

Yashavantrao Chavan Institute of Science, Satara

Department of Statistics

Course Structure under CBCS

M.Sc. (Statistics)

1. TITLE : M. Sc. (Statistics)

2. YEAR OF IMPLEMENTATION: 2018-19

3. PREAMBLE:

The goal of syllabus to make the study of Statistics popular and interesting among the students for job achievements as well as higher studies. The syllabus is prepared after discussion at length with number of faculty members of the subject and experts from industries and research fields. The units of the syllabus are well defined, taking into consideration the level and capacity of students.

Eligibility : B. Sc. with Statistics as principal subject.

4. GENERAL OBJECTIVES OF THE COURSE:

1. The students are expected to understand the principles, concepts and recent developments in the Statistics.

2. To enhance student sense of enthusiasm for Statistics and to involve them in an intellectually stimulating experience of learning in a supportive environment.

3. The practical course is framed in relevance with the theory courses to improve the understanding of the various concepts in Statistics.

5. DURATION: Two Years

6. PATTERN: M. Sc. Statistics program has semester pattern and Choice

Based Credit System. The program consists of 96 credits.

7. MEDIUM OF INSTRUCTION: English

8. STRUCTURE OF COURSE:

Notations:

A six-character code is given to each paper. In MST “M” stands for M.Sc., “S” stands for Statistics, “T” stands for Theory and “P” stands for practical. The first digit following MST is Semester Number. The second digit “0” stands for the core theory course, the digit “1” stands for a practical paper and the third digit indicates the serial number of paper in that semester.

Course Structure with details about instruction hrs per week, credits etc.:

M.Sc. (Statistics) Semester-I (24 credits)

M.Sc. (Statistics) Semester – I

Course Code	Title of the course	Instruction Hrs/week	Duration of Exam(Hrs)	Marks-End Semester Exam	Marks-Internal Assessment	Credits
MST 101	Real Analysis	4	3	80	20	4
MST 102	Linear Algebra	4	3	80	20	4
MST 103	Distribution Theory	4	3	80	20	4
MST 104	Estimation Theory	4	3	80	20	4
MST 105	Optimization Technique	4	3	80	20	4
MSP 116	Practical-I	12	3	-	100	4

M.Sc. (Statistics) Semester – II

Course Code	Title of the course	Instruction Hrs/week	Duration of Exam(Hrs)	Marks-Term End Exam	Marks-Internal Assessment	Credits
MST 201	Probability Theory	4	3	80	20	4
MST 202	Theory of Testing of Hypotheses	4	3	80	20	4
MST 203	Multivariate Analysis	4	3	80	20	4
MST 204	Linear Model and Design of Experiment	4	3	80	20	4
MST 205	Sampling Theory	4	3	80	20	4
MST 216	Practical-II	12	3	-	100	4

3) OTHER FEATURES:

A) LIBRARY:

Reference and Textbooks, Journals and Periodicals, Reference Books for Advanced Books for Advanced studies.

B) SPECIFIC EQUIPMENTS IN LABORATORY:

20 Computers, LCD Projector, Visualizer, Smart board etc.

C) Laboratory softwares:

1. SAS
2. MINITAB
3. SPSS

MST 101: REAL ANALYSIS

Unit1: Set of real numbers, countable and uncountable sets, countability of rationals and uncountability of the interval $(0,1)$ Supremum and Infimum of bounded sets, limit point of a set, open, closed, dense and compact sets. Bolzano-Weierstrass and Heine-Borel Theorems (Statements only). Applications of the theorems. (12L+3T)

Unit2: Sequence of real numbers, convergence, divergence, Cauchy sequence, Convergence of bounded monotone sequence. Limit inferior and limit superior of the sequences. Series of numbers, tests for convergence (without proof) test for absolute convergence, convergence of sequences of non-negative terms. (12L+3T)

Unit3: Real valued function, continuous function, Uniform continuity of sequence of functions, Uniform convergence of series of functions with special emphasis on power series, radius of convergence. Riemann, Riemann-Stieltjes Integrals and their common properties. Integration by parts, Fundamental theorem on calculus, mean value theorem, their applications in finding functional of distributions. (12L+3T)

Unit4: Vector and Matrix differentiation, Maxima, minima of functions of several variables. Constrained maxima, minima, Lagrange's method, Taylor's theorem (without proof), implicit function theorem and their applications. Multiple integrals, Change of variables, Improper integrals, Applications in multivariate distributions. Theorem on differentiation under integral sign (without proof), Leibnitz rule (statement only) and applications. (12L+3T)

Recommended Books:

1. Malik S. C. & Arora S.(1991): Mathematical Analysis- Wiley Eastern Limited-IIInd edition.
2. Goldberg R. R. (1964): Methods of Real Analysis- Blais dell Publishing company, Newyork, U.S.A.
3. Bartle G.R.(1976):Element of Real Analysis-Wiley, 2nd edition.
4. Bartle G.R.& Sherbert D.R.(2000):Introduction to Real Analysis-John Wiley & Son Inc.
5. Royden (1988): Principles of Real Analysis-Mac millian.
6. Widder (1989): Advanced Calculus-Dover Publication.
7. Apostol (1985): Mathematical Analysis-Narosa Publishing House, T. M.

MST102: LINEAR ALGEBRA

Unit 1: Vector space, subspace, linear dependence and independence, basis, dimension of a vector space, example of vector spaces. Null space, Gram- Schmidt orthogonalisation process, Orthonormal basis, orthogonal projection of a vector, Linear transformations, algebra of matrices, row and column spaces of a matrix, elementary operations and elementary matrices, rank and inverse of a matrix, Null space and nullity, partitioned matrices. (12L+3T)

Unit 2: Permutation matrix, reducible/ irreducible matrix, primitive/ imprimitive matrix, idempotent matrix, Kronecker product, Generalized inverse, Moore-Penrose generalized inverse, Solution of a system of homogenous and non-homogenous linear equations, theorem related to existence of solution and examples. (12L+3T)

Unit3: Characteristic roots and vectors of a matrix, algebraic and geometric multiplicities of a characteristic root, right and left characteristic vectors, orthogonal property of characteristic vectors, Caley-Hamilton Theorem and its applications. (12L+3T)

Unit4: Spectral decomposition of a real symmetric matrix, singular value decomposition, Choleskey decomposition, real quadratic forms, reduction and classification, index and signature, extrema of a quadratic form, simultaneous reduction of two quadratic forms. (12L+3T)

Recommended Books:

1. Graybill, F.A (1961) An Introduction to Linear Statistical Models Vol 1, Mc Graw-Hill Book Company Inc.
2. Hadely G. (1962) Linear Algebra, Narosa Publishing House.
3. Harville D. (1997) Matrix Algebra From Statistics Perspective, Springer.
4. Rao A. R. and Bhimasankaram P. (2000), Linear Algebra, Second edition, Hindustan Book Agency.
5. Rao C. R.(2001) Linear Statistical Inference and Its Applications, Second Edition,Wiley.
6. Schott J. (2016) Matrix Analysis for Statistics, Third edition Wiley
7. Searl S. B. (2006) Matrix Algebra Useful for Statistics, Wiley

MST103: DISTRIBUTION THEORY

Unit1: Review of concept of random variable, c.d.f, characteristic properties of c.d.f., p.d.f., p.m.f., absolutely continuous and discrete distributions, mixtures of probability distributions, decomposition of mixture type c.d.f. into discrete and continuous c.d.f.'s., three parameter Weibull, m.g.f., p.g.f., quantiles and Transformations of univariate random variables, moments, Convolutions, compound distributions. Probability Integral transformation. (12L+3T)

Unit2: Moment inequalities:- Markov, Chebychev, Holder, Minkowski and Jensen inequalities with their applications. Basic inequality Liapunov's. Symmetric distributions, properties of symmetric distributions, non-regular families, location and scale families and examples. (12L+3T)

Unit 3: Random vectors, joint distributions, Independence, variance-covariance matrix, joint m.g.f., mixed moments, Conditional expectation and variances, Transformations of bivariate random variables, Bivariate Normal, Bivariate Exponential distributions. Bivariate Poisson distribution. Dirichlet distribution. Results related to generation of random variables from Multinomial, Bivariate Poisson, Bivariate exponential and Bivariate normal distributions.

(12L+3T)

Unit4: Statements and illustrations of theorems on distributions of linear and quadratic forms, Fisher Cochran theorem, non-central χ^2 , noncentral t and F distributions.

(12L+ 3T)

Recommended Books:

1. Rohatagi V. K. & Saleh A. K. Md. E.(2001):Introduction to Probability Theory and Mathematical Statistics-John Wiley and sons Inc.
2. Johnson N. L. & Kotz. S. (1996) : Distributions in Statistics Vol-I,II and III, John Wiley and Sons New york.
3. S. Kotz, N. Balakrishnan, N. L. Johnson: Continuous Multivariate Distributions- Second Edition, Wiley.
4. Casella & Berger (2002) : Statistical Inference -Duxbury advanced series. II nd edition
5. C. R. Rao (1995) Linear Statistical Inference and Its Applications (Wiley Eastern) Second Edition
6. Dasgupta, A. (2010) Fundamentals of Probability: A First Course (Springer)

MST104: ESTIMATION THEORY

Unit 1: Sufficiency principle, factorization theorem, minimal sufficiency, minimal sufficient partition ,construction of minimal sufficient statistics, minimal sufficient statistic for exponential family, power series family, curved exponential family, Pitman family. Completeness, bounded completeness, ancillary statistics, Basu's theorem and applications.

(12L+3T)

Unit 2: Problem of point estimation, unbiased estimators, minimum variance unbiased estimator, Rao- Blackwell theorem and Lehmann-Scheffe theorem and their uses. Necessary and sufficient condition for MVUE and their applications. Fisher information and information matrix, Cramer-Rao inequality, Chapman-Robinson bounds, Bhattacharya bounds, their applications.

(12L+3T)

Unit3: Method of maximum likelihood (MLE) and small sample properties of MLE, method of scoring and application to estimation in multinomial distribution. MLE in non-regular families. Other methods of estimation: method of moments, minimum Chi square. U-Statistics: one and two sample; U-Statistics theorem for one sample and two sample (statements only).

(12L+3T)

Unit4: The concept of prior distributions, various types of priors, non informative, Jeffrey's, least favorable prior, posterior distribution; Posterior distribution conjugate family and standard examples of such families. Bayes estimation under squared error and absolute error loss functions.

(12L+3T)

Recommended Books:

1. Rohatgi, V. K. and Saleh, A. K. MD. E. (2015). Introduction to Probability Theory and Mathematical Statistics-3rd Edition, John Wiley & sons.
2. Lehmann, E. L. (1983). Theory of Point Estimation- John Wiley & sons.
3. Rao, C. R. (1973). Linear Statistical Inference and its Applications, 2nd Edition, wiley.
4. Kale, B. K. and Muralidharan, K. (2015). Parametric Inference: An Introduction, Alpha Science International Ltd.
5. Mukhopadhyay, P. (2015). Mathematical Statistics, Books and Allied (p) Ltd.
6. Dudewicz, E.J. and Mishra, S. N.(1988). Modern Mathematical Statistics, John Wiley and Sons.
7. Casella, G., and Berger, R. L. (2001). Statistical Inference, 2nd Edition, Duxbury press.

MST 105: OPTIMIZATION TECHNIQUES

Unit-1: a) Linear programming problem (LPP): Theorems related to the development of Simplex algorithm, theorems related to a basic feasible solution ; Reduction of a feasible solution to a basic feasible solution, Improvement of a basic feasible solution, Existence of unbounded solution, Optimality conditions and other related theorems (statements only), Examples based on these theorems. Revised Simplex Method.

b) Artificial variable technique: Two phase method, redundancy. (12L+3T)

Unit-2: a) Concept of Duality, related theorems, complementary slackness property and development of dual simplex algorithm.

b) Sensitivity Analysis: Changes in the cost vector, requirement vector and non basic activity vector; addition of new variables and addition of new constraints. (12L+3T)

Unit-3: a) Theory of games: two person zero sum games, minimax and maximin principles, Saddle point, mixed strategies; rules of dominance, solution of 2 x 2 game by algebraic method, Graphical method, Reduction of the game problem as LPP, Minimax and Maximin theorem.

b) Dynamic Programming: The Recursion Equation Approach, Computational Procedure, Characteristics of Dynamic Programming, Solution of L.P.P. by Dynamic Programming.

(12L+3T)

Unit-4: a) Integer Linear Programming Problem (ILPP): The concept of cutting plane, cutting plane method for all ILPP and mixed ILPP, Branch and Bound method.

b) Quadratic programming: Kuhn-Tucker conditions, methods due to Beale, Wolfe. (12L+3T)

Recommended Books:

- 1) Hadley G. (1969): Linear Programming, Addison Wesley.
- 2) Taha H. A. (1971): Operation Research An Introduction- Macmillan

- 3) Kanti Swaroop & Gupta M. M. (1985): Operations Research, Sultan Chand & P. Gupta
- 4) D. S. Hira (2010): Operation Research, Sultan Chand & Co. Ltd.
- 5) J. K. Sharma. (2003): Operation Research Theory and Applications. Macmillan.

MST116: PRACTICAL-I

1. Linear dependence and independence of vectors.
2. Gram-Schmidt orthogonalisation method.
3. Solving systems of equations.
4. Inverse and g-inverse of a matrix.
5. Applications of Caley-Hamilton theorem.
6. Inverse of a Partitioned matrix.
7. Characteristics roots and vectors and their applications.
8. Classifications and reduction of quadratic forms.
9. Sketching of d.f.s.
10. Model sampling from univariate and bivariate.
11. Construction of UMVUE.
12. Methods of Estimation: MML and MLE.
13. Methods of Scoring.
14. Practicals on Bayesian inference.
15. Solution to LPP using simplex method.
16. Revised Simplex method and Dual Simplex Method.
17. Game Theory.
18. Quadratic programming
19. Integer programming.
20. Dynamic Programming

MST201: PROBABILITY THEORY

Unit1: Classes of sets: Sequence of sets: limsup, liminf and limit of sequence of sets field, σ -field, σ -field generated by a class of sets, Borel σ - field. Probability measure, Probability space, properties of a probability measure, continuity, mixture of probability measures. Lebesgue and Lebesgue-Stieltjes measures on \mathbb{R} . Independence of events. (12L+3T)

Unit 2: Measurable function, random variable, distribution function of a random variable, simple random variable, elementary random variable, liminf, limsup and limit of sequence of random variables. Method of obtaining a random variable as a limit of sequence of simple random variables. Integration of a measurable function with respect to a measure, expectation of a random variable, independence. Characteristic function, simple properties. Inversion theorem and uniqueness property (Statement only). (12L+3T)

Unit 3: Monotone convergence theorem, Fatous Lemma, Dominated Convergence theorem, Borel- Cantelli Lemma, (Statements only), and their applications. Convergence of sequence of random variables, Convergence indistribution, Almost sure convergence, a characterizing property, convergence in probability, uniqueness of limit, Yule Slutsky results and preservation under continuous transform. Convergence in r^{th} mean, interrelationships (Statements only), their illustration with examples. (12L+3T)

Unit4: Weak and Strong laws of large numbers, Kolmogorov's three series theorem for almost sure convergence (Statement only), Liapoune's, Lindeberg- Feller Theorems on CLT (Statement only). Applications of the above results. (12L+3T)

Recommended Books:

1. Bhat B. R. (1981): Modern Probability Theory—IIIrd edition: New age international (P) limited
2. Alan Karr,(1993):Probability Theory—Springer Verlag.
- 3.Billingsley P.(1986):Probability &Measure—John Wiley and sons
4. Athreya K. B. and Lahiri S. (2006). Probability Theory vol41, Trim series, (Hindustan Book Agency).
5. Feller, W. (1969).Introduction to Probability and its Applications vol.II (Wiley Eastern Ltd.)
6. Loeve, M.(1978). Probability Theory (Springer Verlag).

MST202: THEORY OF TESTING OF HYPOTHESES

Unit1: Problem of testing of Hypothesis, Simple and composite hypotheses. Randomized and non-randomized tests, Most powerful test, Neyman-Pearson Lemma and its applications. Determination of minimum sample size to achieve the desired strengths. Monotone likelihood ratio property, UMP test, power function of a test, existence of UMP. Tests for one-sided alternatives. Concept of p-value. (12L+3T)

Unit 2: UMP tests for two sided alternatives examples, their existence and non- existence. Generalized Neyman Pearson lemma, unbiased test, UMPU test and their existence in the case of exponential families (Statements of the theorems only). Similar tests, test with Neyman structure. (12L+3T)

Unit3: Problem of confidence intervals, relation with testing of hypotheses problem, shortest length confidence intervals, UMA and UMAU confidence intervals. (12L+3T)

Unit4: Likelihood ratio test and its application to standard distribution. Goodness of fit tests based on Chi-square distribution and application to contingency tables. Spearman's Rank Correlation Test; Kendall's Rank Correlation Test; Kruskal-Wallis Test; Fridman's Two-way analysis of variance by ranks. (12L+3T)

Recommended Books:

1. Rohatgi, V. K. and Saleh, A. K. MD. E. (2015). Introduction to Probability Theory and Mathematical Statistics-3rd Edition, John Wiley & sons.
2. Kale, B. K. and Muralidharan, K. (2015). Parametric Inference: An Introduction, Alpha Science International Ltd.
3. Dudewicz, E. J. and Mishra, S. N. (1988). Modern Mathematical Statistics, John Wiley and Sons.
4. Lehman, E. L. (1987). Theory of testing of hypotheses. Students Edition.
5. Ferguson, T. S. (1967). Mathematical Statistics: A decision theoretical approach. Academic press.
6. Zacks, S. (1971). Theory of Statistical Inference, John Wiley and Sons, New York.
7. Randles, R. H. and Wolfe, D. A. (1979). Introduction to theory of nonparametric Statistics, Wiley.
8. Gibbons J. D. and Chakraborti S. (2010) Nonparametric Statistical Inference, Fifth Edition, CRC Press.

MST203: MULTIVARIATE ANALYSIS

Unit1: Exploratory multivariate data analysis, sample mean vector, sample dispersion matrix, correlation matrix, graphical representation, means, variances, covariances, Partial and multiple correlation coefficients. Correlations of linear transforms. Multivariate normal distribution, two definitions and their equivalence, singular and nonsingular normal distribution, characteristic function, moments, marginal and conditional distributions. (12L+3T)

Unit2: Maximum likelihood estimators of the parameters of the multivariate normal distribution and their sampling distributions. Hotelling's T^2 Statistic and its null distribution. Applications of T^2 statistics and its relationship with Mahalanobis' D^2 statistic. Confidence region for the mean vector, Wishart matrix and its distribution, properties of Wishart distribution, distribution of generalized variance. (12L+3T)

Unit 3: Discrimination and classification. Fisher's discriminant function and likelihood ratio procedure, minimum ECM rule, Rao's U statistics and its use in tests associated with

discriminant function, classification with three populations. Cluster analysis, Hierarchical methods: Single, Complete, average linkage method and non-hierarchical clustering method-k-means clustering. (12L+3T)

Unit4: Canonical correlation analysis, Introduction to principal component analysis and related results, Introduction to factor analysis and estimation. (12L+3T)

Recommended Books:

1. Kshirsagar A.M.(1972):Multivariate Analysis. Marcel- Dekker.
2. Johnson, R. A. and Wichern. D.W(2002):Applied multivariate Analysis. 5th Ad. Prentice–Hall.
- 3.Anderson T. W. (1984): An Introduction to Multivariate Statistical Analysis 2nd Ed. John Wiley.
4. Morrison D. F. (1976): Multivariate Statistical Methods Mc Graw-Hill.

MST204: LINEAR MODELS AND DESIGN OF EXPERIMENTS

Unit1: General linear model: definition, assumptions, concept of estimability, least squares estimation, BLUE, estimation space, error space, Gauss Markov theorem, variances and covariances of BLUEs, Distribution of quadratic forms for normal variables: related theorems (without proof), Tests of hypotheses in general linear models. (12L+3T)

Unit2:Analysis of variance: one way classification, two way classification without interaction and with interaction with equal number of observations per cell, Estimation and tests of hypotheses, multiple comparison procedures: Three types of errors, Tukey, Sheffe and Bonferroni procedure. (12L+3T)

Unit3: Analysis of Covariance: estimation of parameters, related tests of hypothesis. General theory and application of one way and two way setup. (12L+3T)

Unit4: Two way classification with unequal number of observations per cell without interaction model, estimable parametric functions and their BLUEs, tests of hypotheses, incomplete block design, concepts of connectedness, balancedness, and orthogonality, BIBD: Definition, properties and analysis, Symmetric BIBD. (12L+3T)

Recommended Books:

1. Kshirsagar A.M.(1983) Course in Linear Models, Marcel Dekker.
2. Joshi D. D.(1987) Linear Estimation and Analysis of Experiments, Wiley Eastern Ltd.
3. Das, M.N. and N.C. Giri(1986) Design and analysis of experiments, 2nd edition, New Age International(P)Limited Publishers.
4. Searle S.R.(1971)Linear Models, John Wiley & Sons. New York.
5. Chakravarti .M. C.(1962) Mathematics of Design of Experiments, Asia Publishing House,

Bombay.

6. Dey Aloke (2010) Incomplete block design, Hindustan Book Agency.

7. Dean A. M. and Voss D.(1999) Design and Analysis of Experiments, Springer.

MST 205: SAMPLING THEORY

Unit1: Review of concept Simple random sampling with replacement (SRSWR) and Simple random sampling without replacement (SRSWOR), results related to SRSWR and SRSWOR, estimation of sample size. Stratified sampling: Stratification, allocation and estimation problems, comparison with SRS, post stratification, construction of strata, deep stratification, method of collapsed strata, Review of concept of Systematic sampling: linear systematic sampling and circular systematic sampling, Comparison with SRS and Stratified sampling.

(12L+3T)

Unit 2: PPSWR methods: Cumulative total method, Lahiri's method related estimation Problems and PPSWOR methods and related estimation of a finite population mean (Horwitz-Thompson and Des Raj estimators for a general sample size and Murthy's estimator for a sample of size 2, Midzuno sampling, Rao-Hartley-Cochran sampling Strategy.

(12L+3T)

Unit 3: Use of supplementary information for estimation: ratio and regression estimators and their properties. Unbiased and almost unbiased ratio type estimators, Double sampling. Cluster sampling. Two –stage sampling with equal number of Second stage units, multistage-sampling. Stratification estimator, Multiphase sampling.

(12L+3T)

Unit4: Non-sampling errors: Response and non-response errors. Hansen–Hurwitz and Deming's model for the effect of call-backs. Random response techniques, dichotomous population, Warners model, MLE in Warners model, unrelated question model, polychotomous population: use of binary and vector response, binary response and unrelated questions, Multi attribute situations.

(12L+3T)

Recommended Books:

1. Parimal Mukhopadhyay(2008): Theory and methods of survey sampling – 2ndEdition, Prentice Hall of India private limited.
2. Sukhatme P.V., Sukhatme S.& Ashok C (1984): Sampling Theory of surveys and applications–Iowa university press and Indian society of agricultural statistics, New Delhi.
3. Chaudhuri and H. Stenger (2005): SurveySampling: Theory and Methods, 2nd edition, chapman and hall/CRC.
4. Des Raj and Chandhok. P. (1998): Sample Survey Theory - Narosa publication.
5. William G.Cochran. (1977): Sampling Techniques- IIIrd edition– John and Wilely sons Inc.
6. Murthy M.N. (1977): Sampling Theory of Methods -Statistical Publishing Society, Calcutta.
7. Singh D. and Chaudhary F.S(1986).Theory and Analysis of Sample Survey Designs, Wiley Eastern Limited.
8. Singh, S.(2003).Advance Sampling Theory and Applications (Volume I and II), Kluwer Academic Publishers.

MST 216-PRACTICAL-II:

1. Exploratory data analysis.
2. Application of Hotelling's T^2 statistics
3. Discriminant Analysis
4. Cluster Analysis
5. Principle component analysis and Factor Analysis.
6. MP, UMP, and UMPU Tests
7. Likelihood ratio tests.
8. Confidence Intervals.
9. Non-parametric Tests.
10. Linear Estimation: Estimation and Hypothesis testing.
11. ANOVA: One way and two way orthogonal data without interaction.
12. ANOVA: Two way orthogonal data with interaction.
13. Two way non-orthogonal data without interaction
14. Analysis of BIBD.
15. Analysis of general block design.
16. Basic sampling designs.
17. Ratio, regression, Horvitz-Thompson method of estimations.
18. Stratified, Systematic and cluster Sampling.
19. Multi-stage sampling
20. Non-sampling errors.
