Savitribai Phule Pune University

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,

Savitribai Phule Pune University)

(With effect from Academic Year 2018-19)

Date: 30 May, 2018

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 80 credits spread over four semesters. This program is offered at the Department of Statistics, Savitribai Phule Pune University. The program emphasizes both theory and modern 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, PYTHON, 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 and Quality Control, Actuarial Statistics, Stochastic Models in Finance, Statistical Learning and Data Mining, Statistical Analysis of Clinical Trials, Statistical Analysis of Microarray Data, Expert Systems with Applications, Statistical Methods for Bio-Computing, Computer Intensive Statistical Methods 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 of the syllabus, there are four core courses, five optional courses and one project. 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 optional/electives.

3. Introduction: M.A./M.Sc. Statistics program has semester pattern and Choice-based Credit System. The program consists of 80 credits. Credits of a course are specified against the title of the course.
The program has pre-requisites system. Under this system, a student has to clear pre-requisite courses to take some of the 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 theorycourses. The Pre-requisites of a course are indicated in square 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 Statistics as one of the three 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.A./B.Sc. (Second class), with Statistics and Mathematics as two of the three subjects.

5. Examination

A) **Pattern:**

(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, small projects, 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/practical courses the duration for the ETE will be three 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 completes50% 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 may be moderated by a committee of subject experts.

G) **Verification or revaluation:** As per the University rules

6. Structure of the Program

(a) Compulsory papers

Compulsory courses in semester I and semester II are listed below. In the square brackets pre-requisite subjects have been mentioned. **Each course is of 4 credits**.

Semester I

ST-101: Mathematical Analysis
ST-102: Linear Algebra
ST-103: Probability Distributions
ST-104: Calculus
ST-105: Statistical Computing

Total credits 20

Semester II

ST-201: Probability Theory [ST-101]
ST 202: Multivariate Analysis [ST-102]
ST-203: Modern Statistical Inference [ST-103, ST-104]
ST-204: Regression Analysis [ST-102]
ST-205: Practical- I(based on ST-202; ST-204)

Total credits 20

Semester III

ST-301: Stochastic Processes
ST-302: Design and Analysis of Experiments [ST-204]
ST-303: Practical - II (based on ST-301, ST-302)
Optional - I
Optional - II

Total credits 20

Semester IV

ST-401: Survey Sampling and Introduction to Official Statistics
ST-402: Project
Optional - III
Optional - IV
Optional - V

Total credits 20

The core courses have 60 (48 (theory) + 8 (practical) + 4 (project)) credits in all (75% of total credits) and the five optional courses have20 credits (25% of total credits).

A student may be allotted optional courses of Semester III by the Department at the beginning of the second year of the program. The allotment is based on the performance of the student as indicated by her/his rank among all the students admitted in the same year and the courses that the student preferred.

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.

A list of optional courses is given below. **Each optional course is of 4 credits.**

.

List of 22 Optional Courses

ST O1 Advanced Probability [ST 201]
ST O2 Advanced Stochastic Processes [ST 301]
ST O3Inference in Stochastic Processes [ST 205]
ST O4Spatial Processes and their Applications [ST 301]
ST O5 Stochastic Models in Finance [ST 301]
ST O6 Asymptotic Theory of Statistical Inference [ST201, ST 203]
ST O7 Time Series Analysis [ST 204]
ST O8Advances in Generalized Linear Models [ST 204]
ST O9 Astrostatistics [ST 202, ST 203, ST 204]
ST O10 Optimization Techniques [ST 102, ST 104]
ST O11 Reliability and Statistical Quality Control [ST 203]
ST O12 Computer-Intensive Statistical Methods [ST 105 ST 203]
ST O13 Statistical Learning and Data Mining [ST 202, ST 204]
ST O14 Survival Analysis [ST 203, ST 204]
ST O15 Actuarial Statistics [ST 103]
ST O16 Medical and Health Statistics [ST 103, ST 204]
ST O17 Analysis of Clinical Trials [ST 204, ST 302]
ST O18 Statistical Methods in Micro-array Data Analysis [ST 202]
ST O19 Statistical Methods for Bio-computing [ST 105, ST 301]
ST O20 Bayesian Inference [ST203]
ST O21 Expert Systems with Applications [ST 105]
ST O22 Financial Econometrics [ST 204, ST 301]

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

Question papers: In view of academic autonomy given to the Department, question papers are set by the teacher who teaches the course and these may be moderated by a committee of experts, as indicated earlier.

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 in the first year of M.Sc. 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 101: Mathematical Analysis

Unit I

Countability, supremum and infimum of sets of real numbers, Archimedean property, denseness property of rational numbers, 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.

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 toAdvanced 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. Ghorpade, S. R. and Limaye, B. V. (2006). A Course in Calculus and Real Analysis, Springer
5. Goldberg R. R. (1976). Methods of Real Analysis, John Wiley
6. Kumar, A. and Kumaresan, S. (2014). A Basic Course in Real Analysis, CRC Press
7. Rudin, W. (1985). Principles of Mathematical Analysis, McGraw - Hill
8. Trench W. F. (2012). Introduction to Real Analysis, E-book.
9. Yau, D. (2013). A First Course in Analysis, World Scientific

ST 102: Linear Algebra

Unit I

Determinants and their simple properties, partitioned matrices, inverses, Vector operator, 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, Injective, surjective and inverse linear transformations, 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, Cayley-Hamilton theorem.

Matrix Inequalities, Rank, Determinant, and Trace Inequalities, Eigenvalue Inequalities

Quadratic forms with symmetric matrices, 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, Quadratic form inequalities

(12L + 3T)

Unit IV
Derivatives with respect to vectors and matrices

LU Factorization, Cholesky factorization, Spectral Decomposition, Singular value decomposition, Applications

Large scale eigen value problems, Fast Direct Solution of a Large Linear System

(12L + 3T)

Total (48L + 12 T)

Books Recommended

1. Bapat, R.B. (2011). *Linear Algebra and Linear Models*. Springer and Hindustan Book Agency.
2. Beezer, R. A. (2004). *A First Course in Linear Algebra*, Congruent Press, Washigton
3. Hohn, F. E. (1973). *Elements of Matrix Algebra*, Macmillan
4. Kollo, T. and Rosen, D. von (2005). *Advanced Multivariate Statistics withMatrices*, Springer, New York.
5. Kumaresan, S. (2000). Linear Algebra: A Geometric Approach, Prentice Hall
6. Lay, D. C. Lay, S. R. and Mc Donald, J. J. (2016) .*Linear Algebra and ItsApplications*, Fifth Edition, Pearson, Boston.
7. Ramachandra Rao, A. and Bhimasankaram, P. (2000). *Linear Algebra*.Hindustan Book Agency
8. Rao, C. R. (1995). *Linear Statistical Inference and Its Applications*, Wiley
9. Searle, S. R. (1982). *Matrix Algebra Useful for Statistics*, John Wiley, New York.

ST 103: Probability Distributions

Unit I
Cumulative distribution function (cdf), quantile function, symmetry of a distribution, Moment generating function and Stieltjes moment problem, Mixtures of probability distributions, decomposition of a c.d.f. into discrete and continuous c.d.f.s, Truncated distributions (binomial, truncated Poisson, normal etc.), Concept of Survival function, hazard rate and cumulative hazard rates.

(14L + 4T)

Unit II
Random vectors, joint distributions, joint m.g.f., jointp.g.f., mixed moments, variance-covariance matrix, multivariate normal and bivariate exponential distributions. Independence of random variables sums of independent random variables, convolutions, conditional expectation and variances, compound distribution. Multiple and partial correlation coefficients

(14L + 4T)

Unit III

Sampling distributions of statistics from univariate normal random samples, distributions of linear and quadratic forms involving normal random variables, Fisher Cochran theorem, non-central Chi-square, non-central t and F distributions

(12L + 2T)

Unit IV

Order statistics, Joint distribution of order statistics, distribution of r-th order statistics, joint distribution of rth and sth order statistics ($r < s$) and their functions, distribution of spacings, normalized spacings with illustration to exponential case, distribution of sample median and sample range.

(8L + 2T)

Total (48L + 12T)

Books Recommended

1. Berger, R. and Casella G. (2002). *Statistical Inference*, Duxbury Resource Center, Second Edition.
2. Dasgupta, A. (2010) *Fundamentals of Probability: A First Course*, Springer, New York.
3. Hogg, R. V. McKean, J. W. and Craig, T. T. (2005*). Introduction to Mathematical Statistics*, Sixth Edition, Pearson Prentice Hall, New Jersey.
4. Rao, C. R. (2002). *Linear Statistical Inference and Its Applications,* Wiley
5. Rohatgi, V. K. & A. K. M. E Saleh (2001). *Introduction to Probability and Statistics,* Wiley, New York.

ST 104: Calculus

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 thesign 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. Ghorpade, S. R. and Limaye, B. V. (2006). A Course in Calculus and Real Analysis, Springer
4. Kreyszig, E. (1975). Advanced Engineering Mathematics, Wiley Eastern
5. Rudin, W. (1985). Principles of Mathematical Analysis, McGraw-Hill
6. Trench W. F. (2012). Introduction to Real Analysis. E-book

ST 105: Statistical Computing

This course is partly theory and mostly lab oriented. There will be 2 hours of lectures per week and a minimum of 4 hours of lab. (One credit is equal to one hour of lecture or two hours of Lab.) Good programming practices and efficient use of software packages such as R, PYTHON, SYSTAT, MINITAB, MATLAB, MATHEMATICA should be taught during lab sessions.

Unit I

(a) Theory of inverse transformation method (ITM) for random variate generation- definition of quantile function, its properties. Quantile function as a random variable and its distribution function. ITM based algorithms to generate random variates from standard discrete and continuous distributions.

(b) Theory of Acceptance-Rejection method (ARM) for random variate generation - the conditional distribution of Y given that [U<= f(Y)/ Mg(Y)] when Y ~ gand U ~ U(0; 1); where f and g are density functions. Interpretation and optimal choice ofM using exponential tilting, ARM based algorithms for random variate generation.

(c) Generation of random variates using the relationships between distributions, composition and convolution methods. Algorithms for random variate generation from mixture distributions, Chi-square, t and F-distributions.

(d) Random variate generation from bivariate and conditional distributions.

(e) Theory of random number generation -linear, multiplicative and mixed random number generators. Testing a random number generator- runtest, Kolmogrov-Smirnov test, sign test, rank test, gap test, digit frequency test and serial correlation. Selection of a random number generator.

(10L+5Lab)

Unit II

Numerical algorithms such as direct search, grid search, interpolation search, gradient search, Bisection and Newton-Raphson methods, Mullers method, Aitkens extrapolation, Simple applications of the above methods.

(10L+5Lab)

Unit III

Methods to compute integrals- quadrature formula, double integration, Gaussian integration, Monte Carlo Methods: Monte Carlo integration and its application to compute expected values and probabilities, Theory of Importance Sampling with applications to reduce Monte Carlo error and rare-event simulation, Verification of WLLN, CLT and other approximations through simulation.

(10L+5Lab)

Unit IV

Practicum on topics from Linear Algebra:

1. Calculation of determinant of higher order by partitioning method
2. Calculation of rank of a matrix
3. Calculation of equivalent 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

6. Calculation of Moore-Penrose inverse
7. Calculation of a g-inverse
8. Calculation of eigen values and eigen vectors
9. Solution of simultaneous equations
10. Spectral decomposition and powers of a matrix

(15Lab)

Total (30L + 30 Lab)

Books Recommended:

1. Kennedy W. J. and Gentle J. E. (1980). Statistical Computing, Marcel Dekker
2. Law, A.M. and Kelton, W.D. (2000). Simulation, Modeling and Analysis, Tata McGraw Hill, Third Edition
3. Norman Matloff (2011) The Art of R Programming-A Tour of Statistical Software Design, No Starch Press, San Francisco
4. Rajaraman, V. (1993). Computer Oriented Numerical Methods, Prentice-Hall, Fourth edition
5. Schilling, R.J. and Harris, S.L. (2002). Applied Numerical Methods for Engineers Using MATLAB and C, Thomson and Brooks/Cole, Singapore

ST 201: Probability Theory

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. Realand vector-valued random variables

(10L+3T)

Unit II

Distribution functions (d.f.), discrete and continuous random variables and vectors and their distribution functions, decomposition of a d.f. Expectation, Linear properties of expectations, Inequalities.

(10L+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 random variables, Borel zero-one law, Borel-Cantelli Lemma, Kolmogorov zero- one law.

(10L+3T)

Unit IV

Convergence of a sequence of random variables, Various types of convergence and their interrelationships, Cramer's theorem (Slutsky's theorem), monotone convergence theorem and dominated convergence theorem, Fatou's lemma.

Laws of large numbers, weak (with proof) and strong (without proof) law of large numbers,

Characteristic functions, Continuity theorem for characteristic functions.

Central Limit theorem, Liapounov's and Lindeberg'scentral limit theorems (without proof),

(18L+3T)

Total (48L+ 12T)

Books Recommended

1. Athreya, K. B. and Lahiri S. (2006). *Probability Theory*, Hindustan Book Agency,
2. Bhat, B. R. (2007). *Modern Probability Theory: An Introductory Text Book*, New Age International
3. Billingsley, P. (1995). *Probability and Measure*, 3rd Edition, John Wiley, New York
4. Chung, K. L. (2001). *A Course in Probability Theory*, Third Edition, Academic Press, London
5. Gut, Allan (2005), *Probability: A Graduate Course*. Springer, New York
6. Rao, B. V. (2018) Lecture notes on Asymptotic Theory of Statistical Inference, Pre-print.

ST 202: Multivariate Analysis

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)

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 andSolutions, 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. and 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, WileyEastern
11. Timm, N. H. (2002), Applied Multivariate Analysis, Springer, New York

ST 203: Modern Statistical Inference

Unit I

Data reduction, sufficiency, sufficient partition, Neyman factorization theorem, minimal sufficiency, completeness, Ancillarity and Basu's theorem

One-parameter exponential family, Multi-parameter exponential family and Pitman family of distributions, Canonical form, convexity property, minimal sufficiency and completeness

Unbiased Estimator, estimability of parametric functions, Cramer-Rao inequality, uniformly minimum variance unbiased estimators, Rao-Blackwell and Lehmann Scheffe theorems

(12L + 3T)

Unit II

Confidence sets and intervals, shortest expected length confidence intervals, relation with testing of hypotheses.

Test function, MP tests, Neyman- Pearson lemma, UMP tests, nonexistence of UMP tests MLR property.

(10L + 3T)

Unit III

Consistent and asymptotically normal (CAN) estimators for real and vector valued parameters, invariance property under continuous transformation, methods for generating CAN estimators for real and vector valued parameters using method of moments and method of percentiles.

Comparison of consistent estimators, minimum sample size required by the estimator to attain certain level of accuracy

(10 L + 3 T)

Unit IV

Maximum Likelihood Estimation, restricted parameter space, Inconsistent MLEs, MLEs in irregular cases.

Asymptotic distribution of MLE in special class of distributions: Cramer regularity conditions, Cramer- Huzurbazar theorem, Extension to vector-valued parameters,

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.

Asymptotic confidence intervals: Construction and examples.

Applications to categorical data analysis, tests for independence for three dimensional contingency tables

(16L + 3T)

Total (48L + 12T)

Books Recommended

1. Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Duxbury Advanced Series, Second Edition.
2. Efron, B. and Hastie, T. (2016). Computer Age Statistical Inference: Algorithms, Evidence and Data Science. Cambridge University Press
3. Ferguson, T.S. (1996), A Course in Large Sample Theory, Chapman &Hall, London
4. Gupta Anirban Das (2008), Asymptotic Theory of Statistics and Probability, Springer
5. Kale, B.K. & Muralidharan, K. (2015) *Parametric Inference: An Introduction*, Alpha Science International Ltd.
6. Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*. Springer, New York
7. Lehmann, E. L. and Romano, J. (2005). *Testing Statistical Hypotheses*, Springer
8. Rao, C. R. (1995). Linear Statistical Inference and its Applications, Wiley
9. Rohatgi, V. K. and Saleh, A.K. Md. E. (2001). *Introduction to Probability and Statistics*, John Wiley & Sons, New York.
10. Shao, J. (2003). *Mathematical Statistics*, Springer-Verlag, New, New York,

ST 204: Regression Analysis

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 analysis 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 nonlinearity, 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 205: Practical I

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

Unit III
1. Graphical representation of multivariate data
2. Principal Component Analysis, Correspondence analysis
3. Factor Analysis
4. Cluster Analysis
5. Canonical Correlations

Unit IV
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:

1. Khattree, R. &Naik D. N. (2003), Applied Multivariate Statistics with SAS Software, Second Edition (SAS Institute and Wiley), Chapter 2
2. Recommended books as stated in regression analysis, multivariate analysis and stochastic processes.

ST 301: Stochastic Processes

Unit I
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'

(10L+3T)

Unit II
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. 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

(10L+3T)

Unit III

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, Applications of these processes.

(8L+2T)

Unit IV

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

Brownian motion, Hitting Times, Maximum Variable and the Gambler's Ruin Problem

Gaussian Processes, Examples - Ornstein-Uhlenbeck process, Brownian Bridge, Geometric Brownian Motion,

(20L + 4T)

Total (48 L+ 12T)

Books Recommended

1. Bhat B.R. (2000). *Stochastic Models: Analysis and Applications*, New Age International.
2. Medhi, J. (2010) *Stochastic Processes*, New Age Science Ltd.
3. Pinsky M. A. and Karlin, S. (2010). *An Introduction to Stochastic Modeling*, 4thEdn. Academic Press.
4. Ross, S. (2014). *Introduction to Probability Models*, 11th 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 – 302: Design and Analysis of Experiments

Unit I

Review of randomization, replication and local control, Analysis of one wayclassification model. Analysis of two way classification model with equal number ofobservations 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 II

Analysis of 2k 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, $p = 2, 3.$, factional factorial experiments, Resolution of a design, (III, IV & V), aberration of a design.

(12L+3T)

Unit III

Analysis of 3^2 designs: contrasts for linear and quadratic effects, analysis of 3^2 design, confounding and fractional experiments in 3^2design. 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.

(12L+3T)

Unit IV

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 303: Practical II

Unit I

1. Simulation of Marknov chain and estimating the stationary distribution of ergodic Markov chain.
2. Simulation of branching process and estimating its mean and variance.
3. Simulation of Poisson and related processes.
4. Generating birth-death process and its limiting distribution.

Unit II

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.

Unit III

1. 2^k factorial experiments, Analysis of single replicate of 2^k.
2. Total and partial confounding in 2^k factorial experiments.
3. Analysis of 2^k fractional factorial experiments
4. Analysis of 3^2 factorial experiments

Unit IV

1. Random effect and mixed models
2. Analysis of first and second order response surface model. Centralcomposite design. Contour and surface plots, Box-Behnken design
3. Small composite designs, optimality of designs, simplex lattice designs, simplex centroid designs.
4. Taguchi methods: S/N ratio, orthogonal arrays, triangular tables, linear graphs, inner and outer arrays.
5. Split-plot designs

ST 401: Survey Sampling and Introduction to Official Statistics

Unit I

Review of basic\ methods of simple random sampling and stratified random sampling, Use of supplementary information for estimation, ratio and regression estimators with their properties and generalizations,

Systematic sampling, PPS sampling, Estimation problems, Hansen-Horwitz estimator and its properties, Horwitz-Thompson estimator and its properties, Midzuno-Sen method

Cluster sampling, multistage-sampling, Double sampling procedures and their ratio and regression estimators, stratification estimator, Multiphase sampling.

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.

(18L+4T)

Unit II

Inference under super population model, model assisted and model based inference, Robustness of designs and predictors, Bayesian inference, Spatial Smoothing, Sampling on Successive Occasions: Panel Rotation, Non-Response and Not-at-Homes, Weighting Adjustments and Imputation, Time Series Approach in Repeated Sampling, Comparison of strategies under super population models

(10L+3T)

Unit III

Network and Adaptive Procedures, Estimation by Network and by Adaptive Sampling, Constraining Network Sampling and Constraining Adaptive Sampling

(8L+2T)

Unit IV

Elements of Indian Official Statistics, National Accounts – different approaches, Indices for Development, Human Development Index, Evaluation & Monitoring

(12L+3T)

Books Recommended

1. Arnab, R. (2017). Survey Sampling: Theory & Applications, Academic Press
2. Chaudhuri, A. (2014). Modern Survey Sampling, CRC Press
3. Cochran, W.G. (1984). Sampling Techniques,Wiley.
4. Des Raj and Chandhok, P. (1998). Sample Survey Theory, Narosa.
5. Gal, I. and Ograjensˇek, I. (2017). Official Statistics and Statistics Education: Bridging the Gap, Journal of Official Statistics, Vol. 33, No. 1, pp. 79–100
6. Okafor, C (2002). Sample survey Theory with Applications, Snaap Press Ltd.
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.
9. Sukhatme, P.V, Suktatme, B.V.,Sukhatme, S. and Asok, C. (1984).Sampling Theory of Surveys with Applications, Indian Society forAgricultural Statistics, New Delhi.
10. Thmpson, S. K. (2012). Samplig, 3rd Edn., Wiley
11. UNDP (2010) Human Development in India: Analysis to Action
12. UNDP (2015) Training Material for Producing National Human Development Reports
13. UNDP (2016) Human Development Report 2016
14. The 2010 Human Development Index (HDI): Construction and Analysis

Optional Courses

The syllabus of some of the optional courses is given below. These optional courses will be offered subject to the availability of the expert faculty in that area. All these courses are of 4 credits

ST O1 Advanced Probability [ST 201]

Unit I

Measure and Integration, Integral of a measurable function with respect to a measure, its properties, L-p spaces, Hahn – Jordan decomposition, Lebesgue decomposition, Radon – Nikodym derivative, Product measure, Fubini's theorem, Convolutions.

(14L + 3T)

Unit II

Probability spaces, Kolmogorov's consistency

Distribution Functions and their basic properties, Helly–Bray type results

(10L + 3T)

Unit III

Convergence in measure, Almost everywhere convergence, Kolmogorov inequality, Kolmogorov three series criterion and strong law of large numbers, Introduction to weak convergence

The Paul Lévy Continuity Theorem, Convergence in distribution in the multidimensional case - The Cramér-Wold device with examples

(12L + 3T)

Unit IV

Conditional probability and conditional expectations, their simple properties, discrete parameter martingales,

Limit theorems for dependent sequences of random variables, Martingale convergence theorems, central limit theorem for martingales, Mixing sequences, Mixing coefficients, Coupling and covariance inequalities, Central limit theorems for mixing sequences

(12L + 3T)

Books Recommended

1. Ash, R.B. (1972). Real Analysis and Probability. Academic Press
2. Ash, R. B. (2000). Probability & Measure Theory. Academic Press.
3. Athreya, K.B. and Lahiri, S.N. (2006). Measure Theory and Probability Theory. Springer.
4. Billingsley, P. (1986). Probability and Measure. John Wiley
5. Dudley, R. M. (2004). Real Analysis and Probability. Cambridge University Press.
6. Roussas, G. G. (2014) An Introduction to Measure-Theoretic Probability, 2nd Edition, Academic Press
7. Taylor, J. C. (1997). Introduction to Measure and Probability. Springer.
8. Williams, D. (1991). Probability with Martingales. Cambridge University Press.

ST O2: Inference for Stochastic Processes [ST 203, ST 301]

Unit I

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 II

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 III

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 IV

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 O3: Spatial Processes and their Applications [ST 301]

Unit I

Spatial sampling, Smoothing and Interpolation

Spatial models and geo-statistics, Classical Geostatistical Methods, kriging, Variogram and covariance models and estimation,,

Nonstationary Spatial Processes, Non-Gaussian and Nonparametric Models for Continuous Spatial Data

(16L+ 3T)

Unit II

Autocorrelation on spatial network
Random fields, Markov Random fields, Statistical inference in random fields

(10L+ 3T)

Unit III

Spatial Point Process, Spatial Point Process Models (Poisson, Cox, Markov etc.) , Parametric and Nonparametric Methods, Modeling strategies, Multivariate and Marked Point Processes,
Point Process Models and Methods in Spatial Epidemiology, Isotropy for spatial point patterns

(12L+ 3T)

Unit IV

Space-time data, Space-time models, space-time symmetry, parametric and nonparametric methods for assessing space-time symmetry, Spatio-Temporal Processes, Continuous Parameter Spatio-temporal Processes, Dynamic Spatial Models Including Spatial Time, Spatio-temporal Point Processes.

(10L+ 3T)
Total (48 L+ 12T)

Books Recommended

1. Gaetan, C. and Guyon, X. (2010). Spatial Statistics and Modeling, Springer
2. Gelfand, A. E., Diggle, P., Fuentes, M. and Guttorp, P. (2010). Handbook of Spatial Statistics, CRC Press
3. Møller, J. and Waagepetersen, R. P. (2004). Statistical Inference and Simulation for Spatial Point Processes, CRC Press.
4. Ripley, B. D. (2004). Spatial Statistics, Wiley
5. Schinazi, R. B. (2010). Classical and Spatial Stochastic Processes With Applications to Biology, Birkhauser
6. Sherman, M. (2011). Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties , Wiley
7. Tautu, P. (1984). Stochastic Spatial Processes: Mathematical Theories and Biological Applications, Springer-Verlag

ST O 4: Stochastic Models in Finance [ST 103, ST 301]

Unit I

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 II

Behaviour of stock price, 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 III

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 IV

Some exotic equity and foreign exchange derivatives, Greek Letters and hedging, Value-at-risk as a measure of risk, Expected Shortfall, Interest rate derivatives, Black model, Models of the term structure of interest rates: one factor diffusion model, Vasicek, 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. Bodie Z., Kane A., Marcus A. and Mohanty P. (2009). Investments, 8th Edn., McGraw Hill.
4. David, R. (2004). Statistics and Finance: An Introduction. Springer.
5. Hull, J. (2008). Options, futures and other derivatives, International 7th Edn, Pearson Prentice Hall.
6. Joshi, M. S. (2002). The Concepts of Mathematical Finance, Lecture Notes
7. Ross.S. (2003). Introduction to Mathematical Finance, Cambridge University Press.
8. Shreve, S. E. (2004). Stochastic Calculus for Finance I, Springer.
9. Shreve, S. E. (2004). Stochastic Calculus for Finance II, Springer.
10. www.nseindia.com

ST O5: Asymptotic Theory of Statistical Inference [ST 201, ST 203]

Unit I

Classical CAN Theory: Consistency (strong, weak and uniform), CAN estimators, Best Asymptotic Normal (BAN) estimators, Asymptotic efficiency, Hodge's estimator, Delta method, Maximum likelihood estimators (MLE), CAN-ness of MLE in Cramer-regular families, scoring procedures, Inconsistent MLEs (Neyman-Scott problems)

(14L + 3T)

Unit II

Asymptotic Efficiency in Testing, Pitman Efficiencies, Bahadur Slopes and Bahadur Efficiency
Concept of M estimators, Influence curve, asymptotic linearity of estimators, Introduction to functional calculus
Edgeworth Expansions and Cumulants, Expansion for Means, sample percentiles, t-statistics

Cornish-Fisher Expansions

(12L + 3T)

Unit III

Contiguity of probability measures, Local Asymptotic Normality (LAN), Efficiency of estimators, Le Cam's Lemmas, Limitations and redundancy of Cramer's regularity conditions, A class of non-differentiable location models which yield efficient estimators, Quadratic mean differentiable (QMD) families, Applications of LAN theory of QMD families,

(12L + 3T)

Unit IV

Bootstrap Distribution and the Meaning of Consistency, Delta Theorem for the Bootstrap, Failure of the Bootstrap, Bootstrap Confidence Intervals, Bootstrap Confidence Intervals for Quantiles, Bootstrap in Regression, Residual Bootstrap, Problems with bootstrapping in the case of dependent observations

(10L + 3T)

Total (48 L+ 12T)

Books Recommended

1. DasGupta, A. (2008), *Asymptotic Theory of Statistics & Probability*, Springer, New York
2. Ferguson, T. S. (1996), *A Course in Large Sample Theory*, Chapman & Hall, London
3. Fernholz, (1983). Von Mises Calculus for Statistical Functionals, Spribger
4. Le Cam, L. M. and Yang, G. (1990), *Asymptotics in Statistics: Some Basic Concepts*, Springer, New York
5. Lehmann, E. L. (1999), *Elements of Large Sample Theory*, Springer, New York
6. Rajarshi, M. B. (2012), *Statistical Inference for Discrete Time Stochastic Processes*, Springer, New Delhi
7. Roussas, G. G. (1972), *Contiguity of Probability Measures: Some Applications in Statistics*, Cambridge University Press, London
8. van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press, London

ST O6: Time Series Analysis [ST 204]

Unit I

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 II

Stationary processes: General linear processes, moving average (MA), auto regressive (AR), and autoregressive moving average (ARMA), Stationarity and invertibility 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 III

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 IV

Multivariate Time series model, VAR models, Vector ARMA models.

Conditional heteroschedastic models, ARCH and GARCH, properties, examples, estimation & forecasting, extensions of ARCH & GARCH to asymmetric models.

Count time series models, INAR models, Poisson INAR models

(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, Chapman &Hall.
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 Ed. Edward Arnold.
6. Lutkepohl, H. (2005). New Introduction to Multiple Time Series Analysis, Springer
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 O7: Advances in Generalized Linear Models [ST 204]

Unit I

Generalized Linear Models: Model Fitting and Inference, Exponential Dispersion Family Distributions, Likelihood and Asymptotic Distributions, Likelihood-Ratio/Wald/Score Methods of Inference, Parameters, Deviance, Model Comparison, and Model Checking, Goodness of Fit

(12L + 3T)

Unit II

Binary logistic models, Nominal Responses: Baseline-Category Logit Models, Ordinal Responses: Cumulative Logit and Probit Models, Probit and Complementary Log–Log Models, Multinomial Response Models

(10L + 3Lab)

Unit III

Models for Count Data, Poisson GLMs for Counts and Rates, Poisson/Multinomial Models for Contingency Tables, Negative Binomial GLMS, Models for Zero-Inflated Data

Quasi-Likelihood Methods ,Variance Inflation for over dispersed Poisson and Binomial GLMs, Beta-Binomial Models and Quasi-Likelihood Alternatives, Quasi-Likelihood and Model Misspecification

(12L + 3Lab)

Unit IV

Modeling Correlated Responses, Marginal Models and Models with Random Effects
Normal Linear Mixed Models, Fitting and Prediction for Normal Linear Mixed Models, Binomial and Poisson GLMMs, GLMM Fitting, Inference and Prediction
Marginal Modeling and Generalized Estimating Equations (GEE)
Bayesian Generalized Linear Models, Empirical Bayes and Hierarchical Bayes Modeling
Applications in Survival analysis, Insurance, Engineering, Correlated Survey responses etc .

(14L + 3Lab)
Total (48 L+ 12Lab)

Books Recommended

1. Agresti, A. (2015). Foundations of Linear and Generalized Linear Models, Wiley
2. Dobson, A. J. (2002). An Introduction to Generalized Linear Models, 2nd Ed. Chapman & Hall
3. Jiang, J. (2007). Linear and Generalized Linear Mixed Models and their Applications, Springer
4. Jong, P. and Heller, G. Z. (2008) Generalized Linear Models for Insurance Data, Cambridge University Press.
5. Lindsey, J. K. (1997). Applying Generalized Linear Models, Springer
6. McCullagh, P. and Nelder, J. A. (1989). Generalized Linear Models, Chapman & Hall
7. McCulloch, C. E. and Searle, S. R. (2001). Generalized, Linear and Mixed Models, Wiley
8. Stroup, W. W. (2013). Generalized Linear Mixed Models, Modern Concepts, Methods and Applications, CRC Press

ST O8 Astrostatistics [ST 202, ST 203, ST 204]

Unit I

Introduction to Basics of Astronomy, Historical data driven discoveries in astronomy, least squares theory, light and radiation, brightness of stars, magnitude and distance, relation between magnitude and luminosity, stellar parallax and stellar distances, populations, clusters (galactic, globular etc.), galaxies, quasars, pulsars, bursts etc.

(10 L)

Unit II

Astronomical Data, Sources, Properties and Imputation Techniques Sloan Digital Sky Survey, Vizier Service, Data on Eclipsing Binary Stars, Extra Galactic Distance Data Base (EDD), Data on Pulsars, Gamma ray bursts etc.

(10 L)

Unit III

Astronomical Data Analysis using Statistics and Machine Learning Techniques, EM algorithm, advanced regression and measurement error techniques, classification and clustering,

(5L + 15Lab)

Unit IV

Aapplications of time series and spatial processes, Bayesian computing and machine learning algorithms

(5L + 15 Lab)

Total (30L + 30 Lab)

Note: The concepts from above topics may be illustrated using case studies related to Cosmology, Big Bang, LSS, Gravitational Wave, LIGO, TMT, SKA, Stellar Spectra and their Classification, Solar Corona, Helioseismology, Neutron Stars, Pulsars etc.

Recommended Books

1. Andreon, S. and Weaver, B. (2015). *Bayesian Methods for the Physical Sciences: Learning from Examples in Astronomy and Physics*, Springer
2. Babu, G. J. and Fiegelson, E. D. (1996). *Astrostatistics*, CRC Press, London
3. Chattopadhyay, A. K. and Chattopadhyay, T. (2014). *Statistical Methods for Astronomical Data Analysis*, Springer
4. Feigelson, E. D. and Babu, G. J. (2012). *Modern Statistical Methods for Astronomy with R Applications*, Cambridge University Press
5. Feigelson, E. D. and Babu, G. J. (2012). *Statistical Challenges in Modern Astronomy V*, Springer.
6. Hilbe, J. M. (Ed.), (2013). *Astrostatistical Challenges for the New Astronomy*, Springer
7. Jenkins, C. R and Wall, J. V.(2012) *Practical Statistics for Astronomers*, 2nd Edition, Cambridge
8. Sarro, L.M., Eyer, L., O'Mullane, W., De Ridder, J. (Eds.) (2012). *Astrostatistics and Data Mining*, Springer

ST O10 Optimization Techniques [ST 102, ST 104]

Unit I

Linear Programming: Review

Computational complexity of LPP, Ellipsoid method, Polynomial time algorithm, Karmarkar's polynomial time algorithm, Convergence and complexity,

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, cutting planes

Dynamic programming, sensitivity, Bellman's optimality principle, Stochastic dynamic programming

(16L+4Lab)

Unit II

Transportation and Assignment problems, recent developments

(8L+2Lab)

Unit III

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

(12L+3Lab)

Unit IV

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+3Lab)

Total (48 L+ 12Lab)

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. Hiller, F.S. and Lieberman, G.J., (2009). Introduction to Operations Research (9th ed.), McGraw-Hill
6. Kambo, N.S. (1991). Mathematical Programming Techniques. Affiliated East-West press.
7. Panneerselvam, R. (2012). Operations Research, 2nd Edn. Prentice Hall of India.
8. Sinha, S. M. (2006) Mathematical Programming: Theory and Methods, Elsevier's
9. Taha, H. A. (2016) Operations Research: An Introduction, 10th edition, Prentice Hall
10. Winston, W.L., (2003) Introduction to Mathematical Programming (4th ed.), Duxbury Press

ST O11 Reliability and Statistical Quality Control [ST 203]

Unit I

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.

Shape of the system reliability function, applications to relay circuits and safety monitoring systems, Notion of aging and 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.

(14 L + 4 T)

Unit II

Classes of life distributions applicable in replacement models, Shock models, Age replacement and block replacement policies, Renewal theory useful in replacement models, Replacement policy comparisons, preservation of life distribution classes under reliability operations.

(10L + 2 T)

Unit III

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.

Economic designing of control charts, Duncan's model, 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.

(14L+3T)

Unit IV

Process Capability Analysis

Synthetic and 'Group Runs' (GR) control charts, Multi-Attribute control charts, Multivariate control charts for mean vector and covariance matrix.

Acceptance Sampling plans, Chain sampling plans, Bayesian sampling plans

(10L+3T)

Total (48 L+ 12T)

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. Guenther, W. C. (1977). Sampling Inspection in Statistical Quality Control, Alan Stuart.
4. Levenson, W. (2011). Statistical Process Control for Real-World Applications. CRC Press.
5. Montgomery, D. C. (2005). Introduction to Statistical Quality Control, Wiley.
6. Tobias, P. A. and Trindane, D. C. (1995). Applied Reliability, Second edition. CRC Press.

ST O12 Computer-Intensive Statistical Methods [ST 105, ST 203]

Unit I

Review of ARM, Matropolis-Hastings and Gibbs Sampling algorithms. Particle Filtering, Rejection algorithms for Approximate Bayes Computation (ABC-Rejection), Inference in Hidden Markov Models (HMM)

(10L+6Lab)

Unit II

Missing Values and Imputations Techniques: Missing values and types of missingness, imputations methods for missing values, single and multiple imputations. MCMC methods for missing values, EM Algorithm and Applications: EM algorithm for incomplete data, EM algorithm for mixture models, EM algorithm for missing values, stochastic EM algorithm.

(10L+6Lab)

Unit III

Bootstrap methods, estimation of sampling distribution, various types of confidence intervals, variance stabilizing transformation, Jackknife and cross-validation, Permutation tests. Bagging and Boosting methods with applications. Cross validation

(10L + 3Lab)

Unit IV

Smoothing techniques: Kernel estimators, nearest neighbor estimators, orthogonal and local polynomial estimators, wavelet estimators, Splines, Choice of bandwidth and other smoothing parameters. Statistical methods for Big Data analytics

(10L+5Lab)

Total (40 L+ 20 Lab)

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. Effron, B and Hastie, T (2016). Computer-Age Statistical Inference-Algorithms, Evidence and Data Science, Cambridge University Press.
5. Gilks, W. R., Richardson, S., and Spiegelhalter, D. (eds.) (1995) Markov Chain Monte Carlo in Practice. Chapman and Hall.
6. Good, P. I. (2005) Resampling Methods: A Practical Guide to Data Analysis. Birkhauser Bosel.
7. Jim, A. (2009). Bayesian Computation with R, 2nd Edn, Springer.
8. McLachlan, G.J. and Krishnan, T. (2008) The EM Algorithms and Extensions. Wiley.

ST O13 Statistical Learning and Data Mining [ST 202, ST 204]

Unit I

Supervised Learning: K - nearest neighbourhood algorithm, Decision trees, Naïve Bayes and Bayesian networks.

(12L + 5 Lab)

Unit II

Support-Vector Machines and Kernel Methods, Optimal Separating Hyperplane, Soft-Margin Classifier, SVM Criterion as Loss Plus Penalty, Computations and the Kernel Trick, Function Fitting Using Kernels, Kernel Smoothing and Local Regression, Model evaluation techniques, Cost-Benefit analysis using data driven costs

(12L + 5 Lab)

Unit III

Unsupervised Learning: Hierarchical and k-means clustering, Kohonen networks, BIRCH clustering, Measuring cluster goodness

Graphical evaluation of classification, Association rules, Genetic algorithms, Imputation of missing data.

(12L + 4 Lab)

Unit IV

Neural Networks and the Handwritten Digit Problem, Fitting a Neural Network, Autoencoders, Deep Learning, Learning a Deep Network

(8L + 2 Lab)
Total (44 L+ 16Lab))

Books Recommended

1. Alpaydin, E. (2014), Introduction to Machine Learning, 3rd Ed. MIT Press.
2. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth and Brooks.
3. Hastie T., Tibshirani R. and Friedman J. H., (2008). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer.
4. James G., Witten, D., Hastie, T. Tibshirani, R. (2013). An Introduction to Statistical Learning: With Applications in R, Springer
5. Larose, D. T. and Laros, C. (2015). *Data Mining and Predictive Analytics*. Wiley.
6. Mohammad J. Zaki and Wagner Meira. (2014). Data Mining and Analysis. Fundamental Concepts and Algorithms. Cambridge University Press, New York.
7. Ripley,B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press
8. Shmueli, G., Patel, N. Bruce, P. (2010). *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XL Miner*, Wiley.

ST O14 Survival Analysis [ST 203, ST 204]

Unit I

Survival data, Concepts of time, order and random and hybrid censoring,

Life distributions - exponential, gamma, Lognormal, Pareto, linear failure rate, Ageing classes - IFR, IFRA, NBU, NBUE, HNBUE and their duals, Bathtub failure rate.

Parametric inference, point estimation, confidence Intervals, scores, tests based on LR, MLE

(12L+3Lab+2T)

Unit II

Life tables, failure rate, mean residual life and their elementary properties.
Estimation of survival function - Actuarial estimator, Kaplan - Meier estimator, Estimation under the assumption of IFR/DFR

(10L+2Lab+2T)

Unit III

Semi-parametric regression for failure rate - Cox's proportional hazards model, partial likelihood, estimation and inference methods for the Cox models, time-dependent covariates, residuals and model diagnosis, functional forms of the Cox models, goodness-of-fit tests for the Cox models,

Competing risk models, Repair models, Probabilistic models, Joint distribution of failure times
Unconditional tests for the time truncated case, Tests for exponentiality, two sample non-parametric problem.

(10L+3Lab+2T)

Unit IV

Nelson-Aalen estimators, counting processes and martingales, modeling counting processes, Regression models for modeling multiple events,
Frailty models, Shared frailty models, Identifiability of frailty models, Frailty regression models, Bivariate and correlated frailty models, Additive frailty models

(10L+2 Lab+2T)

Total (42 L + 10 Lab + 8 T)

Books Recommended

1. Klein, J. P. and Moeschberger, M. L. (1997). Survival Analysis: Techniques for Censored and Truncated Data , Springer, New York
2. Collett, D. (2003). Modelling Survivaldata in Medical Research, Second Edition, Chapman & Hall/CRC
3. Therneau, T. M. and Grambsch, P. M. (2000). Modeling Survival Data, Extending the Cox Model, Springer, New York.
4. Cox, D.R. and Oakes, D. (1984). Analysis of Survival Data, Chapman and Hall.
5. Deshpande, J.V. and Purohit, S.G. (2005). Life Time Data: Statistical Models and Methods, Word Scientific.
6. Duchateau, L. and Johnson, P. (2008). The Frailty Model. Springer: New York.
7. Hanagal, D. D. (2011). Modeling Survival Data Using Frailty Models. CRC Press.
8. Kalbfleish, JD. and Prentice, RL. (2002). The Statistical Analysis of Failure Time Data. New York: Wiley.
9. Hougaard, P. (2000). Analysis of Multivariate Survival Data. Springer: New York.
10. Wienke, A. (2011). Frailty Models in Survival Analysis, CRC Press: New York.

ST O15 Actuarial Statistics [ST 103]

Unit I

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 II

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 death (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 III

Annuity contracts, annuity certain, discrete annuity, monthly 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 IV

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 monthly 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 O16 Medical and Health Statistics [ST 103, ST 204]

Unit I

Study designs in epidemiology. Measures of disease occurrence and association, variation and bias, Identifying non-causal association and confounding, communicating results of epidemiological studies, ethical issues in epidemiology

(12 L + 3 Lab)

Unit II

Defining and assessing heterogeneity of effects, interaction. Sensitivity and specificity of diagnostic test, Cohort Study designs, statistical power and sample size computations.

(12 L + 3 Lab)

Unit III

Log-linear models, 2xk and 2x2x2 contingency tables, Logistic model, Analysis of binary data. Causal Inference, Longitudinal data Cross-control study designs, matched case-control studies

(12 L + 3 Lab)

Unit IV

Survival data, Product-limit estimator, Proportional hazards model, multivariate survival data Agreement and Reliability
Meta analysis

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

Recommended books

1. Agresti, A. (2002). Categorical Data Analysis, Wiley
2. Armitage, P., Berry, G., Matthews, J. N. S. (20) Statistical Methods in Medical Research, 4th Ed., Blackwell
3. Bland, M. (2015). An Introduction to Medical Statistics, OUP
4. Brookemeyer, R. and Gail, M. H. (1994). AIDS Epidemiology : A Quantitative Approach, OUP
5. Clayton, D. and Hills, M. (2013). Statistical methods in Epidemiology, OUP
6. Daniel, W. W. and Cross, C. L. (2012). Biostatistics: A Foundation for Analysis in the Health Sciences, 10th Edition, Wiley
7. Diggle, P. J., Heagerty, P., Liang, K-Y and Zeger, S. L. (2013). Analysis of longitudinal data, OUP Harris, M. and Taylor, G. (2014). Medical Statistics Made Easy, Vol. 1-3, Scion
8. Matthews, D. E. and Farewell, V. T. (2015). Using and Understanding Medical Statistics, 5th Ed., Karger
9. McCullagh, P. and Nelder, J. A. (1999). Generalized Linear Models, Chapman & Hall
10. Pagano, M. and Gauvreau, K. (2018). Principles of Biostatistics, Taylor & Francis
11. Piantadosi, S. (2017), Clinical Trials, 3rd Ed., Wiley
12. Rao, P. S. R. S. (2017). Statistical Methodologies with Medical Applications, Wiley
13. Rosner, B. (2010). Fundamentals of Biostatistics, Harvard
14. Selvin, S. (2004). Statistical Analysis of Epidemiological Data, 3rd Ed., OUP.
15. Zhou, X. H., Obuchowski, N. A., and McClish, D. K. (2011). Statistical Methods in Diagnostic Medicines, 2nd Ed., Wiley

ST O17 Analysis of Clinical Trials [ST 204, ST 302]

Unit I

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 pharmaco dynamics, two-compartment model

(12L+1Lab+2T)

Unit II

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 III

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 IV

Designs based on clinical endpoints: Weighted least squares method, log-linear models, generalized estimating equations, Drug interaction study, dose proportionality study, steady state analysis. Meta analysis, 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 Ed. CRC Press.
2. Chow S.C. and Liu J.P. (2004). Design and Analysis of Clinical Trials. 2nd Ed. 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 O18 Statistical Methods in Micro-array Data Analysis [ST 202]

Unit I

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, quantile and LOWESS normalization, stage wise normalization, Concordance coefficient and its role in normalization

(12L + 3Lab)

Unit II

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. Paired t –test, Tests for validating assumptions of paired t-test. Wilcoxon signed rank test, Comparison of more than two types of mRNA samples in single channel or two channel microarray experiments. 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 + 3Lab)

Unit III

Multiple hypotheses testing and Principal component analysis Multiple hypotheses testing, 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 rectangular matrix and the concept of ballot, its application to microarray

(12L + 3Lab)

Unit IV

Cluster analysis and Logistic regression, Hierarchical cluster analysis of microarray data, K - means cluster analysis of microarray data, Application of logistic regression for microarray data, Concept of AIC and BIC and its role to identify marker genes

(12L + 3Lab)

Total (48 L+ 12Lab)

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 O19 Statistical Methods for Bio-computing [ST 105, ST 301]

Unit I

Type of genetic 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)

Unit II

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 Lab)

Unit II

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+ 4 Lab)

Unit III

Applications of HMM 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 + 4 Lab)

Total (48 L+ 12 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 O20 Bayesian Inference [ST203]

Unit I

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

(14L + 5 T)

Unit II

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.

(8L + 3 T)

Unit III

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.

(8L + 2 T)

Unit IV

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

(10L + 10 Lab)

Total (40 L+ 10 T + 10 Lab)

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

ST O21 Expert Systems with Applications [ST 105]

Unit I

Introduction of expert system, fuzzy sets, operations on fuzzy sets, fuzzy arithmetic and relations, fuzzy logic and controllers, multi criteria decision making, multi attribute decision making, fuzzy decision making, applications of fuzzy set theory in reliability

(12 L + 3 Lab)

Unit II

Introduction to genetic algorithm (GA), Binary GA, Real coded GA, parameters of GA, different types of crossover operators, multi objective GA, Pareto solutions, Elitist and non-elitist solutions, non-dominated sorting GA

(12 L + 3 Lab)

Unit III

Constrained multi-objective GA, Modified GA, Convergence of GA, Elitist and non-elitist GA under constrained optimization

(12 L + 3 Lab)

Unit IV

Introduction to Neural network, problem of network learning, multi-layer networks and back-propogation, Hebbian learning, Statistical aspects of learning, VC dimensions, Gauss-Newton Method. Hybridization of fuzzy, GA and NN with real world case studies.

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

Recommended books

1. Goldberg, D. E. (1989). Genetic Algorithms, Addison-Wesley Professional
2. Deb, K. (2001). Multi-objective Optimization Using Evolutionary Algorithms, Wiley
3. Mitchell, T. (1997). Machine Learning, McGraw-Hill
4. Langley, P. (1995). Elements of Machine Learning, Morgan Kaufmann Series in Machine Learning
5. Klir, G. J. and Yuan, Bo. (1995). Fuzzy Sets and Fuzzy Logic: Theory and Application, Prentice Hall
6. Donald Waterman. (1986). A Guide to Expert Systems. Pearson India.

ST O22 Financial Econometrics [ST 204, ST 301]

Unit I

Frequency domain analysis of time series, Periodicity, Spectral density, Periodogram and DFT, Spectral representation, Inference, Examples & data analysis

Vector Time Series, Cross correlations, VAR and VARMA models, Stability condition, Impulse response function, Causality & Invertibility, Estimation – least squares and MLE (Gaussian), Order determination, multivariate Portmanteau tests

Testing for unit roots: the DF, ADF, PP and KPSS test statistics.

Granger causality, Cointegration and ECM, Cointegrating VAR, Cointegration tests, 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

(20L + 4 Lab)

Unit II

State-space models, State-space representations, local-trend model, The basic structural model, State-space representation of ARIMA models, Filtering and smoothing, The Kalman recursions, Estimation for State-Space models, Generalized state-space models, Parameter & observation-driven models, Non-Gaussian state-space models, APM with time varying parameters, Examples & data analysis

(12 L + 3 Lab)

Unit III

Nonlinear models, Modeling regime shifts, Markov-switching models, Estimation, Bayesian methods, Diagnostic checking, Forecasting, Examples & data analysis

(8 L + 3 Lab)

Unit IV

Volatility definition and estimation, Volatility forecast evaluation, Stochastic volatility models, MCMC approach, Option pricing with stochastic volatility, Examples

(8L + 3 Lab)

Total (48 L+ 12Lab)

Recommended Books

1. Ait-Sahalia, Y. & Hansen, L. P. (Ed.) (2010). *Handbook of Financial Econometrics: Tools and Techniques*, Vol.1& 2, Elsevier, Amsterdam.
2. Campbell J., Lo A. & McKinley C. (1997). *The Econometrics of Financial Markets*. PrincetonUniversity press. Princeton.
3. Fan J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer, New York.
4. Hamilton (1994). *Time Series Analysis*. Princeton University press. Princeton
5. Knight, J. and Satchell, S. (2007). *Forecasting Volatility in the Financial Markets*, Third Edition, Elsevier, Amsterdam.
6. Lutkepohl, H. and Kratzing, M. (Ed.) (2004). *Applied Time Series Econometrics*, Cambridge University Press, Cambridge
7. Poon Ser-Huang (2005). *A Practical Guide to Forecasting Financial Market Volatility*, Wiley, New York.

8. Rachev. S. T., Mittnik, S., Fabozzi, F. J. ,Focardi, S. M. and Jasic, T. (2007). *Financial Econometrics :From Basics to Advanced Modeling Techniques*, Wiley, New York.
9. Ruppert, D. (2004). *Statistics and Finance: An Introduction*, Springer, New York.
10. Shephard, N. (2004). *Stochastic Volatility: Selected Readings*, Oxford University Press, Oxford.
11. Soderlind, P. (2010). *Lecture Notes in Financial Econometrics*, University of St. Gallen, Switzerland.
12. Tsay, R. S. (2010). *Analysis of Financial Time Series*, Third Edition, Wiley, New York.
13. Wang, P. (2003). *Financial Econometrics: Methods and Models*, Routledge, London.