

# VEER NARMAD SOUTH GUJARAT UNIVERSITY

## M.A. ( Part-II ) Statistics

Effective From: 2002-03 – 2003-04

Paper No.	Title of the paper	Marks			Work load / week
		Int.	Ext.	Total	
Paper- VI	Testing of Hypotheses & Decision Theory	18	42	60	3L + 1T
Paper VII	Linear Models & Design of Experiments	18	42	60	3L +1T
Paper VIII	Optional Paper I	18	42	60	3L +1T
Paper IX	Optional Paper II	18	42	60	3L +1T
Paper X	Optional Paper III	18	42	60	3L + 1T
Practical : Paper 3	Practicals based on Papers VI, VII, VIII & IX	15	35	50	6 hours
Practical : Paper 4	Practicals based on Paper X	15	35	50	6 hours
Viva – Voce		15	35	50	
Project		50		50	

# VEER NARMAD SOUTH GUJARAT UNIVERSITY

## M.A. ( Part-II ) Statistics

### Paper VI Testing of Hypotheses & Decision Theory

#### Section A:( TESTING OF HYPOTHESIS)

Stating the problem of testing of hypotheses as a special case of general decision problem. Simple and composite hypotheses, Critical function and critical region, randomised test, non –randomised test, size of the test, Power function of a test. Generalised Neyman – Pearson’s Lemma, Most Powerful test and Uniformly most Powerful test. UMP tests for families of distributions admitting monotone likelihood ratio, two sided hypothesis, use of least favourable distribution. Likelihood ratio test for simple and composite hypotheses.

Unbiasedness for testing of hypotheses. Similar test, relationship with UMP unbiased test, UMP similar test and its application to one parameter exponential family, Similarity and completeness, tests with Neyman structure, UMP unbiased tests for multiparameter exponential families, Locally most Powerful unbiased tests.

Invariance in testing of hypotheses, maximal invariant test, most powerful invariant test, unbiasedness and invariance.

Sequential testing of hypotheses, Wald’s sequential probability ratio test (SPRT), Properties of SPRT, approximate bounds, OC and ASN functions Efficiency of SPRT, Fundamental identity of sequential analysis and its use to obtain OC and ASN functions of SPRT.

#### REFERENCES

1. T.S.Ferguson : Mathematical Statistics, Academic Press, (1967).
2. M.G. Kendall and A. Stuart: The Advanced Theory of Statistics, Vol. 2. Ed. IV, Charles and Griffin, (1960).
3. E.L.Lehman : Testing Statistical Hypotheses, Wiley eastern, (1976).
4. A. Mood, F. A.Frabil and D.C.Boes: Introduction to the Theory of Statistics, McGraw Hills, International Student Ed. III, (1974).
5. C.R.Rao: Linear Statistical Inference and its Applications, Wiley Eastern, EdII, (1973).
6. Sddney Segial: Nonparametric Methods for Behavioral Sciences, McGraw Hill (1956).
7. S. Wilks: Mathematical Statistics, Wiley New york Ed. II, (1962).
8. S. Zacks: The Theory of Statistical Inference, Wiley New york, (1971)

## Section B : (DECISION THEORY)

Review of basic elements of statistical decision problem. Various inference problems viewed as decision problems. Randomization, Optimal decision rules. Bayes and minimax decision rules. Geometric interpretation for finite parameter space. Generalized Bayes and extended Bayes rules.

Natural ordering of decision rules. Complete and essentially complete classes of decision rules. Admissibility of Bayes rules. Existence of Bayes decision rules and of Minimax complete class when parameter space is finite and the risk set is closed and bounded from below.

Invariant decision problems, Invariant decision rules. Admissible minimax invariant rules.

Introduction, to non-parametric test Sign test, Run test, Median – test, rank test, Wilcoxon & Mann-Whitney test, Kolmogorov – Smirnov test of goodness of fit. Chi-square test of goodness of fit. Fisher's exact probability test for independent samples, Extension of Median test, Kruskal Wallis's test; One way analysis of variance for K-independent samples. Cochran Q test for related samples. Friedman two way analysis of variance. Kendall's rank correlation, Kendall's partial rank correlation, Kendall's coefficient of concordance.

### REFERENCES

1. J.O.Berger: Statistical Decision Theory, Springer –Varlag pub. Co., New york, ( 1980).
2. T.S.Ferguson : Mathematical Statistics, Academic Press, (1967).
3. J.D.Gibbons: Nonparametric Statistical Inference, McGraw Hills, (1971)
4. Sedney Segal : Nonparametric Methods for Behavioral Sciences, McGraw Hill (1956).

# VEER NARMAD SOUTH GUJARAT UNIVERSITY

## M.A. ( Part-II ) Statistics

### Paper VII Linear Models & Design of Experiments

#### Section A : LINEAR MODELS

The general linear model : Gauss-Markoff set –up, estimation of its parameters, least squares, and generalised least squares, Normal equations and least squares estimates. Estimation of linear parametric function, variance and covariances of least squares estimates, estimation of error variance, estimation with correlated observations, least square, estimates with restrictions on parameters, simultaneous estimates of linear parametric functions, Canonical form of the linear hypothesis model and Error and Estimation spaces.

Estimation of scale parameter in the general linear model by quadratic functions, necessary and sufficient conditions for (i) a quadratic form to be distributed as chi-square (ii) two quadratic forms to be independently distributed, Cochran's theorem and its generalization.

Tests of hypotheses regarding parameters of a general linear model, tests involving linear functions of parameters, tests of subhypotheses.

#### REFERENCES

1. Rao, C.R. (1973) : Linear statistical inference and its applications; 2<sup>nd</sup> . Ed., John Wiley and Sons, Inc.
2. Searle, S.R. (1971) : Linear Models.
3. Seber, G.A. F. (1977) : Linear Regression Analysis.
4. Graybill, F. A. (1961) : An Introduction to linear statistical models.
5. Kshirsagar A.M. (1983) : A Course in linear models.

## **Section B : DESIGN AND ANALYSIS OF EXPERIMENTS**

General theory of analysis of experimental designs with one way elimination of heterogeneity (intra-block analysis only)

General properties of incomplete block design; Concepts of connectedness, balance and orthogonality. Balanced incomplete block design and related designs, resolvable and affine resolvable BIB designs, Partially balanced incomplete block designs with two associated classes.

General theory of analysis of experimental designs with two way elimination of heterogeneity ( intra-block analysis only); youden square and Cross over design.

Missing plot technique; its application to randomized block, latin square and balanced incomplete block designs.

General theory of symmetric factorial experiments; concepts of total and partial confounding and  $2^n$  and  $3^n$  confounded experiments.

Construction of (i) orthogonal latin squares, (ii) Balanced incomplete block designs using finite geometries and (iii) total and partially confounded symmetric factorial experiments.

### **REFERENCES**

1. Chakraborti., M.C. (1962): Mathematics of Design of Experiments.
2. Dey, Aloke (1986) : Theory of Block Designs.
3. Raghva Rao, D. (1971) : Construction and Combinatorial Problems in Design of Experiments.
4. Kempthorne, O. (1952) : The Design and Analysis of Experiments.
5. Federer, W.T. (1955) : Experimental Designs.
6. Das, M. N. and Giri, N. (1979) : Design and Analysis of Experiments.
7. Ogawa, J. (1974) : Statistical Theory of the Analysis of Experimental Designs.
8. John , P.W.M. (1971) : Statistical Design and Analysis of Experiments.
9. Joshi, D. D. (1987) : Linear Estimation and Design of Experiments.

The three optional papers in M.Sc. ( Part-II ) are to be selected from out of the following list:

### **List of Optional Papers:**

1. **Mathematical Economics & Econometrics**
2. **Operations Research**
3. **Advanced Computer Programming and FoxPro**
4. **Computer Intensive Statistical Methods**
5. **Data Mining and Pattern Recognition**
6. **Advanced Sample Surveys and Official Statistics**
7. **Advanced Probability Theory**
8. **Advanced Stochastic Processes**
9. **Advanced Regression Analysis**
10. **Bio-Statistics**

# **Mathematical Economics and Econometrics**

## **SECTION – A**

### **MATHEMATICAL ECONOMICS**

1. Input – Output Analysis: Leontief's static models for inter –industry relations. The Walras-Leontief's closed system.
2. Growth Models: Classical and Keynesian simple income determination models. Concepts of multipliers and accelerator. Hicks-Samuelson model, Harrod-Domar and Solow's Growth models. Mahalanobis two and four sector models.
3. Time Series Analysis: Definition, Importance of time series analysis, Components of a time series, Trend, seasonal fluctuations, Determination of trend, moving average method, variate difference method, fitting of polynomial curve, Stationary time series, Box-Jenkin's models, Introduction to AR, MA, ARMA and ARIMA models, Detailed study of AR (1) MA (1) and ARMA (1,1) Models.

#### **BOOKS:**

1. Kendall, M. (1976) : Time Series; Charles Griffin and Company.
2. Box and Jenkins (1970) : Time Series Analysis: Forecasting and Control, Holden Day Pub.
3. Chatfield, C. (1975) : The Analysis of Time Series: Theory and Practice; Chapman and Hall.
4. Waller Vancuels (1983) : Applied Time Series and Box Jenkins Models.

### **Section B: ECONOMETRICS**

The nature and role of econometrics, Introduction to econometric models.

Single Equation method: Revision of the study of Gauss-Markoff theorem, general linear model and its extensions, OLS estimation and Prediction.

Problem of multicollinearity: Detection and consequences, Ridge estimator

Generalized least square method, Heteroscedastic disturbances, Grouping of observations.

Autocorrelation, its consequences and tests. The BLUS procedure. Linear Regression with stochastic regressors, Instrumental variable method Problem of errors, Distributed Lag Models.

Simultaneous Equation method : The Identification problem, Rank and Order conditions.

Estimation methods: Recursive systems, 2SLS Estimators, Limited Information estimators, K-class estimators, Introduction to 3SLS estimators, and Full Information ML method.

#### **REFERENCES**

1. Apte, P.G. (1990): Text Book of Econometrics, Tata McGraw Hill.
2. Chatterjee and Price, B. : Regression Analysis by Example, John Wiley & Sons, (1977) .
3. Cramer (1971): Empirical Econometrics, North Holland.
4. D.Gujarati (1979) : Basic Econometrics, McGraw Hill.
5. H. D. Intriligator (1980) : Econometric Methods, Techniques and Applications, Prentice Hall Pub. Co..
6. J.Jonston (1982) : Econometric Methods, McGraw Hill, Kogakusha Ltd.
7. L.R.Klein: An Introduction to Econometrics, Prentice Hall of India, (1969).
8. Kontsoyiannis, A. (1979): Theory of Econometrics ; Mac Millan press.
9. Malinvar, E. (1966) : Statistical Methods in Econometrics: North Holland.
10. Theil, H. C. (1982) : Introduction to the Theory and Practice of Econometrics, John Wiley.
11. C. F. Charist (1966) : Econometric Models and Methods, John Wiley.

# OPERATIONS RESEARCH

## Section A

1. Linear Programming:
  - 1.1 Definition of linear programming problem (LPP)
  - 1.2 Formulation of LPP
  - 1.3 Solution of LPP by Graphical and Simplex Method ( including Big-M and two – phase methods)
  
2. Duality:
  - 2.1 Definition of Dual Problem
  - 2.2 Rules for converting any Primal into its Dual
  - 2.3 Properties of Duality
  - 2.4 Dual – Simplex Method
  
3. Transportation and Assignment Problems:
  - 3.1 Definition of Transportation problem (TP)
  - 3.2 Special structure of TP
  - 3.3 Methods for Getting Basic feasible solution to TP
  - 3.4 Methods for getting Optimum solution to TP
  - 3.5 Unbalanced TP
  - 3.6 Definition of Assignment Problem ( AP)
  - 3.7 Algorithm for solving an AP
  - 3.8 Unbalanced AP
  - 3.9 Routing Problem
  
4. Inventory Management Systems:
  - 4.1 Definition
  - 4.2 Costs involved in Inventory Problems
  - 4.3 Classical EOQ Models without and with shortages
  - 4.4 Multi-item Deterministic Models
  - 4.5 Probabilistic Inventory Models
  - 4.6 Inventory Models with Price Breaks
  
5. Game Theory:
  - 5.1 Two Person zero sum game
  - 5.2 Game with and without saddle Points
  - 5.3 Mixed Strategies
  - 5.4 Two person zero sum game and its transformation into a linear Programming problem.
  - 5.5 Minimax theorem ( without proof).

### REFERENCE BOOKS

1. K. Swarup, P.K.Gupta and Man Mohan : Operations Research, S. Chand & Co., New Delhi.
2. G. Hadley : Linear Programming, Oxford & IBH Pub. Co., 1969.
3. K.G. Murthy: Linear and Nonlinear Programming.

## Section B

1. Sensitivity Analysis:
  - 1.1 Basic concepts
  - 1.2 Changes in the coefficient of objective function
  - 1.3 Changes in the components of vector b and of Matrix A
  - 1.4 Addition / Deletion of variable in the problem
  - 1.5 Addition / Deletion of constraint in the problem
2. Integer Programming :
  - 2.1 Introduction
  - 2.2 All and mixed integer programming (IPP) Problems
  - 2.3 Gomory 's all –IPP algorithm
  - 2.4 The branch and bound technique
  - 2.5 Zero – one programming
3. Replacement Theory
  - 3.1 Types of Replacement Problem
  - 3.2 Replacement of Items that Deteriorate
  - 3.3 Replacement of Items that fails completely and that of Staff
4. PERT / CPM:
  - 4.1 Basic concepts
  - 4.2 Construction and Time Calculation of the Network
  - 4.3 Determination of Float and of the Critical Path
  - 4.4 Crashing a Project
  - 4.5 Scheduling a Project
  - 4.6 Resource Analysis and Allocation
  - 4.7 Application of PERT/ CPM
5. Sequencing :
  - 5.1 Definition, Notations and Assumptions
  - 5.2 Solution of Sequencing problem
  - 5.3 Problems with n-jobs and 2-machines
  - 5.4 Problems with n-jobs and 3-machines
  - 5.5 Problems with 2-jobs and m-machines
6. Goal Programming :
  - 6.1 Definations and Concepts.
  - 6.2 Formulation of Goal Programming Problem (GPP)
  - 6.3 Solution of GPP by Graphical and Extended Simplex Methods

### REFERNCES BOOKS

1. K. Swarup, P.K.Gupta and Man Mohan: Operations Research, S. Chand & Co., New Delhi.
2. G.Hadley : Linear Programming, Oxford & IBH Pub. Co., 1969.
3. K.G.Murthy : Linear and Nonlinear Programming.

# Advanced Computer Programming and FoxPro

## Section A : Advanced Computer Programming

1. Pointers:
  - 1.1 Basis of Pointers
  - 1.2 Pointer Arithmetic
  - 1.3 Pointer Array
  - 1.4 Call by reference in user defined functions
  - 1.5 Pointer to function
2. Files:
  - 2.1 Reading and writing from a file
  - 2.2 Reading and writing Structures
  - 2.3 Random accessing a file
3. Principles of object oriented Programming :
  - 3.1 Procedure oriented Programming Vs object Oriented Programming
  - 3.2 Benefits of Object oriented Programming
4. Classes and Objects
5. Constructions and Destructions
6. Operators Overloading, Functional Overloading and Type Conversions.
7. Inheritance

### REFERENCES

1. R. Decker and S.Hirshfield (1998) : The Object Concept: An Introduction to Computer Programming using C++ , PWS Publishing.
2. S.B.Lippmann and J.Lajoie ( 1998) : C++ Primer , Third Ed., Addison Wesley.
3. W.J. Savitch (2001): Problem Solving with C++ : The Object of Programming. Third Ed., Addison Wesley.

## **SECTION-B**

### **FOXPRO**

1. Database Package:
  - 1.1 OVERVIEW
    - 1.1.1 Database, Record, Field.
    - 1.1.2 Capabilities and Limitations.
  - 1.2 Operators.
  - 1.3 Database & Array Handling:
    - 1.3.1 Creation,
    - 1.3.2 Modification,
    - 1.3.3 Processing.
  - 1.4 Generating Menu:
    - 1.4.1 POPUP Mcnus,
    - 1.4.2 BAR Menus,
    - 1.4.3 Pull Down Menus.
  - 1.5 Screen Handling :
    - 1.5.1 Generating Screen & Windows.
  - 1.6 Inbuilt Functions, UDFs & Procedures.
  - 1.7 SET Commands.
  - 1.8 REPORT & LABEL Generation.
  - 1.9 Handling Memo Fields.
2. Introduction to Spreadsheet.

### **REFERENCE BOOKS**

1. Illustrated FoxPro 2.0: Robert Granillo BPB Publications.
2. Illustrated FoxPro 1.0: Robert Granillo, BPB Publications.
3. Fox Pro 2.0 the art of Visual Programming : Mukhis
4. Clipper 5 & 5.01 Programming Guide : Spence Rick, BPB Publications.
5. Clipper 5 & 5.01 Developers Guide: BPB Publications.
6. Excel 5 for Windows, Quick & Easy : Jones TECH.
7. Mastering Excel 5 for Windows : Chester , BPB Publications.

# Advanced Sample Surveys and Official Statistics

## Section A : Advanced Sample Surveys

1. Determination of optimum stratification points when (i) stratification variable is same as character under study (ii) stratification variable is different than character under study under (i) Proportional, (ii) Neyman's allocation, and (iii) equal allocation . Stratification for multivariate character under study.
2. Rao-Hartley – Cochran Sampling scheme. Durbin's generalization of varying probability schemes to two – stage (multi-stage ) design. IPPS schemes due to Sen –Midzeno, Brewer, Durbin and JNK Rao ( for  $n = 2$  only).
3. Introduction to unified theory of finite population: Class of linear homogeneous unbiased estimators. Non-existence of uniformly minimum variance estimator in the General class of linear homogeneous unbiased estimators.
4. Non-Sampling errors and biased responses, randomised response for variables, errors in surveys, modelling observation errors, estimation of variance components.
5. Introduction to super population models.

### REFERENCES

1. Chaudhary, A. and JWE Vos (1988) : Unified theory and strategies of survey sapling, North Holland, Amsterdam.
2. Singh, R. and Singh, M. (1986) : Theory and Analysis of Sampling Design; New Age International Publications.
3. Cochran, W.G. (1984) : Sampling techniques, 3 rd Ed., Wiley, New york.
4. Caarsel, Sarandal & Writman.
5. Murthy , M.N. (1977) : Sampling theory, Stat. Pub. House, Kolkata.
6. Sukhatme et al. (1984) : Sampling Theory with Applications; Iowa State Univ. Press& ISAS. New Delhi.
7. Durbin' J. (1967) : Estimation of Sampling Error in Multistage Design; Appl. Stat., 16, 152 – 164.

## **Section B : Official Statistics**

Introduction to Indian and International statistical systems. Role, function and activities of Central and State statistical organizations. Organization of large scale sample surveys. Role of national Sample Survey Organization. General and special data dissemination systems.

Population growth in developed and developing countries, evaluation of performance of family welfare programmes, projections of labour force and manpower. Scope and content of population census of india

System of collection of Agricultural Statistics. Crop forecasting and estimation, productivity, fragmentation of holdings, support prices, buffer stocks, impact of irrigation projects.

Statistics related to industries, foreign trade, balance of payment, cost of living, inflation, educational and other social statistics.

### **REFERENCES**

Basic Statistics Relating to the indian Economy (CSO) 1990.

Guide to Official Statistics (CSO) 1999.

Statistical System in India (CSO) 1995.

Principles and accommodation of National Population Censuses, UNESCO.

Panse, V. G., Estimation of Crop Yields (FAO).

Family Welfare Yearbook. Annual Publication of D/o Family Welfare.

Monthly Statistics of Foreign Trade in india. DGCIS. Calcutta and other govt. Publications.

# **Data Mining and Pattern Recognition**

## **Section A : Data Mining**

Review of classification Methods from multivariate analysis ; classification and decision trees.

Clustering methods form both statistical and data mining viewpoints ; vector quantization

Unsupervised learning form univariate and multivariate data ; dimensional reduction and feature selection.

Supervised learning from moderate to high dimensional input spaces; artificial neural networks and extension of regression models, regression trees.

Introduction to databases, including simple relational databases; data warehouses and introduction to online analytical data processing.

Association rules and prediction; data attributes, applications to electronic commerce.

Note on practicals / tutorials : Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing Note on Prerequisites: This course assumes successful completion of Computer – intensive Statistical Methods I. It can be presented in conjunction with Computer – intensive Statistical Methods II.

### **REFERENCES**

A. Berson and S.J. Smith ( 1997). Data Warehousing, Data Mining, and OLAP. McGraw–Hill.

L. Breiman, J.H. Friedman, R. A. Olshen, and C.J.Stone (1984) . Classification and Regression Trees. Wadsworth and Brooks/ Cole.

J. Han and M. Kambar (2000). Data mining; Concepts and Techniques. Morgan Kaufmann.

T.M. Nitchell (1997) Machine Learning. McGraw – Hill.

B.D. Ripley (1996 ). Pattern Recognition and Neural Networks. Cambridge University Press.

# Computer Intensive Statistical Methods

## Section A

Exploratory data analysis: transforming data, graphical methods of clustering, outliers. (5L)

Linear regression : influential observations and diagnostics, robust methods, collinearity, variable selection. (11L)

Generalized linear models: exponential families and ML estimation, analysis of deviance and variable selection, logistic regression. (10L)

## Section B : PATTERN RECOGNITION

Linear classification : linear discriminant function ( LDF) for minimum squared error, LDF for binary outputs, perception learning algorithm.

Nearest neighbour decision rules: description ,convergence, finite sample considerations, use of branch and bound methods.

Probability of errors: two classes, normal distributions, equal covariance matrix assumptions, Chernoff bounds and Bhattacharya distance, estimation of probability of error.

Feature selection and extraction : interclass distance measures, discriminant analysis, probabilistic distance measures, principal components.

Note on practicals/ tutorials: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing.

Note on prerequisites: This course assumes successful completion of Computer – intensive Statistical Methods I. It can be presented in conjunction with Computer- intensive Statistical Methods II.

## REFERENBCES

R. O. Duda and P.E. Hart (1973). Pattern Recognition and Scene Analysis. Wiley.

K Fukunaga (1990) . introduction to Statistical Pattern Recognition, Second Edition. Academic Press.

G.J McLachlan (1992). Discriminant Analysis and Statistical Pattern recognition. Wiley.

B. D. Ripley ( 1996). Pattern Recognition and Neutal Networks. Cambridge University Press.

# Computer Intensive Statistical Methods

## Section A

Exploratory data analysis: transforming data, graphical methods of clustering, outliers.

Linear regression: Influential observation and diagnostics, robust method, collinearity, variable selection.

Generalized linear models: exponential families and ML estimation, analysis of deviance and variable selection, logistic regression.

Nonlinear regression: estimation, hypothesis testing goodness of fit.

EM algorithm: applications to missing and incomplete data problems, mixture models.

Smoothing with kernels : density estimation, simple nonparametric regression.

Note on practicals: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing.

## REFERENCES

R. Gnanadesikan (1997) ; Methods for Statistical Data Analysis of Multivariate Observations, Second edition. Wiley.

D.A. Belsley E. Kuh, and R.E Welsch (1980) ; Regression Diagnostics. Wiley.

P. McCullagh and J.A. Nelder (1999) ; Generalized Linear Models, Third edition. Chapman and Hall.

G. E. F. Seber and C. J. Wild (1989); Nonlinear Regression Diagnostics. Wiley.

G. J. McLachlan and T. Krishnan (1997); The EM Algorithms and Extensions. Wiley

J.S. Simonoff ( 1996) ; Smoothing Methods in Statistics. Springer.

## Section B

Stochastic simulation : Generating random variables, simulating multivariate distributions, simulating stochastic processes such as simple queues.

Variance reduction : importance sampling for integration, control variates and antithetic variables.

Markov Chain Monte Carlo methods: Gibbs sampling for multivariate simulation, simulated annealing for optimization.

Simulation based testing: simulating test statistics and power functions, permutation tests.

Bootstrap methods: resampling paradigms, bias and standard errors, confidence intervals, bootstrapping in regression.

Jackknife and cross-validation: Jackknife in sample surveys, cross-validation for tuning parameters.

Note on practicals: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing.

### REFERENCES

G. S. Fishman (1996); Monte Carlo : Concepts, Algorithms, and Applications, Springer.

R.Y. Rubinstein (1981): Simulation and the Monte Carlo Method. Wiley.

M.A. Tanner (1996); Tools for Statistical inference, Third Edition. Springer.

B. Efron and R. J. Tibshirani (1993) ; An introduction to the Bootstrap. Chapman and Hall.

J. Shao and D. Tu (1995) ; The Jackknife and the Bootstrap. Springer Verlag.

# Advanced Probability Theory

## SECTION- A

Metric Spaces, topological notions such as limit points, convergence, closed and open sets, continuity, compactness in metric spaces, isometries and homeomorphisms. Complete metric spaces, separable spaces, nowhere – dense sets, the metric spaces  $\mathbb{R}$  (infinity ) and  $C[0,1]$ .

Normed linear spaces, metric defined by a norm, equivalent norms, Banach spaces and examples ( $C[a, b]$   $l_p$ ,  $L_p$  (m) spaces).

Inner product spaces, Hilbert spaces, Bessel's inequality complete orthonormal sets, Parseval's identity,  $L_2$  space. The. Projection theorem for Hilbert spaces.

Bounded linear transformations, Banach Steinhaus theorem ( Principle of uniform boundedness), Open mapping theorem, the Hahn Banach theorem and applications.

The dual of a Banach space, (dual of  $C[a,b]$ ,  $l_p$ ,  $L_p$  spaces). Riesz representation theorem.

### REFERENCES

Functional Analysis (1966) George Bachman and Lawrence Narici, Academic Press.

Real Analysis (1988) H. L. Royden, MacMillan, New York.

Real and complex analysis ( 1966) Walter Rudin, McGraw Hill

Real Analysis and Probability (1972) Robert Ash, Academic Press.

Principles of Functional Analysis ( 1971) Martin Schechter, Academic

Functional Analysis (1973) Walter Rudin McGraw Hill.

## SECTION-B

Probability measures on metric spaces: Measures and integrals, tightness, (measure) determining class. Weak convergence of measures: properties of weak convergence and Portmanteau theorem. Convergence determining class.

Some special cases : Convergence of probability measures and convergence determining class in the Euclidean space, the circle, the space  $\mathbb{R}$  infinity, the product spaces etc.

Random element of a metric space and its distribution, convergence in distribution of sequence of random elements – various equivalent criteria for convergence; convergence in probability; weak convergence and mappings.

Relative compactness and Prohorov's theorem. Weak convergence and tightness in  $C[0,1]$ : wiener measure, Weak convergence and tightness in  $D [0,1]$ ; Donsker's theorem.

### REFERENCES

Billingsley, P. (1968) : Convergence of Probability Measures, Wiley.

Partasarathy, K.R. (1967) : Probability Measures on Metric Spaces, Academic Press.

# Advanced Stochastic Processes

## SECTION-A

Renewal theory : renewal function and its Properties, elementary and key renewal theorems, asymptotic normality of the number of renewals, cost/ rewards associated with renewals regenerative processes. Applications

A review of Markov chains analysis, computing stationary distribution of a large (finite ) state –space, applications.

Poisson Process and Pure Jump Markov processes: a review of main results. Birth and Death processes, Applications to Queuing theory , inventory analysis, communication models, population biology and finance, ecology and epidemilogy (depending upon the interests of students and the teacher). Computational issues.

Brownian motion applications to finance.

Semi-Markov processes and Markov decision processes: an introduction and applications to various areas.

## REFERENCES

Tijms, H. C. (1986) Stochastic Modelling and analysis, Wiley.

Medhi. J. (1982) Stochastic Processes, Wiley Eastern.

Ross, S.M. (1983) Stochastic Processes, Wiley.

Bhat, B.R. (2000) Stochastic Models: Analysis and Applications, New Age international Publications.

Feller, W. (1971) An introduction to Probability theory and its applications, Vol II.

Ross, S.M. (1970) Applied Probability models with optimization applictions. Holden-Day San Franscisco.

Wolff. R. W. (1989) Stochastic Modelling and the Theroy of Queues, Prentice Hall.

## SECTION –B

Point Processes: Definition and examples. Product densities and characteristic functional. Stationary point processes- Khintchine's existence theorem, intensity of point processes. Poisson process, Renewal process, doubly stochastic process.

Stationary processes: wide sense and strict sense stationarity, covariance functions-properties. Bochner's theorem, Herglotz theorem, Karunen's representation theorem. Khintchine's and Cramer's representation of wide sense stationary process. Spectral decomposition for weakly stationary process. Ergodic theorems.

Markov processes- continuous time and continuous state space, time homogeneous Markov processes, Kolmogorov's equations. Diffusion processes with Wiener process and Ornstein –Uhlenbeck process as particular cases. First passage time and reflecting barriers for Wiener process, Relation between Wiener and Uhlenbec processes.

### REFERENCES

Cox, D.R. and Miller, H. D. (1965) : The theory of Stochastic Processes, Mathuen & Co, London.

Cramer, H. and Leadbetter, M.R. (1867) : Stationary and Related Stochastic Proceeses, Wiley.

Daley, D. J. and Vere- Jones (1988) : An introduction to the Theory of Point Processes. Springer- Verlag.

Karlin, S. and Taylor, H. M. (1981): A Second Course in Stochastic Processes Academic Press.

Ross, S. M. (1983): Stochastic Processes, Wiley.

# Advanced Regression Analysis

## SECTION- A

Logistic and Poisson regression : logit model for dichotomous data with single and multiple explanatory variables, ML estimation, Large sample tests about parameter goodness of fit, analysis of deviance, variable selection, extension to polytomous data introduction to Poisson regression.

Log linear models for two and three dimensional contingency tables:interpretation of parameters, comparison with ANOVA and regression, ML estimation of parameters, likelihood ratio tests for various hypotheses including independence, marginal and conditional independence, partial association, models with quantitative levels.

Nonparametric regression and generalized linear models : interpolating and smoothing splines for simple regression , use of cross-validation, application to logistic and Poisson regression introduction to additive models and generalized additive models.

### REFERENCES

D. W. Hosmer and S. Lemeshow (2000) . Applied Logistic Regression, Second edition. Wiley, New York.

A. Agresti (1990). Categorical Data Analysis. Wiley, New York.

R. Christensen (1997). Log- linear Models and Logistic Regression, Second edition. Springer, New York.

P. McCullagh and J.A. Nelder (1999) . Generalized Linear Models, Second edition Chapman and Hall, New York.

P.J. Green and B. W. Silverman (1994). Nonparametric Regression and Generalized Linear Models. Chapman and Hall, New York.

T.J. Hastie and R. J. Tibshirani (1999) . Generalized additive Models. Second edition Chapman and Hall, New York..

## SECTION-B

Residuals and their analysis, influential observations, Power transformations for dependent and independent variables.

Robust and L-1 regression, Estimation of prediction error by cross-validation and boot strap.

Non-linear regression models, Different methods of estimation (Least squares, Maximum Likelihood). Asymptotic properties of estimators

Generalized linear models, Analysis of binary and grouped data by using logistic models, Log-linear models.

Random and mixed effect models, Maximum likelihood, MINQUE and restricted maximum likelihood estimators of variance components, Best linear unbiased predictors (BLUP). Growth curves.

### REFERENCES

Bates, D.M. and Watts, D.G. (1988). Nonlinear Regression Analysis and its Application, Wiley, New York.

Cook, R.D. and Weisberg, S. (1982). Residuals and inference in Regression, Chapman and Hall, London.

Draper, N.R. and Smith, H. (1998); Applied Regression Analysis, 3<sup>rd</sup> Ed. Wiley, New York.

Efron, B. and Tibsirani, J. R. (1993). An introduction to the Bootstrap, Chapman and Hall, New York

Kshirsagar, A. M. (1995). Growth Curves, Marcel and Dekker, New York.

McCullagh, P. and Nelder, J.A. (1989). Generalized Linear Models, 2<sup>nd</sup> Ed., Chapman and Hall, London,

Searle, S. R. ( 1987) . Linear. Models for Unbalanced Data, Wiley, New Yirk .

Seber, G. A. and Wild, G. J. (1989). Nonlinear Regression, Wiley, New York.

## SECTION-A

Types of biological assays; Direct assays; Ratio estimators, asymptotic distributions; Fieller's theorem

Regression approaches to estimating dose-response relationships Logit and probit approaches when dose –response curve for standard preparation is unknown; Quantal responses; Methods of estimation of parameters ; Estimation of extreme quantiles; Dose allocation schemes; polychotomous quantal response

Estimation of points on the quantal response function

Sequential procedures

Estimation of safe doses

Bayesian approach to bioassay

## REFERENCES

Z. Govindnarajulu (2000). Statistical Techniques in Bioassay. S. Kargar.

D. J. Finney (1971) . Statistical Method in Bioassay, Griffin.

D. J. Finney (1971). Probit Analysis ( 3<sup>rd</sup> Ed. ) . Griffin.

G. B. Weatherill (1966). Sequential Methods in Statistics. Methuen.

## SECTION – B

Basic biological concepts in genetics ( relevant to this course ).

Mendel's law. Hardy Weinberg equilibrium. Mating tables, estimation of allele frequency (dominant / co-dominant cases). Approach to equilibrium for X-linked gene, Natural selection. Mutation genetic drift, equilibrium when both natural selection and mutation are operative

Non-random mating, inbreeding phenotypic assortative mating.

Analysis of family data (a) Relative pair data. I. T . O matrices. Identity by descent (b) family data – estimation of segregation ratio under ascertainment bias (c) Pedigree data –Elston – Stewart algorithm for calculation of likelihoods Linkage. Estimation of recombination fraction, inheritance of quantitative traits Models and estimation 'o'

Sequence similarity, homology and alignment. Algorithms for (a) pairwise sequence alignment, (b) multiple sequence alignment, construction of phylogenetic trees, UPGMA, Neighbour joining, maximum parsimony and maximum likelihood algorithms.

### REFERENCES

C. C. Li (1976). First course on population genetics. Boxwood Press, California.

W. J. Ewens (1979) . Mathematical population genetics. Springer Verlag.

T. Nagylaki (1992) . Introduction to theoretical population genetics Springer Verlag.

R. Durbin, S. R. Eddy, A. Krogh , G. Mitchison (1998) . Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids.

### ADDITIONAL REFERENCES

R. C. Flandt