School of Sciences

M. Sc. Statistics

Course Structure (AY 2023 - 24)

Credit Definition

Type	Duration (in Hours)	Credits
Lecture (L)	1	1
Practical (P)	2	1
Tutorial (T)	1	1

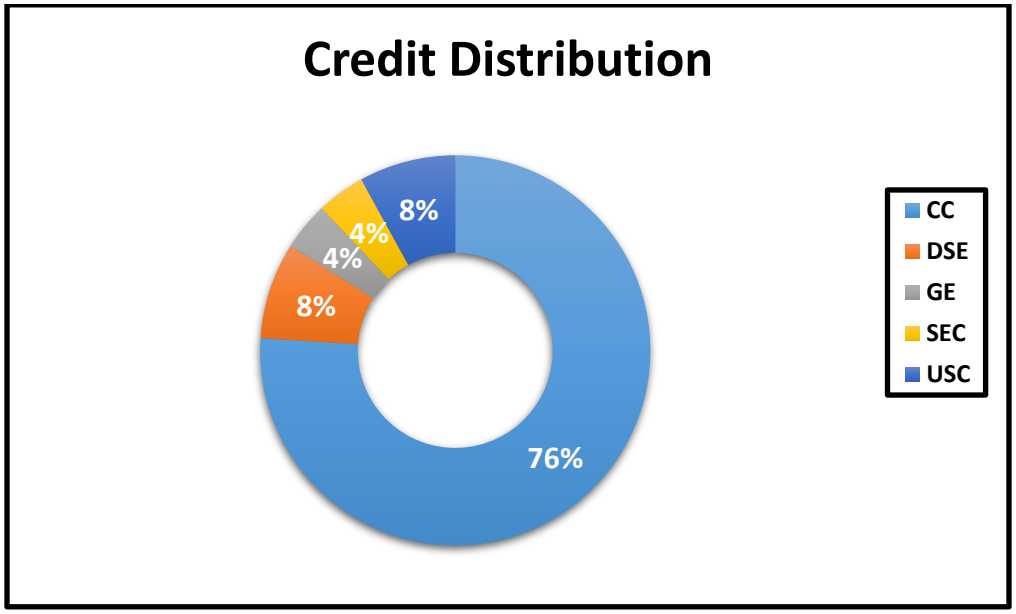
Total Credit

Year	Semester	Hours / Week	Credits
First	1^{st}	31	27
	2^{nd}	35	31
Second	3^{rd}	33	24
	4^{th}	21	18
Total			**100**

Category definition with credit breakup

Definition of Category	Code	No.	Credits
Core Course	CC	1	76
Discipline Specific Elective	DSE	2	8
Generic Elective	GE	3	4
Skill Enhancement Course	SEC	4	4
University Specific Course	USC	5	8
Total			**100**

Category wise Credit Distribution



Subject Codification

Place Value	1	2	3	4	5	6	7	8	9	10	11
Code	Department Name (in short for Core Courses) / Mentored Seminar/ Foreign Language				Program Code	Semester (for semester scheme) / Year (for annual scheme)	Course Category (CC/DSE/GE/ SEC/USC)		Subject type (Theory/Practical/Seasonal/ Project/Internship)	Subject Serial Number	

Values for Place value 1, 2, 3 & 4: Department Name

Sl.	Department	Code
1	STATISTICS	STAT
2	MENTORED SEMINAR	MVMS
3	FOREIGN LANGUAGE	AEFL

Values for Place value 5: Program Code

Department	Program	Program Code
STATISTICS	B. Sc. (Hons.) Statistics	U

STATISTICS	M.Sc. Statistics	P

Values for Place value 6: Semester (for semester scheme)

Value	1	2	3	4
Semester (for semester scheme)	0	1	2	3

Values for Place value 7 & 8: Course Category

Sl.	**Course Category**	**Code**
1	Core Course (CC) / Professional Core Course (PCC)	01
2	Discipline Specific Elective (DSE) / Professional ElectiveCourse (PEC)	02
3	Generic Elective (GE) / Open Electives (OE)	03
4	Skill Enhancement Course (SEC) / Project work, Seminar & Internship or elsewhere (PSE) / Clinical Legal Education (CLE)	05
5	University Specific Course (USC) / Mandatory University Specific Courses (MUS)	06

Values for Place value 9: Subject Type: Theory/Practical/Sessional/Project/Internship

Sl.	**Subject Type**	**Code**
1	Theory	T
2	Practical	P
3	Sessional	S
4	Project/Dissertation	
5	Internship	

Values for Place value 10 & 11: Subject Serial Number

SISTER NIVEDITA UNIVERSITY

SCHOOL OF SCIENCES

DEPARTMENT OF STATISTICS

Program Name: M.SC. STATISTICS

PROGRAMME OUTCOMES (PO)

Our *mission* is to maintain a department that is regarded as equal to any in terms of its relevance of teaching and research, its quality of support and facilities and the learning opportunities and working experience it offers.

PO 01: Statistical Knowledge: Apply the knowledge of statistical and mathematical fundamentals and a statistical specialization to the solution of advanced statistical problems.

PO 02: Problem analysis: Identify, formulate and analyze statistical problems reaching substantiated conclusions using principles of statistics and mathematics.

PO 03: Design/development of solutions: Design solutions for statistical problems and design system or processes that meet the specified needs with appropriate consideration for the application of statistical tools.

PO 04: Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO 05: Modern tool usage: Create, select, and apply appropriate techniques, resources, and tools including prediction and modelling to

PO 06: Problem Solving: The students will be able to critically examine various hypotheses and research queries, and will be able to identify and consult relevant resources to find their rational answers.

PO 07: The Statistician and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant in the field of Statistics.

PO 08: Environment and sustainability: Understand the impact of the statistical knowledge in findings the solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO 09: Ethics: Apply ethical principles and able to identify ethical issues avoid unethical behaviour such as fabrication, falsification or misrepresentation and misinterpretation of data.

PO 10: Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

Course Structure

Semester – I

Sl. No.	Course Name	Code	Credits	Teaching Scheme		
				L	T	P
1	**CC 1** – Analysis I	STATP001T01	4	3	1	0
2	**CC 2** – Measure Theory and Probability	STATP001T02	4	3	1	0
3	**CC 3** – Statistical Inference – I	STATP001T03	4	3	1	0
4	**CC 4** – Linear Models	STATP001T04	4	3	1	0
5	**CC 5** – Statistics Practical – I	STATP001P05	4	0	0	8
6	**DSE 1** – R Programming	STATP002T01	4	3	1	0
7	**SEC 1** – Mentored Seminar – I	MVMSP005S01	1	1	0	0
8	**USC 1** – Foreign Language – I		2	2	0	0
Total Credits (CC: 05, DSE: 01, , USC: 01, SEC: 01)			**27**	**Teaching Hours = 31**		

Semester – II

Sl. No.	Course Name	Code	Credits	Teaching Scheme		
				L	T	P
1	**CC 6** – Sample Survey and Demography	STATP101T01	4	3	1	0
2	**CC 7** – Design of Experiments	STATP101T02	4	3	1	0
3	**CC 8** – Statistical Inference – II	STATP101T03	4	3	1	0
4	**CC 9** – Regression Analysis	STATP101T04	4	3	1	0
5	**CC 10** – Statistics Practical – II	STATP101P05	4	0	0	8
6	**DSE 2** – Operations Research	STATP102T01	4	3	1	0
7	**SEC 2** – Mentored Seminar – II	MVMSP105S02	1	1	0	0
8	**USC 2** – Foreign language – II		2	2	0	0
9	**GE** – Generic Elective		4	4	0	0
Total Credits (CC: 05, DSE: 01, USC: 01, SEC: 01, GE: 01)			**31**	**Teaching Hours = 35**		

Semester – III

Sl. No.	Course Name	Code	Credits	Teaching Scheme		
				L	T	P
1	**CC 11** – Discrete Data Analysis	STATP201T01	4	3	1	0
2	**CC 12** – Time Series Analysis and Stochastic Process	STATP201T02	4	3	1	0
3	**CC 13** – Statistical Computing Using Python	STATP201T03	4	3	1	0
4	**CC 14** – Master Project / Dissertation / Internship	STATP201P04	6	0	0	12
5	**CC 15** – Statistics Practical – III	STATP201P05	3	0	0	6
6	**SEC 3** – Mentored Seminar – III	MVMSP205S03	1	1	0	0
7	**USC 3** – Foreign Language – III		2	2	0	0
Total Credits (CC: 05, USC: 01, SEC: 01)			**24**	**Teaching Hours = 33**		

Course Structure

Semester – IV						
Sl. No.	**Course Name**	**Code**	**Credits**	**Teaching Scheme**		
				L	**T**	**P**
1	**CC 16** – Multivariate Analysis	STATP301T01	4	3	1	0
2	**CC 17** – Statistical Quality Management and Reliability	STATP301T02	4	3	1	0
3	**CC 18** – Biostatistics	STATP301T03	4	3	1	0
4	**CC 19** – Statistics Practical – IV	STATP301P04	3	0	0	6
5	**SEC 4** – Mentored Seminar – IV	MVMSP305S04	2	2	0	0
6	**USC 4** – Foreign Language – IV		1	1	0	0
Total Credits (CC: 04, USC: 01, SEC: 01)			**18**	**Teaching Hours = 21**		

Semester I

Subject Code - STATP001T01	CC1 – Analysis I	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*Undergraduate level of knowledge in real analysis*				
Co-requisites	----				

Learning objectives:

The main objective of the course is to make the students capable to prove elementary facts regarding point set topology, metric spaces, convergence of Fourier series, using the appropriate definition and relevant theorems and solve complex problems.

Course learning outcomes*:*

CO1: **Understand** generalization of the concept of Mean Value Theorem for General Mappings.

CO2: **Understand** the ideas of total variation of a function and **deal with** the functions of bounded variation.

CO3: **Develop** the ability to deal with the problems which are quite significant in the field of real analysis.

CO4: **Apply** the fundamental theorems and principles to solve the complex mathematical problems.

CO5: **Understand** the concept of trigonometric Fourier Series of Functions and the convergence

CO6: **Apply** the concept of Fourier series in Hilbert spaces to solve complex problems.

Course Content:

UNIT I: Differentiation **[10L]**

Derivative Matrix and the Differential; Mean Value Theorem for General Mappings; Chain rule for general mappings; The inverse function and implicit function theorems for general mappings. The related examples.

UNIT II: Metric Spaces **[14L]**

Bounded and totally bounded metric spaces; Compact and separable metric spaces. Functions on metric spaces. Limit, continuity and uniform continuity. Contraction mapping. Banach fixed point theorem. Introductory concepts on Banach space, Hilbert space. Simple properties. Wierstrass approximation theorem.

UNIT III: Functions of Bounded Variation **[10L]**

Functions of bounded variation. Decomposition theorem, Derived function. Derivates. Absolute Continuity.

UNIT IV: Fourier Series **[14L]**

Trigonometric Fourier Series of Functions, convergence at a point. Cesaro summability of Fourier series. Fourier series in Hilbert spaces.

Text &Reference books:

1. H.L. Royden, Real Analysis, 3rd ed., Macmillan Publishing Co., Inc., New York, 1989.
2. E.C. Titchmarsh, Theory of Functions, Clarendon Press, 1932.
3. T.M. Apostal, Mathematical Analysis, Wesley, Reading, 1974.
4. I.P. Natanson, Theory of Functions of a Real Variable, Vols. I & II, Akademie-Verlag, Berlin, 1981.
5. Methods of the theory of Generalized functions – V. S. Vladimirov.

Additional reading:

1. A collection of problems on the equations of mathematical physics - V. S. Vladimirov.

Subject Code - STATP001T02	CC2 – Measure Theory and Probability	C	L	T	P
Version 1.0	**Contact Hours** – **48**	4	3	1	0
Pre-requisites/Exposure	*+2 level Mathematics, UG level Probability Theory.*				
Co-requisites	----				

Learning objectives:

To provide the students with an exposure to an advanced level of Modern Probability and Measure theory. This course is a mandatory for the research enthusiasts in different fields of Statistics as well as Mathematics.

Course learning outcomes*:*

CO1 : **Explain** the introductory concepts of real analysis.

CO2 : **Make use** of the knowledge of different class of sets.

CO3 : **Apply** the concepts of different measures and probability in any area of research and field of experiments.

CO4 : **Build** the idea of measurable functions.

CO5 : **Build** an idea for the requirement of different probability distributions.

CO6 : **Apply** the idea of related convergence theorems in any field of statistical analyses.

Course Content:

UNIT I: Basic Idea of Real Analysis **[10L]**

Set theory, Definition of Function, concepts of image, pre-image, range. Surjection & injection, bijective functions. Left inverse, right inverse, inverse of a function. Equivalent sets, countable and uncountable sets. Real sequences, convergence and divergence of a real sequence. Bounded & unbounded sequence, supremum and infimum, monotonic sequence. Infinite series and properties.

UNIT II: Class of Sets **[12L]**

Class of sets, Union /Intersection over a class of sets, Monotonic sequence of sets and their limits, limit superior, limit inferior. Concept of π-system, λ-system, semi-field, field, sigma-field. Sigma fields generated by class of sets, Borel sigma field on $\mathbb{R}$ and borel sets.

UNIT III: Measurable space and Measures **[12L]**

Definition of measurable space and measure, finite and sigma-finite measures, Counting Measure, Probability Measure and its properties. Continuity theorem and Borel Cantelli lemma (I and II) for probability measure. Caratheodory extension theorem (statement only). Construction of Lebesgue measure over the Borel-sigma field on $\mathbb{R}$. Measurable functions and properties. Concepts of absolutely continuous and singular measures. Concepts of Radon-Nikodyme derivative.

UNIT IV: Random Variable and Probability Distribution **[14L]**

Random variables and probability distributions, distribution function and its properties. Simple random variable and integration: Rieman Integral and Lebesgue Integral. Expectation and moments of a random variable. Monotone Convergence Theorem, Fatou's Lemma, Dominated Convergence Theorem. Sequence of random variables, Convergence in distribution, Convergence in Probability, Convergence in kth mean, Almost sure convergence. Weak Law of large numbers (WLLN), and Strong Law of Large Numbers (SLLN). Central Limit theorem: DeMoivre-Laplace CLT, Lindeberg-Levy and Lindeberg-Feller CLT, Lyapunov's CLT.

Text & Reference books:

1. Sidney I. Resnick: A Probability Path
2. A.K. Basu: Measure theory and Probability
3. B.R. Bhat: Modern Probability Theory
4. P. Billingsley: Probability and Measure
5. J. F. C. Kingman & S. J. Taylor: Introduction to Measure and Probability
6. W. Feller: Introduction to Probability Theory; Vol. I and II
7. R. G. Laha & V. K. Rohatgi: Probability theory

Subject Code - STATP001T03	CC 3 – Statistical Inference - I	C	L	T	P
Version 1.0	**Contact Hours** – **48**	4	3	1	0
Pre-requisites/Exposure	*UG level concepts of Statistical Inference*				
Co-requisites	----				

Learning objectives:

Statistical Inference deals with the theory, methods and practice of forming judgements about the parameters of a population and the reliability of statistical relationships, typically on the basis of random sampling. This course will focus on the inferential problems when the structure of the parent distribution is completely unknown or the exact distribution of the statistic is too complicated or unknown. Main objectives of this course are:

- To develop the concepts regarding the process of using data analysis to infer properties of an underlying distribution of probability.
- Students will learn how to make predictions and draw conclusions about a population based on limited information, which is often the case in real-world problems.

Asymptotic analysis can provide insight into the inference process itself, suggesting what information is available and how this information may be extracted.

Course learning outcomes:

CO1: **Explain** the concepts of sufficiency and completeness.

CO2: **Make use of** the knowledge of unbiasedness and minimum variance in any explanatory data analysis.

CO3**: Apply** the concepts of exponential families, Cramer-Rao inequality, Rao-Blackwell theorem and Lehmann-Scheffe theorems.

CO4: **Build** an idea for the requirement of different optimum tests: MP, UMP, UMPU.

CO5: **Apply** the knowledge of Fundamental Neyman-Pearson lemma, Monotone Likelihood Ratio and Generalized Neyman-Pearson lemma for finding optimum tests.

CO6: **Make use of** the knowledge of maximum likelihood estimator, different large sample tests and their asymptotic behaviours.

Course Content:

UNIT I: Point Estimation **[16L]**

Sufficiency and minimal sufficiency, completeness, complete – sufficient statistic, ancilliary statistic and Basu's theorem. One-parameter and K-parameter exponential families, Unbiasedness and minimum variance, Cramer-Rao inequality and MVB estimators, Rao – Blackwell theorem, Lehman – Scheffe theorem.

UNIT II: Testing Of Hypothesis **[18L]**

Randomized and Non randomized tests, level of significance, size, p-value, power function. Definition of some optimum tests (MP/UMP/UMPU). Construction of MP test – Fundamental Neyman-Pearson Lemma. Monotone Likelihood ratio. Composite null hypothesis, concepts of similar tests, Neyman-structure. Generalized Neyman-Pearson Lemma. UMPU Tests for simple and composite hypotheses.

UNIT III: Asymptotic Inference **[14L]**

Consistency and asymptotic efficiency of estimators, definition of CAN/BAN estimators. Maximum Likelihood estimators and its asymptotic properties. Rao's Score Test, Wald's test, Likelihood ratio tests and their asymptotic properties.

Text & Reference books:

1. Rohatgi V. K. and Saleh A.K. Md. E.: An Introduction to Probability and Statistics. 2nd Edition. (Reprint) John Wiley and Sons, INC.
2. Gun A.M., Gupta M.K. and Dasgupta B.: An Outline of Statistical Theory, Vol. II, The World Press Private Limited, Kolkata.
3. Casella G., Berger RL.: Statistical Inference, Second Edition, Cengage Learning India Private Limited.
4. Kale B.K., Muralidharan K.: Parametric Inference An Introduction, Narosa Publishing House.
5. Mood A.M., Graybill F.A., Boes D.C.: Introduction To The Theory Of Statistics, 3rd Edition, Tata McGraw-Hill Publishing Company Limited.
6. Kendall M.G., Stuart A.: The Advanced Theory of Statistics, Vol. II (Inference and Relationship), CHARLES GRIFFIN. COMPANY LIMITED.
7. Rao C.R.: Linear Statistical Inference and its Applications, 2nd ed., JOHN WILEY & SONS, INC.
8. Lehmann E.L., Casella G.: Theory of Point Estimation, 2nd ed., Springer.
9. Lehmann E.L., Romano J.P.: Testing Statistical Hypotheses, 3rd ed., Springer.

Subject Code - STATP001T04	CC 4 – Linear Models	C	L	T	P
Version 1.0	**Contact Hours** – **48**	4	3	1	0
Pre-requisites/Exposure	*UG level concepts of Statistical Inference*				
Co-requisites	----				

Learning objectives:

To introduce the students with the theory of linear estimation, Gauss-Markov set up, Simple and Multiple Regression Models, ANOVA and ANCOVA Models and finally use some tools to test the model validity.

Course learning outcomes:

CO1: **Demonstrate** the concepts of vector subspaces, orthogonal projection, row-space, column-space, g-inverse etc.
CO2: **Make use of** the knowledge of Gauss-Markov theorem in finding BLUEs.
CO3: **Explain** the fundamental concepts of testing general linear hypothesis, orthogonal splitting and valid error.
CO4: **Apply** the knowledge of different tests in simple and multiple regression set-up.
CO5: **Construct** analysis of variance models to test the significance of several means.
CO6: **Construct** analysis of covariance models to test the significance of several means.

Course Content:

UNIT I: Basic Ideas of Linear Algebra **[16L]**

A brief review of linear algebra: vector and vector spaces, span, spanning set, linear dependence independence and basis. Orthogonal vectors and ortho-complement spaces, orthogonal projection of vectors. Matrices, row space, column space and null space of a matrix, characteristics roots and vectors of square matrices, quadratic forms and their canonical reduction, g-inverse and its properties.

UNIT II: Linear Statistical Inference **[16L]**

Gauss-Markov model: Estimation space and error space, estimable functions, error functions, BLUE and related results, Least Square estimation, Gauss-Markov Theorem. Sum of squares due to a test of linear functions. General Linear Hypothesis problem, description of F test for a general linear hypothesis (without proof). Linear models for correlated errors.

UNIT III: Different Linear Models **[16L]**

ANOVA: fixed, random and mixed effects model, simple and multiple linear regression models, ANCOVA, Multiple comparison, S-method and T-method of multiple comparison.

Text & Reference books:

1. Gun A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. II, 8th Edition, The World Press Private Limited, Kolkata.
2. Scheffe H. (1959): The Analysis of Variance, John Wiley.
3. Rao C.R. (2009): Linear Statistical Inference and its Applications, 2nd Edition, John Wiley & Son, Inc.
4. Faraway J.J. (2014): Linear Models with R, 2nd Edition, CRC Press.
5. Gupta S.C. and Kapoor V.K. (1975): Fundamentals of Applied Statistics: A Modern Approach, S. Chand & Company.

Subject Code - STATP001P05	CC 5 – Statistics Practical - I	C	L	T	P
Version 1.0	**Contact Hours – 96**	4	0	0	8
Pre-requisites/Exposure	*Measure Theory and Probability, Statistical Inference – I, Linear Models and R-programming theory courses*				
Co-requisites	*Use of Scientific Calculators and Statistical Tables/R-Software*				

Learning objectives:

To provide students the hands-on experience of different statistical methods relating the theory courses of M.Sc. 1st semester Statistics.

Course learning outcomes*:*

CO1: **Build** the idea about basic real analysis and different Class of sets by solving problems.

CO2: Apply the idea of Measurable sets and Measurable functions to solve problems.

CO3: Make use of the idea of R programming to solve statistical problems.

CO4: **Solve** numerical problems related to point estimation, hypothesis testing and asymptotic inference.

CO5: **Apply** the knowledge of Gauss-Markov theorem and the theory of linear estimation for solving real-life problems.

CO6: **Apply** the knowledge of ANOVA, ANOCOVA and different Regression models to model various real-life datasets and carryout suitable statistical analyses.

Course Content:

Practical problems related to Measure Theory and Probability, Statistical Inference – I, Linear Models and R programming courses. **(To be executed using Scientific Calculators, MS Excel and R-Software)**

Text & Reference books:

All the references of Measure Theory and Probability, Statistical Inference – I, Linear Models and R programming courses.

Subject Code - STATP002T01	DSE 1 – R Programming	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	Knowledge of Basic Statistics				
Co-requisites					

Learning objectives:

- To help students in building knowledge of R environment.
- To help students work with R packages and their uses for given data. R programming with some basic notations for developing their own simple programs and visualizing graphics in R

Course learning outcomes*:*

CO1: **Build** the idea of basic R language, its origin and its application in real world.

CO2: **Illustrate** how to load data and representation of data by using R.

CO3: **Explain** detailed descriptive statistics, correlation, and lines of regression by R.

CO4: **Build** knowledge of random number generation, fitting of curves and applications of real life problems by using R.

CO5: **Develop** the concepts of regression analysis with their applications and applications of hypothesis testing and p-value by using R.

CO6: **Make use of** the applications of Numerical analysis by using R.

Course Content:

UNIT I: Basics of R and Descriptive Statistics in R [12L]

Basic knowledge about R programming, learn how to load data, plot a graph viz. histograms (equal class intervals and unequal class intervals), box plot, stem-leaf, frequency polygon, pie chart, ogives with graphical summaries of data.

UNIT II: Correlation and Regression in R [12L]

Generate automated reports giving detailed descriptive statistics, correlation and lines of regression and Logistic Regression.

UNIT III: Random number generation and fitting of distributions and Testing of hypothesis in R [14L]

Random number generation and sampling procedures. Fitting of polynomials and exponential curves. Application Problems based on fitting of suitable distribution, Normal probability plot. Simple analysis, create, manage

statistical analysis, projects, import data, Basics of statistical inference in order to understand hypothesis testing and compute p-values and confidence intervals.

UNIT IV: Numerical Analysis Using R **[10L]**

Finding of error, Interpolation, Numerical Solution of algebraic and Transcendental Equations, Numerical Differentiation, Numerical Integration.

Text & Reference books:

1. Gardener M. (2012): Beginning R: The Statistical Programming Language, Wiley Publications.
2. Braun W. J., Murdoch D. J. (2007): A First Course in Statistical Programming with R. Cambridge University Press. New York
3. A simple introduction to R by Arnab Chakraborty (freely available at http://www.isical.ac.in/~arnabc/)
4. R for beginners by Emmanuel Paradis (freely available at https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf)

Semester II

Subject Code – STATP101T01	CC6 – Sample Survey and Demography	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*UG level Sample Survey and Probability*				
Co-requisites	----				

Learning objectives:

- To provide students with an exposure to the applications different survey sampling procedures in real-life situations.
- To introduce the students with the different methodologies of handling data on Demography & Vital Statistics.

Course learning outcomes:

CO1: **Build** knowledge about Simple Random Sampling (SRS), Unequal Probability Sampling (UPS) and their applications.

CO2: **Make use of** the knowledge about Stratified Random Sampling (STRS) and its applications.

CO3: **Make use of** the knowledge of Systematic Sampling and two-stage sampling in life scenarios.

CO4: **Build** knowledge about Ratio & Regression methods of estimation in real-life scenarios.

CO5**: Demonstrate** the concepts of vital statistics and measurements of mortality, life tables.

CO6: **Develop** knowledge about the fertility rates, migration, population growth and population projection.

Course Content:

UNIT I: Simple Random Sampling, Systematic sampling and Unequal Probability Sampling [12L]

Probability sampling from a finite population--–notions of sampling design, sampling scheme, inclusion probabilities; some problems of sampling design construction based on inclusion probabilities. Basic sampling schemes—Simple random sampling with and without replacement. Unequal probability sampling with and without replacement, Systematic sampling. Related estimators of population total/mean, their variances and variance estimators– Mean per distinct UNIT in simple random with replacement sampling, Des Raj and Murthy's estimator (for sample of size two) in unequal probability sampling without replacement.

UNIT II: Stratified sampling, Two-stage sampling, Ratio and Regression Methods [12L]

Allocation problem and construction of strata (optimal, proportional and equal allocation). Ratio, Product, Difference and Regression estimators, Unbiased Ratio estimators – Probability proportional to aggregate size, Sampling and sub-sampling of clusters, Two-stage sampling with equal/unequal number of second stage UNITs and simple random sampling without replacement / unequal probability sampling with replacement at first stage, Ratio estimation in two stage sampling. Double sampling ratio and regression estimators.

UNIT III: Mortality Rates and Life Tables [10L]

Sources of demographic data: census and registration, Measurements of mortality and morbidity, graduation of mortality rates, Life table: Complete and Abridged life tables and their methods of construction.

UNIT IV: Fertility Rates, Growth and Reproduction Rates, Migration, Population estimation and projection [14L]

Measurement of fertility and reproduction and their uses, Stable and stationary population, Internal and international migration, inter-censal and post-censal estimates, population growth, population estimation and projection, Logistic curve and related methods of testing.

Text & Reference books:

1. Gun A.M., Gupta M.K. and Dasgupta B. (2005): Fundamentals of Statistics. Vol. II, 8th Edition. World Press, Kolkata.
2. Mukhopadhyay P. (1999): Applied Statistics, Books and Allied (P) Ltd.
3. Murthy M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta.
4. Cochran W.G. (1984): Sampling Techniques, 3rd Edition, Wiley Eastern.
5. Mukhopadhyay P. (1998): Theory and Methods of Survey Sampling, Prentice Hall.
6. Sukhatme P.V., Sukhatme B.V., Sukhatme S. and Asok C. (1984). Sampling Theory of Surveys With Applications, 3rd Edition (Revised), IOWA State University Press and Indian Society of Agricultural Statistics.
7. Des Raj and Chandhok P. (1998): Sample Survey Theory, Narosa Publishing House.
8. Hansen M.H., Hurwitz W.N., Madow W.G.: Sample Survey Methods and Theory, VOLUME I METHODS AND APPLICATIONS, New York • John Wiley & Sons, Inc., London • Chapman & Hall, Limited.
9. Chattopadhyay A.K. & . Saha A. K.: Demography: techniques and analysis, 1st ed., Viva Books Private.
10. Cox P.R.: Demography, 5th ed., Cambridge University Press.
11. Shryock H. S., Siegel J.S. and Associates : The Methods and Materials of Demography, Vol I.
12. Keyfitz N., Caswell H.: Applied Mathematical Demography, 3rd ed., Springer.

Subject Code – STATP101T02	CC7 – Design of Experiments	C	L	T	P
Version 1.0	**Contact Hours** – **48**	4	3	1	0
Pre-requisites/Exposure	*PG level Linear Models and UG level Design of Experiments*				
Co-requisites	----				

Learning objectives:

To introduce the students with the concepts of basic and advanced Designs of Experiments, Factorial experiments and Confounding which serves as multipurpose tools used in many situations to optimize the output.

Course learning outcomes:

CO1: **Define** the concept of design of experiments and its various related terminologies.
CO2: **Compare** the relative efficiencies of different basic designs.
CO3: **Apply** the knowledge of Balanced Incomplete Block Designs (BIBD).
CO4: **Develop** the concepts of Mutually Orthogonal Latin Squares (MOLS).
CO5: **Demonstrate** the applications of Row-Column designs and Youden Square design.
CO6: **Apply** the knowledge of Factorial Experiments, Confounding and Balancing in real life scenarios.

Course Content:

UNIT I: Review of Basic Experimental Designs [16L]

Terminologies: Experiment – Absolute and Comparative, Design of Experiments, Treatments, Experimental units & Blocks, Experimental error, Precision & Efficiency. Basic principles of Design of Experiments (Fisher). Influence of choice of size and shape of plots and blocks in Agricultural experiments, Uniformity trials, fertility contour maps. Uses in Industrial Experiments. Basic Designs: Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model and statistical analysis, relative efficiency. Analysis with missing observations in RBD and LSD.

UNIT II: Block Designs, BIBD, MOLS and Row Column Designs [20L]

Block Designs: Review of Linear Model methods and general analysis of block designs for one-way heterogeneity set-up, notions of Connectedness, Orthogonality and Balance.
Balanced Incomplete Block Designs (BIBD): definition, applications, properties, analysis, construction and resolvability.
Definition and construction of Mutually Orthogonal Latin Squares (MOLS).
Row-Column designs: applications and general analysis, introduction to Lattice and Youden Square designs.

UNIT III: Factorial Experiments and Confounding [12L]

Factorial designs, advantages, notations and concepts. 2^n experiments: design and analysis. Total and Partial Confounding for 2^n experiments. Balancing in Symmetric Factorial Experiments. Factorial experiments in a single replicate. Introduction to 3^n experiments.

Text & Reference books:

1. Gun A.M., Gupta M.K. and Dasgupta B. (2005): Fundamentals of Statistics. Vol. II, 8th Edition. World Press, Kolkata.
2. Mukhopadhyay P. (1999): Applied Statistics, Books and Allied (P) Ltd.
3. Kempthorne O. (1965): The Design and Analysis of Experiments. John Wiley.
4. Montgomery D. C. (2008): Design and Analysis of Experiments, John Wiley.
5. Cochran W.G. and Cox G.M. (1959): Experimental Design. Asia Publishing House.
6. Das M.N. and Giri N.C. (1986): Design and Analysis of Experiments. Wiley Eastern Ltd.
7. S.C Gupta, V.K Kapoor (2000): Fundamentals of Applied Statistics, 3rd Edition, Sultan Chand and Sons, New Delhi.
8. Faraway J.J. (2014): Linear Models with R, 2nd Edition, CRC Press.
9. Dey A.: Incomplete Block Designs, Hindustan Book Agency.

Subject Code – STATP101T03	CC8 – Statistical Inference - II	C	L	T	P
Version 1.0	**Contact Hours** – **48**	4	3	1	0
Pre-requisites/Exposure	*UG level Statistical Inference, Univariate and Bivariate Probability Distributions*				
Co-requisites	----				

Learning objectives:

Statistical Inference deals with the theory, methods and practice of forming judgements about the parameters of a population and the reliability of statistical relationships, typically on the basis of random sampling. This course will focus on the inferential problems when the structure of the parent distribution is completely unknown or the exact distribution of the statistic is too complicated or unknown. Main objectives of this course are:

- To develop the concepts regarding the process of using data analysis to infer properties of an underlying distribution of probability.
- Students will learn how to make predictions and draw conclusions about a population based on limited information, which is often the case in real-world problems.

Course learning outcomes:

CO1: **Define** the concept of U-statistics and its various related terminologies.
CO2: **Apply** the knowledge of single sample problems of nonparametric testing in real life situations.
CO3: **Apply** the knowledge of two-sample and multi-sample problems of nonparametric testing in real life situations.
CO4: **Develop** the concepts of Bayesian statistical analysis.
CO5: **Demonstrate** the applications of non-informative priors, Jeffreys prior and conjugate priors.
CO6: **Build the** knowledge of different Bayesian inferential techniques and Bayesian computations.

Course Content:

Nonparametric Inference:

UNIT I: Introduction to U-statistic **[12L]**

Nonparametric estimation, concept of U – statistic (single sample and two-sample case), examples. Exact expression of the variance of U-statistic for single samples. Asymptotic properties of U – statistic.

UNIT II: Nonparametric Testing of Hypotheses **[12L]**

Nonparametric testing, single sample location problem, goodness of fit test. Concepts of stochastically larger/smaller distributions, two-sample problems and multi-sample problems. Concept of Linear Rank Statistics.

Bayesian Inference:

UNIT III: Introduction to Bayesian Analysis [14L]

Comparison of classical statistical analysis and Bayesian analysis, prior and posterior distributions. Motivation of choice of different priors: non-informative priors and their advantages and disadvantages, Jeffreys' prior. Conjugate Prior, its advantages, conjugate prior for different families. Derivation of posterior distribution.

UNIT IV: Bayesian Inference and Computations [10L]

Bayesian point estimation: Bayes estimator, Maximum A Priori (MAP) estimator, Bayes factor and Posterior odds ratio, Bayes Credible intervals and High Density Regions (HDR). Hierarchical Bayes, brief discussions on Bayesian computational techniques and applications.

Text & Reference books:

1. Gun A.M., Gupta M.K. and Dasgupta B. (2005): Fundamentals of Statistics. Vol. II, 8th Edition. World Press, Kolkata.
2. Mukhopadhyay P. (1999): Applied Statistics, Books and Allied (P) Ltd.
3. Kempthorne O. (1965): The Design and Analysis of Experiments. John Wiley.
4. Montgomery D. C. (2008): Design and Analysis of Experiments, John Wiley.
5. Cochran W.G. and Cox G.M. (1959): Experimental Design. Asia Publishing House.
6. Das M.N. and Giri N.C. (1986): Design and Analysis of Experiments. Wiley Eastern Ltd.
7. S.C Gupta, V.K Kapoor (2000): Fundamentals of Applied Statistics, 3rd Edition, Sultan Chand and Sons, New Delhi.
8. Faraway J.J. (2014): Linear Models with R, 2nd Edition, CRC Press.
9. Dey A.: Incomplete Block Designs, Hindustan Book Agency.

Subject Code – STATP101T04	CC9 – Regression Analysis	C	L	T	P
Version 1.0	**Contact Hours** – **48**	4	3	1	0
Pre-requisites/Exposure	*PG level Linear Models and Basic concepts of regression*				
Co-requisites	----				

Learning objectives:

Regression analysis is a powerful tool that allows us to examine the relationship between two or more variables of interest and examine the influence of one or more independent variables on a dependent variable.

Main objectives of the course are:

- Learn how to determine the strength of predictors and the predictor with most significant impact.
- Learn how to choose the best model from different prediction models.
- To carryout suitable data analysis using different statistical model validation techniques, diagnostics and tests.

Course learning outcomes:

CO1: **Define** the concept of different transformations of regression model.
CO2: **Develop** the knowledge of model validation and model selection in real life data analysis.
CO3: **Apply** the methods of finding outliers and influential observations in real life data analysis problems.
CO4: **Test** for verifying whether heteroscedasticity and autocorrelation are present in any real life data and give suitable remedies.
CO5: **Test** for normality/non-normality for any real life dataset and suggest further statistical analyses.
CO6: **Demonstrate** the concepts of multicollinearity and the applications of Stepwise regression.

Course Content:

UNIT I: Regression Model **[24L]**

Building regression models: Box-Cox and Box-Tidwell transformations. Model validation: R^2 and Adjusted R^2, Mallows' Cp, AIC and BIC, test of fitness of a model, residuals and their plots, detection of outliers and influential observations: residuals and leverages, DFBETA, DFFIT, Cook's Distance.

UNIT II: Departure from Gauss-Markov setup **[24L]**

Heteroscedasticity, multicollinearity, autocorrelation; non normality: consequences, detection and remedies. An introduction of regression analysis with correlated errors. Stepwise regression.

Text & Reference books:

1. A. M. Gun, M. K. Gupta, B. Dasgupta: Fundamentals of Statistics; Vol-II
2. A. M. Gun, M. K. Gupta, B. Dasgupta: An Outline of Statistical Theory; Vol-II
3. R. D. Cook & S. Weisberg: Residual and its Influence in Regression
4. N. R. Draper & H. Smith: Applied Regression Analysis

5. J. Johnston: Econometric Methods
6. H. D. Vinod & A. Ullah: Recent Advances in Regression Methods
7. D. A. Belsley, Kuh & Welsch: Regression Diagnostics data & sources of collinearity
8. Gujarathi: Basic Econometrics
9. G. G. Judge, R. C. Hill, W. E. Griffith, H. Lutkepohl & T. C. Lee : The Theory and Practice of Econometrics.

Subject Code – STATP101P05	CC 10 – Statistics Practical - II	C	L	T	P
Version 1.0	**Contact Hours** – **96**	4	0	0	8
Pre-requisites/Exposure	*Sample Survey and Demography, Design of Experiments, Statistical Inference – II, Regression Analysis and Operations Research theory courses.*				
Co-requisites	*Use of Scientific Calculators, Statistical Tables and R-Software*				

Learning objectives:

To provide students the hands-on experience of different statistical methods relating the theory courses of M.Sc. 2nd semester Statistics.

Course learning outcomes*:*

CO1: **Build** the idea about different survey sampling methods.

CO2: Make use of the idea of different vital statistics measures to solve real-life problems.

CO3: Apply the knowledge of different Experimental Designs in various aspects of data analysis.

CO4: **Apply** the knowledge of non-parametric methods and Bayesian computations to solve various real-life problems

CO5: **Solve** numerical problems related to Regression Analysis.

CO6: **Apply** the knowledge of Linear Programing Problem, Queuing theory, Inventory and Game theory in various real-life aspects.

Course Content:

Practical problems related to Sample Survey and Demography, Design of Experiments, Statistical Inference – II, Regression Analysis and Operations Research courses **(To be executed using Scientific Calculators, MS Excel and R-Software)**

Text & Reference books:

All the references of Sample Survey and Demography, Design of Experiments, Statistical Inference – II, Regression Analysis and Operations Research courses.

Subject Code – STATP102T01	DSE 2 – Operations Research	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	Basic Mathematics				
Co-requisites	----				

Learning objectives:

To introduce the students with the basic concepts of Linear Programming Problem and Operations Research.

Course learning outcomes*:*

***CO1*: Define** the underlying concept of Operations Research.
***CO2:* Classify** the problems of Linear Programming.
***CO3:* Apply** the concept of Game theory in practical life.
***CO4:* Analyze** the various inventory models.
***CO5:* Explain** the concepts of Queuing theory in our daily life.
***CO6:* Interpret** the various results associated with Operations Research.

Course Content:

UNIT I: Introduction **[3L]**

Definition and scope of Operations Research, models and their solutions, decision making under certainty, uncertainty, risk and competition.

UNIT II: Linear Programming Problem **[17L]**

Linear Programming Problem, Duality, Transportation problem, Assignment Problem, Travelling salesman Problem.

UNIT III: Game Theory **[8L]**

Games in normal form, pure and mixed strategies, solution of 2×2, $m \times 2$, $2 \times n$ and $m \times n$ zero sum games by dominance principles and graphical method, LP formulation of matrix games, fundamental theorem of matrix game.

UNIT IV: Inventory Control **[10L]**

Analytical structure of inventory problems, EOQ formula of Harris and Wilson and its sensitivity analysis, extension of EOQ formula allowing quantity discounts and shortages, models with random demand, static risk models, ABC analysis.

UNIT V: Queuing Theory **[10L]**

Queuing models-specifications and effectiveness measure, M/M/1and M/M/C queues and their steady state solutions, waiting time distribution for M/M/1 queue.

Text & Reference books:

1. Taha H. (1995): Operations Research: An Introduction, Prentice- Hall India.
2. Phillips D. T., Ravindran A. and Solberg J.: Operations Research: Principle and Practice
3. Gross C. and Harris C. M.: Fundamentals of Queuing Theory
4. Swarup K., Gupta P.K. and Mohan M (2010): Operations Research, Sultan Chand & Sons.

Semester III

Subject Code - STATP201T01	CC11 – Discrete Data Analysis	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	Probability II, Descriptive Statistics II, Statistical Inference I, Linear Models and Regression Analysis				
Co-requisites	----				

Learning objectives:

Categorical data, meaning data that is non-numerical by nature, are often collected in scientific studies. This type of data can be used to show relationship between variables and it is invaluable to collect and compile in studies. Discrete data analysis or categorical data analysis is used to gather information from both online and offline surveys or questionnaires as the case may be. The type of categorical data used may differ depending on the aim of data collection. The main objectives of this course are:

- To develop concepts about how to relate some number of continuous and/or categorical predictors to a single outcome variable.
- To allow for more flexible, non-linear relationships by using a different underlying statistical distribution.
- Categorical data analysis has a huge application in epidemiological studies.

Course learning outcomes*:*

CO1: **Apply** the concepts of different types of categorical data.
CO2: **Make use of** the knowledge of different measures of associations in different aspects of data analysis.
CO3: **Build** the knowledge of the Generalized Linear Models (GLM) to analyze more flexible and non-linear relationships among the variables.
CO4: **Explain** the advantages of log regression and logistic regression.
CO5: **Apply** the concepts of modelling binary data, count data and polytomous data in real-life situations.
CO6: **Make use of** the knowledge of Maximum likelihood and Quasi likelihood methods for the purpose of the data modelling.

Course Content:

UNIT I: Categorical Data Analysis **[16L]**

Design of study: Retrospective and Prospective studies, case control study and cohort study. Theory of attributes, data consistency, contingency tables, independence and association of attributes, measures of association for 2×2 case – Yule's measures, Risk Difference, Risk Ratio, Odds ratio. Manifold two-way $(k \times l)$ classification – Pearson's measures, Tschuprow's measure, Cramer's V^2, Goodman-Kruskal gamma, Somer's d, Kendall's measures.

UNIT II: Generalized Linear Models (GLM) **[20L]**

Introduction, departure from a linear model, components of a GLM, uses of different links: Logit link, Probit link, Log-linear link, Logistic regression, Log regression, Maximum Likelihood Estimation and Deviance.

UNIT III: Different Discrete Data and Models **[12L]**

Binary data and count data: grouped and ungrouped. Models with constant coefficient of variation. Polytomous data. Overdispersion and fitting by quasi-likelihood. Zero-inflated Poisson models.

Text & Reference books:

1. Agresti A.: Analysis of Ordinal Categorical Data, John Wiley & Sons, Inc.
2. Agresti A.: Categorical Data Analysis, John Wiley & Sons, Inc.
3. McCullagh P., Nelder A.J.: Generalized Linear Models, 2nd Edition, Chapman and Hall.
4. McCulloch C.E., Searle S.R., Neuhaus John M.: Generalized, Linear, and Mixed Models, 2nd Edition, John Wiley & Sons, Inc.
5. Hastie T., Tibshirani R.: Generalized Additive Models, Chapman & Hall/CRC.
6. Annette J. Dobson: An Introduction To Generalized Linear Models, 2nd Edition, Chapman & Hall/CRC.
7. Faraway Julian J.: Linear Models with R, 2nd Edition, CRC Press Taylor & Francis Group.

Subject Code - STATP201T02	CC12 – Time Series Analysis and Stochastic Process	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*Concept of probability theory and random variables.*				
Co-requisites	----				

Learning objectives:

- To provide students the fundamental knowledge of different statistical methods regarding Time Series Analysis along with their applications.
- To provide the students the knowledge of Stochastic processes, the uses and applications of transition probabilities and Markov chains.

Course learning outcomes*:*

CO1: **Explain** the concepts related to Stationary Time Series, Moving Average and Auto Regressive processes.
CO2: **Build** knowledge of Smoothing techniques and financial forecasting models.
CO3: **Develop** knowledge of stationary process, counting process, Poisson process, finite and countable state space.
CO4: **Make use of** the knowledge of one step and n-step transition probability and stationary distribution of Markov chain.
CO5: **Make use of** continuous time Markov chain, pure Birth and death process.
CO6: **Apply** the concepts of Renewal theorem and Branching Process.

Course Content:

UNIT I: Time Series Modelling **[14L]**

Introduction: Classical Models. Stationary time series, autocorrelation and partial autocorrelation functions, Correlogram analysis. Stationary Processes: Moving Average and Autoregressive processes, Forecasting techniques: Box-Jenkins model. Tests for Unit Roots.

UNIT II: Smoothing Techniques and Volatility **[10L]**

Smoothing techniques: Exponential and Holt-Winters methods. Volatility: ARCH, GARCH models and their variants.

UNIT III: Introduction to Stochastic Process **[10L]**

Definition of stationary process, Counting process. Inter-arrival times and their distributions. Poisson process and its properties.

UNIT IV: Markov Chains and Branching Process **[14L]**

Markov chain with finite state space and countable state space, Classification of states, Chapman-Kolmogorov equation, Calculation of n-step transition probability matrix and its limit, Stationary distribution of Markov chain. Continuous time Markov chain: Pure birth process, pure death process, Birth and death processes, Renewal

theory, Elementary Renewal theorem, Statement and uses of Key Renewal theorem. Branching process: Probability of ultimate extinction and distribution of population size.

Text & Reference books:

1. C. Chatfield: The Analysis of Time Series – An Introduction.
2. G. E. P. Box & G. M. Jenkins: Time Series Analysis–Forecasting and Control.
3. A. Pankratz: Forecasting with Univariate Box-Jenkins Model.
4. N. H. Chen: Time series, applications to finance.
5. J. L. Doob: Stochastic Processes.
6. S. Karlin & H. M. Taylor: A First Course in Stochastic Process, Vol. I.
7. J. Medhi: Stochastic Processes D. R. Cox: Renewal Theory.
8. S. Ross: Stochastic Processes.

Subject Code - STATP201T03	CC13 – Statistical Computing Using Python	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*Knowledge of basic statistics and mathematics*				
Co-requisites	----				

Learning objectives:

Python is a powerful and open-source programming language used for a wide range of applications, from games to web apps, to scientific computing. The main objectives of learning this course are;

- To provide flexibility and extensive support for machine learning frameworks.
- Allows programmers to build intelligent systems, train sophisticated models and apply these models in real-world applications.
- Very helpful for data analysis and data visualization.

Course learning outcomes*:*

CO1: **Create** program in Python IDLE.
CO2: **Build** basic programs using fundamental programming constructs like variables, conditional logic, looping and functions.
CO3: **Apply** OOPs concepts in the programming.
CO4: **Make use** of Arrays and Data structures.
CO5: **Build** an application with the support of graphics in Python.
CO6: **Apply** error handling.

Course Content:

UNIT I: **[24L]**

Basic Idea; Simple Syntax; Basic Operations; Different Libraries; Function; Loop; Array; Data handling and management; Chart and Diagrams.

UNIT II: **[24L]**

Random Number Generation from a known and unknown distribution; Simulation; Application in various Statistical field; Idea of Parallel Computing and/or Efficient Programming.

Text & Reference books:

1. Mueller J.P., Massaron L.: Python for Data Science, 2nd ed., John Wiley & Sons, Inc.
2. Nelli F.: Python Data Analytics, Apress.

Subject Code - STATP201P05	CC 15 – Statistics Practical - III	C	L	T	P
Version 1.0	**Contact Hours** – **72**	3	0	0	6
Pre-requisites/Exposure	*Discrete Data Analysis, Time Series and Stochastic Process, Statistical Computing Using Python courses.*				
Co-requisites	*Use of Scientific Calculators, Statistical Tables and R-Software*				

Learning objectives:

To provide students the hands-on experience of different statistical methods relating the theory courses of M.Sc. 3rd semester Statistics.

Course learning outcomes*:*

CO1: **Build** the idea for solving real life problems involving various discrete type data: binary data, count data.

CO2: Make use of the idea of different types of generalized linear models to model real life datasets.

CO3: Apply the knowledge of different Time series models in various aspects of data analysis.

CO4: **Apply** the knowledge of different stochastic processes and Markov chains to solve various real-life problems.

CO5: **Solve** numerical problems using Python.

CO6: **Apply** the knowledge of Python to solve various statistical problems.

Course Content:

Practical problems related to Discrete Data Analysis, Time Series and Stochastic Process, Statistical Computing Using Python courses **(To be executed using Scientific Calculators, MS Excel and R-Software).**

Text & Reference books:

All the references of Discrete Data Analysis, Time Series and Stochastic Process, Statistical Computing Using Python courses.

Semester - IV

Subject Code - STATP301T01	CC16 – Multivariate Analysis	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*Linear Algebra, Univariate and Bivariate Probability distributions, UG level knowledge of Multivariate Analysis.*				
Co-requisites	----				

Learning objectives:

To provide the students with an exposure to an advanced level of Multivariate statistical analysis. This course is a mandatory for any real life data analysis required in both research and corporate fields.

Course learning outcomes*:*

CO1: **Explain** the introductory concepts of multivariate probability distributions.

CO2: **Make use** of the knowledge of multiple regression in different aspects of data analysis.

CO3: **Build** an idea for the requirement of multivariate normal distribution, Wishart distribution for different inferential techniques and multivariate data analysis in real life scenario.

CO5: **Develop** knowledge about the basic concepts and associated results of principal component analysis.

CO4: **Demonstrate** the basic ideas of factor analysis as a dimension reduction technique.

CO5: **Develop** in-depth knowledge about classification and discriminant analysis tools and cluster analysis.

Course Content:

UNIT I: Multivariate Probability Distributions and Multiple Regression [12L]

Definition of random vectors and multivariate probability distributions, multivariate CDF, discrete and continuous random vectors. Singular and non-singular distributions. Marginal and conditional distributions. Mean vector and dispersion matrix, properties of a dispersion matrix and associated results. Concepts of multiple regression, multiple correlation coefficient, partial correlation coefficient, partial regression coefficients and their significance.

UNIT II: Some Special Multivariate Distributions [12L]

Definition of a multivariate normal distribution (Both non-singular and singular case), MGF of a multivariate normal distribution and its uses. Marginal and conditional distributions. Distributions of quadratic forms. Sampling from a multivariate normal distribution. Definition of Wishart distribution and its properties. Hotelling's T^2 statistic, its distribution and uses.

UNIT III: Principal Component Analysis and Factor Analysis [12L]

Principal Component Analysis: introduction, objectives, geometrical significance, population principal components and associated results, principal components for standardized variables, Scree plot, summarizing sample variation by principal components. Factor Analysis: introduction, Orthogonal Factor Model along with its assumptions, concepts of common factors, specific factors, factor loadings and loading matrix, communality, specific variance. Methods of estimation: Principal Component method. Concept of Factor Rotation - Varimax Rotation.

UNIT IV: Clustering and Classification Methods [12L]

Distance measures for continuous, discrete and mixed data. Clustering: Hierarchical clustering – Agglomerative and Divisive algorithms, Partitioning: K- means or medoids clustering, optimum choice of the number of clusters. Classification and discrimination procedures for discrimination between two multivariate normal populations-sample discriminant function, tests associated with discriminant functions, probabilities of misclassification and their estimation, Bayes and Mini-max rules of classification.

Text & Reference books:

1. Gun A.M., Gupta M.K. and Dasgupta B.: Fundamentals of Statistics, Vol. I, The World Press Private Limited, Kolkata.
2. Gun A.M., Gupta M.K. and Dasgupta B.: An Outline of Statistical Theory, Vol. I, The World Press Private Limited, Kolkata.
3. Gun A.M., Gupta M.K. and Dasgupta B.: An Outline of Statistical Theory, Vol. II, The World Press Private Limited, Kolkata.
4. Anderson T.W.: An Introduction to Multivariate Statistical Analysis, 3rd Edition, John Wiley & Sons, Inc.
5. Johnson R.A. and Wichern D.W.: Applied Multivariate Statistical Analysis, 6th Edition., Pearson Education, Inc.
6. Kshirsagar A.M. (1972): Multivariate Analysis, 1st Edition, Marcel Dekker.

Subject Code - STATP301T02	CC17 – Statistical Quality Management and Reliability	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*Basic knowledge of fundamentals and applied Statistics.*				
Co-requisites	----				

Learning objectives:

To provide the basic concepts of Statistical Quality control and its application in Industry and Application of Reliability theory in practical life.

Course learning outcomes*:*

***CO1*: Define** the underlying concept of a Statistical Quality Management and Reliability.

***CO2:* Classify** the problem of process control and product control.

***CO3:* Apply** various Variable and Attribute Control Charts to comment on the state of control of a given production process.

***CO4:* Analyze** the outcomes of using various charts and curves and **explain** the concepts of Six-Sigma and ISO quality standards.

CO5: Categorize different type of failure models.

***CO6:* Design** System reliability and **estimate** the parameters in Reliability model and construct confidence interval.

Course Content:

UNIT I: Statistical Quality Management **[24L]**

General concept of quality, role of SQM. Rational subgroup and Shewhart's control chart technique (attribute and variable), OC and ARL, modified control chart, cusum control chart for mean. Product control, acceptance sampling plan by attributes, single, double, multiple and sequential sampling plans, OC and ASN functions, AQL and ATI, acceptance sampling by variables, continuous sampling plan. Process capability analysis, Capability indices–Cp, Cpk and Cpm estimation for normally distributed characteristic.

UNIT II: Reliability Theory **[24L]**

Reliability concepts and measures, parallel and series system. Life distribution, reliability function, hazard rate, means remaining life, common life distributions. Life testing experiment involving exponential and Weibull distributions (both complete and censored cases), Estimation and confidence interval for reliability, Two models proportional hazard and the location-scale models for exponential and Weibull distributions.

Text & Reference books:

1. Montogomery, D. C. (2009): Introduction to Statistical Quality Control, 6th Edition, Wiley India Pvt. Ltd.
2. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edn. The World Press, Kolkata.
3. R. E. Barlow and F. Proschan: Statistical Theory of Reliability and Life Testing.
4. J. F. Lawless: Statistical Models and Methods for Lifetime Data.
5. Mukhopadhyay, P (2011):Applied Statistics, 2nd edition revised reprint, Books and Allied(P) Ltd.
6. Montogomery, D. C. and Runger, G.C. (2008): Applied Statistics and Probability for Engineers, 3rd Edition reprint, Wiley India Pvt. Ltd.
7. Ehrlich, B. Harris (2002): Transactional Six Sigma and Lean Servicing, 2nd Edition, St. Lucie Press.
8. Hoyle, David (1995): ISO Quality Systems Handbook, 2nd Edition, Butterworth Heinemann Publication.

Subject Code - STATP301T03	CC18 – Biostatistics	C	L	T	P
Version 1.0	**Contact Hours – 48**	4	3	1	0
Pre-requisites/Exposure	*Basic knowledge of Probability Distributions, Statistical Inference and Regression*				
Co-requisites	----				

Learning objectives:

To make the students aware about the real-life applications of Survival Analysis and Competing Risk Theory. Various survival distributions along with different censoring approaches will be elaborated to provide insights regarding real life problems.

Course learning outcomes*:*

CO1: **Demonstrate** the introductory concepts of Survival Analysis.

CO2: **Develop** concepts of various methods of estimation of survival functions.

CO3: **Build** knowledge about hazard model with one and multiple covariates.

CO4: **Explain** the concepts of Clinical trials and its real-life applications.

CO5: **Apply** the concepts of different multivariate survival models.

CO6: **Make use** of the knowledge of Competing Risk Theory.

Course Content:

UNIT I: Introduction to Survival Analysis, Risk Analysis and Clinical Trials **[6L]**

Concepts of time, various censoring mechanism and likelihood in those cases

UNIT II: Parametric inference **[16L]**

Point estimation, confidence intervals, scores, likelihood ratio tests for selected parametric models. Estimation of survival function from censored data: Acturial estimator, Kaplan-Meier and Nelson – Aalen estimators.

UNIT III: Regression model for Survival data **[10L]**

Cox's proportional hazard model with one and multiple covariates, Accelerated failure time model.

UNIT IV: Competing risk theory and Clinical Trials **[16L]**

Multivariate survival models; Random effects models for survival data analysis. Clinical Trials: General concepts, some useful designs and analysis of clinical trials, use of prognostic factors.

Text & Reference books:

1. R. E. Elandt– Johnson and N. L. Johnson: Survival models and Data Analysis.
2. D. R. Cox and D. Oakes: Analysis of Survival Data.
3. A. J. Gross and V. A. Clark: Survival distribution: Reliability application in the Bio – medical Sciences.
4. R. G. Miller: Survival Analysis E. T. Spurgeon: Life Contingencies.
5. J. L. Fleiss: Design and Analysis of Clinical Experiments.
6. E. T. Lee: Statistical Methods for Survival data Analysis.

Subject Code - STATP301P04	CC 19 – Statistics Practical - IV	C	L	T	P
Version 1.0	**Contact Hours** – **72**	3	0	0	6
Pre-requisites/Exposure	*Multivariate Analysis, Statistical Quality Management and Reliability and Biostatistics courses.*				
Co-requisites	*Use of Scientific Calculators, Statistical Tables and R-Software*				

Learning objectives:

To provide students the hands-on experience of different statistical methods relating the theory courses of M.Sc. 4th semester Statistics.

Course learning outcomes*:*

CO1: **Build** the idea for solving real life problems involving multivariate distributions.

CO2: **Make use of** the idea of multivariate normal and Wishart distribution to solve various real life problems.

CO3: **Apply** the knowledge of process control and product control to solve various industry based statistical problems.

CO4: **Apply** the knowledge of Reliability theory to solve various real-life problems.

CO5: **Solve** different real life problems using clinical trials.

CO6: **Apply** the knowledge of Survival analysis to solve various statistical problems.

Course Content:

Practical problems related to Multivariate Analysis, Statistical Quality Management and Reliability and Biostatistics courses **(To be executed using Scientific Calculators, MS Excel and R-Software).**

Text & Reference books:

All the references of Multivariate Analysis, Statistical Quality Management and Reliability and Biostatistics courses.
