Progressive Education Society's

Modern College of Arts, Science and Commerce Ganeshkhind, Pune 16

(An Autonomous College Affiliated to Savitribai Phule Pune University)

Two Year Degree Program in Statistics

(Faculty of Science & Technology)

M.Sc. (Statistics) Part-I

Choice Based Credit System Syllabus

To be implemented from Academic Year 2022-23

Title of the Course: M. Sc. (Statistics)

Preamble:

M. Sc. Statistics program of semester pattern. There will be 4 semesters and in each semester syllabus of 20 credits will be covered. In this way minimum 80 credits are spread over four semesters. The program is structured so as to increase the employability of the candidates in industry, academics and other by providing knowledge and skills in depth. Accordingly, the program has important features such as individual/ group projects, elective courses and courses on standard software packages such as MATLAB, MINITAB, SPSS, R. Syllabus of the first two semesters covers core courses. The second-year syllabus contains both core, elective and open courses. It is possible for the students to study basic courses from other disciplines such as economics, life sciences, computer science, and mathematics in place of electives.

Introduction:

(a) M. Sc. Statistics program will be conducted under credit system in four semesters. There will be 20 credits in each semester for a total of 80 credits. One creditis taken to be equivalent to 15 clock hours of teaching.

(b) The program consists of core courses which may be compulsory or electives.

(c) In addition, there are lab courses (practical) and a project course.

(d) Some courses are termed Open Courses (O). The open courses are those which can be offered to other departments.

For every course, there will be Continuous Internal Assessment (CIA) conducted by department or college and End of Semester Examination (ESE) conductedby the University at the end of semester.

Structure of the Syllabus:

Following is the structure of two year M.Sc. Statistics program.
T: Theory P: Practical O: Open C: Compulsory E: Elective

M.Sc.(Statistics)-Part I

Course Code	T/ P	O/C/ E	Title	Number of credits	ESE Duration	Marks Assigned
Semester I						
22-ST – 11	**T**	**C**	Basics of Real Analysis and Calculus	**4**	3 Hours	100
22-ST – 12	**T**	**C**	Linear algebra and Numerical methods	**4**	3 Hours	100
22-ST – 13	**T**	**C**	Probability Distribution	**4**	3 Hours	100
22-ST – 14	**T**	**C**	Sampling Theory	**4**	3 Hours	100
22-ST – 15	**P**	**C**	Practical I	**4**	3 Hours	100
			Total	**20**		500
Semester II						
22-ST – 21	**T**	**C**	Probability Theory	**4**	3 Hours	100
22-ST – 22	**T**	**C**	Regression Analysis	**4**	3 Hours	100
22-ST – 23	**T**	**C**	Statistical Inference I	**4**	3 Hours	100
22-ST – 24	**T**	**C**	Multivariate Analysis	**4**	3 Hours	100
22-ST – 25	**P**	**C**	Practical II	**4**	3 Hours	100
			Total	**20**		500

Course Code	T/ P	O/C/ E	Title	Number of credits	ESE Duration	Marks Assigned
Semester I						
22-ST 191	**T**	**C**	Human Rights -I	**1**	2 Hours	25
22-ST 192	**T**	**C**	Cyber Security -I	**1**	2 Hours	25
			Total	**2**		50
Semester II						
22-ST 291	**T**	**C**	Human Rights -II	**1**	2 Hours	25
22-ST 292	**T**	**C**	Cyber Security -II	**1**	2 Hours	25
			Total	**2**		50

M.Sc.(Statistics)-Part II

Course Code	T/ P	O/C/ E	Title	Number of credits	ESE Duration	Marks Assigned
Semester III						
23-ST – 31	**T**	**C**	Statistics Theory Paper I	**4**	3 Hours	100
23-ST – 32	**T**	**C**	Statistics Theory Paper II	**4**	3 Hours	100
23-ST – 33	**T**	**C**	Statistics Theory Paper III	**4**	3 Hours	100
23-ST – 34	**T**	**C/E**	Statistics Theory Paper IV	**4**	3 Hours	100
23-ST – 35	**P**	**C**	Practical I	**4**	3 Hours	100
			Total	**20**		500
Semester IV						
23-ST – 41	**T**	**C**	Statistics Theory Paper I	**4**	3 Hours	100
23-ST – 42	**T**	**C**	Statistics Theory Paper II	**4**	3 Hours	100
23-ST – 43	**T**	**C/E**	Statistics Theory Paper III	**4**	3 Hours	100
23-ST – 44	**T**	**C/O**	Statistics Theory Paper IV	**4**	3 Hours	100
23-ST – 45	**P**	**C**	Practical II	**4**	3 Hours	100
			Total	**20**		500

Course Code	T/ P	O/C/ E	Title	Number of credits	ESE Duration	Marks Assigned
Semester III						
23-ST 391	**T**	**C**	Soft Skills	**2**	4 Hours	50
23-ST 392	**T**	**C**	Cyber Security -III	**1**	2 Hours	25
			Total	**3**		75
Semester IV						
23-ST 491	**P**	**C**	Cyber Security -III	**1**	2 Hours	25
23-ST 492	**P**	**C**	Skill Development	**2**	4 Hours	50
			Total	**3**		75

Semester I

22-ST-11: Basics of Real Analysis and Calculus

Course Outcomes (Cos):

CO1) Students will recall the set theory.
CO2) Students will understand the sequence and series.
CO3) Students will be able to apply the theorems of convergence and divergence of sequence and series.
CO4) Students will be able to find expected values for all types of variables.

Course Content:

Unit1 **(15L)**

- Review of Set theory, Set of real numbers, Supremum and infimum of sets of real numbers,
- realfield, ordered set and field, Archimedean principle
- countable and uncountable sets, countability of rational numbers, uncountability of real numbers,
- Metric space, Properties of metric space
- Neighborhood points, Exterior and interior points, boundary points, limit points, open, closed andcompact sets.
- Bolzano – Weierstrass and Heine- Borel theorem. (Only statement)

Unit 2 **(15L)**

(i) Sequence of real numbers:

- limit of sequence and its properties,
- Convergence and divergenceof sequence
- Cauchy sequence and related theorems (Cauchy criteria of convergence), subsequence and their convergence and divergence,
- convergence of bounded monotone sequence.

(ii) Series of real numbers:

- convergence and divergence of series of real numbers, test for convergence (root test, ratio test),
- absolute convergence (without proof), uniform convergence, power series,
- radius of convergence of power series (Binomial, Exponential, geometric and logseries), term by term differentiation (integration) of absolute convergent series
- change of order of summation of series.

Unit 3 **(15L)**

- Revision: derivative of function of single variable, Mean value theorem, Taylor's series expansion
- Multivariate calculus: explicit and implicit functions, continuity,
- partial derivatives, differentiability, partial derivatives of higher order,
- Hessian matrix,
- Taylor's theorem,
- extreme values, differentiation with respect to vector and matrix,
- Jacobian of transformation.

Unit 4 (15L)

(A) Riemann Integrals:

- **Riemann and Riemann – Stieltje's integral:** Partition of interval, norm of partition,
- finer partition, tagged partition,
- Upper and Lower Riemann and R-S sums, order relation between Upperand Lower Riemann sums.
- Effect of finer partition on the difference between Upper and Lower Riemann sums.
- Necessary and sufficient condition for a function to be R and R-S integrable Properties with proof
  (i) Continuous bounded function is R and R-S integrable
  (ii) Monotonic bounded function is R and R-S integrable
  (iii) Fundamental theorem of integral calculus.
- Application: Mathematical expectation for mixed type of distribution.

(B) Improper integrals:

- Definition, convergence of an integral, types of integral, type I, type II, P- integral, exponential integraltest for convergence (comparison test),
- convergence of beta and gamma integrals, relationship between beta and gamma functions

(C)Double integrals:

- Definition, properties, change of order, iterated integrals, Fubini's theorem, differentiation under integral sign (Leibnitz rule) and transformation of variables.

Books Recommended:

Sr. No.	Title	Name of the author	Publications
1	A Basic Course in Real Analysis	Ajitkumar and S. Kumaresan (2014)	Chapman and Hall
2	Principles of Mathematical Analysis	Rudin W. (1985)	McGraw–Hill
3	Mathematical Analysis: A Modern Approach to Advanced Calculus	Apostol T. M. (1975)	Addison - Wesley)
4	Elements of Real Analysis	Bartle R. G. (1976)	Wiley
5	Mathematical Analysis	Malik S. C. &Arora S. (1991)	Wiley Eastern Limited 2nd edition
6	Methods of Real Analysis	Goldberg R. R. (1964)	Blasdell Publishing company, New York,U.S.A.
7	Introduction to Real Analysis	Bartle G.R. &Sherbert D. R. (2000)	John Wiley & Son Inc.
8	Principles of Real Analysis	Royden (1988)	Macmillian

Semester I

22-ST-12: Linear algebra and Numerical methods

Course Outcomes (Cos):

CO1) Students will be able to recall the basic notions in linear algebra those are often used in Statistical analysis.
CO2) Students will be able to understand the fundamental properties of matrices including determinants, inverse of matrix, matrix factorization, eigenvalues and their transformations.
CO3) Students will be able to learn vector spaces, subspaces and their related results.
CO4) Students will be able learn various properties of canonical forms.
CO5) Students will be able to study of inner product spaces.
CO6) Students will be able to apply concepts of Gram Schmidt orthogonalization process.
CO7) Students will be able to apply numerical methods to obtain approximate solutions to mathematical Problems.

Course Content:

Unit 1 **(15 L)**

- Vector Space, Subspace, Linear dependence and independence, Basis of a vector space, dimension of vector space, orthogonal and orthonormal basis, Gram –Schmidth orthogo-nalization,
- Matrix algebra, special types of matrices, rank, inverse and determinant of a matrix and their properties, Orthogonal and idempotent matrix and their properties
- Projection theorem, linear transformation, linear equations Solution space and null Space.

Unit 2 **(12L)**

- Generalized (g) inverse and Moore-Penrose g- inverse (MP g-inverse), its properties and Examples,
- System of homogeneous and non- homogeneous linear equations, solution space andnull Space.

Unit 3 **(15L)**

- Characteristic roots (Eigen values) of real matrices, right and left characteristics vectors (eigen vectors)
- Properties and utility of eigen vector
- Independence of characteristics vectors corresponding to distinct Characteristic roots, Algebraic and geometric multiplicity, Spectral decomposition, power of a matrix,
- Cayley- Hamilton theorem, singular value decomposition

Unit 4 **(12 L)**

- Introduction of quadratic forms, reduction and classification of a quadratic form, simultaneous reduction of two quadratic forms, Maxima and minima of a quadratic form,
- Properties of a quadratic form for orthogonal and idempotent matrices.

Unit 5 **(6 L)**

- Newton Raphson iterative method for two or more simultaneous transcendental equations
- Newton's bivariate interpolation formula, Simson's, Trapezoidal rule for bivariate integrals

Books Recommended:

Sr. No.	Title of Book	Name of the Author	Publications
1	Linear Algebra and Linear Models	Bapat, R.B.	Springer and Hindustan Book Agency. (2011)
2	A First Course in Linear Algebra	Beezer, R. A.	Congruent Press, Washigton (2004)
3	Linear Algebra: A Geometric Approach	Kumaresan, S.	Prentice Hall, (2000),
4	Linear Algebra and Its Applications	Lay, D. C. Lay, S. R. and Mc Donald, J. J.	Fifth Edition, Pearson, Boston, (2016)
5	Linear Statistical Inference and Its Applications	Rao, C. R.	Wiley,(1995)
6	Matrix Algebra Useful for Statistics	Searle, S. R.	John Wiley, New York,(1982)
7	Introductory methods of Numerical Analysis	Sastri	Prentice Hall : fourth edition), (2009)
8	Computer oriented Numerical Methods	Rajaraman	Prentice Hall, (1993)
9	Linear Algebra	RamachandraRao, A. and Bhimasankaram, P.	Hindustan Book Agency, (2000)
10	Elements of Matrix Algebra	Hohn, F. E.	Macmillan, (1973)
11	Advanced Multivariate Statistics withMatrices	Kollo, T. and Rosen, D. von	Springer, New York.

Semester I

22-ST-13: Probability Distributions

Course Outcomes:

CO1) Students will be able to recall the random variable as a function defined on sample space .

CO2) Students will be able to understand the random variable as a measurable function on probability space using the concepts like sigma field, set function, Borel measurable function.

CO3) Students will know about necessity of existence of absolute moments, uniqueness of m.g.f. Also they can derive the probability distributions using p.g.f., characteristic function and convolutions for sums of independent random variables.

CO 4) Students should use some other probability distributions such that Bivariate Poisson Distribution, Bivariate Exponential Distribution, Dirichlet Distribution, Non- central chi square, F and t distribution.

CO 5) Students will understand the concepts like order statistics, quadratic forms and distribution free statistics and explore the respective applications.

Course Content:

Unit 1 **(15 L)**

- **Random Variable:** Brief introduction to sigma field, probability space, set function, measure, finite measure, probability measure, axioms of probability, Borel field, Borel measurable function, random variable as a measurable function on probability space.
- **Distribution function:** CDF of random variable, continuity theorem of limit of probability (without proof), characteristic properties of CDF (with proof). Decomposition of CDF, mixtureof distributions. Identification of given function whether it is CDF. Results (i) product of distribution function is CDF (ii) If F(.) is a CDF then $F^n(.)$, $1-(1-F(.))^n$ are CDF for n positive integer. (iv) Convex combination of CDFs is CDF.
- Bivariate random variable. CDF of bivariate random variable and its characteristic properties.Identification of given function whether it is CDF.
- **Symmetry**: Symmetric probability distribution around a, Concept, mean median mode need not coincide. Results: (i) Odd ordered central moments of symmetric distribution are zero. (ii) Sum and difference of random variables with symmetric distribution is symmetric.
- **Transformations:** Transformation of random variable of the type (i) one to one on to (ii) monotonic (iii) non monotonic.

Unit2 **(12 L)**

- Expectation of random variable, necessity of existence of absolute moments, uniqueness of moments, existence of r thorder moments if s thorder moment exists ($r \leq s$).
- **MGF:** Existence of MGF, properties, MGF of symmetric distributions.
- **PGF.:** Properties, moments using PGF, probability distributions of X+Y, X-Y when X andY are not identically distributed random variables. Compound distribution and its PGF. Wald's identity $E(X_1+X_2+ - - - + X_N) = E(N).E(X_i)$ where X_i's are i.i.d r.v.s. and independent of N. Var $(X_1+X_2+ - - - + X_N)$.

Results (i) If $P_1(s)$ and $P_2(s)$ are PGF of independent r.v. s then $\{P_1(s) .P_2(s)\}$ is a PGF

(ii) $P_1(P2(s))$ is PGF

(iii)$[P_1(s)]^n$ is a PGF n positive integer. (iv) $P_x(s)= P_{-x}(s)$ if X is symmetric around zero. Convolutions of random variables. Distributions of X+Y, X-Y in case of U(0, 1), Normal, exponential etc.

- **Characteristic function**: characteristic function and properties, conjugate pairs of distributions, Parseval relation, uniqueness theorem
- **Random vector:** Joint and marginal distributions, mixed moments, covariance matrix, conditional mean and variance,

  Results (i)E{E(X|Y)}=E(X)

  (ii)Var(X)=E{Var(X|Y)}+Var{E(X|Y)}.

  (iii)If E(Y|x)=a +bx then b is the regression coefficient of y on x .

  MGF of random vector.
- **Regression:** Conditional expectation as regression function

Unit 3 **(18 L)**

- **Bivariate Poisson Distribution:**

  Definition, derivation of p. m. f, m. g. f of bivariate Poisson distribution, marginal p.m.f of

  variables, c.g.f, correlation coefficient, conditional distribution of X1|X2=x2, problems.

- **Bivariate exponential distribution:**

  Types of bivariate exponential distribution:

  1) Marshall –Olkin model: joint distribution, marginal distribution, lack of memory property, Properties of Marshall –Olkin model
  2) Fruend bivariate exponential model
  3) Moran's bivariate exponential model
  4) Gumblel's bivariate exponential model

  Theorem: Decomposition of bivariate exponential distribution into absolutely continuous and singular parts.

- **<u>Dirichlet Distribution:</u>**

  Definition, derivation of joint p.d.f , particular case of dirichlet distribution, Marginal distribution of X and Y (for particular case k=3) .Conditional distribution of X|Y and Y|X.

Unit 4 **(15 L)**

- **Review of order Statistics:**

  Definition, Marginal pdf of $X_{(r)}$, Joint pdf of $(X_{(r)}, X_{(s)})$, pdf of sample range, corr$((X_{(r)}, X_{(s)})$ when random sample from U(0,1), problems based on exponential . Distributions of spacings and standardized spacings
- **Probability integral transformation.** Quantile function.

- **Distribution free statistics:** Sign test, Wilcoxon sign rank test, Kolmogorov –Smirnov test, Run test. (Derivation of distribution of test statistic)
- **Quadratic forms**: Quadratic forms, Classification of quadratic forms Sampling distribution of quadratic forms and linear forms for random sample from normal distribution. distribution of quadratic forms. All result require for Fisher Cochran theorem, Fisher Cochran theorem.

- **Non-central distribution:**

  Non-central chi square distribution, derivation of p.d.f. , m.g.f. , meanvariance, applications.
  Non-central t-distribution distribution, derivation of pdf , mgf , mean variance, applications.

  Non-central F- distribution, derivation of p.d.f, m.g.f , mean variance, applications.

Books Recommended

Sr. No.	Title of the book	Name of the author	Publication
1	Statistical Inference	Berger R. and Casella G.	Second Edition, Duxbury Resource Center (2002)
2	Fundamentals of Probability: A First Course	Dasgupta A.	Springer, NewYork. (2010)
3	Introduction to Mathematical Statistics	Hogg R. V. ,J. W. and Craig, McKean T.T.	Sixth Edition (2005), Pearson Prentice Hall, New Jersey
4.	Linear Statistical Inference and Its Applications	Rao C. R.	Wiley (2002)
5.	Introduction to Probability and Statistics	Rohatgi V. K. & A. K. M. E Saleh	Wiley, New York (2001)
6.	An Introduction to the Theory of Probability	Parimal Mukhopadhyay	World Scientific Publishing company(2011)
7.	Continuous Multivariate Distributions	Johnson ,Kotz and Balakrishnan	John Wiley and Sons, Vol I (2000)

Semester I

22-ST-14: Sampling Theory

Course Outcomes (COs):

CO1) Students should recall the basic concepts of sampling (Random and non-Random)
CO2) Students will able to understand the basic principles underlying survey design and Estimation
CO3) Students will be able to apply unequal probability sampling designs viz. PPSWR, PPSWOR including Lahiri's and cumulative total method for data.
CO4) Students will able to explain and to compare various allocations using stratified random Sampling.
CO5) Students will use practical applications of ratio and regression method of estimation
CO6) Students will able to apply various sampling methods for real life data.

Course Content:

Unit 1 **(15L)**

- Review of finite population sampling techniques (random and non-random),
- SRSWR and SRSWOR: Inclusion probabilities, related results on estimation of population total, Determination of sample size for pre-specified variance, pre-specified error in the estimation, pre-specified width of the confidence interval, pre-specified relative error in the estimation, Simple random sampling for the proportion**,** Estimation of proportion for the more than two classes**,** Inverse Sampling (Sampling for the rare attribute) and estimator of the population mean and its variance**.**
- Probability Proportional to Size with Replacement (PPSWR) methods:
- Cumulative total method
- Lahiri's method
- For estimation problem, estimation of finite population mean and total.
- Sampling with varying probability without replacement, Ordered and Unordered estimates, Horwitz–Thompson estimator: its variance and properties, Midzuno scheme of sampling, Yates-Grundy sampling estimate, and Murthy's estimate.

Unit 2 **(15L)**

- Stratified sampling, comparison of allocation problem of allocation in stratified sampling, Construction of strata, deep stratification, The method of Collapsed strata,
- Post stratification, estimator of population mean and variance of estimator of population mean under post stratification, stratified random sampling for the proportion, comparison of stratified random sampling with simple random sampling. Use of supplementary information for estimation,

- Ratio estimator of population mean, its bias and mean square error,
- Unbiased ratio type estimators of population mean, variance of estimator of population mean under it,
- Ratio estimator for the stratified random Sampling
- Regression method of estimation, estimator of population mean, its bias and mean square error of the Estimator,
- Comparison of estimator of population mean under ratio, regression and simple random sampling,
- Jack-Knife and bootstrap method of estimation, estimate of bias and standard error.

Unit 3 **(15L)**

Systematic sampling, sample mean and its variance, Comparison of systematic sampling under linear trend, Yates corrected estimator,

- Centered systematic sampling,
- Balanced systematic sampling
- Modified systematic sampling,
- circular systematic sampling,
- two-dimensional systematic sampling (Aligned and Unaligned Systematic sampling),
- Comparison of systematic sampling with random sampling and stratified sampling, PPS systematic sampling.
- Cluster sampling with clusters of equal sizes, estimation of population mean and its standard error, Relative efficiency of cluster sampling w. r. t. SRSWOR, Effect of cluster size on relative efficiency, unbiased estimator of relative efficiency, cluster sampling as a one way ANOVA, Optimum value of the cluster size, cluster sampling for the proportion.
- Cluster sampling with clusters of unequal sizes, bias in estimator of population mean, bias in the estimator and its MSE, unbiased estimator and relative efficiency of unequal cluster sampling, PPS cluster sampling, estimation ofpopulation mean.

Unit 4 **(15L)**

- Two stage sampling with equal and unequal second stage units,
- estimation of population means and its standard error with equal second stage units, optimum value of the number of clusters and second stage units in the two-stage sampling having equal second stage units.
- Two phase sampling, ratio and regression estimator of population mean under two phase sampling, bias in the estimator and its MSE,
- Sampling and non– sampling errors
- Response errors, mathematical model for Response errors,
- Hansen Hurwitz technique, Randomized Response Technique (RRT),
- Warner's randomized response technique.

Books Recommended

Sr. No.	Title of the book	Name of Author	Publication
1	Survey Sampling: Theory & Applications	Arnab, R. (2017).	Academic Press
2	Modern Survey Sampling	Chaudhuri, A. (2014).	CRC Press
3	Sampling Techniques,	Cochran, W.G. (1984).	Wiley.
3	Sampling Techniques,	Cochran, W.G. (1984).	Wiley.
5	Sampling theory and methods	Murthy M.N.(1977)	Statistical Publishing Society
4	Sample Survey Theory	Des Raj and Chandhok, P. (1998).	Narosa.
7	Theory and Analysis of Sample SurveyDesigns	Singh, D. and Chaudhary F.S (1986).	Wiley Eastern Limited.
5	Sampling theory and methods	Murthy M.N.(1977)	Statistical Publishing Society
9	Sampling Theory of Surveys with Applications,	Sukhatme, P.V, Suktatme, B.V., Sukhatme, S. and Asok, C. (1984).	Indian Society for Agricultural Statistics, New Delhi.
6	Sampling theory and methods	S. Sampath (2005)	Narosa
11	Sampilng Theory.	Parimal Mukhopadhyay	New Central Book Agency

Semester I

22-ST-15: Practical I

Teaching Scheme: TH: 8 Lectures /Batch/Week **Credit: 04**

Prerequisite Courses:

- Knowledge of MS-Office, C language, R software
- Basic concepts of linear algebra, sampling theory, probability distributions and numerical analysis

Course Outcomes:

- Students can solve problems from sampling theory, probability distributions, linear algebra etc. using statistical software.

Course Contents:

Sr. No.	Title of the Experiment	No. of Practical
1.	Matrices: Properties of Matrices (rank, transpose, determinant etc.), Getting vectors in row/column space and null space of the given matrix. Eigen values and Eigen Vector of matrix, power of matrix.	1
2.	Inverse of a square matrix (by direct method and partitioning method), g-inverse, MP g- inverse.	1
3.	Gram-Schmidt orthonormalization: Forming an orthogonal matrix of specified orderusing Gram-Schmidt orthogonalization.	1
4.	Solution of System of Linear Equations using Gauss elimination, Gauss Jorden, Gauss-Seidal and Gauss-Jacobbi methods.	2
5.	Classification and Reduction of Quadratic forms, Verification of Cayley-Hamilton theorem	1
6.	Model sampling from discrete, continuous and mixture distribution (Use inversion method if necessary).	1
7.	Model sampling from bivariate probability distribution. Computation of probability ofevents related to bivariate probability distribution	1
8.	Computation of probability of non-central χ^2, t, F-distributions.	1
9.	H-T estimator and PPS, π PS (Midzuno) designs. Confidences Interval of estimator	2
10.	Stratified Random Sampling a. Various kinds of allocation and estimation of population total and mean with S.E. b. Post stratification.	1

11.	Stratified Random Sampling: a. Ratio method of estimation b. Regression method of estimation	1
12.	Circular Systematic Sampling	1
13.	Cluster Sampling with equal and unequal cluster size	1
14.	Jackknife and bootstrap methods of estimation(for Ratio, Regression coefficient,Coefficient of variation, Correlation coefficient)	1
15.	Two stage sampling	1
16.	Numerical methods: (i) Solution to Simultaneous Bivariate equations by Newton Raphson method (ii) Newton's Interpolation for bivariate functions (iii) Evaluation of double integral by Trapezoidal rule and Simpson's rule	1
19.	Review of the Research Paper.	2
Total		**20**

Note:

1. Maximum 4 students are allowed in a group.
2. At least two research papers should be reviewed by each group.
3. At the end of semester, the reviewed research papers should be presented.

Semester II

ST 21: Probability Theory

Course Outcomes (CO's):

CO1) Students will recall the concept of field, measurable space, distribution function.
CO2) Students will understand the sequence of random variables.
CO3) Students will learn convergence in probability, distribution.
CO4) Students will be learn different theorems related to independence of random variables.

Course Content:

Unit1 **(15L)**

- Review of algebra of sets, sequence of sets, lim-sup, lim-inf and limit of a sequence of sets, classes of sets, field, sigma field, minimal sigma field,
- Borel fields, measurable space, monotone classes, Measurable function, Real and Vector valued random variables,
- simple random variable, randomvariable as a limit of sequence of simple random variables,
- Probability measure on a measurable space, probability space, properties of probability measure: continuity, mixture of probability measures,
- Lebesgue and Lebesgue-Steltjes measures.

Unit 2 (15L)

- Distribution function, J-H decomposition of a distribution function
- discrete and continuous type random variable,
- Correspondence theorem,
- Expectation of simple random variable, non-negative random variable,
- Arbitrary random variable, properties of expectation, moments, moment inequalities.

Unit 3 (15L)

- **Revision St 11, St 13**
- Convergence of a sequence of random variables, Examples and real-life situation
- convergence in probability,
- convergence in distribution,
- convergence in r^{th} mean,
- almost sure convergence, their inter-relations,
- Slutkey's Theorem,
- convergence theorem for expectations

Unit 4 (15L)

- Independence of events, class of independent events, independence of classes,
- independence of random variables, expectation of the product of independent random variables,
- equivalent definitions of independence,
- Kolmogorov 0-1 Law, Borel 0-1 criterion,
- Khintchin's WLLN,
- Strong Law of Large Numbers (SLLN) (Statement only),
- Central Limit Theorem (CLT),
- Levy continuity theorem, CLT for i.i.d. random variables,
- Liaponove's form, Lindeberg Feller form and their applications.

Books Recommended:

Sr. No.	Title of the book	Name of Author	Publication
1	Probability Theory	Athreya, K. B. and Lahiri S. (2006)	Hindustan Book Agency
2	Modern Probability Theory: An Introductory	Bhat, B. R. (2007)	New AgeInternational
3	Probability and Measure	Billingsley, P. (1995), 3rdEdition	John Wiley, NewYork
4	A Course in Probability Theory	Chung, K. L. (2001) Third Edition	Academic Press, London
5	Probability: A Graduate Course	Gut, Allan (2005),	Springer, New

Semester II

22-ST-22: Regression Analysis

Course Outcomes (COs):

CO1) Students will recall how to use linear regression models in practice: identify situation where linear regression is appropriate; build and fit linear and multiple regression models with software; interpret estimates and diagnostic statistics; produce exploratory graphs
CO 2) Students will able to apply the theory underlying point estimation, hypothesis and confidence intervals for linear regression models.
CO 3) Students will able to understand the diagnostic measures for Non-linear data such as transformation of data.
CO 4) Students will able to learn the Ridge and Poisson Regression model as real life application.
CO 5) Students will be able to apply regression technique in real life situation.

Course Content:

Unit 1 **(15L)**

- Brief review of simple linear regression: assumptions, least square (LS) estimators of parameters, standard error of estimators, testing of hypothesis for coefficient of regression
- Multiple regression: Standard Gauss-Markov (GM) setup, least square (LS) estimation with and without restrictions on parameters, variance and covariance of LS estimators, GM theorem (statement and proof for $Var(\varepsilon) = \sigma^2 I$ and $Var(\varepsilon) = \sigma^2 V$), estimation of error variance (with and without correlated observations).

Unit 2 **(15L)**

- Confidence intervals and regions, testing of hypothesis for one and more than one linear parametric functions, testing of hypotheses about parallelism (slopes), equality of intercepts, congruence of two simple regression models, lack of fit test
- Polynomial regression model (one and two regressors), orthogonal polynomial regression, cubic spline regression model.

Unit 3 **(15L)**

- Diagnostic checks and correction: graphical techniques, tests for normality (Shapiro test, Anderson-Darling test), uncorrelatedness, homoscedasticity; Criteria for model adequacy: R^2, adjusted R^2, Mallow's Cp etc.
- Durbin Watson test, estimation of parameters in autocorrelation
- Outlier, leverage points, influential points, PRESS statistic, Cook's D statistic
- Multicollinearity: consequences, tools for detection and remedies, Ridge Regression.

Unit 4 **(15L)**

- Non-linear regression: linearization transforms, their uses and limitations. Box and Cox transformations

- Generalized linear model: introduction to link functions such as binomial, inverse binomial, inverse Gaussian and Gamma.
- Logistic regression: Logit transform, ML estimation, tests of hypothesis, Wald test, LR test, score test, test for overall regression
- Poisson regression: log link transform, ML estimation, tests of hypothesis, Wald test, LR test, score test, test for overall regression.

Books Recommended:

Sr. No.	Title of the book	Name of Author	Publication
1	Regression Analysis of Count Data	Cameron, A. C. and P. K. Trivedi (1998).	Cambridge
2	Applied Regression Analysis	Draper, N. R. and Smith, H. (1998).	John Wiley, Third Edition
3	Applied Logistic Regression	Hosmer, D. W. and Lemeshow, S. (1989	Wiley
4	Logistic Regression	Kleinbaum, D. G. & Klein M. 2002	A Self-Learning Text,Springer
5	Generalized Linear Models	McCullough, P. and Nelder, J. A	Chapman& Hall
6	Applied Linear Statistical Models	Neter J.W.and Kutner M.H.	Wiley
7	Introduction to linear Regression Analysis	Montgomery D.C., Peck E.And Vining G.G.(2003)	Wiley
8	Nonlinear Regression Modelling	Ratkowsky, D. A. (1983,	Marcel Dekker, London.
9	Semi parametric Regression	Rupert, D., Wand, M. P. and Carroll, R. J. (2003),	Cambridge University Press.
10	Nonlinear Regression	Seber, G. E. F. and Wild, C. J. (1989)	Wiley.
11	Applied Liner Regression	Weisberg, S. (2005).	Wiley.

Semester II

22-ST-23-Statistical Inference

Course Outcomes (COs):

CO1) Students should recall various terms for Fisher Information, interval estimation to understand the problem of statistical inference.

CO2) Students will be able to compute Cramer – Rao lower bound in order to find most efficient estimator.

CO3) Students will be able to estimate the parameters with multiple criteria

i) Minimum variance Bound Unbiased

ii) Rao-Blackwell Theorem

CO4) Students will be able to analyze the estimation techniques using Confidence Interval and Bayes estimation.

CO5) Students will be able to solve the problems based on testing of hypotheses using various techniques.

Course Content:

Unit1 **(15 L)**

- Fisher information and information matrix
- Concept of Sufficiency, Neyman factorization theorem, likelihood equivalence, minimal sufficiency, construction of minimal sufficient statistics, special classes of distributions: one parameter exponential family, multiparameter exponential family, Pitman family, minimal sufficient statistic for special classes of distributions.

Unit 2 **(15 L)**

- Completeness, bounded completeness, complete sufficient statistics, special classes of distributions admitting complete sufficient statistics, ancillary statistic,
- Basu's theorem and itsapplications, estimability of parametric function.
- Cramer-Rao inequality, minimum variance unbiased estimators (MVUE), necessary and sufficient conditions for existence of MVUE,Minimum variance bound unbiased estimators (MVBUE),
- Rao- Blackwell theorem
- Chapman-Robin Bounds (without proof), Bhattacharya Bounds (without proof),
- Lehman- Scheffe theorem.

Unit3 **(15 L)**

- Problem of testing of hypothesis, simple and composite hypotheses.
- randomized and non- randomized tests
- Neyman-Pearson Lemma (with proof), most powerful test, power function of a test,
- Existence of UMP tests for one-sided alternatives in one parameter exponential family and

Pitman family, UMP tests for two sided alternatives, their existence and non-existence, unbiased test, UMP unbiased test

- Monotone likelihood ratio property and its applications.

Unit4 **(15L)**

- Concept of confidence intervals, relation with testing of hypothesis, shortest expected length confidence intervals (Revision of Pivotal quantity method, normal approximation method , studentized method,), uniformly most accurate confidence bounds.
- Introduction to Bayesian estimation, prior and posterior distributions, loss functions, principle of minimum expected posterior loss, quadratic and other common loss functions, conjugate family of prior distributions and its examples.

Books Recommended

Sr. No.	Title of the book	Name of Author	Publication
1	Statistical Inference	Casella, G. and Berger, R. L.	Duxbury Advanced Series, SecondEdition (2002)..
2	Computer Age Statistical Inference: Algorithms, Evidence andData Science.	Efron, B. and Hastie, T.	Cambridge University Press
3	Parametric Inference: An Introduction	Kale, B.K. & Muralidharan, K.	Alpha Science International Ltd. (2015)
4	Theory of Point Estimation	Lehmann, E.L. and Casella, G.	Springer, NewYork(1998).
5	Testing Statistical Hypotheses	Lehmann, E. L. and Romano, J.	Springer, New York (2005).
6	Linear Statistical Inference and its Applications	Rao, C. R.	Wiley(1995).
7	Introduction to Probability and Statistics	Rohatgi, V. K. and Saleh, A.K. Md. E.	JohnWiley & Sons, New York. (2001).
8	Mathematical Statistics	Shao, J.	Springer-Verlag, New,New York (2003)

Semester II

22-ST 24: Multivariate Analysis

Course Outcomes:

CO1) Students will be able to understand difference between one and multidimensional random variables.

CO2) Students should understand the concepts and distributions such as principal component analysis, factor analysis and Multivariate normal distribution

CO3) Students will be able to estimate MLEs of parameters of multivariate normal distribution and their sampling distribution.

CO 4) Students should be able to understand the concepts like Wishart distribution, Hotteling T^2 statistics.

CO 5) Students will apply MANOVA techniques and their respective applications.

Course Contents:

Unit 1 **(18 L)**

- Exploratory multivariate Data Analysis: Sample mean vector, Dispersion Matrix, Correlation Matrix, Linear transformation and its mean and variance, graphical interpretation.
- Principal component Analysis (by using covariance and correlation method, standardized method), Factor analysis (Their models, rotation types), Canonical correlation with real life examples.
- Cluster analysis (Hierarchical and Non- hierarchical, Agglomerative, Single, complete, average Wald's linkage , K- mean clustering method , qualitative method clustering)

Unit 2 **(12 L)**

- Multivariate normal distribution, Singular and non-singular normal distribution, mean, variance of multivariate normal distribution, independence of variable, M.G.F, Characteristic function, moments.
- Distribution of linear and quadratic form of normal variables, marginal and conditional distribution, multiple and partial correlation coefficient (3 random variable case)with examples on each of the topic.

Unit 3 **(15 L)**

- MLES of parameters of multivariate normal distribution and their sampling distribution, Tests and confidence region for the mean when dispersion matrix is known,
- Wishart distribution (generalized case of chi-square) and its properties,
- Hotelling T^2 statistics and uses of its distribution,
- Beharen- Fishers problem, confidence region for mean vector of multivariate normal distributions.

Unit 4 **(15 L)**

- MANOVA technique, Likelihood ratio test, Test for equality of dispersion matrices,
- Discriminant analysis (by using prior probabilities, by using cost), Fisher Discrimenant analysis, Mahalanobis D^2 Statistics

Books Recommended:

Sr. No.	Title of the book	Name of the author	Publication
1	Introduction to Multivariate Analysis	Anderson T.W.	John Wiley (1984)
2	Symmetric Multivariate and Related distributions	Fang. K., Kotz S., Ng K. W.	Chapman and Hall (1990)
3	Applied Multivariate Statistical Analysis	Härdle W. K. & Simar L.	Springer, Newyork (2012)
4.	Applied Multivariate Statistical Analysis	Johnson R.A. & Wichern D.W.	Prentice Hall , Inc. (1988)
5.	Continuous MultivariateDistributions	Kotz S., Balakrishnan N.,Jhonson N.L.	Volume 1 , John Wiley & Sons (2000)
6.	Multivariate Analysis	Kshirsagar A.M.	Marcel Dekker (1983)
7.	Directional Statistics	Morrison D.F.	Mc Graw Hill Co. (1990)
8	Multivariate Analysis and its Applications	K.C. Bhuyan	New Central Book Agency

Semester II

22-ST-25 Practical II

Teaching Scheme: TH: 8 Lectures /Batch/Week **Credit: 04**

Prerequisite Courses:

- Regression analysis
- Multivariate analysis

Course Outcomes:

- After learning this course student will be able to apply regression analysis and multivariate analysis techniques to real life data.

Course Contents:

Sr. No.	Title of Experiment	No. of Practical
1	Exploratory Multivariate data Analysis and Principal component Analysis (covariance & Correlation technique and their interpretation)	1
2	Factor analysis (PCA., MLE, all Rotations and their interpretation)	1
3	Cluster analysis (Single, Complete, Average, Wards , k- mean linkage method)	1
4	Canonical correlation	1
5	Multivariate Analysis. (Multivariate normality, Marginal, Conditional, Q-Q plot, contour plot)	1
6	Model Sampling from Multivariate Distribution, and computation of MLE's of Parameters.	1
7	Application of Hotelling T^2 Statistic	1
8	MANOVA technique	1
9	Likelihood Ratio Test(Equality of means, Equality of variance, R=0)	1
10	Discriminant Analysis (Fishers linear discriminant function)	1
11	Simple and Multiple Regression and Regression diagnosis	1
12	Selection of variables in Multiple regression and lack of fit	1
13	Transformation and weighting to correct model inadequacies.	1
14	Polynomial regression model (one and two regressors)	1
15	Multicolinarity and ridge regression.	1
16	Spline Regression	1
17	Logistic Regression and Poisson Regression	1
18	Application of Central Limit theorem and Weak law of large number	1
19	Progression Report of a project.	2
Total		**20**

Note:

1. Student should start the project from first semester and should discuss it with respective mentor.
2. At the end of second semester, progression report of project should submitted in given time.