M.Sc. (Applied Statistics and Informatics) Course Curriculum

FIRST YEAR												
Semester I							**Semester II**					
Course No.	**Name**	**L**	**T**	**P**	**C**		**Course No.**	**Name**	**L**	**T**	**P**	**C**
SI 431	Introduction to Data Analysis using R	2	0	2	6		SI 424	Statistical Inference I	3	1	0	8
SI 423	Linear Algebra and Applications	3	1	0	8		SI 404	Applied Stochastic Processes	3	1	0	8
SI 427	Probability I	3	1	0	8		SI 416	Optimization	2	1	0	6
SI 419	Combinatorics	3	1	0	8		SI 422	Regression Analysis	3	1	0	8
SI 429	Real Analysis	3	1	0	8		SI 426	Algorithms	3	1	0	8
	Total Credits	**14**	**4**	**2**	**38**			**Total Credits**	**14**	**5**	**0**	38
SECOND YEAR												
Semester III							**Semester IV**					
Course No.	**Name**	**L**	**T**	**P**	**C**		**Course No.**	**Name**	**L**	**T**	**P**	**C**
ES 200/ HS 200	Environmental Studies /Dept. Elective/Institute Elective	-	-	-	6		SI 509	Time Series Analysis	3	1	0	8
							SI 526	Experimental Designs	2	1	0	6
SI 505	Multivariate Analysis	3	1	0	8			Elective III	2	1	0	6
SI 503	Categorical Data Analysis	3	1	0	8			Elective IV	2	1	0	6
	Elective I	3	1	0	8			Elective V	2	1	0	6
	Elective II	3	1	0	8		SI 598	Project II/Dept. Elective/Institute Elective	-	-	-	6
SI 593	Project I (Optional)	-	-	-	4							
	Total Credits	**12**	**4**	**0**	**38**			**Total Credits**	**11**	**5**	**0**	**38**

Electives – Semester III		**Electives – Semester IV**	
Elective I-II		**Elective III - V**	
SI 515	Statistical Techniques in Data Mining	SI 534	Nonparametric Statistics
SI 541	Statistical Epidemiology	SI 514	Statistical Modeling
SI 507	Numerical Analysis	SI 548	Computational Statistics
SI 513	Theory of Sampling	SI 536	Analysis of Multi-type & Big Data
SI 537	Probability II	SI 527	Introduction to Derivative Pricing
SI 543	Asymptotic Statistics	SI 544	Martingale Theory
		SI 546	Statistical Inference II

1st Semester

SI 431 Introduction to data analysis using R 2 0 2 6

Overview of R software, Data Frames, R Scripts, creating, importing/exporting and merging of datasets, creating matrices and basic matrix operations in R, 2d/3d plotting, programming in R (for, if else, do and while loops), functions, creating report using R markdown

Exploring data using R, Scatter plot, histogram, bar chart, pie chart, box plot, basic statistics computation (mean, median, variance etc.)
Generating random samples from standard distributions (such as Bernoulli, Poisson, Normal, Exponential etc.) and comparing theoretical pdfs/pmfs using histograms/frequency distributions, quantiles of sampling distributions (t, chi and F distribution)
Maximization/minimization of functions in R (some algorithm), MLE estimation.

Polynomial fitting of scatter plot, introducing regression line, least squares estimates, residual plots, testing normality of residuals (qqplot), goodness of fit measures and tests, testing of regression parameters, simulation of regression model, empirical distribution of least square estimator and its comparison with theoretical distribution

Simulation of multivariate normal random vectors, estimation of mean and covariance matrix, eigen values and eigen vector of variance covariance matrix, spectral decomposition covariance matrix.
Generating dependent random variables with some models like (random walk, AR(1), MA(1) etc)

Texts / References:

FOSSEE, Spoken tutorials at https://r.fossee.in/

James, G., Witten, D., Hastie, T., and Tibshirani, R., An introduction to statistical learning with applications in R, Springer, New York, 2013.

Wickham, H., Advanced R, CRC press, New York, 2015.

Wickham, H., and Grolemund, G., R for Data Science, O'Reilly Media Inc, Canada, 2017.

SI 423 Linear algebra and applications 3 1 0 8

Vector spaces (with emphasis over R and C): Subspaces, linear dependence and independence, basis and dimension.

Linear transformations: Rank-nullity theorem, matrix representation of a linear transformation, invertibility and isomorphism, effect of change of basis on the matrix representation of a linear transformation, dual spaces.

Review of elementary properties of determinants, Cramer's rule.

Diagonalization: Eigenvalues and eigenvectors, algebraic and geometric multiplicities of an eigenvalue, diagonalizability, invariant subspaces and Cayley-Hamilton theorem.

Inner product spaces: Gram-Schmidt orthogonalization, adjoint of a linear operator, normal and self-adjoint operator, orthogonal projections and the spectral theorem, singular value decomposition and pseudoinverse, bilinear and quadratic forms.

Canonical forms: Jordan canonical form (with emphasis on computation).

Texts / References:

Rao, A.R., and Bhimashankaram, P., Linear algebra, 2nd edition, Hindustan book agencey, New Delhi, 2000.

Friedber, S.H., Insel, A.J., and Spence, L.E., Linear algebra, 4th edition, PHI learning, New Delhi, 2011.

Strang, G., Linear algebra and its applications, 4th edition, Thomson Learning, Toronto, 2006.

SI 427 Probability I 3 1 0 8

Random phenomena, sample spaces, events, sigma algebra, probability space, properties of probability, conditional probability, independence, Bayes formula, Polya's urn model.

Discrete random variable, probability mass function, independent random variables, sum of random variables, random vector, expectation of discrete random variable, properties of expectation and variance.

Continuous random variable, distribution function, density of a continuous random variable, expectation, change of variable formula, random vector, joint distribution of random variables, joint density, distribution of sums and products of random variables, conditional density, conditional expectation, order statistics, moment generating function, characteristic function, brief introduction to moment problem.

Inequalities: Markov, Chebyshev, Schwarz and Chernoff bound.

Convergence in probability, almost sure convergence, convergence in distribution, relation between these three modes of convergences, weak law of large numbers (WLLN), strong law of large numbers (SLLN), central limit theorem (CLT).

Texts / References:

Athreya, K.B. and Lahiri, S. N., Probability Theory, Hindustan Book Agency, 2006.

Billingsley, P., Probability and Measure, 2nd edition, John Wiley & Sons, New York, 1995.

Hoel, P.G., Port, S.C., and Stone, C.J., Introduction to Probability Theory, Universal Book Stall, New Delhi, 1998.

Karr, A.F., Probability, Springer-Verlag, New York, 2003.

Rosenthal, J.S., A first look at rigorous Probability theory, 2nd edition, World Scientific, 2006.

Ross, S., A first course in Probability, 9th Edition, Pearson, Delhi, 2019.

SI 419 Combinatorics 3 1 0 8

Counting Basic Combinatorial objects: Sets, Multisets, Partitions of sets, Partitions of numbers, Permutations, Trees, Partially ordered sets.

Generating functions, Recurrence relations, Principle of Inclusion-Exclusion.

Graph Theory Graphs and Directed graphs, Paths, Walks, Connectivity, Matchings in bipartite graphs, Network flows, Dilworths theorem.

Texts / References:

Bona, M., A walk through Combinatorics, 4th Edition, World Scientific, New Jersey, 2017.

Nesetril, J., and Matousek, J., Invitation to Discrete Mathematics, 2nd edition, Oxford University Press, Oxford, 2009.

Lehman, E., Leighton, F. T., and Meyer, A. R., Mathematics for Computer science, Samurai Media Limited, Surrey, 2019. (Freely available online)

SI 429 Real analysis 3 1 0 8

Review of sequences and series of real numbers. Limit superior and limit inferior, Cauchy sequences and completeness of R.
Tests for convergence of series of real numbers.

Basic notions of Metric Spaces with emphasis on R n . Heine Borel Theorem.

Continuity and Uniform continuity.
Derivatives. Mean Value Theorem and applications. Functions of bounded variation.

Riemann-Stieltjes integral. Improper integrals and Gamma function.

Sequences and series of functions. Uniform convergence, interchanging limits with integrals and derivatives. Arzela-Ascoli theorem (statement only).

Functions of several variables: Partial derivative, directional derivative, total derivative; Mean value theorem, Taylor's theorem.

Texts / References:

Ajit kumar, and Kumaresan, S., A basic course in Real analysis, CRC Press, Boca Raton, 2014.

Apostol, T.M., Mathematical analysis, 2nd edition, Narosa Publishers, New Delhi, 2002.

Bartle, R.G., and Sherbert, D. R., Introduction to Real analysis, 4th edition, John Wiley, New York, 2011.

Ghorpade, S.R., and Limaye, B.V., A course in Calculus and Real analysis, Springer (India), New Delhi, 2006.

Ross, K.A., Elementary analysis: The theory of Calculus, 2nd edition, Springer (India), New Delhi, 2013.

Tao T., Analysis I, 3rd Edition, Hindustan Book Agency, New Delhi, 2006.

2nd Semester

SI 424 Statistical Inference I 3 1 0 8

Prerequisites: SI 427 (Exposure) (For students from other departments, instructor's permission will be required)

Distributions of functions of random variables, Sampling distributions, Order statistics, Sufficiency, completeness and ancillary statistic, exponential family of distributions, Methods of estimation (Method of Moments, MLE and Bayesian), Unbiased estimators, Evaluating estimators, UMVUEs, Testing, Likelihood Ratio tests, UMP tests, unbiased tests, Interval estimation, Consistent and efficient estimators.

Texts / References:

Casella, G. and Berger, R. Statistical Inference, 1st Edition, Duxbury Press, Pacific Grove, 2002.

Hogg, R., McKean. J. and Craig, A., Introduction to Mathematical Statistics, 8th Edition, Pearson, Boston, 2019.

Lehmann, E. Theory of Point Estimation, 1st Edition, John Wiley & Sons, New York, 1983.

Lehmann, E and Romano, J. Testing Statistical Hypotheses, 3rd Edition, Springer-Verlag New York, 2005.

DeGroot, M. and Schervish, M. Probability and Statistics, 4th Edition, Addison Wesley, Boston, 2002.

SI 404 Applied Stochastic Processes 3 1 0 8

Prerequisite**:** SI 427 (Exposure) (For students from other departments, instructor's permission will be required)

Stochastic processes: description and definition. Markov chains with finite and countably infinite state spaces. Classification of states, irreducibility, ergodicity. Basic limit theorems. Statistical Inference. Random walk in one dimension, Reflection principle, last visits and first passages. Applications to queuing models. Markov processes with discrete spaces. Poisson process, pure birth process, birth and death process.

Texts / References:

Bhat, U. N. and Miller, G.K., Elements of Applied Stochastic Processes, 3rd edition, John Wiley & Sons, New York, 2002.

Feller, W., An Introduction to Probability Theory and its Applications, Vol. I, 3rd edition, John Wiley & Sons, New York, 1968.

Hoel, P.G., Port, S.C. and Stone, C.J., Introduction to Stochastic Processes, Waveland Press, Inc., Long Grove, Illinois, 1986.

Kulkarni, V.G., Introduction to Modeling and Analysis of Stochastic Systems, 2nd Edition, Springer, New York, 2011.

Kulkarni, V.G., Modeling and Analysis of Stochastic Systems, 3rd Edition, Chapman and Hall/CRC, Boca Raton, 2017.

SI 416 Optimization 2 1 0 6

Unconstrained optimization using calculus (Taylor's theorem, convex functions, coercive functions). Unconstrained optimization via iterative methods (Newton's method, Gradient/ conjugate gradient based methods, Quasi-Newton methods). Constrained optimization (Penalty methods, Lagrange multipliers, Karush-Kuhn-Tucker conditions).

Introduction to Linear Programming: Lines and hyperplanes, Convex sets, Convex hull, Formulation of a Linear Programming Problem, Theorems dealing with vertices of feasible regions and optimality, Graphical solution. Simplex method (including Big M method and two-phase method), Dual problem, Duality theory, Dual simplex method, Revised simplex method.

Texts /References:

Beale, E.M.L., and Mackley, L. , Introduction to Optimization, John Wiley & Sons, Hoboken, 1988.

Chavatal, V., Linear Programming, W.H. Reeman and Company, New York, 1983.

Chong, E.P.K. and Zak, S.H., An Introduction to Optimization, 4th Edition, John Wiley & Sons, Hoboken, 2013.

Joshi, M.C., and Moudgalya, K., Optimization: Theory and Practice, Narosa, New Delhi, 2004.

Nocedal, J. and Wright, S. J., Numerical Optimization, 2nd Edition, Springer, New York, 2006.

Vanderbei, R.J., Linear Programming Foundations and Extensions, 3rd Edition, Springer, New York, 2008.

SI 422 Regression Analysis 3 1 0 8

Prerequisites: SI 427 (Exposure) (For students from other departments, instructor's permission will be required)

Simple and multiple linear regression models – estimation, tests and confidence regions. Simultaneous testing methods- Bonferroni method etc.

Analysis of Variance for simple and multiple regression models.

Analysis of residuals. Lack of fit tests. Checks (graphical procedures and tests) for model assumptions: Normality, homogeneity of errors, independence, correlation of covariates and errors.

Multicollinearity, outliers, leverage and measures of influence.
Model selection (stepwise, forward and backward, best subset selection) and model validation. Discussion of algorithms for model selection.

Regression models with indicator variables. Polynomial regression models. Regression models with interaction terms.

Transformation of response variables and covariates. Variance stabilizing transformations, Box-Cox method.

Ridge's regression. Weighted Regression.

Texts / References:

Draper, N. and Smith,H. Applied Regression Analysis, 3rd Edition, John Wiley and Sons Series in Probability and Statistics, New York, 1998.

Montgomery, D., Peck, E., Vining, G. Introduction to Linear Regression Analysis, 5th Edition, John Wiley, New York, 2012.

Sen, A. and Srivastava, M. Regression Analysis – Theory, Methods & Applications, 1st Edition, Springer-Verlag Berlin Heidelberg, New York, 1990.

Kutner, M., Nachtsheim, C., Neter, J. and Li, W. Applied Linear Statistical Models, 5th Edition, McGraw-Hill Companies, Boston, 2005.

SI 426 Algorithms 3 1 0 8

Basics Algorithm analysis and asymptotic notation, Linked lists.

Graphs Breadth first search, Depth first search, Strongly connected com- ponents.

Divide and Conquer Mergesort, Fast Fourier transform.

Greedy Algorithms Dijkstra's algorithm, Minimum spanning tree algo- rithms, Huffman codes and data compression.

Dynamic programming Longest increasing sequences, edit distance, short- est paths.

Network flows Max-flow Min-cut theorem, max flow algorithms, applica- tion to bipartite matchings.

Introduction to Randomized algorithms randomized quick sort, global min-cut, hashing.

Texts / References:

Dasgupta, S., Papadimitriou, C., and Vazirani, U. (2008) Algorithms, Tata McGraw-Hill.

Kleinberg, J. and Tardos, E. (2006) Algorithm design, Pearson.

Cormen, T., Leiserson, C., Rivest, R., and Stein, C. (2009) Introduction to Algorithms, 3rd edition, MIT Press .

3rd Semester

SI 505 Multivariate Analysis 3 1 0 8

Prerequisites: SI 424 (Exposure) (For students from other departments, instructor's permission will be required)

Aspects of Multivariate Analysis: The organization of data – Arrays, descriptive statistics, graphical techniques.

Data displays and pictorial representations – Multiple two-dimensional scatter plots, graphs of growth curves, stars, Chernoff faces.

Random vectors – mean vectors and covariance matrices and related results

The geometry of the sample, random samples – multivariate distributions, expected values of the sample mean and covariance matrix and their algebraic operations, generalized variance.

K-variate normal distribution. Estimation of the mean vector and dispersion matrix. Random sampling from multivariate normal distribution.

Fisher's Discriminant Analysis; classification analysis of several multivariate normal populations, Logistic regression and classification.

canonical analysis.

Principal components, Factor analysis.

Cluster analysis: Hierarchical and non-hierarchical clustering methods.

Textx/References:

Anderson, T.W., An Introduction to Multivariate Statistical Analysis, 2nd Edition, Wiley, New York, 1984.

Gnanadesikan, R., Methods for Statistical Data Analysis of Multivariate Observations, John Wiley, New York, 1997.

Johnson, R.A. and Wichern, D.W., Applied Multivariate Statistical Analysis, 6th Edition, Upper Saddle River, Prentice Hall, New Jersey, 2007.

SI 503 Categorical Data Analysis 3 1 0 8

Categorical Data-nominal and ordinal random variables.

Two-way contingency tables: Table structure for two dimensions. Ways of comparing proportions. Measures of associations-odds ratio. Sampling distributions. Goodness-of-fit tests, testing of independence. Exact and large sample inference.

Three-way contingency tables, Partial associations, Cochran-Mantel-Haenszel methods. Conditional association and related inference.

Generalized Linear Models (GLMs): components of a GLM.

Logisitc regression models for binary data, inference for logistic regression models, multiple logistic regression with qualitative predictors, exact inference for logistic regression, sample size and power of test.

Loglinear models for two-way and three-way contingency tables, inference for loglinear models, the connection between loglinear-logit regression models.

Multicategory logit models for nominal responses, cumulative logit models for ordinal responses.

Texts / References

Agresti, A., Categorical Data Analysis, 3rd Edition, Wiley, New York, 2013.

Agresti, A., An Introduction to Categorical Data Analysis, 3rd Edition, Wiley, New York, 2019.

Andersen, E.B., The Statistical Analysis of Categorical Data, Springer-Verlag, Berlin, 1994.

Santner, T.J. and Duffy, D., The Statistical Analysis of Discrete Data, Springer-Verlag, New York, 1989.

Sen, A.A. and Srivastava, M., Regression Analysis – Theory, Methods and Applications, Springer-Verlag, New York, 1990.

SI 515 Statistical Techniques in Data Mining 3 1 0 8

Pre-requisite: SI 424 (Exposure), SI 416 (Exposure) and SI 422 (Exposure) (For students from other departments, instructor's permission will be required)

Introduction to Data Mining and its Virtuous Cycle. Cluster Analysis: Hierarchical and Non-hierarchical techniques. Classification and Discriminant Analysis Tools: Support Vector Machine, CART, Random forests, Boosting and Bagging, Fisher's discriminant functions and other related rules, Bayesian classification and learning rules. Dimension Reduction and Visualization Techniques: Principal Component Analysis, Multidimensional scaling. Algorithms for data-mining using multiple nonlinear and nonparametric regression. Neural Networks: Multi-layer perceptron, predictive ANN model building using backpropagation algorithm. Exploratory data analysis using Neural Networks – self organizing maps. Genetic Algorithms, Neuro-genetic model building. Discussion of Case Studies.

Texts / References:

Breiman, L., Friedman, J.H., Olschen, R.A. and Stone C.J., Classification of Regression Trees, Brooks/Cole Publishing, Monterey, 1984.

Hand, D.J., Mannila, H. and Smith, P., Principles of Data Mining, MIT Press, Cambridge, 2001.

Hassoun, M.H., Fundamentals of Artificial Neural Networks, Prentice-Hall of India, New Delhi, 1998.

Hastie, T., Tibshirani, R. and Friedman, J. H., The elements of Statistical Learning: Data Mining, Inference & Prediction, 2nd Edition, Springer Series in Statistics, Springer-Verlag, New York, 2009.

James, G., Witten, D., Hastie, T. and Tibshirani, R., An Introduction to Statistical Learning with Applications in R, 2nd Edition, Springer Texts in Statistics, New York, 2021.

Johnson, R.A. and Wichern, D.W., Applied Multivariate Analysis, 6th Edition, Upper Saddle River, Prentice-Hall, New Jersey, 2015.

Press, J. S., Subjective and Objective Bayesian Statistics: Principles, Models, and Applications, 2nd Edition, Wiley Series in Probability and Statistics, Hoboken, 2002.

SI 541 Statistical Epidemiology 3 1 0 8

Epidemiologic approach to clinical trials: observational studies, cross-sectional studies, designing a case control study, bias in a case-control study, matching issues, cohort studies, design of a cohort study, biases in a cohort study, comparing case and cohort studies, randomized trials, selection of subjects, crossover trials, issues on sample size, recruitment. Case studies to explore above topics.

Spatial Epidemiology: Geographical Representation and Mapping, Spatial Interpolation and Smoothing Methods, Estimation and Inference, Spatial Proximity Indices, Disease Clustering, Spatial Regression, Infectious disease modelling

Survival Analysis in Epidemiology: Functions of survival time, censoring mechanisms, nonparametric estimators of survival function, Cox's proportional hazards model, Cases studies using survival analysis methods in health research (5 weeks)

Texts/ References

Lawson, A. B. (2006) Statistical Methods in Spatial Epidemiology, 2nd Edition., Wiley.

Gordis, Leon (2014) Epidemiology, Fifth edition, Elsevier Saunders.

Kalbfleisch J. D. and Ross L. (2000) Statistical Analysis of Failure Time Data, Second Edition, Wiley.

Lee, E.T., Wang, John Wenyu Hoboken (2003) Statistical methods for survival data analysis 3rd Edition, Wiley Interscience.

SI 507 Numerical Analysis 3 1 0 8

Principles of floating-point computations and rounding errors.

Systems of Linear Equations: factorization methods, pivoting and scaling, residual error correction method.

Iterative methods: Jacobi, Gauss-Seidel methods with convergence analysis, conjugate gradient methods.

Eigenvalue problems: only implementation issues.

Nonlinear systems: Newton and Newton like methods and unconstrained optimization.

Interpolation: review of Lagrange interpolation techniques, piecewise linear and cubic splines, error estimates.

Approximation: uniform approximation by polynomials, data fitting and least squares approximation.

Numerical Integration: integration by interpolation, adaptive quadratures and Gauss methods

Initial Value Problems for Ordinary Differential Equations: Runge-Kutta methods, multi-step methods, predictor and corrector scheme, stability and convergence analysis.

Two Point Boundary Value Problems: finite difference methods with convergence results. Lab. Component: Implementation of algorithms and exposure to public domain packages like LINPACK and ODEPACK.

Texts / References

Atkinson, K.E., An Introduction to Numerical Analysis, Wiley, New York, 1989.

Conte, S.D. and De Boor, C., Elementary Numerical Analysis – An Algorithmic Approach, 3rd Edition, McGraw-Hill, New York, 1981.

Eriksson, K., Estep, D., Hansbo, P. and Johnson, C., Computational Differential Equations Cambridge Univ. Press, Cambridge, 1996.

Golub, G.H. and Ortega, J.M., Scientific Computing and Differential Equations: An Introduction to Numerical Methods, Academic Press, San Diego, 2014.

Stoer, J. and Bulirsch, R., Introduction to Numerical Analysis, 2nd ed., Texts in Applied Mathematics, Vol. 12, Springer Verlag, New York, 1993.

SI 513 Theory of Sampling 3 1 0 8

Principals of sample survey, Probability sampling, Non-probability sampling, Simple random sampling, Estimation of population total, Variance estimation, finite population correction, Random sampling with replacement, linear estimators of population mean, Sampling for proportions and percentages, sample size estimation for proportion as well as continuous data in random sampling.

Stratified random sampling, Estimator of population total and its variance, Optimum allocation, comparison between stratified and simple random sampling, Stratified sampling for proportion and sample size estimation, construction of strata, Number of strata, Quota sampling.

Ratio estimator, estimation of variance from sample, comparison between ratio estimator and best linear unbiased estimator, bias of ratio estimates, ratio estimates in stratified sampling. Regression estimators, Large sample comparison with ratio estimate.

Single stage cluster sampling with equal and unequal cluster sizes, Sampling with probability proportion to size, selection with unequal probabilities with and without replacement, the Horvitz Thompson estimator, Brewer's method, Murthy's method, Rao, Hartley and Cochran method. Two stage sampling with units of equal and unequal sizes.

Introduction to randomized response techniques with examples and estimation.

Texts / References

Chaudhuri, A. and Stenger, H., Survey Sampling: Theory and Methods, Chapman & Hall/CRC, Boca Raton, 2005.

Cochran, W.G., Sampling Techniques, 3rd Edition, John Wiley & Sons, New York, 1977.
Des Raj, Sampling Theory, McGraw-Hill Book Co., New York, 1978.

Mukhopadhyay, P., Theory and Methods of Survey Sampling, Prentice-Hall of India, New Delhi, 1998.

SI 537 Probability II 3 1 0 8

Prerequisite: SI 427 (Exposure) (For students from other departments, instructor's permission will be required)

Probability space, random variables (R, R^d valued), distributions of random variables, change of variables formula, expectation of R valued random variable, Jensen's inequality, Holder's inequality, Chebyshev's inequality, Fatou's lemma, monotone convergence theorem, dominated convergence theorem, product measure, Fubini's theorem, notion of independence of sigma-fields and random variables, Kolmogorov's consistency theorem.

Convergence in probability, almost sure convergence, convergence in distribution, convergence in L^{p}, relation between different modes of convergence, Borel-Cantelli lemma, characteristic function, inversion formula, continuity theorems, Scheffe's lemma, uniform integrability, tightness, Helly's selection principle, moment problem.

Weak law of large numbers, strong law of large numbers, central limit theorem.

Radon Nikodym theorem (statement only), condition expectation: definition and its properties.

Texts/ References

Athreya, K.B. and Lahiri, S. N., Measure Theory and Probability Theory, Springer, New York, 2006.

Ash, R. B., Probability and measure theory, Second edition, Academic Press, Burlington, 2000.

Billingsley, P., Probability and Measure, Anniversary Edition, John Wiley & Sons, Hoboken, 2012.

Chung, K. L., A Course in Probability Theory, Third edition, Academic Press, San Diego, 2001.

Durret, R., Probability: Theory and Examples, Fifth edition, Cambridge University Press, Cambridge, 2019.

Pollard, D., A user's guide to Measure Theoretic Probability, Cambridge University Press, Cambridge, 2002.

SI 543 Asymptotic Statistics 3 1 0 8

Prerequisite: SI 427 (Exposure) and SI 424 (Exposure) (For students from other departments, instructor's permission will be required)

Review of modes of stochastic convergences: Almost sure convergence, convergence in probability, convergence in the p-th moment and their relations. Convergence in distribution

Additional topics in stochastic convergence: Portmanteau theorem (Statement only). Convergence in total variation (Scheffe's theorem). Skorohod representation theorem. Hally-Bray theorems (Statement only). Uniform tightness and Prohorov's theorem for random vectors. Characteristic function. Levy's continuity theorem (statement only). Strong law of large numbers (i.i.d. random variables with finite mean). Weak law of large numbers (finite variance). Levy-Lindeberg central limit theorem.

Delta method and variance stabilizing transformations.

Asymptotic properties of moment estimators, M-estimators and Z-estimators. Strong consistency and asymptotic normality of the MLE. Berry-Essen Theorem (without proof). Argmax theorem (statement without proof). Convergence of U-statistics (without proof) and its applications to linear rank statistics.

Glivenko-Cantelli lemma. Convergence of the Kolmogorov-Smirnov statistics to the Brownian bridge (statement only without proof). Convergence of the quantile process. Almost sure and weak convergence results for maximum of i.i.d. random variables. Order statistics. Renyi's representation theorem for the order statistics of the i.i.d. exponential random variables. Bahadur-Rao representation theorem for the sample quantiles.

Efficiency of tests: Asymptotic power function, consistency and asymptotic relative efficiency

Texts/Reference:

DasGupta A., Asymptotic Theory of Statistics and Probability, Springer, New York, 2008.

Serfling R.J., Approximation Theorems of Mathematical Statistics, Wiley, New York, 2009.

van der vaart A. W. and Wellner J. A., Weak Convergence and Empirical Processes, Springer, New York, 1996.

.

4th Semester

SI 509 Time Series Analysis 3 1 0 8

Prerequisites: SI 424 (Exposure) (For students from other departments, instructor's permission will be required)

Stationary processes – strong and weak, linear processes, estimation of mean and covariance functions. Wald decomposition Theorem. Modeling using ARMA processes, estimation of parameters testing model adequacy, Order estimation. Prediction in stationary processes, with special reference to ARMA processes. Frequency domain analysis – spectral density and its estimation, transfer functions.

Nonlinear ARCH and GARCH models.

Discrete-Valued time series models.

Texts / References:

Brockwell P. and Davis R., Introduction to Time Series and Forecasting, Springer, New York, 2000.
Brockwell P. and Davis R., Time Series: Theory and Methods, Springer, New York, 1991.

Box G.E.P., Jenkins G., Reinsel G. and Ljung, Time Series Analysis-Forecasting and Control, 5th Edition, Wiley, New York, 2016.

Chatfield C., The Analysis of Time Series – An Introduction, 6th Edition, Chapman and Hall / CRC, New York, 2016.
Shumway R.H. and Soffer D.S., Time Series Analysis and Its Applications, 4th Edition, Springer, New York, 2016.

Weiss C. H., An Introduction to Discrete-Valued Time Series Data, John Wiley & Sons, Inc., Chichester, 2018.

SI 526 Experimental Designs 2 1 0 6

Prerequisites: SI 424 Statistical Inference (For students from other departments, instructor's permission will be required)

Linear Models and Estimators, Estimability of linear parametric functions. Gauss- Markoff Theorem. One-way classification and two-way classification models and their analyses. Standard designs such as CRD, RBD, LSD, BIBD. Analysis using the missing plot technique. Factorial designs. Confounding. Analysis using Yates' algorithm. Fractional factorial.

A brief introduction to Random Effects models and their analyses.

A brief introduction to special designs such as split-plot, strip-plot, cross-over designs.

Response surface methodology. Applications using SAS software.

Texts / References

Kshirsagar A.M., A First Course in Linear Models, Marcel Dekker, New York, 1983.

Montgomery D.C. (2019) Design and Analysis of Experiments, 10th Edition, John Wiley & Sons, New York, 2019.

Wu C.F.J. and Hamada M., Experiments: Planning Analysis, and Parameter Design Optimization, John Wiley & Sons, New York, 2002.

SI 534 Nonparametric Statistics 2 1 0 6

Prerequisite: SI 424 (Exposure) (For students from other departments, instructor's permission will be required)

Kolmogorov-Smirnov Goodness-of – Fit Test. The empirical distribution and its basic properties. Order Statistics. Inferences concerning Location parameter based on one-sample and two-sample problems. Inferences concerning Scale parameters. General Distribution Tests based on Two or More Independent Samples.

Tests for Randomness and equality of distributions. Tests for Independence. The one-sample regression problem.

Asymptotic Relative Efficiency of Tests. Confidence Intervals and bounds.

Texts/References:

Arnold B.C., Balakrishnan N. and Nagaraja H. N., First Course in Order Statistics, John Wiley & Sons, New York, 1992.

Daniel W.W., Applied Nonparametric Statistics, 2nd Edition, PWS- KENT, 1990.

Ghosh J.K. and Ramamoorthi R.V. , Bayesian Nonparametrics, Springer, New York, 2003.

Gibbons J.D., Nonparametric Statistical Inference, Marcel Dekker, New York, 1985.

Hollandor M., and Wolfe D.A., Non-parametric Statistical Methods, John Wiley & Sons, New York, 1973.

Lehmann E.L, Nonparametric Statistical Methods Based on Ranks, McGraw-Hill, New York, 1975.

Randles R.H. and Wolfe D.A., Introduction to the Theory of Nonparametric Statistics, John Wiley & Sons, New York, 1979.

Sprent P., Applied Nonparametric Statistical Methods, Chapman and Hall, London, 1989.

SI 514 Statistical Modelling 2 1 0 6

Prerequisite: SI 424 (Exposure) (For students from other departments, instructor's permission will be required)

Mixed Effects linear regression models – Estimation and Testing, Kernel based regression, Spline methods, Nonlinear regression, Time Series ARIMAX models.

Texts / References:

Box G.E.P., Jenkins G., and Reinsel G., Time Series Analysis-Forecasting and Control, 3rd Edition, Pearson Education, 1993.

Hardle W., Applied Nonparametric Regression, Cambridge University Press, Cambridge, 1990

Hastie T., and Tibshirani R. Generalized Additive Models, Chapman and Hall, London,1990.

Seber G.A.F., and Wild C.J., Nonlinear Regression, John Wiley & Sons., New York, 1989.

Wasserman L., All of Nonparametric Statistics, Springer Text in Statistics, New York, 2006.

SI 548 Computational Statistics 2 1 0 6

Introduction to Bayesian Theory and methods; non-informative priors and conjugate priors; posterior inference (with special reference to one parameter exponential family)-credible intervals and hypothesis testing; hierarchical and empirical Bayesian models; computational techniques for use in Bayesian analysis, especially the use of simulation from posterior distributions, with emphasis on the WinBUGS package as a practical tool.

MCMC simulation (Markov chains; Metropolis-Hastings algorithm; Gibbs sampling; convergence), EM algorithm, Bootstrap (Bootstrapping; jackknife resampling; percentile confidence intervals).

Permutation tests.

Texts/References:

Efron B. and Tibshirani R.J., An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.

Gentle J.E., Elements of Computational Statistics (ECS), Springer-Verlag, New York, 2002.

Gentle J.E., Computational Statistics, Statistics and Computing Series, Springer-Verlag, New York, 2009.

Gelman A., Carlin J. B., Stern H. S., and, Dunson D. B., Vehtari A., and Rubin D.B., Bayesian Data Analysis, 3rd Edition, CRC Press, Taylor & Francis Group, Boca Raton, 2014.

Givens G. H. and Hoeting J. A., Computational Statistics, 2nd Edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2013.

Lange K., Numerical Analysis for Statisticians, 2nd Edition, Springer-Verlag, New York, 2002.

Little R.J.A. and Rubin D.B., Statistical Analysis with Missing Data, 2nd Edition, Wiley, New York, 2019.

Liu J., Monte Carlo Strategies in Scientific Computing, Springer-Verlag, New York, 2001.

Rice J.A., Mathematical Statistics and Data Analysis, 2nd Edition, Duxbury Press, Belmont, California, 1995.

Schafer J.L., Analysis of Incomplete Multivariate Data, Chapman & Hall, New York, 1997.

SI 536 Analysis of Multi-Type and Big Data 2 1 0 6

Prerequisites: SI 505 (Exposure), SI 515 (Exposure) (For students from other departments, instructor's permission will be required)

Overview of Spatial Data, Structured Data. Structural Equation Modelling.

Introduction to Big Data. Large dimension small size multivariate data analysis, Mean Testing in High-Dimension.

Classification of Big Data; Screening and Variable Selection.

Sparse Linear Regression Models**,** Lasso Regression.

Principal Component version for High-dimensional data**.**

MCMC techniques for Bayesian Modelling in High-dimension.

Texts/References:

Bollen K.A., Structural Equations with Latent Variables, John Wiley & Sons, New York, 1989.

Bollen K.A., Latent Curve Models: A Structural Equation Perspective, John Wiley & Sons, Hoboken, 2006.

B¨uhlmann, P. and van de Geer, S., Statistics for High-Dimensional Data: Methods, Theory and Applications. Springer, Berlin, 2011.

Gamerman, D., Hedibert, F. L., Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2nd Edition., Chapman and Hall/CRC, Boca Raton, 2006.

Hastie, T., Tibshirani, R. and Friedman, J., The Elements of Statistical Learning, Springer, Berlin, 2009.

SI 527 Introduction to Derivative Pricing 2 1 0 6

Introduction to options and markets: types of options, interest rates and present value. Black-Scholes Model: arbitrage, option values, payoffs and strategies, put call parity, Black-Scholes equation, similarity solution and exact formulae for European options. American options: call and put options, free boundary problem. Binomial methods: option valuation, dividend paying stock, general formulation and implementation. Monte-Carlo simulation: valuation by simulation. Finite Difference Methods: explicit and implicit methods with stability and convergence analysis, methods for American option-constrained matrix problem, projected SOR, time stepping algorithms with convergence and numerical examples. Lab Component: Implementation of the option pricing algorithms and Evaluation for Indian companies.

Texts/References:

Clewlow L. and Strickland C. (1998) Implementing Derivative Models, Wiley Series on Financial Engineering, John Wiley and Sons.

Hull J. C. (2009) Options, Futures and Other Derivatives, 7th Edition, Prentice Hall of India.

Wilmott P., Howison S. and Dewynne J. (1995) The Mathematics of Financial Derivatives: A Student Introduction, Cambridge University Press.

Wilmott P., Dewynne J., and Howison S. (1993) Option Pricing: Mathematical Models and Computation, Oxford Financial Press.

SI 544 Martingale theory 2 1 0 6

Prerequisites: SI 427 (Exposure), SI 537 (Exposure) (For students from other departments, instructor's permission will be required)

Review of conditional expectation: Conditional expectation and conditional probability, properties of conditional expectation, regular conditional distributions, disintegration, conditional independence.

Martingales and Stopping times: Stopping times, random time change, martingale property, optional sampling theorem, maximum and up-crossing inequalities, martingale convergence theorem, Martingale central limit theorem.

Texts/References:

Athreya, K.B. and Lahiri, S.N., Probability Theory, Hindustan Book Agency, 2006.

Billingsley, P., Probability and Measure, Anniversary Edition, John Wiley & Sons, Hoboken, 2012.

Chung, K. L., A Course in Probability Theory, Third edition, Academic Press, San Diego, 2001.

Williams, D., Probability with martingales, Cambridge University Press, Cambridge, 1991.

SI 546 Statistical Inference II 2 1 0 6

SI 424 (Exposure))(For students from other departments, instructor's permission will be required.)

Minimaxity and admissibility: Minimax estimation, admissibility and minimaxity in exponential families, admissibility and minimaxity in group families, Simultaneous estimation. Maxmin tests and invariance, Hunt-stein Theorem, Most stringent tests. Multiple testing via Maximin procedures and Scheff\'{e}'s S-method.

U-statistics: Variance computation and projection method. Convergence of U statistics (one sample and two samples). Linear rank statistics. Asymptotic normality under null hypothesis. Pitman's Asymptotic relative efficiency, Noether's theorem for evaluating asymptotic relative efficiency. Bahadur's efficiency. Resampling techniques.

Robust inference: Break-down point in finite sample, Influence curve. M-estimator, L-estimator, R-estimator, minimum distance estimator and Pitman's estimator. Relations to minimax estimator and equivariant estimators. Robust tests and confidence sets.

Texts/References:

Casella G. and Berger R.L., Statistical Inference, Wadsworth, a part of Cengage Learning, Delhi, 2002.

Dasgupta A., Asymptotic Theory of Statistics and Probability, Springer, New York, 2008.

Jurečková J., Sen P.K. and Picek J., Methodology in Robust and Nonparametric Statistics, CRC press, Boca Raton, 2012.

Lehmann E.L., Theory of Point Estimation, Springer, New York, 1998.

Lehmann E.L. and Romano J.P. , Testing of Statistical Hypotheses, Springer, New York, 2011.

Huber P. J. and Ronchetti E.M., Robust Statistics, Wiley, New York, 2009.

Shao J., Mathematical Statistics, Springer, New York, 2003.