SHIVAJI UNIVERSITY, KOLHAPUR

Revised Syllabi of M.Sc. Statistics (CBCS)

(To be implemented in the Department of Statistics, Shivaji University and in P.G. Centres in Affiliated Colleges) (w.e. form 2016-2017)

1) Title of the course: M. Sc. (Statistics)
   M. Sc. Statistics program has semester pattern and Choice Based Credit System. The program consists of 96 credits.
2) Eligibility: B. Sc. (Second class) with Statistics as principal subject.
3) Structure of the course
   The following table gives the scheme of Examination at M.Sc. Statistics according to the Revised Syllabus and pattern of Examination.

   Notations: A six-character code is given to each paper. In MST "M" stands for M.Sc. and "ST" stands for Statistics. The first digit following MST is Semester Number. The second digit "0" stands for the **core** theory paper, the digit "1" stands for a practical paper, the digit "2" stands for an **elective paper**/project work, digit "3" stands for an **open elective** and digit "4" stands for ability enhancement paper. 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)

Course Code	Title of the course	Instruction hrs/week	Duration of Exam (hrs)	Marks - Term End Exam	Marks- Internal Assessment	Credits
MST 101	Statistical Mathematics-I	4	3	80	20	4
MST 102	Statistical Mathematics-II	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	Statistical Computing	3	3	80	20	3
MST 146	Communication Skills	15 per sem	1	-	20	1
MST 117	Practical-I	12	3	-	100	4

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

Course Code	Title of the course	Instruction hrs/week	Duration of Exam	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 Models and Design of Experiments	4	3	80	20	4
MST 205	Sampling Theory	4	3	80	20	4
MST 216	Practical-I	12	3	-	100	4

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

Course Code	Title of the course	Instruction hrs/week	Duration of Exam	Marks - Term End Exam	Marks- Internal Assessment	Credits
MST 301	Asymptotic Inference	4	3	80	20	4
MST 302	Elementary Stochastic Processes	4	3	80	20	4
MST 303	Planning and Analysis of Industrial Experiments	4	3	80	20	4
MST 304	Design and Analysis of Sample Surveys	2	-	-	20	1
MST 316	Practical III	12	3	-	100	4
ANY TWO from the following electives						
MST-321	Reliability Theory	3	3	80	20	3
MST-322	Regression Analysis	3	3	80	20	3
MST-323	Statistical Genetics	3	3	80	20	3
MST-324	Demography	3	3	80	20	3
MST-325	Statistical	3	3	80	20	3

	Methods in Clinical Trials					
MST-326	Actuarial Statistics	3	3	80	20	3
MST-327	Dissertation	3	-	-	100	3
Any ONE open elective of one credit from the list given below:						

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

Course Code	Title of the course	Instruction hrs/week	Duration of Exam	Marks – Term End Exam	Marks-Internal Assessment	Credits
MST 401	Optimization Techniques	4	3	80	20	4
MST 416	Practical IV,Viva and Project Work	12	1.5	-	100	3
ANY FOUR from the following electives of 4 credits each						
MST-421	Generalized Linear models	4	3	80	20	4
MST-422	Survival Analysis	4	3	80	20	4
MST-423	Statistical Quality Control	4	3	80	20	4
MST-424	Time Series Analysis	4	3	80	20	4
MST-425	Statistical Ecology	4	3	80	20	4
MST-426	Econometrics	4	3	80	20	4
MST-427	Data Mining	4	3	80	20	4
Any ONE open elective of one credit from the list given below						

Open Electives for PG students:

Semester	Title of the course	Instruction hrs/week	Intake Capacity	Eligibility	Marks-Internal Assessment	Credits
ODD	An introduction to MINITAB	1	20	Science and Technology	20	1
ODD	An introduction to SPSS	1	20	All Faculties	20	1
ODD	An introduction to MATLAB	1	20	Mathematics, Electronics, Physics,	20	1

				Engineering		
ODD	An introduction to R software	1	20	Mathematics, MCA, Computer Science (Excluding Statistics)	20	1
ODD	Design of Experiments for Applied Research	1	20	All Faculties excluding Statistics)	20	1
EVEN	Design and Analysis of Laboratory Experiments	1	20	All Faculties (Excluding Statistics)	20	1
EVEN	MATLAB Programming	1	20	Mathematics, Electronics, Physics, Engineering	20	1
EVEN	An introduction to SAS	1	20	Mathematics	20	1
EVEN	Advanced course in MSExcel	1	20	All faculties	20	1

Note:-

a) Syllabi for some Elective courses have been given. Depending on need and demand, syllabi for other elective courses listed above or for new elective courses will be submitted for approval.
b) There shall be Continuous Internal Evaluation pattern in which internal examination will be for 20 marks, while University examination will be for 80 marks.

Nature of the Theory and Practical Question Paper at the M.Sc. Statistics course under the Semester Scheme:

1. **Nature of the theory question papers:-**

a) There shall be 7 questions each carrying 16 marks.
b) Question No.1 is compulsory. It consists of 8 questions for 2 marks each.
c) Students have to attempt any 4 questions from question No. 2 to 7.
d) Question No. 2 to 6 shall contain 2 to 4 sub-questions.
e) Question No. 7 shall contain 4 short note type questions, each carrying 4 marks.

2. **Practical Paper:-**

a) Semester I, II, III "Practical MST 116, MST 216, MST 316"

1. There shall be 20 marks for day-to-day performance and journal.
2. Examination (60): Practical Examinations will be conducted in the middle of the term and at the end of the term. Each exam will be of 3 hrs. duration carrying 60 marks. There shall be 8 questions each of 12 marks, of which a student has to attempt any 5 questions. The average of these two tests will be considered as final score out of 60.
3. Practical VIVA will be for 20 marks.

b) For Semester IV: MST-416

1. There shall be 10 marks for day-to-day performance and journal.
2. Examination (30): Practical Examinations will be conducted in the middle of the term and at the end of the term. Each exam will be of 1.5 hrs. (90 Min.) duration carrying 30 marks. There shall be 5 questions each of 10 marks, of which a student has to attempt any 3 questions. The average of these two tests will be considered as final score out of 30.
3. Practical VIVA will be for 20 marks.
4. Project work carries 40 marks. Project work consists of collecting and analyzing the data and preparing report. 20 marks are reserved for project based VIVA. Project report will be evaluated for 20 marks. The project work should be preferably based on field work. The budgetary provision for field work/industrial tour will be utilized to meet expenses towards the field work.

Detailed Syllabi of papers for Semester I:

MST 101: STATISTICAL MATHEMATICS – I

Unit 1: 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 these theorems.

(12 L + 3 T)

Unit 2: 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 absoluteconvergence, convergence of sequences of non-negative terms.

(12 L + 3 T)

Unit 3: 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-Steltjes Integrals and their common properties. Integration by parts, Fundamental theorem on calculus, mean value theorem, their applications in finding functional of distributions. (12 L + 3 T)

Unit 4: 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 inmultivariate distributions. Theorem on differentiation underintegral sign (without proof), Leibnitz rule (statement only) and applications.

(12 L + 3 T)

References:

1. Malik S. C. &Arora S.(1991): Mathematical Analysis- Wiley Eastern LimitedIInd edition.
2. Goldberg R. R. (1964): Methods of Real Analysis- Blaisdell Publishingcompany, New york, 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- JohnWiley & Son Inc.
5. Royden (1988): Principles of Real Analysis - Macmillian.
6. Widder (1989): Advanced Calculus - Dover Publication.
7. Apostol (1985): Mathematical Analysis - Narosa Publishing House, T.M.

MST 102: STATISTICAL MATHEMATICS – II

Unit 1: Vector space, subspace, linear dependence and independence, basis, dimension of a vector space, example of vector spaces. Null space, Gram-Schmidtorthogonalisation 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)

Unit 3: 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)

Unit 4: 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)

References:

1. Graybill, F.A (1961) An Introduction to Linear Statistical Models Vol 1, McGraw-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

MST 103: DISTRIBUTION THEORY

Unit 1: 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. Probability Integral transformation, Moment inequalities: Basic, Holder, Jensen, Tchebysheff . Symmetry of r.v.

(12L+3T)

Unit 2: Brief review of basic distribution theory: Standard discrete and continuous distributions, truncated distributions (Truncated Binomial, Truncated Poisson, Truncated Normal etc.), three parameter Weibull, m.g.f., p.g.f., quantiles and Transformations of univariate random variables, moments, m.g.f., p.g.f., Convolutions, compound distributions.

(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.

(12L+3T)

Unit 4: Order statistics: distribution of r-th order statistics, joint distribution of $r^{th}$and s^{th}order statistics ($r < s$) and their functions, distribution of spacings, normalized spacings with illustration to exponential case, distribution of sample range. Statements and illustrations of theorems on distributions of linear and quadratic forms, Fisher Cochran theorem, non-central χ^2, non central t and F distributions.
(12L+3T)

References:

1. RohatagiV.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, JohnWiley 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.IInd 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)

MST 104: 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, Chapmann-Robinson bounds, Bhattacharya bounds, their applications.

(12L+3T)

Unit 3: Method of maximum likelihood (MLE) and small sample properties of MLE, methodof 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)

Unit 4: The concept of prior distributions, various types of priors, noninformative, 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)

References

1. Rohatgi, V.K. and Saleh, A. K. MD. E. (2015).Introduction to Probability Theory and MathematicalStatistics -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. andMishra,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: STATISTICAL COMPUTING

Unit 1: MSEXCEL: Introduction to MSEXCEL. Data manipulation using EXCEL: sort and filter, find and replace, text to columns. Built-in mathematical and statistical functions for obtaining descriptive statistic, computing PMF/PDF, CDF and quantiles of the well known distributions, rand function, Logical functions: if, true, false, and, or, not. Lookup functions: choose, hlookup, vlookup, index. Formula Errors.Basic charts. Excel add-ins: analysis tool pack. Pivot tables and charts. Introduction to Excel macro. R-software: Introduction to R, data types, data input, data import and export, built in functions for descriptive statistics, random sampling and computation of pdf, cdf and quantiles of well known distribution. Matrix algebra, graphical procedures, frequencies and cross tabulation, built in functions: lm, t.test, prop.test, wilcox.test, ks.test, var.test, chisq.test, aov. Control statements. Elementary programming, user defined functions, R-packages.

(12L+3T)

Unit 2: Concept of simulation. Concept of random number generator, true random number and pseudo random number generators, requisites of a good random number generator.Congruential method of generating uniform random numbers. Algorithms for generating random numbers from well known univariate discrete and continuous distributions, generating random vectors from multinomial, bivariate normal, and bivariate exponential distributions, generating random numbers from mixture of distributions (related results without proofs). Acceptance-Rejection Technique.Use of random numbers to evaluate integrals, to estimate event probabilities and to find expected value of random variables

(12L+3T)

Unit 3: Resampling techniques: Bootstrap methods, estimation of bias and standard errors, estimation of sampling distribution, confidence intervals. Jackknife method: estimation of bias and standard errors, bias reduction method. Solutions to system of linear equation: Jacobi and Gauss Seidel methods with convergence analysis. Solutions to nonlinear equation: Newton Raphson method, Regula falsi method, bisection method. Numerical integration: quadrature formula, trapezoidal rule and Simpson's rule for single integral.

(12L+3T)

References

1. Atkinson K. E. (1989): An Introduction to Numerical Analysis. (Wiley)
2. Devroye L. (1986) : Non- Uniform Random Variate Generation. (Springer- Verlag New York)
3. Efron B. and Tibshirani. R. J. (1994): An Introduction to the Bootstrap. (Chapman and Hall)
4. Morgan B. J. T.(1984) : Elements of Simulation. (Chapman and Hall)
5. Robert C. P. and Casella G. (1999): Monte carlo Statistical Methods. (Springer-verlag New York, Inc.)

6. Ross. S. M. (2006): Simulation. (Academic Press Inc)
7. Rubinstein, R. Y. (1998) Modern Simulation and Modeling. (Wiley Series in Probability and Statistics)
8. William J., Kennedy, James E. Gentle. (1980): Statistical Computing. (Marcel Dekker)

MST 146 - COMMUNICATION SKILLS

Unit 1: Parts of Speech, Phrases, Basic Clause Types, Clauses, Subordination and Coordination, Cohesion. Information, Reality and Belief. Permission and Obligation.

(Section B from a book : *A Communicative Grammar of English*)

(12 L+ 3 T)

References:

1. Leech, Geoffrey (1982) *English Grammar for Today*, Palgrave
2. Leech, Geoffrey (2003) *A Communicative Grammar of English*, Third Edition, Pearson.
3. Quirk, Randolph and Greenbaum Sidney (1973) *University Grammar of English*, Longman: London
4. Halliday, M.A.K. and Ruquyia Husan (1976) *Cohesion in English*, Longman: London

MST 117: 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.fs.

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 Baysian inference.

15 – 20. Programming assignments on MST–105 Course. (Software to be used : R/MINITAB/MATLAB/SAS/SYSTAT depends on availability)

Detailed Syllabi of papers for Semester II:

MST 201: PROBABILITY THEORY

Unit 1: 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 - Steltjes measures on 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 in distribution, 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)

Unit 4: Weak and Strong laws of large numbers, Kolmogorov's three series theorem for almost sure convergence (Statement only), Liaponove's, Lindeberg-Feller Theorems on CLT (Statement only). Applications of the above results.

(12L+3T)

References:

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. AthreyaK. B. and Lahiri S. (2006). Probability Theory vol 41, 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). Fourth edition.

MST 202: THEORY OF TESTING OF HYPOTHESES

Unit 1: 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)

Unit 3: Problem of confidence intervals, relation with testing of hypotheses problem, UMA and UMAU confidence intervals, shortest length confidence intervals. Likelihood ratio test and its application to standard distribution.

(12L+3T)

Unit 4: Goodness of fit tests based on Chi-square distribution and application to contingency tables.Non-parametric tests, One and two sample problem; one sample tests: Sign test, Wilcoxon Signed-Rank test. Two sample tests: Wald-Wolfowitz Runs test, Mann-Whiteny U test, Median test, Kolmogorov Smirnov test. Spearman's Rank Correlation Test; Kendall's Rank Correlation Test; Kruskal-Wallis Test; Fridman's Two-way analysis of variance by ranks.

(12L+3T)

References:

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, NewYork.
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.

MST 203: MULTIVARIATE ANALYSIS

Unit 1: 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)

Unit 2: 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, Heirarchicalmethods : Single, Complete, average linkage method and non-heirarchical clustering method-k-means clustering.

(12L+3T)

Unit 4: Canonical correlation analysis,Introduction to principal component analysisand related results, Introduction to factor analysis and estimation.

(12L+3T)

Reference:-

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 Wiely.
4. Morrison D.F. (1976) : Multivariate Statistical Methods McGraw-Hill.

MST 204: LINEAR MODELS AND DESIGN OF EXPERIMENTS

Unit 1: General linear model: definition, assumptions, concept of estimability, least squares estimation, BLUE, estimation space, error space, Guass 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)

Unit 2: 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.

(2L+3T)

Unit 3: Analysis of Covariance: estimation of parameters, related tests of hypothesis. General theory and application of one way and two way set up, introduction to mixed and random effects models.

(12L+3T)

Unit 4: 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)

References:

1. Kshirsagar A. M.(1983) Course in Linear Models, Marcel Dekker.
2. Joshi D. D.(1987) Linear Estimation and Analysis of Experiments, Wiley Estern 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. DeyAloke (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

Unit 1: Review of concept of population and sample, Need for Sampling, sampling frame, census and sample surveys, basic concepts in sampling and designing of large-scale surveys design, sampling scheme and sampling strategy. 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.

(12L+3T)

Unit 2: Systematic sampling: linear systematic sampling and circular systematic sampling, Comparison with SRS and Stratified sampling. 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.

(12L+3T)

Unit 4: Non - sampling errors: Response and non- response errors. Hansen –Horwitz and Demings 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, Multiattribute situations.

(12L+3T)

References

1. Parimal Mukhopadhyay (2008): Theory and methods of survey sampling – 2nd Edition, 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): Survey Sampling: 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 Wieleysons 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. M.P. 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.

Detailed Syllabi of papers for Semester III:

MST 301: ASYMPTOTIC INFERENCE

Unit 1: Consistency of an estimator, weak and strong consistency, joint and marginal consistency, invariance property under continuous transformations, methods of constructing consistent estimators, asymptotic relative efficiency. Consistent and Asymptotic Normal (CAN) Estimators: Definition of CAN estimator for real and vector valued parameters, invariance of CAN property under non-vanishing differentiable transformation. Methods of constructing CAN estimators: Method of Moments, method of percentiles, comparison of CAN estimators.

(12+ 3T)

Unit 2: CAN and BAN estimators in one parameter and multi-parameter exponential family of distributions, BAN estimators, super efficient estimators, Crammer regularity conditions, Cramer – Huzurbazar results.

(12L+3T)

Unit 3: Variance stabilizing transformations; their existence; their applications in obtaining large sample tests and estimators. Asymptotic Confidence Intervals based on CAN estimators and based on VST, Asymptotic Confidence regions in multi-parameter families.

(12L+3T)

Unit 4: Likelihood ratio test and its asymptotic distribution, Wald test, Rao's Score test, Pearson Chi-square test for goodness of fit, Bartlett's test for homogeneity of variances. Consistent test, comparison of tests: asymptotic relative efficiency of tests (Pitman and Bahadur efficiency). Performance evaluation (based on simulation) of asymptotic tests and confidence intervals.

(12 L +3 T)

References:

1) Kale B.K. (1999): A first course on parametric inference, Narosa Pub.
2) Zacks S. (1971): Theory of statistical inference, Wiley & Sons inc.
3) Rohatagi V.K. and Saleh A. K. Md. E.(2001) : Introduction to Probability Theory and Mathematical Statistics- John Wiley and sons Inc.
4) Ferguson, T.S. (1996): A Course in Large Sample Theory. Chapman and Hall
5) Lehmann E L (1999): Elements of Large Sample Theory, Springer.
6)DasGupta A. (2008): Asymptotic Theory of Statistics and Probability, Springer Texts in Statistics.

MST 302: ELEMENTARY STOCHASTIC PROCESSES.

Unit 1: Definition of stochastic process, classification of stochastic processes according to state space and time domain, finite dimensional distributions. Examples of various stochastic processes. Definition of Markov chain. Examples of Markov chains, Formulation of Markov chain models, initial distribution, transition probability matrix, Chapman-Kolmogorov equations, calculation of n-step transition probabilities. Simulation of Markov Chain.

(12L +3T)

Unit 2: Classification of states, irreducible Markov chain, period of the state, random walk and gambler's ruin problem, first entrance theorem, first passage time distribution. Long-Run proportions and limiting probabilities, relation with mean recurrence time, stationary distribution.

(12L +3T)

Unit 3: Discrete state space continuous time Markov chain, Poisson process and related results. Birth and death processes and associated cases. Renewal and delayed renewal processes, related theorems, key renewal theorem (Without proof) and its application. Simulation of Poisson process and discrete state space Markov processes.

(12L +3T)

Unit 4: Galton-Watson Binaymi Branching process. Generating functions and its properties, moments. Probability of ultimate extinction. Distribution of population size and association results. Simulation of branching process. Basic elements of Queuing model. Steady state probabilities and various average characteristics for the models: M/M/1, M/M/1 with balking, M/M/c and M/G/1.

(12 L+ 3 T)

References:

1. Bhat B. R. (2000). Stochastic Models: Analysis and Applications, (New Age International)
2. Cinlar E. (2013): Introduction to Stochastic Process. (Courier Corporation)
3. Feller W.(2008): An Introduction to Probability Theory and Its Applications. (Wiley)
4. Hoel P. G., Port S. C. and Stone C. J. (1987): Introduction to Stochastic Processes. (Waveland Press)
5. Karlin S. and Taylor H. M. (1968): A First Course in Stochastic Process. (Academic Press)
6. Medhi J. (2009): Stochastic Process, (New Age International Publications)
7. Ross S. (1996): Stochastic Processes. (Wiley)
8. Ross S. (2014): Introduction to Probability Models. (Academic Press)
9. Taylor H. M. and Karlin S. (2014): An Introduction to Stochastic Modeling (Academic Press)

MST-303 PLANNING AND ANALYSIS OF INDUSTRIAL EXPERIMENTS

Unit-1: 2^k factorial Experiments: Concepts of main effects, interaction, their graphical representation, Analysis of full 2^k replicated and unreplicated factorial designs. Concept of Confounding: Total and partial confounding, construction and analysis confounded design.

(12L+3T)

Unit-2: 3^k factorial Experiments: Concepts of main effects, interaction, their graphical representation, linear and quadratic components, Analysis of full 3^k replicated and unreplicated factorial designs. Confounding: construction and analysis confounded design, Factorials with mixed levels.

(12L+3T)

Unit-3: Fractional replication for symmetric factorials, concept of generator, defining contrasts, aliasing, resolution and minimum aberration, construction and analysis of 2^{k-p} and 3^{k-p} fractional designs, Central composite designs.

(12L+3T)

Unit-4: Response surface experiments: linear and quadratic model, test for curvature, stationary point, central ridge systems, Rotatability, Multiple responses.

Taguchi methods: Concept of noise and control factors, inner and outer arrays, concept of loss function, S/N ratio, orthogonal arrays, linear graphs, interaction tables, ANOVA.

(12L+3T)

References:

1. Montgomery D.C. (2013): *Design and Analysis of Experiments*, 8th edition, Wiley India Pvt Ltd.
2. Phadke, M. S.(1989). *Quality Engineering using Robust Design*, Prentice-Hall.
3. Voss, D., Dean, A., and Dean, A.(1999).*Design and Analysis of Experiments*, Springer-verlagGmbh.
4. Wu, C. F., Hamada M. S.(2000). *Experiments : Planning, Analysis and Parameter Design Optimization*, 2nd edition, John Wiley & Sons.

MST-304 DESIGN AND ANALYSIS OF SAMPLE SURVEY

A review of sampling techniques, sample size determination, design of questionnaire, properties of good questionnaire. Planning of data collection. Steps involved in data analysis using statistical software: data coding, feeding, data cleaning and validation, missing data handling, descriptive and inferential statistical analysis. A case study on the above concepts. (Evaluation of this course is based on a case study report)

References:

1. Parimal Mukhopadhyay (2008): Theory and methods of survey sampling – 2nd Edition, 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.

MST-321 RELIABILITY THEORY

Unit-1: Structure function, dual of a structure, cuts and paths, components & systems, coherent systems, redundancy, Pivotal decomposition, Associated random variables and their properties. Birnbaum's measure of structural importance. Reliability concepts and measures, reliability of coherent systems, bounds on system reliability, Modular decomposition.

(12L+3T)

Unit-2: Life time distributions, survival functions, failure rate function, cumulative hazard function, residual life time, survival function of residual life time, mean residual life time, Computation of these functions for Common life time distributions: exponential, Weibull, Gamma, Makeham, Pareto, Rayleigh.

(12L+3T)

Unit-3: Notion of ageing: IFR, DFR, IFRA, DFRA, DMRL, NBU, NWU, NBUE, NWUE classes, ageing properties of common life time distributions, closure properties under formation of coherent structures, convolutions and mixtures of these classes. Damage model, cumulative damage model, univariate shock models and life distributions arising from shock models, bivariate exponential distribution.

(12L+3T)

References:

1) Barlow R.E. and Proschan F. (1975): Statistical Theory of Reliability & Life testing, Holt, Reinhart and Winston.
2) Lawless J.F.(1982): Statistical Models & Methods of Life Tome Data, John Wiley.
3) Miller R.C. (1981): Survival Analysis. John Wiley
4) Bain L.J (1978): Statistical Analysis of Reliability & Life Testing, Models, *Marcel Dekker.*
5) Martz H.F. and Waller R.A (1982): Bayesian Reliability Analysis*, John Wiley.*

MST 322: REGRESSION ANALYSIS.

Unit-1: Multiple regression model, Least squares estimate (LSE), Properties of LSE, Hypothesis testing, confidence and prediction intervals, General linear hypothesis testing. Dummy variables and their use in regression analysis. Residuals and their properties, residual diagnostics. Transformation of Variables: VST and Box-Cox power transformation.

(12L+3T)

Unit-2: Variable Selection Procedures: R-square, adjusted R-square, Mallows' Cp, forward, backward selection methods, AIC, BIC. Multicollinearity: Consequences, detection and remedies, ridge regression. Nonlinear regression models: Non linear least squares, Transformation to a linear model, Parameter estimation in a non linear system, Statistical inference in non linear regression.

(12L+3T)

Unit-3: Robust Regression: Influential observations, leverage, outliers, methods of detection of outliers and influential observations, estimation in the presence of outliers: M-estimator, Huber loss function, breakdown point, influence function, efficiency, Asymptotic distribution of M-estimator (Statement only), Mallows' class of estimators.

(12L+3T)

References

1. Draper N.R. and Smith, H. (1998): Applied Regression Analysis. 3rd ed Wiley
2. Wiesberg, S. (1985): Applied Linear Regression, Wiley.
3. Kutner, Neter, Nachtsheim and Wasserman (2003): Applied Linear Regression Models, 4th Edition, McGraw-Hill
4. Montgomery, D.C., Peck, E.A.,and Vining, G.(2012): Introduction to Linear Regression Analysis, 5th Ed . Wiley
5. Cook R.D. &WiesbergS.(1982): Residuals and Influence in Regression. Chapman and Hall.
6. Birkes, D and Dodge, Y. (1993). Alternative methods of regression, John Wiley & Sons.
7. Huber, P. J. and Ronchetti, E. M (2011) Robust Statistics, Wiley, 2nd Edition.
8. Seber, G. A., Wild, C. J. (2003). Non linear Regression, Wiley.

MST -325 STATISTICAL METHODS IN CLINICAL TRIALS

Unit-1: Introduction to clinical trials: the need and ethics of clinical trials, bias and random error in clinical studies, conduct of clinical trials, overview of Phase I-IV trials, multi-center trials. Data management: data definitions, case report forms, database design, data collection systems for good clinical practice.

(12L+3T)

Unit-2: Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. Longitudinal designs, review of factorial designs, objectives and endpoints of clinical trials, design of Phase I trials, design of single-stage and multi-stage Phase II trials, design and monitoring of Phase III trials with sequential stopping, design of bioequivalence trials.

(12L+3T)

Unit-3: Reporting and analysis: analysis of categorical outcomes from Phase I - III trials, analysis of survival data from clinical trials.

(12L+3T)

References:

1. S. Piantadosi (1997). Clinical Trials: A Methodologic Perspective, Wiley and Sons.
2. C. Jennison and B. W. Turnbull (1999): Group Sequential Methods with Applications to Clinical Trials, CRC Press.
3. L. M. Friedman, C. Furburg, D. L. Demets (1998). Fundamentals of Clinical Trials,SpringerVerlag.
4. J. L. Fleiss (1989). The Design and Analysis of Clinical Experiments.Wiley and Sons.

MST -326 ACTUARIAL STATISTICS

Unit-1: Introduction to Insurance Business, Insurance and utility theory, Risk models for Insurance: Individual and aggregate Risk models for short term, Distribution of aggregate claims, compound Poisson distribution and its applications. Survival function and Life tables: Survival function, Distribution function, Density functions and Force of mortality. Time-until-death random variable and Curtate-future lifetime random variable.

(12L+3T)

Unit-2: Life tables, Select and ultimate life tables. Assumptions for fractional ages and some analytical laws of mortality.Life Insurance: Principles of compound interest: Nominal and effective rates of interest and force of interest and discount, compound interest, accumulation factor, continuous compounding. Insurance payable at the moment of death and at the end of the year of death, level benefit insurance, Whole life insurance, endowment insurance, deferred insurance and varying benefit insurance. Recursion equations and commutation functions.

(12L+3T)

Unit-3: Annuities: Annuities certain, Continuous and Discrete life annuities. Life annuities with monthly payments and approtionable annuities. Recursion equations. Net premium: Fully continuous and discrete premiums, True monthly payment premiums, apportionable premiums and accumulation type benefits. Insurance model including expenses.

(12L+3T)

References:

1. Robin Cunningham, Thomas N. Herzog, Richard L. Models for Quantifying Risk, 4th Edition, ACTEX Publications, 2011.
2. Browers, Newton L et al., Actuarial Mathematics 2nd . Society of Actuaries, 1997.
3. Dickson, David C. M., Hardy, Mary R. and Waters, Howard R., Actuarial Mathematics for life contingent risks, International series on actuarial science, Cambridge 2009.
4. Deshmukh S. R., An Introduction to Actuarial Statistics, University Press, 2009
5. Narang, Uma, Insurance Industry in India: Features, Reforms and Outlook, New Century Publications

MST 316- PRACTICAL –III

1. Construction of Consistent/CAN Estimators
2. Confidence interval based on CAN and VST.
3. Realization from a Markov Chain.
4. Classification of t.p.m and computation of n-step probability matrix.
5. Classification of State: Computations of absorption probabilities.
6. Stationary distribution and recurrence times.
7. Realization from discrete state space Markov Processes and related estimation.
8. Analysis of full replicated unconfounded 2^n and 3^2 factorial experiments.
9. Analysis of single replicated 2^n and 3^2 factorial experiments.
10. Analysis of confounded 2^n and 3^2 factorial experiments: total and partial confounding.
11. Analysis of fractional factorial 2^{n-k} and 3^{n-k} factorial experiments.
12. Analysis of response surface 1^{st} and 2^{nd} order experiments.

13 - Onwards: At least three practical each on the optional courses.

Detailed Syllabi of papers for Semester IV:

MST 401: OPTIMIZATION TECHNIQUES

Unit 1: Convex Sets and Functions: Convex sets, supporting and separating hyperplanes, convex polyhedra and polytope, extreme points, convex functions. Linear programming problem (LPP): Definition and applications, methods of solving LPP: Graphical method, Simplex method, 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. Artificial variable technique: Two phase method, Big M method, degeneracy.

(12L+3T)

Unit 2: Concept of Duality, related theorems, complementary slackness property and development of dual simplex algorithm. 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: Integer Linear Programming Problem (ILPP): The concept of cutting plane, cutting plane method for all ILPP and mixed ILLP, Branch and Bound method.
Quadratic programming: Kuhn-Tucker conditions, methods due to Beale, Wolfe.

(12L+3T)

Unit 4: 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.
Dynamic Programming: The Recursion Equation Approach, Computational Procedure, Characteristics of Dynamic Programming, Solution of L.P.P. by Dynamic Programming.

(12L+3T)

References:

1) Hadley G.(1969): Linear Programming, Addison Wesley
2) Taha H. A. (1971): Operation Research: An Introduction, Macmillan N.Y.
3) Kanti Swaroop & Gupta M. M.(1985): Operations Research, Sultan Chand & Co. ltd.
4) P.Gupta & D.S.Hira(2010): Operation Research, Sultan Chand & Co. ltd.
5) J. K. Sharma. (2003): Operation Research: Theory and Applications. Macmillan.

MST 421: GENERALIZED LINEAR MODELS

Unit-1: Generalized linear models: concept of generalized linear model, Link function, ML estimation, Quasi-likelihood estimation, large sample tests about parameters, goodness of fit, analysis of deviance. Residual analysis, types of residuals: raw, Pearson, deviance, Anscombe, quantile; residual plots. Variable selection: AIC and BIC.

(12L+3T)

Unit-2: Logistic regression: logit, probit and cloglog model for dichotomous data with single and multiple explanatory variables, ML estimation, large sample tests about parameters. Hosmer-Lemeshow test, ROC curve. Multilevel logistic regression, Logistic regression for Nominal response: Baseline Category model and ordinal response: Proportional odds model.

(12L+3T)

Unit-3: Poisson regression: ML and Quasi-likelihood estimation of parameters, testing significance of coefficients, goodness of fit, power family of link functions, over dispersion: Types, causes and remedies. Negative Binomial regression: NB-2 model.

(12L+3T)

Unit-4: Generalized linear mixed models (GLMM): Structure of the model, consequences of having random effects, estimation by maximum likelihood, marginal versus conditional models, estimation by generalized estimating equations and conditional likelihood, tests of hypothesis: LRT, asymptotic variance, Wald and score test.

(12L+3T)

References:

1. Hosmer D.W. and Lemeshow S. (2000): Applied Logistic regression, 2nd ED. Wiley New York.
2. Agresti A. (1990) : Categorical Data Analysis. Wiley , New York.
3. R. Christensen (1997) Log-Linear Models and Logistic Regression, Springer. New York.
4. Hilbe, J. (2011): Negative Binomial regression, Cambridge University, Press, 2nd Edition.
5. McCulloch, C. E., & Searle, S. R. (2003). Generalized, linear, and mixed models, Wiley series in probability and statistics, New York.

MST 422 SURVIVAL ANALYSIS

Unit 1: Concept of censoring, various types of censoring, type-I, type-II, random censoring, progressive censoring. Writing likelihood function under all these censoring schemes, estimation and testing of parameters under above types of censoring.

(12L+3T)

Unit 2: Estimation of survival function: Actuarial Estimator, Kaplan Meier product limit estimator, properties: self-consistency, and asymptotic normality, redistribution to the right algorithm. Nelson Aalen estimator, estimation of the mean of the distribution. Concept of TTT Transform and its applications. Test for exponentiality against alternatives IFRA, NBU and NBUE.

(12L+3T)

Unit 3: Two-sample problem: Gehen test, Log rank test, Mantel Haenszel test. Competing risk models, parametric and nonparametric inference for this model.

(12L+3T)

Unit 4: Semi parametric regression for failure rate – Cox's proportional hazards model with one and several covariates, related estimation and test procedures. Introduction to accelerated time models: Linear rank tests, Least squares, Miller, Buckley-James and Koul-Susara Van-ryzin estimators.

(12L+3T)

References:

1. Bain L. J. (1978): Statistical Analysis of Reliability and Life Testing Models. (Marcel Dekker)
2. Barlow R. E. and Proschan F. (1975): Statistical theory of Reliability and Life testing: Probability Models. (Holt, Rinehart and Winston Inc.)
3. Barlow R. E. and Proschan F. (1996): Mathematical Theory of Reliability. (John Wiley)
4. Cox D. R. and Oakes D. (1984). Analysis of Survival Data, Chapman and Hall.
5. Crowder M. J. (2001): Classical Competing Risks. (CRC Press).
6. Deshpande J. V. and Purohit S. G. (2005): Life Time Data: Statistical Models and Methods. (Word Scientific).
7. Kalbfleisch J. D. and Prentice R. L. (2001): The Statistical Analysis of Failure Time Data (John Wiley and Sons)
8. Lawless J. F.(2011): Statistical Models and Methods of Failure Time Data. (John Wiley and Sons)
9. Miller R. G.(1981): Survival Analysis. (John Wiley and Sons)
10. Nelson W. B. (2005): Applied Life Data Analysis. (John Wiley and Sons)

MST 423: STATISTICAL QUALITY CONTROL

Unit 1: Quality Improvement Tools: affinity diagram, interrelationship digraph, tree diagram, prioritization matrix, matrix diagram, process decision program chart, activity network diagram, stem-and-leaf display, dot diagrams, boxplot, and normal probability plot, Engineering Process Control. Shewhart Control charts: basic statistical principles and assumptions, phase I and phase II applications, benefits from the use of control charts, concept of rational subgroups, performance measures of a control chart, $\overline{X}$, *R, S,* S^2*, p, c* and *D* charts, σ-control limits and probability control limits, over dispersion. Modifications to control chart procedures: warning limits, sensitizing rules, adaptive design parameters, integration of two charts. Concept of economic design of a control chart.

(12L+3T)

Unit 2: Alternatives to Shewhart control charts: CUSUM and EWMA charts, Shewhart–EWMA Chart. Multivariate Control Charts: multivariate chart versus individual charts, Hotelling's T^2 control chart, multivariate CUSUM Charts, multivariate EWMA charts, Regression adjustment. Other Control Charts: SPRT chart, GLR Chart, charts for autocorrelated data, nonparametric control charts, Bayesian control charts. The change point model for process monitoring.

(12L+3T)

Unit 3: Process capability Analysis: process capability, process capability indices (C_p, C_{pk}, C_{pm}, C_{pmk}), point and interval estimation of C_p and C_{pk}, Nonparametric Capability Indices: robust capability indices, capability indices based on fitted distributions, data transformation, capability indices computed using resampling methods. Multivariate Process Capability Indices. Six Sigma Methodology: components of a Six Sigma methodology, the DIMAC process, Six Sigma applications, Six Sigma concept for customer satisfaction, Six Sigma training, Lean Six Sigma.

(12L+3T)

Unit 4: Acceptance sampling plans for attributes: single sampling plan, double and multiple sampling plans, sequential sampling. Performance measures of sampling plans. Acceptance sampling plans for variables: Advantages and Disadvantages of Variables Sampling, Sampling inspection plans by variables for one or two sided specifications, Sequential Sampling by Variables, Rectifying inspection of lots, the Deming inspection criterion, Continuous sampling plans, skip-lot sampling plans.

(12L+3T)

References:

1. Guenther, W. C. (1977). Sampling Inspection in statistical quality control. Macmillan.
2. Kenett, R. S. and Zacks, S. (2014). Modern Industrial Statistics with applications in R, MINITAB and JMP. John Wiley & Sons.
3. Montgomery, D. C. (2010). Statistical Quality Control: A Modern Introduction, 6th Edition. Wiley India Pvt Ltd.
4. Ryan, T. P. (2011). Statistical Methods for Quality Improvement. John Wiley & Sons.

MST-424 TIME SERIES ANALYSIS

Unit-1: Time series as a discrete parameter stochastic process, Auto - Covariance, Auto-correlation functions and their properties. Partial auto covariance function. Stationary time series, Exploratory time series analysis, Exponential and moving average smoothing, Holt – Winter smoothing, forecasting based on smoothing.

(12 L + 3 T)

Unit-2: Wold representation of linear stationary processes, linear time series models: Autoregressive, Moving Average, Autoregressive Moving Average models. Concept of Causality, invertibility, computation of π -weights and ψ- weights, computation of ACVF, ACF and PACF. Autoregressive Integrated Moving Average models.

(12 L + 3 T)

Unit-3: Estimation of ARMA models: Yule-Walker estimation for AR Processes, Maximum likelihood and least squares estimation for ARMA Processes, Discussion (without proof) of estimation of mean, Auto-covariance and auto-correlation function under large samples theory, Residual analysis and diagnostic checking. Minimum mean squared errorForecasting for ARIMA models. Introduction to SARIMA models.

(12 L + 3 T)

Unit-4: Introduction to spectral analysis, Spectral Representation of the ACVF, Spectral density of an ARMA process, its computation for simple models. Introduction to ARCH and GARCH models. Properties and estimation under ARCH(1) and GARCH(1,1) model.

(12L + 3T)

References:

1. Box, G.E.P and Jenkins G.M. (1970) Time Series Analysis, Forecasting & Control, Holden-Day.
2. Brockwell, P.J and Davis R.A. (1987) Time Series: Theory and Methods, Springer-
3. TsayR. S. Analysis of Financial Time Series, 3rd Ed. (Wil. Ser. in Prob. and Statistics)
4. Kendall, M.G. (1978) Time Series, Charler Graffin
5. Chatfield, C. (2004) The Analysis of Time Series - An Introduction, Sixth edition, Chapman and Hall.

MST-426: ECONOMETRICS

Unit 1: Introduction: Definition and scope of econometrics, Methodology of econometrics. Nature and source of Data for econometric analysis, Types of data: cross section, time series, panel data, dummy variable, instrumental variable. Basic concepts of estimation: Review of general linear model, Ordinary least squares, generalized least squares.

(12L+ 3T)

Unit 2: Heteroscedasticity: consequences and tests: White test, Goldfeld-Quandt test; Estimation: estimation with grouping of observations, estimation of the heteroscedasticity relation. Linear regression with stochastic regressors, Instrumental variable estimation, Errors in variables. Autocorrelation, Autoregressive linear regression, Distributed lag models.

(12L+ 3T)

Unit 3: Simultaneous linear equations model. Example, Identification problem, Restrictions on structural parameters-rank and order conditions. Estimation in simultaneous equations model.Recursive systems. Two-Stage Least Squares estimators. Limited information estimators.

(12L+ 3T)

Unit 4: Definition of causality, Granger causality, Granger test for causality. Application of econometric methods: estimation of demand and supply function – production and cost function, consumption and investment functions.

(12L+ 3T)

References:

1. Apte, P.G. (1990): Text book of Econometrics. Tata McGraw Hill.
2. Gujarati, D.N. (2003): Basic Econometrics, McGraw Hill.
3. Johnston, J. (2006). Econometric Methods, third edition, McGraw Hill
4. Marno Verbeek, (2012): A guide to Modern Econometrics, 4/e,Wiley and Sons.
5. Nachane, D. M. (2006). Econometrics: Theoretical Foundations and Empirical Perspective, Oxford University Press
6. Ramanathan, R. (2002). Introductory Econometrics with applications, 5/e, Thomson Asia Private Limited
7. Wooldridge, J. (2012). Introductory Econometrics: A Modern Approach, 5/e, South-Western.

MST 427: DATA MINING

Unit-1: Data understanding and data cleaning, concept of supervised and unsupervised learning. Problem of classification, classification techniques: k-nearest neighbor, decision tree, Naïve Bayesian, classification based on logistic regression, Bayesian belief Network.

(12L+3T)

Unit-2: Model evaluation and selection: Metrics for Evaluating Classifier Performance, Holdout Method and Random Subsampling, Cross-Validation, Bootstrap, Model Selection Using Statistical Tests of Significance, Comparing Classifiers Based on Cost–Benefit and ROC Curves. Techniques to Improve Classification Accuracy: Introduction to Ensemble Methods, Bagging, Boosting and AdaBoost, Random Forests, Improving Classification Accuracy of Class-Imbalanced Data.

(12L+3T)

Unit-3: ANN and SVM: Artificial Neutral Network (ANN): Introduction to ANN, types of activation function, McCulloch-Pitts AN model, single layer network, multilayer feed forward network model, training methods, ANN & regression models. Convexity and optimization: Convexity, unconstrained and constrained optimization, KKT conditions. Support vector machine: Introduction to support vector machine, loss functions, soft margin, optimization hyperplane, support vector classification, support vector regression, linear programming support vector machine for classification and regression.

(12L+3T)

Unit-4: Unsupervised learning: Clustering: k-mediods, CLARA, DENCLUE, DBSCAN, Probabilistic model based clustering. Market Basket Analysis: Association rules and prediction, Apriori Algorithm, data attributes, applications to electronic commerce.

(12L+3T)

References:

1. Berson and Smith S.J. (1997) : Data warehousing, Data Mining, and OLAP, McGraw-Hill.
2. Breiman J.H Friedman, R.A. Olshen and stone C.J. (1984) : Classification and Regression Trees, Wadsworth and Brooks / Cole.
3. Han, J. and Kamber, M. and Pei, J. (2012) : Data Mining: Concepts and Techniques. MorganGaufmann.3rd Edition.
4. Mitchell T.M. (1997) : Machine Learning , McGraw-Hill.
5. Ripley B.D. (1996) : Pattern Recognition and Neural Networks. Cambridge University Press.
6. Vapnik V.N. The nature of Statistical learning theory, Springer.

7. Cristianini N. and Shawe-Taylor J. An Introduction to support vectormachines.
8. Data set source: http://www.ICS.uci.edu/~mlearn/MLRepository.html
9. Mehrika, K., Mohan, C., and Ranka (1997) Elements of Artificial neural networks. Penram international.
10. Hastie T, Tibshirani R, Friedmant J, (2009): The elements of statistical Learning, Springer.

MST 416- PRACTICAL –IV

1.Solution to LPP using simplex method.

2. Revised Simplex method and Dual Simplex Method.

3. Game Theory.

4. Quadratic programming

5. Integer programming.

6. Dynamic Programming

7-14 Two Practicals each on the Four Optional papers.

Open Electives

AN INTRODUCTION TO MINITAB

Getting started with Minitab: Menu bar, Worksheet, Minitab commands and sub commands, Data entry into worksheet, Importing data, Saving, Retrieving, Printing file.
Operators and functions: Arithmetical operations, Mathematical functions, Column and Row statistics, Comparisons and Logical operations, Coding, Computing ranks, Sorting data, Stacking and Unstacking columns.
Exploratory data analysis: Tallying data, Describing data, Histograms, Box plots, Bar charts Pie charts. Correlation and Regression: Correlation: Scatter plots, Karl Pearson correlation coefficient, Regression: Simple and Multiple regression.
Matrix operation and Macros: Creating matrices, Commands for matrix operations, Writing simple macros.

Text Books / Reference Books:

1. David Moore, George McCabe: MINITAB Manual Introduction to the Practice of Statistics, Michael Evans University of Toronto.
2. Barbara F. Ryan, Brian L. Joiner, Jonathan D. Cryer: Minitab Handbook- Updated for Release 14, Cengage Learning, 2005.
3. Ruth K. Meyer, David D. Krueger: A Minitab Guide to Statistics, Prentice Hall, 1998.
4. Thomas Arthur Ryan, Brian L. Joiner, Barbara F. Ryan: Minitab student handbook, Duxbury Press, 1976.

AN INTRODUCTION TO SPSS

Getting started with SPSS: Data editor, Output viewer, Syntax editor, Script window, Variable view. Charts and Graphs: Line chart, Scatter Plots, Histogram, Bar chart, Box Plot, Pie chart.
Exploratory data analysis: Sum, Mean, Standard deviation, Variance, Minimum value, Maximum value and Range. Correlation and Regression: Correlation: Scatter plots, Karl Pearson correlation coefficient, Partial correlation, Spearman correlation, Regression: Simple and Multiple regression.
Testing of hypotheses: Chi square test for association, Chi square test goodness of fit, Independent sample t test, Paired sample t test, One sample t test, Report generation.

Text Books / Reference Books:

1. William C. Rinaman: Workshop Statistics: SPSS Software Companion Manual, Key College Publishing, 2004.
2. Ton J. Cleophas, Ton J. M. Cleophas, Aeilko H. Zwinderman: Cookbook for Starters on SPSS, Springer.
3. EelkoHuizingh: Applied Statistics with SPSS, SAGE, 2007.

AN INTRODUCTION TO MATLAB

To Launch the Matlab application to get a command line prompt, Import/export of data from/to external files, Creating and manipulating new variables from the command line, Using the built-in help documentation.
Simple linear algebra with matrices: transpose, products, powers, elment-by-element products, determinants, inverse, g-inverse , Characteristic roots and related commands, Sort, minimum, maximum.
Prepare simple macros in the form of M-files: if-else-end, :for-end, while –end and other statements.
Using built-in functions and tool-boxes, creating new function files, Numerical integration, Roots of polynomials and solving complicated equations.
2-D and 3-D Mathematical and Statistical plots.
Generating random numbers from different probability distributions, Descriptive Statistics, Hypothesis testing, Linear regression analysis, ANOVA.

Recommended Texts:

1. Enander, Eva Part (1996) Matlab handbook.
2. Hanselman, D (1998) Mastering matlab 5: Prentice Hall
3. Etter, Delores M (1997) Engineering problem solving with matlab : by Delores M Ettet : Prentice Hall
4. Kwon, Young W (1997) The Finite Element Method Using Matlab : CRC Pub.
5. Pratap, Rudha(2006) Getting started with matlab 7: A quick introduction for scientist and engineers Oxford Uni. Pr
6. Sigh, Y. Kiran (2007) Matlab programming: Prentice Hall
7. Elden, Lars (2006) Introduction to numerical computation: Analysis and MATLAB illustation Pearson Education
8. Mathews, J.H. (2007) Numerical methods using MATLAB :Pearson Education Palm, William J. (2005) Introduction to MATLAB 7 for engineers : McGraw Hill
9. Driscoll, Tobin A. (2009) Learning Matlab -- Society for Industrial & Applied Mathematics 10. Siciliano, Antonio, (2008) MatlabR : Data analysis and visualization -- World Scientific

AN INTRODUCTION TO R SOFTWARE

Getting Started with R: download and installation, introduction to components of R.
Working with data in R: input from keyboard, import file-.xlx, .xlsx, .txt, .csv, etc.Creating vectors, performing arithmetic operations, adding elements to a vector, creating a matrix, matrix operations, extracting elements from a matrix, creating data frame, extracting elements from a data frame, Dealing with missing data. Programming in R: understanding the flow, operators-comparison and logical, looping, for loop, while loop, repeat loop, if loop.

Creating your own function in R and export data from R to another format like .xlx, .xlsx, .csv, .txt.
Statistical analysis in R: descriptive statistics, creating tables and graphs, correlation and regression, performing t-tests, anova.

Recommended Texts:

1. Larry Pace (2012), Beginning R: An Introduction to Statistical Programming, A press.
2. S.R. Deshmukh and S. Purohit. (2007) Microarray Data: Statistical Data Analysis using R, Alpha Science International.

DESIGN OF EXPERIMENTS FOR APPLIED RESEARCH

Need of experimental designs, basic concepts; response variable, factors, levels, randomization, replication, blocking etc. Introduction, model and analysis of basic designs: CRD, RBD, LSD. Simple two level factorial designs: concepts of main effects, interactions, testing significance of factorial effects, Concepts of confounding and regular fractions. Analyzing real data Using MINITAB and interpreting the results.
Response Surface Designs: Basic terminology, first and second order models, search for steepest ascent/descent direction, tests for curvature, search for an optimal level combination, Analyzing real data Using MINITAB and interpreting the results.

References:

1. Montgomery D.C. (2013): *Design and Analysis of Experiments*, 8th edition, Wiley India Pvt Ltd.
2. Voss, D., Dean, A., and Dean, A.(1999).*Design and Analysis of Experiments*, Springer-verlagGmbh.
3. Wu, C. F., Hamada M. S.(2000). *Experiments : Planning, Analysis and Parameter Design Optimization*, 2nd edition, John Wiley & Sons.
4. Mathews Paul (2010) Design of Experiment with Minitab, New Age International Ltd.

MATLAB PROGRAMING

MATLAB basics - The MATLAB environment, operators and simple calculations, Formulas and functions , MATLAB toolboxes. Matrix and linear algebra review - Vectors and matrices in MATLAB - Matrix operations and functions in MATLAB.

Computer programming in MATLAB: MATLAB scripts and functions (m-files) - Simple sequential algorithms - Control structures (if…then, loops) ; Reading and writing data, file handling - Personalized functions; MATLAB graphic ; Numerical methods and simulations ; Random number generation; Monte-carlo methods.

References:

1. Enander, Eva Part (1996) Matlab handbook.
2. Hanselman, D (1998) Mastering matlab 5: Prentice Hall
3. Etter, Delores M (1997) Engineering problem solving with matlab : by Delores M Ettet : Prentice Hall
4. Kwon, Young W (1997) The Finite Element Method Using Matlab : CRC Pub.
5. Pratap, Rudha(2006) Getting started with matlab 7: A quick introduction for scientist and engineers Oxford Uni. Pr
6. Sigh, Y. Kiran (2007) Matlab programming: Prentice Hall
7. Elden, Lars (2006) Introduction to numerical computation: Analysis and MATLAB illustation Pearson Education
8. Mathews, J.H. (2007) Numerical methods using MATLAB :Pearson Education Palm, William J. (2005) Introduction to MATLAB 7 for engineers : McGraw Hill
9. Driscoll, Tobin A. (2009) Learning Matlab -- Society for Industrial & Applied Mathematics 10. Siciliano, Antonio, (2008) MatlabR : Data analysis and visualization -- World Scientific

BASICS OF SAS

Introduction to SAS, Common structure of SAS, Getting familier with interface with Base SAS System, Accessing Data, Use FORMATTED, LIST and COLUMN input to read raw data files, Use of INFILE statement options to control processing when reading raw data files, Use various components of an INPUT statement to process raw data files including column and line pointer controls, and trailing @ controls, Combine SAS data sets using the DATA step.

Creating Data Structures, Create temporary and permanent SAS data sets, Create and manipulate SAS date values, Use DATA Step statements to export data to standard and comma delimited raw data files, Control which observations and variables in a SAS data set are processed and output.

Managing Data : PROC step, Investigate SAS data libraries using base SAS utility procedures, Sort observations in a SAS data set, Conditionally execute SAS statements, Use assignment statements in the DATA step, Modify variable attributes using options and statements in the DATA step, Accumulate sub-totals and totals using DATA step statements, Use SAS functions to manipulate character data, numeric data, and SAS date values, Use SAS functions to convert character data to numeric and vice versa, Process data using DO LOOPS, Process data using SAS arrays.

Generating Reports : Generate list reports using the PRINT and REPORT procedures, Generate summary reports and frequency tables using base SAS procedures, Enhance reports through the use of labels, SAS formats, user-defined formats, titles, footnotes, and SAS System reporting options, Generate HTML reports using ODS statement. Error handling.

There shall be hands-on sessions Practice Sessions where students get familie with SAS

with simple SAS programs.

Recommended Texts:

1. Delwiche, Lora D. (2008), *The Little SAS Book: Primer*, Fourth Edition. Cary, NC: SAS Publishing. ISBN: 9781599947259.
2.Cody, Ronald P. and Jeffrey K. Smith (2006), *Applied Statistics and the SAS ProgrammingLanguage*, Fifth Edition, Upper Saddle River, NJ: Pearson Prentice Hall. ISBN 0-13-146532-5.
3.SAS Institute Inc. (2009), *SAS Certification Prep Guide: Base Programming for SAS 9*, 2nd Edition, Cary, NC: SAS Institute, Inc. ISBN: 978-1-60764-045-5 .

ADVANCED COURSE IN MSEXCEL

Formulae: Techniques for creating complicated formulas, Text functions, Lookup functions, Reference functions, Error functions, Logical Function, Array and Summarizing functions, Database Functions, Date and Time Functions, Information Functions,
Working with data: Data validation, data consolidation, what-if analysis, Pivot tables and charts, advanced filter, subtotals and outlines, securing sheets/workbook, dashboards.
MACRO: Definition and use, record a macro, assign a macro, run a macro and store a macro, Introduction to VBA Programming.

References:

1. Help manual of EXCEL.