SHIVAJI UNIVERSITY, KOLHAPUR.

Accredited By NAAC with 'A' Grade

Revised Syllabus For

Master of Science

Part- II

Applied Statistics and Informatics

CBCS PATTERN

Syllabus to be implemented from

June, 2019 onwards.

SHIVAJI UNIVERSITY, KOLHAPUR

Revised Syllabi of M.Sc. Applied Statistics and Informatics (CBCS) sem III and IV

(To be implemented in the Department of Statistics, Shivaji University (w.e. from 2019-2020)

1) Title of the Programme: M. Sc. (Applied Statistics and Informatics)
   M. Sc. **Applied Statistics and Informatics** program has semester pattern and Choice Based Credit System. The program consists of 96 credits.
2) Year of Establishment of Course: June 2013(Started under UGC Innovative Programme: Teaching and Research in Interdisciplinary and Immerging Areas)
3) Eligibility: B. Sc. (Second class) with Statistics as principal subject.
4) **Equivalence:** This Programme is equivalent to M. Sc.(Statistics) Programme since it covers more than 75% of the core papers of M.Sc.(Statistics) course of Shivaji University. Certificate of the same is available on the university website.
5) Structure of the course
   The following table gives the scheme of Examination at M.Sc. **Applied Statistics and Informatics** according to the Revised Syllabus and pattern of Examination.

   Notations: A seven-character code is given to each paper. In MASI "M" stands for M.Sc. and "ASI" stands for Applied Statistics and Informatics. The first digit following MASI is Semester Number. The second digit "0" stands for the **core** theory paper, the digit "1" stands for a practical paper/project work, the digit "2" stands for an **elective paper**, 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.

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

M.Sc. (Applied Statistics and Informatics) 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
MASI -101	Fundamentals of Computer Programming	4	3	80	20	4
*MASI-102	Statistical Mathematics	4	3	80	20	4
*MASI-103	Distribution Theory	4	3	80	20	4
*MASI-104	Estimation Theory	4	3	80	20	4
*MASI-105	Statistical Computing	3	3	80	20	3
*MASI -146	Communication Skills	1	1	-	20	1
MASI -117	Practical-I	12	3	-	100	4

M.Sc. (Applied Statistics and Informatics) Semester-II(24 credits)

Course Code	Title of the course	Instruction hrs/week	Duration of Exam (hrs)	Marks - Term End Exam	Marks- Internal Assessment	Credits
MASI-201	Advanced Data Structure with C++	4	3	80	20	4

*MASI-202	Theory of Testing of Hypotheses	4	3	80	20	4
*MASI-203	Multivariate Analysis	4	3	80	20	4
*MASI-204	Linear Models and Design of Experiments	4	3	80	20	4
*MASI-205	Sampling Theory	4	3	80	20	4
MASI-216	Practical-II	12	3	-	100	4

M.Sc. (Applied Statistics and Informatics) Semester-III (24 credits)

Course Code	Title of the course	Instruction hrs/week	Duration of Exam (hrs)	Marks - Term End Exam	Marks- Internal Assessment	Credits
MASI-301	Data Base Management System	4	3	80	20	4
*MASI-302	Elementary Stochastic Processes	4	3	80	20	4
*MASI-303	Planning and Analysis of Industrial Experiments	4	3	80	20	4
*MASI-304	Design and Analysis of Sample Surveys	1	-	-	20	1
MASI-316	Practical III	12	3	-	100	4
ANY TWO from the following electives						
*MASI-321	Reliability Theory	3	3	80	20	3
*MASI-322	Regression Analysis	3	3	80	20	3
*MASI-323	Statistical Genetics	3	3	80	20	3
*MASI-324	Demography	3	3	80	20	3
*MASI-325	Statistical Methods in Clinical Trials	3	3	80	20	3
*MASI-326	Actuarial Statistics	3	3	80	20	3
*MASI-327	Dissertation	3	-	-	100	3
MASI-328	Computer Graphics	3	3	80	20	3
MASI-329	Computer Incentive Statistical Methods	3	3	80	20	3
Any ONE open elective of one credit from the list given below:						

M.Sc. (Applied Statistics and Informatics) Semester-IV (24 credits)

Course Code	Title of the course	Instruction hrs/week	Duration of Exam (hrs)	Marks – Term End Exam	Marks- Internal Assessment	Credits
*MASI-401	Optimization Techniques	4	3	80	20	4

MASI-402	Python for Data Science	4	3	80	20	4
MASI-416	Practical IV and Project Work	12	1.5	-	100	3
ANY THREE from the following electives of 4 credits each						
*MASI-421	Generalized Linear models	4	3	80	20	4
*MASI-422	Survival Analysis	4	3	80	20	4
*MASI-423	Statistical Quality Control	4	3	80	20	4
*MASI-424	Time Series Analysis	4	3	80	20	4
*MASI-425	Statistical Ecology	4	3	80	20	4
*MASI-426	Econometrics	4	3	80	20	4
*MASI-427	Data Mining	4	3	80	20	4
MASI-428	Bioinformatics	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, Engineering	20	1
ODD	An introduction to R software	1	20	Mathematics, MCA, Computer Science (Excluding Statistics, ASI)	20	1
ODD	Design of Experiments for Applied Research	1	20	All Faculties excluding Statistics, ASI)	20	1
EVEN	Design and Analysis of Laboratory Experiments	1	20	All Faculties (Excluding Statistics, ASI)	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

*These courses are common to the M. Sc. (Statistics) and M. Sc. (Applied Statistics and Informatics) programs. Question papers of these courses for both the programs will be common in all examinations.

Note:-

a) A syllabus for some elective courses is framed. 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. Applied Statistics and Informatics 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 MASI-117, MASI-216 and MASI-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) Semester IV, Practical MASI-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 understanding the domain of the problem, formulation of the problem, collection of the relevant data, Analysis of the data and report writing. They are expected to use software for which they are trained. 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 or problem in industry. The budgetary provision for field work/industrial tour will be utilized to meet expenses towards the field work.

Syllabus and the nature of question paper for each of the courses of M.Sc. (Statistics) programme mentioned in the first column of the following table is exactly same as that of the course of M. Sc. (Applied Statistics and Informatics) programme mentioned against it in the second column. The question papers of such two matched courses will be common in all examinations of M.Sc. (Statistics) and M. Sc. (Applied Statistics and Informatics) programs.

M. Sc. Statistics	M. Sc. Applied Statistics and Informatics
SEMESTER I	
MST 102: Statistical Mathematics-II	MASI-102: Statistical Mathematics
MST 103: Distribution Theory	MASI-103: Distribution Theory
MST 104: Estimation Theory	MASI-104: Estimation Theory
MST 105: Statistical Computing	MASI-105: Statistical Computing
SEMESTER II	
MST 202: Theory of Testing of Hypotheses	MASI-202: Theory of Testing of Hypotheses
MST 203: Multivariate Analysis	MASI-203: Multivariate Analysis
MST 204: Linear Models and Design of Experiments	MASI-204: Linear Models and Design of Experiments
MST 205: Sampling Theory	MASI-205: Sampling Theory
SEMESTER III	
MST 302: Elementary Stochastic Processes	MASI-302: Elementary Stochastic Processes
MST 303: Planning and Analysis of Industrial Experiments	MASI-303: Planning and Analysis of Industrial Experiments
MST 304: Design and Analysis of Sample Surveys	MASI-304: Design and Analysis of Sample Surveys
MST 321: Reliability Theory	MASI-321: Reliability Theory
MST 322: Regression Analysis	MASI-322: Regression Analysis
MST 323: Statistical Genetics	MASI-323: Statistical Genetics
MST 324: Demography	MASI-324: Demography
MST 325: Statistical Methods in Clinical Trials	MASI-325: Statistical Methods in Clinical Trials
MST 326: Actuarial Statistics	MASI-326: Actuarial Statistics
MST 327: Dissertation	MASI-327: Dissertation
SEMESTER IV	
MST 401: Optimization Techniques	MASI-401: Optimization Techniques
MST 421: Generalized Linear models	MASI-421: Generalized Linear models
MST 422:Survival Analysis	MASI-422:Survival Analysis
MST 423:Statistical Quality Control	MASI-423:Statistical Quality Control
MST 424:Time Series Analysis	MASI-424:Time Series Analysis
MST 425:Statistical Ecology	MASI-425:Statistical Ecology
MST 426: Econometrics	MASI-426: Econometrics
MST 427:Data Mining	MASI-427:Data Mining

Detailed Syllabi of papers for Semester III

MASI-301: DATA BASE MANAGEMENT SYSTEM

Unit-I: Introduction to Databases and Data Models: Concept of database system, purpose of database system, view of data, different sources of databases, relational databases, database architecture, importance of data models, basic building blocks, business rules, the evolution of data models, data abstraction, database users and administrators. (12L+3T)

Unit-II: Database design and ER Model: Overview, ER-Model, Constraints, ER-Diagrams, ERD issues, weak entity sets, Codd's rules, relational schema, introduction to UML relational database model: logical view of data, keys, and integrity rules. Relational Database design: features of good relational database design,atomicdomainandnormalization(1NF,2NF,3NF,BCNF). (12L+3T)

Unit III: Relational algebra: Introduction of RDBMS, Selection and projection, set operations, renaming, Joins, Division, syntax, semantics. Operators, grouping and ungrouping, relational comparison.
Calculus: tuple relational calculus, Domain relational Calculus, calculus vs algebra, computational capabilities. Constraints: Concept and types of constrains. Views: Introduction to views, data independence, security, updates on views, comparison between tables and views. SQL: Basics of SQL, DDL, DML, DCL, structure: creation and alteration, defining constraints: primary key, foreign key, unique, not null, check, IN operator, Functions: aggregate functions, built-in functions, numeric, date, string functions, set operations, sub-queries, correlated sub-queries, use of group by, having, order by, join and its types, exist, any, all , view and its types joined relations, Triggers. (12L+3T)

Unit-IV: Transaction management: ACID properties, serializability and concurrency control, lock based concurrency control (2PL, Deadlocks), time stamping methods, optimistic methods, database recovery management, data dictionary.
NoSQL: Overview, and history of NoSQL databases, definition of the four types of NoSQL database, introduction to Big Data. (12L+3T)

References:

1. Abraham Silderschutz, H. Korth and S. Sudarshan: "Database systems concepts ",6th Edition, McGraw Hill Education
2. Peter Rob, Carlos Coronel: Database Systems: Design, Implementation, &Management.
3. Oracle installation and user manual
4. Ivan Bayross: SQL, PL/SQL The programming language of oracle, 4th edition, BPB publication

MASI-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 BinaymiBranching 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)

MASI- 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.

MASI-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. ParimalMukhopadhyay (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.

MASI-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.*

MASI-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.

MASI- 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.

MASI-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

MASI-328: COMPUTER GRAPHICS

Unit 1: Introduction: a) Input/output devices: Keyboard, Mouse, Trackball, Joysticks, Data Glove, Digitizers, Light pen, Touch panels, Image scanners, Printers and plotters. b) Logical Input Devices: Locator, Stroke, String, Valuator, Choice and Pick. c) Video Display Devices: Refresh Cathode-Ray Tubes, Raster-Scan Displays, Random-Scan Displays, Color CRT Monitors, Direct-View Storage Tubes, Flat Panel Displays d) Raster-Scan Systems: Video Controller, Raster-Scan Display Processor, Random-Scan Systems. (12L +3T)

Unit 2:A) Output Primitives:a) Line, Circle, Ellipse and Curve generation algorithm, b) Polygon filling algorithm c) Windowing and clipping : Window to Viewport transformation, line clipping and polygon clipping, B) 2D and 3D transformations : a) 2D basic transformation, other transformation, composite transformation, matrix representation and homogeneous transformation, b) 3D concepts : Display models, parallel and perspective projections c) 3D basic transformation, other transformation & composite transformation. (12L+3T)

Unit 3: Colors in computer graphics: Chromatic and achromatic light, properties of light, color lookup tables, Color models : XYZ, RGB, CMY, HSV, HLS, B) Curve generation: a) Bezier curve, properties of Bezier curve, Cubic Bezier Curve, b) B-Spline curves: i)Uniform, Periodic B Spline, ii) Cubic, periodic B-spline, iii) Open, uniform B- Spline iv)Non-uniform B-spline, c) Beta-Spline : Beta spline continuity conditions, cubic periodic beta spline, matrix Representation, d) Introduction to fractal (Koch and Hilberts curve) (12L+3T)

Unit 4:A) Illumination model and shading methods:a) Basic illumination models : Ambient light, diffuse reflection, specular reflection and its Phong model, shadows and transparency, ray tracing, displaying continuous tone images, halftone pattern and Dithering techniques, aliasing and antialiasing b)Phong rendering methods: Constant intensity shading, Gouroud shading, Phong and Fast Phong shading B) Visible surface detection methods :Classification of visible surface detection algorithm, Back-face detection, depth-buffer method, A-buffer method and Painter's algorithm C) Design and implementation of Application s/w :Study of advance software platform viz. 3-D studio max, Animator Pro, Introduction to OPEN GL, comparison with the facilities provided by conventional IDEs viz. CC++, Visual computing environment.

(12L+3T)

Reference:
1. Donald Heran and M Pauline Baker: Computer Graphics
2. Roger Stevens: Advanced Graphics Programming in C and C++ 3) F. S. Hill: Computer Graphics
3. Newmann Sproul: Principles of Interactive Computer Graphics
4. Foley J D & van Dam: Fundamentals of Interactive Computer Graphics
5. Plastock& Kelly: Theory and Problems of Computer Graphics
6. Devidrogers: Mathematical Elements of Computer Graphics

MASI-329: COMPUTER INCENTIVE STATISTICAL METHODS

Unit 1: General Principles in Random Variate Generation: The Inversion Method, the Rejection Method, Composition, the Ratio-of-Uniforms Method (RoU), Almost-Exact Inversion, Using Special Properties of the Distribution. (12L +3T)

Unit 2: MCMC methods for performing inference; Importance sampling and Monte –Carlo Integration; Monte –Carlo Optimization: Exploratory, probabilistic approach for missing data; Simulated Annealing / Metropolis Algorithm for Optimization, Application to missing & Censored data; EM: Theoretical core, Genetic linkage, Application to censored and missing data; Monte–Carlo EM (MC- EM): genetic linkage, Applications to censored data; EM & MC-EM standard errors. (12L +3T)

UNIT 3: MCMC algorithm Essentials: AR (1), Stochastic Processes & their properties M-H Algorithm: Working principle of MCMC algorithm, properties, convergence, Independent M-H Algorithm, Random walk Metropolis algorithm; Gibbs Sampler Algorithm: Two stage Gibbs sampler, connection of Gibbs sampler to E-M algorithm and Applications, Hamersley-Clifford theorem, multistage Gibbs sampling. (12L +3T)

Reference: Hormann, W., Leydold, J., &Derflinger, G. (2013). Automatic Non uniform random variate generation. Springer Science & Business Media.

C. Rabat & G. Casella (2004) Monte Carlo statistical methods, Springer

MASI 316- PRACTICAL –III

1. Design Database, Creation of tables with constraints
2. Implementation of SQL queries 1
3. Implementation of SQL queries 2
4. SQL queries for triggers, and insertion , updation and deletion using the referential integrity constraints
5. Realization from a Markov Chain.
6. Classification of t.p.m and computation of n-step probability matrix.
7. Classification of State: Computations of absorption probabilities.
8. Stationary distribution and recurrence times.
9. Realization from discrete state space Markov Processes and related estimation.
10. Analysis of full replicated unconfounded 2^n and 3^2 factorial experiments.

11. Analysis of single replicated 2^n and 3^2 factorial experiments.
12. Analysis of confounded 2^n and 3^2 factorial experiments: total and partial confounding.
13. Analysis of fractional factorial 2^{n-k} and 3^{n-k} factorial experiments.
14. 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:

MASI-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) KantiSwaroop& 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.

MASI-402: PYTHON FOR DATA SCIENCE

Unit 1: Introduction, Installation and Working with Pythonhttp://stanford.edu/class/cme193/python_course/lectures/lecture1.pdf, variables, Operators, data types input/output, data types, lists, dictionaries, data import and export, tuples, operations Functions, Functional Programming, Control Structures, Classes, Objects, Inheritance, basic statistical analysis, Libraries: Numpy, Pandas, Scipy, matplotlip. (12L+3T)

Unit 2: Overview of neural Network concept, optimizers (Stochastic gradient descent, Adaptive Gradient Algorithm (AdaGrad)), introduction to deep learning, understanding different types of layers in sequential method: Dense, Convolutional Layers, Recurrent Layers, Normalization Layers. Deep Neural Network architecture design, optimization for Deep NN, regularization methods for deep NN, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), Deep learning Python libraries: tensorflow (Low level) and Keras (High Level). (12L+3T)

Unit 3: Introduction to Natural Language Processing (NLP), Natural Language Toolkit (NLTK) in Python. Data Preparation:punctuation removal, stop-words removal, numeric value removal, frequent words removal, rare words removal, spelling correction, tokenization, stemming, lemmatization. Feature Engineering:count vectors as features, Term Frequency-Inverse Document Frequency (TF-IDF), TF-IDF vectors as features, word level TF-IDF, N-Gram level TF-IDF, Character level TF-IDF, world cloud, Inverse Document Frequency, word embedding as features, Text/NLP based features, Topic Models as features, word2vec, sentiment analysis. (12L+3T)

Unit 4: Machine learning using scikit-learn library: Classification, Regression, and Clustering. Computer vision: Introduction, OpenCv library in Python, Getting Started with images, Basic Operations on Images, Arithmetic Operations on Images,Image Preprocessing: changing colorspaces, geometric transformations, thresholding, smoothing, morphological transformations, gradients, Canny Edge detection, image pyramids, image segmentation with Watershed algorithm, Feature Detection and Description. Image Detection and recognition examples. (12L+3T)

Reference:

1. Bird, S., Klein, E., &Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. " O'Reilly Media, Inc.".
2. Goodfellow, I., Bengio, Y., Courville, A., &Bengio, Y. (2016). Deep learning (Vol. 1). Cambridge: MIT press.
3. Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., &Mathur, I. (2016). Natural Language Processing: Python and NLTK. Packt Publishing Ltd.
4. Lutz, M. (2013). Learning Python: Powerful Object-Oriented Programming. " O'Reilly Media, Inc.".
5. McKinney, W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.".
6. Mueller, J. P., &Massaron, L. (2015). Python for data science for dummies. John Wiley & Sons.
7. Shanmugamani R. (2018). Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras. “Packt Publishing Ltd”
8. Solem, J. E. (2012). Programming Computer Vision with Python: Tools and algorithms for analyzing images. " O'Reilly Media, Inc.".
9. VanderPlas, J. (2016). Python data science handbook: essential tools for working with data. " O'Reilly Media, Inc.".

MASI-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.

MASI- 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)

MASI- 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.

MASI-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 error forecasting 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, CharlerGraffin
5. Chatfield, C. (2004) The Analysis of Time Series - An Introduction, Sixth edition, Chapman and Hall.

MASI-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.

MASI-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. CambridgeUniversity 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 statisticalLearning, Springer.

MASI-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. Introduction to Python
8. Deep learning using Python
9. Natural language processing using Python
10. Machine learning algorithms using Python
11. Computer Vision using Python

12-19 Two Practicals each on the Four Optional papers.

Syllabus of 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 illustationPearson 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 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:Help manual of EXCEL.

Equivalence for papers

M.Sc. (Applied Statistics and informatics) Semester-I

Paper No.	Old Title of the course	Paper No.	New Title of the course
MASI-101	Fundamentals of Computer Programming	MASI -101	Fundamentals of Computer Programming
MASI-102	Statistical Mathematics	MASI-102	Statistical Mathematics
MASI-103	Probability Distributions	MASI-103	Distribution Theory
MASI-104	Statistical Inference	MASI-104	Estimation Theory
MASI-105	Statistical Computing and Numerical Methods.	MASI-105	Statistical Computing
MASI-116	Practical-I	MASI -117	Practical-I

M.Sc. (Applied Statistics and informatics) Semester-II

Paper No.	Old Title of the course	Paper No.	New Title of the course
MASI-205	Advanced Data Structure with C++	MASI-201	Advanced Data Structure with C++
MASI-201	Probability theory	**	*
MASI-202	Multivariate Analysis	MASI-203	Multivariate Analysis
MASI-204	Linear Models and Design of Experiments	MASI-204	Linear Models and Design of Experiments
MASI-203	Sampling Theory	MASI-205	Sampling Theory
MASI-216	Practical –II and Project-I	MASI-216	Practical-II

*: In case any repeaters could not clear the paper in stipulated number of attempts and want to re-appear for the paper, then it is recommended to set a question paper on the old paper for such candidates.

M.Sc. (Applied Statistics and informatics) Semester-III

Paper No.	Old Title of the course	Paper No.	New Title of the course
MASI-32422	Data Base Managment System (elective)	MASI-301	Data Base Management System
MASI-301	Stochastic Processes	MASI-302	Elementary Stochastic Processes
MASI-302	Computer Graphics	MASI-328	Computer Graphics (elective)
MASI-32531	Reliability and Survival Analysis (elective)	MASI-321	Reliability Theory (elective)
MASI-32311	Regression Analysis (elective)	MASI-322	Regression Analysis (elective)
MASI-32532	Analysis of Clinical Data (elective)	MASI-325	Statistical Methods in Clinical Trials (elective)

MASI-32533	Actuarial Statistics (elective)	MASI-326	Actuarial Statistics (elective)
MASI-32312	Computer Intensive Statistical Methods (elective)	MASI-329	Computer Incentive Statistical Methods (elective)
MASI-336	Practical -III	MASI-316	Practical III

M.Sc. (Applied Statistics and informatics) Semester-IV

Paper No.	**Old Title of the course**	**Paper No.**	**New Title of the course**
MASI-401	Optimization Techniques	MASI-401	Optimization Techniques
MASI-402	Discrete Data Analysis	MASI-421	Generalized Linear models
MASI-42341	Time Series Analysis	MASI-424	Time Series Analysis
MASI-42342	Planning and Analysis of Industrial Experiments (elective)	MASI-303	Planning and Analysis of Industrial Experiments
MASI-42451	Industrial Statistics	MASI-423	Statistical Quality Control
MASI-42561	Data Warehousing and Data Mining	MASI-427	Data Mining
MASI-436	Practical –III and Project -II	MASI-416	Practical IV and Project Work