Savitribai Phule Pune University

(Formerly known as University of Pune)

Two Year M. A. / M. Sc. Degree Program in Statistics

Revised Syllabi of M. A. / M. Sc. in Statistics (Credit System)

(To be implemented in the Department of Statistics, SPPU)

(With effect from Academic Year 2013-2014)

Submitted by

Prof. S. R. Deshmukh
Chairperson,
Board of Studies in Statistics

1. Title of the course: M. A. / M. Sc. in Statistics

2. Preamble of the syllabus: M. A. / M. Sc. Statistics program is of minimum 100 credits spread over four semesters. This program is offered at the Department of Statistics, Savitribai Phule Pune University of. The program emphasizes both theory and applications of statistics and is structured to provide knowledge and skills in depth necessary for the employability of students in industry, other organizations, as well as in academics. The program has some unique features such as independent projects, a large number of elective courses, pre-requisite system and extensive computer training of statistical computations including standard software packages such as MATHEMATICA, MATLAB, MINITAB, R, S-PLUS, GAUSS and SYSTAT. The department has the academic autonomy and it has been utilized to add the new and need based elective courses. In the past courses such as *Time Series Analysis, Survival Analysis, Reliability, Actuarial Statistics, Multiple Decrement Models in Insurance, Stochastic Models in Finance, Data Mining, Statistical Analysis of Clinical Trials, Statistical Analysis of Microarray Data, Analysis of Longitudinal Data, Frailty models, Statistical Bio-Computing and Natural Language Processing* have been introduced. The independent project work is one of the important components of this program. The syllabus of the first year (two semesters) covers most of the core courses. In the second year syllabus there are five core courses and five elective courses. The syllabus has been framed to have a good balance of theory, methods and applications of statistics.
It is possible for the students to study basic courses from other disciplines such as economics, life sciences, computer science, and mathematics in place of electives.

3. Introduction: M. A. / M. Sc. Statistics program has semester pattern and credit system with variable credits. The program consists of 100 credits. Credits of a course are specified against the title of the course. A course with T in brackets indicates that it is a theory course whereas a course with P in brackets indicates that it is a practical course. Some of the practical courses are linked with a theory course and in such a case, both the courses will have the same number with T and P, indicating a theory and a practical course respectively. A student can enroll for a practical course if

(i) the student has enrolled for the corresponding theory course (as indicated) in the same term

or

(ii) the student has passed the corresponding theory course in an earlier term

or

(iii) terms for the corresponding theory course have been granted in an earlier term.

The program has pre-requisites system. Under this system, a student has to clear pre-requisite courses to take up some courses in following semesters. Unless a student passes the pre-requisites of a course, (s)he will not be allowed to enroll in the said theory courses. Pre-requisites of a course are indicated in curly brackets against the course. The Departmental committee may change the pre-requisites of a course.

4. Eligibility: For M. A. in Statistics following candidates are eligible.

(i) B. A. (Second class) with Statistics as principal and Mathematics at subsidiary level,
(ii) B. A. (Second class) with Mathematics as principal and Statistics at subsidiary level,
(iii) M. A. (Second class) in Mathematics
(iv) B. A. (Second class) in Actuarial Science with Mathematics and Statistics at subsidiary level,
(v) B. A. (Second class), with three subjects at third year with Statistics as one of the subjects.

For M. Sc. in Statistics following candidates are eligible.

(i) B.Sc. (Second class) with Statistics as principal and Mathematics at subsidiary level,
(ii) B.Sc. (Second class) with Mathematics as principal and Statistics at subsidiary level,
(iii) M.Sc. (Second class) in Mathematics,
(iv) B.Sc. (Second class) in Actuarial Science with Mathematics and Statistics at subsidiary level,
(v) B.Sc. (Second class), with three subjects at third year with Statistics as one of the subjects.

5. Examination

A) (i) Pattern of examination: There would be continuous internal assessment (CIA) and an end of term examination (ETE) for each course. Both CIA and ETE have 50% weightage. The CIA includes class tests, assignments, viva-voce and presentations. There would be assignments or minor projects for some of the elective courses in ETE also.

(ii) Pattern of the question paper: For a theory course with 4 credits, duration for the ETE will be three hours. For a theory course with 3 credits, duration for the ETE will two and half hours, for a theory course with 2 credits, duration for the ETE will two hours and for a practical course with 2/3 credits, duration for the ETE will be 2 hours

B) Standard of passing: A student has to obtain 40% marks in the combined grading of the ETE and the CIA for passing the course, with a minimum passing of 30% in both CIA and ETE separately.

C) ATKT rules: A student can register for the third semester, if (s)he completes 50% credits of the total credits expected to be completed within first two semesters, subject to the pre-requisite system mentioned above. Once registered, a student must complete M. A. / M. Sc. within a period as specified by the University.

D) Award of class: As per the University rules.

E) External students: Not applicable

F) Setting of question paper: The Department has been given autonomy for setting the question papers and grading the answer papers. A teacher teaching a course sets the question paper of the course and assesses the answer books. Question papers are moderated by a committee of subject experts after a preliminary round of internal moderation.

G) Verification or revaluation: As per the University rules

6. Structure of the course

(a) Compulsory papers

Compulsory courses in semester I and semester II are listed below.

Semester I

ST 1 Mathematical Analysis (4)
ST 2 (T) Linear Algebra (3)
ST 2 (P) Linear Algebra (2)
ST 3 Univariate Probability Distributions (2)
ST 4 Multivariate Probability Distributions (4)
ST 5 (T) Elements of Statistical Computing (4)
ST 5 (P) Elements of Statistical Computing (3)
ST 6 Calculus (4)

Total credits 26

Semester II

ST 7 Probability theory (4) {ST 1, ST 6}
ST 8 Limit Theorems in Probability (2) {ST 1, ST 6}
ST 9 (T) Multivariate Analysis (4) { ST 2 (T), ST 3, ST 4, ST 6}
ST 9 (P) Multivariate Analysis (2) { ST 2 (T), ST 3, ST 4, ST 6}
ST 10 Point &Interval Estimation (3) {ST 3, ST 4}
ST 11 Testing of Hypotheses (3) {ST 3, ST 4}
ST 12 (T) Regression Analysis (4) {ST 2 (T), ST 3, ST 4}
ST 12 (P) Regression Analysis (2) {ST 2 (T), ST 3, ST 4}

Total credits 24

In semesters III and IV some courses are compulsory and some are optional. List of optional courses is given in 6(b).

Semester III

ST 13 Stochastic Processes (4)
ST 14 (T) Asymptotic Inference (4) {ST 7, ST 8, ST 10, ST 11}
ST 14 (P) Asymptotic Inference (2) {ST 7, ST 8, ST 10, ST 11}
ST 15 (T) Design of Experiments and Analysis of Variance (4) {ST 12(T)}
ST 15 (P) Design of Experiments and Analysis of Variance (2) {ST 12(T)}
Elective course I (4)
Elective course II (4)

Total credits 24

Semester IV

ST 16 (T) Sampling Methods (4)
ST 16 (P) Sampling Methods (2)
ST 17 Bayesian Inference (4) {ST 10, ST 11}
ST 18 (P) Project (4) {ST 9 (T), ST 12 (T)}
Elective course III (4)
Elective course IV (4)
Elective course V (4)

Total credits 26

The compulsory courses in all the semesters are the core courses. These have 80 (61 (theory) + 15 (practical) + 4 (project)) credits in all. There are five elective courses having 20 credits.

A student may be allotted a module by the Department at the beginning of the second year of the program. The allotment is based on the performance of student as indicated by her/his rank among all the students given admission in the same year. A total of 12 credits should be obtained from the courses listed as module specific courses. The remaining 8 credits need to be obtained from the optional courses listed in 6(b) or from courses from other departments, as allowed by the University. Departmental committee may allow a student to complete M.A./M.Sc. without any specific module.

A student can opt for a theoretical project in lieu of an optional course or two optional courses with the permission of the Head of the Department. A copy of rules and regulations regarding completion and submission of the project work by a student and assessment of the project work is available in the Department.

Below is given a list of 5 modules. Courses specific to a given module are also listed. Each of these courses, called as a module-specific course, has 4 credits.

A project can have either 4 credits or 8 credits. A project with 4 credits has to be completed within a semester. A project with 8credits will be spread over two semesters; work in each semester is to be counted for 4 credits. A student can opt for two different projects also, each having 4 credits.

Modules

M1. Probability and Mathematical Statistics
M2. Industrial Statistics
M3. Financial and Actuarial Statistics
M4. Bio-Statistics
M5. Computational Statistics

List of module specific courses

Module 1: Probability & Mathematical Statistics (M1)

ST P1 Measure Theory and Probability {ST 7, ST 8}
ST P2 Advanced Stochastic Processes {ST 7, ST 8, ST 13}
ST P3 Inference in Stochastic Processes {ST 13, ST 14 (T)}
ST P4 Advanced Probability {ST 7, ST 8}
ST P5 Advanced Inference {ST 10, ST 11}

Module 2: Industrial Statistics (M2)

ST I1 Optimization Techniques {ST 2 (T)}
ST I2 Statistical Methods for Quality Control {ST 3, ST 4}
ST I3 Statistical Methods for Reliability {ST 3, ST 4}
ST I4 Time Series Analysis {ST 12 (T)}
ST I5 Stochastic Models in Queues and Inventories {ST 13}

Module 3: Financial and Actuarial Statistics (M3)

ST F1 Actuarial Statistics {ST 3, ST 4}
ST F2 Time Series Analysis {ST 12 (T)}
ST F3 Stochastic Models in Finance {ST 13}
ST F4 Survival Analysis {ST 14 (T)}
ST F5 Multiple Decrement Models in Insurance {ST F1}

Module 4: Bio-Statistics (M4)

ST B1 Survival Analysis {ST 14 (T)}
ST B2 Statistical Analysis of Clinical Trials {ST 12 (T), ST 15 (T)}
ST B3 Medical and Health Statistics {ST 3, ST 4}
ST B4 Statistical Methods in Micro-array Data Analysis {ST 9 (T)}
ST B5 Population Growth Models {ST 3, ST 4, ST 10, ST 11}
ST B6 Statistical Methods for Bio-computing {ST 5 (T), ST 13}

Module 5: Computational Statistics (M5)

ST C1 Statistical Simulations {ST 3, ST 4, ST 5 (T)}
ST C2 Computer-Intensive Statistical Methods {ST 5 (T)}
ST C3 Bayesian Computing {ST 5 (T)}
ST C4 Statistical Learning and Data Mining {ST 9 (T), ST 12 (T)}
ST C5 Statistics in Natural Language Processing {ST 3, ST 4}
ST C6 Statistical Methods for Bio-computing {ST 5 (T), ST 13}

The Departmental committee may decide (i) not to offer modules if sufficient faculty is not available and (ii) to offer two modules, where courses from these five modules will be merged.

(b) Elective Courses: (Depending on the availability of faculty elective courses will be offered from the following list.) A student may choose elective courses from the courses which are listed as compulsory for a module other than the module allotted to her/him or from courses offered by other Departments subject to the approval of the Head of the Department. A student may opt for a project (or two projects) in lieu of an elective course (or two elective courses), as described earlier.

List of Proposed Optional Courses

ST E01 Stochastic Models {ST 3, ST 4, ST 13}
ST E02 Empirical Processes {ST 7, ST 8, ST 13}
ST E03 Sequential Analysis {ST 10, ST 11}
ST E04 Nonparametric Inference {ST 10, ST 11}
ST E05 Discrete Data Analysis {ST 14 (T)}
ST E06 Stochastic Models in Epidemiology {ST 3, ST 4, ST 13}
ST E07 Advanced Time Series {ST I4 / ST F2}
ST E08 Longitudinal Data Analysis {ST 12 (T)}
ST E09 Financial Econometrics {ST I4 / ST F2}
ST E10 Directional Data Analysis {ST 9 (T)}
ST E11 Total Quality Management and Six Sigma {ST I2}
ST E12 Statistics in General Insurance {ST 3, ST 4}
ST E13 Generalized Linear Models {ST 12 (T)}
ST E14 Frailty Models {ST 14 (T)}

The Head of the Department may introduce additional elective courses on recommendations of the Departmental Committee. The syllabus of the electivel courses will be prepared by the concerned teacher and will be flexible to accommodate new developments in that area. Whenever such an elective course is floated, the concerned syllabus will be discussed and approved in the Departmental committee.

(c) (i) Question papers: In view of academic autonomy given to the Department, question papers are set by the teacher who teaches the course and these are moderated by a committee of experts, as indicated earlier.
(ii) Medium of instructions: English

7) Equivalence: The Head will take the decisions as and when the case arises.

8) University Terms: Time table for the beginning and end of the terms as announced by the University will be followed.

11) Qualification of Teacher: As per the University rules.

9) and 10) Subject wise detailed syllabus and recommended books: Detailed syllabi, along with the list of recommended books of the compulsory courses and some elective courses are given below.

***L**: Lectures; **T**: Tutorials.*

Tutorial session is equivalent to one Lecture wherein the students are expected to solve pre-circulated problems on specified topics.

ST 1: Mathematical Analysis (4 credits)

Unit I

Countability, supremum and infimum of sets of real numbers. Archimedean property, denseness property of rationals. Metric spaces, limit points and interior points of a set, open sets, closed sets etc. (These concepts will be introduced through metric spaces and R^n will be considered as a special case).
(12L + 3T)

Unit II

Compactness, Bolzano-Weierstrass theorem, Heine-Borel theorem. Sequences of real numbers, Cauchy sequence, limit superior, limit inferior, limit and convergence of a sequence of real numbers. Cauchy criterion for convergence. Completeness of R.
(12L + 3T)

Unit III

Series of real numbers, convergence of series, tests for convergence of series, absolute convergence, Cauchy product of two series and its convergence. Power series and radius of convergence, examples and problems on these concepts. **(12L + 3T)**

Unit IV

Continuous functions, uniform continuity, uniform convergence of sequences and series of functions, term by term differentiation and integration, applications to power series. **(12L + 3T)**
Total (48L + 12T)

Books Recommended

1. Apostol, T. M. (1975). Mathematical Analysis: A Modern Approach to Advanced Calculus. (Addison - Wesley)
2. Bartle R. G. & Sherbert D. R., (2007), Introduction to Real Analysis (Wiley
3. Bartle, R. G. (1976). Elements of Real Analysis (John Wiley)
4. Goldberg R. R. (1976). Methods of Real Analysis, (John Wiley & Sons, Inc.)
5. Rudin, W. (1985). Principles of Mathematical Analysis (McGraw - Hill) Student Edition)
6. Trench W. F. (2012). Introduction to Real Analysis. E-book.

ST 2 (T): Linear Algebra (3 credits)

Unit I

Determinants and their simple properties, inverses of partitioned matrices, special types of matrices, orthogonal and idempotent matrices, quadratic forms, vector spaces, inner product of vector spaces, linear dependence and linear independence of vectors, bases, an orthogonal basis, basis and dimension, properties and uses of a basis. **(12L + 3T)**

Unit II

Linear transformation and their matrix representations, rank of a matrix, linear equations, solution space and null space, generalized inverse, echelon forms, canonical forms, Gram-Schmidt orthogonalization, projection theorem. **(12L + 3T)**

Unit III

Characteristic roots of real matrices, right and left characteristic vectors, linear independence of characteristic vectors corresponding to distinct characteristic roots, algebraic and geometric multiplicities, spectral decomposition theorem, Cayley-Hamilton theorem. Definiteness of a real quadratic form, reduction of quadratic forms, simultaneous reduction of two quadratic forms, maxima and minima of ratios of two quadratic forms. **(12L + 3T)**

Total (36L + 9T)

Books Recommended

1. Bapat, R.B. (2011). Linear Algebra and Linear Models. Springer and Hindustan Book Agency
2. Bellman, R. (1970). Introduction to Matrix Analysis. (Tata McGraw Hill)
3. Graybill, F.E. (1961). Introduction to Matrices with Applications in Statistics (Wadsworth Pub. Co.)
4. Hohn, F. E. (1973). Elements of Matrix Algebra (Macmillan)
5. Ramachandra Rao, A. and Bhimasankaram, P. (2000). Linear Algebra. (Hindustan Book Agency)
6. Rao, C. R. (1995). Linear Statistical Inference and Its Applications. (Wiley Eastern)

7. Sadun L. (2008): Applied Linear Algebra: The Decoupling Principle. (American Mathematical Society) Second Edition.
8. Searle, S. R. (1982). Matrix Algebra Useful for Statistics. (John Wiley)

Additional Reference:

http://aix1.uottawa.ca/~jkhoury/app.htm

ST 2 (P): Linear Algebra Practical (2 credits)
Software such as R, SYSTAT, MATLAB etc will be used.

Unit I
1. Calculation of determinant of higher order by partitioning method.
2. Calculation of a rank of a matrix.
3. Calculation of equivalence canonical form by using elementary row and column operations.
4..Calculation of inverses of symmetric matrices of higher order by partitioning method.
5. Calculation of inverse of matrices of higher order.

Unit II
1. Calculation of Moore-Penrose inverse.
2. Calculation of a g- inverse.
3. Calculation of eigen values, eigen vectors,
4. Solution of simultaneous equations.
5. Spectral decomposition and powers of a matrix

ST 3: Univariate Probability Distributions (2 credits)

Unit I

Review of the concept of a random variable, c.d.f, characteristic properties of c.d.f. p.d.f., p.m.f., absolutely continuous and discrete distributions, m.g.f., p.g.f., quantiles and symmetry of a distribution, Stieltjes moment problem. Mixtures of probability distributions, decomposition of a c.d.f. into discrete and continuous c.d.f.s.

(12L+3T)

Unit II

Brief review of standard discrete and continuous distributions, truncated distributions (truncated binomial, truncated Poisson, truncated normal etc.)
Concept of hazard rate and cumulative hazard rates. Introduction to IFR, DFR, IFRA and DFRA classes of distributions. Transformations (monotone, non-monotone) of random variables and their distributions

(12L+3T)

ST 4: Multivariate Probability Distributions (4 credits)

Unit I

Random vectors, joint distributions, joint m.g.f., joint p.g.f., mixed moments, variance-covariance matrix, multivariate Normal and bivariate exponential distributions. **(12L + 3T)**

Unit II

Independence of random variables sums of independent random variables, convolutions, conditional expectation and variances, compound distribution.

Regression function, best predictor, best linear predictor, multiple and partial correlation coefficients. **(12L + 3T)**

Unit III

Sampling distributions of statistics from univariate normal random samples, such as linear and quadratic forms, Fisher Cochran theorem, non-central χ^2, non central t and F distributions.

(12L + 3T)

Unit IV

Order statistics: Probability Integral transformation, joint distribution of order statistic, distribution of r-th order statistic, 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.

(12L + 3T)

Total (48L + 12T)

Books Recommended

1. Berger, R. and Casella G. (2002) Statistical Inference (Duxbury Resource Center) Second Edition.
2. Cramer H., (1946) Mathematical Methods of Statistics (Springer).
3. Dasgupta, A. (2010) Fundamentals of Probability: A First Course (Springer)
4. Hogg, R. V. and Craig, T. T. (1978) Introduction to Mathematical Statistics (Collier-McMillan) Fourth Edition
5. Rao B.L.S. P. (2009). A First course in probability and Statistics, (World
6. Rao, C. R. (1995) Linear Statistical Inference and Its Applications (Wiley Eastern) Second Edition
7. Rohatgi, V. K. (1988) Introduction to Probability Theory and Mathematical Statistics (Wiley Eastern Scientific, Singapore)

ST 5 (T): Elements of Statistical Computing (4 Credits)

Unit I

Random number generation. Requisites of a good random number generator, methods of random number generation such as linear congruential, mixed

congruential and multiplicative congruential. Testing of random number generator, run test, Kolmogrov-Smirnov test, sign test, rank test, gap test, digit frequency test and serial correlation, selection of a random number generator. Methods of generating random variables such as inverse transforms, composition, convolution and acceptance –rejection. **(12L+3T)**

Unit II

Simple optimization method, direct search, grid search, interpolatory search, gradient search. Newton-Raphson method, Muller's method, Aitken's extrapolation. Simple problems, case studies and applications to bacterial chemostat, industrial high temperature oven etc.

(12L+3T)

Unit III

Methods to compute integrals: quadrature formula, double integration, singularity, Gaussian integration. Monte Carlo Methods: Monte Carlo integration and Simple case studies, applications of Monte Carlo methods to compute expected values of functions of random variables, such as Laplace transform, fourier transform etc. Case studies such as change in entropy, R-C network.

(12L+3T)

Unit IV

Approximating probabilities and percentage points in selected probability distribution, verification of WLLN and CLT using random number generator, simulating null distribution of various test statistics. simple applications and case studies. **(12L+3T)**

Total (48 L+ 12T)

Books Recommended

1. Kennedy W. J. & Gentle J. E. (1980). Statistical Computing (Marcel Dekker)
2. Krishnamurthy V. & Sen (1993). Numerical Algorithm Computation in Science and Engineering. (Affiliated East West Press), Second edition
3. Law, A.M. and Kelton, W.D. (2000). Simulation, Modeling and Analysis Third Edition. (Tata McGraw Hill)
4. Rajaraman V. (1993). Computer Oriented Numerical Methods, (Prentice-Hall). Fourth edition
5. Ripley B. D.(1987) Stochastic Simulation. (John Wiley)
6. Ross, S. (2000). Introduction to Probability Models. (Academic Press)
7. Schilling, R.J. and Harris, S.L.(2002). Applied Numerical Methods for Engineers Using MATLAB and C, (Thomson and Brooks/Cole, Singpore)
8. Thisted R. A.(1988). Elements of Statistical Computing, (Chapman and Hall)

ST 5 (P): Elements of Statistical Computing Practical (3 Credits)

Elementary Statistics as listed below using software packages: SYSTAT, R, MINITAB, S^+, MATLAB, MATHEMATICA

Unit I

1. Computation of integrals by Riemann & RS sums.
2. Calculation of double integrals.
3. Calculation of p-value for standard tests of hypotheses.
4. Plotting of curves (standard & non standard), Box Plot, Q-Q plot

Unit II

1. Limits of functions
2. Computing integrals by statistical methods
3. Computing expectations of complicated functions, mean and variance of estimators
4. Empirical level of significance and empirical power
5. Empirical confidence coefficient of a confidence interval

Unit III

1. Test for randomness, Sign test, Kolmogrov-Smirnov test, Wilcoxon signed rank test.
2. Iterative solution of a function using Newton-Raphson method, Muller's method.
3. Verification of WLLN for i.i.d. random variables.
4. Verification of CLT for i.i.d. random variables.

Books Recommended

1. Boswell, M.T., Gore, S.D., Patil G. P. and Tallie C. (1993). Handbook of Statistics - Vol 4 (The Art of Computer Generation of Random Variables)
2. Kennedy W. J. & Gentle J. E. (1980). Statistical Computing (Marcel Dekker)
3. Krishnamurthy V. & Sen (1993). Numerical Agorithm Computation in Science and Engineering 2ndEd. (Affiliated East West Press)
4. Purohit, S.G. ,Gore, S.D. and Deshmukh, S.R., (2008): Statistics using R: Narosa Publications
5. Rajaraman V. (1993). Computer Oriented Numerical Methods, (Prentice-Hall). Fourth Edition.
6. Ripley B. D. (1987). Stochastic Simulation (Wiley)
7. Ross, S. (2000). Introduction to Probability Models.(Academic Press)
8. Thisted R. A.(1988). Elements of Statistical Computing, (Chapman and Hall)

ST 6: Calculus (4 credits)

Unit I

Review of calculus of one variable: differentiability, mean value theorem and Taylor series expansion. Functions of several variables: Continuity, directional derivatives, differentials of functions of several variables, the gradient vector **(12L+3T)**

Unit II

Differentials of composite functions and the chain rule, the mean value theorem, a sufficient condition for the existence of the differential, partial derivatives of higher order and Taylor's formula. Applications of partial differentiation, Jacobians. **(12L+3T)**

Unit III

Inverse function theorem (without proof), implicit function theorem (without proof), extremum problems. Riemann and Riemann–Stieltjes integrals, integration by parts, mean value theorem. **(12L+3T)**

Unit IV

Improper Riemann – Stieltjes integrals: Improper integrals of first and second kind for one variable. Uniform convergence of improper integrals, differentiation under the sign of integral – Leibnitz rule. Multiple Integrals and evaluation of multiple integrals by repeated integration. Mean value theorem for multiple integrals. (12L+3T)

Total (48 L+ 12T)

Books Recommended

1. Apostol T.M. (1975). Mathematical Analysis: A modern approach to advanced calculus. (Addison-Wesley)
2. Bartle, R. G. (1976). Elements of Real Analysis. (John Wiley)
3. Kreyszig, E. (1975). Advanced Engineering Mathematics (Wiley Eastern)
4. Rudin, W. (1985). Principles of Mathematical Analysis (McGraw-Hill)
5. Trench W. F. (2012). Introduction to Real Analysis. E-book

ST 7: Probability Theory (4 credits)

Unit I

Classes of sets, fields and sigma-fields, limit of sequences of subsets, sigma-field generated by a class of subsets, Borel fields. Probability measure on a sigma-field, probability space, continuity of a probability measure. Real and vector-valued random variables (r.v.s). **(12L+3T)**

Unit II

Distribution functions (d.f.), discrete r.v.s, r.v.s of the continuous type, decomposition of a d.f. Expectation of a real r.v. Linear properties of expectations, Inequalities, Characteristic functions, their simple properties, uniqueness theorem (without proof). **(12L+3T)**

Unit III

Independence of two events and n (> 2) events, sequence of independent events, independent classes of events, π-system and λ-system of events, Dynkin's theorem (without proof) independence of r.v.s, Borel zero-one law, Borel - Cantelli Lemma, Kolmogorov zero- one law.

(12L+3T)

Unit IV

Convergence of a sequence of r.v.s., convergence in probability, almost sure convergence, convergence in quadratic mean and convergence in distribution, and their interrelationships, examples and problems on these concepts.

(12L+3T)

Total (48L+ 12T)

ST 8: Limit Theorems in probability (2 credits)

Unit I

Cramer's theorem (Slutsky's theorem) on algebraic properties of convergence in distribution and convergence in probability. Monotone convergence theorem and dominated convergence theorem, Fatou's lemma. Problems and examples on these concepts. **(12L+3T)**

Unit II

Laws of large numbers, weak law of large numbers , strong law of large numbers. Khintchin weak law of large numbers, Kolmogorov strong law of large numbers (without proof). ,Continuity theorem for characteristic functions. Lindeberg's CLT (without proof) and its particular cases: Liapounov's theorem, CLT for i.i.d random variables.

(12L+3T)

Total (24L + 6T)

Books Recommended

1. Athreya K. B. and Lahiri S. (2006). Probability Theory Vol 41, Trim series, (Hindustan Book Agency).
2. Bhat, B. R. (1985). Modern Probability Theory (New Age International)
3. Billingsley, P. (1986). Probability and Measure (John Wiley)
4. Feller, W. (1969). Introduction to Probability and its Applications Vol.II (Wiley Eastern Ltd.)
5. Gnedenko, B. V. (1988). Probability Theory (Mir.Pub.)
6. Gut, Allan (2005), Probability: A Graduate Course. (Springer, New York)
7. Loeve, M. (1978). Probability Theory (Springer Verlag). Fourth edition.

ST 9 (T): Multivariate Analysis (4 Credits)

Unit I

Exploratory multivariate data analysis, sample mean vector, sample dispersion matrix, correlation matrix, graphical representation, means, variances, covariances, correlations of linear transforms. Introduction to principle component analysis, factor analysis. Canonical correlation coefficients and canonical variables. **(12L + 3T)**

Unit II

Cluster Analysis and multidimensional scaling.

Multivariate normal distribution, pdf and mgf, singular and nonsingular normal distributions, distribution of a linear form and a quadratic form of normal variables, marginal and conditional distributions.

(12L + 3T)

Unit III

MLE's of the parameters of multivariate normal distribution and their sampling distributions, properties of the Wishart Distribution, tests of hypothesis about the mean vector of a multivariate normal distribution, Hotelling's T^2-statistic and its distribution, applications of Hotelling's T^2-statistic. Goodness of fit of multivariate normal distribution, simultaneous confidence interval for the linear functions of the mean. Tests of significance for multiple and partial correlation coefficients.

(12L + 3T)

Unit IV

Classification problem. Discriminant analysis, Mahalanobis D^2-statistic, methods and applications of MANOVA (without derivation of the distribution of Wilks' lambda).

Likelihood ratio tests, introduction to Non-Gaussian multivariate distributions such as multivariate beta, t, F distributions, Introduction to copula and its applications.

Introduction to directional data analysis.

(12L + 3T)

Total (48L + 12T)

Books Recommended

1. Anderson, T. W. (1984). Introduction to Multivariate Analysis (John Wiley)
2. Fang ,K., Kotz, S., Ng K. W. (1990). Symmetric Multivariate and Related Distributions (Chapman and Hall)
3. Härdle, W. K. & Simar, L. (2012). Applied Multivariate Statistical Analysis (Springer, New York)

4. Härdle, W. K., Hlávka, Z. (2007). Multivariate Statistics: Exercises and Solutions (Springer, New York)
5. Johnson R.A. & Wichern, D.W. (1988). Applied Multivariate Statistical Analysis (Prentice Hall Inc.)
6. Kotz, S., Balakrishnan N. and Johnson N. L. (2000). Continuous Multivariate Distributions, Volume 1, Models and Applications (John Wiley & Sons)
7. Kshirsagar, A. M. (1983). Multivariate Analysis (Marcel Dekker)
8. Mardia ,K. V. & Jupp, P. E. (2000), Directional Statistics (John Wiley & Sons)
9. Morrison, D.F. (1990). Multivariate Statistical Methods (McGraw Hill Co.)
10. Rao, C. R. (1995). Linear Statistical Inference and its Applications (Wiley Eastern)
11. Tim, N. H. (2002), Applied Multivariate Analysis (Springer, New York)

ST 9 (P): Multivariate Analysis Practical (2 Credits)

Unit I

1. Graphical representation of multivariate data
2. Principal Component Analysis
3. Factor Analysis
4. Cluster Analysis
5. Canonical Correlations

Unit II:

1. Model Sampling from multivariate normal distribution
2. Likelihood ratio tests
3. Applications of Hotelling's T^2
4. MANOVA
5. Discriminant Analysis

Books Recommended

Khattree, R. & Naik D. N. (2003), Applied Multivariate Statistics with SAS Software, Second Edition (SAS Institute and Wiley), Chapter 2

ST 10: Point &Interval Estimation (3 credits)

Unit I

Point estimation, sufficiency, sufficient partition, Neyman Factorization theorem. Minimal Sufficiency, one parameter exponential family, canonical form of exponential family, convexity property, minimal sufficiency in one parameter exponential family, ancillary statistic, Basu's theorem .

(12L + 3T)

Unit II

Pitman family of distributions, minimal sufficiency in Pitman family, completeness of family, complete statistic, completeness of one parameter

exponential family and Pitman family. Unbiased Estimator, estimability of parametric functions, Cramer-Rao inequality, uniformly minimum variance unbiased estimators, Rao-Blackwell Theorem, Lehmann Scheffe theorem,

(12L + 3T)

Unit III

Multi-parameter exponential family, complete sufficient statistic for multi-parameter exponential family. Confidence sets and intervals, shortest expected length confidence intervals, relation with testing of hypotheses.

(12L + 3T)

Total (36L + 9T)

ST 11: Testing of Hypotheses (3 credits)

Unit I

Test function, most powerful test function, Neyman- Pearson lemma for most powerful test functions. Uniformly most powerful tests, UMP tests for simple null against one sided alternative for one parameter exponential class of densities, UMP tests for one sided null against one sided alternative for one parameter exponential class of densities. **(12L + 3T)**

Unit II

Non existence of UMP tests for simple null against two sided alternative for one parameter exponential class of densities. Monotone likelihood ratio property, UMP tests for one sided null against one sided alternative for distributions having MLR property, UMP Unbiased tests in exponential family.

(12L + 3T)

Unit III

Non parametric tests: derivation of the test statistics and its null distribution. Tests of goodness of fit, Kolmogorov-Smirnov one-sample test, Applications of the Kolmogorov-Smirnov one-sample test, Lilliefors's test for normality. Tests based on sample quantiles: Sign test, Wilcoxon signed-rank test. General two-sample problem: Wald-Wolfowitz runs test, Kolmogorov-Smirnov two sample test, Mann-Whitney U test. Linear rank tests for the location problem: Wilcoxon rank-sum test. Linear rank tests for the scale problem: Sukhatme test. Tests of the equality of k Independent samples: Kruskal-Wallis test.

(12L + 3T)

Total (36L + 9T)

Books Recommended

1. Casella, G. and Berger, R. L. (2002) Statistical Inference. (Duxbury Advanced Series). Second Edition.
2. Dudewicz, E. J. and Mishra, S.N.(1988) Modern Mathematical Statistics. (John Wiley)
3. Gibbons, J.D. and Chakraborty, S. (2003). Nonparametric Statistical Inference. (Marcel Dekker Inc., New York), Fourth edition

4. Kale, B.K. (1999) A First Course on Parametric Inference. (Narosa).
5. Lehmann, E. L. (1986) Testing of Statistical hypothesis (John Wiley)
6. Lehmann, E. L. (1988) Theory of Point Estimation (John Wiley)
7. Lehmann, E.L. and Casella, G. (1998). Theory of Point Estimation. (Springer-Verlag, New York) Second edition
8. Rohatgi, V. K. and Saleh, A.K. Md. E. (2001) Introduction to Probability and Statistics. (John Wiley & Sons)
9. Shao, J. (2003). Mathematical Statistics. (Springer-Verlag, New York) Second edition.
10. Silvey, S. D. (1975) Statistical Inference. (Chapman and Hall)

. **ST 12 (T): Regression Analysis (4 credits)**

Unit I

Simple linear regression, assumptions, inference related to regression parameters, standard error of prediction, tests on intercepts and slopes, extrapolation, diagnostic checks and correction: graphical techniques, tests for normality, uncorrelatedness, homoscedasticity, lack-of-fit testing, polynomial regression, transformations on Y or X (Box-Cox, square root, log etc.), method of weighted least squares, inverse regression.

(12L+3T)

Unit II

Multiple regression: standard Gauss Markov setup, least squares (LS) estimation, variance-covariance of LS estimators, estimation of error variance, regression ananlysis with correlated observations, LS estimation with restriction on parameters. Simultaneous estimation of linear parametric functions. Test of hypotheses for one and more than one linear parametric functions, confidence intervals and regions. Variable selection problems, multicollinearity and ridge regression, LASSO.

(12L+3T)

Unit III

Logistic Regression: Logit transform. ML estimation. Tests of hypotheses, Wald test, LR test, score test, test for overall regression, multiple logistic regression, forward, backward method, interpretation of parameters, relation with categorical data analysis. Generalized Linear model: link functions such as Poisson, binomial, inverse binomial, inverse Gaussian, gamma. **(12L+3T)**

Unit IV

Non Linear Regression (NLS): Linearization transforms, their uses & limitations, examination of non linearity, initial estimates, iterative procedures for NLS, grid search, Newton-Raphson, steepest descent, Marquardt's methods. Introduction to semiparametric regression models, additive regression models. Introduction to nonparametric regression methods.

(12L+3T)

Total (48 L+ 12T)

Books Recommended

1. Cameron, A. C. and P. K. Trivedi (1998). Regression Analysis of Count Data (Cambridge)
2. Draper, N. R. and Smith, H. (1998). Applied Regression Analysis (John Wiley) Third Edition.
3. Hosmer, D. W. and Lemeshow, S. (1989). Applied Logistic Regression (Wiley).
4. Kleinbaum, D. G. & Klein, M. (2002). Logistic Regression: A Self-Learning Text (Springer).
5. McCullagh, P. and Nelder, J. A. (1989). Generalized Linear Models (Chapman & Hall).
6. Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003). Introduction to Linear Regression Analysis (Wiley).
7. Neter, J., W., and Kutner, M. H. (1985). Applied Linear Statistical Models (Wiley).
8. Ratkowsky, D. A. (1983). Nonlinear Regression Modelling, Marcel Dekker, London.
9. Ruppert, D., Wand, M. P. and Carroll, R. J. (2003) Semiparametric Regression (Cambridge University Press).
10. Seber, G. E. F. and Wild, C. J. (1989). Nonlinear Regression (Wiley).
11. Weisberg, S. (2005). Applied Liner Regression (Wiley).
12. Yan, X. and Su, X. G. (2009). Linear Regression Analysis: Theory & Computing, (World Scientific).

ST 12 (P): Regression Analysis Practical (2 Credits)

Unit I

1. Simple Linear Regression
2. Multiple Regression
3. Variable Selection Problem
4. Multicollinearity and Ridge Regression

Unit II

1. Nonlinear regression
2. Nonparametric regression
3. Logistic regression (binary and multiple)
4. Poisson/Negative binomial regression

ST 13: Stochastic Processes (4 Credits)

Unit 1

Notion of stochastic processes, Markov chain, one step transition probabilities, Chapman-Kolmogorov equations, evaluation of higher step transition

probabilities, classification of states, periodicity of a Markov chain, concept of closed class, minimal closed class, stationary distribution. Some examples such as gamblers ruin problem and one dimensional random walk. Concept of absorption probabilities. Use of these to compute probability of winning the game by a gambler having initial capital 'a'.

(12L+3T)

Unit 2

Branching process, classification of states, identification of criticality parameter. extinction probability, relationship between criticality parameter and extinction probability of the process. Expression for mean and variance of the process. Method to compute the extinction probability. Some epidemiological applications.

Introduction to Markov chain in continuous time, concept of intensity rate, relationship between intensity matrix and transition probability matrix. Kolmogorov's forward and backward equations.

(12L+3T)

Unit 3

Introduction to birth process, birth and death process. Linear birth and death process. Growth model with immigration and related results. Expression for mean and variance of a birth process and, birth and death process. Simple applications of these processes.

(12L+3T)

Unit 4

Poisson process, two definitions and their equivalence. Distribution of inter arrival times, conditional joint distribution of inter arrival times. Compound Poisson process. Some applications.

Introduction to renewal process, relationship with Poisson process, key and elementary renewal theorems associated with renewal processes. Some applications. Introduction to Weiner process and its simple properties. (**12L+3T)**

Total (48 L+ 12T)

Books Recommended

1. Bhat B.R. (2000). *Stochastic Models: Analysis and Applications*, New Age International.
2. Medhi, J. (1982) *Stochastic Processes*, Wiley Eastern.
3. Pinsky M. A. and Karlin, S. (2010). *An Introduction to Stochastic Modeling*, 4th Edn. Academic Press.
4. Ross, S. (2007). *Introduction to Probability Models*, 9th Edn., Academic Press.

Additional Books for Reference

1. Feller, W. (1972). *An Introduction to Probability Theory and its Applications*, Vol. 1, Wiley Eastern.

2. Hoel, P.G. Port, S.C. & Stone, C.J. (1972). *Introduction to Stochastic Processes*, Houghton Mifflin.
3. Karlin, S & Taylor, H.M. (1975). *A First Course in Stochastic Processes* (Second. Edition), Academic Press.
4. Serfozo, R. (2009). *Basics of Applied Stochastic Processes*, Springer.

ST 14 (T): Asymptotic Inference (4 Credits)

Unit 1

Consistent estimator for real and vector valued parameter, joint consistency and marginal consistency and equivalence between them, invariance property under continuous transformation of consistent estimators for real and vector valued parameters

Methods for generating consistent estimators for real and vector valued parameters:

(i) method of moments: generating consistent estimator using weak law of large numbers

(ii) method of percentiles: consistency of sample percentiles for population percentiles

Comparison of consistent estimators based on (i) mean squared error, (ii) minimum sample size required by the estimator to attain certain level of accuracy

(12 L + 3 T)

Unit 2

Consistent and asymptotically normal (CAN) estimators: Definition of CAN estimator for real and vector valued parameter, invariance property of CAN estimator (delta method), and variance stabilizing transformation

Generating CAN estimators for real and vector valued parameter:

(i) method of moments: generating CAN estimator using central limit theorem

(ii) method of percentiles: generating CAN estimator using asymptotic distribution of sample percentiles

Comparison of CAN estimators based on asymptotic variance **(12 L + 3 T)**

Unit 3

Maximum Likelihood Estimation:

(i) MLE in case of restricted parameter space, MLE of parametric function, Inconsistent MLEs, MLEs in irregular cases.

(ii) Asymptotic distribution of MLE in special class of distributions: Cramer regularity conditions, Cramer- Huzurbazar theorem, Extension to vector-valued parameters, exponential family with natural parameters (canonical set up), its extension to the (usual) exponential family set up and multinomial distribution

(iii) Super efficient estimators, Hodges-Lecam shrinkage technique

(12 L + 3 T)

Unit 4

Asymptotic theory of tests of hypotheses: Tests based on MLEs. Likelihood ratio tests, asymptotic distribution of log likelihood ratio, Bartlett correction, Wald test, Score test, Pearson's chi-square test and LR test, Consistent Test

Comparison of tests: Asymptotic relative efficiency of tests, Pitman's asymptotic relative efficiency.
Asymptotic confidence intervals: Construction and examples
Applications: Applications to categorical data analysis, tests for independence for three dimensional contingency tables.

(12 L + 3 T)
Total (48 L+ 12T)

Books Recommended

1. Ferguson, T.S. (1996), *A Course in Large Sample Theory*, Chapman & Hall, London, Chapters 1, 2, 4.
2. Gupta Anirban Das (2008), *Asymptotic Theory of Statistics and Probability*, Springer, New York. Chapters 1, 4, 7, 13, 16, 21, 22, 27.
3. Kale, B. K. (1999), *A First Course in Parametric Inference*, Narosa, Chapters 5, 6, 7, 9.
4. Lehmann, E. L. and Casella, G. (1998), *Theory of Point Estimation,* Springer, New York, Chapter 6.
5. Rao, C. R. (1995). *Linear Statistical Inference and its Applications*, Wiley, New York, Chapters 5c-5g, 6a-6e.

ST 14 (P): Asymptotic Inference (2 Credits)

R will be used for the practical.

Unit 1

1. Verification of consistency of the estimators
2. Verification asymptotic normality of the estimators
3. Comparing methods of estimation, MSE and sample size considerations
4. Maximum likelihood estimation under various set up (includes censoring)

Unit 2

1. Analysis of higher dimensional contingency tables.
2. Verification of consistency of test procedure
3. Likelihood ratio tests
4. Verification of asymptotic distribution of likelihood ratio test statistic in case of Cramer family

ST – 15(T): Design and Analysis of Experiments (4 Credits)

Unit 1

Review of randomization, replication and local control. Analysis of one way classification model. Analysis of two way classification model with equal number of observations per cell with and without interactions. Analysis of two way classification model with unequal number of observations per cell without interactions. Notion of connectedness, balance and orthogonality. Analysis of BIBD. Analysis of covariance in one way and two way classification models. Testing of hypotheses for estimable parametric functions.

(12L+3T)

Unit 2

Analysis of 2^k full factorial experiments: diagrammatic presentation of main effects and first and second order interactions, model, analysis of single as well as more than one replicates, using ANOVA. Total confounding of 2^k design in 2^p blocks, $p \geq 2$. Partial confounding in 2^p blocks, $p = 2, 3$. Fractional factorial experiments. Resolution of a design, (III, IV & V), aberration of a design.

(12L+3T)

Unit 3

Analysis of 3^2 designs: contrasts for linear and quadratic effects, analysis of 3^2 design, confounding and fractional experiments in 3^2 design. Response surface methodology (RSM): linear and quadratic model, stationary point, central composite designs(CCD), ridge systems, multiple responses, concept of rotatable designs, Box-Behnken design, optimality of designs, simplex lattice designs, simplex centroid designs. **(12+3T)**

Unit 4

Taguchi methods: concept of noise factors, concept of loss function, S/N ratio, orthogonal arrays, triangular tables, linear graphs, inner and outer arrays. Random effect models and mixed models.

(12L+3T)

Total (48 L+ 12T)

Books Recommended

1. Dean, A. and Voss, D. (1999). *Design and Analysis of Experiments*, Springer.
2. George E. P. Box, Draper N.R. (1987). *Empirical Model-Building and Response Surfaces*, Wiley.
3. Hicks, C.R., Kenneth V. and Turner, Jr. (1999). *Fundamental Concepts in the Design of Experiments*, Oxford University Press.
4. Kshirsagar A.M. (1983). *Linear Models*, Marcel Dekker
5. Montgomery, D.C. (2001). *Design and Analysis of Experiments*, Wiley.
6. Ogawa, J. (1974). *Statistical Theory of the Analysis of Experimental Design*, Marcel Dekker.

7. Phadke, M.S. (1989). Quality Engineering using Robust Design, Prentice Hall, Englewood Cliffs, New Jersey
8. Wu, C.F. Jeff and Hamada M. (2000). Experiments: Planning ,Analysis and Parameter Design Optimization, John Wiley and Sons

ST 15 (P): Design and Analysis of experiments (2 Credits)

Unit 1

1. One way classification. Multiple comparison tests
2. Two way classification with equal number of observations per cell (Model with interaction). Two way classification with unequal number of observations per cell (Model without interaction)
3. Analysis of LSD and BIBD.
4. Analysis of covariance in one way and two way model.
5. 2^k factorial experiments, Analysis of single replicate of 2^k.
6. Total and partial confounding in 2^k factorial experiments.

Unit 2

7. Analysis of 2^k fractional factorial experiments
8. Analysis of 3^2 factorial experiments
9. Random effect and mixed models
10. Analysis of first and second order response surface model. Central composite design. Contour and surface plots, Box-Behnken design
11. Small composite designs, optimality of designs, simplex lattice designs, simplex centroid designs.
12. Taguchi methods: S/N ratio, orthogonal arrays, triangular tables, linear graphs, inner and outer arrays.

ST 16 (T): Sampling Methods (4 credits)

Unit 1

Review of basic methods of sample selection, simple random sampling with replacement (SRSWR), simple random sampling without replacement (SRSWOR). Use of supplementary information for estimation, ratio and regression estimators with their properties and generalizations, systematic sampling. **(12L+3T)**

Unit2

Review of stratification: Allocation problems and estimation problems.
Formation of strata and number of strata. method of collapsed strata, Post stratification, probability proportional sampling with and without replacement, estimation problems, Hansen-Horwitz estimator and its properties, Horwitz-Thompson estimator and its properties, Midzuno-Sen method. **(12L+3T)**

Unit 3

Cluster sampling, multistage-sampling. Double sampling procedures and their ratio and regression estimators, stratification estimator. Multiphase sampling. **(12L+3T)**

Unit 4

Non-sampling errors, response and non-response errors, Hansen and Hurwitz's model and their treatments, methods of imputation, randomized response, Warner's model, Franklin's model. Jackknife technique. **(12L+3T)**

Total (48 L+ 12T)

Books Recommended

1. Cochran, W.G. (1984). *Sampling Techniques*, Wiley.
2. Des Raj and Chandhok, P. (1998). *Sample Survey Theory*, Narosa.
3. Murthy, M.N. (1977). *Sampling Theory and Methods*, Statistical Publishing Society
4. Okafore, C (2002). *Sample survey Theory with Applications*, Snaap Press Ltd.
5. Singh, D. and Chaudhary F.S (1986).*Theory and Analysis of Sample Survey Designs*, Wiley Eastern Limited.
6. Singh, S. (2003). *Advance Sampling Theory and Applications* (Volume I and II), Kluwer Academic Publishers.
7. Sukhatme P.V, Suktatme, B.V.,Sukhatme S. and Asok C. (1984). *Sampling Theory of Surveys with Applications*, Indian Society for Agricultural Statistics, New Delhi.

ST 16 (P): Sampling Methods (2 credits)

Unit 1

1. SRSWOR, SRSWR,
2. Stratified random sampling, various kinds of allocation, post stratification, use of auxiliary information.
3. Ratio and regression methods of estimation.
4. Probability proportional to sizes sampling design.
5. Double sampling.

Unit 2

6. Two stage sampling.
7. Systematic sampling.
8. Cluster sampling
9. Randomized response technique.

10. Two practicals of consolidated nature each one of which would use theory of one or more of the above topics or a sampling problem under a real set up (including estimation)

ST-17: Bayesian Inference (4 credits)

Unit 1

Basics of minimaxity. Subjective and frequentist probability, Bayesian inference, prior distributions, posterior distribution, loss function, principle of minimum expected posterior loss, quadratic and other common loss functions, advantages of being a Bayesian. Improper priors, common problems of Bayesian inference, Point estimators. Bayesian HPD confidence intervals, testing, credible intervals, prediction of a future observation,

(12L + 3 T)

Unit 2

Bayesian analysis with subjective prior, robustness and sensitivity, classes of priors, conjugate class, neighborhood class, density ratio class, different methods of construction of objective priors: Jeffrey's prior, probability matching prior, conjugate priors and mixtures, posterior robustness: measures and techniques

(12L + 3 T)

Unit 3

Model selection and hypothesis testing based on objective probabilities and Bayes factors

Large sample methods**:** Limit of posterior distribution, consistency of posterior distribution, asymptotic normality of posterior distribution

(12L + 3 T)

Unit 4

Bayesian Computations: Analytic approximation, E-M Algorithm, Monte Carlo sampling, Markov Chain Monte Carlo Methods, Metropolis-Hastings Algorithm, Gibbs sampling, examples, convergence issues

(12L + 3 T)

Total (48 L+ 12T)

Books Recommended

1. Bolstad, W. M. (2007). *Introduction to Bayesian Statistics*, 2nd Edn. Wiley.
2. Christensen R, Johnson, W., Branscum A. and Hanson T. E. (2011). *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*, Chapman & Hall.
3. Congdon, P. (2006). *Bayesian Statistical Modeling*, Wiley
4. Ghosh, J. K., Delampady M. and T.Samantha (2006). *An Introduction to Bayesian Analysis: Theory & Methods*, Springer.
5. Jim, A. (2009). *Bayesian Computation with R*, 2nd Edn, Springer.
6. Lee, P. M. (2004). *Bayesian Statistics: An Introduction*, Hodder Arnold.

7. Rao. C.R. and Day. D. (2006). *Bayesian Thinking, Modeling & Computation*, *Handbook of Statistics*, Vol. 25. Elsevier
8.

Elective Courses

The syllabi of some module specific and elective courses, which are being taught during the last few years is given below. These module specific and elective courses will be taught in the coming years, depending on the availability of the faculty. All these courses are of 4 credits.

Module Specific and Elective Courses

ST P1: Measure Theory and Probability

Unit 1

Ring, σ-ring, Measure, measure space, Caratheodory Extension theorem, Lebesgue measure. **(12L + 3T)**

Unit 2

Integral of a measurable function with respect to a measure, its properties. Hahn – Jordan decomposition, Lebesgue decomposition, Radon – Nikodym derivative. product measure, Fubini's theorem. **(12L + 3T)**

Unit 3

Convergence in measure, almost everywhere convergence, Kolmogorov inequality. Kolmogorov three series criterion and strong law of large numbers. Introduction to weak convergence. **(12L + 3T)**

Unit 4

Conditional probability and conditional expectations, their simple properties. Martingales, martingale convergence theorems (SLLN, CLT). **(12L + 3T)**

Total (48 L+ 12T)

Books Recommended

1. Ash, R.B. (1972). *Real Analysis and Probability.* Academic Press
2. Athreya, K.B. and Lahiri, S.N. (2006). *Measure Theory and Probability Theory*. Springer.
3. Billingsley, P. (1986). *Probability and Measure.* John Wiley
4. Taylor, J. C. (1997). *Introduction to Measure and Probability.* Springer.
5. Williams, D. (1991). *Probability with Martingales*. Cambridge University Press.

Additional Books for Reference

1. Ash, R. B. (2000). *Probability & Measure Theory*. Academic Press.
2. Dudley, R. M. (2004). *Real Analysis and Probability*. Cambridge University Press.
3. Halmos, P. R. (1978). *Measure Theory*, Springer.

ST P2: Advanced Stochastic Processes

Unit 1

Markov Chains: Taboo probabilities and ratio limit theorems. Invariant measures: [1]. ch. 11, Sections 1-4. **(12L + 3T)**

Unit 2

Markov Sequences: Definitions, transition densities, stationary distribution, normal Markov sequences [2] Ch. VI Section 11, Ch. III Section 8. Markov pure jump processes [4], [1]. **(12L + 3T)**

Unit 3

Stochastic processes (general Theory): Probability spaces appropriate for stochastic processes, Kolmogorov's extension theorem. (only sketch of the proof), separability, progressive measurability and strong Markov property of stochastic processes [3], Ch. 4 Sections 1 and 2. **(12L + 3T)**

Unit 4

Diffusion processes: Definition, elementary properties, infinitesimal parameters, standard process and Dynkin's theorem, continuity and non-differentiability of diffusion processes. Modeling based on diffusion processes, standard Brownian motion, Ornstein-Ulhenbeck process and other processes, transformation of processes, distribution of hitting times and related problems, scale function and speed density, Kolmogorov's backward differential equations, forward differential equations (without proof), transition density and stationary distribution of a diffusion process. **(12L + 3T)**

Total (48 L+ 12T)

Note: Section and chapter numbers are from the book by Karlin, S. and Taylor H.M. (1981). *A Second course in Stochastic Processes.* Academic Press.

Books Recommended

1. Adke, S.R. & Manjunath, S. M. (1984). Finite Markov Processes, Wiley.
2. Ash, R.B. and Gardner, M. F. (1975). *Topics in Stochastic Processes*. Academic Press.
3. Athreya, K.B. and Lahiri, S.N. (2006). *Measure Theory and Probability Theory*. Springer.

4. Feller, W. (1969). *An Introduction to Probability Theory*. (Vol. II)
5. Karlin, S. and Taylor H.M. (1981). *A Second course in Stochastic Processes.* Academic Press.

ST P3: Inference for Stochastic Processes (4 Credits)

Unit 1

Inference in Markov chains, estimation of transition probabilities, testing for order of a Markov chain, estimation of functions of transition probabilities, parametric models and their goodness of fit.

Markov sequences, estimation of parameters based on likelihood and conditional least squares, auto-regressive time series. Models for higher order Markov chains. (Raftery's long memory model). Statement of martingale strong law of large numbers and CLT for martingales, CAN property of the MLEs from a general sequence of dependent random variables, Fisher information. Applications to Markov chains and sequences. **(12L + 3T)**

Unit 2

Inference for Poisson process. Likelihood of Poisson and other pure Jump Markov processes from first principles, CAN property of MLEs, testing for a Poisson process, Inference for non-homogeneous Poisson process.

Inference for parametric pure jump processes, such as birth process, birth-death process, birth-death-immigration processes. **(12L + 3T)**

Unit 3

Diffusion processes and their likelihood, properties of estimators (without proof)

Branching processes: ergodic and non-ergodic processes, inconsistency of MLE/moment estimators, properties of estimators on the non-extinction path, estimation of asymptotic distribution theory. **(12L+ 3T)**

Unit 4

Elements of semi-parametric and non-parametric analysis, theory and applications of optimal estimating functions, estimation of transition and stationary density, intensity function of a counting process. Methods based on estimating functions, panel data, introduction to spatial models. **(12L + 3T)**

Total (48 L+ 12T)

Books Recommended

1. Adke, S.R. and Manjunath.S.M. (1984). *An introduction to Finite Markov Processes*, Wiley Eastern.
2. Basawa, I.V. and Prakasa Rao, B.L.S.(1980). *Statistical Inference for Stochastic Processes*, Academic Press.
3. Bhat, B. R. (2000). *Stochastic Models: Analysis and Applications*. New Age International.

4. Billingsley, P. (1962). *Statistical Inference for Markov Chains*, Chicago University Press.
5. Guttorp, P. (1991). *Statistical Inference for Branching Processes*, Wiley.
6. Guttorp, P.(1995). *Stochastic Modelling for Scientific Data*, Springer.
7. Prakasa Rao, B.L.S. and Bhat, B.R.(1996). *Stochastic Processes and Statistical Inference*, New Age International.
8. Rajarshi M.B, (2013). *Inference for Discrete Parameter Stochastic Processes*, Springer India.

ST I1: Optimization Techniques

Unit 1

Linear Programming: Review of simplex algorithm and simplex method, artificial variable technique methods. degeneracy, duality in linear programming, duality theorems, dual simplex method with justification.

Integer linear programming problem: pure and mixed integer programming problem, Gomory's all Integer programming method. Fractional cut method- all integer and mixed integer linear programming problem, branch and bound method, dynamic programming, sensitivity. Bellman's optimality principle.

(12L+3T)

Unit 2

Review of transportation and assignment problems: Balance and degeneracy in transportation problem. Transshipment problem, duality theory of testing optimality of solution in transportation problem and transshipment problem, Hungarian method of assignment, maximization, prohibitions and other variations of assignment problems, duality theory of assignment problems.

(12L+3T)

Unit 3

Nonlinear programming: Karush-Kuhn-Tucker conditions, Quadratic programming, Wolfes, Beales and Fletchers algorithms for solving quadratic programming problems. Convex problems, mixed integer models.

(12L+3T)

Unit 4

Networking models: Network flows, maximal flow in the network. Transportation problems, transshipment problems and assignment problems as networking problems. Network scheduling by PERT/CPM Techniques. Resource Analysis in network scheduling. **(12L+3T)**

Total (48 L+ 12T)

Books Recommended

1. Bertsekas, D. (1999). *Nonlinear Programming*, 2nd Edn. Athena Scientific.
2. Chong, E. K. P. and Zak, S. (2004). *An Introduction to Optimization*, Wiley.
3. Fletcher, R. (2000). *Practical Methods of Optimization*, Wiley
4. Hadley, G. (1987). *Linear Programming*. Addison-Wesley.

5. Kambo, N.S. (1991). *Mathematical Programming Techniques*. Affiliated East-West press.
6. Panneerselvam, R. (2012). *Operations Research*, 2nd Edn. Prentice Hall of India.
7. Taha, H.A. (1992). *Operations Research*, 5th ed. Macmillan.

ST I2: Statistical Methods for Quality Control

Unit 1

Review: Quality, dimensions of quality, seven SPC tools. Concepts of stable industrial processes, Systematic variation, random variation. Variable & attribute control charts, $\overline{X}$ chart and R chart. X-MR chart, C chart, U chart, nP chart, P chart. Demerit control chart. CUSUM chart for process mean, CUSUM chart for process variability, Tabular CUSUM. EWMA chart for process mean. EWMA chart for process variability. Comparison of Shewhart control charts with CUSUM chart and EWMA chart.

(12L+3T)

Unit 2

General ideas on economic designing of control charts. Duncan's model for the economic control chart. Concepts of Conforming Run Length (CRL), CRL chart. Properties of CRL chart, Average Run Length (ARL), Average Time to Signal (ATS), ARL and ATS models to obtain the design parameters. Process capability & performance indices, C_p, C_{pk}. Estimation & confidence intervals for estimators of C_p. Connection between proportion of defectives & C_p. Non normal situations.

(12L+3T)

Unit 3

Synthetic control chart to detect increases in fraction non-conforming: Motivation, derivation of ATS. Operation and performance of the synthetic control chart. Synthetic control chart to detect shifts in the process mean: Motivation, derivation of ATS. Operation and performance of the synthetic control chart

Side Sensitive synthetic(SSS) control chart for variables: Motivation. Operation TPM and performance of SSS chart..

'Group Runs' (GR) control chart to detect increases in fraction non-conforming: Motivation, derivation of ATS. Operation and performance of the Group Runs control chart. Comparisons of various control charts. 'Group Runs' GR control chart to detect shifts in the process mean: Motivation, derivation of ATS. Operation and performance of the GR control chart. Comparisons of various control charts. **(12L+3T)**

Unit 4

Multi-Attribute control charts: 'Jolayemi's Multi-Attribute Control Chart' (J-MACC), drawbacks of J-MACC, 'Exact Independent Attribute Control Chart' (E-IACC).

Multivariate control charts for mean vector: The need. Hotelling T^2 chart, Multivariate synthetic control chart.

Acceptance Sampling plans: Single, double & multiple sampling plans for attributes. Curtailed double sampling plans. Operating characteristic functions &

other properties of the sampling plan. Use of sampling plans for rectification. Dodge-Romig acceptance sampling plans. Chain sampling plans, Continuous sampling plans CSP-I & CSP – II. Hamaker's conjecture.

Acceptance sampling plan for variables: Designing variable acceptance sampling plans. AQL based sampling plans. Bayesian sampling plans, Minimax regret and other sampling plans. **(12L+3T)**

Total (48 L+ 12T)

Books Recommended

1. Guenther, W. C. (1977). *Sampling Inspection in Statistical Quality Control*, Alan Stuart.
2. Levenson, W. (2011). *Statistical Process Control for Real-World Applications*. CRC Press.
3. Montgomery, D. C. (2005). *Introduction to Statistical Quality Control*. Wiley.

ST I3: Statistical Methods for Reliability

Unit 1

Coherent structures, representation of coherent systems in terms of paths and cuts, modules of coherent systems. Reliability of system of independent components, association of random variables, bounds on system reliability, improved bounds on system reliability using modular decompositions.

(12L+1Lab+2T)

Unit 2

Shape of the system reliability function, applications to relay circuits and safety monitoring systems. Notion of aging, life distributions of coherent systems, Distributions with increasing failure rate average arising from shock models, preservation of life distribution classes under reliability operations. Reliability bounds, Mean life series and parallel systems. **(12L+1Lab+2T)**

Unit 3

Classes of life distributions applicable in replacement models, NBU, NBUE, NWU & NWUE classes of life distributions and their implications. Shock models leading to NBU. Age replacement and block replacement policies. Renewal theory useful in replacement models. **(12L+1Lab+2T)**

Unit 4

Replacement policy comparisons, preservation of life distribution classes under reliability operations. Reversed hazard rate, cumulative reversed hazard function, relation between hazard function and reversed hazard function.

(12L+1Lab+2T)

Total (48 L+ 8T + 4Lab)

Books Recommended

1. Barlow, R. E. and Proschan F. (1975). *Statistical theory of Reliability and Life testing: Probability Models*. Holt, Rinehart and Winston Inc.
2. Barlow, R. E. and Proschan F. (1996*). Mathematical Theory of Reliability*. John Wiley.
3. Tobias, P. A. and Trindane, D. C. (1995). *Applied Reliability*. Second edition. CRC Press.

ST I4/ST F2: Time Series Analysis

Unit 1

Exploratory time series analysis, tests for trend and seasonality. Exponential and Moving average smoothing. Holt -Winters smoothing. Forecasting based on smoothing, adaptive smoothing. Time - series as a discrete parameter stochastic process. Auto covariance and autocorrelation functions and their properties, Portmanteau tests for noise sequences, transformation to obtain Gaussian series. **(12L + 3Lab)**

Unit 2

Stationary processes: General linear processes, moving average (MA), auto regressive (AR), and autoregressive moving average (ARMA). Stationarity and inevitability conditions. Nonstationary and seasonal time series models: Auto regressive integrated moving average (ARIMA) models, Seasonal ARIMA (SARIMA) models, Transfer function models (Time series regression) **(12L + 3Lab)**

Unit 3

Forecasting in time series models, Durbin-Levinson algorithm, innovation algorithm (without proof). Estimation of mean, auto covariance and autocorrelation functions, Yule-Walker estimation, Estimation of ARIMA model parameters, maximum likelihood method, large sample theory (without proofs). Choice of AR and MA periods, FPE, AIC, BIC, residual analysis and diagnostic checking. Unit-root non stationarity, unit-root tests. **(12L + 3Lab)**

Unit 4

Multivariate Time series model, VAR models, Vector ARMA models. Conditional heteroschedastic models, ARCH and GARCH, properties, examples, estimation & forecasting, extensions of ARCH & GARCH. **(12L + 3Lab)**

Total (48 L+ 12Lab)

Books Recommended:

1. Brockwell, P.J. and Davis, R. A. (2003). *Introduction to Time Series Analysis*, Springer
2. Chatfield, C. (2001). *Time Series Forecasting*, Chapmann & hall, London
3. Fuller, W. A. (1996). *Introduction to Statistical Time Series*, 2nd Ed. Wiley.
4. Hamilton N. Y. (1994). *Time Series Analysis*. Princeton University press.

5. Kendall, M. and Ord, J. K. (1990). *Time Series, 3rd Edn.* Edward Arnold.
6. Lutkepohl, H. and Kratzing, M. (Ed.) (2004). *Applied Time Series Econometrics*, Cambridge University Press.
7. Shumway, R. H.and Stoffer D. S. (2010). *Time Series Analysis & Its Applications*, Springer.
8. Tsay, R. S. (2010). *Analysis of Financial Time Series*, Wiley.

ST- F1: Actuarial Statistics

Unit 1

Future life time random variable, its distribution function and density function, concept of force of mortality, curtate future life time random variable its probability mass function, deferred probabilities, all these functions in terms of international actuarial notation. Analytical laws of mortality such as Gompertz' law and Makeham's law, Single decrement life table, select and ultimate life table. **(12L + 3T)**

Unit 2

Concept of compound interest rate, discount factor, present value of the money, nominal rate of interest, force of interest. Assurance contracts with level and varying benefits, such as whole life insurance, term insurance endowment insurance. Means and variances of the present value random variables of the payments under these contracts under the assumption of constant force of interest, when the benefit payments are made at the end of year of eath(discrete set up) or when it is paid at the epoch of death(continuous set up). Actuarial present value of the benefit. Net single premiums **(12L + 3T)**

Unit 3

Annuity contracts, annuity certain, discrete annuity, m-thly annuity, continuous annuity, deferred annuity, present values and accumulated values of these annuities.

Continuous life annuity, discrete life annuity, such as whole life annuity, temporary life annuity, n-year certain and life annuity, life annuities with mthly payments.

Present value random variables for these annuity payments, their means and variances. Actuarial present value of the annuity. **(12L + 3T)**

Unit 4

Loss at issue random variable, various principles to decide net premiums for insurance products and annuity schemes defined in unit II and III, fully continuous premiums and fully discrete premiums, True m-thly payment premiums. Extended equivalence principle to decide gross premiums.

Concept of reserve, prospective & retrospective approach. Fully continuous reserve. Fully discrete reserve. **(12L + 3T)**

Total (48 L+ 12T)

Books Recommended

1. Bowers, JR. N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. (1997). *Actuarial Mathematics*, 2nd Edn., The Society of Actuaries.
2. Deshmukh S.R. (2009). *Actuarial Statistics: An Introduction Using R*, Universities Press.
3. Harriett, E.J. and Dani, L. L.(1999). *Principles of Insurance: Life,Health, and Annuities*, 2nd Edn., Life Office Management Association.
4. Neill, Alistair (1977). *Life Contingencies*, The Institute of Actuaries.
5. Palande, P. S., Shah, R. S. and Lunawat, M. L. (2003). *Insurance in India - Changing Policies and Emerging Opportunities*, Response Books.

ST F3: Stochastic Models in Finance

Unit 1

Derivatives hedging: forward and future contracts. Markets, prices, arbitrage and hedging

Complete market, market risk and credit risks in the use of derivatives.

Options markets, properties of stock option prices. American and European options. Binomial model: One-step and two-step models, Binomial trees. Risk neutral valuation. **(12L + 3T)**

Unit 2

Behaviour of stock prices: Conditional expectation, martingales, Brownian motion and Geometric Brownian motion, Markov property, Ito integral, Ito/diffusion and mean reverting processes process, Ito Lemma. **(12L + 3T)**

Unit 3

Black Scholes model: Distribution of returns, volatility, risk neutral pricing, equivalent martingale measure, Black-Scholes-Merton differential equation. Estimating volatility (historical data, implied volatility). Options on stock indices, currencies and futures. **(12L + 3T)**

Unit 4

Some exotic equity and foreign exchange derivatives. Greek Letters and hedging. Value-at-risk as a measure of risk. Interest rate derivatives, Black model. Models of the term structure of interest rates: one factor diffusion model, Vesicle, Cox-Ingersoll-Ross and Hull white models **(12L + 3T)**

Total (48 L+ 12T)

Books Recommended

1. Baxter, M. and Rennie,.A. (1996). *Financial Calculus*, Cambridge University Press.
2. Bingham, N. and Keisel, R. (1998). *Risk-Neutral Valuation*, Springer.
3. Hull John , (2008). *Options, futures and other derivatives*, International 7th Edn, Pearson Prentice Hall.

4. Ross.S. (2003). *Introduction to Mathematical Finance*, Cambridge University Press.

Additional Reference Books

1. Bodie Z., Kane A., Marcus A. And Mohanty P. (2009). *Investments,* 8th Edn., McGraw Hill.
2. David, R. (2004). *Statistics and Finance: An Introduction*. Springer.
3. Shreve, S. E. (2004). *Stochastic Calculus for Finance I*. Springer.
4. Shreve, S. E. (2004). *Stochastic Calculus for Finance II*. Springer.
5. www.nseindia.com

ST F4/ST B3: Survival Analysis

Unit 1

Concepts of time, order and random censoring. Life distributions - exponential gamma, Lognormal, pareto, linear failure rate. Parametric inference, point estimation, confidence Intervals, scores, tests based on LR, MLE.

(12L+1Lab+2T)

Unit 2

Life tables, failure rate, mean residual life and their elementary properties. Ageing classes - IFR, IFRA, NBU, NBUE, HNBUE and their duals, Bathtub failure rate. Estimation of survival function - Actuarial estimator, Kaplan - Meier estimator. Estimation under the assumption of IFR/DFR.

(12L+1Lab+2T)

Unit 3

Semi-parametric regression for failure rate - Cox's proportional hazards model with one and several covariates. Competing risk models. Repair models. Probabilistic models. Joint distribution of failure times. Unconditional tests for the time truncated case. **(12L+1Lab+2T)**

Unit 4

Concept of frailty. Shared frailty models. Identifiability of frailty models. Various frailty models. Gamma, positive stable, inverse Gaussian, power variance function, compound Poisson and compound negative binomial shared frailty models. Frailty regression models. Bivariate and correlated frailty models. Additive frailty models. **(12L+1Lab+2T)**

Total (48 L+ 8T + 4 Lab)

Books Recommended

1. Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data*, Chapman and Hall.

3 Deshpande, J.V. and Purohit S.G. (2005). *Life Time Data: Statistical Models and Methods*, Word Scientific.
4 Duchateau, L. and Johnson, P. (2008). *The Frailty Model*. Springer: New York.
5 Gross A.J. and Clark, V. A. (1975) *Survival Distributions: Reliability Applications in the Biomedical Sciences*, John Wiley and Sons.
6 Hanagal, D. D. (2011). *Modeling Survival Data Using Frailty Models*. CRC Press: New York.
7 Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer: New York.
9 Wienke, A. (2011). *Frailty Models in Survival Analysis*, CRC Press: New York.

ST-F5: Multiple Decrement Models in Insurance

Unit 1

Multiple life contracts, joint life status, last survivor status, future life time of the status, curt ate future life time of the status, actuarial present value of the benefit payments, annuity payments and net premiums in a group life insurance contracts. **(12L + 3T)**

Unit 2

Multiple decrement models, distribution of time to decrement and cause of decrement random variables. Multiple decrement table, associated single decrement model, their application in the calculation of monetary functions when cause of death/decrement is involved. Premiums and reserves for insurance contracts involving cause of death/decrement.

(12L + 3T)

Unit 3

Application of multiple decrement models in defined benefit pension plan, calculation of actuarial present value of the benefit payments in a pension plan and determination of the contribution to the pension plan. Various methods of pension funding such as accrued benefit cost method for an individual, accrued benefit cost method for a group, aggregate actuarial cost method

(12L + 3T)

Unit 4

Multistate Markov models for cash flows contingent on competing risks, actuarial present value of the cash flows, determination of premium in disability income insurance contract in employee benefit schemes, determination of premium in continuing care retirement communities model in health insurance.

Stochastic interest rate, various models such as random interest scenario, parametric models, time series models and its application in the calculation of monetary functions in a variety of insurance contracts. **(12L + 3T)**

Total (48 L+ 12T)

Books Recommended

1. Bowers, JR. N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. (1997). *Actuarial Mathematics*, Second Edition, The Society of Actuaries.
2. Deshmukh S.R. (2012). *Multiple Decrement Models in Insurance: An Introduction Using R*, Springer India.
3. Harriett, E.J. and Dani, L.L.(1999).*Principles of Insurance: Life, Health, and Annuities*, Second Edition, Life Office Management Association
4. Neill, Alistair (1977). *Life Contingencies*, Institute of Actuaries, London.

ST B2: Statistical Analysis of Clinical Trials

Unit 1

Introduction to clinical trials: 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. Bioavailibility, pharmacokinetics and pharmacodynamics, two-compartment model.

(12L+1Lab+2T)

Unit 2

Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal 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 bio-equivalence trials. Inference for 2x2 crossover design: Classical methods of interval hypothesis testing for bioequivalence, Bayesian methods, nonparametric methods. **(12L+1Lab+2T)**

Unit 3

Power and sample size determination, multiplicative (or log-transformed) model, ML method of estimation, assessment of inter and intra subject variabilities, detection of outlying subjects. Optimal crossover designs: Balaam's design, Two-sequence dual design. Optimal four period designs. Assessment of bioequivalence for more than two drugs, Williams design. **(12L+1Lab+2T)**

Unit 4

Designs based on clinical endpoints: Weighted least squares method, log-linear models, generalized estimating equations. Drug interaction study, dose proportionality study, steady state analysis. Interim analysis and group sequential tests, alpha spending functions. Analysis of categorical data.

(12L+1Lab+2T)
Total (48 L+ 8T + 4 Lab)

Books Recommended

1. Chow S.C. and Liu J.P.(2009). *Design and Analysis of Bioavailability and bioequivalence*. 3rd Edn. CRC Press.
2. Chow S.C. and Liu J.P. (2004). *Design and Analysis of Clinical Trials*. 2nd Edn. Marcel Dekkar.
3. Fleiss J. L.(1989). *The Design and Analysis of Clinical Experiments*. Wiley.
4. Friedman L. M.Furburg C. Demets D. L.(1998). *Fundamentals of Clinical Trials*, Springer.
5. Jennison .C. and Turnbull B. W. (1999). *Group Sequential Methods with Applications to Clinical Trails*, CRC Press.
6. Marubeni .E. and Valsecchi M. G. (1994). *Analyzing Survival Data from Clinical Trials and Observational Studies*, Wiley.

ST B4: Statistical Methods in Microarray Data Analysis

Unit 1

Background of Microarrays and Normalization techniques

Introduction to Biology relevant to microarray experiment. Microarray experimental set up and quantification of information available from microarray experiments.

Data cleaning, transformation of data.

Between array & within array normalization, in particular quantile and LOWESS normalization, stage wise normalization.

Concordance coefficient and its role in normalization.

(12L + 3T)

Unit 2

Statistical Inference procedures in comparative experiments

Inference procedures for single channel microarray data. Application of two sample t –test. Tests for validating assumptions of two sample t-test. Application of Welch test and Wilcoxon rank sum test.

Inference procedures for two channel microarray data. Application of paired t – test. Tests for validating assumptions of paired t-test. Application of Wilcoxon signed rank test.

Inference procedures for comparing more than two types of mRNA samples in single channel or two channel microarray experiments. Application of one way ANOVA F test, one way ANOVA Welch F test, Kruskal-Wallis test, pairwise t-test, pairwise Welch test and pairwise Wilcoxon rank sum test. Strip charts and its role to decide the profile of differentially expressed genes

(12L + 3T)

Unit 3

Multiple hypotheses testing problem and Principal component analysis

Multiple hypotheses testing problem. Adjustments for multiple hypotheses testing, adjusted p-values.

False discovery rate and its application to microarray data analysis.

Principal component analysis for microarray data, scree plot, plot of scores to display sets of differentially expressed genes. Singular value decomposition of a

rectangular matrix and the concept of ballot. Its application to microarray data analysis.

(12L + 3T)

Unit 4

Cluster analysis and Logistic regression

Hierarchical cluster analysis of microarray data to identify groups of genes and outlying genes

K - means cluster analysis of microarray data to identify groups of genes

Application of logistic regression for microarray data. Concept of AIC and BIC and its role to identify marker genes.

(12L + 3T)

Total (48 L+ 12T)

R software will be heavily used in applications of all the statistical methods to microarray data to identify differentially expressed genes in two or more biological samples.

Books Recommended

1. Amartunga D. and Cabrera J. (2004). *Exploration and Analysis of DNA Microarray and Protein Array Data*. Wiley.
2. Deshmukh S.R. and Purohit S.G. (2007). *Microarray Data: Statistical Analysis Using R*, Narosa.
3. Draghici, S. (2003). *Data Analysis Tools for DNA Microarrays*, Chapman and Hall/CRC.
4. Dov, S. (2003). *Microarray Bioinformatics*, Cambridge University Press,
5. McLachlan, G.J.; Do, K.A. and Ambroise, C. (2004). *Analyzing Microarray Gene Expression Data*, Wiley.
6. Simon, R.M ; Korn, E.L. ; McShane, L.M. ; Radmacher, M.D. ; Wright, G.W. and Zhao, y. (2003). *Design and Analysis of DNA Microarray Investigations*. Springer.
7. Speed, T. (2003). *Statistical Analysis of Gene Expression Microarray Data*, Chapman and Hall/CRC.

ST-C2: Computer Intensive Statistical Methods

Note: It is recommended that this course be conducted in Computer Laboratory. R will be used for computing purpose.

Unit 1

Resampling Techniques: Re sampling paradigms, bias-variance trade-off. Bootstrap methods, estimation of sampling distribution, confidence interval, variance stabilizing transformation. Jackknife and cross-validation. Jackknife in sample surveys. Jackknife in regression under heteroscedasticity. Permutation tests. **(12L + 3T)**

Unit 2

Missing Values and Imputations Techniques: Missing values and types of missingness, imputations methods for missing values, single and multiple imputations.

EM Algorithm and Applications: EM algorithm for incomplete data, EM algorithm for mixture models, EM algorithm for missing values, stochastic EM algorithm.

(12L+3T)

Unit 3

Smoothing techniques: Kernel estimators, nearest neighbor estimators, orthogonal and local polynomial estimators, wavelet estimators. Splines. Choice of bandwidth and other smoothing parameters.

(12L+3T)

Unit 4

Bayesian computing, Markov Chain Monte Carlo. Simulation using MCMC, Particle filtering, MCMC methods for missing values. **(12L+3T)**

Total (48 L+ 12T)

Books Recommended

1. Buuren, Stef van (2012). *Flexible Imputation of Missing Data.* Chapman and Hall.
2. Chihara, L. and Hesterberg, T. (2011) *Mathematical Statistics with Resampling and R*. Wiley.
3. Davison, A.C. and Hinkley, D.V. (1997) *Bootstrap methods and their Applications*. Chapman and Hall.
4. Efron, B. and Tibshirani. R.J. (1994); *An Introduction to the Bootstrap*. Chapman and Hall.
5. Christensen R, Johnson, W., Branscum A. and Fishman, G.S. (1996) *Monte Carlo: Concepts, Algorithms, and Applications*. Springer.
6. Gilks, W. R., Richardson, S., and Spiegelhalter, D. (eds.) (1995) *Markov Chain Monte Carlo in Practice*. Chapman and Hall.
7. Good, P. I. (2005) *Resampling Methods: A Practical Guide to Data Analysis*. Birkhauser Bosel.
8. Hanson T. E. (2011). *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*, Chapman & Hall.
9. Jim, A. (2009). *Bayesian Computation with R*, 2nd Edn, Springer.
10. Kennedy W. J. & Gentle J. E. (1980) *Statistical computing*. Marcel Dekker.
11. McLachlan, G.J. and Krishnan, T. (2008) *The EM Algorithms and Extensions*. Wiley.
12. Rubinstein, R.Y. (1981); *Simulation and the Monte Carlo Method*. Wiley.

13. Shao J. and Tu, D. (1995); *The Jackknife and the Bootstrap*. Springer Verlag.
14. Tanner, M.A. (1996); *Tools for Statistical Inference*, Third edition. Springer.

ST C4: Statistical Learning and Data Mining

Note: It is recommended that this course be conducted in Computer Laboratory.

Unit 1

Supervised Learning: Linear methods for classification, linear discriminant analysis (LDA), logistic regression, Bayes classifier, nearest neighbor classifier. Packages in R for these methods. **(12L + 3T)**

Unit 2

Neural network (NN), support vector machine (SVM). Packages in R for these methods. **(12L + 3T)**

Unit 3

Regression and classification trees (CART). Assessment and model selection: Bias-variance trade off, training error rate, AIC, BIC, CIC, DIC (information criterion), cross-validation. Ada boosting. **(12L + 3T)**

Unit 4

Unsupervised learning: Clustering procedures- k-means, hierarchical, self-organizing map, EM algorithm. Feature selection: principal component analysis. Association rules. Software packages for these methods. **(12L + 3T)**

Total (48 L+ 12T)

Books Recommended

1. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth and Brooks.
2. Daniel T. Larose, (2006). *Data Mining Methods and Models*. Wiley.
3. Galit Shmueli, Nitin Patel, Peter Bruce, (2010). *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XL Miner* , Wiley
4. Hastie T., Tibshirani R. and Friedman J. H., (2003). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer.
5. Mitchell Tom, (1997). Machine Learning. McGraw-Hill.
6. Ripley,B. D. (1996). *Pattern Recognition and Neural Networks.* Cambridge University Press.

ST B6/ST C6 : Statistical Methods for Bio-Computing

Unit 1

Type of genetic data:- Molecular and morphological data. Differences and advantages of molecular data on, morphological data, Character data and distance data, their relative merits and demerits.

Concept of entropy, entropy as a measure of uncertainty, entropy of single and combined scheme/s, Measure of information content based on entropy. Relative entropy its similarity with likelihood ratio. Applications of these to biological sequences.

(12L + 4 Lab/ T)

Unit 2 (Alignment of biological sequences): Pairwise and local alignment of biological Sequences (DNA/protein sequences). How biological sequences are different from mathematical sequences? The scoring matrices for alignment algorithms PAM and BLOSUM matrices. Algorithm for global alignment (Needleman–Wunch algorithm). Local alignment algorithms (Smith - Waterman) Gap Model, dynamic programming algorithms for alignment with gaps such as linear gap model, affine gap model. Introduction to heuristic alignment algorithms such as BLAST, FASTA.

(12L + 4 T/Lab)

Unit 3

Molecular phylogeny Analysis: Tree of life, gene and species tree. Distance based methods for reconstruction of phylogenetic tree such as UPGMA, weighted UPGMA, transformed distance method, nearest – neighbor joining method. Comparison of trees generated using different distance function Requisites of a good distance function.

Character based methods for molecular phylogeny, maximum likelihood method and maximum parsimony method. Assessing trees via bootstrap. Probabilistic approach to phylogeny. Probabilistic models of evolution, Felsentein's algorithm for likelihood computation. Juke – Canter model and Kimura and other probabilistic models for evolution.

(12L+4T/Lab)

Unit 4

Applications of Markov and Hidden Markov models to biological sequence Analysis.

Markov chain as a classifier, use of Markov chain Model for demarcation of a region in Biological sequence analysis. Application of these in genetic sequence analysis such as detection of CPG Island. Testing whether given stretch of sequence is coming from CPG Island (use of Markov model for discrimination) Markov model based classification & clusterization, testing order of a Markov model, testing homogeneity of two Markov models, Use of these test to design clustering algorithm. Hidden Markov/chains. Difference between these and

simple Markov chains. Analysis of Hidden Markov Models/chains. Verterbi's algorithm, Forward and backward algorithm for hidden Markov model. Parameter estimation in hidden Markov model when path is known as well as unknown, Baum – Welch algorithm.

(12L + 4T/lab)
Total (48 L+ 12T/lab)

Recommended Books

1. Alexander Isaac: (2001). *Introduction to Mathematical Methods Bioinformatics*. Springer.
2. Durbin R., Eddy S. Krogh A. Michelson G. (1998). *Biological Sequence Analysis*, Cambridge University Press.
3. Robin S., Rudolph F, Schboth S. (2003) *DNA Words and models Statistics of Exceptional Words*, Cambridge University Press.

ST-E9: Financial Econometrics

Unit 1

Unit roots, Cointegration and VAR models: difference stationary and trend stationary processes. Testing for unit roots: the DF, ADF, PP and KPSS test statistics, VAR, ML estimation Granger causality, Cointegrating VAR, Applications to the PPP (purchasing power parity), Applications to the net present value model of stock prices, market microstructure and the efficient market hypothesis, examples & data analysis **(12L+3T)**

Unit 2

Frequency domain analysis of time series: Fourier transform and discrete Fourier Transform (DFT), periodicity, Spectral density, periodogram and DFT, spectral representation, inference, examples & data analysis **(12L+3T)**

Unit 3

Structural equation modeling, state-space models, Kalman filter: State-space representations, the basic structural model, state-space representation of ARIMA models, The Kalman recursions, estimation for State-Space models, Generalized state-space models, parameter & observation-driven models, examples & data analysis. **(12L+3T)**

Unit 4

Stochastic volatility models: Volatility definition and estimation, Volatility forecast evaluation, Stochastic volatility models, MCMC approach, examples.
Modeling regime shifts - Markov-switching models: Representation (AR & MA), estimation, diagnostic checking, forecasting, examples & data analysis

(12L+3T)

Total (48 L+ 12T)

Recommended Books:

1. Ait-Sahalia, Y. & Hansen, L. P. (Ed.) (2010). *Handbook of Financial Econometrics: Tools and Techniques*, Vol.1 & Vol. 2, Elsevier.
2. Campbell J., Lo A. & McKinley C. (1997). *The Econometrics of Financial Markets*. Princeton University press.
3. Hamilton (1994). *Time Series Analysis*. Princeton University press.
4. Knight, J. and Satchell, S. (2007). *Forecasting Volatility in the Financial Markets,* 3rd Edn., Elsevier.
5. Lutkepohl, H. and Kratzing, M. (Ed.) (2004). *Applied Time Series Econometrics*, Cambridge University Press.
6. Poon Ser-Huang (2005). *A Practical Guide to Forecasting Financial Market Volatility*, Wiley.
7. Rachev. S. T., Mittnik, S., Fabozzi, F. J. , Focardi, S. M. and Jasic, T. (2007). *Financial Econometrics:From Basics to Advanced Modeling Techniques,* Wiley.
8. Ruppert, D. (2004). *Statistics and Finance: An Introduction*, Springer.
9. Shephard, N. (2004). *Stochastic Volatility: Selected Readings*, Oxford University Press.
10. Soderlind, P. (2010). *Lecture Notes in Financial Econometrics*, University of St. Gallen.
11. Tsay, R. S. (2010). *Analysis of Financial Time Series*, 3rd Edn. Wiley.
12. Wang, P. (2003). *Financial Econometrics: Methods and Models*, Routledge.