible data analysis.
- *Effective Communication and Data Storytelling*: The ability to convert complex statistical results into actionable insights, clear reports, and compelling data visualisations (charts, dashboards) for non-technical stakeholders and business executives.
- *Problem-Solving and Critical Thinking*: Graduates will be able to translate real-world business challenges into structured statistical investigations, critically evaluating model assumptions and validity.
- *Ethical Data Management*: Demonstrated understanding of data ethics, privacy laws, data integrity, and ethical decision-making in predictive modelling and business strategies.
- *Industry-Specific Domain Knowledge*: Proficiency in applying statistical tools to specific sectors, including bioinformatics (clinical trials), finance (risk modelling), marketing (consumer insights), or operations research.
- *Further Studies*: Solid theoretical foundation to pursue advanced degrees in Statistics, Biostatistics, or Data Science.
- *Independent Research*: Ability to design, implement, and document a research project from data collection to final report.
- *Entry-Level Readiness*: Students are prepared for roles such as data analyst, statistical analyst, business analysts, junior data scientist, or research officer.
- *Independent and Collaborative Work*: The skill to function efficiently both as an independent statistician and as a team member in professional, multidisciplinary settings.
- *Continuous Learning*: Development of a self-driven approach to stay updated with evolving technologies and methodologies in data science and analytics.

Course Outcomes (CO):

Semester	Course	Course Outcomes
I	STAT-H-CC1-1 STAT-H-MC1-1 STAT-MD-CC1-1 (Descriptive Statistics I and Probability I)	At the end of this course a student should be able to understand • **CO1:** Different types of data and the art of data handling. • **CO2:** Techniques of summarization and identification of the salient features of the data through graphical displays and other descriptive measures. • **CO3:** The Computation and interpretation of descriptive measures for a single variable. • **CO4:** Random experiments, sample spaces, and events, and the application of algebra of events in probabilistic reasoning. • **CO5:** The difference between the classical, statistical, and axiomatic definitions of probability, and identify situations where each is appropriate. • **CO6:** The axiomatic framework of probability and its use in deriving the fundamental probability properties. • **CO7:** The concept of conditional probability and the application of Bayes' theorem.
	STAT-H-SEC1-1 (Numerical Computations with C)	At the end of this course a student should be able to understand • **CO1:** Numerical approximation and analyze computational errors. • **CO2:** Interpolation methods to approximate functions. • **CO3:** The use of numerical differentiation and numerical integration techniques and their applications. • **CO4:** Techniques to obtain numerical solutions of equations and assessing their accuracy and convergence. • **CO5:** The components of a C program and appropriate use of data types and operators. • **CO6:** Input/output operations and implementation of decision-making and branching statements. • **CO7:** Looping constructs to solve iterative problems. • **CO8:** The application of arrays for data storage and manipulation. • **CO9:** The structure and use of user-defined functions to develop modular C programs.
	STAT-MD-SEC1-1 (An Introduction to R)	At the end of this course a student should be able to understand: • **CO1:** The **R environment** and the use of **R as a calculator** for basic numerical computations. • **CO2:** The **different types of numbers in R** and their appropriate use in computations. • **CO3:** The use of **R as a graphing calculator** for visualizing data and functions. • **CO4: Matrix operations in R** and their applications. • **CO5:** Methods for **loading data from files** into R.

		• **CO6:** The application of **numerical integration** techniques using R. • **CO7:** Techniques for obtaining **numerical solutions of equations** using R. • **CO8:** The role of **simulation in R** for statistical and numerical problem solving.
	STAT-H-IDC1-1 STAT-MD-IDC1-1 (Statistics for Practitioners)	At the end of this course a student should be able to understand: • **CO1: Understand and Apply Univariate Analysis:** Demonstrate the ability to identify types of data, select appropriate data collection methods, and describe data distribution using summary measures (mean, median, mode), measures of dispersion (range, SD, CV), and visual tools for skewness and kurtosis. • **CO2: Analyze Bivariate Relationships:** Interpret bivariate data using correlation techniques (Pearson's and Spearman's rank) and perform simple and multiple linear regression analyses to model associations between variables. • **CO3: Perform Statistical Inference (Testing of Hypothesis):** Apply concepts of sampling and test hypotheses about population means and proportions using normal and binomial distributions, including decision-making based on critical values and p-values. • **CO4: Apply Advanced Statistical Methods:** Utilize one-way/two-way ANOVA and Kruskal-Wallis tests to compare sample means and medians, and understand the application of logistic regression for binary response data. • **CO5: Perform Data Preparation and Estimation:** Execute exploratory data analysis, determine necessary sample sizes, and estimate population mean and variability for finite populations.
II	STAT-H-CC2-2 STAT-H-MC2-2 STAT-MD-CC2-2 (Descriptive Statistics II and Probability II)	At the end of this course a student should be able to understand • **CO1:** Methods to summarize and visually represent bivariate data using appropriate statistical and graphical techniques, and interpret relationships between variables. • **CO2:** The computation and interpretation of measures of correlation for bivariate data, including the strength and direction of association. • **CO3:** The application of rank correlation methods (Spearman's and Kendall's) to analyze ordinal data and data with tied ranks, and interpretation of results. • **CO4:** The analysis of categorical data using contingency tables and summary measures to study association and independence between attributes. • **CO5:** The classification of random variables (discrete and continuous) and their role in modelling random phenomena. • **CO6:** The construction and interpretation of the probability mass function (p.m.f.), probability density function (p.d.f.), and cumulative distribution function (c.d.f.) of a random variable.

		• **CO7:** The computation and interpretation of expectation and variance of discrete and continuous random variables. • **CO8:** The use of the Binomial and Poisson distributions to model count data arising from Bernoulli trials and rare events. • **CO9:** The properties of the Normal distribution, including standardization and the computation of probabilities.
	STAT-H-SEC2-2 (Artificial Intelligence for Everyone)	At the end of this course a student should be able to • **CO1:** Define and explain the fundamental concepts and subfields of AI. • **CO2:** Identify real-world applications of AI across various industries. • **CO3:** Analyze the ethical, social, and economic implications of AI. • **CO4:** Recognize the potential of AI to drive innovation and transformation in different domains.
	STAT-MD-SEC2-2 (An Introduction to R)	Same as STAT-MD-SEC1-1
	STAT-H-IDC2-2 STAT-MD-IDC2-2 (Statistics for Practitioners)	Same as STAT-H-IDC1-1 STAT-MD-IDC1-1
III	STAT-H-CC3-3 (Real Analysis I)	At the end of this course a student should be able to understand • **CO1:** The structure and properties of real numbers, including order, bounds, supremum and infimum, and their use in analytical arguments. • **CO2:** Sequences and series of real numbers, including tests for convergence and divergence. • **CO3:** The properties of real valued functions, such as limit, continuity, differentiability. • **CO4:** The concepts of differentiation to analyze real-valued functions, including monotonicity, extrema, and mean value theorems. • **CO5:** The theory of Riemann integration, including integrability, computation of integrals of real-valued functions and integral properties and theorems.
	STAT-H-CC4-3 STAT-MD-CC3-3 (Statistical Inference I)	At the end of this course a student should be able to understand • **CO1:** The basic concepts of statistical inference, including population, sample, and sampling distributions. • **CO2:** The principles and methods of point estimation and the properties of estimators. • **CO3:** The formulation and procedure of statistical hypothesis testing and the use of appropriate test statistics. • **CO4:** The construction and interpretation of confidence intervals for population parameters.

	STAT-H-SEC3-3 (Introduction to R)	At the end of this course a student should be able to understand • **CO1:** The R environment and its role in statistical computing and data analysis. • **CO2:** Data types and data structures in R and their appropriate usage. • **CO3:** Methods for importing and exporting data from various file formats. • **CO4:** Basic and advanced data manipulation techniques using R. • **CO5:** Principles of data visualization and create appropriate graphical displays in R. • **CO6:** Programming fundamentals in R, including control structures and functions. • **CO7:** The application of statistical analysis techniques using R.
	STAT-H-MC1-3 STAT-MD-MC1-3 (Descriptive Statistics I and Probability I)	Same as STAT-H-CC1-1 STAT-H-MC1-1 STAT-MD-CC1-1
	STAT-MD-SEC3-3 (An Introduction to R)	Same as STAT-MD-SEC1-1
	STAT-H-IDC3-3 STAT-MD-IDC3-3 (Statistics for Practitioners)	Same as STAT-H-IDC1-1 STAT-MD-IDC1-1
IV	STAT-H-CC5-4 (Linear Algebra)	Upon successful completion of this course, students will be able to: • **CO1: Define and analyze fundamental concepts of vector spaces** in R^n, including linear combinations, spanning sets, linear independence/dependence, basis, and dimension, and apply these to determine the structure of subspaces. • **CO2: Perform inner product operations** and apply orthogonalization techniques (Gram-Schmidt process) to construct orthonormal bases and find orthogonal projections in Euclidean spaces. • **CO3: Apply matrix algebra** to linear transformations, including computing rank, nullity, inverses, and determinants, and solve systems of linear equations using matrix methods. • **CO4: Compute and interpret eigenvalues and eigenvectors** for square matrices to analyze diagonalizability, and perform spectral decomposition for symmetric matrices. • **CO5: Identify and evaluate quadratic forms** and apply the Cayley-Hamilton theorem to compute matrix powers and inverses. • **CO6: Demonstrate understanding of subspace relationships**, such as intersections, sums, and direct sums, and apply foundational inequalities (Cauchy-Schwarz and Triangle inequality) in geometric contexts.

	STAT-H-CC6-4 (Probability III)	Upon successful completion of this course, students will be able to: • **CO1: Theoretical Understanding of Distributional Measures:** Analyze the shape and characteristics of data distributions by computing moments, quartiles, skewness, and kurtosis, and apply these measures specifically to Binomial, Poisson, and Normal distributions. • **CO2: Generating Functions and Inequalities:** Derive and apply probability generating functions (PGF) and moment generating functions (MGF) for analyzing distributions, and apply Markov and Chebyshev's inequalities to estimate probabilities, including a conceptual understanding of the characteristic function. • **CO3: Discrete Probability Distributions & Approximations:** Identify, model, and analyze real-world scenarios using discrete distributions (Uniform, Geometric, Negative Binomial, Hypergeometric), and understand the conditions under which these distributions approximate each other (limiting cases) and handle truncated data. • **CO4: Continuous Probability Distributions & Properties:** Examine and apply continuous distributions (Uniform, Exponential, Cauchy, Beta, Gamma, Lognormal, Logistic, Double Exponential, Pareto, Weibull) by deriving their moments and analyzing their properties, including truncated continuous distributions. • **CO5: Bivariate and Multivariate Random Variables:** Analyze two-dimensional random variables by deriving joint, marginal, and conditional distributions, compute correlation and regression, understand independence, and apply the theorems on the sum and product of expectations. • **CO6: Specialized Distributions (Trinomial & Bivariate Normal):** Understand the properties, derivation, and applications of the Trinomial distribution and Bivariate Normal Distribution in multi-dimensional data analysis.
	STAT-H-CC7-4 (Sampling Distributions and Statistical Inference II)	At the end of this course a student should be able to • **CO1: Understand the Foundation of Sampling Distributions:** Explain the concepts of random sampling, sampling distributions, standard error, and reproductive properties of sampling distributions. • **CO2: Analyze Order Statistics and Joint Distributions:** Derive and apply the distribution of the r-th order statistic, sample median, and joint distribution of order statistics for continuous random variables. • **CO3: Derive Standard Sampling Distributions:** Derive sampling distributions of χ^2, t and F statistics from univariate normal populations, and analyze sample means, variances, and regression coefficients from bivariate normal populations.

		• **CO4: Apply Point Estimation Techniques:** Apply the factorization theorem to identify sufficient statistics, calculate Fisher's Information, and construct Minimum Variance Bound (MVB) estimators using the Cramer-Rao inequality and Rao-Blackwell theorem. • **CO5: Evaluate Hypothesis Testing Procedures:** Construct Most Powerful (MP) and Uniformly Most Powerful (UMP) tests using the Neyman-Pearson Lemma, and analyze the properties of Likelihood Ratio Tests. • **CO6: Formulate Interval Estimation:** Develop confidence intervals for population parameters and understand the concept of Uniformly Most Accurate (UMA) confidence sets.
	STAT-H-CC8-4 STAT-MD-CC4-4 (Design of Experiments I and Sample Survey I)	After successful completion of this course, students will be able to: • **CO1:** Understand the fundamental principles of design of experiments (randomization, replication, and local control) and distinguish between fixed, random, and mixed-effects models. • **CO2:** Apply Analysis of Variance (ANOVA) techniques to analyze one-way and two-way classified data with equal observations (fixed-effects models). • **CO3:** Layout and analyze experimental data using Completely Randomized Design (CRD), Randomized Block Design (RBD), and Latin Square Design (LSD), including handling missing data in RBD and LSD. • **CO4:** Evaluate and compare the relative efficiency of different experimental designs through uniformity trials. • **CO5:** Differentiate between complete enumeration and sampling, and analyze sampling/non-sampling errors in survey research. • **CO6:** Apply simple random sampling (with and without replacement) to estimate population parameters (mean, total, proportion) and calculate standard errors. • **CO7:** Implement stratified random sampling techniques, including proportional and optimum allocation, and compare their efficiency with Simple Random Sampling (SRS). • **CO8:** Design and analyze two-stage sampling techniques (with equal-sized primary units) and understand non-probability sampling methods like snowball and purposive sampling.
	STAT-H-MC2-4 STAT-MD-MC2-4 (Descriptive Statistics II and Probability II)	Same as STAT-H-CC2-2 STAT-H-MC2-2 STAT-MD-CC2-2

	STAT-MD-CC5-4 (Descriptive Statistics III and Probability III)	Upon successful completion of this course, the student will be able to: • **CO1: Analyze Multivariate Relationships.** Apply multiple regression and compute multiple and partial correlation coefficients to quantify and model the dependencies between three or more variables. • **CO2: Evaluate Distributional Characteristics.** Utilize moments, quantiles, skewness, and kurtosis to describe the shape and spread of distributions, specifically for Binomial, Poisson, and Normal models. • **CO3: Implement Generating Functions.** Derive and apply Probability Generating Functions (PGF) and Moment Generating Functions (MGF) to find moments and identify the distribution of transformed random variables. • **CO4: Model Discrete and Continuous Scenarios.** Categorize and apply standard distributions (Geometric, Hypergeometric, Gamma, Beta, etc.) to real-world data and understand their limiting/ approximation behaviours. • **CO5: Solve Bivariate Probability Problems.** Compute joint, marginal, and conditional probability functions and apply theorems of expectation and independence to two-dimensional random variables. • **CO6: Analyze Specialized Bivariate Distributions.** Demonstrate proficiency in the properties and applications of the Trinomial and Bivariate Normal distributions in statistical modelling.
V	STAT-H-CC9-5 (Multivariate Analysis I)	At the end of this course a student should be able to • **CO1:** Understand and apply notation for multivariate data structures, including mean vectors, dispersion matrices, and correlation matrices. • **CO2:** Analyze the properties of random vectors, including joint, marginal, and conditional distributions, and the concept of independence. • **CO3:** Compute and interpret multiple, partial, and canonical correlations to evaluate relationships between sets of variables. • **CO4:** Formulate mathematical models for Multivariate Normal, Multinomial, and Dirichlet distributions and derive their fundamental properties. • **CO5:** Perform Maximum Likelihood Estimation (MLE) for multivariate parameters and assess the distribution of sample mean vectors and covariance matrices. • **CO6:** Execute statistical hypothesis tests for regression coefficients and correlation measures in complex multivariate datasets.

	STAT-H-CC10-5 (Statistical Inference III)	At the end of this course a student should be able to • **CO1:** Analyze sequences of random variables using various modes of convergence and apply the Central Limit Theorem to approximate distributions. • **CO2:** Employ the Delta Method to derive asymptotic standard errors and handle distributions of sample moments and quantiles. • **CO3:** Apply variance stabilizing transformations (such as Sin^{-1} and z-transformation) to prepare non-normal data for large-sample testing. • **CO4:** Critique estimators using asymptotic properties, including consistency, efficiency, and the theory of CAN and BAN estimators. • **CO5:** Construct and interpret Large Sample tests using Likelihood Ratio, Wald's, and Rao's Score criteria under the normality assumption. • **CO6:** Implement and justify the use of nonparametric tests (e.g., Wilcoxon, Kruskal-Wallis, K-S) when distributional assumptions are violated. • **CO7:** Develop Sequential Probability Ratio Tests (SPRT) and evaluate their performance using OC and ASN functions.
	STAT-H-CC11-5 (Linear Models)	At the end of this course a student should be able to • **CO1:** Formulate General Linear Models and apply the Method of Least Squares to estimate parameters in a Gauss-Markov setup. • **CO2:** Evaluate the estimability of linear parametric functions and apply the Gauss-Markov Theorem to identify Best Linear Unbiased Estimators (BLUE). • **CO3:** Perform orthogonal splitting of total variation (ANOVA) to partition sums of squares and select appropriate error terms for hypothesis testing. • **CO4:** Analyze Fixed, Random, and Mixed effect models and perform statistical inference on variance components in balanced designs. • **CO5:** Build and interpret Simple and Multiple Linear Regression models, including parameter estimation and diagnostic hypothesis testing. • **CO6:** Execute Analysis of Covariance (ANCOVA) using concomitant variables to ad- just treatment effects and reduce experimental error.

	STAT-H-CC12-5 (Demography, Index Numbers and Psychometry)	At the end of this course a student should be able to • **CO1:** Compute and interpret vital rates (Mortality and Fertility) and reproduction rates (GRR, NRR) to evaluate population health and growth. • **CO2:** Construct and utilize Life Tables to derive actuarial functions (l_x, d_x, q_x, e_x) for demographic and insurance-based projections. • **CO3:** Model population trends using arithmetic, geometric, and exponential growth models and differentiate between stable and stationary populations. • **CO4:** Construct simple and weighted Index Numbers and validate their consistency using Time and Factor Reversal tests. • **CO5:** Perform advanced economic adjustments including base shifting, splicing, and deflating of price/value index series. • **CO6:** Design standardized psychological measurements using specialized scales (Z-scale, T-scale) and appropriate levels of measurement. • **CO7:** Appraise the quality of educational and psychological tests by calculating reliability (Cronbach's α) and verifying multi-faceted validity.
	STAT-H-MC3-5 STAT-MD-MC3-5 (Statistical Inference I)	Same as STAT-H-CC4-3 STAT-MD-CC3-3
	STAT-H-MC4-5 (Design of Experiments I and Sample Survey I)	Same as STAT-H-CC8-4 STAT-MD-CC4-4
	STAT-MD-CC6-5 STAT-MD-MC4-5 (Project)	Upon successful completion of this project course, students will be able to: • **CO1: Identify and Formulate** relevant problems based on real-world scenarios that can be solved using statistical methods. • **CO2: Design and Implement** appropriate data collection strategies, including surveys, experiments, or secondary data extraction, while ensuring data quality and ethical standards. • **CO3: Clean, Organize, and Manipulate** raw data using data processing techniques, handling missing values, outliers, and data formatting issues. • **CO4: Perform Comprehensive Data Analysis** by applying appropriate descriptive statistics, inferential tests (e.g., ANOVA, regression), and statistical modelling. • **CO5: Interpret and Visualize** findings from analysis to extract meaningful insights, using graphical representations like graphs, charts, and plots. • **CO6: Develop and Document** a structured, technical report and present findings to a non-technical audience, drawing sound conclusions and limitations.

	STAT-MD-CC7-5 (Applications of Statistics I)	Upon completion of this course, students will be able to: • **CO1 (Demography):** Identify and evaluate various sources of population data (Census, Registration) and apply techniques to measure mortality (SDR, IMR), fertility (TFR, GFR), and population growth (reproduction rates). • **CO2 (Demography):** Construct and interpret complete life tables and differentiate between stable and stationary population models for demographic analysis. • **CO3 (Index Numbers):** Construct and compare price, quantity, and value index numbers using various formulas (e.g., Laspeyres, Paasche, Fisher) and analyze their limitations. • **CO4 (Index Numbers):** Apply key economic indices like the Consumer Price Index (CPI) and Index of Industrial Production (IIP) for economic monitoring and policy formulation. • **CO5 (Psychometry):** Apply scaling procedures – including Z, T, and percentile scaling – to standardize psychological and educational data. • **CO6 (Psychometry):** Develop and implement attitude scaling techniques (Likert and Thurstone scaling) and analyze the concept of intelligence tests and IQ.
VI	STAT-H-CC13-6 (Applied Multivariate Analysis)	At the end of this course a student should be able to • **CO1:** Distinguish between dependence and interdependence methods to select appropriate multivariate strategies for complex datasets. • **CO2:** Execute Principal Component Analysis (PCA) for dimensionality reduction and interpret Scree plots to determine the proportion of variance explained. • **CO3:** Model latent structures using Factor Analysis (FA) and apply Varimax rotation to enhance the interpretability of factor loadings. • **CO4:** Evaluate similarity and dissimilarity using Euclidean and Mahalanobis distances to facilitate robust cluster formation. • **CO5:** Perform Hierarchical and Non-hierarchical (K-means) clustering to identify natural groupings and interpret results via Dendrograms. • **CO6:** Construct Fisher's Linear Discriminant Function for classification tasks and assess the efficiency of classification rules across groups.

	STAT-H-CC14-6 (Regression Analysis)	At the end of this course a student should be able to • **CO1:** Apply the Gauss-Markov model in correlated setups and perform least squares estimation under linear restrictions on parameters. • **CO2:** Construct flexible regression models using Box-Cox transformations, dummy variables, and spline-based smoothing techniques. • **CO3:** Identify and assess outliers and influential observations using residuals, leverages, DFBETA, DFFIT, and Cook's Distance. • **CO4:** Evaluate and select optimal regression models using R^2, Adjusted R^2, Mallows' C_p, and Akaike Information Criterion (AIC). • **CO5:** Detect multicollinearity using Variance Inflation Factors (VIF) and Condition Numbers, and implement remedial measures. • **CO6:** Formulate and interpret Logistic Regression models for binary response data in both grouped and ungrouped contexts.
	STAT-H-CC15-6 (Time Series Analysis)	At the end of this course a student should be able to • **CO1:** Define the objectives of time series and distinguish between stationary and evolutionary (non-stationary) processes. • **CO2:** Decompose time-indexed data into trend, seasonal, and cyclical variations using additive and multiplicative models. • **CO3:** Estimate deterministic trends and perform deseasonalization using moving averages and the least squares method. • **CO4:** Apply seasonal adjustment techniques, including the Ratio to Moving Average and Link Relative methods, to normalize periodic fluctuations. • **CO5:** Analyze Autocorrelation (ACF) and Partial Autocorrelation (PACF) functions to detect stationarity and identify model orders. • **CO6:** Formulate and fit AR, MA, ARMA, and ARIMA models for forecasting and evaluate model adequacy using the Ljung-Box test and AIC.
	STAT-H-MC3-6 (Statistical Inference I)	Same as STAT-H-CC4-3 STAT-MD-CC3-3
	STAT-H-MC4-6 STAT-MD-MC5-6 (Design of Experiments I and Sample Survey I)	Same as STAT-H-CC8-4 STAT-MD-CC4-4
	STAT-MD-CC7-6 (Applications of Statistics I)	Same as STAT-MD-CC7-5

	STAT-MD-CC8-6 (Applications of Statistics II)	At the end of this course a student should be able to • **CO1:** Identify and decompose time series components using additive and multiplicative models to analyze historical data trends. • **CO2:** Apply moving averages, curve fitting, and deseasonalization techniques to forecast and interpret stationary and non-stationary data. • **CO3:** Construct and interpret Process Control Charts $(p, c, np, \overline{X}, R)$ to distinguish between common and assignable causes of variation. • **CO4:** Evaluate Product Control risks (Producer's vs. Consumer's) and design Single/Double Acceptance Sampling plans using OC, ASN, and ATI curves. • **CO5:** Distinguish between Supervised and Unsupervised learning frameworks and implement Clustering algorithms like K-means. • **CO6:** Execute classification and regression algorithms (Decision Trees, Random Forests, etc.) and apply Principal Component Analysis (PCA) for dimension reduction.
	STAT-MD-MC6-6 (Descriptive Statistics III and Probability III)	Same as STAT-MD-CC5-4
VII	STAT-H-CC16-7 (Real Analysis II and Multivariate Analysis II)	At the end of this course a student should be able to understand • **CO1:** Topology and sequences in R^n. • **CO2:** Limits, continuity, and uniform continuity of functions of several variables. • **CO3:** The concept of differentiation of multivariable functions. • **CO4:** The concept of multiple integration techniques. • **CO5:** Sampling distributions of quadratic forms. • **CO6:** Properties of the Wishart Distribution and use of Hotelling's T^2 for hypothesis testing and confidence region construction in single and two-sample multivariate cases. • **CO7:** Extension of ANOVA concepts to multivariate settings using MANOVA, apply hypothesis testing, and interpret results.
	STAT-H-CC17-7 (Probability IV)	At the end of this course a student should be able to understand • **CO1:** The Foundations of measure-theoretic probability • **CO2:** Probability measures and their properties • **CO3:** The concept of independence of events and limiting behaviour of sequences of events. • **CO4:** Random variables and different types of distribution functions. • **CO5:** Joint distributions and expectation theory. • **CO6:** Modes of convergence and laws of large numbers. • **CO7:** The concept and classification of stochastic processes, including counting processes.

	STAT-H-CC18-7 (Statistical Inference IV)	At the end of this course a student should be able to understand / achieve • **CO1: Advanced Point Estimation Theory** - **Mastery of Sufficiency and Completeness:** Gain the ability to identify minimal sufficient statistics and verify the concept of complete families of distributions. - **Exponential Families:** Understand the properties and applications of both one-parameter and multi-parameter exponential families. - **Optimal Estimation:** Develop skills to derive Minimum Variance Unbiased Estimators (MVUE) using the Lehmann-Scheff´e Theorem. - **Distributional Independence:** Apply Basu's Theorem to understand the independence of ancillary and sufficient statistics. - **Variance Bounds:** Utilize the Bhattacharya system of lower bounds and the Method of Covariances for estimation. • **CO2: Rigorous Hypothesis Testing** - **Neyman-Pearson Framework:** Master the fundamental Neyman-Pearson lemma, including its existence, sufficiency, and necessity parts. - **UMP Test Construction:** Construct Uniformly Most Powerful (UMP) tests, specifically for distributions with varying supports and those with the Monotone Likelihood Ratio (MLR) property. - **Advanced Testing Methods:** Construct Uniformly Most Powerful Unbiased (UMPU) tests for simple and composite hypotheses, as well as Locally Most Powerful (LMP) tests. • **CO3: Theory of Confidence Sets** - **Interval Estimation:** Understand the theory and construction of Uniformly Most Accurate Unbiased (UMAU) confidence intervals. - **Simultaneous Inference:** Grasp the concept and construction of simultaneous confidence intervals.

	STAT-H-CC19-7 (Linear Models and Regression Analysis II)	Upon successful completion of this course, students will be able to: • **CO1: Analyze Gauss-Markov Models:** Understand and apply the Gauss-Markov model, including alternative approaches to solving normal equations and the theoretical foundations of BLUEs (Best Linear Unbiased Estimators). • **CO2: Perform Theoretical Linear Modelling:** Implement linear sets, orthogonal splitting, and evaluate general linear hypotheses to derive related sampling distributions. • **CO3: Apply Projection Matrix Theory:** Utilize projection matrices to determine estimability, derive least squares estimators, and apply these concepts to ANOVA models. • **CO4: Conduct Statistical Inference & Comparison:** Derive F-tests, compare them with likelihood ratio tests, construct confidence intervals/regions, and apply multiple comparison techniques (Scheffe and Tukey). • **CO5: Evaluate Regression Models & Variables:** Perform variable selection, examine properties of OLS estimators for model subsets, and apply model selection criteria (e.g., AIC, BIC). • **CO6: Detect and Remediate Multicollinearity:** Detect multicollinearity using diagnostics (e.g., VIF) and implement remedial measures including Ridge Regression and Principal Components. • **CO7: Perform Diagnostic Analytics:** Identify outliers and influential observations using residuals, leverage, DFBETA, DFFIT, and Cook's Distance. • **CO8: Analyze Error Structure & Assumptions:** Identify and address consequences of heteroscedasticity and autocorrelation, and validate model assumptions using Q-Q plots, normal probability plots, and the Shapiro-Wilk test.
	STAT-H-CC20-7 (Design of Experiments II and Sample Survey II)	Upon successful completion of this course, students will be able to: • **CO1: Design and Analyze Complex Experimental Layouts:** Construct and analyze split-plot designs in RBD and strip-plot arrangements, along with groups of experiments involving RBD and LSD, to handle complex research scenarios. • **CO2: Implement Factorial Experiments ($2^n, 3^n$):** Develop, analyze, and interpret 2^n and 3^n full factorial designs, assessing main effects and interaction effects, and apply techniques of partial/complete confounding and balancing. • **CO3: Evaluate and Construct Incomplete Block Designs (BIBD):** Define, analyze, and construct Balanced Incomplete Block Designs (BIBD) using finite projective and Euclidean geometries, and verify resolvability using Bose's inequality. • **CO4: Apply Advanced Sampling Techniques:** Implement Linear and Circular Systematic sampling, and apply ratio and

		regression methods of estimation, including double sampling for improved precision. • **CO5: Analyze Complex Surveys and Unequal Probability Sampling:** Estimate population parameters using Cluster sampling and PPS sampling (WR & WOR) utilizing intraclass correlation coefficients, and evaluate estimators using the Horvitz-Thompson technique. • **CO6: Perform Error Analysis and Estimation:** Estimate variances through interpenetrating subsampling techniques and construct Hansen-Hurwitz and Desraj estimators for PPS sampling.
VIII	STAT-H-CC21-8 (Research Methodology I: Research Methodology with Survey Methods)	At the end of this course a student should be able to understand / achieve • **CO1: Mastery of Research Fundamentals** - **Problem Formulation**: Ability to define research problems, conduct systematic literature reviews, and formulate testable hypotheses. - **Design Selection**: Proficiency in choosing between exploratory, descriptive, and causal research designs based on the study's conceptual blueprint. - **Logical Reasoning**: Understanding the application of deductive versus inductive reasoning in both social and scientific research contexts. • **CO2: Ethical and Professional Integrity** - **Ethical Frameworks**: Comprehension of the branches of ethics (Meta-ethics, Normative, and Applied) and their application in bioethics and business. - **Academic Honesty**: Recognition and avoidance of research misconduct, specifically Falsification, Fabrication, and Plagiarism (FFP). - **Professional Standards**: Understanding the ethics of authorship, peer review, and the implications of redundant publication. • **CO3: Survey Methodology and Quality Control** - **Total Survey Error (TSE)**: Ability to use the TSE framework as a tool for survey quality, identifying coverage, non-response, and measurement errors. - **Data Collection Techniques**: Competency in various collection modes such as Face-to-face, Telephone, and Self-administered surveys. - **Questionnaire Design**: Skill in drafting, sequencing, and reliability testing for designing effective questionnaires, including the use of CAPI over PAPI. • **CO4: Advanced Sampling and Estimation** - **Sampling Techniques**: Expertise in applying probability sampling (SRS, Systematic, Stratified, Multi-stage Cluster, and PPS) and non-probability techniques like Convenience or Quota sampling. - **Statistical Application**: Capacity to estimate sample sizes and analyze the trade-off between bias and variance using the Design Effect (*Deff*).

		- **Population Estimates**: Ability to construct and interpret Survey Weights to produce accurate population-level estimates.
	STAT-H-CC22-8 (Research Methodology II: Research Methodology with Statistical Learning)	At the end of this course a student should be able to understand • **CO1:** The nature and role of literature review in research, identify research gaps, and formulate relevant research questions using academic sources. • **CO2:** The use of major academic databases to conduct literature searches, and evaluate research impact using bibliometric indicators. • **CO3:** The role of data science and statistical learning in research. • **CO4:** Formulate and solve classification problems. • **CO5:** Apply and interpret advanced supervised learning models. • **CO6:** Unsupervised learning and clustering. • **CO7:** Academic integrity, apply proper citation practices, and avoid plagiarism in research writing. • **CO8:** Standard citation styles and distinguish between reference lists and bibliographies for proper academic documentation.
	STAT-H-CC23-8 (Statistical Inference V)	At the end of this course a student should be able to understand / achieve • **CO1: Mastery of Statistical Decision Theory** Students will be able to frame statistical problems as a functional game between a researcher and nature. Key competencies include: - Defining and calculating **Loss and Risk functions** for various decision rules. - Evaluating the **Admissibility** of rules and understanding the hierarchy of complete and essentially complete classes. - Identifying optimal strategies using **Minimax rules** and **Bayes rules** to minimize risk. • **CO2: Proficiency in Bayesian Analysis** Students will transition from classical methods to a Bayesian framework, learning to incorporate prior information into statistical models. Outcomes include: - Differentiating between **Classical and Bayesian** approaches and justifying the choice of different priors. - Deriving **Posterior distributions** for standard models, such as the Binomial-Beta and Normal- Normal conjugate systems. - Performing Bayesian point estimation (Posterior Mean), hypothesis testing, and constructing **Credible Sets**. • **CO3: Understanding Asymptotic Inference** Students will gain the theoretical tools to analyze the behavior of estimators and tests as sample sizes increase. Outcomes include:

		- Applying the concepts of **Consistency and Asymptotic Efficiency** to evaluate estimator performance. - Utilizing **Maximum Likelihood Estimators (MLE)** and calculating observed **Fisher Information** to estimate asymptotic variance. - Implementing and comparing large-sample tests, specifically the **Likelihood Ratio Test (LRT)**, **Rao's Score Test**, and **Wald's Test**.
	STAT-H-WR-CC23-8 (Research Internship)	On successful completion of the research internship, students will be able to: • **CO1: Formulate and Define Research Problems:** Identify real-life or academic research problems and formulate statistical hypotheses, objectives, and appropriate study designs. • **CO2: Perform Data Collection and Management:** Apply appropriate sampling techniques to collect data, or efficiently source data from existing databases, ensuring data cleaning, structuring, and validation. • **CO3: Apply Advanced Statistical Computing:** Utilize statistical software and programming languages (e.g., R, Python) for data manipulation, analysis, and simulation. • **CO4: Analyze and Interpret Results:** Apply descriptive and inferential statistical methods – including regression, ANOVA, and multivariate techniques – to interpret data, draw valid conclusions, and report findings. • **CO5: Communicate Findings Professionally:** Present statistical analysis results effectively through professional reports, research papers, data visualization, and presentations. • **CO6: Practice Research Ethics:** Demonstrate understanding and application of research ethics, data integrity, and professional accountability in research. • **CO7: Develop Collaboration and Teamwork Skills:** Demonstrate the ability to work in teams, adapt to professional research environments, and collaborate with mentors or industry professionals.
	STAT-H-CC24-8 (Discrete Data Analysis)	Upon completion of this course, students will be able to: • **CO1:** Understand the principles of alternative study designs and analyze contingency tables using measures of association, conditional/marginal independence, and interpret Simpson's paradox. • **CO2:** Comprehend the components of Generalized Linear Models (GLM), including the link function, maximum likelihood estimation (MLE), the triad of tests (Wald, Score, Likelihood Ratio), and the analysis of deviance and residuals. • **CO3:** Apply GLMs to model binary (grouped/ungrouped) and count data, including negative binomial regression for overdispersed data, zero-inflated models, and joint modelling of mean and variance.

		• **CO4:** Analyze nominal and ordinal polytomous categorical data using appropriate regression models and association measures. • **CO5:** Understand the concept of Generalized Linear Mixed Models (GLMM) for correlated data, including estimation procedures (BLUP) for random effects. • **CO6:** Implement and interpret GLM and categorical data models using statistical software, checking for model goodness-of-fit and assessing overdispersion.
	STAT-H-WR-CC24-8 (Dissertation)	Dissertation focuses on applying theoretical knowledge to real-world data, independent research capability, and technical proficiency. Key course outcomes include: • **CO1: Research Design and Data Collection:** Ability to understand basic concepts of research, formulate research questions, and design data collection methods, including sampling techniques. • **CO2: Application of Statistical Techniques:** Practical application of statistical methodologies, such as regression analysis, time series analysis, or multivariate analysis, to real-life situations. • **CO3: Data Analysis Proficiency:** Competence in using statistical software tools (such as R or Python) for data cleaning, visualization, and analysis. • **CO4: Report Generation and Interpretation:** Ability to analyze results, interpret findings, and document the research in a structured report. • **CO5: Independent Research Skills:** Development of independent thinking, critical analysis of data, and the ability to draw meaningful conclusions. • **CO6:** Prepare students for professional data analytics roles or research. • **CO7:** Aims for potential publication of findings in peer-reviewed journals or presentation at seminars.
	STAT-H-CC25-8 (Project)	Project Work in Statistics at the Undergraduate (UG) level are designed to bridge theoretical knowledge with practical application. After completing the project, students should be able to: • **CO1: Formulate Research Questions:** Identify real-world problems and translate them into actionable statistical questions, defining appropriate populations and sample frameworks. • **CO2: Design Data Collection Mechanisms:** Develop valid data collection instruments (e.g., questionnaires, surveys, experimental layouts) to gather structured data while addressing sampling and non-sampling errors. • **CO3: Perform Exploratory Data Analysis (EDA):** Clean, organize, summarize, and visualize raw data using

		appropriate statistical graphs (histograms, box plots, scatter plots) and descriptive measures to identify patterns. • **CO4: Apply Statistical Modelling Techniques:** Select and apply suitable statistical inferential methods (e.g., regression analysis, ANOVA, hypothesis testing, Time Series) to analyze data and address the research objectives. • **CO5: Utilize Statistical Software Tools:** Employ computational tools such as R, Python or MS Excel to perform data management and advanced statistical analyses efficiently. • **CO6: Interpret and Communicate Findings:** Interpret statistical results in the context of the real-world problem and present findings through a systematic, written project report and oral presentation. • **CO7: Evaluate Ethical and Methodological Limitations:** Critically evaluate the limitations of the chosen methodology, ethical considerations in data handling, and the robustness of the conclusions.

UNIVERSITY OF CALCUTTA

Syllabus

for

Four-year B.Sc. (Honours & Honours with Research) Courses of Studies

in

STATISTICS

(Under Curriculum & Credit Framework, 2022)

With effect from the Academic Session

2023-24

Structure of Statistics Major Courses

Semester	Course / Paper Code	Course Name
I	STAT-H-CC1-1-Th / P	Descriptive Statistics I & Probability I
	STAT-H-SEC1-1-Th / P	Numerical Computations with C
II	STAT-H-CC2-2-Th / P	Descriptive Statistics II & Probability II
	STAT-H-SEC2-2-Th / TU	Artificial Intelligence for Everyone
III	STAT-H-CC3-3-Th / TU	Real Analysis I
	STAT-H-CC4-3-Th / P	Statistical Inference I
	STAT-H-SEC3-3-P	Introduction to R
IV	STAT-H-CC5-4-Th / P	Linear Algebra
	STAT-H-CC6-4-Th / P	Probability III
	STAT-H-CC7-4-Th / P	Sampling Distributions and Statistical Inference II
	STAT-H-CC8-4-Th / P	Design of Experiments I and Sample Survey I
V	STAT-H-CC9-5-Th / P	Multivariate Analysis I
	STAT-H-CC10-5-Th / P	Statistical Inference III
	STAT-H-CC11-5-Th / P	Linear Models
	STAT-H-CC12-5-Th / P	Demography, Index Numbers and Psychometry
VI	STAT-H-CC13-6-Th / P	Applied Multivariate Analysis
	STAT-H-CC14-6-Th / P	Regression Analysis
	STAT-H-CC15-6-Th / P	Time Series Analysis
VII	STAT-H-CC16-7-Th / P	Real Analysis II and Multivariate Analysis II
	STAT-H-CC17-7-Th / TU	Probability IV
	STAT-H-CC18-7-Th / P	Statistical Inference IV
	STAT-H-CC19-7-Th / P	Linear Models and Regression Analysis II
	STAT-H-CC20-7-Th / P	Design of Experiments II and Sample Survey II
VIII	*Honours and Honours With Research*	
	STAT-H-CC21-8-Th / Viva	Research Methodology I (Research Methodology with Survey Methods)
	STAT-H-CC22-8-Th / Viva	Research Methodology II (Research Methodology with Statistical Learning)
	Honours	*Honours With Research*
	STAT-H-CC23-8-Th / P Statistical Inference V	STAT-H-WR-CC23-8-Th / Viva Research Internship
	STAT-H-CC24-8-Th / P Discrete Data Analysis	STAT-H-WR-CC24-8-Th / Viva Dissertation *
	STAT-H-CC25-8-Th / Viva Project	

* Carrying 8 Credits (6 Th + 2 Viva)

Structure of Statistics Minor Courses

Semester	Course / Paper Code	Course Name
I	STAT-H-MC1-1-Th / P	Descriptive Statistics I & Probability I
II	STAT-H-MC2-2-Th / P	Descriptive Statistics II & Probability II
III	STAT-H-MC1-3-Th / P	Descriptive Statistics I & Probability I
IV	STAT-H-MC2-4-Th / P	Descriptive Statistics II & Probability II
V	STAT-H-MC3-5-Th / P	Statistical Inference I
	STAT-H-MC4-5-Th / P	Design of Experiments I and Sample Survey I
VI	STAT-H-MC3-6-Th / P	Statistical Inference I
	STAT-H-MC4-6-Th / P	Design of Experiments I and Sample Survey I

Structure of Statistics Interdisciplinary Course

Semester	Course / Paper Code	Course Name
I	STAT-H-IDC1-1-Th / P	Statistics for Practitioners
II	STAT-H-IDC2-2-Th / P	
III	STAT-H-IDC3-3-Th / P	

Question Patterns

SEMESTER - I

STATISTICS MAJOR

STAT-H-CC1-1-Th **3 Credits**

(Descriptive Statistics I & Probability I) **THEORY**

Statistics: Definition and scope. Concepts of statistical population and sample.

Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous.

Scales of measurement: nominal, ordinal, interval and ratio.

Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays. (10)

Measures of Central Tendency: Mean, Median, Mode.

Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Ginis Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers. (15)

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications. (20)

STAT-H-CC1-1-P **1 Credit**

(Descriptive Statistics I & Probability I) **PRACTICAL**

List of Suggested Practical

- Diagrammatic representation of data.
- Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations, stem and leaf plot.
- Problems based on measures of central tendency.
- Problems based on measures of dispersion.
- Problems based on combined mean and variance and coefficient of variation.
- Problems based on moments, skewness and kurtosis.
- Problems related to quantiles and measures based on them, construction of box plot.
- Application problems based on Classical Definition of Probability.
- Application problems based on Bayes' Theorem.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STAT-H-SEC1-1-Th **2 Credits**

(Numerical Computations with C) **THEORY**

Numerical Analysis:

Approximation of numbers and functions. Absolute and Relative errors.

Interpolation: Polynomial approximation, Weierstrass Theorem (Statement). 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 and Simpson's 1/3rd rules.

Numerical solution of equations: method of fixed-point iteration and Newton-Raphson method in one unknown, conditions of convergence (statement only). (15)

C-programming:

Components, basic structure programming, character set, C/C++ tokens, Keywords and Identifiers and execution of a C/C++ program.

Data types: Basic data types, enumerated data types, derived data types. Constants and variables: declaration and assignment of variables, symbolic constants, overflow and underflow of data.

Operators and expressions: library functions.

Managing input and output operations: reading and printing formatted and unformatted data.

Decision making and branching: if...else, nesting of if...else, else if ladder, switch.

Looping in C/C++: for, nested for, while, do...while, and jumps in and out of loops.

Arrays: Declaration and initialization of one-dim and two-dim arrays. User-defined functions.

(15)

STAT-H-SEC1-1-P **2 Credits**

(Numerical Computations with C) **PRACTICAL**

List of Suggested Practical

- Finding values of a function $y = f(x)$ for given values of x.
- Roots of a quadratic equation (with imaginary roots also).
- Sorting of an array and hence finding median.
- Mean, median and mode of a grouped frequency Data.
- Variance and coefficient of variation of a grouped frequency data.
- Preparing a frequency table.
- Numerical methods: Interpolation by Lagrange's formula, Solving one-variable equations using Newton-Raphson and iteration methods.
- Trapezoidal and Simpson's 1/3rd rule for numerical integration with convergence.
- Storing the C output in a file.

Reference Books:

- Kernighan, B.W. and Ritchie, D.: C Programming Language, 3rd Edition, Prentice Hall.
- Balagurusamy, E.: Programming in ANSI C, 6th Edition, Tata McGraw Hill.
- Gottfried, B.S.: Schaums Outlines: Programming with C, 2nd Edition, Tata McGraw Hill.
- Jain, M. K., Iyengar, S. R. K. and Jain, R. K.: Numerical methods for scientific and engineering computation, New age International Publisher, India.
- Mukherjee, Kr. Kalyan: Numerical Analysis, New Central Book Agency.
- Sastry, S.S.: Introductory Methods of Numerical Analysis, 3rd edition, Prentice Hall of India Pvt. Ltd., New Delhi.
- Scarborough, J.B.: Numerical Mathematical Analysis, Oxford and IBH Publishing.

STATISTICS MINOR

STAT-H-MC1-1-Th **3 Credits**

(Descriptive Statistics I & Probability I) **THEORY**

Statistics: Definition and scope. Concepts of statistical population and sample.
Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous.
Scales of measurement: nominal, ordinal, interval and ratio.
Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays. (10)

Measures of Central Tendency: Mean, Median, Mode.
Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Ginis Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers. (15)

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications. (20)

STAT-H-MC1-1-P **1 Credit**

(Descriptive Statistics I & Probability I) **PRACTICAL**

List of Suggested Practical

- Diagrammatic representation of data.
- Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations, stem and leaf plot.
- Problems based on measures of central tendency.
- Problems based on measures of dispersion.
- Problems based on combined mean and variance and coefficient of variation.
- Problems based on moments, skewness and kurtosis.
- Problems related to quantiles and measures based on them, construction of box plot.
- Application problems based on Classical Definition of Probability.
- Application problems based on Bayes' Theorem.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STATISTICS INTERDISCIPLINARY

STAT-H-IDC1-1-Th **2 Credits**

(Statistics for Practitioners) **THEORY**

Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (arithmetic mean, median, mode), dispersion (range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (only through diagrams), Exploratory Data Analysis. (8)

Understanding bivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation (no tie), measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). (8)

Statistical Inference (*testing of hypothesis*): Basic idea of binomial and normal populations (graphical idea only, derivation of the properties excluded). Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and p-value for some standard testing problems like test for proportion/proportions, mean based on single (normal) sample, test on comparing means based on two-sample and paired sample data. (7)

Miscellaneous discussion: Applications of one-way and two-way ANOVA with one observation per cell (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimation of population mean and variability for finite population, idea and application of logistic regression for binary response data. (7)

STAT-H-IDC1-1-P **1 Credit**

(Statistics for Practitioners) **PRACTICAL**

List of Suggested Practical

- Measures of mean, median, mode, range, QD, SD, CV for univariate data case.
- Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearsons correlation coefficients, Spearman's Rank correlation, measures of association of attributes through contingency table.
- Tests for proportion/proportions, tests of means for single sample, two-sample, and paired sample data on normal response using p-value approach.
- Applications of ANOVA and Kruskal-Wallis test.
- Sample size determination, estimation of population mean and variability for finite population.
- Fitting of logistic regression for binary response data.

Reference Books:

- Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
- Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis, PHI.
- Hardle W. and Simar, L.: Applied Multivariate Statistical Analysis.
- Kutner, M.H. et.al.: Applied Linear Statistical Models.
- Belsley D.A. et.al.: Regression Diagnostics.
- Draper N.R. and Smith, H.: Applied Regression Analysis.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - II

STATISTICS MAJOR

STAT-H-CC2-2-Th **3 Credits**

(Descriptive Statistics II & Probability II) **THEORY**

Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, fitting of polynomial and exponential curves, correlation ratio, correlation index, intraclass correlation.

Rank correlation: Spearman's and Kendall's measures. (15)

Analysis of Categorical Data: Contingency table, independence & association of attributes. (5)

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.) and probability density function (p.d.f.). Expectation and Variance.

Standard probability distributions: Discrete Uniform, Binomial, Poisson, and Normal. (25)

STAT-H-CC2-2-P **1 Credit**

(Descriptive Statistics II & Probability II) **PRACTICAL**

List of Suggested Practical

- Problems based on analysis of bivariate data.
- Problems based on measures of rank correlation.
- Problems based on analysis of categorical data.
- Finding expectation, variance from a given probability distribution.
- Fitting of binomial distributions for n and p = q = 1/2.
- Fitting of binomial distributions for given n and p.
- Fitting of binomial distributions after computing mean and variance.
- Fitting of Poisson distributions for given value of mean.
- Fitting of Poisson distributions after computing mean.
- Application problems based on binomial distribution.
- Application problems based on Poisson distribution.
- Problems based on area property of normal distribution.
- To find the ordinate for a given area for normal distribution.

- Application based problems using normal distribution.
- Fitting of normal distribution when parameters are given.
- Fitting of normal distribution when parameters are not given.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Agresti, A.: Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, 4th Edition, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Hogg, R.V., Tanis, E.A. and Rao J.M.: Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
- Myer, P.L.: Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STAT-H-SEC2-2-Th **2 credits**

(Artificial Intelligence for Everyone) **THEORY**

Unit 1: Introduction to Artificial Intelligence (6 Lectures)

- Definition and scope of AI
- Historical overview and key milestones
- Differentiating AI from human intelligence

Unit 2: AI Subfields and Technologies (6 Lectures)

- Machine learning: Supervised, unsupervised, and reinforcement learning
- Deep learning and neural networks
- Natural language processing (NLP) and computer vision

Unit 3: Applications of AI (8 Lectures)

- AI in healthcare: Diagnosis, treatment, and medical imaging
- AI in finance: Fraud detection, algorithmic trading, and risk assessment
- AI in transportation: Autonomous vehicles and traffic optimization
- AI in customer service and chatbots
- AI in education: Personalized learning and intelligent tutoring systems

Unit 4: Ethical and Social Implications of AI (5 Lectures)

- Bias and fairness in AI systems
- Privacy and data protection concerns
- Impact of AI on employment and the workforce
- AI and social inequality

Unit 5: Other Important Issues

- Ethical guidelines and responsible AI practices
- AI and Innovation
- Emerging trends and future directions in AI
- AI and creativity: Generative models and artistic applications

STAT-H-SEC2-2-TU **2 credits**

(Artificial Intelligence for Everyone) **TUTORIAL**

Term Paper In lieu of Tutorial paper: Students are required to submit a report which can be based on any topic as given below or as prescribed by the respective faculty/teachers.

Suggested problems for the reports (Term Paper) on Artificial intelligence (A.I.):

- A comparative study on Human Intelligence and Artificial Intelligence in Man-Made Machines.
- Chatbot Development for customer service support and integration of NLP for realistic/human type interactions.
- Development of a sentiment analysis tool that can classify and analyze the sentiment (positive, negative, neutral) in social media posts, reviews, or any text data.
- Use of AI in Image Recognition and Classification.
- Application of Linear regression and A.I/Machine learning on house price prediction- Estimating the selling price of a house based on its features, location, and market trends.
- Application of Linear regression and A.I/Machine learning on the data of Climate prediction.
- Application of Linear regression and A.I/Machine learning on the data of Bank related works.
- Application of Linear regression and A.I/Machine learning on Sales Forecasting: Predicting future sales based on historical sales data and trends.
- Application of Linear regression and A.I/Machine learning on Weather Forecasting, predicting future weather conditions by analyzing historical weather data.
- Application of Linear regression and A.I/Machine learning on Forecasting future energy usage based on past consumption patterns and environmental factors.
- Application of Linear regression and A.I/Machine learning on Risk Assessment: Predicting the likelihood of an event occurring, such as loan default or insurance claims, based on historical data and risk factors.
- Application of Linear regression and A.I/Machine learning on Health Outcome Prediction: Estimating patient outcomes or disease progression based on historical medical data and patient characteristics.
- Different types of Machine learning algorithms and its applications.
- Smart home automation solutions.
- Using AI in autonomous ground and Aerial vehicles.
- Application of AI in healthcare industry, Industrial automation, Smart farming and Internet of Things (IoT).
- Use of AI and its impact in future of education.
- Application of AI in Robotics and Computer Vision.

Note: *The topics listed here are merely examples. Students are encouraged to write a term paper in the form of a report on subjects not limited to these topics; additional suggestions can be provided by teachers, faculty members, or the department*.

STATISTICS MINOR

STAT-H-MC2-2-Th **3 Credits**

(Descriptive Statistics II & Probability II) **THEORY**

Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, fitting of polynomial and exponential curves, correlation ratio, correlation index, intraclass correlation.

Rank correlation: Spearman's and Kendall's measures. (15)

Analysis of Categorical Data: Contingency table, independence & association of attributes. (5)

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.) and probability density function (p.d.f.). Expectation and Variance.

Standard probability distributions: Discrete Uniform, Binomial, Poisson, and Normal. (25)

STAT-H-MC2-2-P **1 Credit**

(Descriptive Statistics II & Probability II) **PRACTICAL**

List of Suggested Practical

- Problems based on analysis of bivariate data.
- Problems based on measures of rank correlation.
- Problems based on analysis of categorical data.
- Finding expectation, variance from a given probability distribution.
- Fitting of binomial distributions for n and p = q = 1/2.
- Fitting of binomial distributions for given n and p.
- Fitting of binomial distributions after computing mean and variance.
- Fitting of Poisson distributions for given value of mean.
- Fitting of Poisson distributions after computing mean.
- Application problems based on binomial distribution.
- Application problems based on Poisson distribution.
- Problems based on area property of normal distribution.
- To find the ordinate for a given area for normal distribution.
- Application based problems using normal distribution.
- Fitting of normal distribution when parameters are given.
- Fitting of normal distribution when parameters are not given.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Agresti, A.: Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, 4th Edition, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Hogg, R.V., Tanis, E.A. and Rao J.M.: Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
- Myer, P.L.: Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STATISTICS INTERDISCIPLINARY

STAT-H-IDC2-2-Th **2 Credits**

(Statistics for Practitioners) **THEORY**

Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (arithmetic mean, median, mode), dispersion (range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (only through diagrams), Exploratory Data Analysis. (8)

Understanding bivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation (no tie), measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). (8)

Statistical Inference (*testing of hypothesis*): Basic idea of binomial and normal populations (graphical idea only, derivation of the properties excluded). Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and p-value for some standard testing problems like test for proportion/proportions, mean based on single (normal) sample, test on comparing means based on two-sample and paired sample data. (7)

Miscellaneous discussion: Applications of one-way and two-way ANOVA with one observation per cell (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimation of population mean and variability for finite population, idea and application of logistic regression for binary response data. (7)

STAT-H-IDC2-2-P **1 Credit**

(Statistics for Practitioners) **PRACTICAL**

List of Suggested Practical

- Measures of mean, median, mode, range, QD, SD, CV for univariate data case.
- Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearsons correlation coefficients, Spearman's Rank correlation, measures of association of attributes through contingency table.
- Tests for proportion/proportions, tests of means for single sample, two-sample, and paired sample data on normal response using p-value approach.
- Applications of ANOVA and Kruskal-Wallis test.
- Sample size determination, estimation of population mean and variability for finite population.
- Fitting of logistic regression for binary response data.

Reference Books:

- Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
- Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis, PHI.
- Hardle W. and Simar, L.: Applied Multivariate Statistical Analysis.
- Kutner, M.H. et.al.: Applied Linear Statistical Models.
- Belsley D.A. et.al.: Regression Diagnostics.
- Draper N.R. and Smith, H.: Applied Regression Analysis.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - III

STATISTICS MAJOR

STAT-H-CC3-3-Th **3 Credits**

(Real Analysis I) **THEORY**

Introduction to Real Numbers

Axiomatic definition of the real number system: Field Axioms, Order Axioms, Archimedean property: Intuition, bounded and unbounded sets, Supremum, Infimum, axiom of LUB. Rational and Irrational Numbers, Absolute value function. Intervals: Types of bounded, Unbounded Intervals. Neighbourhoods: ε-Neighbourhood, Open Neighbourhood. (10)

Sequences and Series

Definition of sequences, limits and Convergence: Uniqueness of Limits, Algebra of Limits, Squeeze Theorem. Monotone sequences and Monotone Convergence Theorem (MCT). Subsequences: Properties, Cauchy sequences. Bolzano-Weierstrass theorem. Series of real numbers: Cauchy criterion, Convergence tests: Comparison Test, Ratio Test, Root Test, Alternating Series Test (Leibniz's Test) (statements). Absolute and conditional convergence. (12)

Functions, Limits and Continuity

Bounded and unbounded functions, Monotone functions, Limits of functions, ε-δ definition, one-sided limits. Properties of limits: Uniqueness of Limits, Algebra of Limits: Sequential criterion, infinite limits, limits of +- infinity. Continuous functions: definition of continuity at a point (sequential characterization), continuity on a set, Algebraic Properties, Composition. Continuity on intervals, extreme value theorem. Intermediate value theorem. (10)

Differentiation

Definition of the derivative; interpretations of the derivative. Criteria for extrema. Theorems involving differentiable functions: Rolle's Theorem, Lagrange's Mean Value Theorem (MVT), Cauchy's Mean Value Theorem, Higher order derivatives, Taylor's Theorem (Statement only). (8)

Integration

Riemann integration: Definition, properties: Linearity, Additivity, Interval Reversal, Monotonicity, Integrability of monotone functions, continuous functions (without proof), Fundamental theorem of calculus. Improper integrals. (5)

STAT-H-CC3-3-TU **1 Credit**

(Real Analysis I) **TUTORIAL**

Reference Books

- Malik S.C. and Savita Arora: Mathematical Analysis, Second Edition, Wiley Eastern Limited, New Age International Limited, New Delhi.
- Somasundram, D. And Chaudhary, B.: A First Course in Mathematical Analysis, Narosa Publishing House, New Delhi.
- Gupta S.L. and Nisha Rani: Principles of Real Analysis, Vikas Publ. House Pvt. Ltd., New Delhi.
- Apostol, T.M: Mathematical Analysis, Second Edition, Narosa Publishing House, New Delhi.
- Shanti Narayan: A course of Mathematical Analysis, 12th revised Edition, S. Chand & Co. (Pvt.) Ltd., New Delhi.
- Singal, M.K. and Singal A.R: A First Course in Real Analysis, 24th Edition, R. Chand& Co., New Delhi.
- Bartle, R. G. and Sherbert, D. R.: Introduction to Real Analysis (3rd Edition), John Wiley and Sons (Asia) Pte. Ltd., Singapore.
- Ghorpade, Sudhir R. and Limaye, Balmohan V.: A Course in Calculus and Real Analysis, Undergraduate Texts in Mathematics, Springer (SIE), Indian reprint.
- Chakraborty, Arnab: Real Analysis, volumes 1,2,3, second edition. Sarat Book House.
- Walter Rudin: Principles of Mathematical Analysis
- Stephen Abbott: Understanding Analysis
- Robert G. Bartle and Donald R. Sherbert: Introduction to Real Analysis

STAT-H-CC4-3-Th **3 Credits**

(Statistical Inference I) **THEORY**

Basic concepts of Statistical Inference: population & sample, parameter & statistic, population distribution and sampling distribution. Point estimation, interval estimation and testing of hypothesis. Three useful distributions for statistical Inference: χ^2, t and F (derivations excluded). (5)

Point Estimation: Concepts of estimation, requirements of a good estimator, notions of mean square error, unbiasedness, bias-variance trade off, best linear unbiasedness and minimum variance unbiasedness. Properties of uniformly minimum variance unbiased estimators (UMVUE). Comparison of Estimators, Efficiency. Methods of Estimation: Method of moments, method of maximum likelihood estimation and statements of their small sample properties. Point estimators of the parameters of Binomial, Poisson, and univariate Normal distributions. (15)

Elements of hypothesis testing: Null and alternative hypotheses, simple & composite hypotheses, critical region, type I and type II errors, level of significance, size, power, p-value. Exact tests and confidence intervals: classical and p-value approaches. Tests relating to Binomial and Poisson distributions, Fisher's exact test. Chi-square tests for association, homogeneity and goodness of fit. Tests of hypotheses for the parameters of normal distribution (one sample and two sample problems), paired t-test. Combination of probabilities in tests of significance. (20)

Interval Estimation: Confidence Interval and Confidence Coefficient, Exact confidence interval for mean(s) and variance(s) for one and two sample problems under the Normal set-up. (5)

STAT-H-CC4-3-P **1 Credit**

(Statistical Inference I) **PRACTICAL**

List of Suggested Practical

- Maximum Likelihood Estimation.
- Estimation by the method of moments.
- Test of significance for single proportion and difference of two proportions.
- Test of significance for single Poisson mean and difference of two Poisson means.
- Chi square tests for association, homogeneity and goodness of fit.
- Test of significance and confidence intervals for single mean and difference of two means.
- Test of significance and confidence intervals for single variance and ratio of two variances.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory, Vol-1. World Press.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Hogg, R.V., Tanis, E.A. and Zimmerman, D. L.: Probability and Statistical Inference. Pearson Education.
- Johnson, R.A. and Bhattacharya, G.K.: Statistics-Principles and Methods, 4th Edn. John Wiley and Sons.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Gupta, S.C. and Kapoor, V.K. (2020): Fundamentals of Mathematical Statistics. Sultan Chand and Sons.
- Ramachandran, K.M. and Tsokos, C.P.: Mathematical Statistics with Applications. Academic Press.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STAT-H-SEC3-3-P **4 Credits**

(Introduction to R) **PRACTICAL**

Introduction to R:

Overview of R: History and Applications. Installation and Setup (R and RStudio). Basic R Syntax and Commands: Comments, Assignment, Printing Output, installing a Package, Loading a Package. R as a Calculator: Addition, Subtraction, Multiplication, Division, Exponentiation, Modulus (Remainder), Integer Division, Order of Operations (PEMDAS/BODMAS).

Data Types and Structures:

Basic Data Types (Numeric, Integer, Character, Logical). Vectors, Matrices, Arrays: Creation, Operations, Indexing. Lists and Data Frames: Structure, creating a list, Accessing Elements, adding an element, removing an element, creating a data frame, accessing columns, accessing rows, accessing specific elements, adding a new column, removing a column, Adding a new row. Factors and Categorical Data: Creating a factor, checking levels, converting a vector to a factor, creating an ordered factor, converting factor to numeric, Converting factor to character, Data frame with a factor variable.

Data Import and Export:

Reading Data from Files (CSV, Excel, etc.): Basic usage, Example with options, reading an Excel file Using “readxl”, “openxlsx” packages, Example with options, Reading TXT Files, Reading Data from R's Built-in Datasets. Writing Data to Files: Writing the data frame to a CSV file, Writing the data frame to an Excel file Using “writexl”, “openxlsx” packages, Writing the text vector to a file. Checking for missing values in a vector, removing rows with missing values, removing columns with missing values, Mean/Median/Mode Imputation, Handling “NA”s in common functions.

Basic Data Manipulation:

Subsetting Data: Extracting a specific column of a Data Frame, Extracting multiple columns of a Data Frame, Extracting specific rows and columns of a Data Frame, Extracting specific rows only of a Data Frame, Extracting rows of a Data Frame that meet a condition, Combine multiple conditions, Extracting specific elements of a Vector, Extracting elements of a Vector that meet a condition, Finding indices where the condition is true Using the “which()” Function. Sorting and Ordering Data: Sorting Vectors in Ascending Order, Descending Order, ordering a data frame by one or more columns use the “order()” function, Using “dplyr” package for Sorting. Merging and Joining Data Frames: Using “merge()” Function, Using “dplyr”, “data.table” Packages.

Data Visualization:

Base R Graphics (Histograms, Boxplots, Scatterplots): Creating histograms, Boxplots, Scatterplots. Introduction to ggplot. Customizing Plots (Titles, Labels, Legends): Basic Plot Customization like adding X-axis label, Y-axis label, Colour of points, Type of point, Adding Legends. Creating Multi-panel Plots: Using "par()" function for Base R Graphics.

Programming Basics:

Functions and Loops (for, while): basic syntax for defining a function, calling the function, For Loop, While Loop, using loops inside functions or vice versa. Conditional Statements (if, else): "if" Statement, "if-else" Statement, chaining multiple conditions using "if-else if-else" Statement, Nested "if" Statements, applying conditions to each element of a vector using "ifelse". Simulation in the context of Loop, Acceptance-Rejection method, illustration of repeated sampling and large sample properties through simulation. Numerical integration, Numerical solution of equations.

Statistical Analysis with R:

Descriptive Statistics: Mean, Median, Standard Deviation, Variance, Summary Statistics, Quantiles, Range, Interquartile Range. Probability Distributions and Random Number Generation: Density, Cumulative Distribution Function, Quantile Function, Random Generation from Normal Distribution, Binomial Distribution, Poisson Distribution, Exponential Distribution, and Uniform Distribution, Setting Seed for Reproducibility. Hypothesis Testing (t-tests, Chi-square tests): One-sample t-test, Two-sample t-test (independent), Paired t-test, Chi-square test for independence between two categorical variables. Correlation and Regression Analysis: Pearson Correlation, Spearman Correlation, Simple Linear Regression, Plotting the data and regression line, Multiple Linear Regression. ANOVA (Analysis of Variance): Conducting one-way ANOVA, Viewing ANOVA table, conducting two-way ANOVA with interaction, Viewing ANOVA table.

Advanced Data Manipulation:

Applying Functions to Data (apply, lapply, sapply).

Reference Books

- Gardener, M (2012) Beginning R: The Statistical Programming Language, Wiley Publications
- Braun W J, Murdoch D J (2007): A First Course in Statistical Programming with R. Cambridge University Press. New York
- A simple introduction to R by Arnab Chakraborty
  (freely available at http://www.isical.ac.in/~arnabc/)
- R for beginners by Emmanuel Paradis
  (freely available at https://cran.rproject.org/doc/contrib/Paradisrdebuts_en.pdf)
- Michael J. Crawley (2012): The R Book, Wiley
- Garrett Grolemund (2014): Hands-On Programming with R, O'Reilly Media

STATISTICS MINOR

STAT-H-MC1-3-Th **3 Credits**

(Descriptive Statistics I & Probability I) **THEORY**

Statistics: Definition and scope. Concepts of statistical population and sample.
Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous.
Scales of measurement: nominal, ordinal, interval and ratio.
Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays. (10)

Measures of Central Tendency: Mean, Median, Mode.
Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Ginis Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers. (15)

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications. (20)

STAT-H-MC1-3-P **1 Credit**

(Descriptive Statistics I & Probability I) **PRACTICAL**

List of Suggested Practical

- Diagrammatic representation of data.
- Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations, stem and leaf plot.
- Problems based on measures of central tendency.
- Problems based on measures of dispersion.
- Problems based on combined mean and variance and coefficient of variation.
- Problems based on moments, skewness and kurtosis.
- Problems related to quantiles and measures based on them, construction of box plot.
- Application problems based on Classical Definition of Probability.
- Application problems based on Bayes' Theorem.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STATISTICS INTERDISCIPLINARY

STAT-H-IDC3-3-Th **2 Credits**

(Statistics for Practitioners) **THEORY**

Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (arithmetic mean, median, mode), dispersion (range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (only through diagrams), Exploratory Data Analysis. (8)

Understanding bivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation (no tie), measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). (8)

Statistical Inference (*testing of hypothesis*): Basic idea of binomial and normal populations (graphical idea only, derivation of the properties excluded). Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and p-value for some standard testing problems like test for proportion/proportions, mean based on single (normal) sample, test on comparing means based on two-sample and paired sample data. (7)

Miscellaneous discussion: Applications of one-way and two-way ANOVA with one observation per cell (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimation of population mean and variability for finite population, idea and application of logistic regression for binary response data. (7)

STAT-H-IDC3-3-P **1 Credit**

(Statistics for Practitioners) **PRACTICAL**

List of Suggested Practical

- Measures of mean, median, mode, range, QD, SD, CV for univariate data case.
- Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearsons correlation coefficients, Spearman's Rank correlation, measures of association of attributes through contingency table.
- Tests for proportion/proportions, tests of means for single sample, two-sample, and paired sample data on normal response using p-value approach.
- Applications of ANOVA and Kruskal-Wallis test.
- Sample size determination, estimation of population mean and variability for finite population.
- Fitting of logistic regression for binary response data.

Reference Books:

- Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
- Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis, PHI.
- Hardle W. and Simar, L.: Applied Multivariate Statistical Analysis.
- Kutner, M.H. et.al.: Applied Linear Statistical Models.
- Belsley D.A. et.al.: Regression Diagnostics.
- Draper N.R. and Smith, H.: Applied Regression Analysis.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - IV

STATISTICS MAJOR

STAT-H-CC5-4-Th **3 Credits**

(Linear Algebra) **THEORY**

Vector Algebra:

Vectors as points in Euclidean spaces; subspaces of R^n; linear combinations and span; generators of subspaces; linear dependence and independence (of finite sets of vectors); intersections, sums and direct sums of subspaces; complements and projections; basis, characterizations of bases, extensions of linearly independent sets to bases; dimension, connection between dimensions of a pair of subspaces and those of their sum and intersection; (Euclidean) norm and inner product; triangle inequality and Cauchy-Schwarz inequality, orthogonality and orthonormality, Gram-Schmidt process, orthocomplements and orthogonal projections. (20)

Matrix Algebra:

Linear transformations and matrices (w.r.t. standard bases), examples, properties, transposes, compositions, elementary operations, diagonal, triangular, symmetric and skew-symmetric matrices, Echelon matrix, Kronecker product, range and rank, null space and nullity, rank-nullity theorem, ranks of products and sums, row and column spaces, invertibility, operations with partitioned matrices, orthogonal matrices, construction of non-singular and orthogonal matrices with given rows/columns. (12)

Trace and determinant of square matrices, properties, computations. (4)

Systems of simultaneous linear equations: consistency, solution spaces of homogeneous and non-homogeneous systems, existence and uniqueness of solutions, methods of solution. (4)

Eigenvalues and eigenvectors of square matrices, multiplicities, properties, characteristic polynomial, similarity, spectral decomposition of symmetric matrices. Quadratic forms, connections with eigenvalues, idempotent matrices. Cayley-Hamilton Theorem (Statement only) (5)

STAT-H-CC5-4-P **1 Credit**

(Linear Algebra) **PRACTICAL**

List of Suggested Practical

- Problems related to vectors and vector spaces – linear dependence and independence, spanning set and bases, orthogonality, projections.
- Problems on matrix operations and reductions.
- Problems on linear transformations.
- Problems on determinants and inverses.
- Problems on rank of matrices.
- Problems on solutions to system of linear equations.
- Problems on quadratic forms.
- Problems on eigen values and eigen vectors.

Reference Books

- Lay David C.: Linear Algebra and its Applications, Addison Wesley.
- Schaum's Outlines: Linear Algebra, Tata McGraw-Hill Edition, 3rd Edition.
- Krishnamurthy, V., Mainra V.P. and Arora J.L.: An Introduction to Linear Algebra (II, III, IV, V).
- Biswas, S.: A Textbook of Matrix Algebra, New Age International.
- Gupta, S.C.: An Introduction to Matrices (Reprint). Sultan Chand & Sons.
- Artin, M: Algebra. Prentice Hall of India.
- Datta, K.B: Matrix and Linear Algebra. Prentice Hall of India Pvt. Ltd.
- Hadley, G: Linear Algebra. Narosa Publishing House (Reprint).
- Searle, S.R.: Matrix Algebra Useful for Statistics. John Wiley &Sons.
- Chakraborty, Arnab: Linear Algebra, First Edition. Sarat Book House.
- Goon A.M.: Vectors and Matrices, World Press
- Das Roy, A. and Roy, D.: Linear Algebra with Application to Statistics, World Press

STAT-H-CC6-4-Th **3 Credits**

(Probability III) **THEORY**

Moments, Quantiles, Skewness and Kurtosis (including discussions of these measures for Binomial, Poisson and Univariate Normal Distributions). Probability inequalities (Markov's and Chebyshev's). Generating functions – probability generating function and moment generating function. Characteristic function (definition only). (10)

Discrete probability distributions: Uniform, Geometric, negative binomial, hypergeometric. Limiting/approximation cases. Truncated distributions. (5)

Continuous probability distributions: uniform, exponential, Cauchy, beta, gamma, lognormal, logistic, double exponential, Pareto and Weibull distributions along with their properties. Truncated distributions. (15)

Two dimensional random variables: joint, marginal and conditional distributions, properties of c.d.f., independence of variables, Theorems on sum and product of expectations of random variables, Conditional Expectation, Correlation and Regression. (10)

Trinomial distribution and its properties. Bivariate Normal Distribution and its properties. (5)

STAT-H-CC6-4-P **1 Credit**

(Probability III) **PRACTICAL**

List of Suggested Practical

- Fitting of negative binomial distribution.
- Application problems based on geometric, negative binomial and hypergeometric distributions.
- Fitting of continuous distributions.
- Application Problems based on continuous distributions.
- Application based Problems on trinomial distributions.
- Application based Problems on bivariate normal distributions.

Reference Books

- Hoel, P.G., Port, S.C. and Stone, C.J.: Introduction to Probability Theory. Houghton Mifflin Company.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Mood A.M, Graybill F.A. and Boes D.C: Introduction to the Theory of Statistics, McGraw Hill.
- Dasgupta, A.: Fundamentals of Probability – A First Course. Springer.
- Goon A.M., Gupta M.K. and Das Gupta. B.: Outline of Statistics, Vol. 1, World Press, Calcutta.
- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Casella, G. and Berger R.L.: Statistical Inference, 2nd Edition. Thomson Learning.
- Pitman, J.: Probability. Springer.
- Mukhopadhyay, P.: Mathematical Statistics. Books and Allied Publishers, Kolkata.
- Mukhopadhyay, N.: Probability and Statistical Inference. Marcel Dekker, New York.
- Chandra, T.K. and Chatterjee, D.: A First Course in Probability. Alpha Science.
- Cacoullos, T.: Exercises in Probability. Springer Verlag.
- Johnson, N.L., Kotz, S. and Balakrishnan, N.: Continuous Univariate Distributions, Volume 1 and 2. John Wiley and Sons.
- Johnson, N.L., Kemp, A.W. and Kotz, S.: Univariate Discrete Distributions. John Wiley and Sons.
- Ross, S.: A First Course in Probability. Pearson Education.

STAT-H-CC7-4-Th **3 Credits**

(Sampling Distributions and Statistical Inference II) **THEORY**

Concepts of random sampling. Sampling distribution of a statistic and its standard error. Distributions of statistics – illustrations using different distributions, reproductive properties of the distributions. *Order Statistics*: Introduction, distribution of the *r*-th order statistic, smallest and largest order statistics. Joint distribution of order statistics, distribution of sample median. (10)

Derivation of some standard sampling distributions: χ^2, t and F, sample mean and variance of a random sample from a univariate normal population, distributions of sample means, sample variances and sample correlation coefficient (null case) of a random sample from a bivariate normal population, distribution of the simple regression coefficient (for non-stochastic independent variable). (10)

Point Estimation: Sufficiency, factorization theorem (proof for discrete case only). Fisher's information (for single parameter only). Cramer-Rao inequality and minimum variance bound (MVB) estimators, Rao-Blackwell theorem and its applications. (10)

Theory of hypothesis testing: Test function, randomized and non-randomized tests, most powerful (MP) test, uniformly most powerful (UMP) test, Neyman - Pearson Lemma (statement and proof of sufficiency part only) and its applications to construct MP and UMP tests, uniformly most powerful unbiased (UMPU) tests (definition only).
Likelihood ratio test, properties of likelihood ratio test (without proof). (10)

Interval Estimation: Confidence intervals, Concepts of Uniformly Most Accurate (UMA) confidence sets, relationship with tests of hypotheses. (5)

STAT-H-CC7-4-P **1 Credit**

(Sampling Distributions and Statistical Inference II) **PRACTICAL**

List of Suggested Practical

- Most powerful critical region.
- Uniformly most powerful critical region.
- Unbiased critical region.
- Power curves.
- Confidence intervals and UMA confidence sets.
- Likelihood ratio tests for simple null hypothesis against simple alternative hypothesis.
- Likelihood ratio tests for simple null hypothesis against composite alternative hypothesis.

Reference Books:

- Goon A.M., Gupta M.K.: Das Gupta. B.: An Outline of Statistical Theory, Vol. I & II, World Press, Calcutta.
- Rohatgi V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
- Miller, I. and Miller, M.: John E. Freund's Mathematical Statistics (6th addition, low price edition), Prentice Hall of India.
- Mood A.M, Graybill F.A. and Boes D.C: Introduction to the Theory of Statistics, McGraw Hill.
- Casella, G. and Berger R.L.: Statistical Inference, 2nd Edn. Thomson Learning.
- Bhattacharya, P. K. and Burman, P.: Theory and Methods of Statistics. Academic Press.
- Mukhopadhyay, N.: Probability and Statistical Inference. Marcel Dekker, New York.
- Kale, B. K.: A First Course on Parametric Inference. Narosa Publishing House.
- Lehmann, E. L. and Romano, J. P.: Testing Statistical Hypothesis, 3rd Edition. Springer.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Mukhopadhyay, P.: Mathematical Statistics. Books and Allied Publishers, Kolkata.
- Rajagopalan, M. and Dhanavanthan, P.: Statistical Inference. Prentice Hall India.
- Srivastava, M.K. and Srivastava, N.: Statistical Inference: Testing of Hypotheses. Prentice Hall India.

STAT-H-CC8-4-Th **3 Credits**

(Design of Experiments I and Sample Survey I) **THEORY**

Design of Experiments I:

Analysis of Variance: Factors, types and effects; Fixed, random and mixed effects models; Analysis of one-way and two-way classified data with equal number of observations in each cell (Fixed Effects Models only). (10)

Experimental designs: Treatments, Experimental units & Blocks, Experimental error, Basic principles of Design of Experiments (Fisher). (2)

Basic designs: Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model, Applications of the techniques of ANOVA to the analysis of the above designs. Missing plot techniques in RBD and LSD. Uniformity trial experiments and comparison of designs. (10)

Sample Survey I:

Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, random numbers, procedure of selecting a sample, estimates of population mean, total and proportion, standard errors of these estimates, estimates of their standard errors. (15)

Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. (4)

Two-stage sampling (with primary units of equal size and equal selection probability at each stage): unbiased estimation of population mean and total. Ideas of snowball sampling, purposive sampling. (4)

STAT-H-CC8-4-P **1 Credit**

(Design of Experiments I and Sample Survey I) **PRACTICAL**

List of Suggested Practical

- Analysis of Variance of a one-way classified data (fixed effects model).
- Analysis of Variance of a two-way classified data with one observation per cell (fixed effects model).
- Analysis of Variance of a two-way classified data with more than one observation per cell (fixed effects model).
- Analysis of a CRD.
- Analysis of an RBD.
- Analysis of an LSD.
- Analysis of an RBD with one missing observation.
- Analysis of an LSD with one missing observation.
- To select a simple random sample with and without replacement.
- Simple random sampling – estimation of population mean, total and proportion; estimation of related standard error.
- Estimate the sample size for SRSWOR.
- Stratified Sampling –estimation of population mean and total, allocation of sample to strata by proportional and Neyman's methods, Comparison of the efficiencies of the above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Two-stage Sampling – estimation of population mean and total.

Reference Books:

- Renchner, A. C. And Schaalje, G. B.: Linear Models in Statistics (Second edition), John Wiley and Sons.
- Scheffe, H.: The Analysis of Variance, John Wiley.
- Cochran, W.G. and Cox, G.M.: Experimental Design. Asia Publishing House.
- Das, M.N. and Giri, N.C.: Design and Analysis of Experiments. Wiley Eastern Ltd.
- Kempthorne, O.: The Design and Analysis of Experiments. John Wiley.
- Montgomery, D. C.: Design and Analysis of Experiments, John Wiley.
- Wu, C. F. J. And Hamada, M.: Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Dean, A.M. and Voss, D.: Design and Analysis of Experiments. Springer Texts in Statistics.
- Goon, A.M., Gupta, M.K., Das Gupta, B.: An Outline of Statistical Theory, Vol-II, World Press, Calcutta.
- Goon A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol-II, World Press.
- Cochran, W.G.: Sampling Techniques (3rd Ed.), Wiley Eastern.
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C.: Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics.
- Murthy, M.N.: Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
- Des Raj and Chandhok P.: Sample Survey Theory, Narosa Publishing House.

STATISTICS MINOR

STAT-H-MC2-4-Th **3 Credits**

(Descriptive Statistics II & Probability II) **THEORY**

Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, fitting of polynomial and exponential curves, correlation ratio, correlation index, intraclass correlation.

Rank correlation: Spearman's and Kendall's measures. (15)

Analysis of Categorical Data: Contingency table, independence & association of attributes. (5)

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.) and probability density function (p.d.f.). Expectation and Variance.

Standard probability distributions: Discrete Uniform, Binomial, Poisson, and Normal. (25)

STAT-H-MC2-4-P **1 Credit**

(Descriptive Statistics II & Probability II) **PRACTICAL**

List of Suggested Practical

- Problems based on analysis of bivariate data.
- Problems based on measures of rank correlation.
- Problems based on analysis of categorical data.
- Finding expectation, variance from a given probability distribution.
- Fitting of binomial distributions for n and p = q = 1/2.
- Fitting of binomial distributions for given n and p.
- Fitting of binomial distributions after computing mean and variance.
- Fitting of Poisson distributions for given value of mean.
- Fitting of Poisson distributions after computing mean.
- Application problems based on binomial distribution.
- Application problems based on Poisson distribution.
- Problems based on area property of normal distribution.
- To find the ordinate for a given area for normal distribution.
- Application based problems using normal distribution.
- Fitting of normal distribution when parameters are given.
- Fitting of normal distribution when parameters are not given.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Agresti, A.: Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, 4th Edition, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Hogg, R.V., Tanis, E.A. and Rao J.M.: Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
- Myer, P.L.: Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - V

STATISTICS MAJOR

STAT-H-CC9-5-Th **3 Credits**

(Multivariate Analysis I) **THEORY**

Introduction to Multivariate Analysis: Concept of multivariate data. Notation and data structures (mean vector, dispersion and correlation matrices). Multiple regression, multiple correlation and partial correlation – their properties and related results. (12)

Random Vector: Joint Probability mass and density functions, Distribution Function, Mean vector, Dispersion matrix and Correlation Matrix, Marginal and Conditional Distributions, Independence of random vectors, Ellipsoid of Concentration, Multiple Regression, Multiple Correlation, Partial Correlation. Linear combinations of random variables and vectors. (17)

Canonical correlation: Definition, Conceptual overview, Mathematical formulation and computations, canonical weights and scores. Applications of canonical correlations. (4)

Multivariate Distributions: Multinomial, Dirichlet, Multivariate Normal distributions and their properties.

Sampling from multivariate normal distribution, MLEs, distributions and independence of sample mean vector and variance-covariance matrix.

Tests for Multiple and partial correlation coefficients. Tests for regression coefficients in multiple regression model. (12)

STAT-H-CC9-5-P **1 Credit**

(Multivariate Analysis I) **PRACTICAL**

List of Suggested Practical

- Problems on Multiple Linear Regression.
- Problems on Multiple and Partial Correlation Coefficients.
- Problems based on Tests for Multiple Correlation.
- Problems based on Tests for Partial Correlation.
- Problems on canonical correlations.
- Problems relating to Multinomial Distribution.
- Problems relating to Multivariate Normal Distribution.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2013): Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Anderson, T.W. (2003): An Introduction to Multivariate Statistical Analysis, 3rd Edn., John Wiley
- Muirhead, R.J. (1982): Aspects of Multivariate Statistical Theory, John Wiley.
- Mukhopadhyay, P. (1996): Mathematical Statistics. New Central Book Agency.
- Rohatgi, V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Rencher, A.C. and Christensen, W.F. (2012). Methods of Multivariate Analysis. Wiley.
- Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). Multivariate Analysis. Academic Press.
- Johnson, R.A. and Wichern, D.W. (2007) Applied Multivariate Statistical Analysis. 6th ed. Pearson Prentice Hall.

STAT-H-CC10-5-Th **3 Credits**

(Statistical Inference III) **THEORY**

Limit laws (Only statements and applications): Sequence of random variables, convergence in probability, convergence in mean square and convergence in distribution and their interrelations, Weak law of large numbers, Slutsky's Theorem, De-Moivre Laplace Limit theorem, Central Limit Theorem (C.L.T.) for i.i.d. variates. (7)

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, correlation coefficient and their uses in large sample tests under normality assumption.
Large sample distribution of sample quantiles.

Transformations of Statistics to stabilize variance: derivation and uses of Sin^{-1}, square root, logarithmic and z-transformations.
Large sample tests for binomial proportions, Poisson means (single and two independent samples cases) and correlation coefficients.
Large Sample distribution of Pearsonian χ^2- statistic.
Consistency and Asymptotic Efficiency of Estimators, CAN and BAN estimators, Large sample properties of maximum likelihood estimates and their uses: Likelihood Ratio test, Wald's test and Rao's Score Test. (20)

Nonparametric Tests: Introduction and Concept, Test for randomness based on total number of runs, Empirical distribution function, One Sample Tests: Kolmogrov- Smirnov, Sign, Signed rank. Two-Sample Tests: Wilcoxon rank sum and Wilcoxon-Mann-Whitney test. Multi-sample test: Kruskal-Wallis test. (12)

Sequential Probability Ratio Tests: Introduction, Wald's formulation of SPRT, Operating Characteristic Function, Average Sample Number. (6)

STAT-H-CC10-5-P **1 Credit**

(Statistical Inference III) **PRACTICAL**

List of Suggested Practical

- Test of significance and confidence intervals for single proportion and difference of two proportions using CLT.
- Test of significance and confidence intervals for single Poisson mean and difference of two Poisson means using CLT.
- Test of significance and confidence intervals concerning sample standard deviation, coefficient of variation and correlation coefficient (both single sample and two sample cases).
- Test of significance and confidence intervals using variance stabilizing transformations.
- Test for randomness based on total number of runs.
- Kolmogrov Smirnov test for one sample.
- Sign test.
- Wilcoxon-Mann-Whitney U-test.
- Kruskal-Wallis test.
- SPRT and approximate calculations of OC and ASN functions.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2013): Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2016): An Outline of Statistical Theory, Vol-1. World Press.
- Ferguson, Thomas S. (1996). A Course in Large Sample Theory, 1st Edition. Chapman & Hall.
- Serfling R.J. (1980): Approximation Theory of Mathematical Statistics, John Wiley.
- Chandra T.K. (1999): A First Course in Asymptotic Theory in Statistics, Narosa.
- Rohatgi, V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Mood, A.M., Graybill, F.A. and Boes, D.C. (2017): Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill.
- Hogg, R.V., McKean, J.W. and Craig, A.T. (2021): Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Gibbons, J.D. and Chakraborti, S. (2003): Nonparametric Statistical Inference. 4th ed. Marcel Dekker.

STAT-H-CC11-5-Th **3 Credits**

(Linear Models) **THEORY**

Gauss-Markov set-up: General linear model, General linear hypothesis, Theory of linear estimation, Estimability of linear parametric functions, Method of least squares, Gauss-Markov theorem, Estimation of error variance. Fundamental Theorems on least squares (statements only), Orthogonal splitting of total variation, selection of valid error. (10)

Analysis of variance: Fixed, random and mixed effect models. Analysis of variance in one-way balanced random effect model. Two-way classified data with equal number of observations per cell for random and mixed effect models. Inference on variance components. (10)

Regression analysis: Simple Linear regression, Multiple Linear Regression. Estimation and hypothesis testing in case of simple and multiple regression models. (15)

Analysis of covariance: Analysis of covariance (with one concomitant variable) in one-way and two-way classified data with equal number of observations per cell. (10)

STAT-H-CC11-5-P **1 Credit**

(Linear Models) **PRACTICAL**

List of Suggested Practical

- Analysis of Variance of a one-way classified data (random effects model).
- Analysis of Variance of a two-way classified data with equal number of observations per cell (random effects model).
- Analysis of Variance of a two-way classified data with equal number of observations per cell (mixed effects model).
- Problems on testing of hypotheses relating to simple linear regression.
- Problems on testing of hypotheses relating to multiple linear regression.
- Analysis of covariance (with one concomitant variable) in one-way classified.
- Analysis of covariance (with one concomitant variable) in two-way classified data with equal number of observations per cell.

Reference Books:

- Goon A.M., Gupta M.K.: Das Gupta. B.: An Outline of Statistical Theory, Vol. II, World Press, Calcutta.
- Weisberg, S. (2005): Applied Linear Regression (Third edition). Wiley.
- Wu, C. F. J. And Hamada, M. (2009): Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Renchner, A. C. And Schaalje, G. B. (2008): Linear Models in Statistics (Second edition), John Wiley and Sons.
- Scheffe, H. (1959): The Analysis of Variance, John Wiley.
- Chatterjee S., Hadi A.S., Price B.: Regression Analysis by Example, 3rd Edn, John Wiley & Sons.

STAT-H-CC12-5-Th **3 Credits**

(Demography, Index Numbers and Psychometry) **THEORY**

Demography:

Basic Concepts: Definition and scope of demography. Measurement of population: rates and ratios. (2)

Measures of Mortality and Fertility: Crude Death Rate (CDR), Age-Specific Death Rate (ASDR), Standardized Mortality rate (SMR), Infant Mortality Rate (IMR). Crude Birth Rate (CBR), Age-Specific Fertility Rate (ASFR), General Fertility Rate (GFR), Total Fertility Rate (TFR).

Reproduction Rates: Gross Reproduction Rate (GRR), Net Reproduction Rate (NRR). (6)

Life Tables: Description and uses of life table functions: $l_x, d_x, q_x, L_x, T_x, e_x$. Construction of a life table. (4)

Measures of Population Change: Measurement of population growth: Arithmetic, Geometric, Exponential. Population projection: Concept of stable and stationary population. (3)

Index Numbers:

Introduction to Index Numbers: Meaning and uses of index numbers. Price, quantity, and value index numbers. Simple and weighted index numbers. Errors in index numbers. (4)

Methods of Construction: Laspeyres', Paasche's, Marshall-Edgeworth, Fisher's Ideal, and Kelly's methods. Cost of Living Index (Consumer Price Index), Wholesale Price Index. Time Reversal and Factor Reversal Tests. (9)

Chain Base and Fixed Base Index: Link relative method. Chain index and fixed base index. Base Shifting and Deflating: Base shifting of index numbers. Deflating economic series. (2)

Psychometry:

Basics of Psychometry: Concept of psychological and educational measurement. Attributes: Qualitative vs Quantitative. Scales of measurement: Nominal, Ordinal, Interval, Ratio. (2)

Different types of scaling: Z-scale, T-scale, Percentile-Scale, Equivalent Scale. (3)

Test Construction: Characteristics of a good test: Validity, Reliability, Objectivity, Usability. Methods of reliability estimation: Test-retest, Parallel forms, Split-half, Kuder-Richardson method, Cronbach's alpha. Types of validity: Content, Construct, Criterion-related. (7)

Attitude and Personality Measurement: Likert scale, Thurstone scale. Intelligence and aptitude testing. (3)

STAT-H-CC12-5-P **1 Credit**

(Demography, Index Numbers and Psychometry) **PRACTICAL**

List of Suggested Practical

- Calculate Crude Birth Rate (CBR), General Fertility Rate (GFR), and Age-Specific Fertility Rate (ASFR) using given population and birth data.
- Compute Crude Death Rate (CDR), Infant Mortality Rate (IMR), and Age-Specific Death Rate (ASDR) from a tabulated dataset.
- Construct an abridged life table from given mortality data - calculate $l_x, d_x, q_x, L_x, T_x, e_x$.
- Compute Arithmetic, Geometric, and Exponential growth rates of population using data from two or more time points.
- Construct Price Index Numbers using Laspeyres', Paasche's, and Fisher's Ideal formulae from given price and quantity data.
- Verify Time Reversal and Factor Reversal Tests for Fisher's Index using the above calculation.
- Compute Cost of Living Index from household data.
- Construct Fixed Base and Chain Base Index Numbers.
- Perform Base Shifting and Deflation on a given price index series and interpret results.
- Calculate reliability using the Split-Half Method from a given dataset.
- Compute the Difficulty Index and Discrimination Index for a set of objective test items with student responses.
- Given the inter-item correlations, compute Cronbach's alpha for internal consistency estimation (small dataset).

Reference Books:

- Bhende, A. A. & Kanitkar, T. (2010): Principles of population studies (22nd ed.). Himalaya Publishing House.
- Pathak, K. B. & Ram, F. (1998): Techniques of demographic analysis. Himalaya Publishing House.
- Bogue, D. J. (1969): Principles of demography. John Wiley & Sons.
- Rowland, D. T. (2003): Demographic methods and concepts. Oxford University Press.
- Goon, A. M., Gupta, M. K. & Dasgupta, B. (1986): Fundamentals of statistics (Vol. 2) (6th ed.). World Press.
- Gupta, S. C. & Kapoor, V. K. (2014): Fundamentals of applied statistics (4th ed.). Sultan Chand & Sons.
- Spiegel, M. R., Schiller, J. & Srinivasan, R. A. (2017): Schaum's outline of statistics (5th ed.). McGraw- Hill Education.
- Kendall, M. G. & Stuart, A. (1977): The advanced theory of statistics (Vol. 1) (4th ed.). Charles Griffin & Company.
- Guilford, J. P. (1954): Psychometric methods (2nd ed.). McGraw-Hill.
- Nunnally, J. C. & Bernstein, I. H. (1994): Psychometric theory (3rd ed.). McGraw-Hill.
- Anastasi, A. & Urbina, S. (1997): Psychological testing (7th ed.). Prentice Hall.
- Kline, P. (2000): The handbook of psychological testing (2nd ed.). Routledge.
- Singh, A. K. (2016): Tests, measurements and research methods in behavioural sciences (6th ed.). Bharati Bhawan.

STATISTICS MINOR

STAT-H-MC3-5-Th **3 Credits**

(Statistical Inference I) **THEORY**

Basic concepts of Statistical Inference: population & sample, parameter & statistic, population distribution and sampling distribution. Point estimation, interval estimation and testing of hypothesis. Three useful distributions for statistical Inference: χ^2, t and F (derivations excluded). (5)

Point Estimation: Concepts of estimation, requirements of a good estimator, notions of mean square error, unbiasedness, bias-variance trade off, best linear unbiasedness and minimum variance unbiasedness. Properties of uniformly minimum variance unbiased estimators (UMVUE). Comparison of Estimators, Efficiency. Methods of Estimation: Method of moments, method of maximum likelihood estimation and statements of their small sample properties. Point estimators of the parameters of Binomial, Poisson, and univariate Normal distributions. (15)

Elements of hypothesis testing: Null and alternative hypotheses, simple & composite hypotheses, critical region, type I and type II errors, level of significance, size, power, p-value. Exact tests and confidence intervals: classical and p-value approaches. Tests relating to Binomial and Poisson distributions, Fisher's exact test. Chi-square tests for association, homogeneity and goodness of fit. Tests of hypotheses for the parameters of normal distribution (one sample and two sample problems), paired t-test. Combination of probabilities in tests of significance. (20)

Interval Estimation: Confidence Interval and Confidence Coefficient, Exact confidence interval for mean(s) and variance(s) for one and two sample problems under the Normal set-up. (5)

STAT-H-MC3-5-P **1 Credit**

(Statistical Inference I) **PRACTICAL**

List of Suggested Practical

- Maximum Likelihood Estimation.
- Estimation by the method of moments.
- Test of significance for single proportion and difference of two proportions.
- Test of significance for single Poisson mean and difference of two Poisson means.
- Chi square tests for association, homogeneity and goodness of fit.
- Test of significance and confidence intervals for single mean and difference of two means.
- Test of significance and confidence intervals for single variance and ratio of two variances.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory, Vol-1. World Press.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Hogg, R.V., Tanis, E.A. and Zimmerman, D. L.: Probability and Statistical Inference. Pearson Education.
- Johnson, R.A. and Bhattacharya, G.K.: Statistics-Principles and Methods, 4th Edn. John Wiley and Sons.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Gupta, S.C. and Kapoor, V.K. (2020): Fundamentals of Mathematical Statistics. Sultan Chand and Sons.
- Ramachandran, K.M. and Tsokos, C.P.: Mathematical Statistics with Applications. Academic Press.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STAT-H-MC4-5-Th **3 Credits**

(Design of Experiments I and Sample Survey I) **THEORY**

Design of Experiments I:

Analysis of Variance: Factors, types and effects; Fixed, random and mixed effects models; Analysis of one-way and two-way classified data with equal number of observations in each cell (Fixed Effects Models only). (10)

Experimental designs: Treatments, Experimental units & Blocks, Experimental error, Basic principles of Design of Experiments (Fisher). (2)

Basic designs: Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model, Applications of the techniques of ANOVA to the analysis of the above designs. Missing plot techniques in RBD and LSD. Uniformity trial experiments and comparison of designs. (10)

Sample Survey I:

Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, random numbers, procedure of selecting a sample, estimates of population mean, total and proportion, standard errors of these estimates, estimates of their standard errors. (15)

Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. (4)

Two-stage sampling (with primary units of equal size and equal selection probability at each stage): unbiased estimation of population mean and total. Ideas of snowball sampling, purposive sampling. (4)

STAT-H-MC4-5-P **1 Credit**

(Design of Experiments I and Sample Survey I) **PRACTICAL**

List of Suggested Practical

- Analysis of Variance of a one-way classified data (fixed effects model).
- Analysis of Variance of a two-way classified data with one observation per cell (fixed effects model).
- Analysis of Variance of a two-way classified data with more than one observation per cell (fixed effects model).
- Analysis of a CRD.
- Analysis of an RBD.
- Analysis of an LSD.
- Analysis of an RBD with one missing observation.
- Analysis of an LSD with one missing observation.
- To select a simple random sample with and without replacement.
- Simple random sampling – estimation of population mean, total and proportion; estimation of related standard error.
- Estimate the sample size for SRSWOR.
- Stratified Sampling –estimation of population mean and total, allocation of sample to strata by proportional and Neyman's methods, Comparison of the efficiencies of the above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Two-stage Sampling – estimation of population mean and total.

Reference Books:

- Renchner, A. C. And Schaalje, G. B.: Linear Models in Statistics (Second edition), John Wiley and Sons.
- Scheffe, H.: The Analysis of Variance, John Wiley.
- Cochran, W.G. and Cox, G.M.: Experimental Design. Asia Publishing House.
- Das, M.N. and Giri, N.C.: Design and Analysis of Experiments. Wiley Eastern Ltd.
- Kempthorne, O.: The Design and Analysis of Experiments. John Wiley.
- Montgomery, D. C.: Design and Analysis of Experiments, John Wiley.
- Wu, C. F. J. And Hamada, M.: Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Dean, A.M. and Voss, D.: Design and Analysis of Experiments. Springer Texts in Statistics.
- Goon, A.M., Gupta, M.K., Das Gupta, B.: An Outline of Statistical Theory, Vol-II, World Press, Calcutta.
- Goon A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol-II, World Press.
- Cochran, W.G.: Sampling Techniques (3rd Ed.), Wiley Eastern.
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C.: Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics.
- Murthy, M.N.: Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
- Des Raj and Chandhok P.: Sample Survey Theory, Narosa Publishing House.

SEMESTER - VI

STATISTICS MAJOR

STAT-H-CC13-6-Th **3 Credits**

(Applied Multivariate Analysis) **THEORY**

Types of multivariate techniques: Dependence vs. interdependence methods. (2)

Principal Component Analysis (PCA): Objectives and geometrical interpretation. Computation of principal components. Standardization and correlation matrix-based PCA. Scree plot, proportion of variance explained. Applications and interpretation. Dimensionality reduction and visualization. (13)

Factor Analysis (FA): Basic model: common and unique factors. Estimation of factor loadings (Principal Component Method). Varimax Rotation method (concept only). Factor scores and interpretation. Applications and limitations. (10)

Cluster Analysis: Similarity and dissimilarity measures. Distances: Euclidean, Manhattan, Mahalanobis (conceptual). Hierarchical clustering: single-linkage, complete- linkage, average-linkage. Dendrogram and its interpretation. Non-hierarchical clustering: k-means algorithm. Choosing the number of clusters. Applications in market segmentation, social sciences, etc. (12)

Discriminant Analysis: Problem of classification. Fisher's linear discriminant function. Assumptions and classification rule. Classification with two groups. Cost of misclassification. Applications in biology, marketing, etc. (8)

STAT-H-CC13-6-P **1 Credit**

(Applied Multivariate Analysis) **PRACTICAL**

List of Suggested Practical (Using R)

Data Handling for Multivariate Data:

- Reading multivariate datasets from various file formats (e.g., CSV, Excel) into R.
- Visualizing multivariate data using techniques like scatterplot matrices, and correlation heatmaps.

Principal Component Analysis (PCA):

- Performing PCA on a given dataset using both the covariance matrix and the correlation matrix.
- Visualizing the scree plot to determine the number of principal components to retain.
- Interpreting the principal components based on their loadings and the proportion of variance explained.
- Plotting the data on the first few principal components for dimensionality reduction and visualization.

Factor Analysis (FA):

- Applying Factor Analysis to a dataset, focusing on the Principal Component Method for estimation of factor loadings.
- Performing Varimax rotation on the factor loadings.
- Interpreting the rotated factor loadings to identify underlying factors.
- Calculating and interpreting factor scores for each observation.
- Visualizing factor loadings (e.g., loading plots).

Cluster Analysis:

- Computing various distance/dissimilarity matrices between observations (e.g., Euclidean distance, Manhattan distance).
- Performing different hierarchical clustering methods (e.g., single-linkage, complete-linkage, average- linkage) and visualizing the results using dendrograms.
- Interpreting dendrograms to identify potential clusters.
- Applying non-hierarchical clustering algorithms like k-means clustering.
- Methods for choosing the number of clusters (e.g., Elbow method, Silhouette score).
- Visualizing the resulting clusters (e.g., scatter plots with cluster assignments).

Discriminant Analysis:

- Applying Fisher's linear discriminant analysis to a dataset with two known groups.
- Developing a classification rule based on the discriminant function.
- Constructing and interpreting a confusion matrix.
- Calculating and interpreting the misclassification rate and considering the cost of misclassification.
- Visualizing group separation based on the discriminant function.

Reference Books

- Anderson, T. W. (2003): An Introduction To Multivariate Statistical Analysis (3rd ed.). Hoboken, NJ: Wiley.
- Chatfield, C. & Collins, A. J. (1980): Introduction to Multivariate Analysis. London: Chapman and Hall.
- Everitt, B. S. & Hothorn, T. (2011): An Introduction to Applied Multivariate Analysis with R. New York, NY: Springer.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019): Multivariate Data Analysis (8th ed.).

STAT-H-CC14-6-Th **3 Credits**

(Regression Analysis) **THEORY**

Gauss Markov Model: Correlated set-up, least squares estimate with restriction on parameters. (3)

Building a regression model: Transformations – Box-Cox models. (7)

Dummy variables, piecewise regression, splines and scatter plot smoothing. (6)

Detection of outliers and influential observations: Residuals and leverages, DFBETA, DFFIT, Cook's Distance. (4)

Regression Diagnostics: Model checking: Prediction from a fitted model. Variable selection (adjusted R^2, Cp criteria, AIC). (12)

Multicollinearity: Detection (condition number and VIF) and remedial measures. (8)

Binary data regression: Logistic Regression (Grouped and ungrouped data). (5)

STAT-H-CC14-6-P **1 Credit**

(Regression Analysis) **PRACTICAL**

List of Suggested Practical

- Problems on transformation
- Problems on using dummy variables
- Problems on outlier detection
- Problems on variable selection
- Problems on multicollinearity – detection and remedial measures
- Problems on fitting logistic regression

Reference Books

- C.R. Rao: Linear Statistical Inference and its Applications.
- R. B. Bapat: Linear Algebra and Linear Model.
- H. Scheffe: The Analysis of Variance.
- S.R. Searle: Linear Models.
- G.A.F. Seber: Linear Regression Analysis.
- N.C. Giri: Analysis of Variance.
- N.R. Draper & H. Smith: Applied Regression Analysis.
- D.W. Belsley, E. Kuh & R.E. Welsch: Regression Diagnostics – identifying Influential data & sources of collinearity.
- J. Rousseeuw & A.M. Leroy: Robust Regression & Outlier Detection.
- R.D. Cook & S. Weisberg: Residual and its Influence in Regression.
- S. Chatterjee & A.S. Hadi: Regression Analysis by Example.
- J. Johnston: Econometric Methods (3rd ed.).
- G.G. Judge, W.E. Griffith, et. al.: The Theory and Practice of Econometrics (2nd ed.).
- T.P. Ryan: Modern Regression Methods (2nd ed.).
- A. Agresti: Analysis of Ordinal Categorical Data.
- A. Agresti: Categorical Data Analysis.

STAT-H-CC15-6-Th **3 Credits**

(Time Series Analysis) **THEORY**

Introduction to Time Series: Definition and applications of time series. Objectives of time series analysis. Stationary and evolutionary processes. Trend and Seasonality. Additive and multiplicative models. Graphical representation of time series data. (5)

Trend Analysis: Deterministic trend models: linear, quadratic, exponential. Estimation of trend using: Moving averages (Simple, Centered), Least squares method. Deseasonalization. (8)

Seasonal and Cyclical Components: Seasonal variation: Identification. Methods of seasonal adjustment: Ratio to moving average method, Ratio to trend method, Link relative method. Concept of cyclical variation and its distinction from seasonality. (9)

Stationarity and Autocorrelation: Concept of stationarity. Autocorrelation function (ACF) and partial auto-correlation function (PACF). White noise and random walk models. (8)

Time Series Models: AR (Auto-Regressive), MA (Moving Average), ARMA, and ARIMA Models. Forecasting. Test of randomness: Pormanteau and Ljung-Box. Concept of model selection (AIC: brief idea). (15)

STAT-H-CC15-6-P **1 Credit**

(Time Series Analysis) **PRACTICAL**

List of Suggested Practical

- Graphical Representation and Component Identification: Plotting given time series data and identifying trend, seasonal, and irregular components.
- Fitting Deterministic Trend Models: Fit a linear trend model to time series data using the least squares method. Fit a quadratic trend model using least squares method. Fit an exponential trend model by transforming data logarithmically and applying least squares.
- Estimation of Trend Using Moving Averages: Compute simple moving average (odd and even periods). Apply centered moving average to smooth data. Extract trend values and interpret the result.
- Deseasonalization: Remove seasonality from given data using seasonal indices.
- Estimation of Seasonal Indices: Compute seasonal indices using: Ratio to Moving Average method, Ratio to Trend method, Link Relative method.
- Calculation and Plotting of ACF and PACF: Compute the autocorrelation coefficients up to lag 5 or 10 (using formula). Plot ACF and PACF and interpret lag behavior (e.g., identify AR or MA process nature).
- Forecasting Using Fitted ARIMA Models: Given AR model parameters, compute one-step ahead or two-step ahead forecasts. Compute forecasts using a fitted ARIMA(1,0,0) or ARIMA(0,1,1) model.

Reference Books:

- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015): Time series analysis: Forecasting and control (5th ed.). Hoboken, NJ: Wiley.
- Chatfield, C. (2004): The analysis of time series: An introduction (6th ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Anderson, T. W. (1971): The statistical analysis of time series. New York, NY: Wiley.
- Kendall, M. G., & Ord, J. K. (1990): Time series (3rd ed.). London, UK: Edward Arnold.
- Shumway, R. H., & Stoffer, D. S. (2017): Time series analysis and its applications: With R examples (4th ed.). Cham, Switzerland: Springer.

STATISTICS MINOR

STAT-H-MC3-6-Th **3 Credits**

(Statistical Inference I) **THEORY**

Basic concepts of Statistical Inference: population & sample, parameter & statistic, population distribution and sampling distribution. Point estimation, interval estimation and testing of hypothesis. Three useful distributions for statistical Inference: χ^2, t and F (derivations excluded). (5)

Point Estimation: Concepts of estimation, requirements of a good estimator, notions of mean square error, unbiasedness, bias-variance trade off, best linear unbiasedness and minimum variance unbiasedness. Properties of uniformly minimum variance unbiased estimators (UMVUE). Comparison of Estimators, Efficiency. Methods of Estimation: Method of moments, method of maximum likelihood estimation and statements of their small sample properties. Point estimators of the parameters of Binomial, Poisson, and univariate Normal distributions. (15)

Elements of hypothesis testing: Null and alternative hypotheses, simple & composite hypotheses, critical region, type I and type II errors, level of significance, size, power, p-value. Exact tests and confidence intervals: classical and p-value approaches. Tests relating to Binomial and Poisson distributions, Fisher's exact test. Chi-square tests for association, homogeneity and goodness of fit. Tests of hypotheses for the parameters of normal distribution (one sample and two sample problems), paired t-test. Combination of probabilities in tests of significance. (20)

Interval Estimation: Confidence Interval and Confidence Coefficient, Exact confidence interval for mean(s) and variance(s) for one and two sample problems under the Normal set-up. (5)

STAT-H-MC3-6-P **1 Credit**

(Statistical Inference I) **PRACTICAL**

List of Suggested Practical

- Maximum Likelihood Estimation.
- Estimation by the method of moments.
- Test of significance for single proportion and difference of two proportions.
- Test of significance for single Poisson mean and difference of two Poisson means.
- Chi square tests for association, homogeneity and goodness of fit.
- Test of significance and confidence intervals for single mean and difference of two means.
- Test of significance and confidence intervals for single variance and ratio of two variances.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory, Vol-1. World Press.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Hogg, R.V., Tanis, E.A. and Zimmerman, D. L.: Probability and Statistical Inference. Pearson Education.
- Johnson, R.A. and Bhattacharya, G.K.: Statistics-Principles and Methods, 4th Edn. John Wiley and Sons.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Gupta, S.C. and Kapoor, V.K. (2020): Fundamentals of Mathematical Statistics. Sultan Chand and Sons.
- Ramachandran, K.M. and Tsokos, C.P.: Mathematical Statistics with Applications. Academic Press.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STAT-H-MC4-6-Th **3 Credits**

(Design of Experiments I and Sample Survey I) **THEORY**

Design of Experiments I:

Analysis of Variance: Factors, types and effects; Fixed, random and mixed effects models; Analysis of one-way and two-way classified data with equal number of observations in each cell (Fixed Effects Models only). (10)

Experimental designs: Treatments, Experimental units & Blocks, Experimental error, Basic principles of Design of Experiments (Fisher). (2)

Basic designs: Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model, Applications of the techniques of ANOVA to the analysis of the above designs. Missing plot techniques in RBD and LSD. Uniformity trial experiments and comparison of designs. (10)

Sample Survey I:

Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, random numbers, procedure of selecting a sample, estimates of population mean, total and proportion, standard errors of these estimates, estimates of their standard errors. (15)

Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. (4)

Two-stage sampling (with primary units of equal size and equal selection probability at each stage): unbiased estimation of population mean and total. Ideas of snowball sampling, purposive sampling. (4)

STAT-H-MC4-6-P **1 Credit**

(Design of Experiments I and Sample Survey I) **PRACTICAL**

List of Suggested Practical

- Analysis of Variance of a one-way classified data (fixed effects model).
- Analysis of Variance of a two-way classified data with one observation per cell (fixed effects model).
- Analysis of Variance of a two-way classified data with more than one observation per cell (fixed effects model).
- Analysis of a CRD.
- Analysis of an RBD.
- Analysis of an LSD.
- Analysis of an RBD with one missing observation.
- Analysis of an LSD with one missing observation.
- To select a simple random sample with and without replacement.
- Simple random sampling – estimation of population mean, total and proportion; estimation of related standard error.
- Estimate the sample size for SRSWOR.
- Stratified Sampling –estimation of population mean and total, allocation of sample to strata by proportional and Neyman's methods, Comparison of the efficiencies of the above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Two-stage Sampling – estimation of population mean and total.

Reference Books:

- Renchner, A. C. And Schaalje, G. B.: Linear Models in Statistics (Second edition), John Wiley and Sons.
- Scheffe, H.: The Analysis of Variance, John Wiley.
- Cochran, W.G. and Cox, G.M.: Experimental Design. Asia Publishing House.
- Das, M.N. and Giri, N.C.: Design and Analysis of Experiments. Wiley Eastern Ltd.
- Kempthorne, O.: The Design and Analysis of Experiments. John Wiley.
- Montgomery, D. C.: Design and Analysis of Experiments, John Wiley.
- Wu, C. F. J. And Hamada, M.: Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Dean, A.M. and Voss, D.: Design and Analysis of Experiments. Springer Texts in Statistics.
- Goon, A.M., Gupta, M.K., Das Gupta, B.: An Outline of Statistical Theory, Vol-II, World Press, Calcutta.
- Goon A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol-II, World Press.
- Cochran, W.G.: Sampling Techniques (3rd Ed.), Wiley Eastern.
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C.: Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics.
- Murthy, M.N.: Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
- Des Raj and Chandhok P.: Sample Survey Theory, Narosa Publishing House.

SEMESTER - VII

STATISTICS MAJOR

STAT-H-CC16-7-Th **3 Credits**

(Real Analysis II and Multivariate Analysis II) **THEORY**

Bounded and unbounded sets, neighbourhoods, open, closed and compact sets in one and several dimensions; closures, interiors and boundaries.

Sequences in several dimensions: Boundedness, convergence, Bolzano-Weierstrass Theorem.

Functions of several variables: Limit, continuity, uniform continuity. (9)

Differentiation of multivariable functions: Directional, partial derivatives, total differentiability, chain rule and mean value theorem, higher order derivatives, equality of mixed partial derivatives, Taylor's theorems. Maxima and minima of multivariable functions, Hessian matrix and second derivative test.

Multiple Integral: Double and triple integrals, iterated integrals, Fubini's theorem, change of variables and Jacobian. (11)

Univariate sampling distributions based on normal: Non-central χ^2, t & F distributions – definitions and properties (statements only). Distribution of quadratic forms, Fisher-Cochran theorem, applications; independence of quadratic and affine forms. (8)

Wishart distribution, properties of Wishart distribution, distribution of conditional dispersion matrix.

Hotelling T^2: Definition, distribution and applications in testing in single and two-sample cases and confidence set construction. (11)

Multivariate Analysis of Variance (MANOVA): Generalization of ANOVA. Hypothesis testing in MANOVA. Wilks' lambda. Assumptions and interpretation. (6)

STAT-H-CC16-7-P **1 Credit**

(Real Analysis II and Multivariate Analysis II) **PRACTICAL**

List of Suggested Practical

- Problems on Maxima and Minima of multivariable functions.
- Problems based on non-central χ^2, t, and F Distributions.
- Problems based on Quadratic Forms and applications of Fisher–Cochran Theorem.
- Problems based on Wishart Distribution and Hotelling's T^2.
- Problems based on MANOVA for Small Dataset.

Reference Books

- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2016): An Outline of Statistical Theory, Vol-2. World Press.
- Anderson, T.W. (2003): An Introduction to Multivariate Statistical Analysis, 3rd Edn., John Wiley.
- Muirhead, R.J. (1982): Aspects of Multivariate Statistical Theory, John Wiley.
- Mukhopadhyay, P. (1996): Mathematical Statistics. New Central Book Agency.
- Rencher, A.C., & Christensen, W.F. (2012): Methods of Multivariate Analysis. Wiley.
- Mardia, K.V., Kent, J.T., & Bibby, J.M. (1979): Multivariate Analysis. Academic Press.
- Seber, G.A.F. (2004): Multivariate Observations. John Wiley.
- Giri, N.C. (2004): Multivariate Statistical Analysis. 2nd ed. Marcel Dekker.
- Rao, C.R. (1973): Linear Statistical Inference and Its Applications (2nd ed.). New York, Wiley.
- Apostol, T. M. (1974): Mathematical Analysis (2nd ed.). Reading, MA: Addison-Wesley.
- Rudin, W. (1976): Principles of Mathematical Analysis (3rd ed.). McGraw-Hill.
- Khuri, A. I. (1998): Advanced Calculus with Applications in Statistics. Wiley.
- Bartle, R. G., & Sherbert, D. R. (2011): Introduction to Real Analysis (4th ed.). Wiley.
- Ghorpade, S. R., & Limaye, B. V. (2010): A Course in Multivariable Calculus and Analysis. Springer.

STAT-H-CC17-7-Th **3 Credits**

(Probability IV) **THEORY**

Semi-fields, Fields and sigma-fields, Borel sigma-field on R.

limsup and liminf of sequences of sets.

Probability measure and properties, Continuity of probability, Statement and idea of Caratheodory Extension Theorem for probability measure. (9)

Independence of events, Borel-Cantelli lemmas.

Probability space, Random variables and generated sigma-fields, distribution functions and types of distributions. (8)

Product σ-field, Product probability space (finite case only), Joint distributions and marginal distributions, Independence of random variables.

General definition of expectation and properties, Expectation of product of independent random variables. Jensen's Inequality, Statement and application of Fatou's Lemma, Statement and application of Monotone Convergence Theorem and Dominated Convergence Theorem. (12)

Almost sure convergence, relationships among modes of convergence.

Strong laws of large numbers – Statements and illustrative examples. (6)

Introduction to Stochastic Processes, classifications of stochastic processes, counting processes.

Discrete-time Markov chains with countable state space, time-homogeneity, one-step & multi-step transition probabilities.

Classification of states: Communicating classes, Transient states, Recurrent states, Positive and null recurrence (briefly), absorbing states, irreducibility, decomposition of state space into irreducible classes, Periodicity.

Stationary distribution. Statement of Ergodic Theorem for Markov Chains. (10)

STAT-H-CC17-7-TU **1 Credit**

(Probability IV) **TUTORIAL**

Reference Books

- Billingsley, P. (1995): Probability and Measure (3rd ed.). Wiley.
- Durrett, R. (2019): Probability: Theory and Examples (5th ed.). Cambridge University Press.
- Athreya, K. B., & Lahiri, S. N. (2006): Measure Theory and Probability Theory. Springer.
- Feller, W. (1968): An Introduction to Probability Theory and Its Applications, Vol. 1 (3rd ed.). Wiley.
- Ross, S. M. (2014): Introduction to Probability Models (11th ed.). Academic Press.
- Athreya, S., & Sunder, V. S. (2009): Measure and Probability: Theory and Applications. Routledge.
- Goswami, A., & Rao, B. V. (2025): Measure Theory for Analysis and Probability. Springer.
- Bose, A. (2025): Elements of Measure and Probability. Springer.
- Ash, R., & Doleans-Dade, C. (2000): Probability and Measure Theory. Academic Press.
- Basu, A. K. (2010): Measure Theory and Probability. PHI Learning.
- Grimmett, G., & Stirzaker, D. (2001): Probability and Random Processes (3rd ed.). Oxford University Press.
- Medhi, J. (2010): Stochastic Processes (4th ed.). New Age International.
- Goon, A. M., Gupta, M. K., & Dasgupta, B. (2012): Outline of Statistical Theory, Vol. 1. World Press.
- Goon, A. M., Gupta, M. K., & Dasgupta, B. (2012): Outline of Statistical Theory, Vol. 2. World Press.
- Bhat, B. R. (2014): Modern Probability Theory: An Introductory Textbook (2nd ed.). New Age International.

STAT-H-CC18-7-Th **3 Credits**

(Statistical Inference IV) **THEORY**

Point Estimation:

Review of sufficiency. Minimal sufficient statistics and the concept of complete families of distributions. Properties and applications of one-parameter and multi-parameter Exponential families. (5)

Minimum Variance Unbiased Estimators (MVUE). Lehmann-Scheffe Theorem. (4)

Basu's Theorem and its applications: Independence of ancillary and sufficient statistics. (3)

Bhattacharya system of lower bounds and the Method of Covariances. (3)

Testing of Hypotheses:

Fundamental Neyman-Pearson lemma including existence, sufficiency, and necessity parts. (5)

Construction of Uniformly Most Powerful (UMP) tests, with focus on distributions with varying sup- ports. (5)

Monotone Likelihood Ratio property and its role in UMP test construction. (4)

Construction of UMP Unbiased (UMPU) tests for simple and composite hypotheses, and Locally Most Powerful (LMP) tests. (8)

Confidence Sets:

Theory of Confidence sets. Construction of Uniformly Most Accurate Unbiased (UMAU) confidence intervals. (5)

Concept and construction of Simultaneous confidence intervals. (3)

STAT-H-CC18-7-P **1 Credit**

(Statistical Inference IV) **PRACTICAL**

List of Suggested Practical

- Finding minimal sufficient statistics for standard distributions and verifying completeness.
- Application of Lehmann-Scheffe Theorem to derive MVUEs.
- Construction of UMP tests for populations with MLR property.
- Construction of UMAU and Simultaneous intervals.
- Construction of LMP tests.

Reference Books:

- Casella, G., & Berger, R. L. (2002): *Statistical inference* (2nd ed.). Duxbury.
- Goon, A. M., Gupta, M. K., & Dasgupta, B. (2005): *An outline of statistical theory* (Vol. 2). World Press.
- Lehmann, E. L., & Casella, G. (1998): *Theory of point estimation* (2nd ed.). Springer. Lehmann, E. L., & Romano, J. P. (2005): *Testing statistical hypotheses* (3rd ed.). Springer.
- Rohatgi, V. K., & Saleh, A. K. Md. E. (2015): *An introduction to probability and statistics* (3rd ed.). Wiley.

STAT-H-CC19-7-Th **3 Credits**

(Linear Models and Regression Analysis II) **THEORY**

Linear Models II (20)

Gauss Markov Model: recap, Alternative approaches for solution of normal equations. (4)

Linear Set and orthogonal splitting, General linear hypothesis and derivation of related sampling distribution. (4)

Projection matrices and properties. Connection to estimability, Least Squares and BLUEs. Applications to ANOVA models. (4)

Confidence intervals and regions. Multiple comparison techniques due to Scheffe and Tukey. (4)

Derivation of F-tests and equivalence to likelihood ratio tests. (4)

Regression Analysis II (25)

Variable selection: Properties of OLS estimators of regression coefficients and error variance based on subsets. Model selection criteria with derivation. (5)

Multicollinearity – detection and remedial measures including Principal Components, Ridge Regression, Dropping variables and Use of external data, Sample phenomenon. (5)

Detection of outliers and influential observations based on residuals and leverage. Derivation of DFBETA, DFFIT, Cook's Distance. (5)

Heteroscedasticity and Autocorrelation – consequences, detection and remedies. (5)

Checking for normality: Q-Q plots, Normal Probability plot, Shapiro-Wilks test. (5)

STAT-H-CC19-7-P **1 Credit**

(Linear Models and Regression Analysis II) **PRACTICAL**

List of Suggested Practical (Using R)

- Estimability and projection
- Simultaneous confidence intervals and confidence ellipsoid
- Problems on multicollinearity: detection and remedies
- Problems on Heteroscedasticity: detection and remedies
- Problems on Autocorrelation: detection and remedies

Reference Books

- C.R. Rao: Linear Statistical Inference and its Applications.
- S.R. Searle: Linear Models.
- G.A.F. Seber: Linear Regression Analysis.
- N.R. Draper & H. Smith: Applied Regression Analysis.
- S. Chatterjee & A.S. Hadi: Regression Analysis by Example.

STAT-H-CC20-7-Th **3 Credits**

(Design of Experiments II and Sample Survey II) **THEORY**

Design of Experiments II (20)

Split Plot Design in RBD and Strip arrangements, Groups of experiments with RBD and LSD. (2)

Factorial Designs: Concept and applications of factorial designs, Design and analysis of 2^n and 3^n full factorial designs, Concept of confounding (partial and complete) and balancing in 2^n and 3^n factorial designs. Analysis of 2^n and 3^n confounded factorial designs. (8)

Balanced Incomplete Block Designs: Definition, application, properties, analysis, constructions (using finite Projective and Euclidean geometry only), Resolvability: Bose's Inequality. (10)

Sample Survey II (25)

Linear and Circular Systematic sampling: Estimation of mean, Estimation of variance through interpenetrating subsampling technique. (4)

Ratio and Regression methods of estimation, Double sampling and its applications in ratio and regression methods of estimation. (6)

Cluster sampling: Estimation of mean and variance and comparison with SRS – use of intraclass correlation coefficient. (3)

Unified Sampling Theory: Concept of inclusion probability, Horvitz-Thompson estimator, Variance of the estimator and its estimation. Probability Proportional to Size (PPS) sampling (WR & WOR) – Hansen-Hurwitz and Desraj estimator. (12)

STAT-H-CC20-7-P **1 Credit**

(Design of Experiments II and Sample Survey II) **PRACTICAL**

List of Suggested Practical

- Analysis of Split Plot and Strip Plot designs.
- Analysis of Groups of experiments in RBD and LSD
- Analysis of 2^2 and 2^3 factorial design.
- Analysis of 3^2 and 3^3 factorial design.
- Layout and analysis of a completely confounded factorial design.
- Layout and analysis of a partially confounded factorial design.
- Analysis of a 2^n design in a single replicate.
- Systematic sampling
- Ratio and regression methods of estimation
- Cluster sampling
- Horvitz-Thompson estimator
- PPS sampling

Reference Books

- Cochran, W.G. and Cox, G.M.: Experimental Design. Asia Publishing House.
- Das, M.N. and Giri, N.C.: Design and Analysis of Experiments. Wiley Eastern Ltd.
- Kempthorne, O.: The Design and Analysis of Experiments. John Wiley.
- Montgomery, D. C.: Design and Analysis of Experiments, John Wiley.
- Wu, C. F. J. And Hamada, M.: Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Dean, A.M. and Voss, D.: Design and Analysis of Experiments. Springer Texts in Statistics.
- Alok Dey: Theory of Block Designs
- Goon, A.M., Gupta, M.K., Das Gupta, B.: An Outline of Statistical Theory, Vol-II, World Press.
- Goon A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol-II, World Press.
- Cochran, W.G.: Sampling Techniques (3rd Ed.), Wiley Eastern.
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C.: Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics.
- Murthy, M.N.: Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
- Des Raj and Chandhok P.: Sample Survey Theory, Narosa Publishing House.

SEMESTER - VIII

STATISTICS MAJOR

STAT-H-CC21-8-Th **3 Credits**

(Research Methodology I: Research Methodology with Survey Methods) **THEORY**

Fundamentals of Research Process & Design: Defining the research problem, conducting a literature review, and formulating hypotheses. Conceptualizing the research blueprint; Exploratory, Descriptive, and Causal research designs. Identifying independent, dependent, and confounding variables; levels of measurement (Nominal, Ordinal, Interval, Ratio). Deductive vs. Inductive reasoning in social and scientific research. Understanding control groups, randomization, and observational studies. (6)

Research Ethics & Integrity: Definition of ethics and morality; Branches: Meta-ethics, Normative, and Applied ethics (Bioethics and Business ethics). Research integrity vs. misconduct; Falsification, Fabrication, and Plagiarism (FFP). Peer review, redundant publication, and the ethics of authorship. (8)

Foundations of Survey Sampling & Total Survey Error (*TSE*): Target population, sampling frame, and the distinction between sampling and non-sampling errors. In-depth study of coverage error, non-response error, and measurement error. Understanding the framework of TSE as a tool for survey quality. Construction, interpretation, and application of Survey Weights in producing population-level estimates. (9)

Modes of Data Collection & Questionnaire Design: Different modes of data collection, like Face-to-face, Telephone, and Self-administered surveys. Aligning mode-specific biases with the TSE framework (with practical examples). Advantages of CAPI (Computer-Assisted Personal Interviewing) over PAPI (Paper-and-Pen Interviewing). Principles of drafting, sequencing, and reliability testing for designing Questionnaire. (10)

Sampling Techniques & Estimation: Simple Random (SRS), Systematic, Stratified, and Multi-stage Cluster sampling. Probability Proportional to Size sampling and its practical relevance. Convenience, Quota, and Purposive techniques. The Design Effect (*Deff*), sample size estimation, and the trade-off between bias and variance. (12)

STAT-H-CC21-8-Viva **1 Credit**

(Research Methodology I: Research Methodology with Survey Methods) **VIVA**

Reference Books

- Biemer, P. P., & Lyberg, L. E. (2003). Introduction to Survey Quality. John Wiley & Sons.
- Groves, R. M., et al. (2009). Survey Methodology. Wiley-Interscience.
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2017). Applied Survey Data Analysis. Chapman and Hall/CRC.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Wiley.
- Fowler Jr, F. J. (2013). Survey Research Methods. Sage Publications.
- Couper, M. P. (2008). Designing Effective Web Surveys. Cambridge University Press.
- Cochran, W. G. (1977). Sampling Techniques. John Wiley & Sons.
- Lohr, S. L. (2021). Sampling: Design and Analysis. Chapman and Hall/CRC.
- Kish, L. (1965). Survey Sampling. Wiley.
- Kalton, G. (2020). Introduction to survey sampling. Sage Publications.
- Sarndal, C. E., Swensson, B., & Wretman, J. (2003). Model assisted survey sampling. Springer Science & Business Media.
- Beatty, P., Robert M. Groves, Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau (2005). Survey Methodology. Hoboken, NJ: Wiley–Interscience. 2004.
- Motheram, A, Chowdhury, SR, Pramanik, S (2024). Statistical validation of a large-scale web survey during the COVID-19 pandemic in India Medical Research Archives, 12(12).
- Bradley, V. C., Kuriwaki, S., Isakov, M., Sejdinovic, D., Meng, X. L., & Flaxman, S. (2021). Unrepresentative big surveys significantly overestimated US vaccine uptake. Nature, 600 (7890), 695-700.
- Pramanik, S., Neerad Deshmukh, Bijay Chouhan, Manjistha Banerji, Reem Ashraf, Dinesh Kumar Tiwari and Sonalde Desai (2022). Coverage and Nonresponse Bias in Telephone Surveys during the COVID-19 Lockdown in India. NCAER Working Paper No. WP 136.

- Deshmukh, Neerad, Sonalde Desai, Santanu Pramanik, and Dinesh Tiwari (2020). Improving Measurement of Women's Work Participation: Sensitivity of Work Participation Rates to Questionnaire Design. NCAER National Data Innovation Centre Measurement Brief No. 2020-1. NCAER, New Delhi.
- Ravi, Shamika (2023). Our national surveys are based on faulty sampling. Express Opinion, The Indian Express, July 7, 2023.
  Available at https://indianexpress.com/article/opinion/columns/shamika-ravi-writes-our-national-surveys-are-based-on-faulty-sampling-8799300/.
- Resnik, D. B. (2018). The Ethics of Science: An Introduction. Routledge.
- Oliver, P. (2010). The Student's Guide to Research Ethics. Open University Press.
- Steneck, N. H. (2007). ORI Introduction to the Responsible Conduct of Research. Government Printing Office.
- Fink, A. (2019). Conducting Research Literature Reviews: From the Internet to Paper. Sage Publications.
- Cronin, B., & Sugimoto, C. R. (2014). Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact. MIT Press.
- Thelwall, M. (2004). Link Analysis: An Information Science Approach. Academic Press.

STAT-H-CC22-8-Th **3 Credits**

(Research Methodology II: Research Methodology with Statistical Learning) **THEORY**

Data-Driven Research: Role of data science and statistical learning in modern research, Types of learning: supervised, unsupervised and semi-supervised learning. Brief idea of reinforcement learning.

Model Assessment and Validation in Research: Training error vs test error, Bias–variance trade-off, Overfitting & underfitting. (6)

Supervised Learning Methods for Classification: Classification problems in empirical research, Mathematical formulation of classification, Class-specific errors and average misclassification probability.

Nearest Neighbour Methods: k-Nearest Neighbour classifier, Choice of k.

Logistic Regression and Regularization Methods: Multinomial logistic regression, lasso regression, Variable selection interpretation, Shrinkage vs. subset selection. (12)

Tree-Based Modelling and Ensemble Methods: Decision trees for classification, tree construction and interpretation, bagging, random forests. (8)

Artificial Neural Network: Single-layer and multi-layer perceptron. Backpropagation and other optimization methods.

Support Vector Machines: Linear support vector machines, Maximal margin classifier, support vector classifiers, relationship to logistic regression. (10)

Unsupervised Learning for Exploratory Research: Introduction to unsupervised classification or clustering. Discussion on different types of distances. k-means and k-medoids algorithm, Hierarchical clustering based on different types of linkages.

Research Documentation: Literature review in data-driven research, Citation and bibliography, Reference management tools, Reproducible and ethical research practices. (9)

STAT-H-CC22-8-Viva **1 Credit**

(Research Methodology II: Research Methodology with Statistical Learning) **VIVA**

Reference Books

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning: with Applications in R (2nd ed.). Springer.
- Vapnik, V. N. (1998). Statistical Learning Theory. Wiley.
- Wasserman, L. (2004). All of Statistics: A Concise Course in Statistical Inference. Springer.
- Tibshirani, R., Wainwright, M., & Hastie, T. (2015). Statistical Learning with Sparsity: The Lasso and Generalizations. CRC Press.
- Efron, B., & Hastie, T. (2016). Computer Age Statistical Inference. Cambridge University Press.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). Classification and Regression Trees. CRC Press.
- Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press.
- Thomas, C. G. (2021). Research Methodology and Scientific Writing (2nd ed.). Springer.
- Booth, W. C., Colomb, G. G., Williams, J. M., Bizup, J., & Fitzgerald, W. T. (2016). The Craft of Research (4th ed.). University of Chicago Press.
- Turabian, K. L. (2018). A Manual for Writers of Research Papers, Theses, and Dissertations: Chicago Style for Students and Researchers (9th ed.). University of Chicago Press.

STAT-H-CC23-8-Th **3 Credits**

(Statistical Inference V) **THEORY**

Decision Theory:

Decision problems viewed as a two-person game; Non-randomized and randomized decision rules. (3)

Definition of Loss and Risk functions; Concept of Admissibility of decision rules. (2)

Complete, minimal complete, essentially complete, minimal essentially complete classes. (Definitions only) (4)

Essential completeness based on sufficient statistics; Statements of conditions for essential complete- ness of the class of non-randomized rules. (3)

Bayes rules, Generalized Bayes rules, Limit of Bayes, Extended Bayes rules, Admissibility of Bayes rules, Minimax rules and methods for finding Minimax rules. (6)

Bayesian Analysis:

Comparison between Classical and Bayesian analysis; Motivation and choice of different priors. (5)

Bayesian point estimation, hypothesis testing, Credible Sets. Detailed derivation of the Normal con- jugate prior (both parameters unknown), idea of predictive distributions and examples. (10)

Asymptotic Inference:

Consistency and Asymptotic Efficiency; Maximum Likelihood Estimators (MLE) and their large sample properties. (7)

Asymptotic null distributions of Likelihood Ratio Tests (LRT), Rao's Score test, and Wald's tests for simple hypotheses. (5)

STAT-H-CC23-8-P **1 Credit**

(Statistical Inference V) **PRACTICAL**

List of Suggested Practical

- *Analysis of Loss and Risk Functions*: Given a discrete parameter space $\Theta = \{\theta_1, \theta_2\}$ and a finite set of non-randomized decision rules, calculate the Risk $R(\theta, d)$ for each rule under a given loss function (e.g., Squared Error or Absolute Error).
- *Graphical Determination of Admissibility*: Plotting the risk points (R_1, R_2) on a graph to identify the *Risk Set*. Use the lower-left boundary to visually identify and justify the set of admissible rules.
- *Finding Minimax Rules in Finite Games*: For a 2×2 or 2×3 loss matrix, find the minimax rule. This includes checking for a saddle point or using the property that for a minimax rule $d^*, R(\theta_1, d^*) = R(\theta_2, d^*)$.
- *Bayes Rule for Discrete Prior*: Given a prior distribution $\pi(\theta)$ and a small data sample, compute the posterior risk for different actions and select the action that minimizes it.
- *Posterior Distribution for Binomial-Beta Model*: Given a small sample (e.g., $n = 10, x = 3$ successes) and a $Beta(\alpha, \beta)$ prior, calculate the parameters of the posterior distribution and find the Bayesian point estimate under Squared Error Loss (Posterior Mean).
- *Normal-Normal Conjugate Analysis*: Assume a Normal population $N(\mu, \sigma^2)$ where σ^2 is known. Given a prior $N(\mu_0, \tau^2)$ and a small sample, calculate the posterior mean and posterior variance.
- *Construction of Bayesian Credible Intervals*: Using the posterior distribution and Standard Normal Tables, calculate the 95% Symmetric (Equal-Tailed) Credible Interval for the parameter μ.
- *Maximum Likelihood Estimation* (MLE) *and Information*: Given a sample from a distribution (e.g., Exponential or Poisson), calculate the MLE $\hat{\theta}$ and the observed Fisher Information $I(\hat{\theta})$. Use this to estimate the asymptotic variance.
- *Wald and Rao's Score Test*: Compute the Wald Statistic and the Score Statistic. Verify that for large n, the three statistics (LRT, Wald, Score) yield similar conclusions.
- *Consistency Demonstration*: Calculate the estimate $\bar{X}_n$ for increasing sub-samples (e.g., $n =$ 5, 10, 15) from a provided dataset to observe the convergence toward the true parameter (Asymptotic Consistency).

Reference Books:

- Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis* (2nd ed.). Springer.
- Ghosh, J. K., Delampady, M., & Samanta, T. (2006). *An introduction to Bayesian analysis: Theory and methods*. Springer.
- Hogg, R. V., McKean, J., & Craig, A. T. (2018). *Introduction to mathematical statistics* (8th ed.). Pearson.
- Lehmann, E. L., & Casella, G. (1998). *Theory of point estimation* (2nd ed.). Springer.
- Rao, C. R. (1973). *Linear statistical inference and its applications* (2nd ed.). Wiley.
- Goon, A. M., Gupta, M. K., & Dasgupta, B.: *An outline of statistical theory* (Vol. 2). World Press.

STAT-H-WR-CC23-8-Th / Viva **3 + 1 Credits**

(Research Internship) **(Only for Honours with Research Course)**

Students pursuing Honours with Research Course shall have to complete one Research Internship carrying 4 credits (3 Th + 1 Viva) during the final semester as per the stipulated modalities of the university.

STAT-H-CC24-8-Th **3 Credits**

(Discrete Data Analysis) **THEORY**

Alternative study designs, Analysis of contingency tables, Alternative measures of association for nominal categorical data: estimators, standard errors and properties.

Conditional and marginal association; Simpson's paradox. (6)

Generalized Linear Models (GLM): Introduction, Components of a GLM, Canonical link function, Maximum Likelihood estimation and triad of tests, Deviance; Types of residuals. (15)

Applications to binary (grouped and ungrouped) and count data.

Models with constant coefficient of variation. (6)

Negative binomial regression for overdispersed counts, Overdispersion and fitting by quasi-likelihood for GLMs, Zero inflated Poisson models, Joint modelling of mean and variance. (10)

Alternative measures of association for ordinal × nominal and nominal × nominal polytomous categorical data, Regression models for ordinal and nominal polytomous data. (4)

Concept of Generalized Linear Mixed Models (GLMM) – Examples, Fitting procedures and BLUP for LMM and extension to GLMM (concepts only). (4)

STAT-H-CC24-8-P **1 Credit**

(Discrete Data Analysis) **PRACTICAL**

List of Suggested Practical (Using R)

- Measures of association for categorical data.
- Fitting, parameter interpretation and diagnostics for binary and count data.
- Polytomous regression.
- GLMM.

Reference Books

- A. Agresti: Analysis of Ordinal Categorical Data.
- A. Agresti: Categorical Data Analysis.
- P. McCullagh & A.J. Nelder: Generalized Linear Models.
- C.E. McCullough & S.R. Searle: Generalized, Linear and Mixed Models, 2nd ed.

STAT-H-WR-CC24-8-Th / Viva **6 + 2 Credits**

(Dissertation) **(Only for Honours with Research Course)**

Students pursuing Honours with Research Course shall have to complete one Dissertation carrying 8 credits (6 Th + 2 Viva) during the final semester as per the stipulated modalities of the university.

STAT-H-CC25-D-8-Th / Viva **3 + 1 Credits**

(Project) **(For Honours Course)**

Students pursuing Honours Course shall have to complete one Project carrying 4 credits (3 Th + 1 Viva) during the final semester as per the stipulated modalities of the university.

UNIVERSITY OF CALCUTTA

Syllabus

for

Three-year B.Sc. Multidisciplinary Courses of Studies

in

STATISTICS

(Under Curriculum & Credit Framework, 2022)

With effect from the Academic Session

2023-24

Structure of Core Courses in Statistics for MDC

Semester	Course / Paper Code	Course Name
I	STAT-MD-CC1-1-Th / P	Descriptive Statistics I & Probability I
II	STAT-MD-CC2-2-Th / P	Descriptive Statistics II & Probability II
III	STAT-MD-CC3-3-Th / P	Statistical Inference I
IV	STAT-MD-CC4-4-Th / P	Design of Experiments I and Sample Survey I
	STAT-MD-CC5-4-Th / P	Descriptive Statistics III & Probability III
V	STAT-MD-CC6-5-P	Project Work
V OR VI	STAT-MD-CC7-5-Th / P OR STAT-MD-CC7-6-Th / P	Applications of Statistics I
VI	STAT-MD-CC8-6-Th / P	Applications of Statistics II

Structure of Minor Courses in Statistics for MDC

Semester	Course / Paper Code	Course Name
III	STAT-MD-MC1-3-Th / P	Descriptive Statistics I & Probability I
IV	STAT-MD-MC2-4-Th / P	Descriptive Statistics II & Probability II
V	STAT-MD-MC3-5-Th / P	Statistical Inference I
	STAT-MD-MC4-5-P	Project Work
VI	STAT-MD-MC5-6-Th / P	Design of Experiments I and Sample Survey I
	STAT-MD-MC6-6-Th / P	Descriptive Statistics III & Probability III

Structure of Skill Enhancement Courses in Statistics for MDC

Semester	Course / Paper Code	Course Name
I	STAT-MD-SEC1-1-Th / TU	An Introduction to R
II	STAT-MD-SEC2-2-Th / TU	
III	STAT-MD-SEC3-3-Th / TU	

Structure of Interdisciplinary Courses in Statistics for MDC

Semester	Course / Paper Code	Course Name
I	STAT-MD-IDC1-1-Th / P	Statistics for Practitioners
II	STAT-MD-IDC2-2-Th / P	
III	STAT-MD-IDC3-3-Th / P	

Question Patterns

SEMESTER - I

STATISTICS MAJOR

STAT-MD-CC1-1-Th **3 Credits**

(Descriptive Statistics I & Probability I) **THEORY**

Statistics: Definition and scope. Concepts of statistical population and sample.

Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous.

Scales of measurement: nominal, ordinal, interval and ratio.

Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays. (10)

Measures of Central Tendency: Mean, Median, Mode.

Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Ginis Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers. (15)

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications. (20)

STAT-MD-CC1-1-P **1 Credit**

(Descriptive Statistics I & Probability I) **PRACTICAL**

List of Suggested Practical

- Diagrammatic representation of data.
- Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations, stem and leaf plot.
- Problems based on measures of central tendency.
- Problems based on measures of dispersion.
- Problems based on combined mean and variance and coefficient of variation.
- Problems based on moments, skewness and kurtosis.
- Problems related to quantiles and measures based on them, construction of box plot.
- Application problems based on Classical Definition of Probability.
- Application problems based on Bayes' Theorem.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Gupta, S.C., Kapoor, V.K.: Fundamentals of Mathematical Statistics, Sultan Chand & Sons

STAT-MD-SEC1-1-Th **3 Credits**

(An Introduction to R) **THEORY**

Introduction to R: Installation, command line environment, overview of capabilities, brief mention of open source philosophy.
R as a calculator: The four basic arithmetic operations. Use of parentheses nesting up to arbitrary level. The power operation. Evaluation of simple expressions. Quotient and remainder operations for integers. Standard functions, e.g., sin, cos, exp, log. (10)

The different types of numbers in R: Division by zero leading to Infor -Inf. NaN. NA. (No need to go into details). Variables. Creating a vector using c(), seq() and colon operator. How functions map over vectors. Functions to summarise a vector: sum, mean, sd, median etc. Extracting a subset from the vector (by index, by property).
R as a graphing calculator: Introduction to plotting. Plot(), lines(), abline(). No details about the graphics parameters except colour and line width. Barplot, Pie chart and Histogram. Box plot. (15)

Matrix operations in R: Creation. Basic operations. Extracting submatrices.
Loading data from a file: read.table() and read.csv(). Mention of head=TRUE and head=FALSE. Dataframes. Mention that these are like matrices, except that different columns may be of different types. (8)

Numerical Integration in R: Trapezoidal and Simpson's 1/3-rd rules.
Numerical solution of equations in R: Method of fixed-point iteration and Newton-Raphson method in one unknown.
Simulation in R: Simulating a coin toss, a die roll and a card shuffle. Finding probabilities of events related to such experiments using simulation. (12)

STAT-MD-SEC1-1-TU **1 Credit**

(An Introduction to R) **TUTORIAL**

Reference Books:

- Gardener, M.: Beginning R - The Statistical Programming Language, Wiley Publications.
- Braun W J, Murdoch D J: A First Course in Statistical Programming with R. Cambridge University Press. New York.
- A simple introduction to R by Arnab Chakraborty
  (freely available at http://www.isical.ac.in/~arnabc/)
- R for beginners by Emmanuel Paradis
  (freely available at https://cran.r-project.org/doc/contrib/Paradisrdebuts_en.pdf)

STATISTICS INTERDISCIPLINARY

STAT-MD-IDC1-1-Th **2 Credits**

(Statistics for Practitioners) **THEORY**

Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (arithmetic mean, median, mode), dispersion (range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (only through diagrams), Exploratory Data Analysis. (8)

Understanding bivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation (no tie), measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). (8)

Statistical Inference (*testing of hypothesis*): Basic idea of binomial and normal populations (graphical idea only, derivation of the properties excluded). Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and p-value for some standard testing problems like test for proportion/proportions, mean based on single (normal) sample, test on comparing means based on two-sample and paired sample data. (7)

Miscellaneous discussion: Applications of one-way and two-way ANOVA with one observation per cell (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimation of population mean and variability for finite population, idea and application of logistic regression for binary response data. (7)

STAT-MD-IDC1-1-P **1 Credit**

(Statistics for Practitioners) **PRACTICAL**

List of Suggested Practical

- Measures of mean, median, mode, range, QD, SD, CV for univariate data case.
- Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearson's correlation coefficient, Spearman's Rank correlation, measures of association of attributes through contingency table.
- Tests for proportion/proportions, tests of means for single sample, two-sample, and paired sample data on normal response using p-value approach.
- Applications of ANOVA and Kruskal-Wallis test.
- Sample size determination, estimation of population mean and variability for finite population.
- Fitting of logistic regression for binary response data.

Reference Books:

- Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
- Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis, PHI.
- Hardle W. and Simar, L.: Applied Multivariate Statistical Analysis.
- Kutner, M.H. et.al.: Applied Linear Statistical Models.
- Belsley D.A. et.al.: Regression Diagnostics.
- Draper N.R. and Smith, H.: Applied Regression Analysis.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - II

STATISTICS MAJOR

STAT-MD-CC2-2-Th **3 Credits**

(Descriptive Statistics II & Probability II) **THEORY**

Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, fitting of polynomial and exponential curves, correlation ratio, correlation index, intraclass correlation.

Rank correlation: Spearman's and Kendall's measures. (15)

Analysis of Categorical Data: Contingency table, independence & association of attributes. (5)

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.) and probability density function (p.d.f.). Expectation and Variance. Standard probability distributions: Discrete Uniform, Binomial, Poisson, and Normal. (25)

STAT-MD-CC2-2-P **1 Credit**

(Descriptive Statistics II & Probability II) **PRACTICAL**

List of Suggested Practical

- Problems based on analysis of bivariate data.
- Problems based on measures of rank correlation.
- Problems based on analysis of categorical data.
- Finding expectation, variance from a given probability distribution.
- Fitting of binomial distributions for n and p = q = 1/2.
- Fitting of binomial distributions for given n and p.
- Fitting of binomial distributions after computing mean and variance.
- Fitting of Poisson distributions for given value of mean.
- Fitting of Poisson distributions after computing mean.
- Application problems based on binomial distribution.
- Application problems based on Poisson distribution.
- Problems based on area property of normal distribution.
- To find the ordinate for a given area for normal distribution.

- Application based problems using normal distribution.
- Fitting of normal distribution when parameters are given.
- Fitting of normal distribution when parameters are not given.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Agresti, A.: Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, 4th Edition, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Hogg, R.V., Tanis, E.A. and Rao J.M.: Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
- Myer, P.L.: Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Gupta, S.C., Kapoor, V.K.: Fundamentals of Mathematical Statistics, Sultan Chand & Sons

STAT-MD-SEC2-2-Th **3 Credits**

(An Introduction to R) **THEORY**

Introduction to R: Installation, command line environment, overview of capabilities, brief mention of open source philosophy.

R as a calculator: The four basic arithmetic operations. Use of parentheses nesting up to arbitrary level. The power operation. Evaluation of simple expressions. Quotient and remainder operations for integers. Standard functions, e.g., sin, cos, exp, log. (10)

The different types of numbers in R: Division by zero leading to Infor -Inf. NaN. NA. (No need to go into details). Variables. Creating a vector using c(), seq() and colon operator. How functions map over vectors. Functions to summarise a vector: sum, mean, sd, median etc. Extracting a subset from the vector (by index, by property).

R as a graphing calculator: Introduction to plotting. Plot(), lines(), abline(). No details about the graphics parameters except colour and line width. Barplot, Pie chart and Histogram. Box plot. (15)

Matrix operations in R: Creation. Basic operations. Extracting submatrices.

Loading data from a file: read.table() and read.csv(). Mention of head=TRUE and head=FALSE. Dataframes. Mention that these are like matrices, except that different columns may be of different types. (8)

Numerical Integration in R: Trapezoidal and Simpson's 1/3-rd rules.

Numerical solution of equations in R: Method of fixed-point iteration and Newton-Raphson method in one unknown.

Simulation in R: Simulating a coin toss, a die roll and a card shuffle. Finding probabilities of events related to such experiments using simulation. (12)

STAT-MD-SEC2-2-TU **1 Credit**

(An Introduction to R) **TUTORIAL**

Reference Books:

- Gardener, M.: Beginning R - The Statistical Programming Language, Wiley Publications.
- Braun W J, Murdoch D J: A First Course in Statistical Programming with R. Cambridge University Press. New York.
- A simple introduction to R by Arnab Chakraborty
  (freely available at http://www.isical.ac.in/~arnabc/)
- R for beginners by Emmanuel Paradis
  (freely available at https://cran.r-project.org/doc/contrib/Paradisrdebuts_en.pdf)

STATISTICS INTERDISCIPLINARY

STAT-MD-IDC2-2-Th **2 Credits**

(Statistics for Practitioners) **THEORY**

Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (arithmetic mean, median, mode), dispersion (range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (only through diagrams), Exploratory Data Analysis. (8)

Understanding bivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation (no tie), measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). (8)

Statistical Inference (*testing of hypothesis*): Basic idea of binomial and normal populations (graphical idea only, derivation of the properties excluded). Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and p-value for some standard testing problems like test for proportion/proportions, mean based on single (normal) sample, test on comparing means based on two-sample and paired sample data. (7)

Miscellaneous discussion: Applications of one-way and two-way ANOVA with one observation per cell (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimation of population mean and variability for finite population, idea and application of logistic regression for binary response data. (7)

STAT-MD-IDC2-2-P **1 Credit**

(Statistics for Practitioners) **PRACTICAL**

List of Suggested Practical

- Measures of mean, median, mode, range, QD, SD, CV for univariate data case.
- Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearson's correlation coefficient, Spearman's Rank correlation, measures of association of attributes through contingency table.
- Tests for proportion/proportions, tests of means for single sample, two-sample, and paired sample data on normal response using p-value approach.
- Applications of ANOVA and Kruskal-Wallis test.
- Sample size determination, estimation of population mean and variability for finite population.
- Fitting of logistic regression for binary response data.

Reference Books:

- Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
- Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis, PHI.
- Hardle W. and Simar, L.: Applied Multivariate Statistical Analysis.
- Kutner, M.H. et.al.: Applied Linear Statistical Models.
- Belsley D.A. et.al.: Regression Diagnostics.
- Draper N.R. and Smith, H.: Applied Regression Analysis.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - III

STATISTICS MAJOR

STAT-MD-CC3-3-Th **3 Credits**

(Statistical Inference I) **THEORY**

Basic concepts of Statistical Inference: population & sample, parameter & statistic, population distribution and sampling distribution. Point estimation, interval estimation and testing of hypothesis. Three useful distributions for statistical Inference: χ^2, t and F (derivations excluded). (5)

Point Estimation: Concepts of estimation, requirements of a good estimator, notions of mean square error, unbiasedness, bias-variance trade off, best linear unbiasedness and minimum variance unbiasedness. Properties of uniformly minimum variance unbiased estimators (UMVUE). Comparison of Estimators, Efficiency. Methods of Estimation: Method of moments, method of maximum likelihood estimation and statements of their small sample properties. Point estimators of the parameters of Binomial, Poisson, and univariate Normal distributions. (15)

Elements of hypothesis testing: Null and alternative hypotheses, simple & composite hypotheses, critical region, type I and type II errors, level of significance, size, power, p-value. Exact tests and confidence intervals: classical and p-value approaches. Tests relating to Binomial and Poisson distributions, Fisher's exact test. Chi-square tests for association, homogeneity and goodness of fit. Tests of hypotheses for the parameters of normal distribution (one sample and two sample problems), paired t-test. Combination of probabilities in tests of significance. (20)

Interval Estimation: Confidence Interval and Confidence Coefficient, Exact confidence interval for mean(s) and variance(s) for one and two sample problems under the Normal set-up. (5)

STAT-MD-CC3-3-P **1 Credit**

(Statistical Inference I) **PRACTICAL**

List of Suggested Practical

- Maximum Likelihood Estimation.
- Estimation by the method of moments.
- Test of significance for single proportion and difference of two proportions.
- Test of significance for single Poisson mean and difference of two Poisson means.
- Chi square tests for association, homogeneity and goodness of fit.
- Test of significance and confidence intervals for single mean and difference of two means.
- Test of significance and confidence intervals for single variance and ratio of two variances.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory, Vol-1. World Press.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Hogg, R.V., Tanis, E.A. and Zimmerman, D. L.: Probability and Statistical Inference. Pearson Education.
- Johnson, R.A. and Bhattacharya, G.K.: Statistics-Principles and Methods, 4th Edn. John Wiley and Sons.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Gupta, S.C. and Kapoor, V.K. (2020): Fundamentals of Mathematical Statistics. Sultan Chand and Sons.
- Ramachandran, K.M. and Tsokos, C.P.: Mathematical Statistics with Applications. Academic Press.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STATISTICS MINOR

STAT-MD-MC1-3-Th **3 Credits**

(Descriptive Statistics I & Probability I) **THEORY**

Statistics: Definition and scope. Concepts of statistical population and sample.

Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous.

Scales of measurement: nominal, ordinal, interval and ratio.

Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays. (10)

Measures of Central Tendency: Mean, Median, Mode.

Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Ginis Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers. (15)

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications. (20)

STAT-MD-MC1-3-P **1 Credit**

(Descriptive Statistics I & Probability I) **PRACTICAL**

List of Suggested Practical

- Diagrammatic representation of data.
- Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations, stem and leaf plot.
- Problems based on measures of central tendency.
- Problems based on measures of dispersion.
- Problems based on combined mean and variance and coefficient of variation.
- Problems based on moments, skewness and kurtosis.
- Problems related to quantiles and measures based on them, construction of box plot.
- Application problems based on Classical Definition of Probability.
- Application problems based on Bayes' Theorem.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Gupta, S.C., Kapoor, V.K.: Fundamentals of Mathematical Statistics, Sultan Chand & Sons

STAT-MD-SEC3-3-Th **3 Credits**

(An Introduction to R) **THEORY**

Introduction to R: Installation, command line environment, overview of capabilities, brief mention of open source philosophy.

R as a calculator: The four basic arithmetic operations. Use of parentheses nesting up to arbitrary level. The power operation. Evaluation of simple expressions. Quotient and remainder operations for integers. Standard functions, e.g., sin, cos, exp, log. (10)

The different types of numbers in R: Division by zero leading to Infor -Inf. NaN. NA. (No need to go into details). Variables. Creating a vector using c(), seq() and colon operator. How functions map over vectors. Functions to summarise a vector: sum, mean, sd, median etc. Extracting a subset from the vector (by index, by property).

R as a graphing calculator: Introduction to plotting. Plot(), lines(), abline(). No details about the graphics parameters except colour and line width. Barplot, Pie chart and Histogram. Box plot. (15)

Matrix operations in R: Creation. Basic operations. Extracting submatrices.

Loading data from a file: read.table() and read.csv(). Mention of head=TRUE and head=FALSE. Dataframes. Mention that these are like matrices, except that different columns may be of different types. (8)

Numerical Integration in R: Trapezoidal and Simpson's 1/3-rd rules.

Numerical solution of equations in R: Method of fixed-point iteration and Newton-Raphson method in one unknown.

Simulation in R: Simulating a coin toss, a die roll and a card shuffle. Finding probabilities of events related to such experiments using simulation. (12)

STAT-MD-SEC3-3-TU **1 Credit**

(An Introduction to R) **TUTORIAL**

Reference Books:

- Gardener, M.: Beginning R - The Statistical Programming Language, Wiley Publications.
- Braun W J, Murdoch D J: A First Course in Statistical Programming with R. Cambridge University Press. New York.
- A simple introduction to R by Arnab Chakraborty
  (freely available at http://www.isical.ac.in/~arnabc/)
- R for beginners by Emmanuel Paradis
  (freely available at https://cran.r-project.org/doc/contrib/Paradisrdebuts_en.pdf)

STATISTICS INTERDISCIPLINARY

STAT-MD-IDC3-3-Th **2 Credits**

(Statistics for Practitioners) **THEORY**

Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (arithmetic mean, median, mode), dispersion (range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (only through diagrams), Exploratory Data Analysis. (8)

Understanding bivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation (no tie), measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). (8)

Statistical Inference (*testing of hypothesis*): Basic idea of binomial and normal populations (graphical idea only, derivation of the properties excluded). Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and p-value for some standard testing problems like test for proportion/proportions, mean based on single (normal) sample, test on comparing means based on two-sample and paired sample data. (7)

Miscellaneous discussion: Applications of one-way and two-way ANOVA with one observation per cell (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimation of population mean and variability for finite population, idea and application of logistic regression for binary response data. (7)

STAT-MD-IDC3-3-P **1 Credit**

(Statistics for Practitioners) **PRACTICAL**

List of Suggested Practical

- Measures of mean, median, mode, range, QD, SD, CV for univariate data case.
- Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearson's correlation coefficient, Spearman's Rank correlation, measures of association of attributes through contingency table.
- Tests for proportion/proportions, tests of means for single sample, two-sample, and paired sample data on normal response using p-value approach.
- Applications of ANOVA and Kruskal-Wallis test.
- Sample size determination, estimation of population mean and variability for finite population.
- Fitting of logistic regression for binary response data.

Reference Books:

- Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
- Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis, PHI.
- Hardle W. and Simar, L.: Applied Multivariate Statistical Analysis.
- Kutner, M.H. et.al.: Applied Linear Statistical Models.
- Belsley D.A. et.al.: Regression Diagnostics.
- Draper N.R. and Smith, H.: Applied Regression Analysis.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

SEMESTER - IV

STATISTICS MAJOR

STAT-MD-CC4-4-Th **3 Credits**

(Design of Experiments I and Sample Survey I) **THEORY**

Design of Experiments I:

Analysis of Variance: Factors, types and effects; Fixed, random and mixed effects models; Analysis of one-way and two-way classified data with equal number of observations in each cell (Fixed Effects Models only). (10)

Experimental designs: Treatments, Experimental units & Blocks, Experimental error, Basic principles of Design of Experiments (Fisher). (2)

Basic designs: Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model, Applications of the techniques of ANOVA to the analysis of the above designs. Missing plot techniques in RBD and LSD. Uniformity trial experiments and comparison of designs. (10)

Sample Survey I:

Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, random numbers, procedure of selecting a sample, estimates of population mean, total and proportion, standard errors of these estimates, estimates of their standard errors. (15)

Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. (4)

Two-stage sampling (with primary units of equal size and equal selection probability at each stage): unbiased estimation of population mean and total. Ideas of snowball sampling, purposive sampling. (4)

STAT-MD-CC4-4-P **1 Credit**

(Design of Experiments I and Sample Survey I) **PRACTICAL**

List of Suggested Practical

- Analysis of Variance of a one-way classified data (fixed effects model).
- Analysis of Variance of a two-way classified data with one observation per cell (fixed effects model).
- Analysis of Variance of a two-way classified data with more than one observation per cell (fixed effects model).
- Analysis of a CRD.
- Analysis of an RBD.
- Analysis of an LSD.
- Analysis of an RBD with one missing observation.
- Analysis of an LSD with one missing observation.
- To select a simple random sample with and without replacement.
- Simple random sampling – estimation of population mean, total and proportion; estimation of related standard error.
- Estimate the sample size for SRSWOR.
- Stratified Sampling –estimation of population mean and total, allocation of sample to strata by proportional and Neyman's methods, Comparison of the efficiencies of the above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Two-stage Sampling – estimation of population mean and total.

Reference Books:

- Renchner, A. C. And Schaalje, G. B.: Linear Models in Statistics (Second edition), John Wiley and Sons.
- Scheffe, H.: The Analysis of Variance, John Wiley.
- Cochran, W.G. and Cox, G.M.: Experimental Design. Asia Publishing House.
- Das, M.N. and Giri, N.C.: Design and Analysis of Experiments. Wiley Eastern Ltd.
- Kempthorne, O.: The Design and Analysis of Experiments. John Wiley.
- Montgomery, D. C.: Design and Analysis of Experiments, John Wiley.

- Wu, C. F. J. And Hamada, M.: Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Dean, A.M. and Voss, D.: Design and Analysis of Experiments. Springer Texts in Statistics.
- Goon, A.M., Gupta, M.K., Das Gupta, B.: An Outline of Statistical Theory, Vol-II, World Press, Calcutta.
- Goon A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol-II, World Press.
- Cochran, W.G.: Sampling Techniques (3^{rd} Ed.), Wiley Eastern.
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C.: Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics.
- Murthy, M.N.: Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
- Des Raj and Chandhok P.: Sample Survey Theory, Narosa Publishing House.

STAT-MD-CC5-4-Th **3 Credits**

(Descriptive Statistics III and Probability III) **THEORY**

Descriptive Statistics III (15)

Multivariate frequency distribution: multivariate data, multiple regression, multiple correlation, partial correlation coefficient, Particular case with three variates.

Probability III (30)

Moments, Quantiles, Skewness and Kurtosis (including discussions of these measures for Binomial, Poisson and Univariate Normal Distributions). Generating functions – probability generating function and moment generating function.

Discrete probability distributions: Uniform, Geometric, negative binomial, hypergeometric. Limiting/approximation cases.

Continuous probability distributions: uniform, exponential, beta, gamma, lognormal distributions along with their properties.

Two dimensional random variables (*discrete and continuous cases*): Joint p.m.f. and p.d.f., marginal and conditional distributions, properties of c.d.f., independence of variables, Theorems on sum and product of expectations of random variables, Conditional Expectation, Correlation and Regression. Trinomial distribution and its properties. Bivariate Normal Distribution and its properties.

STAT-MD-CC5-4-P **1 Credit**

(Descriptive Statistics III and Probability III) **PRACTICAL**

- Problems on trivariate data: regression line, multiple & partial correlation.
- Application problems based on discrete distributions.
- Application Problems based on continuous distributions.

Reference Books:

- Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol-1), World Press
- Johnson, N.I. & Kotz S.: Distributions in Statistics, John Wiley
- Ross S.M.: Introduction to Probability Models, Academic Press
- Mood A.M., Graybill F. & Boes D.C.: An Introduction to the Theory of Statistics (3rd ed), McGraw Hill
- Rao C.R.: Advanced Statistical Methods in Biometric Research, John Wiley
- Hogg R.V. & Craig A.T.: Introduction to Mathematical Statistics
- Rohatgi V.K.: An Introduction to Probability Theory & Mathematical Statistics, John Wiley
- Stuart G & Ord J.K.: Advanced Theory of Statistics (Vol 2), Charles Griffin
- Goon A. M., Gupta M. K. and Dasgupta B.: Fundamentals of Statistics (V-1), World Press.
- Bhattacharya, G.K. & Johnson R. A.: Concepts & Methods of Statistics, John Wiley
- Gupta, S.C., Kapoor, V.K.: Fundamentals of Mathematical Statistics, Sultan Chand & Sons

STATISTICS MINOR

STAT-MD-MC2-4-Th **3 Credits**

(Descriptive Statistics II & Probability II) **THEORY**

Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, fitting of polynomial and exponential curves, correlation ratio, correlation index, intraclass correlation.

Rank correlation: Spearman's and Kendall's measures. (15)

Analysis of Categorical Data: Contingency table, independence & association of attributes. (5)

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.) and probability density function (p.d.f.). Expectation and Variance. Standard probability distributions: Discrete Uniform, Binomial, Poisson, and Normal. (25)

STAT-MD-MC2-4-P **1 Credit**

(Descriptive Statistics II & Probability II) **PRACTICAL**

List of Suggested Practical

- Problems based on analysis of bivariate data.
- Problems based on measures of rank correlation.
- Problems based on analysis of categorical data.
- Finding expectation, variance from a given probability distribution.
- Fitting of binomial distributions for n and p = q = 1/2.
- Fitting of binomial distributions for given n and p.
- Fitting of binomial distributions after computing mean and variance.
- Fitting of Poisson distributions for given value of mean.
- Fitting of Poisson distributions after computing mean.
- Application problems based on binomial distribution.
- Application problems based on Poisson distribution.
- Problems based on area property of normal distribution.
- To find the ordinate for a given area for normal distribution.
- Application based problems using normal distribution.
- Fitting of normal distribution when parameters are given.
- Fitting of normal distribution when parameters are not given.

Reference Books:

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. I, The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory (Vol-1), World Press.
- Miller, Irwin and Miller, Marylees: John E. Freunds Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.: Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Agresti, A.: Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
- Freedman, D., Pisani, R. and Purves, R.: Statistics, 4th Edition, W. W. Norton & Company.
- Chung, K.L.: Elementary Probability Theory with Stochastic Process, Springer / Narosa.
- Feller, W.: An Introduction to Probability Theory & its Applications, John Wiley.
- Parzen, E.: Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.: Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.: Exercises in Probability, Narosa.
- Rahman, N.A.: Practical Exercises in Probability and Statistics, Griffin.
- Ross, S.: A First Course in Probability, Prentice Hall.
- Hogg, R.V., Tanis, E.A. and Rao J.M.: Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
- Myer, P.L.: Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
- Roychowdhury, S., Bhattacharya, D.: Statistics Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.
- Gupta, S.C., Kapoor, V.K.: Fundamentals of Mathematical Statistics, Sultan Chand & Sons

SEMESTER - V

STATISTICS MAJOR

STAT-MD-CC6-5-P **4 Credits**

(Project Work) **PRACTICAL**

Objective: The aim of the course is to initiate students to write and present a statistical report, under the supervision of a faculty, on some area of human interest. The project work will provide hands on training to the students to deal with data emanating from some real-life situation and propel them to dwell on some theory or relate it to some theoretical concepts.

STAT-MD-CC7-5-Th **3 Credits**

(Applications of Statistics I) **THEORY**

Demography:

Introduction: Sources of Population Data – Census data, Registration 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. (6)

Life tables: Descriptions of Complete Life Table 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. (12)

Index Numbers:

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. Uses of Index numbers. (12)

Psychometry:

Scaling procedures: percentile scaling, Z scaling, T scaling, method of equivalent scores, Intelligence tests & IQ, Scaling ordinal data: Likert scaling, Thurstone Scaling. (15)

STAT-MD-CC7-5-P **1 Credit**

(Applications of Statistics I) **PRACTICAL**

- Mortality rates
- Life table
- Fertility rates and population growth
- Index Number
- Problems using different scaling

Reference Books:

- Goon, A.M., Gupta, M.K., Dasgupta, B. (2001): Fundamentals of Statistics (V-2), World Press.
- Spiegelman M. (1980): Introduction to Demography, Harvard University Press.
- Cox P.R. (1976): Demography, Cambridge University press.
- Shryock, H. S., Siegel, J. S. & Associates. (1976): The methods and materials of demography (4th ed.). U.S. Bureau of the Census.
- Preston, S. H., Heuveline, P. & Guillot, M. (2001): Demography: Measuring and modeling population processes. Blackwell Publishers.
- Siegel, J. S. (2002): Applied demography: Applications to business, government, law, and public policy. Academic Press.
- Croxton, F.E. and Cowden, D.J.: Applied General Statistics. Prentice-Hall.
- Allen, R. G. D. (1975): Index numbers in theory and practice. Macmillan.
- Mudgett, B. D. (1951): Index numbers. Wiley.
- Crocker, L. & Algina, J. (2008): Introduction to classical and modern test theory. Cengage Learning.
- McDonald, R. P. (1999): Test theory: A unified treatment. Psychology Press.
- Kline, P. (2000): The handbook of psychological testing (2nd ed.). Routledge.

STATISTICS MINOR

STAT-MD-MC3-5-Th **3 Credits**

(Statistical Inference I) **THEORY**

Basic concepts of Statistical Inference: population & sample, parameter & statistic, population distribution and sampling distribution. Point estimation, interval estimation and testing of hypothesis. Three useful distributions for statistical Inference: χ^2, t and F (derivations excluded). (5)

Point Estimation: Concepts of estimation, requirements of a good estimator, notions of mean square error, unbiasedness, bias-variance trade off, best linear unbiasedness and minimum variance unbiasedness. Properties of uniformly minimum variance unbiased estimators (UMVUE). Comparison of Estimators, Efficiency. Methods of Estimation: Method of moments, method of maximum likelihood estimation and statements of their small sample properties. Point estimators of the parameters of Binomial, Poisson, and univariate Normal distributions. (15)

Elements of hypothesis testing: Null and alternative hypotheses, simple & composite hypotheses, critical region, type I and type II errors, level of significance, size, power, p-value. Exact tests and confidence intervals: classical and p-value approaches. Tests relating to Binomial and Poisson distributions, Fisher's exact test. Chi-square tests for association, homogeneity and goodness of fit. Tests of hypotheses for the parameters of normal distribution (one sample and two sample problems), paired t-test. Combination of probabilities in tests of significance. (20)

Interval Estimation: Confidence Interval and Confidence Coefficient, Exact confidence interval for mean(s) and variance(s) for one and two sample problems under the Normal set-up. (5)

STAT-MD-MC3-5-P **1 Credit**

(Statistical Inference I) **PRACTICAL**

List of Suggested Practical

- Maximum Likelihood Estimation.
- Estimation by the method of moments.
- Test of significance for single proportion and difference of two proportions.
- Test of significance for single Poisson mean and difference of two Poisson means.
- Chi square tests for association, homogeneity and goodness of fit.
- Test of significance and confidence intervals for single mean and difference of two means.
- Test of significance and confidence intervals for single variance and ratio of two variances.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B.: Fundamentals of Statistics, Vol. 1. The World Press, Kolkata.
- Goon, A.M., Gupta, M.K. & Dasgupta, B.: An Outline of Statistical Theory, Vol-1. World Press.
- Rohatgi, V. K. and Saleh, A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edn. (Reprint). John Wiley and Sons.
- Hogg, R.V., Tanis, E.A. and Zimmerman, D. L.: Probability and Statistical Inference. Pearson Education.
- Johnson, R.A. and Bhattacharya, G.K.: Statistics-Principles and Methods, 4th Edn. John Wiley and Sons.
- Mood, A.M., Graybill, F.A. and Boes, D.C.: Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill.
- Hogg, R.V., McKean, J.W. and Craig, A.T.: Introduction to Mathematical Statistics, 8th Edition. Pearson.
- Gupta, S.C. and Kapoor, V.K. (2020): Fundamentals of Mathematical Statistics. Sultan Chand and Sons.
- Ramachandran, K.M. and Tsokos, C.P.: Mathematical Statistics with Applications. Academic Press.
- Roychowdhury, S., Bhattacharya, D.: Probability and Statistical Inference Theory and Practice, U.N. Dhur & Sons. Pvt. Ltd.

STAT-MD-MC4-5-P **4 Credits**

(Project Work) **PRACTICAL**

Objective: The aim of the course is to initiate students to write and present a statistical report, under the supervision of a faculty, on some area of human interest. The project work will provide hands on training to the students to deal with data emanating from some real-life situation and propel them to dwell on some theory or relate it to some theoretical concepts.

SEMESTER - VI

STATISTICS MAJOR

STAT-MD-CC7-6-Th **3 Credits**

(Applications of Statistics I) **THEORY**

Demography:

Introduction: Sources of Population Data – Census data, Registration 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. (6)

Life tables: Descriptions of Complete Life Table 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. (12)

Index Numbers:

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. Uses of Index numbers. (12)

Psychometry:

Scaling procedures: percentile scaling, Z scaling, T scaling, method of equivalent scores, Intelligence tests & IQ, Scaling ordinal data: Likert scaling, Thurstone Scaling. (15)

STAT-MD-CC7-6-P **1 Credit**

(Applications of Statistics I) **PRACTICAL**

- Mortality rates
- Life table
- Fertility rates and population growth
- Index Number
- Problems using different scaling

Reference Books:

- Goon, A.M., Gupta, M.K., Dasgupta, B. (2001): Fundamentals of Statistics (V-2), World Press.
- Spiegelman M. (1980): Introduction to Demography, Harvard University Press.
- Cox P.R. (1976): Demography, Cambridge University press.
- Shryock, H. S., Siegel, J. S. & Associates. (1976): The methods and materials of demography (4th ed.). U.S. Bureau of the Census.
- Preston, S. H., Heuveline, P. & Guillot, M. (2001): Demography: Measuring and modeling population processes. Blackwell Publishers.
- Siegel, J. S. (2002): Applied demography: Applications to business, government, law, and public policy. Academic Press.
- Croxton, F.E. and Cowden, D.J.: Applied General Statistics. Prentice-Hall.
- Allen, R. G. D. (1975): Index numbers in theory and practice. Macmillan.
- Mudgett, B. D. (1951): Index numbers. Wiley.
- Crocker, L. & Algina, J. (2008): Introduction to classical and modern test theory. Cengage Learning.
- McDonald, R. P. (1999): Test theory: A unified treatment. Psychology Press.
- Kline, P. (2000): The handbook of psychological testing (2nd ed.). Routledge.

STAT-MD-CC8-6-Th **3 Credits**

(Applications of Statistics II) **THEORY**

Time Series Analysis:

Introduction: Examples of time series from various fields, Components of a times series, Additive and Multiplicative models. Stationary and Non-stationary time series.

Trend and Seasonal Components: Estimation of trend by moving averages and curve fitting - polynomial, exponential, Detrending. Estimation of seasonal component by ratio to moving average method, ratio to trend method, Deseasonalization. (7)

Stationary time series, autocorrelation function, moving average process of general order, Autoregressive process of order 1. (10)

Statistical Quality Control:

Introduction: Concepts 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 $(\bar{X}, R)$. (6)

Product Control: Producer's Risk, Consumer's Risk, process average, acceptable quality level, Acceptance Sampling Plan, Single and Double sampling plans by attributes, their OC, ASN and ATI, LTPD and AOQL.

Single sampling plan for inspection by variables (one-sided specification, known and Unknown σ case). (8)

Statistical learning:

Overview of statistical learning, unsupervised and supervised learning, Validation, Training and Testing, Dimensionality reduction, Grouping of observations.

Unsupervised learning: cluster analysis, Hierarchical Methods and partitioning (K-means clustering method).

Supervised learning: Idea of classification, Regression based algorithms (linear classifiers, Decision trees, random forest, logistic regression, linear, polynomial, decision tree regression).

Dimension reduction Method: Principal component Analysis. (14)

STAT-MD-CC8-6-P **1 Credit**

(Applications of Statistics II) **PRACTICAL**

- Trend determination by method of moving averages and by fitting polynomial of appropriate degree.
- Computation of seasonal indices by ratio to trend & ratio to moving average method.
- Control charts for attribute and variable.
- Sampling Inspection plan.
- Cluster Analysis.
- Logistic regression.
- Principal component analysis.

Reference Books:

- Goon, A.M., Gupta, M.K., Dasgupta, B. (2001): Fundamentals of Statistics (V-2), World Press.
- Kendall M.G. (1976): Time Series, Charles Griffin.
- Chatfield C. (1980): The Analysis of Time Series –An Introduction, Chapman & Hall.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015): Time series analysis: Forecasting and control (5th ed.). Wiley.
- Hyndman, R. J., & Athanasopoulos, G. (2021): Forecasting: Principles and practice (3rd ed.). OTexts.
- Montgomery, D. C. (2019): Introduction to statistical quality control (8th ed.). Wiley.
- Duncan A.J. (1953): Quality Control and Industrial Statistics, Richard D Irwin.
- Cowden D.J. (1957): Statistical Methods in Quality Control, Prentice Hall.
- Jiawei Han, Micheline Kamber, Jian Pei: Data Mining concepts and technique.
- James, G., Witten, D., Hastie, T., &Tibshirani, R. (2023): An introduction to statistical learning: With applications in Python (2nd ed.). Springer.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009): The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

STATISTICS MINOR

STAT-MD-MC5-6-Th **3 Credits**

(Design of Experiments I and Sample Survey I) **THEORY**

Design of Experiments I:

Analysis of Variance: Factors, types and effects; Fixed, random and mixed effects models; Analysis of one-way and two-way classified data with equal number of observations in each cell (Fixed Effects Models only). (10)

Experimental designs: Treatments, Experimental units & Blocks, Experimental error, Basic principles of Design of Experiments (Fisher). (2)

Basic designs: Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model, Applications of the techniques of ANOVA to the analysis of the above designs. Missing plot techniques in RBD and LSD. Uniformity trial experiments and comparison of designs. (10)

Sample Survey I:

Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, random numbers, procedure of selecting a sample, estimates of population mean, total and proportion, standard errors of these estimates, estimates of their standard errors. (15)

Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. (4)

Two-stage sampling (with primary units of equal size and equal selection probability at each stage): unbiased estimation of population mean and total. Ideas of snowball sampling, purposive sampling. (4)

STAT-MD-MC5-6-P **1 Credit**

(Design of Experiments I and Sample Survey I) **PRACTICAL**

List of Suggested Practical

- Analysis of Variance of a one-way classified data (fixed effects model).
- Analysis of Variance of a two-way classified data with one observation per cell (fixed effects model).
- Analysis of Variance of a two-way classified data with more than one observation per cell (fixed effects model).
- Analysis of a CRD.
- Analysis of an RBD.
- Analysis of an LSD.
- Analysis of an RBD with one missing observation.
- Analysis of an LSD with one missing observation.
- To select a simple random sample with and without replacement.
- Simple random sampling – estimation of population mean, total and proportion; estimation of related standard error.
- Estimate the sample size for SRSWOR.
- Stratified Sampling –estimation of population mean and total, allocation of sample to strata by proportional and Neyman's methods, Comparison of the efficiencies of the above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Two-stage Sampling – estimation of population mean and total.

Reference Books:

- Renchner, A. C. And Schaalje, G. B.: Linear Models in Statistics (Second edition), John Wiley and Sons.
- Scheffe, H.: The Analysis of Variance, John Wiley.
- Cochran, W.G. and Cox, G.M.: Experimental Design. Asia Publishing House.
- Das, M.N. and Giri, N.C.: Design and Analysis of Experiments. Wiley Eastern Ltd.
- Kempthorne, O.: The Design and Analysis of Experiments. John Wiley.
- Montgomery, D. C.: Design and Analysis of Experiments, John Wiley.

- Wu, C. F. J. And Hamada, M.: Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
- Dean, A.M. and Voss, D.: Design and Analysis of Experiments. Springer Texts in Statistics.
- Goon, A.M., Gupta, M.K., Das Gupta, B.: An Outline of Statistical Theory, Vol-II, World Press, Calcutta.
- Goon A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol-II, World Press.
- Cochran, W.G.: Sampling Techniques (3rd Ed.), Wiley Eastern.
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C.: Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics.
- Murthy, M.N.: Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
- Des Raj and Chandhok P.: Sample Survey Theory, Narosa Publishing House.

STAT-MD-MC6-6-Th **3 Credits**

(Descriptive Statistics III and Probability III) **THEORY**

Descriptive Statistics III (15)

Multivariate frequency distribution: multivariate data, multiple regression, multiple correlation, partial correlation coefficient, Particular case with three variates.

Probability III (30)

Moments, Quantiles, Skewness and Kurtosis (including discussions of these measures for Binomial, Poisson and Univariate Normal Distributions). Generating functions – probability generating function and moment generating function.

Discrete probability distributions: Uniform, Geometric, negative binomial, hypergeometric. Limiting/approximation cases.

Continuous probability distributions: uniform, exponential, beta, gamma, lognormal distributions along with their properties.

Two-dimensional random variables (*discrete and continuous cases*): Joint p.m.f. and p.d.f., marginal and conditional distributions, properties of c.d.f., independence of variables, Theorems on sum and product of expectations of random variables, Conditional Expectation, Correlation and Regression. Trinomial distribution and its properties. Bivariate Normal Distribution and its properties.

STAT-MD-MC6-6-P **1 Credit**

(Descriptive Statistics III and Probability III) **PRACTICAL**

- Problems on trivariate data: regression line, multiple & partial correlation.
- Application problems based on discrete distributions.
- Application Problems based on continuous distributions.

Reference Books:

- Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol-1), World Press
- Johnson, N.I. & Kotz S.: Distributions in Statistics, John Wiley
- Ross S.M.: Introduction to Probability Models, Academic Press
- Mood A.M., Graybill F. & Boes D.C.: An Introduction to the Theory of Statistics (3rd ed), McGraw Hill
- Rao C.R.: Advanced Statistical Methods in Biometric Research, John Wiley
- Hogg R.V. & Craig A.T.: Introduction to Mathematical Statistics
- Rohatgi V.K.: An Introduction to Probability Theory & Mathematical Statistics, John Wiley
- Stuart G & Ord J.K.: Advanced Theory of Statistics (Vol 2), Charles Griffin
- Goon A. M., Gupta M. K. and Dasgupta B.: Fundamentals of Statistics (V-1), World Press.
- Bhattacharya, G.K. & Johnson R. A.: Concepts & Methods of Statistics, John Wiley
- Gupta, S.C., Kapoor, V.K.: Fundamentals of Mathematical Statistics, Sultan Chand & Sons

QUESTION PATTERNS

Four-year B.Sc. (Honours and Honours with Research) Course
&
Three-year B.Sc. Multidisciplinary Course
(Under Curriculum & Credit Framework, 2022)

Theoretical

Full Marks: 75 (Applicable for all papers with 3 Credits)

(a) Objective type: 5 questions × 2 marks = 10 marks

(5 questions are to be attempted out of 8 choices carrying 2 marks each)

(b) Short answer type: 4 questions × 5 marks = 20 marks

(4 questions are to be attempted out of 6 choices carrying 5 marks each)

(c) Broad answer type: 3 questions × 15 marks = 45 marks

(3 questions are to be attempted out of 5 choices carrying 15 marks each)

Full Marks: 50 (Applicable for all papers with 2 Credits)

(a) Objective type: 5 questions × 2 marks = 10 marks

(5 questions are to be attempted out of 8 choices carrying 2 marks each)

(b) Short answer type: 4 questions × 5 marks = 20 marks

(4 questions are to be attempted out of 6 choices carrying 5 marks each)

(c) Broad answer type: 2 questions × 10 marks = 20 marks

(2 questions are to be attempted out of 3 choices carrying 10 marks each)

Practical

Full Marks: 25 (Applicable for all papers with 1 Credit)

Duration: *2 Hours*

(a) Problems: 16 marks

(b) Viva-voce: 5 marks

(c) Practical Note Book: 4 marks

Full Marks: 50 (Applicable for all papers with 2 Credits)

Duration: *4 Hours*

(a) Problems: 40 marks

(b) Viva-voce: 5 marks

(c) Practical Note Book: 5 marks

Full Marks: 100 (Applicable for all papers with 4 Credits)

Duration: *5 Hours*

(a) Problems: 80 marks

(b) Viva-voce: 10 marks

(c) Practical Note Book: 10 marks

Guidelines for Tutorial, Practical and Dissertation/Project

Guidelines for Tutorial Evaluation:

The tutorial component is designed to assess students' conceptual understanding, analytical ability, and communication skills through a combination of written assignment and viva voce.

The tutorial evaluation shall have to be done by internal examiners only and shall consist of the following two parts:

- **Assignment Submission** : 10 Marks
- **Viva Voce** : 15 Marks

Guidelines for Practical Examinations:

1. The candidates will be required to answer all the questions in the Practical Papers.

2. Practical Examination for B.Sc. Four-year Statistics Major to be conducted in away centres simultaneously on the basis of single question paper in presence of external examiners

3. Practical Examination for B.Sc. Four-year Statistics Major papers STAT-H-SEC1-1, STAT-H-SEC3-3, STAT-H-CC13-6, STAT-H-CC19-7 and STAT-H-CC24-8 to be conducted in several batches, as may be required, in away centres using Computers only.

4. Practical Examination of each applicable Major paper for B.Sc. Four-year Statistics, except those mentioned in No. 3 above, to be conducted in a single batch using non-programmable scientific Calculators only.

5. Practical Examination for B.Sc. Four-year Statistics Minor and IDC to be conducted in home centres by internal examiners only.

6. Practical Examination for Three-year Statistics Major, Minor and IDC to be conducted in home centres by internal examiners only.

Guidelines for Dissertation/Project Work:

The Dissertation/Project work is designed to provide students with hands-on experience in handling real-life data and developing the ability to prepare and present a structured statistical report under faculty supervision.

1. **Nature and Type of Dissertation/Project**
   a) The project must be individual work only. Group projects are not permitted.
   b) The project must be based on real-life data (either primary or secondary). Purely theoretical projects are not allowed. However, theoretical study through simulation may be carried out.

2. **Marks Distribution (Total: 200 Marks for Honours with Research or 100 Marks for others)**

Project Report (Notebook)	: 150 Marks (6 Credits)	or	75 Marks (3 Credits)
Presentation and Viva Voce	: 50 Marks (2 Credits)	or	25 Marks (1 Credit)

3. **Mode of Evaluation**

   For Four-year Statistics Major, the project evaluation, presentation, and viva voce shall be conducted at home centres in the presence of External Examiner(s). Project Examination for Three-year Statistics Major and Minor to be conducted at home centres by internal examiners only. The evaluation will assess both the written report and the oral presentation of the project.

4. **Dissertation/Project Report Requirements**

   The project report should be typed and well-structured. The length of the report must be preferably between 3000 to 5000 words.

5. **Structure of the Dissertation/Project Report**

   The report must include the following components:
   a) *Index or contents*
   b) *Executive Summary* - A concise abstract summarizing objectives, methodology, and key findings.
   c) *Introduction* - Background of the study, Objectives and scope of the project.
   d) *Methodology* - Description of data (source, type: primary/secondary), Statistical tools and techniques used, Theoretical concepts applied.
   e) *Results and Analysis* - Presentation of findings using tables, graphs, and statistical outputs, Interpretation of results.
   f) *Conclusion* - Key insights and implications of the study.

g) *Further / future scope of studies* - if any.

h) *References* - Proper citation of sources using standard referencing style.

i) *Acknowledgement* - Recognition of guidance and support received.

j) *Appendix* - Source of data, Original dataset and supplementary material.

6. **Software Usage**

Students are required to use appropriate statistical software for data analysis.

7. **Presentation and Viva Voce**

Students must present their work clearly using suitable aids (slides/charts). Presentation file format must be PPT or PDF.