Syllabi of M.Sc. (Statistics-OR)

January 9, 2017

Contents

Note: All Courses are of FOUR credits

I Semester (Statistics-O.R.)

1 Basic Real Analysis (Code: ST401)

Real number system and its structure, infimum, supremum, Dedekind cuts. (Proofs omitted)

Sequences and series of real numbers, subsequences, monotone sequences, limit inferior, limit superior, convergence of sequences and series, Cauchy criterion, root and ratio tests for the convergence of series, power series, product of series, absolute and conditional convergence, metric spaces, limits in metric spaces.

Functions of a single real variable, limits of functions, continuity of functions, uniform continuity, continuity and compactness, continuity and connectedness, types of discontinuities, monotonic functions, infinite limit and limit at infinity.

Differentiation, properties of derivatives, chain rule, Rolle's theorem, mean value theorems, l'Hospital's rule, derivatives of higher order, Taylor's theorem.

Sequences and series of functions, point-wise and uniform convergence, continuity of the uniform limit of continuous functions, uniform convergence and differentiability, Dini's theorem, equicontinuity, point-wise and uniform boundedness, Arzela-Ascoli's theorem, Weierstrass approximation theorem, space filling curve, continuous but nowhere differentiable function.

Fourier series: Dirichlet kernel, point-wise convergence of Fourier series.

References

[1] Goldberg, Richard R., *Methods of Real Analysis*, second edition, John Wiley & Sons, Inc., New York-London-Sydney, 1976.

[2] Rudin, Walter, *Principles of Mathematical Analysis*, third edition, International Series in Pure and Applied Mathematics. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.

[3] Bartle, Robert G., *The Elements of Real Analysis*, second edition, John Wiley & Sons, New York-London-Sydney, 1976. (for Fourier Series)

[4] Ross, Kenneth A., *Elementary Analysis. The Theory of Calculus*, second edition, in collaboration with Jorge M. López, Undergraduate Texts in Mathematics, Springer, New York, 2013.

2 Linear Algebra and Matrix Theory (Code: ST402)

Matrices: Elementary operations, reduced row-echelon form, consistency of system of equations, solutions of systems of equations, homogeneous system, inverse of a matrix, determinants, Cramer's rule.

Vector spaces and subspaces, linear independence of vectors, basis, linear transformations and matrices, kernel, nullity theorem, rank of a matrix.

Inner Product spaces, C-S inequality, triangle inequality, orthonormal basis, Gram-Schmidt construction of orthonormal basis.

Basics of theory of matrices and determinants, row and column spans of real matrices, rank, elementary matrices, idempotent matrices, eigenvalues and eigenvectors of a matrix, spectral decomposition of a real symmetric matrix, definiteness-positive, non-negative of real matrices, g-inverses: existence and definition, some important results on g-inverses, bilinear and quadratic forms, extrema of quadratic forms.

References

[1] Hoffman, K. and Kunze, R., *Linear Algebra*, second edition, Prentice-Hall, New Delhi, 1978.

[2] Rao, A. R. and Bhimashankaram, P., *Linear Algebra*, second edition, TRIM-Hindustan Book Agency, 2000.

[3] Rao, C. R., *Linear Statistical Inference and Its Applications*, second edition, Wiley Eastern, 1973.

3 Discrete Mathematical Structures (Code: ST403)

Sets and propositions: Introduction, combinations of sets, finite and infinite sets, uncountably infinite sets, mathematical induction, principle of inclusion and exclusion, pigeonhole principle, multisets.

Statement calculus, truth tables, validity, consequence.

Predicate calculus: Propositional logic, propositional equivalence, predicates and quantifiers, rules of inference, proof of methods.

Permutations, combinations: Introduction, the rules of sum and product, permutations, combinations, generation of permutations and combinations.

Relations and functions: Introduction, properties of binary relations, equivalence relation and partitions, partial ordered relations.

Groups and rings: Introduction, semi-groups, groups, subgroups, generators and evaluation of powers, cosets and Lagrange's theorem, permutation groups, normal sub-groups, quotient groups, group homomorphisms, automorphisms, isomorphisms, fundamental theorems of group homomorphisms, Cayley's theorem, group actions, Burnside's theorem.

Sylow's first, second and third theorems and their applications, structure theorem for finite abelian groups, composition series Jordan-Hölder theorem, nilpotent and solvable groups.

References

[1] Liu, C., *Elements of Discrete Mathematics*, McGraw-Hill Education (India) Pvt Limited, 2008.

[2] Tremblay, J. P. and Manohar, R., *Discrete Mathematics Structures with Applications to Computer Science*, McGraw Hill, 1997.

[3] Rosen, K. H., *Discrete Mathematics and Its Applications*, sixth edition, Tata McGraw Hill Education, New Delhi, 2008.

[4] Grimaldi, Ralph P., *Discrete and Combinatorial Mathematics*, fifth edition, Pearson Education, New Delhi, 2005.

[5] Kolman, B.; Busby, R. C. and Ross, S., *Discrete Mathematics*, fifth edition, Prentice Hall, 1996.

[6] Gossett, Eric, *Discrete Mathematics with Proof*, second edition, Wiley India Pvt. Ltd, 2009.

[7] Koshy, Thomas, *Discrete Mathematics and Applications*, second edition, Elsevier Publication, New Delhi, 2003.

[8] Mott, J. L.; Candell, A. and Bekar, I., *Discrete Mathematics for Computer Scientists and Mathematicians*, PHI, 1986.

4 Elements of Probability and Statistics (Code: ST405)

Random experiments, sample spaces, sets, events, algebras; elements of combinatorial analysis; classical definition and calculation of probability, independence of events.

Random variables, distribution functions, moments, probability and moment generating functions, independence of random variables, inequalities.

Introduction to various discrete and continuous random variables, limiting distributions of some random variables, distributions of functions of random variables.

Bi-Variate distributions, conditional and marginal distributions, conditional expectation and variance, co-variance and correlation co-efficient, bivariate moment generating functions.

Elementary understanding of data: frequency curves, empirical measures of location, spread, empirical moments, analysis of bivariate data; fitting of distributions.

Sampling distributions, Chi-square, t, F.

References

[1] Feller, W., *Introduction to Probability Theory and its Applications*, third edition, Wiley Eastern, 1978.

[2] Ross, S., *A First Course in Probability*, sixth edition, Pearson Education, 2007.

[3] Prakasa Rao, B. L. S., *A First Course in Probability and Statistics*, World Scientific, 2009.

5 Statistical Methods (Code: ST407)

Topics in Statistical computing, using EXCEL, R and other available packages.

Probability distributions: Generating random numbers from standard, non-standard univariate and mixture distributions (through use of uniform random variables); plotting of distributions, failures rates.

Descriptive statistics: Learning data input, descriptive statistics generation and interpretation.

Parametric inference: t (one, two, paired), Chi-square and F-tests.

Model fitting: Simple and multiple regression model, ANOVA.

Dealing with discrete data: Descriptive statistics and association analysis.

Non-parametric inference: Wilcoxon test, Wilcoxon signed-rank test, Kruskal Wallis test., rank correlation

Generating from simple stochastic processes.

Examples of central limit theorem through sampling.

Introductory topics in measure theory: Monotone classes and sigma algebras on sample spaces; outer measure and mu star measurability, Carathéodory's extension and uniqueness theorem, probability measure on the sigma algebra; measurable functions and random variables, distribution function of a random variable - its properties.

References

[1] Tanner, M. A., *Tools for Statistical Inference*, Springer-Verlag, 2011.

[2] Snedcor, G. W. and Cochran, W. G., *Statistical Methods*, seventh edition, Iowa State University Press, 1982.

[3] Conover, W. J., *Practical Non-Parametric Statistics*, third edition, John Wiley, NY, 2007.

[4] Ripley, B.D., *Stochastic Simulations*, Wiley-Interscience, 2006.

[5] Kennedy, W. J. and Gentle, J. E., *Statistical Computing*, Taylor & Francis, 1980.

II Semester (Statistics-O.R.)

6 First Course on Operations Research (Code: ST456)

Linear programming models and their applications.

Pivot operations on systems of linear equations; LP solutions, basic feasible solutions, feasible and optimal bases, and bases; LP in standard form, geometry of polyhedra.

Solving LPs: the graphical and simplex methods, geometry of simplex method, the revised simplex method.

Duality in LP; fundamental theorems of duality and optimality conditions for LP; the dual simplex algorithm.

Sensitivity analyses in LP - basic concepts.

The transportation and assignment models and algorithm for solving them.

Review of convex sets.

Convex functions and their generalizations, methods for checking convexity of functions.

Optimization of convex functions.

References

[1] Murty, K. G., *Operations Research - Deterministic Optimization Models*, Prentice-Hall, 1995.

[2] Murty, K. G., *Linear Programming*, John Wiley, NY, 1983.

[3] Bazaraa, M.S.; Jarvis, J. J. and Shirali, H. D., *Linear Programming and Network Flows*, second edition, John Wiley, Singapore, 2003.

[4] Saigal, R., *Linear Programming - A Modern Integrated Analysis*, Springer, 2012.

[5] Bertsimas, D. and Tsitsiklis, J. N., *Introduction to Linear Optimization*, Athena Scientific, 1997.

[6] Fourer, R.; Gay, D. M. and Kernighan, B. W., *AMPL- A Modeling Language for Mathematical Programming*, Brooks/Cole, 1999.

7 Probability Theory (Code: ST452)

Probability measure: The tail sigma algebras, Borel-Cantelli lemmas; Convergence of sequences of real valued random variables-in probability and almost surely; applications of Borel-Cantelli lemmas - the weak and the strong laws of large numbers; some applications-Monte Carlo methods for Riemann-integrals of real valued functions.

Integration: Integration of a non-negative measurable function with respect to a measure; monotone convergence theorem; integrability and integral of any measurable function, monotone convergence theorem, Lebesgue dominated convergence theorem, Fatou's lemma; expected values of random variables, existence.

The product Boolean algebra and sigma algebra on the cartesian product of two sample spaces, transition measure and the transition probability measure, construction of more measures on the product spaces, Fubini's theorem.

The convolution of two or more measures on the real line, convolution of distribution functions and density functions.

The characteristic function of random variables and its properties; the inversion and uniqueness theorem; Helly-Bray theorems; Central limit theorems: Laplace deMoivre, Lindeberg, Lyapunov.

References

[1] Parthasarathy, K. R., *Introduction to Probability and Measure*, TRIM - Hindustan Book Agency, 2005.

[2] Athreya, K. B. and Lahiri, S. N., *Probability Theory*, TRIM - Hindustan Book Agency, 2006.

[3] Billingsley, P., *Probability and Measure*, John Wiley & Sons, NY, 1979.

8 Theory of Inference-I (Code: ST453)

PROBLEM OF POINT ESTIMATION:

Properties of point estimation: Unbiasedness, closeness, mean squared error, consistency and BAN. Methods of finding estimators: Method of moments, maximum likelihood, other methods; sufficient, complete and ancillary statistics.

unbiased estimation: Lower bound for the variance of an estimator; location and scale invariant estimators; optimum properties of maximum likelihood estimation.

PARAMETRIC INTERVAL ESTIMATION:

Definition of confidence interval, pivotal quantity.

Sampling from the normal distribution : Confidence interval for the mean, confidence interval for the variance, simultaneous confidence region for the mean and variance, confidence interval for difference in means.

Methods of finding confidence intervals: Pivotal quantity method, statistical method, large sample confidence intervals.

References

[1] Rohatgi, V. K., *An Introduction to Probability and Statistics*, Wiley Eastern, 1985.

[2] Casella, G. and Berger, R. L., *Statistical Inference*, second edition, Academic Internet Pub., 2002.

[3] Mood, A. M.; Graybill, F. A. and Boes, D. C., *Introduction to the Theory of Statistics*, third edition, McGraw-Hill, 1974.

[4] Lehmann, E. L., *Theory of Point Estimation*, John Wiley, NY, 1983.

9 Linear Models (Code: ST454)

The linear Model, Gauss- Markov Model - Estimability of linear parametric functions and their BLUEs under both uncorrelated and correlated observations; the properties of BLUEs and the residual sum of squares under normality.

Hypothesis tests for one or more linear parametric functions, estimation under linear constraints on the parameters.

The ANOVA models-one-way, two-way with one observation per cell and with 'c' observations per cell, determining tests for the testable linear hypotheses.

Linear regression models-estimation of model parameters, multiple correlation coefficient, tests for model parameters-partial correlation coefficient.

Estimation of variance components, BLIMB estimation.

References

[1] Rao, C. R., *Linear Statistical Inference and Its Applications*, second edition, Wiley Eastern, 1973.

[2] Rao, A. R. and Bhimasankaram, P., *Linear Algebra*, second edition, TRIM-Hindustan Book Agency, 2000.

[3] Sengupta, D. and Jammalamadaka Rao, S., *Linear Models-An Integrated Approach*, World Scientific, 2003.

10 Theory of Sampling (Code: ST455)

The role of sampling theory, simple random sampling for mean, proportions and percentages, estimation of sample size.

Stratified sampling for mean and proportions, optimum allocation, relative precision of stratified random and simple random sampling, effect of deviations from the optimum allocation, effect of errors in the stratum sizes.

Systematic sampling for linear and circular cases, variance of estimated mean, comparison of systematic sampling with simple random sampling and stratified sampling for population with linear trend.

Procedure of selecting a sample with varying probabilities, estimation of the population mean and variance of the estimated mean, sampling with varying probabilities and without replacement, ordered estimates, the Horvitz-Thompson Estimator, Some IPPS sampling without replacement procedures, Rao Hartley and Cochran's procedure.

Ratio Estimator, bias and mean squared error of the ratio estimator and its approximation, ratio estimates in stratified sampling, comparison of the ratio estimate and the mean per unit.

Regression estimator, bias and mean squared error of the regression estimator, efficiency of the regression estimator, regression estimates in stratified sampling, comparison of the regression estimate with the ratio estimate and the mean per unit.

References

[1] Cochran, W. C., *Sampling Techniques*, second edition, third edition, Wiley Eastern, 1977.

[2] Sukhatme, P. V.; Sukhatme, B. V.; Sukhatme, S. and Asok, C., *Sampling Theory of Surveys with Applications*, Indian Society of Agricultural Statistics, 1954.

[3] Raj, Des, *Sampling Theory*, Tata McGraw-Hill, 1968.

III Semester (Statistics-O.R.)

11 Multivariate Analysis (Code: ST501)

Bivariate - Binomial, Poisson and exponential distributions, multinomial and Dirichlet distributions; conditional, marginal distributions.

Multivariate normal distribution, (singular and nonsingular) marginal and conditional distributions; linear transformations, characteristics function; maximum likelihood estimators of the parameters and their sampling distributions.

Tests of hypothesis about the mean vector of normal distribution, Hotelling's T^2- statistics, its distribution and applications; relationship between T^2 and Mahalanobis D^2-statistics; Wishart distribution and its properties; confidence regions.

Introduction to principle components, canonical correlation and canonical variables; cluster analysis; classification problems; discriminant analysis; elliptical distributions.

References

[1] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, John Wiley, 2003.

[2] Kshirsagar, A. M., *Multivariate Analysis*, Marcel Dekker, 1972.

[3] Johnson, R. A. and Wichern, D. W., *Applied Multivariate Statistical Analysis*, fifth edition, Pearson Education, Singapore, 2002.

12 Theory of Inference-II (Code: ST502)

Hypothesis Testing: Some fundamental notions of hypothesis testing, UMP test, Neyman-Pearson Lemma, monotone likelihood ratio, unbiased test, generalized likelihood ratio test, asymptotic distribution of generalized likelihood ratio, chi square goodness-of-fit test, test of equality of two multinomial distributions and generalization; test of independence in contingency tables, test of hypothesis and confidence intervals.

Sequential Test of Hypotheses: Definition of sequential probability ratio test, approximate sequential probability ratio test, approximate expected expected sample size, operational characteristic (OC) function or curve of sequential probability ratio test.

Nonparametric Methods: Some fundamental notion of nonparametric/distribution free methods.

Some applications of order statistics: Point and interval estimates of a quantile, test of hypotheses concerning quantiles, tolerance limits and coverage; sample or empirical cumulative distribution function, Kolmo-Smirnov goodness-of-fit test, confidence bands for cumulative distribution function, Wilcoxon Signed-rank test.

Some two sample test: Kolmogorov-Smirnov test, run test, median test, Mann-Whitney-Wilcoxon test, test of independence, Kendall's Tau and Spearman rank correlation coefficient.

Decision Theory: Posterior distribution, loss-function approach, Minimax estimator, Bayes estimator.

References

[1] Rohatgi, V. K., *An Introduction to Probability and Statistics*, Wiley Eastern, 1985.

[2] Casella, G. and Berger, R. L., *Statistical Inference*, second edition, Academic Internet Pub., 2002.

[3] Mood, A. M. , Graybill, F.A., and Boes, D.C., *Introduction to the Theory of Statistics*, third edition, McGraw-Hill, 1974.

[4] Lehmann, E. L., *Theory of Point Estimation*, John Wiley, NY, 1983.

[5] Gibbons, J. D. and Chakaraborti, *Nonparametric Statistical Inference*, fourth edition, Marcel Dekker, 2003.

13 Design and Analysis of Experiments (Code: ST503)

Principles of Experimental Design: Need for designed experiments, how data are obtained, difference between data obtained from designed experiments and sampling, randomization, replication and blocking-the need for them and how to achieve these principles in an experiment.

Analysis of different designs, their strengths and when to employ which design, completely randomized design, fixed effects and random effects, randomized block design-without and with interaction; Latin Square design, repeated LSs; incomplete block designs, balanced incomplete block design(BIBD); analysis of covariance, factorial designs-2^k factorial experiments, confounding, 2^{k-1} fractional factorial experiments, response surface method to determine optimum factor level combination with data obtained from a 2^k factorial design; split plot design; partially balanced incomplete block design (PBIBD).

References

[1] Montgomery D. C., *Design and Analysis of Experiments*, fifth edition, John Wiley, 2008.

[2] Das, M.N. and Giri, N. C., *Design and Analysis of Experiments*, second edition, Wiley Eastern, 1991.

[3] Dey, A., *Incomplete Block Designs*, Hindustan Book Agency, 2010.

14 Regression Theory and Analysis (Code: ST504)

Simple regression with one independent variable(X), assumptions, estimation of parameters, standard error of estimator, testing of hypothesis about regression parameters, standard error of prediction, testing of hypotheses about parallelism, equality of intercepts, congruence, extrapolation, optimal choice of X.

Model adequacy checking, diagnostic checks and correction: graphical techniques, tests for normality, uncorrelatedness, homoscedasticity, lack of fit, modifications like polynomial regression, transformations on Y or X, WLS.

Multiple regression: Standard Gauss Markov setup, least square(LS) estimation, error and estimation spaces, variance- covariance of LS estimators, estimation of error variance, case with correlated observations, simultaneous estimation of linear parametric functions, test of Hypotheses for one and more than one linear parametric functions, confidence intervals and regions.

Use of indicator variables and ANOVA.

Variable selection and model building, multicollinearity, non-linear regression, validation of regression models.

References

[1] Montgomery, D. C.; Peck, E. A. and Vining, G. G., *Introduction to Linear Regression Analysis*, third edition, John Wiley, 2003.

[2] Draper, N. R. and Smith, H., *Applied Regression Analysis*, third edition, John Wiley & Sons, NY, 2005.

[3] McCullagh, P. and Nelder, J. A., *Generalized Linear Models*, second edition, CRC Books, London, 2000.

[4] Seber, G. E. F. and Wild, C. J., *Nonlinear Regression*, John Wiley, NY, 1989.

[5] Neter, J.; Kutner, M.; Wasserman, W. and Nachtsheim, C., *Applied Linear Statistical Models.*, fourth edition, McGraw-Hill/Irwin, 1996.

15 Stochastic Processes (Code: ST505)

Introduction, Poisson process - Introduction, homogeneous and the non-homogeneous Poisson processes, renewal process, key renewal theorem and some applications. Markov Chains : classification of states, existence of stationary distribution, expected time between successive visits to a state - positive recurrent states, Random walk, Markov Processes, Continuous time discrete state Markov processes, embedded Markov chains and semi-Markov processes, various birth and death processes-Yule Process, the various Markovian queues and their steady state distribution when they exist.

Martingale processes : martingale theorem and some applications, Doob's inequality.

Brownian motion process as a limit of the random walk, its properties, Brownian bridge.

References

[1] Ross, S., *Stochastic processes*, second edition, John Wiley, 1996.

[2] Goswami, A. and Rao, B. V., *A Course in Applied Stochastic Processes*, TRIM-Hindustan Book Agency, 2006.

IV Semester (Statistics-O.R.)

16 Theory of Inference-III (Code: ST574)

Review of modes of convergence, laws of large numbers and central limit theorem, Maximum Likelihood Estimates (MLEs); MLEs in multi parameter exponential family and their asymptotic behavior, strong consistency and asymptotic normality of MLE, asymptotic efficiency, super efficient estimator.

Asymptotic normality of posterior distribution, asymptotic distribution of LRT-statistic and chi square statistic, Wald test, score test, U-statistic asymptotic comparison of tests, asymptotic relative efficiency (Pitman's).

Nonparametric estimation of density and distribution functions, various methods of density estimation, finite and asymptotic properties of kernel method estimators.

Statistical inference for stochastic processes, likelihood function, estimation in finite Markov chain , branching process, continuous time Markov point processes.

References

[1] Ferguson, T. S., *A course in Large Sample theory*, Taylor & Francis, 1996.

[2] Lehmann, E. L., *Theory of Point Estimation*, John Wiley, NY, 1983.

[3] Rao, C. R., *Linear Statistical Inference and Its Applications*, second edition, Wiley Eastern, 1973.

[4] Bhat, B. R., *Stochastic Models*, (New Age International Pvt. Limited), 2004.

[5] Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, 1986.

[6] Basawa, I. V. and Prakasa Rao, B. L. S., *Statistical Inference for Stochastic Processes*, Academic Press, 1980.

17 Stochastic modeling in finance (Code: ST583)

Information set (sigma algebra), filtration, stochastic process, Wiener process (Brownian motion).

Conditional expectation , conditional density.

Stochastic integral, stochastic differential equation.

Ito formula, derivation of Black-Scholes partial differential equation.

Martingales, examples of martingales, sub martingales, super martingales.

Girasanov theorem, risk-neutral probabilities, application in derivates pricing.

Derivatives pricing on incompletes markets, stochastic volatility.

Change of numeraire, forward probability measure, alternative derivation of Black-Scholes formula.

Poisson process, jump-diffusion process, generalization of Ito formula, application in derivates pricing.

Numerical solutions of stochastic differential equation. (Euler and Milstein approximation).

Stopping time, local martingales, semi martingales, Levy processes, applications.

References

[1] Durrett, R., *Stochastic Calculus - a Practical Introduction*, Taylor & Francis, 1996.

[2] Hunt, P. and Kennedy, J. E., *Financial Derivatives in Theory and Practice*, John Wiley & Sons, 2004.

18 Time Series (Code: ST571)

Introduction: Examples, simple descriptive techniques, trend, seasonality, the correlogram.

Probability models for time series: stationarity, moving average (MA), autoregressive (AR), ARMA and ARIMA models.

Estimating the autocorrelation function and fitting ARIMA models.

Forecasting: Exponential smoothing, forecasting from ARIMA models.

Stationary processes in the frequency domain: The spectral density function, the periodogram, spectral analysis.

State-space models: Dynamic linear models and the Kalman filter.

References

[1] Box, G. E. P. and Jenkins, G. M., *Time Series Analysis - Forecasting and control*, third edition, Englewood Cliffs Prentice Hall, 1994.

[2] Fuller, W. A., *Introduction to Time Series Analysis*, John Wiley, NY, 1976.

[3] Kendall, Stuart, A. and Ord, J. K., *The Advanced Theory of Statistics*, Vol.3, Charles-Griffin, London, 1977.

19 Econometrics (Code: ST572)

Econometric modelling, data and methodology, simple linear regression model.

Multiple linear regression model: Estimates, statistical inference, asymptotic properties, units of measurement, functional forms, predictions and analysis of residuals, analysis with qualitative information dummy variables.

Problems with specification and data, multicollinearity.

Generalized linear model: Heteroskedasticity, logistic growth model, estimates of demand.

Basic time series regression analysis, dynamic models, non-stationary, time series and spurious regression, autocorrelation and heteroskedasticity in time series analysis.

Static and dynamic neoclassical production function.

Duration models.

References

[1] Dhrymes, P. J., *Introductory Econometrics*, Springer-Verlag, NY, 1985.

[2] Greene, W. H., *Econometric Analysis*, fifth edition, Pearson Education, NY, 2006.

[3] Johnston, J., *Econometric Methods*, McGraw-Hill, 1991.

20 Sequential Analysis (Code: ST577)

Sequential procedures, gain from sequential sampling, illustrations, sampling inspections, Stein two-stage procedure, sequential tests between three hypotheses,

SPRT; properties and extensions, unbiased estimation in sequential binomial sampling plans, lower bond for variance.

Risk of a sequential procedure, backward induction, optimal bounded sequential procedure.

References

[1] Wetherill, C. ., *Sequential Methods in Statistics*, Chapman and Hall, 1979.

[2] De Groot, M. H., *Optimal Statistical Decisions*, Wiley, 2005.

[3] Rohatgi, V. K., *An Introduction to Probability and Statistics*, Wiley Eastern, 1985.

[4] Lehmann, E. L., *Theory of Point Estimation*, John Wiley, NY, 1983.

21 Reliability and Survival Analysis (Code: ST578)

Structure function, coherent structure Reliability and their bonds, the notion of aging; failure rate, classes of life time distributions and their relationships,

Availability, maintenances through repairs and spares, multi state systems, multivariate monotone failure distributions.

Estimation and testing under censoring for parametric models, testing for class properties, nonparametric models, Kaplan-Meier estimator and its properties.

Cox proportional hazard model, conditional and partial likelihoods, asymptotic normality of estimators.

References

[1] Barlow, R. E. and Proschan, F., *Statistical Theory of Reliability and Life Testing*, Holt, Renehart and Winston Inc., 1975.

[2] Miller, R. G., *Survival Analysis*, John Wiley, NY, 1981.

[3] Deshpande, J. V. and Purohit, S. G., *Life Time Data - Statistical Models and Methods*, World Scientific, 2005.

22 Order Statistics (Code: ST579)

Introduction: The subject of Order Statistics.

Basic distribution Theory: Distribution of s single order statistic, joint distribution of two or more order statistics, distribution of the range and other systematic statistics, discrete order statistics, some properties of order statistics, independence results - order statistics as a Markov chain.

Expected values and moments: Basic formulae, order statistics from some specific distributions - Bernoulli distribution, binomial distribution, three point distribution, Poisson distribution, geometric distribution, exponential distribution, uniform distribution, normal distribution.

Some identities and recurrence relations.

Bounds and approximations for moments of order statistics.

Some characterizations using order statistics.

Order statistics in statistical inference: Types of order statistics data, order statistics and sufficiency, maximum likelihood estimation, linear estimation of location and scale parameters, life testing, with special emphasis on the exponential distribution.

Record Values: Definitions and preliminary results on record statistics.

References

[1] David, H. A., *Order Statistics*, second edition, John Wiley, NY, 1981.

[2] Arnold, B. C.; Balakrishnan, N. and Nagaraja, H. N., *A First Course in Order Statistics*, John Wiley, NY, 1992.

23 Actuarial Modelling (Code: ST580)

Actuarial models: Motivation via modeling examples.

Random variables: Random variables encountered in actuarial work, key functions and four models.

Basic distributional quantities: Moments, quantiles, generating functions and sums of random variables.

Characteristics of actuarial models: The role of parameters.

Continuous models: Creating new distributions, selected distributions and their relationships, the linear exponential family.

Discrete distributions and processes: The Poisson distribution, the negative binomial distribution, the binomial distribution, the $(a, b, 0)$ class.

Frequency and severity with coverage modifications: Deductibles, the loss elimination ratio and the effect of inflation for ordinary deductibles, policy limits.

Aggregate loss models: Model choices, the compound model for aggregate claims, analytic results, computing the aggregate claims distribution, the recursive method, the impact of individual policy modifications on aggregate payments, inversion methods.

Estimation for complete data: The empirical distribution for complete, individual and grouped data.

Estimation for modified data: Kernel density models, approximation for large data sets (Kaplan-Meier type approximations).

Parameter estimation: Methods of moments and percentile matching, maximum likelihood estimation, interval estimation, Bayesian estimation.

Model Selection: Graphical comparison of the density and distribution functions, hypothesis tests, selecting a model.

Credibility: Limited fluctuation credibility theory, greatest accuracy credibility theory, empirical Bayes parameter estimation.

Simulation: Basics of simulation, examples of simulation in actuarial modeling.

References

[1] Klugman, S. A.; Panjer, H. H. and Willmot, G. E., *Loss Models - From Data to Decisions*, third ediion , John Wiley, Canada, 2008.

24 Industrial Statistics (Code: ST575)

Statistical process control: Chance and Assignable causes of quality variation, Statistical basis of control chart.

Control charts for variables: Control charts for $\overline{X}$ and R - Statistical basis of the charts, development and use of $\overline{X}$ and R charts, chart based on standard values, the operating-characteristic function and the average run length for the $\overline{X}$ chart, control charts for $\overline{X}$ and S with constant and variable sample size, the S^2 control chart, control charts for individual measurements.

Control charts for attributes: the control chart for fraction nonconforming with constant and variable sample size,the operating -characteristic function and average run length calculations, control charts for non-conformities (defects) with constant and variable sample size and OC-function.

Cumulative sum and exponentially weighted moving average control charts. Other statistical process control techniques: statistical process control for short production runs, modified and acceptance control charts.

Process capability analysis.

Acceptance sampling: Single sampling plans for attributes - methods of choosing sampling plans.

Double, multiple and sequential sampling plans.

Acceptance sampling by variables.

References

[1] Montgomery, D. C., *Introduction to Statistical Quality Control*, fourth edition, John Wiley, 2003.

[2] Schilling, E. G., *Acceptance Sampling in Quality Control*, Marcel Dekker, 1982.

[3] Wetherill, G. B., *Sampling Inspection and Quality Control*, John Wiley, 1977.

25 Non-Linear Programming (Code: ST551)

Introduction to applications of nonlinear programming: optimal control problems, structural design, mechanical design, electrical networks, water resources management, stochastic resource allocation, location of facilities, financial engineering problems.

Review of convex functions and convex optimization.

Nonlinear programming problems, unconstrained problems, problems with inequality and equality constraints, second-order necessary and sufficient optimality conditions for constrained problems (Fritz John and Karush-Kuhn-Tucker conditions).

Duality and optimality conditions in nonlinear programming.

Algorithms for solving NLPs: The line search methods, method of feasible directions.

Focus on special application in one of the following areas: Financial engineering, supply chain management, airline optimization, production planning.

References

[1] Bazaraa, M. S.; Jarvis, J. J. and Shirali, H. D., *Linear Programming - and Network Flows*, second edition, John Wiley, Singapore, 2003.

[2] Ravindran, A. Ravi (Eds.), *Operations Research and Management Science*, Hand Book , CRC Press, 2009.

[3] Cottle, R. W. and Lemke, C. E. (Eds), *Nonlinear Programming*, American Mathematical Society, Providence, RI, 1976.

[4] Bertsekas, D. P., *Nonlinear Programming*, Athena Scientific, 1999.

26 Generalised Linear Models (Code: ST552)

Generalised linear models, and ordinary linear models.

How to build models (with design matrices).

The exponential family of distributions (Normal, binomial, Poisson etc.).

Checking models: residuals etc.

How GLMs are fitted to data.

Dealing with over parameterized models.

Inference (sampling distribution of parameters and analysis of deviance).

Specific types of GLM

References

[1] Krzanowski, W., *An Introduction to Statistical Modelling*, John Wiley, 2010.

[2] Dobson, A.J., *An Introduction to Generalized Linear Models*, Chapman and Hall, 2010.

27 Decision Theory and Bayesian Analysis (Code: ST581)

Formulation of a decision problem : Randomized and non-randomized decision rules, risk function, optimality of decision rules; utility theory and loss function.

Subjective probability and selection of prior distribution for Bayesian analysis, Bayesian analysis for statistical inference problems of estimation, testing hypotheses, confidence interval and prediction; Bayesian decision theory; admissible and minimax decision rules; complete class of decision rules.

References

[1] Berger, J. O., *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, NY, 1985.

[2] De Groot, M. H., *Optimal Statistical Decisions*, Wiley, 2005.

[3] Ferguson, T. S., *Mathematical Statistics - A Decision Theoretic Approach*, Academic Press, 1967.

28 Biostatistics (Code: ST576)

Basics of biological data, regression and classification, simple and generalized linear models; multiple linear regression, recursive partitioning and nonparametric tree models; exploratory data analysis; clustering and unsupervised learning - parametric and non-parametric methods; classification and supervised learning techniques; survival analysis; longitudinal analysis; principal component analysis and multi-dimensional scaling; Programming in R.

References

[1] Vittinghoff, E.; Glidden, David C.; Shiboski, Stephen C. and McCulloch, Charles E., *Regression Methods in Biostatistics - Linear, Logistic, Survival, and Repeated Measures Models*, Springer, 2012.

[2] Everitt, B. and Hothorn, T., *An Introduction to Applied Multivariate Analysis with R*, Springer, 2011.

29 High-dimensional and Big Data Analytics (Code: ST582)

Basic properties of high-dimensional and big data; data cleaning and visualization; data smoothing and generalized additive models; density estimation and non-parametric methods; dimensionality reduction methods; feature selection techniques - shrinkage, regularization and optimization; bagging and boosting; fast classification and clustering techniques; discriminant analysis; outlier analysis and anomaly detection; association analysis; stream and sensor data analysis; multiple hypotheses testing; MapReduce fundamentals for big data; social network analysis.

References

[1] Hastie, T., Tibshirani, R. and Friedman, J. H., *The Elements of Statistical Learning - Data Mining, Inference and Prediction*, Springer-Verlag, Berlin, 2011.

[2] Bühlmann, P. and van de Geer, S., *Statistics for High-Dimensional Data - Methods, Theory and Applications*, Springer, 2011.

[3] Rajaraman, A. and Ullman, J. D, *Mining of Massive Datasets*, Cambridge University Press, 2011.

30 Machine Learning Using R(Code : ST 573)

Machine learning is the science originated at the intersection of statistics, data-mining, and computer science. It is a powerful tool, capable of finding insight in the large data sets. In general, data used for machine learning contains features of the concepts to be learned, further we summarize this data in the form of a model, which is then used for predictive or descriptive purpose. This course provides a broad introduction to data-mining, pattern recognition and machine learning. We will discuss numerous case studies and application, so that you'll learn how to apply learning algorithms to disease classification (cancer classification), text understanding (mobile phone spam data, letter data), quality control (Wine quality data) and other areas such as Medical expenses data, Concrete data, and Groceries data. In this course, you will learn about most effective machine learning techniques, and gain practice implementing them and getting them to work for yourself. More importantly, you'll learn about not only the theoretical underpinnings of learning, but also gain the practical know-how needed to quickly and powerfully apply these techniques to new problems. R is an open source programming language and software environment for statistical computing and graphics. It provides wide support for machine learning. We will use R platform to learn various machine learning techniques.

Syllabus : Basics of R : Data Structure, File handling, Graphics and Data Visualization, Programming. Managing and understanding data : Data description, Data processing, Dimension Reduction. Evaluating model performance : Measures performance for the classification problem. Introduction to machine learning. Application to Machine Learning : Classification using Nearest Neighbors, Naive Bayes, Regression methods, Neural networks, Support Vector Machine, Market basket analysis using Association Rules, Finding group of data - clustering with K-means and K-medoids. Resampling methods : Leave-One-Out Cross-Validation, k-Fold Cross Validation Dataset used : Breast cancer data, Mobile phone spam data, Medical expenses data, Wine Quality data, Concrete data, Letter data, Groceries data,

References : 1. James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. An Introduction to Statistical Learning: With Applications in R. New York: Springer 2. Lantz, B. 2013. Machine learning with R. Packt Publishing, Birmingham. 3. Maindonald J, Braun J. Data Analysis and Graphics Using R. Cambridge University Press: Cambridge, 2003. 4. Crawley MJ, (2005) Statistics: an introduction using R. Volume 1. 1st edition. New York: John Wiley & Sons. 5. Seefeld, K. & Linder, E. (2007), Statistics Using R with Biological Examples . Department of Mathematics & Statistics, University of New Hampshire, Durham, NH, USA. 6. Vinod, H.D. (Ed.) (2010). Advances in social science research using R, Springer, ISBN 978-1- 4419-1763-8, New York 7. Kleiber C, Zeileis A (2006). Applied Econometrics with R. Springer-Verlag, New York. Forthcoming.

31 Mathematical Finance(Code : ST 585)

1. Financial markets and derivatives; arbitrage and no-arbitrage condition
2. Fundamental theorem of asset pricing
3. State space vectors and risk neutral measures
4. Binomial trees
5. Brownian motion, random walks, the martingale and Markov properties
6. The Itos lemma
7. The Black-Scholes-Merton formula and pricing under risk-neutral measure
8. Multivariate stochastic calculus, Levys characterization of the Brownian motion, Gaussian process and Brownian bridge
9. Existence of risk-neutral measure, no arbitrage, and the first fundamental theorem of asset pricing
10. Uniqueness of risk measures, completeness, and the second theorem of asset pricing
11. Options pricing
12. Dupire local volatility, Heston stochastic volatility, and jump processes

References

[1] K. Back, A Course in Derivative Securities: Introduction to Theory and Computation, Springer, 2005.

[2] M. Baxter and A. Rennie, Financial Calculus: An Introduction to Option Pricing, Cambridge, 1996.

[3] T. Bjrk, Arbitrage Theory in Continuous Time, Oxford, 2004.

[4] J. C. Hull, Options, Futures, and other Derivatives, 6th Edition, Prentice Hall, 2006.

[5] P. Hunt and J. Kennedy, Financial Derivatives in Theory and Practice, Wiley, 2004.

[6] M. Jackson and M. Staunton, Advanced Modelling in Finance using Excel and VBA, Wiley, 2001.

[7] R. Jarrow and S. Turnbull, Derivative Securities, 2nd edition, South-Western College,1999.

[8] I. Karatzas and S. E. Shreve, Brownian Motion and Stochastic Calculus, Springer, 1997.

[9] D. G. Luenberger, Investment Science, Oxford, 1997.

[10] S. R. Pliska, Introduction to Mathematical Finance: Discrete Time Models, Blackwell, 1997.

[11] S. E. Shreve, Stochastic calculus and Finance I: The Binomial Asset pricing Model, Springer, 2004.

[12] S. E. Shreve, Stochastic calculus and Finance II: Continuous Time Models, Springer, 2004.

[13] J. M. Steele, Stochastic calculus and financial applications, Springer, 2000.

[14] P. Wilmott, Paul Wilmott on Quantitative Finance, 2nd edition, 3 volume set, Wiley, .