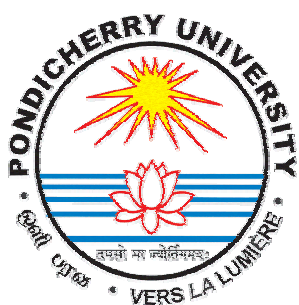


PONDICHERRY UNIVERSITY

DEPARTMENT OF STATISTICS



**SYLLABUS FOR M.Sc - STATISTICS
(CBCS – Five Year Integrated)
Effective from the Academic Year 2009-2010**

PONDICHERY UNIVERSITY
CHOICE BASED CREDIT SYSTEM
SYLLABUS FOR THE M.Sc. INTEGRATED STATISTICS COURSE

Effective from the Academic Year 2009 – 2010

COURSE CODE	TITLE OF THE COURSE	NATURE OF THE COURSE	NO. OF CREDITS
I SEM			
STAT 111	Basic Statistics	Hard Core	3
II SEM			
STAT 121	Basic Probability Theory	Hard Core	3
III SEM			
STAT 231	Probability Distributions	Hard Core	3
STAT 232	Sampling Theory	Hard Core	3
IV SEM			
STAT 241	Basic Estimation Theory	Hard Core	3
STAT 242	Practical - I	Hard Core	2
V SEM			
STAT 351	Elements of Testing Statistical Hypotheses	Hard Core	3
STAT 352	Statistical Quality Control and Operations Research	Hard Core	3
STAT 353	Practical - II	Hard Core	2
VI SEM			
STAT 361	Principles of Experimental Design	Hard Core	3
STAT 362	Applied Statistics	Hard Core	3
STAT 363	Practical - III	Hard Core	2
VII SEM			
STAT 471	Real Analysis and Linear Algebra	Hard Core	4
STAT 472	Advanced Sampling Theory	Hard Core	4
STAT 473	Probability - I	Hard Core	4
STAT 474	Advanced Distribution Theory	Hard Core	4
STAT 475	Practical – IV	Hard Core	2
VIII SEM			
STAT 481	Probability - II	Hard Core	4
STAT 482	Quality Control and Operations Research	Hard Core	4
STAT 483	Elements of Stochastic Processes	Hard Core	4
STAT 484	Statistical Inference - I	Hard Core	4
STAT 485	Practical – V	Hard Core	2
IX SEM			
STAT 591	Multivariate Analysis	Hard Core	4
STAT 592	Statistical Inference - II	Hard Core	4
STAT 593	Linear Models and Regression Theory	Hard Core	4
	Elective – I	Soft Core	4
STAT 594	Practical – VI	Hard Core	2
X SEM			
STAT 595	Design of Experiments	Hard Core	4
	Elective – II	Soft Core	4
	Elective – III	Soft Core	4
STAT 596	Practical - VII	Hard Core	2
STAT 597	Project and Viva-Voce	Hard Core	4

Electives:

STAT 601 Time Series Analysis	STAT 606 Econometrics
STAT 602 Advanced Operations Research	STAT 607 Actuarial Statistics
STAT 603 Reliability	STAT 608 Data Base Management Systems
STAT 604 Survival Analysis	STAT 609 Bayesian Inference
STAT 605 Data Mining Methods	STAT 610 Bio–Statistics

Unit I

Convergence of infinite numerical sequences and series (review only) – Absolute and conditional convergence – Sequences and series of functions – Pointwise and Uniform convergence – Tests for Uniform convergence – Properties of Uniform convergence

Unit II

Riemann - Stieltjes integral: Definition and properties – Integrals with step function and monotonic functions as integrators and their properties – Mean value theorem, Taylor's theorem – Evaluation of Riemann - Stieltjes integral – Fundamental theorem

Unit III

Functions of several variables : Limits and continuity – Partial derivatives and Differentiability - Properties of differentiable functions – Higher order derivatives and differentials – Young and Schwartz theorems – Taylor's theorem - Maxima and Minima – Extrema under constraints

Unit IV

Vector space and sub-space – Linear independence and orthogonality – Dimension and basis of a vector space – Orthonormal basis – Gram-Schmidt orthogonalization – Inner product space – Simultaneous linear equations (homogeneous and non-homogeneous)

Unit V

Matrices: Rank, inverse, trace and their properties – Characteristic roots and vectors – Idempotent and partitioned matrices – G-inverse and Moore Penrose inverse - their properties – Reduction of a matrix into diagonal, echelon, canonical and triangular forms – Quadratic forms – reductions of different types – Definite quadratic forms – Cochran's theorem

Books for Study:

1. Rudin.W(1976): Principles of Mathematical Analysis, Mc Graw Hill
2. Malik .S.C. and Arora(1987): Mathematical Analysis , Wiley Eastern Ltd.,
3. Datta, K.E (1991): Matrix and Linear Algebra, Prentice-Hall of India Private Ltd.,
4. Rao, C.R (1973): Linear Statistical Inference and its Applications, Wiley Eastern Ltd.,
5. Searle, S.R (1982): Matrix Algebra useful for Statistics, John Wiley, NY.

Books for Reference:

1. Goldberg .R.R. (1970) : Methods of Real Analysis , Oxford and IBH Publishing Co.(P) Ltd., New Delhi.
2. Shanti Narayan (1993): Mathematical Analysis , Sultan Chand and Co.
3. Gilbert and Gilbert (2005): Linear Algebra and Matrix Theory, Elsevier Publications.
4. Graybill, F.A.(1983): Matrices and applications in statistics, Wadsworth Publishing Company, Belmont, California, USA.
5. Ramachandra Rao, A.and Bhimasankaran, P.(1992): Linear Algebra , TMH.

Unit I

Sampling and its need – Review of SRSWR and SRSWOR – Stratified Random sampling and systematic sampling – estimation of population mean and variance .

Unit II

Ratio estimates and their properties for Simple Random and Stratified Random sampling – Ratio estimator and Multivariate Ratio estimator - Regression Estimators – Regression estimates with pre assigned “b” – sample estimate of variance – Bias – Regression estimators in Stratified Sampling - Multivariate Regression Estimator.

Unit III

Varying probability Sampling: Cumulative total method and Lahiri’s method - Estimation in pps sampling with replacement, pps sampling without replacement; General selection procedures, Narian’s Scheme of sample selection and Sen-Midzuna method - Ordered estimator; Des Raj Unordered estimators: Horvitz – Thompson estimator and Murthy’s estimator.

Unit IV

Cluster Sampling: Equal cluster sampling – Estimators of mean and variance, optimum cluster size, Unequal cluster sampling – Estimators of mean and variance, varying probability cluster sampling - Two stage sampling – variance of the estimated mean - Three stage sampling – variance of the estimated mean.

Unit V

Multiphase sampling: Double sampling for stratification - Optimum allocation - Estimated variance in Double sampling for stratification. Sources of errors in Surveys - Mathematical model for the effects of call-backs and the errors of measurement – Interpenetrating sub sampling method.

Books for Study:

1. Cochran, W.G(1977): Sampling Techniques, Wiley Eastern Ltd,. (Chapter 6 for Unit I, Chapter 7 for Unit II and Chapter 13 for Unit V)
2. Singh, D and Choudhary, F.S(1986): Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd,. (Chapter 5 for Unit III and Chapter 8 for Unit IV)

Books for Reference:

1. Desraj and Chandok (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. Sharon L Lohr (1999): Sampling : Design and Analysis, Duxbary Press
5. Sukhatme et al (1984): Sampling Theory of Surveys with applications, Iowa State University Press and IARS.

Unit I

Algebra of sets – Fields – Sigma fields – Inverse function – Measurable function – Random Variables – Induced Sigma fields – Limits of Random variables.

Unit II

Probability – Definition – Simple properties – Discrete Probability Space – General Probability Space – Induced Probability Space – Conditional Probability - Distribution Function of a Random Variable – Decomposition of distribution functions – Distribution function of random vectors.

Unit III

Expectation and moments – Definition and properties – Moment Generating Function – Moment Inequalities – C_r , Holder, Jensen and basic inequalities.

Unit IV

Convergence: Modes of convergence – Convergence in probability – Convergence in distribution – Convergence in r^{th} mean – Almost sure convergence and their interrelationships.

Unit V

Convergence theorem for expectation: Monotone Convergence theorem – Fatou's theorem – Dominated Convergence theorem - Definition of product space – Fubini's Theorem (statement only) - Independence: Definition – Multiplication properties – Zero-one law.

Books for Study:

1. Bhat .B.R. (1999) : Modern Probability theory , 3/e, New Age International (P) Ltd., (Relevant sections of chapters 1,2,3,4,5,6 and 9)

Books for Reference:

1. Ash, R.B (1972): Real Analysis and Probability, Academic Press.
2. Burriel,C.W (1972): Measure, Integration and. Probability, Mc Graw Hill International.
3. Chow, Y.S and Teicher, H (1979): Probability Theory, Springer, Narosa.
4. Loeve, M (1985):Probability Theory, 3/e, Von Nostrand.

Unit I

Brief review of distribution theory, functions of random variables and their distributions - Laplace, Cauchy, lognormal, logarithmic series and power series distributions - Multinomial distribution

Unit II

Bivariate Binomial – Bivariate Poisson - Bivariate Exponential - Compound, truncated and mixture of distributions, concepts of convolution

Unit III

Multivariate normal distribution – marginal and conditional distributions – characteristic function and other properties - Sampling distributions: Non-central chi-square, t and F distributions and their properties

Unit IV

Distributions of quadratic forms under normality-independence of quadratic forms and linear form- Cochran's theorem - 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 distributions, exponential, Weibull and extreme value distributions, Mills ratio, distributions classified by hazard rate.

Books for Study:

1. M. Mood, F.A.Graybill and D.C.Boes(2001) : Introduction to the Theory of Statistics, Tata McGraw-Hill, New Delhi.
2. Johnson, N.L &Kotz.S (1969): Discrete Distributions, Houghton Mifflin, Boston MA
3. Johnson, N.L & Kotz.S (1970): Continuous Distributions, Houghton Mifflin, Boston MA

Books for Reference:

1. Bhattacharya and Johnson(1977): Statistical Concepts and Methods, John Wiley.
2. Pitman J. (1993): Probability Distributions, Narosa Publishing House.
3. Rao C. R.,(1973): Linear Statistical Inference and its Applications, Wiley Eastern Ltd, New Delhi.
4. Dudewicz, E.J and Mishra, S.N(1980): Mathematical Statistics, John Wiley, NY.

I 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 sampling – SRS, PPSWR, PPSWOR
5. Cluster sampling – of equal sizes.
6. Ratio, Regression and Difference estimation.

II. Computations based on SYSTAT software (40 marks)

1. Random number generation. (i) Binomial, (ii) Poisson, (iii) Normal
2. Descriptive Statistics
3. Computation of
 - (i) Simple and multiple correlation coefficients
 - (ii) Simple and multiple regression coefficients
4. Statistical Tests
 - (i) Z test and Confidence Interval
 - (ii) One sample t-test and Confidence Interval
 - (iii) Two sample t-test and Confidence Interval
 - (iv) Chi-square test for independence of attributes
 - (v) Tests for homogeneity of proportions, variances and correlation coefficients
5. Non – Parametric Tests
 - (i) Run test
 - (ii) Sign test
 - (iii) Wilcoxon one sample test
6. Design of Experiments
 - (i) One-way ANOVA
 - (ii) Two-way ANOVA
 - (iii) Analysis of Covariance (ANCOVA)

Unit I

Characteristic function - Definition and properties – Inversion formula and its application – Characteristic Function and Moments – Bochner's theorem (statement only) – Simple problems

Unit II

Weak and complete convergence of distribution functions – Helly's First and Second limit theorems

Unit III

Law of large numbers: Kolmogorov Inequality – Weak law of large numbers (Khinchin's and Kolmogorov) - Kolmogorov Strong law of large numbers – Glivenko-Cantelli Theorem (statement only)

Unit IV

Central Limit Theorem : iid case – Lindeberg-Levy and Liapounov's form - Lindeberg - Feller form – Infinitely Divisible distributions – definition, elementary properties and examples – canonical representation (without proof)

Unit V

Conditioning: Radon Nikodym theorem and derivative (without proof) - Conditional expectation – definition – properties (probability and expectation properties) - conditional probability and its applications – Definition and properties of Martingales and Sub-martingales – Martingale convergence theorem

Books for Study:

1. Bhat .B.R. (1999) : Modern Probability theory , 3/e , New Age International (P) Ltd., (Relevant sections of Chapters 7,8, 10, 11, 12)

Books for Reference:

1. Ash, R.B (1972) : Real Analysis and Probability, Academic Press .
2. Billingsley . P.(1979): Probability and Measure , Wiley .
3. Kingman and Taylor (1966) : Probability Theory , Narosa.
4. Tucker. H.G. (1967) : A Graduate course in probability , Academic Press .
5. Loeve. M.(1985) : Probability theory , 3/e, Von Nostrand .
6. Burrill, C.W. (1972): Measure, Integration and Probability, Mc Graw Hill .

Unit I

Quality improvement: Meaning of quality and quality improvement – Different types of Quality costs and their management

Control charts: Review of \bar{X} , R, p, c, d charts - Modified control charts for mean – CUSUM chart – technique of V-mask – Weighted Moving average charts – Stopping control charts and group control charts

Unit II

Process Capability analysis: Meaning, Estimation technique for capability of a process – Capability Indices: C_p , capability ratio and C_{pk} index – Estimation of natural tolerance limit of a process

Acceptance Sampling plans for attributes: Single, double, multiple and continuous sampling plans for attributes (Dodge type)

Unit III

Acceptance Sampling plans for variables: one sided and two sided specification – Standardized plans (ANSI/ANSQ Z1.9) and MIL-STD-414

Reliability: Concept, Definition and need - Concepts of Hazard rate, IFR and DFR - Relevance of exponential distribution in Reliability – Failure models – Taguchi's approach in Quality and reliability - Six sigma approach

Unit IV

Review of LPP – Simplex and revised simplex methods - Duality in LPP – Dual Simplex method – Some important theorems on duality - Sensitivity Analysis – Variation in cost vector 'c' – Variation in the requirement vector 'b' – Addition and deletion of single variable – Addition and deletion of single constraint

Unit V

Replacement problem – Replacement of policy when value of money change/does not change with time – Replacement of equipment that fails suddenly – Group replacement – Inventory – Various types of inventory – Costs associated with inventory models – Deterministic inventory models with and without shortages

Books for Study:

1. Douglas C. Montgomery(1985): Introduction to Statistical Quality Control, John Wiley and Sons, New York.
2. Hamdy A. Taha (1995): Operations Research – An Introduction, 5/e, Prentice Hall of India Private Ltd, New Delhi.
3. Sinha S M (2006): Mathematical Programming : Theory and Methods, Elsevier Publications.
4. Mittage, H.J and Rinne, H(1993): Statistical Methods of Quality Assurance, Chapman Hall, London, UK
5. Balagurusamy, B(1995): Reliability Engineering, Tata Mc Graw Hill Book Company, New Delhi.

Books for Reference:

1. Lewis, E.E(1995): Introduction to Reliability Engineering, John Wiley and Sons Inc, New York.
2. Mahajan,M(1998): Statistical Quality Control, Dhanpat Rai & Co Private Ltd., New Delhi.
3. Zacks, Z(1995): Reliability Analysis, Springer Verlag.
4. Gupta,H.D (1984): Quality assurance through ISO 9000, South Asia Publication, New Delhi
5. Smith, G.M(1991): Statistical Process Control and Quality Improvement, 3/e, Printice Hall, NY

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 process and its 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 (1988): Stochastic Processes, 3/e, New age Publication.
3. Sidney I. Resnick(1992):Adventures in Stochastic Processes, Birkhauser, Boston.

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, HallLondon,Third Edition
5. Prabu N.U. (1965): Stochastic Processes Macmillan.
6. Ross S.M (1983): Stochastic Process Wiley.

Unit I

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

Unit II

The information measure – Cramer - Rao (CR) inequality - Kiefer – Chapman - Robbins (KCR)
inequality - Bhattacharya inequality - 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- 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 (minimum
average width) - Constructions of shortest confidence intervals.
Notion of Bayes estimation – Concepts of prior, posterior and conjugate priors. Simple
problems involving quadratic error loss function - Elementary notions of minimax estimation -
Simple illustrations.

Books for Study:

1. Rohatgi, V.K(1986): Statistical Inference, Wiley Eastern Ltd,.
2. Kale .B.K. (1999): A First course on parametric inference , Narosa Publishing House.
3. Zacks,S (1981): Parametric Statistical Inference, John Wiley, NY.

Books for Reference:

1. Goon, A.M, Gupta,M.K, and Das Gupta, B.C(1980) : An outline of Statistical Theory, Vol. II,The World Press, Calcutta.
2. Lehmann, E.L(1983) : Theory of Point Estimation, Wiley Eastern Ltd, 1983.
3. Mood, A.M., Graybill, F.A and Boers, D.C(1974) : Introduction to Theory of Statistics, Mc Graw-Hill Book Company.
4. Rao, C.R(1998): Linear Statistical Inference and its Applications, Wiley Eastern Ltd,.
5. Casella, G and Berger, R.L(2002):Statistical Inference , Duxubury Process, Belmont, USA.

I Estimation

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 Statistical Quality Control

Control charts:

1. CUSUM chart
2. Modified Control chart
3. Moving Average Control chart
4. Exponentially Weighted Moving Average chart
5. Sloping Control Chart

Acceptance sampling:

1. Single sampling plans and double sampling plans (for attributes)
2. Variable Sampling plans (Single and double specifications)
3. Standard plans

Unit I

Maximum likelihood estimation of the parameters of Multivariate Normal and their sampling distributions – Inference concerning the mean vector when covariance matrix is known - Total , Partial, Multiple correlation in the Multivariate setup – MLEs of Total, Partial and Multiple correlation coefficients and their sampling distributions in the null case

Unit II

Hotelling T^2 distribution and its applications - derivation of generalized T^2 statistic and its distribution - Uses of T^2 statistic - optimum properties of T^2 statistic - Mahalanobis D^2 statistic and its distribution - relation between T^2 and D^2 – Test based on T^2 statistic

Unit III

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

Unit IV

Classification problems - Classification into one of two populations (known and unknown dispersion matrix) - Classification in to one of several populations – Linear discriminant function – Multivariate analysis of variance (MANOVA) – One- Way classification .

Unit V

Principal components - Definition- Maximum likelihood estimates of the principal components and their variances – Extraction of Principal components and their variances. Factor analysis - Mathematical model- Estimation of Factor Loadings – Canonical correlation – Estimation of canonical correlation and variates

Books for Study:

- 1.Anderson, T.W(1984) : An Introduction to Multivariate Statistical Analysis, Wiley Eastern Ltd.,.
- 2.Giri, N.C(1977): Multivariate Statistical Inference, Academic Press, NY

Books for Reference:

1. Morrison, F(1985): Multivariate Statistical Methods, Mc Graw Hill Book Company.
2. Rao, C.R(1965): Linear Statistical Inference and its Applications, Wiley Eastern Ltd.,
3. Johnson, Richard A and. Wichern D.W (2003): Applied Multivariate Statistical Analysis, Prentice-Hall of India Private Ltd., New Delhi.

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 regions and complete sufficient statistics, Tests with Neyman structure, Uniformly most powerful similar tests, 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 - k- sample problem: Extension of Median test, Kruskal Wallis test, Friedman test – Notion of ARE.

Unit V

Sequential methods: Sequential unbiased estimation – Application to Normal distribution - Sequential test - 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. Lehmann, E.L and Joseph P. Romano (2005): Testing Statistical Hypotheses, 3/e, Springer.
2. Rohatgi, V.K (1986): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern Ltd., New Delhi.
3. Kale .B.K. (1999): A First course on Parametric Inference, Narosa Publishing House.
4. Gibbons .J.D. (1985) : Non Parametric Statistical Inference, 2/e, Marcel Dekker.

Books for Reference:

1. Rao, C.R (1998): Linear Statistical Inference and its Application, John Wiley, Second Edition.
2. Casella, G & Berger, R.L (1990): Statistical Inference, Duxbury Press, Belmont, USA
3. Wald, A (1949): Sequential Analysis, John Wiley, NY
4. Ghosh, B.K (1970): Sequential Tests of Statistical Hypotheses, Addison Wesley.

Unit I

Full rank linear model – least square estimators of the parameters and their properties – 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

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

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. Draper, N and Smith, H(1981) : Applied Regression Analysis , Second Edition.
2. Alvin C. Rencher (2000): Linear Models in Statistics, John Wiley & Sons, New York (Chapters 7,8 & 9 for Unit I & II)

Books for Reference:

1. Chatterjee, S and Price, B (1971): Regression Analysis by Example, John Wiley.
2. Gunst, R.F and Mason, R.L (1980): Regression Analysis and Applications – A Data Oriented Approach, Marcel Dekker.
3. Guttman, I (1982): Linear Models - An Introduction , John Wiley.
4. Kshirsagar, A.M (1972): Course in Linear Models, Marcel Dekker.
5. Wetherill, G.B (1986): Regression Analysis.
6. Searle, S.R.(1971): Linear Models, John Wiley.

I Multivariate analysis

1. Maximum likelihood estimators – Mean vector and dispersion matrix
2. Test for Mean Vectors (Σ is known)
3. Hotelling's T^2 statistic
4. MANOVA (one way)
5. Test for covariance matrix
6. Discriminant analysis
7. Principal component analysis
8. Canonical correlation and canonical variables

II Testing of hypothesis

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, Median test for k-sample problem, Kruskal Wallis test and Friedman's test

III Linear Models and Regression Analysis

1. Multiple linear regression
2. Linear Models

Unit I

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

Unit II

Principles of scientific experimentation – Pen and Plot techniques - Basic Design: CRD, RBD and LSD, Analysis of RBD (with one observation per cell, More than one but equal number of observations per cell) – Multiple comparison test: LSD, SNK, DMR, Tukey tests.

Unit III

Factorial experiments: 2^n and 3^n experiments and their analysis – Complete and Partial Confounding - Fractional Replication in Factorial Experiments – Split plot and strip plot design and their analysis .

Unit IV

BIBD - Types of BIBD - Simple construction methods - Concept of connectedness and balancing – Intra Block analysis of BIBD – Recovery of InterBlock information – Partially Balanced Incomplete Block Design with two associate classes – intra block analysis only.

Unit V

Youden square and lattice design and their analysis – Mutually Orthogonal Latin Squares – Analysis of Covariance with one concomitant variable – Analysis for CRD and RBD only –Response Surface Designs – Method of Steepest Ascent

Books for Study:

1. Das, M.N. and Giri, N.C(1979): Design and Analysis of Experiments, Wiley Eastern Ltd, (Relevant Chapters for Units II, III, IV and V)
2. Douglas C. Montgomery (1984) : Design and Analysis of Experiments, John Wiley and Sons, (Chapter 16 for Parts of Unit IV and Unit V)
3. Graybill, F.A(1961) : An Introduction to Linear Statistical Models, Mc Graw Hill Book Company,(Chapter 5 & Parts of Chapter 6 for Unit I)
4. Pearce .S.C. (1984) : Design of Experiments , Wiley , New York .

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. Chakrabarthi .M.C. (1962) : Mathematics of design and analysis of experiments ,Asia Publication company.
7. Cobb G.W.(1998): Introduction to Design and Analysis of Experiments.

I. Design of Experiments (20 marks)

1. Multiple Comparison tests
2. 2^4 , 3^2 , 3^3 factorial experiment
3. Complete and partial confounding in 2^4 , 3^2 , 3^3 factorial experiments
4. Fractional factorial
5. Split plot design
6. Strip plot design
7. BIBD
8. PBIBD with two associate classes
9. Youden Square Design
10. Analysis of Covariance – CRD and RBD

II. Computations based on SPSS software (40 marks)

1. Statistical Quality Control and Regression Analysis
 - (i) Simple Linear Regression
 - (ii) Multiple regression
 - (iii) \bar{X} and R charts
 - (iv) p and np chart
 - (v) c chart
2. Multivariate Analysis
 - (i) Tests based on Hotelling T^2
 - (ii) MANOVA
 - (iii) Discriminant Analysis
 - (iv) Principal Component Analysis
 - (v) Factor Analysis
3. Parametric and Non-parametric tests
 - (i) Tests based on mean
 - (ii) Chi-square test for independence of attributes
 - (iii) Run test
 - (iv) Sign test and Wilcoxon signed rank test
 - (v) Kruskal Wallis and Friedman test
 - (vi) Kolmogorov-Smirnov test
 - (vii) Median test
4. Design of Experiments
 - (i) One-way and two-way ANOVA
 - (ii) Factorial experiments
 - (iii) ANOCOVA

STAT 597 – PROJECT AND VIVA-VOCE**CREDITS: 4**

1. A project work is compulsory and shall be offered in semester X. It will have 4 credits.
2. A project work may be taken individually or by a group of two students.
3. Project work shall be supervised by a faculty member assigned by the Head of the Department in the beginning of the semester.
4. 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.
5. At the end of the semester, before the last working day, 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 report separately (not copy of the same report).
6. 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 and related references should also be given.
7. The project work will be assessed for a maximum of 100 marks. Each student will give a seminar before the end of the semester on their project work which will be evaluated internally for a maximum of 30 marks. There will be viva-voce examination for a maximum of 10 marks by an internal and an external examiner. The project report will be valued by the same external and internal examiner for a maximum of 60 marks.

ELECTIVES

STAT 601 - TIME SERIES ANALYSIS

CREDITS: 4

Unit I

Stochastic Time Series models – Stationary models and their autocorrelation properties – Estimation of autocorrelation and its standard error – Periodogram – Spectrum and spectral densities – link between autocorrelation and sample spectrum

Unit II

General linear stationary models – stationarity and invertability – Autoregressive and moving average processes and their autocorrelation functions – mixed autoregressive moving average processes

Linear nonstationary models – ARIMA processes and their explicit forms – Integrated MA processes

Unit III

Forecasting: MMSE forecasts and their properties – Forecasts and their updating – Forecast of functions and forecast weights – examples

Unit IV

Model Identification – Identification techniques – Initial estimates for different processes – MA, AR, ARMA – choice between stationary and nonstationary models – model multiplicity

Unit V

Model estimation – Likelihood and sum of squares functions – Nonlinear estimation – estimation for special processes AR, MA, mixed processes – separation of linear and nonlinear components in estimation – estimation using Bayes' theorem

Books for Study:

1. Box G E P, Jenkins G M and Reinsel G C (2004): Time Series Analysis – Forecasting and Control, Pearson Education.

Books for Reference:

1. Kendall, Sir Maurice and Ord J K (1990): Time Series, Edward Arnold.
2. Chatfield C (1996): The Analysis of Time Series: Theory and Practice, fifth edition, Chapman and Hall.
3. Brockwell P J and Davis R A (2002): Introduction to Time Series and Forecasting, Springer.
4. Montgomery D C and Johnson L A (1977): Forecasting and Time Series Analysis, McGraw Hill.

Unit I

Parameter Programming – Parameterization of the Cost Vector ‘c’ -Parameterization of requirement vector ‘b’ – All integer programming problem- Gomory’s cutting plane algorithm – Mixed integer programming problem – Branch and Bound technique.

Unit II

Inventory models with one or two price breaks - Multi item deterministic problem – Constraints on storage and investment – Probabilistic Inventory models – Periodic Review systems – Fixed order quantity system

Unit III

Non-linear programming problem – Kuhn Tucker conditions – Quadratic programming problem (QPP) - Wolfe’s and Beale’s algorithms for solving QPP – Geometric programming

Unit IV

Dynamic programming problem (DPP) - Bellman’s principle of optimality - General formulation - computation methods and application of DP - Solving LPP through DP approach - Convex programming

Unit V

Queuing theory – Basic characteristics of queuing models – Arrival and service distribution – steady state solution of M/M/1 and M/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 result.

Books for Study:

1. Hamdy A. Taha (1995): Operations Research – An Introduction , Prentice Hall of India Private Ltd, New Delhi, 5/e.
2. Sinha S M (2006): Mathematical Programming : Theory and Methods, Elsevier Publications.
3. Gross and Harris(1982): Fundamentals of Queueing Theory, John Wiley .

Books for Reference:

1. Sharma .S.D: Operation Research , Kedar Nath Ram Nath & co, Meerut , 1999. 12 th.
2. Kanti Swarup, P.K. Gupta and Man Mohan : Operations Research, Sultan Chand and Sons, New Delhi, 2004.

Unit I

Reliability concepts and measures; components and systems; coherent systems; reliability of coherent systems; cuts and paths; modular decomposition; bounds on system reliability; structural and reliability importance of components.

Unit II

Life distributions; reliability function ; hazard rate; common life distributions – exponential , Weibull, Gamma, etc. Estimation of parameters and tests in these models.

Unit III

Notions of ageing: IFR, IFRA, NBU, DMRL, and NBUE Classes and their duals and implications

Unit IV

Common bivariate exponential distributions and their properties

Unit V

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. (1982) Statistical Models and Methods of Life Time Data; John Wiley.

Books for Reference:

1. Bain L.J. and 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 , Reliability Analysis, Springer Verlag.

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. Ageing classes – and their properties , Bathtub Failure 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.

Semi- parametric regression for failure rate – Cox’s proportional hazards model with one and several covariates. Rank test for the regression coefficients.

Unit V

Competing risks model, parametric and non- parametric inference for this model. Multiple decrement life table.

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, New York.

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, John Wiley.
4. Lawless J.F. (1982) Statistical Models and Methods of Life Time Data; John Wiley.

Unit I

Introduction to Data Mining – Different data mining tools – Different data types – Distance metrics – Information measures- Entropy, Gini Index

Unit II

Basic Clustering : Hierarchical clustering methods – partitioning clustering methods – k-means, k-medoids, fuzzy k-means – clustering validity measures – Purity, precision, recall, silhouette

Unit III

Advanced Clustering : Density Based Clustering – DBSCAN – DENCLUE , Genetic clustering – Model based clustering

Unit IV

Decision trees- rule deduction using decision trees – k-nn classification, naïve Bayesian classification – vector support machines and simple applications

Unit V

Association analysis – Market Basket Analysis – A priori algorithm –Pruning and Candidate generation – Rule Mining

Books for Study:

Tan, Steinbach and Kumar :Introduction to Data Mining , Pearson Education
(Portions of Chapters 2,4,5,6,8,9)

Books for Reference:

Rajan Chattamvelli (2009): Data Mining Methods, Narosa Publishing House, New Delhi

Unit I

Aitken's generalized least squares (GLS) estimator – Heteroscedasticity – Auto-correlation – Test of auto-correlation.

Unit II

Multi collinearity - Tools for handling multicollinearity – The idea of ridge regression – Properties of ridge regression.

Unit III

Linear regression with stochastic regressors – Errors in variable models and instrumental variable estimation – Independent stochastic linear regression – Auto regressive linear regression – Contemporaneously uncorrelated linear regression – Distributed lag models.

Unit IV

Statistical analysis of disturbances – Quick review of results on generalized least squares and linear constraints – Constrained generalized least squares estimation and prediction – BLUE residual vector – Tests against non-linearity – Techniques for detection of outliers.

Unit V

Simultaneous linear equations model – Structure of linear equations model – Identification problem – Rank and order conditions – Single equation and simultaneous equations - Methods of estimation – Indirect least squares – Least variance ratio and two – stage least squares.

Books for Study:

1. Gujarathi, D : Basic Econometrics, Mc Graw Hill.
2. Johnston, Econometric Methods, Mc Graw Hill, Third Edition, 1984.

Books for Reference:

1. Schmidt Peter: Econometrics, Marcel Dekker, 1976.
2. Theil, H : Principles of Econometrics, John Wiley, 1971.
3. Theil, H : Introduction to Econometrics, Prentice Hall of India Private Ltd, 1981.
4. Wetherill, G.B : Regression analysis with applications, Chapman & Hall, 1986.

Unit-I

Life tables and Insurance; Utility Theory and Insurance- Elements of Insurance and Optimal Insurance- Models for Individual Claims and their Sums.

Life Tables and Survival functions: Probability for the age at death and time until death for a person, curtate future life time and force of mortality- Relation between Life Table function and Survival function

Unit-II

Life Insurance: Principals of Compound Interest- Nominal and Effective rates of interest and Discount- Force of Interest and Discount- Accumulation factor and continuous compounding-Insurance payable at the time of death-Level-Benefit Insurance-Endowment insurance-Deferred Insurance- Varying benefit Insurance-Insurance payable at the end of the year of death-recursions and commutation functions.

Life-Annuities: Single payment –continuous and discrete Life annuities. Life annuities with monthly payment-varying annuities-apportion able annuities- Due and complete annuities immediate.

Unit-III

Benefit premiums: Fully continuous and fully discrete premiums, true monthly payment premiums, apportionable premiums and accumulation benefits

Benefit reserves: Fully continuous benefit reserves, benefit reserves on a semi-continuous basis, benefit reserves based on monthly benefit premiums, benefit reserves on an apportionable or discounted continuous basis.

Analysis of Benefit Reserves: benefit reserves for general insurance, reserves at fractional durations, allocation of loss to policy years, recursive formulae and differential equations for reserves and commutation functions.

Unit-IV

Multiple Life Functions: Joint distribution of future life times, the joint life status, the last-survivor status, dependent lifetime models, common shock, capulas, insurance and annuity benefits, survival status, special two life annuities, reversionary annuities.

Gompertz and Macham laws, uniform distribution, simple contingent functions and their evaluation.

Multiple decrement Models: Deterministic and random survivor groups, associated single decrement tables, central rates of multiple decrement, net single premiums and their evaluation. Construction of multiple decrement table.

Applications of Multiple Decrement Theory: Actuarial present value and their numerical evaluation, withdrawal benefit patterns that can be ignored in evaluating premiums and reserves, valuation of pension plans, demographic assumptions, projecting payment and contribution rates, defined-benefit plans, defined-contribution plans, disability benefits with individual life insurance, disability income benefits waiver premium benefits.

UNIT-V

Collective risk Model for a Single Period: Distribution of aggregated plans, selection of basic distributions, distribution of individual claim amount, properties of certain compound distributions and approximations to the distribution of aggregate claims.

Collective Risk models over an extended period: Discrete and Continuous models, ruin probabilities and the claim amount distribution, first surplus below the initial level, maximum aggregate loss.

Books for Study:

1. N.L. Bowers, H.U. Gerber, J.C.Hickman, D.A. Jones and C.J. Nesbitt, (1997), “Actuarial Mathematics”, Society of Actuaries, Ithaca, Illioid, U.S.A. Second edition. (Chapters: 1, 2, 3, 4,5,6,7, 8, 9, 11, 13, 14).

Books for Reference:

1. Spurgeon E.T. (1972): Life Contingencies, Cambridge University Press.
2. Neill, A. (1977): Life Contingencies, Heineman.
3. Ammann.M(2002): Credit Risk Valuation, Springer Verlag
4. Bol, G, G Nakhaeizadeh, S.T Rachev. et al (2003): Credit Risk, Springer Verlag.
5. Bielecki, T.R(2002): Credit Risk: Modeling valuation and Hedging, Springer Verlag

UNIT-I

Introduction to DB Systems: Overview-historical perspective-file system versus DBMS-Advantages of DBMS-Storing Data in DBMS-Queries in DBMS-Structure of a DBMS-Use of DBMS.

ER Model: Overview-Entities -Attributes and Entity Sets-Relationships and Relationship Sets-Features of ER Model-

UNIT-II

Relational Model: Introduction-Integrity Constraints over relations -. Enforcing Integrity Constraints-Querying Relational Data-Logical Database Design: ER to Relational - View: Introduction- Altering/Destroying tables and views.

UNIT-III

Database Design:-Introduction to Schema refinement-functional dependencies-Reasoning for function Dependencies-Normal Forms-Decompositions-Normalization.

UNIT-IV

Database Security: Introduction-Access Control-Discretionary Access Control-Mandatory Access Control-Additional Issues Related to Security – Introduction to Transactions.

UNIT-V

Design of ER Model, Design of Database, Normalization and Queries for the Case Studies: Student Information System-Library Information System-Railway Reservation System.

Books for study:

Raghu Ramakrishnan and Johannes Gehrke (2007): Database Management System, 2/e.

Books for Reference:

1. Elmasri Navathe (2000): Fundamentals of Database System, 3/e, Addison Wesley.
2. Jeffrey D.Ullman and Jeffrey Widom (2001): A First course in Database System, Pearson Education, Asia.
3. Selbetschhatz and Sudharshan (2002): Database system Concepts, 4/e, McGraw Hill

Unit I

Subjective Interpretation of probability in terms of fair odds. Evaluation of (i) Subjective probability of an event using a subjectively unbiased coin (ii) Subjective prior distribution of a parameter - Bayes theorem and computation of the posterior distribution.

Unit II

Natural Conjugate family of priors for a model. Hyper parameters of a prior from conjugate family. Conjugate families for (i) exponential family models. (ii) models admitting sufficient statistics of fixed dimension. Enlarging the natural conjugate family by (i) enlarging hyper parameter space (ii) 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: as a prediction problem from posterior distribution. Bayes estimators for (i) absolute error loss (ii) squared error loss (iii) 0 -1 loss. Generalization to convex loss functions. Evaluation of the estimate in terms of the posterior risk.- Bayesian interval estimation : Credible intervals. Highest posterior density regions - Interpretation of the confidence coefficient of an interval and its comparison with the interpretation of the confidence coefficient for a classical confidence interval.

Unit IV

Bayesian Testing of Hypothesis: Specification of the appropriate form of the prior distribution for a Bayesian testing of hypothesis problem - Prior odds, Posterior odds, Bayes factor for various types of testing hypothesis problems depending upon whether the null hypothesis and the alternative hypothesis are simple or composite. Specification of the Bayes tests in the above cases. Discussion of Lindley's paradox for testing a point hypothesis for normal mean against the two sided alternative hypothesis.

Unit V

Bayesian prediction problem - Large sample approximations for the posterior distribution - Bayesian calculations for non conjugate priors: (i) Importance sampling, (ii) Obtaining a large sample of parameter values from the posterior distribution using Acceptance – Rejection methods, Markov Chain Monte Carlo methods and other computer simulation methods.

Books for Study:

1. Berger, J.O. Statistical Decision Theory and Bayesian Analysis, Springer Verlag.
2. Robert C.P. and Casella, G. Monte Carlo Statistical Methods, Springer Verlag.
3. Leonard T. and Hsu, J.S.J. Bayesian Methods , Cambridge University Press.
4. Bansal , Bayesian Inference, Narosa Publications

Books for Reference:

1. DeGroot M.H. Optimal Statistical Decisions. Mc Graw Hill.
2. Bernardo. J.M. and Smith, A.F.M. Bayesian Theory, John Wiley and sons.
3. Robert, C.P. The Bayesian Choice : A decision Theoretic Motivation , Springer.
4. Gemerman, D. Markov Chain Monte Carlo : Stochastic Simulation for Bayesian Inference. Chapman Hall.
5. Box, G.P. and Tiao, G.C. Bayesian Inference in Statistical Analysis, Addison – Wesley.

Unit I

Bioassays : Quantitative and qualitative response – Dose response relation – Estimation of modern effective dose - Estimation of unknown concentration of potency – Probit and logit transformations .

Unit II

Parallel line and slope ratio assay – Potency ratio – Feller's theorem – Tests for non validity – Symmetric and asymmetric assays – Toxic action of mixtures.

Unit III

Types of mating – Random mating – Hardy Weinberg equilibrium – Random mating in finite population – Inbreeding through generation matrix approach.

Unit IV

Segregation and linkages - Estimation of segregation and linkage parameters. Concept of gene frequencies - Estimation of gene frequencies.

Unit V

Quantitative inheritance – Genetic parameters - Heritability – Genetic correlation and repeatability – Methods of estimation – Selection and its effect – Selection index – Dialed and partially dialed crosses.

Books for Study:

1. Daniel : Bio statistics – A foundation for analysis in health sciences, John Wiley, Third Edition

Books for Reference:

1. Elandt – Johnson : Probability models and statistical methods in genetics, 1971.
2. Jain, J.R : Statistical techniques in quantitative genetics, Tata Mc Graw Hill, 1982.
3. Poti, S.J : Quantitative study in life sciences, Vikas Publishing Ltd, 1983.

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.