PONDICHERRY UNIVERSITY

DEPARTMENT OF STATISTICS

SYLLABUS FOR M.Sc. STATISTICS
(CBCS Pattern)
Effective from the Academic Year 2016-2017

PONDICHERRY UNIVERSITY
PUDUCHERRY 605 014

M.Sc. STATISTICS (CBCS - Semester Pattern)

REGULATIONS

Aim of the Course

The Degree of Master of Science in Statistics aims to train the students in the development and applications of Statistical techniques for analyzing data arising in the scientific investigation of problems in various disciplines. It is also proposed to provide first hand practical experience in handling modern statistical software in the analysis of such data.

Eligibility for admission

Candidates for admission to the first year of the M.Sc. (Statistics) degree programme shall be required to have passed the B.Sc. degree examination of any Indian University recognized by the University Grants Commission with Statistics as the main subject or Mathematics as the main subject with Mathematical Statistics as one of the minor subject and a minimum of 55% marks in the main and allied subjects.

Duration of the Course

The course shall be of two years duration spread over four semesters. The maximum duration to complete the course shall not be more than 8 semesters.

Eligibility for admission to Examination

A candidate shall be permitted to appear for the M.Sc. examination in a subject of study only if he/she secures not less than 70% attendance in the subject concerned.

Medium : The medium of instruction shall be English.

Passing Minimum and Weight age of marks

The weight age of marks for Continuous Internal Assessment (CIA) and end semester examinations shall be 40 and 60 respectively. As per the Choice Based Credit System regulations of the Pondicherry University, a student is declared as pass in a given subject he / she secures

(a) A minimum of 40% marks in end-semester exam and

(b) A minimum of 50% marks in aggregate when Internal assessment and End-Semester marks are added together

Supplementary Exam

(a) A failed student who meets the attendance requirement (70%) and has a minimum of 40% in the Internal Assessment marks may be permitted to register for the next End Semester examination in the following semester itself

(b) Students who have failed due to insufficient attendance and / or less than 40% in the Internal Assessment marks should repeat the course as and when it is offered.

Continuous Internal Assessment

The weightage of 40 marks for Continuous Internal Assessment component shall consist of the following:

a)	Internal Assessment Tests (two)	(2 x 15)	= 30 marks
b)	Seminars/Assignments/Presentations/Viva etc.	(1 x 10)	= 10 marks
	Internal Total		= 40 marks

Choice Based Credit System (CBCS)

The M.Sc. Statistics program is offered through a unique CBCS. The salient feature of the CBCS is that the program is offered through credit based courses. Subjects are divided into Hard Core and Soft Core. Hard Core subjects are compulsory. The students have the choice to select from among the list of soft core subjects. Soft core subjects are similar to elective subjects.

A student is expected to complete a minimum of 72 credits within four semesters. Students are assessed and awarded letter grades based on their performances in the respective courses.

PONDICHERRY UNIVERSITY
CHOICE BASED CREDIT SYSTEM
M.Sc. STATISTICS SYLLABUS
Effective from the Academic Year 2016 – 2017

SEMESTER	COURSE CODE	TITLE OF THE COURSE	NATURE OF THE COURSE	NO. OF CREDITS
I	STAT 411	Linear Algebra and Matrix Theory	Hard Core	4
	STAT 412	Probability Theory	Hard Core	4
	STAT 413	Distribution Theory	Hard Core	4
	STAT 414	Programming in R (Lab. Based)	Hard Core	4
		Soft Core	Soft Core	3
II	STAT 421	Theory of Estimation	Hard Core	4
	STAT 422	Sampling Theory	Hard Core	4
	STAT 423	Stochastic Processes	Hard Core	4
	STAT 424	Statistical Laboratory – I with Excel (where ever feasible) Based on STAT 421 & STAT 422	Hard Core	3
		Soft Core	Soft Core	3
III	STAT 531	Multivariate Statistical Analysis	Hard Core	4
	STAT 532	Testing of Statistical Hypotheses	Hard Core	4
	STAT 533	Linear Models and Regression Analysis	Hard Core	4
	STAT 534	Statistical Laboratory – II with Excel (where ever feasible) Based on STAT 531, STAT 532 & STAT 532)	Hard Core	3
		Soft Core	Soft Core	3
IV	STAT 541	Design and Analysis of Experiments	Hard Core	4
	STAT 542	Statistical Laboratory – III with Excel (where ever feasible) Based on STAT 541	Hard Core	3
	STAT 543	Project Work and Viva-Voce/ Dissertation	Hard Core	4
		Soft Core	Soft Core	3
		Soft Core	Soft Core	3

Soft Core Papers

	Semester I		**Semester II**
STAT 415	Operations Research	STAT 425	Decision Theory
STAT 416	Statistical Quality Control	STAT 426	Total Quality Management
STAT 417	Demographic Techniques	STAT 427	Econometrics
		STAT 428	Actuarial Statistics
	Semester III		**Semester IV**
STAT 535	Reliability Theory	STAT 544	Survival Analysis
STAT 536	Time Series Analysis	STAT 545	Biostatistics
STAT 537	Elements of Queuing Theory	STAT 546	Statistical Data Mining Methods
STAT 538	Simulation Techniques	STAT 547	Bayesian Inference

STAT411 - LINEAR ALGEBRA AND MATRIX THEORY CREDITS: 4

Unit I

Vector Spaces, Sub-spaces, Basis of a vector space – Vector spaces with inner products - Gram-Schmidt orthogonalization.

Unit II

Linear transformation (LT) – Properties – Matrix of a linear transformation – Matrix of composite transformation – Matrix of an inverse transformation – Change of basis - Orthogonal transformation - Dual space.

Unit III

Linear equations – Solution space and null space – Sylvester's law of nullity – Generalized inverse of a matrix – Moore – Penrose inverse

Unit IV

Eigen values and Eigen vectors of an LT – left Eigen vectors, right Eigen vectors, Diagonalizable LT – Lambda matrix, Composition of lambda matrices, Operator polynomial, Cayley-Hamilton theorem and minimal polynomial for an LT – Eigen values of matrix polynomials.

Unit V

Bilinear forms - Canonical reduction – Sylvester's law of inertia-Definitions of quadratic form - Lagrange's reduction – Kronecker's reduction -Reduction involving the Eigen values of the matrix, Generalized Eigen value problem.

Books for Study

1. Biswas S. (2012), Text book of Matrix Algebra, Third Edition, PHI Learning Private Limited, New Delhi.
2. Bhattacharya P.B., Jain S.K., Nagpaul S.K. (2012), First Course in Linear Algebra, New Age International (P) Ltd, New Delhi.
3. Parashar B.P. (1989), Linear Algebra, CBS Publishers and Distributors, Delhi.
4. Rao C.R. (2009), Linear Statistical Inference and its Applications, Second Edition, John Wiley and Sons

Books for Reference

1. Friedberg S.H., Insel A.J. and Spence L.E. (2014), Linear Algebra, Pearson Education.
2. Gilbert J. and Gilbret L. (2005), Linear Algebra and Matrix Theory, Academic Press.
3. Lipschutz S. and Lipson M. (2009), Schaum's outlines, Linear Algebra, Fourth Edition, McGraw Hill Company.
4. Rao A.R. and Bhimasankaram P. (2000), Linear Algebra, Hindustan.
5. Searle S.R. and Khuri A.I. (2017), Matrix Algebra useful for Statistics, Second Edition, John Wiley and Sons, New Jersey.
6. Searle S.R. and Gruber MHI (2016), Linear Models, Second Edition, John Wiley and Sons, New Jersey.

STAT412 - PROBABILITY THEORY **CREDITS: 4**

Unit I

Algebra of sets - fields and sigma-fields, Inverse function – Measurable function – Probability measure on a sigma field – simple properties - Probability space - Random variables and Random vectors – Induced Probability space – Distribution functions – Decomposition of distribution functions.

Unit II

Expectation and moments – definitions and simple properties – Moment inequalities – Holder, Jenson, Chebyshev, Markov Inequalities– Characteristic function – definition and properties – Inversion formula.

Unit III

Convergence of a sequence of random variables - convergence in distribution, convergence in probability, almost sure convergence and convergence in quadratic mean - Weak convergence of distribution functions – Slustky theorem - Helly-Bray theorem.

Unit IV

Definition of product space – Fubini's theorem (statement only) - Independence of two events – Independence of classes – Independence of random variables – properties – Borel zero –one law.

Unit V

Law of large numbers - Khintchin's weak law of large numbers, Kolmogorov strong law of large numbers (statement only) – Central Limit Theorem – Lindeberg – Levy theorem, Linderberg – Feller theorem (statement only), Liapounov theorem – Relation between Liapounov and Linderberg – Feller forms – Radon Nikodym theorem and derivative (without proof) – Conditional expectation – definition and simple properties.

Books for Study

1. Bhat, B. R. (2007): Modern Probability Theory, 3rd edition, New Age International Pvt. Ltd.
2. Ash, R.B. (1972): Real Analysis and Probability, Academic Press.
3. Rohatgi, V.K. and Saleh (2002): An Introduction to Probability Theory and Mathematical Statistics, John Wiley

Books for Reference

1. Athreya K B and Lahiri S N (2005):Measure Theory, Hindustan Book Agency.
2. Tucker, H.G. (1967): A Graduate course in Probability, Academic Press.
3. Burill, C.W. (1972): Measure, Integration and Probability, McGraw Hill.
4. Chow, Y.S. and Teicher, H. (1979): Probability Theory, Springer.
5. Loeve, M. (1985). Probability Theory, 3rd edition, Springer..
6. Resnick S.I. (2001): A Probability Path, Birkauser.
7. Basu A K. and A Bandopadhyay (2012): Measure Theory and Probability, PHI Learning Pvt. Ltd.

STAT 413 – DISTRIBUTION THEORY **CREDITS: 4**

Unit I

Brief review of distribution theory, distribution of functions of random variables - Laplace, Cauchy, Inverse Gaussian, Lognormal, Logarithmic series and Power series distributions - Multinomial distribution

Unit II

Bivariate Binomial – Bivariate Poisson – Bivariate Normal- Bivariate Exponential of Marshall and Olkin - Compound, truncated and mixture of distributions, Concept of convolution

Unit III

Multivariate normal distribution (Definition and Concept only) - Sampling distributions: Non-central chi-square, t and F distributions and their properties - Distributions of quadratic forms under normality-independence of quadratic form and a linear form- Cochran's theorem

Unit IV

Order statistics, their distributions and properties- Joint and marginal distributions of order statistics - Distribution of range and mid range - Extreme values and their asymptotic distributions (concepts only)

Unit V

Empirical distribution function and its properties - Kolmogorov-Smirnov distributions -Life time distributions - Exponential and Weibull distributions - Mills ratio -Distributions classified by hazard rate.

Books for Study

1. Mood M., Graybill F.A. and Boes D.C.(2001) : Introduction to the Theory of Statistics, Tata McGraw-Hill, New Delhi.
2. Johnson, N.L,Kotz, S. and Balakrishnan, N. (1994): Continuous Univariate Distributions, Vol.1 &2, Wiley Series in Probabilty and Statistics.
3. Johnson, N.L , Kemp A.W. & Kotz, S. (1994): Univariate Discrete Distributions, Wiley Series in Probabilty and Statistics
4. David H. A. and Nagaraja H.N.(2003): Order Statistics, 3/e, John Wiley & Sons.

Books for Reference

1. Rao C. R.,(1973): Linear Statistical Inference and its Applications, Wiley Eastern Ltd, New Delhi.
2. Dudewicz, E.J and Mishra, S.N(1980): Mathematical Statistics, John Wiley, NY.
3. Kocherlakota S and Kocherlakota K(1992): Bivariate Discrete distributions, M. Dekker.
4. Balakrishnan N and Lai C.D.(2009): Continuous Bivariate Distributions, Springer.
5. Rohatgi, V.K. and Saleh (2002): An Introduction to Probability Theory and Mathematical Statistics, John Wiley.
6. Parimal Mukhopadhyay(2006):Mathematical Statistics, 3/e, Books and Allied (P) Ltd, Kolkata.

STAT 414–PROGRAMMING IN R **CREDITS: 4**
(Lab. Based)

Unit I
R language Essentials: Expressions and objects, Assignments, creating vectors, vectorized arithmetic, creating matrices, operations on matrices, lists, data frames – creation, indexing, sorting and conditional selection ; examples.

Unit II
R Programming: conditional statements – if and if else; loops – for, while, do-while; functions – built-in and user defined; Data entry – reading from text file, data editor; examples.

Unit III
Descriptive Statistics and Graphics: Obtaining summary statistics; generating tables; Bar plots, Pie charts, Box plots, Histogram; exercises.

Unit IV
Probability and Distributions: Random sampling and combinatory; obtaining density, cumulative density and quantile values for discrete and continuous distributions; generating samples from discrete and continuous distributions; Plotting density and cumulative density curves; Q-Q plot.

Unit V
Correlation: Pearson, Spearman and Kendall's correlation; Regression – fitting, obtaining residuals and fitted values; one and two sample tests for mean and variance – one way and two way ANOVA.

Books for Study

1. Michael J.Crawley (2007), The R Book, John Wiley and Sons Ltd.
2. Peter Dalgaard (2008), Introductory Statistics with R, 2nd edition, Springer.

Lab Exercises:
1. Operations on vectors and matrices
2. Creating and manipulating data frames.
3. Writing user defined functions for finding arithmetic mean, median, factorial, matrix addition and multiplication.
4. Bar and Pie charts.
5. Box plots for single and multiple groups.
6. Density and cumulative density plots for Binomial, Poisson, Normal and exponential distributions.
7. Checking Normality using Histogram and Q-Q plot.
8. Correlation coefficient – Pearson's, Spearman and Kendall's Tau.
9. Fitting simple linear and multiple linear regressions.
10. One sample and two sample t test.
11. One way and two way ANOVA.

STAT 421 - THEORY OF ESTIMATION **CREDITS: 4**

Unit I

Parametric point estimation – properties of estimators – Consistency and its different forms Sufficient condition for consistency- Unbiasedness – sufficient statistics – Factorization theorem – Distributions admitting sufficient statistic – Exponential and Pitman families procedure for finding minimal sufficient statistic.

Unit II

The information measure – Cramer – Rao (CR) inequality – Chapman – Robbins (KCR) inequality (single parameter case only) – Bhattacharya inequality (single parameter case only) – minimum variance bound estimator- Invariant (equivariant) estimators (concepts only)

Unit III

Uniformly minimum variance unbiased estimators (UMVUE)- condition for the existence of UMVUE- Completeness and Bounded completeness- Relation between complete statistic and minimal sufficient statistic- Rao – Blackwell Theorem- Lehmann – Scheffe's theorem.

Unit IV

Methods of estimation – method of moments and its properties – method of maximum likelihood and its properties-Large sample properties of MLE - Method of minimum chi-square and its properties – Methods of least squares – Optimum properties of least square estimates in linear model.

Unit V

Interval estimation – Pivotal method of construction – shortest confidence intervals and their construction (minimum average width) – Construction of shortest confidence intervals in large samples. Decision Theory: Simple problems involving quadratic error loss function – Elementary notions of minimax estimation – Simple illustrations.

Books for Study

1. Rajagopalan M and Dhanavanthan P (2012): Statistical Inference, PHI Learning, New Delhi.
2. Casella, G. and Berger, R.L. (2002):Statistical Inference, Duxubury Process, Belmont, USA.
3. Rohatgi, V.K. (2003): Statistical Inference, Dover Publications, New York.

Books for Reference

1. Lehmann, E.L and Casella G(1998) :Theory of Point Estimation, 2/e, Wiley Eastern Ltd.
2. B.K.Kale and K.Muralidharan (2015), Parametric Inference – An Introduction, Narosa Publishing House
3. Kale, B.K. (1999): A First course on Parametric Inference , Narosa Publishing House.
4. Zacks,S. (1981): Parametric Statistical Inference, John Wiley, NY.
5. Srivastava, Khan and Srivastava (2014), Statistical Inference: Theory of Estimation, PHI, India

STAT 422 – SAMPLING THEORY **CREDITS: 4**

Unit I

Preliminaries – Sampling Designs – Simple random sampling– Stratified Random Sampling – Allocation problems – Systematic Sampling Schemes – Linear, Circular, Balanced and Modified systematic sampling methods

Unit II

Probability Proportional to size sampling- Inclusion Probabilities – Horvitz-Thompson estimator – Yates –Grundy Form –Midzuno Sampling design – PPSWOR- Des-Raj's Ordered estimator – Murty's unordered estimators

Unit III

Ratio estimators and their properties in Simple Random Sampling – Ratio estimators in Stratified Random sampling – Regression Estimators, Regression estimators in Stratified Random Sampling – Multivariate Ratio estimators and Multivariate Regression Estimators

Unit IV

Cluster Sampling: Equal cluster sampling – Estimators of mean and variance, optimum cluster size, Unequal cluster sampling – Estimators of mean and variance – Two stage sampling – variance of the estimated mean – Double Sampling for stratification and Ratio estimation

Unit V

Randomized response methods – Warner's, Simmon's and Two Stage response methods – Sources of errors in Surveys – Mathematical model for the effects of call-backs and the errors of measurement – Official Statistical Systems in India – Role of NSSO and CSO and their activities – Organization of Large Scale Sample Surveys.

Books for Study

1. Cochran, W.G. (1977): Sampling Techniques, 3/e, Wiley Eastern Ltd,.
2. Gupta, A. K. and Kabe D.G, (2011): Theory of Sample Surveys, World Scienific Publishing Co. Pte. Ltd., Singapore
3. Singh, D. and Choudhary, F.S (1986): Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd,.
4. Sukhatme PV. Etal. (1984): Sampling Theory of Surveys with Applications, Iowa State University Press and ISARI Publications, New Delhi

Books for Reference

1. Desraj and Chandhok P.(1998): Sampling Theory, Narosa Publications, New Delhi
2. Kish, L(1995) : Survey Sampling, John Wiley and Sons.
3. Murthy, M.N (1979): Sampling Theory and Methods, Statistical Publishing Society, Calcutta.
4. Sarjinder Singh (2004): Advanced Sampling – Theory with Applications, Kluwer Publications

STAT 423 – STOCHASTIC PROCESSES **CREDITS: 4**

Unit I

Stochastic processes and their classification – Markov chain– Examples (Random walk, Gambler's ruin problem)- classification of states of a Markov chain-Recurrence-Basic limit theorem of Markov chains-Absorption probabilities and criteria for recurrence.

Unit II

Markov chains continuous in time – General pure birth processes and Poisson process, birth and death processes, finite state continuous time Markov chains.

Unit III

Branching processes discrete in time – Generating functions relations – Mean and variance – Extinction probabilities – Concept of Age dependent Branching process

Unit IV

Renewal processes – Definition and examples – key renewal theorem – Study of residual life time process

Unit V

Stationary process – weakly and strongly stationary process – Moving average and Autoregressive processes and their covariance functions – Brownian Motion process – Joint probabilities for Brownian motion process – Brownian motion as a limit of random walk

Books for Study

1. Karlin, S and Taylor, H.M(1975): A First Course in Stochastic Processes, Academic Press, New York.
2. Medhi,J (2009): Stochastic Processes, 3/e, New age International.
3. Bhat B.R.(2004): Stochastic Models: Analysis and Applications, New Age Publications

Books for Reference

1. Bhattacharya and Waymire, E.C. (1992): Stochastic Process with Applications John Wiley and sons.
2. Jones,P.W and Smith,P(2001): Stochastic Processes: An Introduction, Arnold Press.
3. Cinlar, E(1975): Introduction to Stochastic Processes, Prentice-Hall Inc., New Jersey.
4. Cox, D.R and Miller, H.D(1983) : Theory of Stochastic Processes – Chapman and Hall, London,Third Edition
5. Prabu N.U. (1965): Stochastic Processes Macmillan.
6. Ross S.M (1983): Stochastic Process, Wiley.
7. Sidney I. Resnick (1992):Adventures in Stochastic Processes, Birkauser, Boston.
8. G. Grimmett and D. Stirzaker (2001): Probability and Random Processes, 3/e, OUP Oxford.

STAT 424 - STATISTICAL LABORATORY – I **CREDITS: 3**

(with Excel (where ever feasible) and SPSS)

Based on STAT 421 and STAT 422

I. Estimation (20 marks)

1. MLE and Standard error of ML estimators.
2. MLE through the method of successive approximation.
3. MLE for truncated distribution.
4. Method of Moments
5. Method of Minimum Chi-square
6. Method of Least square
7. Interval estimation: Confidence interval for mean, difference of means, variance and ratio of variances.

II. Sampling Theory (20 marks)

1. Simple random sampling methods of drawing sample – Estimation of the population total and variance estimation.
2. PPSWR – Hurwitz Thompson estimator - Des Raj ordered estimator – Murthy's unordered estimator – Midzuno scheme.
3. Linear and circular systematic sampling.
4. Stratified random sampling Allocation Problems
5. Cluster sampling – of equal size.
6. Ratio and Regression estimation.

III. Computations based on SPSS software (20 marks)

1. Graphical plots: Box-Whisker plots, Histograms and Population Pyramids
2. Test for normality : P-P Plot, Q-Q Plot
3. Random number generation.(i) Binomial, (ii) Poisson, (iii) Normal
4. Computation of simple, multiple, partial and rank correlation coefficients.
5. Computation of simple regression and test for intercept and slope.
6. Fitting of curves – parabola, cubic and exponential.
7. Testing of Hypothesis
   (i) Two sample t-test and confidence interval
   (ii) Paired Samples t-test and confidence interval
   (iii) Chi square for independence of attributes
8. One way ANOVA with post hoc tests (Dunnett's test and Duncan's Multiple Range Test (DMRT)).
9. Two Way ANOVA with post hoc tests (Tukey's and Bonferonni'e test)).

STAT 531 – MULTIVARIATE STATISTICAL ANALYSIS CREDITS: 4

Unit I

Multivariate normal distribution– Marginal and conditional distributions – characteristic function. Maximum likelihood estimation of the parameters of Multivariate Normal and their sampling distributions – Inference concerning the mean vector when covariance matrix is known

Unit II

Total, Partial, Multiple correlation in the Multivariate setup – MLEs of Total, Partial and Multiple correlation coefficients. Sampling distributions of Total and Multiple Correlation in the null case. Hotelling T^2 statistic, derivation and its distribution –Uses of T^2 statistic - relation between T^2 and D^2– Mahalanobis D^2 statistic and its distribution

Unit III

Generalized variance – Wishart distribution (statement only) – Properties of Wishart distribution – Test for covariance matrix – Test for equality of covariance matrices

Unit IV

Classification problems – Classification into one of two populations (known and unknown dispersion matrix) – Classification in to one of several populations – Fisher's Linear discriminant function

Unit V

Principal components –properties, Extraction of Principal components and their variances Canonical correlation – Estimation of canonical correlation and variates. Factor analysis – Mathematical model- Estimation of Factor Loadings — Concept of factor rotation – Varimax criterion

Books for Study

1. Anderson, T.W. (2003) : An Introduction to Multivariate Statistical Analysis, Wiley Eastern Ltd.
2. Johnson, R. A and. Wichern D.W (2007): Applied Multivariate Statistical Analysis, 6 /e, Prentice-Hall of India Private Ltd., New Delhi.

Books for Reference

1. Giri, N.C(2003): Multivariate Statistical Inference, Academic Press, NY
2. Morrison, F(1985): Multivariate Statistical Methods, Mc Graw Hill Book Company.
3. Rao, C.R(1998): Linear Statistical Inference and its Applications, Wiley Eastern Ltd,.
4. Alvin C. Rencher(2002): Methods of Multivariate Analysis, 2/e, Wiley Interscience
5. Srivastava M.S. and Khatri C.G.(1979):Introduction to Multivariate Analysis, Elsevier

STAT 532 – TESTING OF STATISTICAL HYPOTHESES CREDITS: 4

Unit I

Randomized and non-randomized tests, Neyman – Pearson fundamental lemma, Most powerful tests, Uniformly most powerful test, Uniformly most powerful test for distributions with monotone likelihood ratio, Generalization of fundamental lemma and its applications

Unit II

Unbiasedness for hypothesis testing, Uniformly most powerful unbiased tests, Unbiased tests for one parameter exponential family, Similar test and complete sufficient statistics, Similar tests with Neyman structure, Locally most powerful tests.

Unit III

Invariant tests, maximal invariants, Uniformly most powerful invariant tests, Consistent tests, Likelihood ratio test, its properties and its asymptotic distribution, Applications of the LR method.

Unit IV

Non-parametric tests: Goodness of fit test : Chi-square and Kolmogorov Smirnov test - Test for randomness, Wilcoxon Signed rank test – Two sample problem: Kolmogrov-Smirnov test, Wald-Wolfowitz run test, Mann-Whitney U test, Median test, Kruskal Wallis test and Friedman's test

Unit V

Sequential tests: Basic Structure of Sequential tests – Sequential Probability Ratio Test (SPRT) and its applications – Determination of the boundary constants – Operating Characteristic and expected sample size of SPRT – Optimum properties of SPRT.

Books for Study

1. Rajagopalan M and Dhanavanthan P (2012): Statistical Inference, PHI Learning, New Delhi.
2. Lehmann, E.L and Joseph P. Romano (2005): Testing Statistical Hypotheses, 3/e, Springer
3. Rohatgi, V.K.(2003): Statistical Inference, Dover Publications,.
4. Gibbons, J.D. (1985) : Non Parametric Statistical Inference , 2/e , Marckel Decker.

Books for Reference

1. Casella, G & Berger, R.L (1990):Statistical Inference , Duxubury Press, Belmont. USA
2. Ghosh,B.K(1970): Sequential Tests of Statistical Hypotheses, Addison Wesley.
3. Parimal Mukhopadhyay(2006):Mathematical Statistics, 3/e, Books and Allied (P) Ltd, Kolkata.
4. Manoj Kumar Srivastava and Namita Srivastava (2009): Statistical Inference – Testing of Hypotheses, Prentice Hall of India

STAT 533 –LINEAR MODELS AND REGRESSION ANALYSIS CREDITS: 4

Unit I

Full rank linear model – least square estimators of the parameters and their properties – Gauss-Markov theorem – Model in centered form – Estimators under normality assumption and their properties – Coefficient of determination – Generalized least squares – misspecification of the error structure and the model.

Unit II

Test for overall regression and for a subset of the parameters – test in terms of R^2 – General Linear Hypothesis testing – special cases – confidence region for the parameters and the mean – prediction intervals – likelihood ratio tests for the parameters – study of the residual outliers and influential observations

Unit III

Selection of input variables and model selection – Methods of obtaining the best fit – Stepwise regression, Forward selection and backward elimination – Multicollinearity – Collinearity diagnostics – Causes, Consequences and Remedy –Departure from normality

Unit IV

Introduction to general non-linear regression – Least squares in non-linear case – Estimating the parameters of a non-linear system – Reparametrisation of the model – Non-linear growth models – Concept of non-parametric regression

Unit V

Robust regression – Linear absolute deviation regression – M estimators – Robust regression with rank residuals – Resampling procedures for regression models – methods and its properties (without proof) - Jackknife techniques and least squares approach based on M-estimators.

Books for Study

1. Alvin C. Rencher (2000): Linear Models in Statistics, John Wiley & Sons, New York (Chapters 7,8 & 9 for Unit I & II)
2. Draper, N and Smith, H (1998): Applied Regression Analysis, 3rd Edition, Wiley-Interscience.
3. Elizabeth C. Peck, Douglas C. Montgomery, G. Geoffry Vinning (2006): Introduction to Linear Regression Analysis, 3/e, John Wiley & Sons.

Books for Reference

1. Chaterjee, S, Ali S. Hadi (2013): Regression Analysis by Example, 5thedition, John Wiley.
2. Searle, S.R. (1997): Linear Models, John Wiley.
3. Thomas P.Ryan(2006): Modern Regression Methods, John Wiley and Sons,Inc.
4. Seber G.A.F and Wild C.J. (2003): Nonlinear Regression, John Wiley & Sons

STAT 534 – STATISTICAL LABORATORY – II **CREDITS: 3**

(with Excel (where ever feasible) and SPSS)

Based on STAT 531, STAT 532 and STAT 533

I Testing of Hypotheses **(30 marks)**

1. Construction of randomized and nonrandomized MP, UMP and UMPU tests of hypotheses and drawing the power curves.
2. Construction of SPRT and its OC and ASN curves.
3. Non parametric tests:
   Kolmogorov Smirnov test, Mann-Whitney U test, Median test for k-sample problem, Kruskal Wallis test and Friedman's test

COMPUTATIONS BASED ON SPSS (30 MARKS)

II Multivariate Statistical Analysis

1. Test for equality of mean vectors when covariance matrix is unknown (Hotelling's T^2 test)
2. Test for Covariance matrices
3. Discriminant Analysis
4. Canonical correlation and canonical variables
5. One Way MANOVA with Post hoc tests (DMRT and Tukey's).
6. Principal Component Analysis
7. Factor Analysis

III Linear Models and Regression Analysis

1. Fitting of Multiple linear regression model
2. Residual Analysis for model adequacy, detection of outliers and influential observations
3. Variable Selection procedures
4. Collinearity Diagnostics

STAT 541 – DESIGN AND ANALYSIS OF EXPERIMENTS CREDITS: 4

Unit I

Notion of design matrix- general analysis of design models (Inter and Intra Block analysis) – C Matrix and its properties – Expected Mean Squares (EMS) and its uses- Algorithm for calculating EMS – Two way elimination of heterogeneity – Orthogonality – Connectedness and resolvability

Unit II

Principles of scientific experimentation – Basic Designs: Overview of Completely Randomized Design (CRD), Randomized Block Design (RBD) and Latin Square Design (LSD) – Analysis of RBD (with one observation per cell, more than one but equal number of observations per cell) – Derivation of one and two missing values: Iterative and non-iterative methods – Loss of Efficiency due to missing values- Multiple comparison test: Least Significant Difference, Student Newman Kuel , Duncan's Multiple Range, Tukey tests.

Unit III

Balanced Incomplete Block Design (BIBD)– Types of BIBD – Simple construction methods – Concept of connectedness and balancing – Intra Block analysis of BIBD – Recovery of Inter Block information – Partially Balanced Incomplete Block Design with two associate classes – intra block analysis only - Split plot and strip plot design and their analysis.

Unit IV

Youden square and lattice design and their analysis – Analysis of Covariance with one concomitant variable in CRD and RBD

Unit V

Factorial experiments: 2^2, 2^3, 2^4 and 3^2, 3^3 experiments and their analysis – Complete and Partial Confounding - Fractional Replication in Factorial Experiments

Books for Study

1. Das, M.N. and Giri, N.C(1979): Design and Analysis of Experiments, Wiley Eastern Ltd,
2. Douglas C. Montgomery (2009) : Design and Analysis of Experiments, 7/e, John Wiley and Sons,
3. Graybill, F.A(1961) : An Introduction to Linear Statistical Models, Mc Graw Hilll Book Company

Books for Reference

1. John, P.W.M (1971): Statistical Design and Analysis of Experiments, Mc Graw Hill Book Company.
2. Kempthorne, O(1966): The Design and Analysis of Experiments, John Wiley and Sons.
3. Ragahavarao, D(1971): Constructions and Combinatorial Problems in Design of Experiments, John Wiley and Sons.
4. Searle, S.R(1987) : Linear Models, John Wiley and Sons.
5. Cochran .W.G. and Cox .G.M. (1995): Experimental designs, 4/e, Wiley.
6. Cobb G.W.(1998): Introduction to Design and Analysis of Experiments.
7. Parimal Mukhopadhyay(2005):Applied Statistics, 2/e, Books and Allied (P) Ltd, Kolkata.
8. R. Paneerselvam (2012), Design and Analysis Experiments, PHI Learning Pvt. Ltd
9. K.Krishnaiah and P. Shahabudeen (2013), Applied Design of Experiments and Taguchi Mehtods, PHI Learning Pvt. Ltd.

STAT 542 – STATISTICAL LABORATORY – III **CREDITS: 3**
(Based on STAT 541)

I. Design of Experiments (60 marks)

1. Multiple Comparison tests (Least Significant Difference (LSD) test, SNK test, Tukey test, Duncan's test)
2. Missing Data Analysis- one and two observations in RBD
3. Missing Data Analysis- one and two observations in LSD
4. 2^4 , 3^2 factorial experiments
5. Fractional factorial experiments
6. Complete confounding in 2^4, 3^2 factorial experiments
7. Partial confounding in 2^4, 3^2 factorial experiments
8. Split plot design
9. BIBD
10. Youden Square Design
11. Analysis of Covariance – CRD – One Concomitant Variable
12. Analysis of Covariance – RBD – One Concomitant Variable

STAT 543 – PROJECT WORK AND VIVA-VOCE/DISSERTATION CREDITS: 4

1. A project work is compulsory and shall be offered in semester IV. Project submission is in Semester IV but the allocation of students should be done at the end of II semester.
2. A project work may be taken individually or by a group of students (not more than three per batch).
3. Project work shall be supervised by faculty members assigned by the Head of the Department at the end of second semester.
4. The orientation of Project work shall be neither of Theory paper nor of a lab/practical paper. Hence, the work allotment of each teacher will be one hour per week per batch of students subject to the maximum of four hours per faculty member in a week.
5. Review meetings are to be done periodically (fortnightly/monthly) to the allocated students by the respective supervisor
6. The project work should be selected in such a way that there is enough scope to apply and demonstrate the statistical techniques learnt in the course.
7. At the end of the semester, a report on the work done should be submitted (two copies). If a team of two students jointly do a project work then they must submit individual reports separately (not copy of the same report).
8. The project report shall clearly state the selected problem, the statistical methodologies employed for data collection and analysis and the conclusions arrived at. Details of previous studies in the area of work and related references should also be given.
9. The project work will be assessed for a maximum of 100 marks. Each student shall give a seminar before the end of the semester on their project work which will be evaluated internally for a maximum of 40 marks. There will be an external viva-voce examination for a maximum of 20 marks by an internal and an external examiner. The parameters for viva voce include (i) Clarity of presentation (ii) Clarity of the content / concept (iii) response to the queries and (iv) relevance of topic for carrying out the project. The project report will be valued by the same external and internal examiner for a maximum of 40 marks.

SOFT CORE PAPERS

SEMESTER I

STAT 415 – OPERATIONS RESEARCH **CREDITS: 3**

Unit I

(Review of Linear Programming Problem (LPP)–Simplex, Big M and Two Phase methods) – Revised simplex method- Duality in LPP – Dual Simplex method –Some important theorems on duality- Sensitivity Analysis–Variation in cost vector and requirement vector– Addition and deletion of single variable and single constraint

Unit II

Integer Programming Problem (IPP) - Gomory's cutting plane algorithm– Mixed IPP – Branch and Bound technique - Dynamic programming problem (DPP) - Bellman's principle of optimality - General formulation - computation methods and application of DPP - Solving LPP through DPP approach

UnitIII

Inventory models – Deterministic inventory models – Classic EOQ model – EOQ with price breaks – EOQ with storage limitations – Probabilistic Inventory models – Continuous review model – Single period model - No setup model – setup model (s-S policy)

Unit IV

Non-linear programming problem–KuhnTucker conditions–Quadratic Programming Problem(QPP) - Wolfe's and Beale's algorithms for solving QPP – Convex programming

UnitV

Queuingtheory–Basiccharacteristicsofqueuingmodels–Arrivalandservicedistribution– steady state solution of M/M/1andM/M/C models with associated distribution of queue length and waiting time - M/G/1 queue-steady results using embedded Markov chain Methods- Pollazcek Khinchin formula.

Books for Study

1. Hillier FS andLibermannGJ(2002):IntroductiontoOperationsResearch,7th Edition, McGraw Hill
2. KantiSwarup,P.K.GuptaandManMohan(2004):OperationsResearch,SultanChand and Sons, New Delhi.
3. Gross D, Shortle J.F. , Thompson J.M. and Harris C.M. (2011): Fundamentals of Queuing Theory, John Wiley & Sons

Books for Reference

1. SinhaSM(2006):MathematicalProgramming: TheoryandMethods,ElsevierPublications.
2. Devi Prasad (2015), Operations Research, Narosa Publishing House
3. KapoorV.K.(2008):OperationsResearch,8/e,SultanChand&Sons
4. Sharma .S.D(1999): Operation Research , Kedar Nath RamNath & Co., Meerut.
5. HamdyA.Taha(1987):OperationsResearch–AnIntroduction, 4/e,PrenticeHallofIndia,PrivateLtd,NewDelhi.
6. Sujit K. Bose (2012), Operations Research Methods, 2/e, Narosa Publishing House
7. K. Chandrasekhara Rao and Shanti Lata Misra (2012), Operations Research, Narosa Publishing House

STAT 416 – STATISTICAL QUALITY CONTROL **CREDITS: 3**

Unit I

Modified control charts for mean – CUSUM chart – technique of V-mask – Weighted Moving average charts – multivariate control charts – Hotellings T^2 control charts and Economic design of X-bar chart

Unit II

Process Capability analysis: Meaning, Estimation technique for capability of a process – Capability Indices: Process capability ratios C_p; C_{pk}, C_{pm}, C_{mk}, C_{pc} – Process capability analysis using a control chart – Process capability analysis using design of experiments

Unit III

Acceptance sampling – Terminologies – Attribute sampling plan by attributes – Single sampling plan and Double sampling plan – OC, ASN, AOQ, AOQL and ATI curves – MILSTD -105E Tables

Unit IV

Acceptance sampling variables for process parameter – Sequential plans for process parameter (σ known and unknown) – Sampling variables for proportion non-conforming - $\bar{X}$ method, K method –

Unit V

Double specification limits – M-method, Double sampling by variables - MILSTD -414 Tables – Continuous Sampling plan – CSP-1, CSP-2, CSP-3, Wald and Wolfowitz SP-A and SP- B

Books for Study

1. Douglas C. Montgomery (2009): Introduction to Statistical Quality Control, 6/e, John Wiley and Sons, New York.
2. Edward G. Schilling, Dean V. Neubauer, (2009), Acceptance Sampling in Quality Control, Second Edition, Taylor & Francis
3. Oakland, J.S.(1989): "Total Quality Management", Butterworth–Hcinemann Ltd., Oxford

Books for Reference

1. Mittage, H.J and Rinne, H(1993): Statistical Methods of Quality Assurance, Chapmann Hall, London, UK
2. Zairi (1991): "Total Quality Management for Engineers", Wood Head Publishers.
3. Juran J.M and Frank M.Gryna Jr .(1982): "Quality Planning and Analysis", TMH, India.

STAT 417 - DEMOGRAPHIC TECHNIQUES **CREDITS: 3**

Unit I
Sources of demographic Statistics, Basic demographic measures: Ratios, Proportions and percentages, Population Pyramids, Sex ratio Crude rates, Labour force participation rates, Density of population, Probability of dying.

Unit II
Life tables: Construction of a life table, Graphs of l_x, q_x, d_x, Funtions L_x, T_x, and E_x. Abridged life tables Mortality: Rates and Ratios, Infant mortality, Maternal mortality, Expected number of deaths, Direct and Indirect Standardization, Compound analysis, Morbidity.

Unit III
Fertility: Measures of Fertility, Reproductivity formulae, Rates of natural increase, Fertility Schedules, Differential fertility, Stable Populations, Calculation of the age distribution of a stable population

Unit IV
Population estimates, Population Projections: Component method, Mortality basis for projections, Fertility basis for projections, Migration basis for projections.

Unit V
Ageing of the population, Estimation of demographic measures from incomplete data.

Books for Study:

1. Pollard, A. H. Yusuf, F. and Pollard, G.N. (1990). Demographic Techniques, Pergamon Press, Chapters 1-8, 12.

Books for Reference:

1. Keyfitz, N. (1977) Applied Mathematical Demography A Willey-Interscience Publication.
2. Keyfilz, N. (1968) Introduction to the Mathematic of Population Ready, Mass: Addition-Wesley.
3. Keyfilz, N. and Caswell, H. (2005) Applied Mathematical Demography, Third edition, Springer.

SEMESTER – II

STAT 425–DECISION THEORY **CREDITS: 3**

Unit I

Basic elements of a decision problem - Randomized and non-randomized decision rules - Estimation and testing of hypothesis as decision problems - Bayes approach to inference and decision -

Unit II

Loss functions - Prior and Posterior distributions and its analysis for Bernoulli, Poisson, and normal processes - Decision principles and Baye's risk–

Unit III

Utility theory - axioms, construction of utility functions, sufficiency, equivalence of classical and Bayesian sufficiency, complete and essentially complete classes of decision rules

Unit IV

Minimax analysis - Basic elements of game theory - General techniques of solving games - Finite games - Supporting and separating hyper plane theorems - Minimax theorem - Minimax estimation for normal and Poisson means

UNIT V

Admissibility of Bayes and minimax rules, General theorems on admissibility, Robustness of Bayes rules, Invariant decision rules, Location parameter problems, Confidence and credible sets.

Books for Study:

1. James O. Berger (1980): Statistical Decision Theory and Bayesian Analysis, Springer Verlag
2. M.H. DeGroot (1970): Optimal Statistical Decisions, John Wiley
3. H. Raiffa and R. Schlaifer (2000): Applied Statistical Decision Theory, Wiley

Books for Reference:

1. Zellener (1971): An Introduction to Bayesian Inference in Econometrics, Willey
2. Hayes J. G and Winkler R I (1976): Probability, Statistics and Decision, Dower
3. Anthony O' Hagan (1994): Kendall's Advanced theory of Statistics Vol. 2B, Bayesian Inference, John Wiley
4. Casella, G. and Berger, R.L. (2002):Statistical Inference, Duxubury Process, Belmont, USA.

STAT 426 -TOTAL QUALITY MANAGEMENT CREDITS: 3

Unit I

Need for TQM, evolution of quality, Definition of quality, TQM philosophy – Contributions of Deming, Juran, Crosby, Taguchi and Ishikawa.

Unit II

Vision, Mission, Quality policy and objective, Planning and Organization for quality, Quality policy Deployment, Quality function deployment, Analysis of Quality Costs.

Unit III

Customer focus, Leadership and Top management commitment, Employee involvement – Empowerment and Team work, Supplier Quality Management, Continuous process improvement, Training, performance Measurement and customer satisfaction.

Unit IV

PDSA, The Seven QC Tools of Quality, New Seven management tools, Concept of six sigma, FMEA, Bench Marking, JIT, POKA YOKE, 5S, KAIZEN, Quality circles.

Unit V

Need for ISO 9000 Systems, clauses, Documentation, Implementation, Introduction to QS 9000, Implementation of QMS, Case Studies.

Books for Study

1. Dale H.Besterfiled(2002): "Total Quality Management", Pearson Education Asia
2. Oakland.J.S(1989): "Total Quality Management", Butterworth–Hcinemann Ltd., Oxford

Books for Reference

1. Narayana V. and Sreenivasan, N.S.(1996): "Quality Management – Concepts and Tasks", New Age International.
2. Zeiri(1991): "Total Quality Management for Engineers", Wood Head Publishers.
3. Juran J.M and Frank M.Gryna Jr.(1982): "Quality Planning and Analysis", TMH, India.
4. Brain Rethery(1993): ISO 9000, Productivity and Quality Publishing Pvt.Ltd.
5. D.Mills(1993): Quality Auditing, Chapman and Hall.

STAT 427 – ECONOMETRICS **CREDITS: 3**

Unit I

Nature and Scope of Econometrics - Review of General Linear Model (GLM), Ordinary Least Squares (OLS), Generalized Least Squares (GLS) and Multicolloinearity

Unit II

Heteroscedasticity - Autocorrelation, its consequences and tests - Ridge regression -Linear regression with stochastic regressors - Instrumental variable estimation - Errors in variables - Autoregressive linear regression - Distributed lag models.

Unit III

Simultaneous linear equations model - Identification problem - Restrictions on structural parameters - rank and order conditions - Restrictions on variances and covariances - Estimation in simultaneous equations model

Unit IV

Forecasting - Univariate forecasting methods - Forecasting in regression models - Forecasting with Simultaneous equations model - Evaluation of forecasts - Combination of forecasts

Unit V

Definition of casuality - Granger casuality - testing of casuality - Cointegration, Bivariate cointegration tests - multivariate cointegration.

Books for Study:

1. Damodar Gujarati and Dawn Porter (2009): Basic Econometrics, McGraw Hill.
2. Johnston, J. (1984): Econometric methods, 3/e, McGraw Hill.
3. Nachane. D.M. (2006): Econometrics: Theoretical Foundations and Empirical Perspective, Oxford University Press.

Books for Reference:

1. Apte, P.G. (1990): Text book of Econometrics. Tata McGraw Hill.
2. Intrulligator, M.D. (1980): Econometric models - Techniques and Applications, Prentice Hall of India.
3. Kleiber, C. and Zeileis, A. (2008): Applied Econometrics with R, Springer, NY.
4. A. Koutsoyiannis (2001): Theory of Econometrics, 2/e, Palgrave Macmillan Ltd.

STAT 428 - ACTUARIAL STATISTICS CREDITS: 3

Unit I

Basic deterministic model: Cash flows, discount function, interest and discount rates, balances and reserves, internal rate of return, The life table: Basic definitions, probabilities, construction of life tables, life expectancy, Life annuities: Introduction, calculating annuity premium, interest and survivorship discount function, guaranteed payments, deferred annuities.

Unit II

Life insurance: Introduction, calculation of life insurance premiums, types of life insurance, combined benefits, insurances viewed as annuities, Insurance and annuity reserves: The general pattern reserves, recursion, detailed analysis of an insurance, bases for reserves, non forfeiture values, policies involving a return of the reserve, premium difference and paid-up formula.

Unit III

Fractional durations: Life annuities paid monthly, immediate annuities, fractional period premium and reserves, reserves at fractional durations, Continuous payments: Continuous annuities, force of discount, force of mortality, Insurance payable at the moment of death, premiums and reserves. The general insurance – annuity identity, Select morality: Select an ultimate tables, Changed in formulas.

Unit IV

Multiple life contracts: Joint life status, joint annuities and insurances, last survivor annuities and insurances, moment of death insurances. The general two life annuity and insurance contracts, contingent insurances

Unit V

Multiple decrement theory: Basic model, insurances, Determination of the models from the forces of decrement. Stochastic approach to insurance and annuities; Stochastic approach to insurance and annuity benefits, deferred contracts, Stochastic approach to reserves and premiums, variance formula.

Books for Study

1. Promislow, S.D(2006): Fundamentals of Actuarial Mathematics, John Willey

Books for Reference

1. Neill, A. (1977): Life contingencies, Heinemann, London.
2. Newton L. Bowers, Jr, Hans U. Gerber, James C. Hickmann, Donald A. Jones and Cecil J. Nesbitt (1997): Actuarial Mathematics, The Society of Actuaries.
3. King, G. Institute of Actuaries Text Book. Part 11, Second edition, Charles and Edwin Layton, London.

SEMESTER III

STAT 535 – RELIABILITY THEORY **CREDITS: 3**

Unit I

Introduction to Reliability and its needs; Structural properties of coherent system: components and systems, coherent structures, representation of coherent systems in terms of paths and cuts, relevant & irrelevant structure; Modules of coherent systems; Reliability of a coherent systems; Reliability importance of components; Bounds on System Reliability.

Unit II

Life Distributions: Concept of distribution function, hazard function, Reliability function, MTTF, Bathtub failure rate; loss of memory property of Exponential distribution - parametric families of some common life distributions – Exponential, Weibull and Gamma and its characterization - Reliability estimation of parameters in these models.

Unit III

Notions of Ageing; Classes of life distributions and their duals - preservation of life distribution classes for reliability operation - Formation of coherent systems, convolutions and mixtures.

Unit IV

Univariate stock models and life distributions arising out of them: cumulative damage model, shock models leading to univariate IFR, Successive shock model; bivariate shock models; common bivariate exponential distributions due to shock and their properties.
Maintenance and replacement policies; availability of reparable systems; modeling of a repairable system by a non-homogeneous Poisson process.

Unit V

Stress-Strength reliability - Concepts and its estimation for exponential, weibull and gamma distributions; Reliability growth models; probability plotting techniques; Hollander – Proschan and Deshpande tests for exponentiality – Basic ideas of accelerated life testing.

Books for Study:

1. Barlow, R.E. and Proschan F. (1985) Statistical Theory of Reliability and Life Testing; Rinehart and Winston.
2. Lawless, J.F. (2003): Statistical Models and Methods of Life Time Data; John Wiley.

Books for Reference:

1. Bain L.J. and Max Engelhardt (1991): Statistical Analysis of Reliability and Life Testing Models; Marcel Dekker.
2. Nelson, W (1982): Applied Life Data Analysis; John Wiley.
3. Zacks, S(1992): Introdcution to Reliability Analysis, Springer Verlag.
4. Marshall, A.W. and Olkin I(2007): Life Distributions, Springer

STAT 536 - TIME SERIES ANALYSIS **CREDITS: 3**

Unit I

Exploratory Time Series Analysis: Forecasting trend and seasonalitybased on smoothing. Methods of Exponential and moving average smoothing; Types and implications of interventions; Outliers, additive and innovational outliers, procedure for detecting outliers

Unit II

Stationary Stochastic models: weak and strong stationarity, Deseasonalising and detrending an observed time series, Auto-covariance, autocorrelation function (ACF), partial autocorrelation function (PACF) and their properties, Conditions for stationarity and invertibility,

Unit III

Models for Time Series: Time series data, Trend, seasonality, cycles and residuals, Stationary, White noise processes, Autoregressive (AR), Moving Average (MA), Autoregressive and Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) processes, Choice of AR and MA periods

Unit IV

Spectral analysis and decomposition: Spectral analysis of weakly stationary process, Periodogram and correlogram analysis, Spectral decomposition of weakly AR process and representation as a one-sided MA process – necessary and sufficient conditions, implication in prediction problems.

Unit V

Modeling Seasonal Time Series: seasonal ARIMA models, estimation and forecasting, Fitting ARIMA models with Box-Jenkins procedure, Identification, Estimation, Verification, Test for white noise, Forecasting with ARMA models.

Books for Study:

1. Nicholas T. Thomopoulos, 1980, Applied Forecasting Methods, Prentice Hall
2. BoxGEP,JenkinsGMandReinselGC(2004):TimeSeriesAnalysis–Forecasting and Control, Pearson Education.
3. BrockwellPJandDavisRA(2002):IntroductiontoTimeSeriesandForecasting, Springer.
4. Montgomery D C and Johnson L A (1977): Forecasting and Time Series analysis, McGraw Hill.

Books for Reference:

1. ChatfieldC(1996):TheAnalysisofTimeSeries:TheoryandPractice,fifthedition, Chapman and Hall.
2. NachaneD.M.(2006):Econometrics:TheoreticalFoundationsandEmpiricalPerspective, Oxford University Press
3. Diggle, P.J Time Series: A Biostatistical Introduction, Oxford University Press (1990).
4. Hamilton, J., 1994, Times Series Analysis, Princeton University Press.
5. Harvey, A.C., 1993, Time Series Models, MIT Press.
6. Kendall, Sir Maurice and Ord J K (1990): Time Series, Edward Arnold.
7. Tsay, R., 2002, Analysis of Financial Time Series, Wiley Series

STAT 537 - ELEMENTS OF QUEUEING THEORY CREDITS: 3

Unit I

Queuing Models: Basic characteristics of a Queuing Model – Role of Poisson and Exponential distributions, Stochastic Processes, Markov chains, Poisson Processes, Generalized Birth and Death Processes, steady state Birth and death processes.

Unit II

Poisson Queuing Models with single server: Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/M/1): (∞/FIFO) and (M/M/1): (N/FIFO) Models, simple numerical problems

Unit III

Poisson Queuing Models with multiple server: Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/M/C): (∞/FIFO), (M/M/C): (N/FIFO) and (M/M/C): (C/FIFO) Models, simple numerical problems

Unit IV

Non Poison Queuing Models (Erlangian): Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/Ek/1), (Ek/M/1), simple numerical problems

Unit V

General Queuing Models: Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/G/1), (G/M/1), Simple numerical problems

Books for Study

1. Donald Gross & Carl M Harris (1998): Fundamentals of Queuing theory, John Wiley & Sons, Inc.
2. Hamdy A.Taha(2006): Operations Research – An Introduction, 8/e , Prentice Hall of India Private Ltd., New Delhi

Books for Reference

1. S.D.Sharma (2003)Operations Research , Kedar Nath Ram Nath & Co, Meerut, India
2. Kanthi Swarup, P.K.Gupta and Man Mohan (2004), Operations Research, Sultan Chand & Sons, New Delhi
3. Prabhu N.U. (1965) Applied Stochastic Processes, Mc.Millan
4. J.Medhi (2009), Stochastic Processes, 3/e, New Age International
5. Bhat B.R.(2004): Stochastic Models: Analysis and Applications, New Age Publications

STAT 538 – SIMULATION TECHNIQUES **CREDITS: 3**

Unit I

Introduction, need of simulation, physical versus digital simulation, Bufferies needle problem, Use of simulation in defense, inventory problems and other fields.

Unit II

Random number generation: Congruential generators, Metropolis Hasting algorithm, Statistical tests for pseudo random numbers. Random number generation from mixture of distributions, compound distributions.

Unit III

Random Variate Generation Methods: Inverse transform method, acceptance rejection method, Composition method. Generating random variables from continuous and discrete uni variate distributions. Generation of random vectors from multivariate distributions.

Unit IV

Monte Carlo integration and variance reduction techniques. Hit or Miss Monte Carlo Method, Sample mean Monte Carlo method – MCMC.

Unit V

Re-sampling techniques: Bootstrap, Jackknife techniques and their applications. Machine Learning: Supervised and unsupervised learning techniques.

Books for Study

1. Rubinstein, R.Y. and D P Kroese (2007): Simulation and Monte Carlo Method, 2/e, John Wiley and Sons.
2. Gorden, G. (2001): System Simulation, Prentice Hall of India, New Delhi.
3. Fishman, G.S. (1996): Monte Carlo Concepts, Algorithms and Applications, Springer Verlag, New York.

Books for Reference

1. Deo, Narshingh (1991): System simulation with digital computer, Prentice Hall of India Pvt. Ltd, New Delhi.
2. Lewis, P.A.W. and Orav, E.J. (1988): Simulation Methodology for Statisticians, Operations Analysis and Engineering, Wadsworth and Brooks Cole Advanced Books and Software, Vol. 1.
3. Sethi, I.K. and Jain, A.K. (1991): Artificial Neural Networks and Statistical pattern recognition, North-Holland, Amsterdom.
4. McLachlan, G. and Krishnan, T. (2007): The EM Algorithm and Extensions, John Wiley and Sons.

STAT 544 - SURVIVAL ANALYSIS **CREDITS: 3**

Unit I

Concepts of time, Order and random Censoring, likelihood in these cases. Life distributions- Exponential, Gamma, Weibull , Lognormal , Pareto , Linear Failure rate. Parametric inference (Point estimation, scores, MLE)

Unit II

Life tables, failure rate, mean residual life and their elementary properties. Concept of Ageing, Types of Ageing classes and their properties and relationship between them , Bathtub Failure rate, Concept of Inverse Hazard rate.

Unit III

Estimation of survival function Actuarial Estimator, Kaplan- Meier Estimator, Estimation under the assumption of IFR / DFR . Tests of exponentiality against non- parametric classes- Total time on test, Deshpande test.

Unit IV

Two sample problem- Gehan test, Log rank test. Mantel Haenszel test, Tarone Ware tests. Introduction to Semi- parametric regression for failure rate, Cox's proportional hazards(PH) model with one and several convariates and estimation problems in Cox's PH Model. Rank test for the regression coefficients.

Unit V

Introduction to Competing risks analysis and estimation problems in competing risk model for parametric and non- parametric semi parametric set up. Ideas of Multiple decrement life table and its applications.

Books for Study:

1. Miller, R.G. (1981) : Survival analysis (John Wiley).
2. Cox, D.R. and Oakes, D. (1984) : Analysis of Survival Data, Chapman and Hall, NewYork.
3. Elisha T Lee, John Wenyu Wang and Timothy Wenyu Patt(2003): Statistical Methods for Survival data Analysis, 3/e, Wiley Inter Science.

Books for Reference:

1. Gross, A.J. and Clark, V.A. (1975) : Survival distribution : Reliability applications in the Biomedical Sciences, John Wiley and Sons.
2. Elandt Johnson, R.E. Johnson N.L.: Survival Models and Data Analysis, John Wiley and sons.
3. Kalbfleisch J.D. and Prentice R.L.(1980), The Statistical Analysis of Failure Time Data, JohnWiley.
4. Klelin P. John and Moeschberger(2003): Survival Analysis: Techniques for Censored andTruncated Data, 2/e, Springer.
5. Lawless J.F. (1982) Statistical Models and Methods of Life Time Data; John Wiley & Sons.

STAT 545 - BIOSTATISTICS **CREDITS: 3**

Unit I

Statistical Methods in Clinical Trials: Introduction to clinical trial and it's phases I, II, III and IV, statistical designs-fixed sample trials: simple randomized design, stratified randomized crossover design; Sequential design - open and close sequential design. Randomization-Dynamic randomization, Permuted block randomization; Blinding-Single, double and triple.

Unit II

Biological Assays: Introduction, parallel-line assay, slope- ratio assays and quantal- response assay, Feller's theorem. Dose-response relationships-qualitative and quantitative response, dose response relation- estimation of median effective dose – PK-PD Analysis.

Unit III

Categorical Data Analysis: Categorical response data, logistic regression-odds ratio, Wald's statistic, logistic regression and its diagnostics, - Poison regression – Estimation of relative risk and its applications.

Unit IV

ROC Curve analysis - Estimation of Binormal Model and the Area under the Curve, its applications – Properties of ROC curve - Kullback –Leibler Divergence (KLD)– definition – functional relationship between Kullback –Leibler Divergence and the slope of the ROC curve – derivations of KLD expressions for Bi-normal ROC model

Unit V

Repeated Measures ANOVA – One Way and Two Classified Data –Measures of disease frequency – incidence – prevalence – relative risk – Epidemiological study designs – Cohort study design and its analysis – Case control study design and its analysis – concept of bias – information bias and selection bias

Books for Study

1. Elisa T.Lee & John Wenyu Wang (2003): Statistical methods for Survival Data analysis, 3rd Edition, John Wiley
2. Jerrold H. Zar (1999): Biostatistical Analysis, 4th edition, Pearson
3. Armitage, P, Berry G and Mathews J.N.S (2002): Statistical Methods in Medical Research, 4/e, Blackwell Scientific Publications
4. Krzanowski, W and Hand, D.J.(2009): ROC Curves for Continuous Data, Chapman and Hall

Books for Reference

1. Hosmer and Lemeshow (2000): "Applied Logistic Regression", 2/e, Wiley Series
2. Alan Agresti (2002): Categorical Data analysis, 2/e, John Wiley
3. Sylvia Wasserthial and Smoller, (2004): Biostatistics and Epidemiology – A Primer for Health and Biomedical professionals, 3rd Edition, Springer
4. Rastogi, V.B. (2006): Fundamentals of Biostatistics, ANE Books, India

STAT 546 - STATISTICAL DATA MINING METHODS CREDITS: 3

Unit I

Introduction to data mining – data types – Measures of similarity and dissimilarity – Data mining tools – supervised and unsupervised learning – Introduction to Cluster Analysis – Types of clustering – Agglomerative Hierarchical clustering algorithm – Issues – strength and weaknesses.

Unit II

Basic k-means algorithm – Issues – Bisecting k-means – fuzzy clustering – fuzzy c means algorithm - cluster evaluation – unsupervised and supervised measures - Introduction to classification – Decision Trees – Building a decision tree – Tree induction algorithm – Splitting of nodes based on information gain and Gini index – model over fitting – Evaluating the performance of a classifier

Unit III

Nearest Neighbor classifiers – kNN algorithm – Naïve Bayesian classifier – Binary logistic regression – odds ratio – Interpreting logistic regression coefficients – Multiple logistic regression

Unit IV

Association rules mining – Basics – Apriori algorithm – Pruning and candidate generation – Rule mining.

Unit V

Case studies based on k means clustering - fuzzy c means clustering - kNN classification - Binary logistic regression using R programming language.

Books for Study

1. Tan, T., Steinbach, M. and Kumar, V. (2006): Introduction to Data Mining, Pearson Education.
2. Gupta, G.K. (2008): Introduction to Data Mining with case studies, Prentice – Hall of India Pvt. Ltd.
3. Daniel T. Larose (2006): Data Mining: Methods and Models, John Wiley and sons.

Books for Reference

1. Han, J. and Kamber, M. (2006): Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann Publishers.
2. Paolo Gludici (2003): Applied Data Mining: Statistical Methods for Business and Industry, John Wiley and sons.
3. Rajan Chattamvelli (2009): Data Mining Methods, Narosa Publishing House, New Delhi.

STAT 547 -BAYESIAN INFERENCE **CREDITS: 3**

Unit I

Introduction about Thomas Baye's-Motivations and Contributions - Evaluation of Subjective probability of an event using a subjectively unbiased coin - Subjective prior distribution of a parameter – Baye's theorem and computation of the posterior distribution.

Unit II

Introduction of Prior Distributions, Types of Prior Distributions, Proper Prior-Enlarging the natural conjugate family by enlarging hyper parameter space - mixtures from conjugate family - choosing an appropriate member of conjugate prior family - Non informative, improper and invariant priors - Jeffrey's invariant prior

Unit III

Bayesian point estimation: Prediction problem from posterior distribution - Bayes estimators for absolute error loss, squared error loss, linex loss function, Jeffrey's and 0 -1 loss - Generalization to convex loss functions - Evaluation of the estimate in terms of the posterior risk

Unit IV

Bayesian interval estimation : Credible intervals - Highest posterior density regions - Interpretation of the confidence coefficient of an interval.

Unit V

Bayesian Testing of Hypothesis: Prior and Posterior odds - Bayes factor for various types of testing hypothesis problems -Monte-Carlo Integration and Basic Concepts on Markov chain Monte Carlo techniques (MCMC)(without proof) – Gibb's sampling.

Books for Study

1. Bansal A.K.(2007): Bayesian Parametric Inference, Narosa Publications
2. Berger, J.O.(1985): Statistical Decision Theory and Bayesian Analysis, 2/e, Springer Verlag.
3. Sinha S K (1998): Bayesian Estimation, New Age International(P) Ltd, New Delhi

Books for Reference

1. Robert C.P. and Casella, G.(2004): Monte Carlo Statistical Methods, 2/e,Springer Verlag.
2. DeGroot, M.H.(2004): Optimal Statistical Decisions, Wiley-InterScience.
3. Gamerman, D. and Lobes H.F. (2000): Stochastic Simulation for Bayesian Inference, Taylor and Francis.
4. Box, G.P. and Tiao, G.C.(1973): Bayesian Inference in Statistical Analysis, Addison – Wesley.

SOFT CORE COURSE FOR OTHER DEPARTMENTS

STAT 418 - STATISTICAL METHODS **Credits: 3**

Unit I

Definition of statistics – Scope and limitations of statistics – Primary and Secondary data and its sources - Simple Random, Stratified and Systematic sampling techniques - preparation of a questionnaire -Collection and classification of data – Frequency tables – Diagrammatic and Graphical representation of data

Unit II

Measures of central tendency – Mean, Median and Mode – Measures of dispersion – Range, Quartile deviation and Standard deviation – Coefficient of variation and skewness

Unit III

Study of relationship between variables: Quantitative: Correlation and Regression – Partial and Multiple correlation (three variables only) – Qualitative: Contingency tables – Measures of Association.

Unit IV

Elementary Probability theory:Addition theorem – Conditional probability and Multiplication theorem - Bayes' Theorem – Random variables and probability distributions – Binomial, Poisson , Normal (simple applications of the distribution) – Sampling distributions: t, F and chi-square (definition only)

Unit V

Hypothesis testing:Basic concepts in Hypothesis Testing – Types of error – Tests for Mean and Proportion based on Normal and Student t-distribution - Chi-square test for independence of attributes – One-way and two-way Analysis of Variance

Books for Study

1. Hooda.R.P.(2003) : Statistics for Business and Economics , 3/e, Mac Millan .
2. Medhi.J. (1992) : Statistical Methods an Introductory Text , Wiley Eastern Ltd.,.
3. Kapoor.V.K. and Gupta.S. (1978): Fundamentals of Applied Statistics,Sultan Chand and Sons.
4. Sharma J.K.(2004): Business Statistics, Pearson Education

Books for Reference

1. Agarwal.B.L(1996): Basic statistics , 3/e, New Age International (P) Ltd.,.
2. Anderson.R, Sweeney.J and Williams.A (2002): Statistics for Business and Economics, 8/e, Thomson.
3. Sheldon M.Ross (2006): Introductory Statistics, 2/e, Elsevier Publications.
4. Murray R. Spiegel and Larry J. Stephens (2005): Schaum's Outline of Theory and Problems of Statistics, 3/e, Tata Mc Graw Hill Publishing Company Ltd, New Delhi.