SHIVAJI UNIVERSITY, KOLHAPUR.

****

Accredited By NAAC with 'A' Grade

CHOICE BASED CREDIT SYSTEM

Syllabus For

M.Sc. Part - I Applied Statistics and Informatics

SEMESTER I AND II
(Syllabus to be implemented from June, 2018 onwards.)

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

(To be implemented in the Department of Statistics, Shivaji University (w.e. from 2018-2019)

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) **Admission Process:** There will be Entrance Examination. While preparing the merit list for M. Sc. Admission, the performance at B. Sc Part III (Statistics) and the performance at the Entrance Examination will be given equal weightage (50-50).
6) **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	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	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	Project work and scientific report writing	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	3	3	80	20	3

	Clinical Trials					
*MASI-326	Actuarial Statistics	3	3	80	20	3
*MASI-327	Dissertation	3	-	-	100	3
MASI-328	Computer Graphics	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,	20	1

				Electronics, Physics, Engineering		
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) programmes. Question papers of these courses for both the programmes 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) programmes.

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
MST 302: Elementary Stochastic Processes	MASI-302: Elementary Stochastic Processes
SEMESTER III	
MST 303: Planning and Analysis of Industrial Experiments	MASI-303: Planning and Analysis of Industrial Experiments
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

Syllabi of papers for Semester I:

MASI-101: FUNDAMENTALS OF COMPUTER PROGRAMMING

Unit 1: Overview of Computer programming, Algorithms: The concept and features of the algorithm, ways of writing the algorithm, writing step by step procedure, Problem redefinition, Flow charts/Decision Trees, Advantages and drawbacks of flowcharts, concept of Tracing and Testing of Algorithm/flowchart. Sequential flow of Logic, Control structures, Iterative method to reach the solution, Solutions to the simple problems: Pseudo code generation.

(12L+3T)

Unit 2: Variable, Constants and Data types, Implementation of sequential, selection and iterative structures. Device (Files) interfacing. Solutions to the complex problems: Structured programming, Modular programming, TOP DOWN/BOTTOM UP approach, Recursive algorithms, Examples, Illustrating structured program development methodology and use of block structured algorithmic language to solve specific problem. Syntax and semantics, documentation and debugging of a program.

(12L+3T)

Unit 3: Advanced Data Types and their implementation: Arrays, Strings, Records, Pointers, Structures, Union, and Applications in the record keeping of real life system. Dynamic memory allocation.

(12L+3T)

Unit 4: Sorting: Insertion sort, selection sort, bubble sort, heap sort, quick sort, merge sort and radix sort. Searching: linear search, binary search. Design and Analysis of algorithms, Divide and conquer, greedy algorithms, Dynamic programming, Backtracking.

(12L+3T)

Note: Emphasis should be given on better programming styles and implementation is expected through C compiler.

References

1. Rajesh K. Shukla (2015): Analysis and Design of Algorithms, A beginners approach- Wiley.
2. Balagurusamy, E. (2016): Programming in ANSI C, McGraw Hill Education
3. D. Ravichandran: Programming with C++, 3rd edition, McGraw Hill Education
4. Horowitz & Sahani (1998): Fundamentals of Computer Algorithms, Galgotia Publications,
5. D. E. Knuth(1997): The Art of Computer Programming: Volume 1: Fundamental - Narosa Publishing House,
6. Robert L. Kruse(2006): Data structures and program design C, 2nd edition- PEARSON

MASI-101: STATISTICAL MATHEMATICS

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

(12L+3T)

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

(12L+3T)

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, Cayley-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

MASI-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. Rohatagi V. K. & Saleh A. K. Md. E.(2001) : Introduction to Probability Theory and Mathematical Statistics- John Wiley and sons Inc.
2. Johnson N. L. &Kotz. S. (1996) : Distributions in Statistics Vol-I,II and III, 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)

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

(12L+3T)

Unit 4: The concept of prior distributions, various types of priors, non-informative, Jeffrey's, least favorable prior, posterior distribution; Posterior distribution conjugate family and standard examples of such families. Bayes estimation under squared error and absolute error loss functions.

(12L+3T)

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.

MASI-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)

MASI-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+ 1T+2Group Discussion)

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

MASI-117: PRACTICAL –I

1. Algorithms and flowcharts
2. Implementation of control structures and functions in C
3. Implementation of array, pointer, structure, union
4. Implementation of sorting and searching algorithms
5. Linear dependence and independence of vectors.
6. Gram-Schmidt orthogonalisation method.
7. Solving systems of equations.
8. Inverse and g-inverse of a matrix.
9. Applications of Cayley-Hamilton theorem.
10. Inverse of a Partitioned matrix.
11. Characteristics roots and vectors and their applications.
12. Classifications and reduction of quadratic forms.
13. Sketching of d.fs.
14. Model sampling from univariate and bivariate.
15. Construction of UMVUE
16. Methods of Estimation: MLE and MME
17. Methods of Scoring
18. Practical on Bayesian inference.

19 – 24. Programming assignments on MASI–105 Course. (Software to be used : R/MINITAB/MATLAB/SAS/SYSTAT depends on availability)

Syllabi of papers for Semester II:

MASI-201: ADVANCED DATA STRUCTURE WITH C++

Unit 1: C++ programming basics. Data types, Variables, Statements, Expressions, Control structures, Looping, Functions, Arrays, Pointers. Object oriented programming: Concept of OOP, class and objects, inheritance, polymorphism (function and operator overloading).

(12L+3T)

Unit 2: Data Structures: definitions, operations, implementations and applications of basic data structures. Stack, Applications of Stack, Queue, Priority Queue, circular queue Applications of queue, Linked lists, doubly linked list, circular list, dynamic memory allocation, implementation of linked list, further operations, implementation of sparse matrices.

(12L+3T)

Unit 3: Trees, binary trees, binary search trees, operations on binary search tree, applications of binary trees, threaded binary trees. General trees, using binary trees to represent general trees. AVL tree, operations on AVL tree. 2-3 trees, operations on 2-3 tree. Multi-way search tress, B-trees, B-tree indexing, operations on B trees. Huffman tree.

(12L+3T)

Unit 4: Graphs, representation, operations on graph, Applications of graph: shortest path problem, Dijkstra's algorithm, topological ordering, minimum spanning tree, DFS and BFS spanning tree, Kruskal's algorithm. Tables, Hash tables: hashing techniques, collision resolution techniques, closed hashing, open hashing.

(12L+3T)

References:

1. D. Samantha: Classic Data Structures, PHI.
2. Kruse, Leung, Tondo: Data Structure and Program Design in C (PHI).
3. Sartaj Sahani: Data Structure Algorithms and Applications in C++ Macgraw Hill.
4. Mark Allan Welss: Data Structure and Algorithm analysis in C++ Addison weslay.
5. Decker, R and Hirshfield, S. (1998): The Object Concept: An Introduction to Computer Programming using C++. (PWS Publishing).
6. Lippmann, S.B. and Lajoie, J. (1998): C++ Primer. Third edition. (Addison-Wesley).
7. Naughton, P. (1996): The Java Handbook. (Tata McGraw-Hill).
8. Savitch, W.J. (2001): Problem Solving with C++: The Object of Programming. Third edition. (Addison-Wesley Longman).

MASI-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.

MASI-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, Hierarchical methods: Single, Complete, average linkage method and non-heirarchical clustering method-k-means clustering.

(12L+3T)

Unit 4: Canonical correlation analysis, Introduction to principal component analysis and 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.

MASI-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.

MASI-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, and 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.

MASI-216: PRACTICAL –II

1. Programming in C++
2. Object Oriented Programming using C++
3. Implementation of Stack, Queue and Linked list Data Structures
4. Implementation of Tree Data Structures
5. Exploratory data analysis.
6. Application of Hotelling's T^2 statistics
7. Discriminant Analysis
8. Cluster Analysis
9. Principle component analysis and Factor Analysis.
10. M.P. UMP, and UMPU Tests
11. Likelihood ratio tests
12. Confidence Intervals
13. Non-parametric Tests
14. Linear Estimation: Estimation and Hypothesis testing
15. ANOVA : One way and two way orthogonal data without interaction
16. ANOVA: Two way orthogonal data with interaction
17. Two way non-orthogonal data without interaction
18. Analysis of BIBD.
19. Analysis of general block design.
20. Basic sampling designs.
21. Ratio, regression, Horvitz-Thompson method of estimations.
22. Stratified, Systematic and cluster Sampling.
23. Multi-stage sampling
24. Non-sampling errors.

(Each practical should consist of problems to be solved using each of the following software EXCEL/ R/ MINITAB/ MATLAB/ SYSTAT /SAS wherever applicable.)