AGRICULTURAL STATISTICS

Course Structure - at a Glance

1. Syllabus of Service Courses for M. Sc. and Ph. D. programmes of other disciplines under Faculty of Agriculture and Faculty of Horticulture both

CODE	COURSE TITLE	CREDITS
STAT 500	ELEMENTARY STATISTICAL METHODS	3+1
STAT 501	STATISTICAL METHODS FOR RESEARCH WORKERS	2+1
STAT 502	STATISTICAL METHODS FOR BIOLOGY	2+1
STAT 503	EXPERIMENTAL DESIGN FOR RESEARCH WORKERS	2+1
STAT 504	STATISTICAL METHODS FOR SOCIAL SCIENCES	2+1
STAT 505	TIME SERIES ANALYSIS	2+1
STAT 506	LINEAR PROGRAMMING	2+1
STAT 507	ECONOMETRICS	2+1
STAT 508	BIOMETRICAL GENETICS	2+1

2. M. Sc. (Agricultural Statistics)

STAT 551	MATHEMATICAL METHODS-I	3+0
STAT 552	MATHEMATICAL METHODS-II	2+0
STAT 560	PROBABILITY THEORY	2+0
STAT 561	STATISTICAL METHODS	2+1
STAT 562	STATISTICAL INFERENCE	2+1
STAT 563	MULTIVARIATE ANALYSIS	2+1
STAT 564	DESIGN OF EXPERIMENTS	2+1
STAT 565	SAMPLING TECHNIQUES	2+1
STAT 566	STATISTICAL GENETICS	2+1
STAT 567	REGRESSION ANALYSIS	1+1
STAT 568	STATISTICAL COMPUTING	1+1
STAT 569	TIME SERIES ANALYSIS	1+1
STAT 570	ACTUARIAL STATISTICS	2+0
STAT 571	BIOINFORMATICS	2+0
STAT 572	ECONOMETRICS	2+0
STAT 573	STATISTICAL QUALITY CONTROL	2+0
STAT 574	OPTIMIZATION TECHNIQUES	1+1
STAT 575	DEMOGRAPHY	2+0
STAT 576	STATISTICAL METHODS FOR LIFE SCIENCES	2+0
STAT 577	STATISTICAL ECOLOGY	2+0
STAT 591	MASTER'S SEMINAR	1+0
STAT 599	MASTER'S RESEARCH	10+0

Note:

1. STAT 551 and STAT 552 are supporting courses. These are compulsory for all the students of Agricultural Statistics.
2. STAT 560 - STAT 569 are core courses to be taken by all the students of Agricultural Statistics.
3. STAT 591 and STAT 599 are compulsory for all the students of Agricultural Statistics.
4. A student has to take a minimum of 36 credits course work, excluding the supporting courses, seminar and research.

3. Ph. D. (Agricultural Statistics)

STAT 601	ADVANCED STATISTICAL COMPUTING	2+1
STAT 602	SIMULATION TECHNIQUES	1+1
STAT 611	ADVANCED STATISTICAL METHODS	2+0
STAT 612	ADVANCED STATISTICAL INFERENCE	3+0
STAT 613	ADVANCED DESIGN OF EXPERIMENTS	2+0
STAT 614	ADVANCED SAMPLING TECHNIQUES	2+0
STAT 615	ADVANCED STATISTICAL GENETICS	2+0
STAT 616	STATISTICAL MODELING	1+1
STAT 617	ADVANCED TIME SERIES ANALYSIS	2+0
STAT 618	STOCHASTIC PROCESSES	2+0
STAT 619	SURVIVAL ANALYSIS	2+0
STAT 620	ADVANCED BIOINFORMATICS	2+0
STAT 621	ADVANCED ECONOMETRICS	2+0
STAT 651	RECENT ADVANCES IN THE FIELD OF SPECIALIZATION	1+0
STAT 691	DOCTORAL SEMINAR I	1+0
STAT 692	DOCTORAL SEMINAR II	1+0
STAT 699	DOCTORAL RESEARCH	45+0

Note:

1. STAT 601 and STAT 602 are supporting courses. These are compulsory for all the students of Agricultural Statistics.
2. STAT 691, STAT 692, STAT 651 and STAT 699 are compulsory for all the students of Agricultural Statistics.
3. A student has to take a minimum of 18 credits course work, excluding the supporting courses, seminar and research.
4. A student has to take two seminars.

STAT 500 : Elementary Statistical Methods 3+1

(For those students who do not have sufficient statistical background)

Probability	:	Elementary concepts of probability; Addition theorem; Conditional Probability; Multiplication theory; Independence of events.
Statistical Methods	:	Population and its parameters; Sample and its statistics; Frequency distribution; Graphical representation; Measures of central tendency; Measures of dispersion; Moments; Simple correlation and regression.
Probability Distributions	:	Binomial; Poisson & Normal
Sample Survey	:	Elementary concept; Advantages of sample survey over census; Simple random sampling (SRS); SRSWR and SRSWOR; Drawing of random sample & estimation of average, total etc.; Sampling and non-sampling errors; Concept of stratified random sampling.
Design of Experiments	:	One way and two way classification (orthogonal); Principles of design; Uniformity trial and fertility contour map; Lay-out and analysis of CRD, RBD and LSD.
Tests of Significance	:	Hypotheses; Two types of errors; Exact small sample tests: z, t, χ^2 and F-tests.
Practicals	:	Based on above topics.

STAT 501 : Statistical Methods for Research Workers 2+1

Probability and Distribution	:	Preliminaries; Bayes' theorem; Repeated trials; Random variable- Mathematical expectation and its laws; variance, covariance etc.; Distribution: Binomial, Poisson, Normal.
Statistical Methods	:	Rank correlation; Correlation ratio; Intra-class correlation; Multiple Regression involving three variables; Multiple and partial correlation co-efficients; Stepwise multiple regression analysis; Concept of auto-correlation function (ACF).
Tests of Significance	:	t, F, χ^2 -tests and large sample tests; Confidence intervals; Transformation of Variables: Z-transformation.
Sample Survey	:	Stratified random sampling; Systematic sampling and cluster sampling.
Design of Experiments	:	LSD; Uses of repeated Latin squares; Missing plot techniques in RBD and LSD; Split-plot design; Multiple comparison tests.
Practicals	:	Based on above topics.

STAT 502: Statistical Methods for Biology 2+1

Probability and Distribution	:	Random variable and its expectation, variance etc., Binomial Poisson, Normal, Negative Binomial and Log normal distributions.
Statistical Methods	:	Multiple and partial correlation; Multiple regression; Reproduction and mortality rates and their estimation; Techniques for estimation of population number and growth.
Tests of Significance	:	Z, t, F and χ^2 -tests.
Design of Experiments	:	CRD, RBD and LSD and Split-plot design; Multiple comparison tests; Missing plot techniques in RBD and LSD; Elementary bio-assay and probit analysis.
Practicals	:	Based on above topics.

STAT 503: Experimental Design for Research Workers 2+1

Uniformity trails	:	Size and shape of pots and blocks; Lay-out and analysis of CRD and RBD; Use of Repeated LSD's; Efficiency of blocking; Missing plot techniques and analysis of covariance in RBD and LSD Multiple comparison tests.
Factorial Experiments	:	Interpretation of main effects and interaction; Orthogonality and partitioning of degrees of freedom; Analysis of 2^2, 2^3, 3^2 experiments; Concept of confounding and analysis of some confounded factorial experiments; Split plot and strip plot designs; Transformations; Analysis of groups of experiments.
Practicals	:	Based on above topics.

STAT 504: Statistical Methods for Social Sciences 2+1

Introduction	:	Frequency distribution; Principles governing their formation and standard distributions.
Concept of Sampling	:	SRS and stratified random sampling; Sampling and non-sampling errors and their remedial measures.
Tests of Significance	:	t, F, χ^2 -tests and large sample tests; Confidence intervals; Transformation of Variables; Z-transformation; Distribution-free statistics- run test, sign test; Wilcoxon sign-rank test, Mann-Whitney U-test; Wald – Wolfowitz run test; Median test etc.
Statistical Methods	:	Simple and multiple regression and prediction equations.
Application of Multivariate Analysis	:	Factor analysis, Cluster analysis; Discriminant function and D^2 statistics; Principal component analysis.
Praticals	:	Based on above topics.

STAT 505: Time series Analysis 2+1

Time series and its components; Estimation/Elimination of different components; Variate-difference method. Concept of harmonic analysis; Correlogram and periodogram analysis; Introduction of spectral analysis; Economic application of multivariate time serice; Forecasting: Concept and different methods.

Practical : Based on above topics

STAT 506: Linear Programming 2+1

Convex sets; Programming problems; Graphical methods and simplex method for solution; Duality in liner programming; Revised simplex and dual simplex method; Transportation and assignment problems. Introduction to integer programming; Quadratic programming and their application and uses. Elements of game theory; two person-zero-sum game; Relationship between game theory and linear programming.

Practical : Based on above topics

STAT 507: Econometrics 2+1

Classical Linear Regression Models	:	Assumptions; BLUE and least square estimates and their properties; Prediction probems.
Autocorrelation and Heteroscedasticity	:	Definition; Causes of such problems and their remedies.
Multicollinearity	:	The meaning and consequences of its existence; Tests of identifying the existence of multicollinearity; Remedies necessary for analysis.
Simultaneous Equation Models	:	Definitions; OLS; ILS, 2 SLS methods and their applications.
Praticals	:	Based on above topics.

STAT 508: Biometrical Genetics 2+1

Estimation of linkage from back-cross; F2 and F3 data using method of MLE and other methods; Disturbed segregation; Estimation of additive genetic dominance and environmental components of variation; Plant Breeding trials and their use in the estimation of genetic variation and variability; Simple ideas of discriminant function for plant selection. Path analysis; Genotypic correlation; Path coefficients. North – Carolina Mating Designs; NC Design I, II, III, Line X Tester design

STAT 551: MATHEMATICAL METHODS - I 3+0

Objective

This course lays the foundation of all other courses of Statistics/ Agricultural Statistics discipline by preparing them to understand the importance of mathematical methods in research. The students would be exposed to the basic mathematical tools of real analysis, calculus

differential equations and numerical analysis. This would prepare them to study their main courses that involve knowledge of Mathematics.

Theory

UNIT I

Real Analysis: Convergence and divergence of infinite series, use of comparison tests -D'Alembert's Ratio - test, Cauchy's nth root test, Raabe's test, Kummer's test, Gauss test. Absolute and conditional convergence. Riemann integration, concept of Lebesgue integration, power series, Fourier, Laplace and Laplace -Steiltjes' transformation, multiple integrals.

UNIT II

Calculus: Limit and continuity, differentiation of functions, successive differentiation, partial differentiation, mean value theorems, Taylor and Maclaurin's series. Application of derivatives, L'hospital's rule. Integration of rational, irrational and trigonometric functions. Application of integration.

UNIT III

Differential equation: Differential equations of first order, linear differential equations of higher order with constant coefficient.

UNIT IV

Numerical Analysis: Simple interpolation, Divided differences, Numerical differentiation and integration.

Suggested Readings

Bartle RG. 1976. *Elements of Real Analysis.* John Wiley.

Chatterjee SK. 1970. *Mathematical Analysis.* Oxford & IBH.

Gibson GA. 1954. *Advanced Calculus*. Macmillan.

Henrice P. 1964. *Elements of Numerical Analysis*. John Wiley.

Hildebrand FB. 1956. *Introduction to Numerical Analysis*. Tata McGraw Hill.

Priestley HA. 1985. *Complex Analysis.* Clarenton Press.

Rudin W. 1985. *Principles of Mathematical Analysis*. McGraw Hill.

Sauer T. 2006. *Numerical Analysis With CD-Rom.* Addison Wesley.

Scarborough JB. 1976. *Numerical Mathematical Analysis*. Oxford & IBH.

Stewart J. 2007. *Calculus.* Thompson.

Thomas GB Jr. & Finney RL. 1996. Calculus. 9th Ed. Pearson Edu.

STAT 552/ BST 552 **MATHEMATICAL METHODS - II** **2+0**

Objective

This is another course that supports all other courses in Statistics / Agricultural Statistics. The students would be exposed to the advances in Linear Algebra and Matrix theory. This would prepare them to study their main courses that involve knowledge of Linear Algebra and Matrix Algebra.

Theory

UNIT I

Linear Algebra: Group, ring, field and vector spaces, Sub-spaces, basis, Gram Schmidt's orthogonalization, Galois field - Fermat's theorem and

primitive elements. Linear transformations. Graph theory: Concepts and applications

UNIT II

Matrix Algebra: Basic terminology, linear independence and dependence of vectors. Row and column spaces, Echelon form. Determinants, rank and inverse of matrices. Special matrices - idempotent, symmetric, orthogonal. Eigen values and eigen vectors. Spectral decomposition of matrices

UNIT III

Unitary, Similar, Hadamard, Circulant, Helmert's matrices. Kronecker and Hadamard product of matrices, Kronecker sum of matrices. Sub-matrices and partitioned matrices, Permutation matrices, full rank factorization, Grammian root of a symmetric matrix. Solutions of linear equations, Equations having many solutions.

UNIT IV

Generalized inverses, Moore-Penrose inverse, Applications of g-inverse. Spectral decomposition of matrices, Inverse and Generalized inverse of partitioned matrices, Differentiation and integration of matrices, Quadratic forms.

Suggested Readings

Aschbacher M. 2000. *Finite Group Theory*. Cambridge University Press.

Deo N. 1984. *Graph Theory with Application to Engineering and Computer Science*. Prentice Hall of India.

Gentle JE. 2007. *Matrix Algebra: Theory, Computations and Applications in Statistics*. Springer.

Graybill FE.1961. *Introduction to Matrices with Applications in Statistics*. Wadsworth Publ.

Hadley G. 1969. *Linear Algebra*. Addison Wesley.

Harville DA. 1997. *Matrix Algebra from a Statistician's Perspective*. Springer.

Rao CR. 1965. *Linear Statistical Inference and its Applications*. 2nd Ed. John Wiley.

Robinson DJS. 1991. *A Course in Linear Algebra with Applications*. World Scientific.

Searle SR. 1982. *Matrix Algebra Useful for Statistics*. John Wiley.

Seber GAF. 2008. *A Matrix Handbook for Statisticians*. John Wiley.

STAT 560/ BST 560 **PROBABILITY THEORY** **2+0**

Objective

This is a fundamental course in Statistics. This course lays the foundation of probability theory, random variable, probability distribution, mathematical expectation, etc. which forms the basis of basic statistics. The students are also exposed to law of large numbers and central limit theorem. The students also get introduced to stochastic processes.

Theory

UNIT I

Basic concepts of probability. Elements of measure theory: class of sets, field, sigma field, minimal sigma field, Borel sigma field in R, measure, probability measure. Axiomatic approach to probability. Properties of probability based on axiomatic definition. Addition and multiplication

theorems. Conditional probability and independence of events. Bayes theorem.

UNIT II

Random variables: definition of random variable, discrete and continuous, functions of random variables. Probability mass function and Probability density function, Distribution function and its properties. Notion of bivariate random variables, bivariate distribution function and its properties. Joint, marginal and conditional distributions. Independence of random variables. Transformation of random variables (two dimensional case only).

Mathematical expectation: Mathematical expectation of functions of a random variable. Raw and central moments and their relation, covariance, skewness and kurtosis. Addition and multiplication theorems of expectation. Definition of moment generating function, cumulating generating function, probability generating function and statements of their properties.

UNIT III

Conditional expectation and conditional variance. Characteristic function and its properties. Inversion and uniqueness theorems. Functions, which cannot be characteristic functions.

Chebyshev, Markov, Cauchy-Schwartz, Jenson, Liapounov, holder's and Minkowsky's inequalities. Sequence of random variables and modes of convergence (convergence in distribution, in probability, almost surely, and quadratic mean) and their interrelations. Statement of Slutsky's theorem. Borel -Cantelli lemma and Borel 0-1 law.

UNIT IV

Laws of large numbers: WLLN, Bernoulli and Kintchin's WLLN. Kolmogorov inequality, Kolmogorov's SLLNs.

Central Limit theorems: Demoviere- Laplace CLT, Lindberg - Levy CLT, Liapounov CLT, Statement of Lindeberg-Feller CLT and simple applications. Definition of quantiles and statement of asymptotic distribution of sample quantiles.

UNIT V

Classification of Stochastic Processes, Examples. Markov Chain and classification of states of Markov Chain.

Suggested Readings

Ash RB. 2000. *Probability and Measure Theory*. 2nd Ed. Academic Press.

Billingsley P. 1986. *Probability and Measure*. 2nd Ed. John Wiley.

Capinski M & Zastawniah. 2001. *Probability Through Problems*. Springer.

Dudewicz EJ & Mishra SN. 1988. *Modern Mathematical Statistics*. John Wiley.

Feller W. 1972. *An Introduction to Probability Theory and its Applications*. Vols. I., II. John Wiley.

Loeve M. 1978. *Probability Theory*. 4 th Ed. Springer.

Marek F. 1963. *Probability Theory and Mathematical Statistics*. John Wiley.

Rohatgi VK & Saleh AK Md. E. 2005. *An Introduction to Probability and Statistics*. 2nd Ed. John Wiley.

STAT 561 / BST 561 **STATISTICAL METHODS** **2+1**

Objective

This course lays the foundation of probability distributions and sampling distributions and their application which forms the basis of Statistical Inference. Together with probability theory, this course is fundamental to the discipline of Statistics. The students are also exposed to correlation and regression, and order statistics and their distributions. Categorical data analysis is also covered in this course.

Theory

UNIT I

Descriptive statistics: probability distributions: Discrete probability distributions ~ Bernoulli, Binomial, Poisson, Negative-binomial, Geometric and Hyper Geometric, uniform, multinomial ~ Properties of these distributions and real life examples. Continuous probability distributions ~ rectangular, exponential, Cauchy, normal, gamma, beta of two kinds, Weibull, lognormal, logistic, Pareto. Properties of these distributions. Probability distributions of functions of random variables.

UNIT II

Concepts of compound, truncated and mixture distributions (definitions and examples). Pearsonian curves and its various types. Sampling distributions of sample mean and sample variance from Normal population, central and non-central chi-Square, *t* and *F* distributions, their properties and inter relationships.

UNIT III

Concepts of random vectors, moments and their distributions. Bivariate Normal distribution - marginal and conditional distributions. Distribution of quadratic forms. Cochran theorem. Correlation, rank correlation, correlation ratio and intra-class correlation. Regression analysis, partial and multiple correlation and regression.

UNIT IV

Sampling distribution of correlation coefficient, regression coefficient, correlation ratio, intra class correlation coefficient. Categorical data analysis - loglinear models, Association between attributes. Variance Stabilizing Transformations.

UNIT V

Order statistics, distribution of *r*-th order statistics, joint distribution of several order statistics and their functions, marginal distributions of order statistics, distribution of range, median, etc.

Practical

Fitting of discrete distributions and test for goodness of fit; Fitting of continuous distributions and test for goodness of fit; Fitting of truncated distribution; Computation of simple, multiple and partial correlation coefficient, correlation ratio and intra-class correlation; Regression coefficients and regression equations; Fitting of Pearsonian curves; Analysis of association between attributes, categorical data and log-linear models.

Suggested Readings

Agresti A. 2002. *Categorical Data Analysis.* 2nd Ed. John Wiley.

Arnold BC, Balakrishnan N & Nagaraja HN. 1992. *A First Course in Order Statistics*. John Wiley.

David HA & Nagaraja HN. 2003. *Order Statistics*. 3rd Ed. John Wiley.

Dudewicz EJ & Mishra SN. 1988. *Modern Mathematical Statistics*. John Wiley.

Huber PJ. 1981. *Robust Statistics*. John Wiley.

Johnson NL, Kotz S & Balakrishnan N. 2000. *Continuous Univariate Distributions*. John Wiley.

Johnson NL, Kotz S & Balakrishnan N. 2000. *Discrete Univariate Distributions*. John Wiley.

Marek F. 1963. *Probability Theory and Mathematical Statistics*. John Wiley.

Rao CR. 1965. *Linear Statistical Inference and its Applications*. John Wiley.

Rohatgi VK & Saleh AK Md. E. 2005. *An Introduction to Probability and Statistics*. 2nd Ed. John Wiley.

STAT 562/ BST 562 **STATISTICAL INFERENCE** **2+1**

Objective

This course lays the foundation of Statistical Inference. The students would be taught the problems related to point and confidence interval estimation and testing of hypothesis. They would also be given the concepts of nonparametric and sequential test procedures and elements of decision theory.

Theory

UNIT I

Concepts of point estimation: MSE, unbiasedness, consistency, efficiency and sufficiency. Statement of Neyman's Factorization theorem with applications. MVUE, Rao-Blackwell theorem, completeness, Lehmann-Scheffe theorem. Fisher information, Cramer-Rao lower bound and its applications.

UNIT II

Moments, minimum chi-square, least square and maximum likelihood methods of estimation and statements of their properties. Interval estimation-Confidence level, CI using pivots and shortest length CI. CI for the parameters of Normal, Exponential, Binomial and Poisson distributions.

UNIT III

Fundamental notions of hypothesis testing-statistical hypothesis, statistical test, critical region, types of errors, test function, randomized and non-randomized tests, level of significance, power function, most powerful tests: Neyman-Pearson fundamental lemma, MLR families and UMP tests for one parameter exponential families. Concepts of consistency, unbiasedness and invariance of tests. Likelihood Ratio tests, statement of asymptotic properties of LR tests with applications (including homogeneity of means and variances).Relation between confidence interval estimation and testing of hypothesis.

UNIT IV

Notions of sequential vs fixed sample size techniques. Wald's SPRT for testing simple null hypothesis vs simple alternative. Termination property

of SPRT, SPRT for Binomial, Poisson, Normal and Exponential distributions. Concepts of loss, risk and decision functions, admissible and optimal decision functions, estimation and testing viewed as decision problems, conjugate families, Bayes and Minimax decision functions with applications to estimation with quadratic loss.

UNIT V

Non-parametric tests: Sign test, Wilcoxon signed rank test, Runs test for randomness, Kolmogorov - Smirnov test for goodness of fit, Median test and Wilcoxon-Mann-Whitney U-test. Chi-square test for goodness of fit and test for independence of attributes. Kruskal -Wallis and Friedman's tests. Spearman's rank correlation and Kendall's Tau tests for independence.

Practical

Methods of estimation - Maximum Likelihood, Minimum χ^2 and Moments; Confidence Interval Estimation; MP and UMP tests; Large Sample tests; Non-parametric tests, Sequential Probability Ratio Test; Decision functions.

Suggested Readings

Box GEP & Tiao GC. 1992. *Bayesian Inference in Statistical Analysis*. John Wiley.

Casela G & Berger RL. 2001. *Statistical Inference*. Duxbury Thompson Learning.

Christensen R. 1990. *Log Linear Models*. Springer.

Conover WJ. 1980. *Practical Nonparametric Statistics*. John Wiley.

Dudewicz EJ & Mishra SN. 1988. *Modern Mathematical Statistics*. John Wiley.

Gibbons JD. 1985. *Non Parametric Statistical Inference*. 2nd Ed. Marcel Dekker.

Kiefer JC. 1987. *Introduction to Statistical Inference*. Springer.

Lehmann EL. 1986. *Testing Statistical Hypotheses*. John Wiley.

Lehmann EL. 1986. *Theory of Point Estimation*. John Wiley.

Randles RH & Wolfe DS. 1979. *Introduction to the Theory of Nonparametric Statistics*. John Wiley.

Rao CR. 1973. *Linear Statistical Inference and its Applications*. 2nd Ed. John Wiley.

Rohatgi VK & Saleh AK. Md. E. 2005. *An Introduction to Probability and Statistics*. 2nd Ed. John Wiley.

Rohtagi VK. 1984. *Statistical Inference*. John Wiley

Sidney S & Castellan NJ Jr. 1988. *Non Parametric Statistical Methods for Behavioral Sciences*. McGraw Hill.

Wald A. 2004. *Sequential Analysis*. Dover Publ.

STAT 563 / BST 563 **MULTIVARIATE ANALYSIS** **2+1**

Objective

This course lays the foundation of Multivariate data analysis. Most of the data sets in agricultural sciences are multivariate in nature. The exposure provided to multivariate data structure, multinomial and multivariate normal distribution, estimation and testing of parameters, various data

reduction methods would help the students in having a better understanding of agricultural research data, its presentation and analysis.

Theory

UNIT I

Concept of random vector, its expectation and Variance-Covariance matrix. Marginal and joint distributions. Conditional distributions and Independence of random vectors. Multinomial distribution. Multivariate Normal distribution, marginal and conditional distributions. Sample mean vector and its distribution. Maximum likelihood estimates of mean vector and dispersion matrix. Tests of hypothesis about mean vector.

UNIT II

Wishart distribution and its simple properties. Hotelling's T^2 and Mahalanobis D^2 statistics. Null distribution of Hotelling's T^2. Rao's U statistics and its distribution.

Wilks' λ criterion and statement of its properties. Concepts of discriminant analysis, computation of linear discriminant function, classification between k (≥ 2) multivariate normal populations based on LDF and Mahalanobis D^2.

UNIT III

Principal Component Analysis, factor analysis (simple and multi factor models). Canonical variables and canonical correlations. Cluster analysis, similarities and dissimilarities, Hierarchical clustering. Single and Complete linkage methods.

UNIT IV

Path analysis and computation of path coefficients, introduction to multidimensional scaling, some theoretical results, similarities, metric and non metric scaling methods. Concepts of analysis of categorical data.

Practical

Maximum likelihood estimates of mean-vector and dispersion matrix; Testing of hypothesis on mean vectors of multivariate normal populations; Cluster analysis, Discriminant function, Canonical correlation, Principal component analysis, Factor analysis; Multivariate analysis of variance and covariance, multidimensional scaling.

Suggested Readings

Anderson TW. 1984. *An Introduction to Multivariate Statistical Analysis*. 2nd Ed. John Wiley.

Arnold SF. 1981. *The Theory of Linear Models and Multivariate Analysis*. John Wiley.

Giri NC. 1977. *Multivariate Statistical Inference*. Academic Press.

Johnson RA & Wichern DW. 1988. *Applied Multivariate Statistical Analysis*. Prentice Hall.

Kshirsagar AM. 1972. *Multivariate Analysis*. Marcel Dekker.

Muirhead RJ. 1982. *Aspects of Multivariate Statistical Theory*. John Wiley.

Rao CR. 1973. *Linear Statistical Inference and its Applications*. 2nd Ed. John Wiley.

Rencher AC. 2002. *Methods of Multivariate Analysis*. 2nd Ed. John Wiley.

Srivastava MS & Khatri CG. 1979. *An Introduction to Multivariate Statistics*. North Holland.

STAT 564/ BST 564 **DESIGN OF EXPERIMENTS** **2+1**

Objective

Design of Experiments provides the statistical tools to get maximum information from least amount of resources. This course is meant to expose the students to the basic principles of design of experiments. The students would also be provided with mathematical background of various basic designs involving one-way and two way elimination of heterogeneity and their characterization properties. This course would also prepare the students in deriving the expressions for analysis of experimental data.

Theory

UNIT I

Elements of linear estimation, Gauss Markoff Theorem, relationship between BLUEs and linear zero-functions. Aitken's transformation, test of hypothesis, analysis of variance, partitioning of degrees of freedom.

UNIT II

Orthogonality, contrasts, mutually orthogonal contrasts, analysis of covariance; Basic principles of design of experiments, uniformity trials, size and shape of plots and blocks.

UNIT III

Basic designs - completely randomized design, randomized complete block design and Latin square design; orthogonal Latin squares, mutually orthogonal Latin squares (MOLS), Youden square designs, Graeco Latin squares.

UNIT IV

Balanced incomplete block (BIB) designs - general properties and analysis without and with recovery of intra block information, construction of BIB designs. Partially balanced incomplete block designs with two associate classes - properties, analysis and construction, Lattice designs, alpha designs, cyclic designs, augmented designs, general analysis of block designs.

UNIT V

Factorial experiments, confounding in symmetrical factorial experiments (2^n and 3^n series), partial and total confounding, fractional factorials, asymmetrical factorials.

UNIT VI

Designs for fitting response surface; Cross-over designs. Missing plot technique; Split plot and Strip plot design; Groups of experiments; Sampling in field experiments.

Practical

Determination of size and shape of plots and blocks from uniformity trials data; Analysis of data generated from completely randomized design, randomized complete block design; Latin square design, Youden square design; Analysis of data generated from a BIB design, lattice design, PBIB designs; $2^n, 3^n$ factorial experiments without and with confounding; Split and strip plot designs, repeated measurement design; Missing plot techniques, Analysis of covariance; Analysis of Groups of experiments, Analysis of clinical trial experiments. Sampling in field experiments.

Suggested Readings

Chakrabarti MC. 1962. *Mathematics of Design and Analysis of Experiments*. Asia Publ. House.

Cochran WG & Cox DR. 1957. *Experimental Designs*. 2nd Ed. John Wiley.

Dean AM & Voss D. 1999. *Design and Analysis of Experiments*. Springer.

Dey A & Mukerjee R. 1999. *Fractional Factorial Plans*. John Wiley.

Dey A 1986. *Theory of Block Designs*. Wiley Eastern.

Hall M Jr. 1986. *Combinatorial Theory*. John Wiley.

John JA & Quenouille MH. 1977. *Experiments: Design and Analysis*. Charles & Griffin.

Kempthorne, O. 1976. Design and Analysis of Experiments. John Wiley.

Khuri AI & Cornell JA. 1996. *Response Surface Designs and Analysis*. 2nd Ed. Marcel Dekker.

Kshirsagar AM 1983. *A Course in Linear Models*. Marcel Dekker.

Montgomery DC. 2005. *Design and Analysis of Experiments*. John Wiley.

Raghavarao D. 1971. *Construction and Combinatorial Problems in Design of Experiments*. John Wiley.

Searle SR. 1971. *Linear Models*. John Wiley.

Street AP & Street DJ. 1987. *Combinatorics of Experimental Designs*. Oxford Science Publ.

Design Resources Server. *Indian Agricultural Statistics Research Institute(ICAR), New Delhi-110012, India*. www.iasri.res.in/design.

STAT 565 / BST 565 **SAMPLING TECHNIQUES** **2+1**

Objective

This course is meant to expose the students to the techniques of drawing representative samples from various populations and then preparing them on the mathematical formulations of estimating the population parameters based on the sample data. The students would also be exposed to the real life applications of sampling techniques and estimation of parameters.

Theory

UNIT I

Sample survey vs complete survey, probability sampling, sample space, sampling design, sampling strategy; Inverse sampling; Determination of sample size; Confidence-interval; Simple random sampling, Estimation of population proportion, Stratified random sampling, Number of strata and optimum points of stratification.

UNIT II

Ratio and regression methods of estimation, Cluster sampling, Systematic sampling, Multistage sampling with equal probability, Separate and combined ratio estimator, Double sampling, Successive sampling -two occasions**.**

UNIT III

Non-sampling errors - sources and classification, Non-response in surveys, Imputation methods, Randomized response techniques, Response errors - interpenetrating sub-sampling.

UNIT IV

Sampling with varying probabilities with and without replacement, PPS sampling, Cumulative method and Lahiri's method of selection, Horvitz-

Thompson estimator, Ordered and unordered estimators, Sampling strategies due to Midzuno-Sen and Rao-Hartley-Cochran. Inclusion probability proportional to size sampling, PPS systematic sampling, Multistage sampling with unequal probabilities, Self weighting design PPS sampling.

UNIT V

Unbiased ratio and regression type estimators, Multivariate ratio and regression type of estimators, Design effect, Bernoulli and Poisson sampling.

Practical

Determination of sample size and selection of sample; Simple random sampling, Inverse sampling, Stratified random sampling, Cluster sampling, systematic sampling; Ratio and regression methods of estimation; Double sampling, multi-stage sampling, Imputation methods; Randomized response techniques; Sampling with varying probabilities.

Suggested Readings

Cassel CM, Sarndal CE & Wretman JH. 1977. *Foundations of Inference in Survey Sampling*. John Wiley.

Chaudhari A & Stenger H. 2005. *Survey Sampling Theory and Methods*. 2nd Ed. Chapman & Hall.

Chaudhari A & Voss JWE. 1988. *Unified Theory and Strategies of Survey Sampling*. North Holland.

Cochran WG. 1977. *Sampling Techniques*. John Wiley.

Hedayat AS & Sinha BK. 1991. *Design and Inference in Finite Population Sampling*. John Wiley.

Kish L. 1965. *Survey Sampling*. John Wiley.

Murthy MN. 1977. *Sampling Theory and Methods*. 2nd Ed. Statistical Publ. Society, Calcutta.

Raj D & Chandhok P. 1998. *Sample Survey Theory*. Narosa Publ.

Sarndal CE, Swensson B & Wretman J. 1992. *Models Assisted Survey Sampling*. Springer.

Sukhatme PV, Sukhatme BV, Sukhatme S & Asok C. 1984. *Sampling Theory of Surveys with Applications*. Iowa State University Press and Indian Society of Agricultural Statistics, New Delhi.

Thompson SK. 2000. *Sampling*. John Wiley.

STAT 566 / BST 566 **STATISTICAL GENETICS** **2+1**

Objective

This course is meant to prepare the students in applications of statistics in quantitative genetics and breeding. The students would be exposed to the physical basis of inheritance, detection and estimation of linkage, estimation of genetic parameters and development of selection indices.

Theory

UNIT I

Physical basis of inheritance. Analysis of segregation, detection and estimation of linkage for qualitative characters. Amount of information about linkage, combined estimation, disturbed segregation.

UNIT II

Gene and genotypic frequencies, Random mating and Hardy -Weinberg law, Application and extension of the equilibrium law, Fisher's fundamental theorem of natural selection. Disequilibrium due to linkage for two pairs of genes, sex-linked genes, Theory of path coefficients.

UNIT III

Concepts of inbreeding, Regular system of inbreeding. Forces affecting gene frequency - selection, mutation and migration, equilibrium between forces in large populations, Random genetic drift, Effect of finite population size.

UNIT IV

Polygenic system for quantitative characters, concepts of breeding value and dominance deviation. Genetic variance and its partitioning, Effect of inbreeding on quantitative characters, Multiple allelism in continuous variation, Sex-linked genes, Maternal effects - estimation of their contribution.

UNIT V

Correlations between relatives, Heritability, Repeatability and Genetic correlation. Response due to selection, Selection index and its applications in plants and animals improvement programmes, Correlated response to selection.

UNIT VI

Restricted selection index. Variance component approach and linear regression approach for the analysis of GE interactions. Measurement of stability and adaptability for genotypes. Concepts of general and specific combining ability. Diallel and partial diallel crosses - construction and analysis.

Practical

Test for the single factor segregation ratios, homogeneity of the families with regard to single factor segregation; Detection and estimation of linkage parameter by different procedures; Estimation of genotypic and gene frequency from a given data. Hardy-Weinberg law; Estimation of changes in gene frequency due to systematic forces, inbreeding coefficient, genetic components of variation, heritability and repeatability coefficient, genetic correlation coefficient; Examination of effect of linkage, epistasis and inbreeding on mean and variance of metric traits; Mating designs; Construction of selection index including phenotypic index, restricted selection index. Correlated response to selection.

Suggested Readings

Bailey NTJ. 1961. *The Mathematical Theory of Genetic Linkage*. Clarendon Press.

Balding DJ, Bishop M & Cannings C. 2001. *Hand Book of Statistical Genetics*. John Wiley.

Crow JF & Kimura M. 1970. *An Introduction of Population Genetics Theory*. Harper & Row.

Dahlberg G. 1948. *Mathematical Methods for Population Genetics*. Inter Science Publ.

East EM & Jones DF. 1919. *Inbreeding and Outbreeding*. J B Lippincott.

Ewens WJ. 1979. *Mathematics of Population Genetics*. Springer.

Falconer DS. 1985. *Introduction to Quantitative Genetics*. ELBL.

Fisher RA. 1949. *The Theory of Inbreeding*. Oliver & Boyd.
Fisher RA. 1950. *Statistical Methods for Research Workers.* Oliver & Boyd.
Fisher RA. 1958. *The Genetical Theory of Natural Selection.* Dover Publ.
Kempthorne O. 1957. *An Introduction to Genetic Statistics*. The Iowa State Univ. Press.
Lerner IM. 1950. *Population Genetics and Animal Improvement*. Cambridge Univ. Press.
Lerner IM. 1954. *Genetic Homeostasis*. Oliver & Boyd.
Lerner IM. 1958. *The Genetic Theory of Selection*. John Wiley.
Li CC. 1982. *Population Genetics*. The University of Chicago Press.
Mather K & Jinks JL. 1977. *Introduction to Biometrical Genetics*. Chapman & Hall.
Mather K & Jinks JL. 1982. *Biometrical Genetics*. Chapman & Hall.
Mather K. 1949. *Biometrical Genetics*. Methuen.
Mather K. 1951. *The Measurement of Linkage in Heredity*. Methuen.
Narain P. 1990. *Statistical Genetics*. Wiley Eastern.

STAT 567 / BST 567 **REGRESSION ANALYSIS** **1+1**

Objective

This course is meant to prepare the students in linear and non-linear regression methods useful for statistical data analysis. They would also be provided a mathematical foundation behind these techniques and their applications in agricultural data.

Theory

UNIT I

Simple and Multiple linear regressions: Least squares fit, Properties and examples. Polynomial regression: Use of orthogonal polynomials.

UNIT II

Assumptions of regression; diagnostics and transformations; Examination of residuals ~ Studentized residuals, applications of residuals in detecting outliers, identification of influential observations. Lack of fit, Pure error. Testing homoscedasticity and normality of errors, Durbin-Watson test. Use of R^2 for examining goodness of fit.

UNIT III

Concepts of Least median of squares and its applications; Concept of multicollinearity, Analysis of multiple regression models, estimation and testing of regression parameters, sub-hypothesis testing, restricted estimation.

UNIT IV

Weighted least squares method: Properties, and examples. Box-Cox family of transformations. Use of dummy variables, Selection of variables: Forward selection, Backward elimination. Stepwise and Stagewise regressions.

UNIT V

Introduction to non-linear models, nonlinear estimation: Least squares for nonlinear models.

Practical

Multiple regression fitting with three and four independent variables; Estimation of residuals, their applications in outlier detection, distribution of residuals; Test of homoscedasticity, and normality, Box-Cox transformation; Restricted estimation of parameters in the model, hypothesis testing, Step wise regression analysis; Least median of squares norm, Orthogonal polynomial fitting.

Suggested Readings

Barnett V & Lewis T. 1984. *Outliers in Statistical Data*. John Wiley.

Belsley DA, Kuh E & Welsch RE. 2004. *Regression Diagnostics- Identifying Influential Data and Sources of Collinearity*. John Wiley.

Chatterjee S, Hadi A & Price B. 1999. *Regression Analysis by Examples.* John Wiley.

Draper NR & Smith H. 1998. *Applied Regression Analysis*. 3 Ed.rdJohn Wiley.

McCullagh P & Nelder JA. 1999. *Generalized Linear Models*. 2 nd Ed. Chapman & Hall.

Montgomery DC, Peck EA & Vining GG. 2003. *Introduction to Linear Regression Analysis*. 3 Ed. John Wiley.

Rao CR. 1973. *Linear Statistical Inference and its Applications*. 2 nd Ed. John Wiley.

STAT 568/ BST 568 **STATISTICAL COMPUTING** **1+1**

Objective

This course is meant for exposing the students in the concepts of computational techniques. Various statistical packages would be used for teaching the concepts of computational techniques.

Theory

UNIT I

Introduction to statistical packages and computing: data types and structures, pattern recognition, classification, association rules, graphical methods. Data analysis principles and practice

UNIT II

ANOVA, regression and categorical data methods; model formulation, fitting, diagnostics and validation; Matrix computations in linear models. Analysis of discrete data.

UNIT III

Numerical linear algebra, numerical optimization, graphical techniques, numerical approximations, numerical integration and Monte Carlo methods.

UNIT IV

Spatial statistics; spatial sampling; hierarchical modeling. Analysis of cohort studies, case-control studies and randomized clinical trials, techniques in the analysis of survival data and longitudinal studies, Approaches to handling missing data, and meta-analysis.

Practical

Data management, Graphical representation of data, Descriptive statistics; General linear models ~ fitting and analysis of residuals, outlier detection;

Categorical data analysis, analysis of discrete data, analysis of binary data; Numerical algorithms; Spatial modeling, cohort studies; Clinical trials, analysis of survival data; Handling missing data.

Suggested Readings

Agresti A. 2002. *Categorical Data Analysis.* 2nd Ed. John Wiley.

Everitt BS & Dunn G. 1991. *Advanced Multivariate Data Analysis*. 2nd Ed. Arnold.

Geisser S. 1993. *Predictive Inference: An Introduction.* Chapman & Hall.

Gelman A & Hill J. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge Univ. Press.

Gentle JE, Härdle W & Mori Y. 2004. *Handbook of Computational Statistics - Concepts and Methods*. Springer.

Han J & Kamber M. 2000. *Data Mining: Concepts and Techniques.* Morgan.

Hastie T, Tibshirani R & Friedman R. 2001. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer.

Kennedy WJ & Gentle JE. 1980. *Statistical Computing*. Marcel Dekker.

Miller RG Jr. 1986. *Beyond ANOVA, Basics of Applied Statistics*. John Wiley.

Rajaraman V. 1993. *Computer Oriented Numerical Methods*. Prentice-Hall.

Ross S. 2000. *Introduction to Probability Models.* Academic Press.

Ryan BF & Joiner BL. 1994. *MINITAB Handbook*. 3rd Ed. Duxbury Press.

Simonoff JS. 1996. *Smoothing Methods in Statistics*. Springer.

Snell EJ. 1987. *Applied Statistics: A Handbook of BMDP Analyses*. Chapman & Hall.

Thisted RA. 1988. *Elements of Statistical Computing*. Chapman & Hall.

Venables WN & Ripley BD. 1999. *Modern Applied Statistics With S-Plus.* 3rd Ed. Springer.

STAT 569 TIME SERIES ANALYSIS 1+1

Objective

This course is meant to teach the students the concepts involved in time series data. They would also be exposed to components of time series, stationary models and forecasting/ projecting the future scenarios based on time series data. It would also help them in understanding the concepts involved in time series data presentation, analysis and interpretation.

Theory

UNIT I

Components of a time-series. Autocorrelation and Partial autocorrelation functions, Correlogram and periodogram analysis.

UNIT II

Linear stationary models: Autoregressive, Moving average and Mixed processes. Linear non-stationary models: Autoregressive integrated moving average processes.

UNIT III

Forecasting: Minimum mean square forecasts and their properties, Calculating and updating forecasts.

UNIT IV

Model identification: Objectives, Techniques, and Initial estimates. Model estimation: Likelihood function, Sum of squares function, Least squares

estimates. Seasonal models. Intervention analysis models and Outlier detection.

Practical

Time series analysis, autocorrelations, correlogram and periodogram; Linear stationary model; Linear non-stationary model; Model identification and model estimation; Intervention analysis and outliers detection.

Suggested Readings

Box GEP, Jenkins GM & Reinsel GC. 2007. *Time Series Analysis: Forecasting and Control*. 3rd Ed. Pearson Edu.

Brockwell PJ & Davis RA. 2002. *Introduction to Time Series and Forecasting*. 2nd Ed. Springer.

Chatterjee S, Hadi A & Price B.1999. *Regression* Analysis by Examples. John Wiley.

Draper NR & Smith H. 1998. *Applied Regression Analysis*. 3rd Ed. John Wiley.

Johnston J. 1984. *Econometric Methods*. McGraw Hill.

Judge GG, Hill RC, Griffiths WE, Lutkepohl H & Lee TC. 1988. *Introduction to the Theory and Practice of Econometrics*. 2nd Ed. John Wiley.

Montgomery DC & Johnson LA. 1976. *Forecasting and Time Series Analysis*. McGraw Hill.

Shumway RH & Stoffer DS. 2006. *Time Series Analysis and its Applications: With R Examples*. 2nd Ed. Springer.

STAT 570 / BST 570 **ACTUARIAL STATISTICS** **2+0**

Objective

This course is meant to expose to the students to the statistical techniques such as probability models, life tables, insurance and annuities. The students would also be exposed top practical applications of these techniques in computation of premiums that include expenses, general expenses, types of expenses and per policy expenses.

Theory

UNIT I

Insurance and utility theory, models for individual claims and their sums, survival function, curtate future lifetime, force of mortality.

UNIT II

Life table and its relation with survival function, examples, assumptions for fractional ages, some analytical laws of mortality, select and ultimate tables.

UNIT III

Multiple life functions, joint life and last survivor status, insurance and annuity benefits through multiple life functions evaluation for special mortality laws. Multiple decrement models, deterministic and random survivorship groups, associated single decrement tables, central rates of multiple decrement, net single premiums and their numerical evaluations.

UNIT IV

Distribution of aggregate claims, compound Poisson distribution and its applications.

UNIT V

Principles of compound interest: Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding.

UNIT VI

Insurance payable at the moment of death and at the end of the year of death-level benefit insurance, endowment insurance, deferred insurance and varying benefit insurance, recursions, commutation functions.

UNIT VII

Life annuities: Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities-immediate and apportionable annuities-due.

UNIT VIII

Net premiums: Continuous and discrete premiums, true monthly payment premiums, apportionable premiums, commutation functions, accumulation type benefits. Payment premiums, apportionable premiums, commutation functions, accumulation type benefits. Net premium reserves: Continuous and discrete net premium reserve, reserves on a semi-continuous basis, reserves based on true monthly premiums, reserves on an apportionable or discounted continuous basis, reserves at fractional durations, allocations of loss to policy years, recursive formulas and differential equations for reserves, commutation functions.

UNIT IX

Some practical considerations: Premiums that include expenses-general expenses types of expenses, per policy expenses. Claim amount distributions, approximating the individual model, stop-loss insurance.

Suggested Readings

Atkinson ME & Dickson DCM. 2000. *An Introduction to Actuarial Studies*. Elgar Publ.

Bedford T & Cooke R. 2001. *Probabilistic Risk Analysis*. Cambridge.

Booth PM, Chadburn RG, Cooper DR, Haberman S & James DE. 1999. *Modern Actuarial Theory and Practice*. Chapman & Hall.

Borowiak Dale S. 2003. *Financial and Actuarial Statistics: An Introduction. 2003*. Marcel Dekker.

Bowers NL, Gerber HU, Hickman JC, Jones DA & Nesbitt CJ. 1997. *Actuarial Mathematics.* 2nd Ed. Society of Actuaries, Ithaca, Illinois.

Daykin CD, Pentikainen T & Pesonen M. 1994. *Practical Risk Theory for Actuaries*. Chapman & Hall.

Klugman SA, Panjer HH, Willmotand GE & Venter GG. 1998. *Loss Models: From data to Decisions*. John Wiley.

Medina PK & Merino S. 2003. *Mathematical Finance and Probability: A Discrete Introduction.* Basel, Birkhauser.

Neill A. 1977. *Life Contingencies*. Butterworth-Heinemann.

Rolski T, Schmidli H, Schmidt V & Teugels J. 1998. *Stochastic Processes for Insurance and Finance*. John Wiley.

Rotar VI. 2006. *Actuarial Models. The Mathematics of Insurance*. Chapman & Hall/CRC.

Spurgeon ET. 1972. *Life Contingencies*. Cambridge Univ. Press.

STAT 571 BIOINFORMATICS 2+0

Objective

Bioinformatics is a new emerging area. It is an integration of Statistics, Computer applications and Biology. The trained manpower in the area of Bioinformatics is required for meeting the new challenges in teaching and research in the discipline of Agricultural Sciences. This course is meant to train the students on concepts of basic biology, statistical techniques and computational techniques for understanding bioinformatics principals.

Theory

UNIT I

Basic Biology: Cell, genes, gene structures, gene expression and regulation, Molecular tools, nucleotides, nucleic acids, markers, proteins and enzymes, bioenergetics, single nucleotide polymorphism, expressed sequence tag. Structural and functional genomics: Organization and structure of genomes, genome mapping, assembling of physical maps, strategies and techniques for genome sequencing and analysis.

UNIT II

Computing techniques: OS and Programming Languages - *Linux*, *perl*, *bioperl*, *cgi*, *MySQL*, *phpMyAdmin*; Coding for browsing biological databases on web, parsing & annotation of genomic sequences; Database designing; Computer networks - Internet, World wide web, Web browsers - EMBnet, NCBI; Databases on public domain pertaining to Nucleic acid sequences, protein sequences, SNPs, etc.; Searching sequence databases, Structural databases.

UNIT III

Statistical Techniques: MANOVA, Cluster analysis, Discriminant analysis, Principal component analysis, Principal coordinate analysis, Multidimensional scaling; Multiple regression analysis; Likelihood approach in estimation and testing; Resampling techniques - Bootstrapping and Jack-knifing; Hidden Markov Models; Bayesian estimation and Gibbs sampling;

UNIT IV

Tools for Bioinformatics: DNA Sequence Analysis - Features of DNA sequence analysis, Approaches to EST analysis; Pairwise alignment techniques: Comparing two sequences, PAM and BLOSUM, Global alignment (The Needleman and Wunsch algorithm), Local Alignment (The Smith-Waterman algorithm), Dynamic programming, Pairwise database searching; Sequence analysis- BLAST and other related tools, Multiple alignment and database search using motif models, ClustalW, Phylogeny; Databases on SNPs; EM algorithm and other methods to discover common motifs in biosequences; Gene prediction based on Neural Networks, Genetic algorithms, Hidden Markov models. Computational analysis of protein sequence, structure and function; Design and Analysis of microarray experiments.

Suggested Readings

Baldi P & Brunak S. 2001. *Bioinformatics: The Machine Learning Approach.* 2nd Ed. (Adaptive Computation and Machine Learning). MIT Press.

Baxevanis AD & Francis BF. (Eds.). 2004. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. John Wiley.

Bergeron BP. 2002. *Bioinformatics Computing*. Prentice Hall.
Duda RO, Hart PE & Stork DG. 1999. *Pattern Classification*. John Wiley.
Ewens WJ & Grant GR. 2001. *Statistical Methods in Bioinformatics: An Introduction* (*Statistics for Biology and Health*). Springer.
Hunt S & Livesy F. (Eds.). 2000. *Functional Genomics: A Practical Approach* (*The Practical Approach Series, 235*). Oxford Univ. Press.
Jones NC & Pevzner PA. 2004. *An Introduction to Bioinformatics Algorithims*. MIT Press.
Koski T & Koskinen T. 2001. *Hidden Markov Models for Bioinformatics.* Kluwer.
Krane DE & Raymer ML. 2002. *Fundamental Concepts of Bio-informatics*. Benjamin / Cummings.
Krawetz SA & Womble DD. 2003. *Introduction to Bioinformatics: A Theoretical and Practical Approach*. Humana Press.
Lesk AM. 2002. *Introduction to Bio-informatics*. Oxford Univ. Press.
Percus JK. 2001. *Mathematics of Genome Analysis*. Cambridge Univ. Press.
Sorensen D & Gianola D. 2002. *Likelihood, Bayesian and MCMC Methods in Genetics.* Springer.
Tisdall JD. 2001. *Mastering Perl for Bioinformatics*. O'Reilly & Associates.
Tisdall JD. 2003. B*eginning Perl for Bioinformatics*. O'Reilly & Associates.
Wang JTL, Zaki MJ, Toivonen HTT & Shasha D. 2004. *Data Mining in Bioinformatics*. Springer.
Wu CH & McLarty JW. 2000. *Neural Networks and Genome Informatics*. Elsevier.
Wunschiers R. 2004. *Computational Biology Unix/Linux, Data Processing and Programming*. Springer.

STAT 572 ECONOMETRICS 2+0

Objective

This course is meant for training the students in econometric methods and their applications in agriculture. This course would enable the students in understanding the economic phenomena through statistical tools and economics principles.

Theory

UNIT I

Representation of Economic phenomenon, relationship among economic variables, linear and non linear economic models, single equation general linear regression model, basic assumptions, Ordinary least squares method of estimation for simple and multiple regression models; summary statistics correlation matrix, co-efficient of multiple determination, standard errors of estimated parameters, tests of significance and confidence interval estimation. BLUE properties of Least Squares estimates. Chow test, test of improvement of fit through additional regressors. Maximum likelihood estimation.

UNIT II

Heteroscedasticity, Auto-correlation, Durbin Watson test, Multicollinearity. Stochastic regressors,

Errors in variables, Use of instrumental variables in regression analysis. Dummy Variables. Distributed Lag models: Koyck's Geometric Lag scheme, Adaptive Expectation and Partial Adjustment Mode, Rational Expectation Models and test for rationality.

UNIT III

Simultaneous equation model: Basic rationale, Consequences of simultaneous relations, Identification problem, Conditions of Identification, Indirect Least Squares, Two-stage least squares, K-class estimators, Limited Information and Full Information Maximum Likelihood Methods, Three stage least squares, Generalized least squares, Recursive models, SURE Models. Mixed Estimation Methods, use of instrumental variables, pooling of cross-section and time series data, Principal Component Methods.

UNIT IV

Problem and Construction of index numbers and their tests; fixed and chain based index numbers; Construction of cost of living index number.

UNIT V

Demand analysis - Demand and Supply Curves; Determination of demand curves from market data. Engel's Law and the Engel's Curves, Income distribution and method of its estimation, Pareto's Curve, Income inequality measures.

Suggested Readings

Croxton FE & Cowden DJ. 1979. *Applied General Statistics*. Prentice Hall of India.

Johnston J. 1984. *Econometric Methods*. McGraw Hill.

Judge GC, Hill RC, Griffiths WE, Lutkepohl H & Lee TC. 1988. *Introduction to the Theory and Practice of Econometrics*. 2nd Ed. John Wiley.

Kmenta J. 1986. *Elements of Econometrics*. 2nd Ed. University of Michigan Press.

Koop G. 2007. *Introduction to Econometrics*. John Wiley.

Maddala GS. 2001. *Introduction to Econometrics*. 3rd Ed. John Wiley.

Pindyck RS & Rubinfeld DL. 1998. *Econometric Models and Economic Forecasts*. 4th Ed. McGraw Hill.

Verbeek M. 2008. *A Guide to Modern Econometrics*. 3rd Ed. John Wiley.

STAT 573 STATISTICAL QUALITY CONTROL 2+0

Objective

This course is meant for exposing the students to the concepts of Statistical Quality Control and their applications in agribusiness and agro-processing industries. This course would enable the students to have an idea about the statistical techniques used in quality control. students who do not have sufficient background of Statistical Methods.

Theory

UNIT I

Introduction to Statistical Quality Control; Control Charts for Variables - Mean, Standard deviation and Range charts; Statistical basis; Rational subgroups.

UNIT II

Control charts for attributes- 'np', 'p' and 'c' charts.

UNIT III

Fundamental concepts of acceptance, sampling plans, single, double and sequential sampling plans for attributes inspection.

UNIT IV

Sampling inspection tables for selection of single and double sampling plans.

Suggested Readings

Cowden DJ. 1957. *Statistical Methods in Quality Control*. Prentice Hall of India.

Dodge HF & Romig HG. 1959. *Sampling Inspection Tables*. John Wiley.

Duncan A.J. 1986. *Quality Control and Industrial Statistics*. 5th Ed. Irwin Book Co.

Grant EL & Leavenworth RS. 1996. *Statistical Quality Control*. 7th Ed. McGraw Hill.

Montgomery DC. 2005. *Introduction to Statistical Quality Control*. 5th Ed. John Wiley.

Wetherhil G.B. 1977. *Sampling Inspection and Quality Control*. Halsted Press.

STAT 574/ BST 574 **OPTIMIZATION TECHNIQUES** **1+1**

Objective

This course is meant for exposing the students to the mathematical details of the techniques for obtaining optimum solutions under constraints for desired output. They will be taught numerical methods of optimization, linear programming techniques, non-linear programming and multiple objective programming. Students will also be exposed to practical applications of these techniques.

Theory

UNIT I

Classical Optimization Techniques: Necessary Conditions for an Extremum. Constrained Optimization: Lagrange Multipliers, Statistical Applications. Optimization and Inequalities. Classical Inequalities, like Cauchy-Schwarz Inequality, Jensen Inequality and Markov Inequality.

UNIT II

Numerical Methods of Optimization: Numerical Evaluation of Roots of Equations, Direct Search Methods, Sequential Search Methods -- Fibonacci Search Method. Random Search Method - Method of Hooke and Jeeves, Simplex Search Method. Gradient Methods, like Newton's Method, and Method of Steepest Ascent. Nonlinear Regression and Other Statistical Algorithms, like Expectation - Maximization Algorithm.

UNIT III

Linear programming Techniques - Simplex Method, Karmarkar's Algorithm, Duality and Sensitivity Analysis. Zero-sum Two-person Finite Games and Linear Programming. Integer Programming. Statistical Applications.

UNIT IV

Nonlinear Programming and its Examples. Kuhn-Tucker Conditions. Quadratic Programming. Convex Programming. Basics of Stochastic Programming. Applications. Elements of Multiple Objective Programming. Dynamic Programming, Optimal Control Theory - Pontryagin's Maximum Principle, Time-Optimal Control Problems.

Practical

Problems based on classical optimization techniques; Problems based on optimization techniques with constraints; Minimization problems using numerical methods; Linear programming (LP) problems through graphical method; LP problem by Simplex method; LP problem using simplex method (Two-phase method); LP problem using primal and dual method; Sensitivity analysis for LP problem; LP problem using Karmarkar's method; Problems based on Quadratic programming; Problems based on Integer programming; Problems based on Dynamic programming; Problems based on Pontryagin's Maximum Principle.

Suggested Readings

Rao SS. 2007. *Engineering Optimization: Theory and Practice*. 3rd Ed. New Age.

Rustagi JS. 1994. *Optimization Techniques in Statistics*. Academic Press.

Taha HA. 2007. *Operations Research: Introduction with CD*. 8th Ed. Pearson Edu.

Zeleny M. 1974. *Linear Multiobjective Programming*. Springer.

STAT 575 DEMOGRAPHY 2+0

Objective

This course is meant for training the students in measures of demographic indices, estimation procedures of demographic parameters. Students would also be exposed to population projection techniques and principles involved in bioassays.

Theory

UNIT I

Introduction to vital statistics, crude and standard mortality and morbidity rates, Estimation of mortality, Measures of fertility and mortality, period and cohort measures.

UNIT II

Life tables and their applications, methods of construction of abridged life tables, Increment-Decrement Life Tables.

UNIT III

Stationary and stable populations, Migration and immigration. Application of stable population theory to estimate vital rates, migration and its estimation. Demographic relations in Nonstable populations. Measurement of population growth, Lotka's model(deterministic) and intrinsic rate of growth, Measures of mortality and morbidity, Period and

UNIT IV

Principle of biological assays, parallel line and slope ratio assays, choice of doses and efficiency in assays quantal responses, probit and logit transformations, epidemiological models.

Suggested Readings

Cox DR. 1957. *Demography*. Cambridge Univ. Press.

Finney DJ. 1981. *Statistical Methods in Biological Assays*. Charles Griffin.

Fleiss JL. 1981. *Statistical Methods for Rates and Proportions*. John Wiley.

Lawless JF. 1982. *Statistical Models and Methods for Lifetime Data*. John Wiley.

MacMahon B & Pugh TF. 1970. *Epidemiology- Principles and Methods*. Little Brown, Boston.

Mann NR, Schafer RE & Singpurwalla ND. 1974. *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley.

Newell C. 1988. *Methods and Models in Demography*. Guilford Publ.

Preston S, Heuveline P & Guillot M. 2001. *Demography: Measuring and Modeling Population Processes*. Blackwell Publ.

Rowland DT. 2004. *Demographic Methods and Concepts*. Oxford Press.

Siegel JS & Swanson DA. 2004. *The Methods and Material of Demography*. 2nd Ed. Elsevier.

Woolson FR. 1987. *Statistical Methods for the Analysis of Biomedical Data*. John Wiley.

STAT 576 STATISTICAL METHODS FOR LIFE SCIENCES 2+0

Objective

This course focuses on statistical methods for discrete data collected in public health, clinical and biological studies including survival analysis. This would enable the students to understand the principles of different statistical techniques useful in public health and clinical studies conducted.

Theory

UNIT I

Proportions and counts, contingency tables, logistic regression models, Poisson regression and log-linear models, models for polytomous data and generalized linear models.

UNIT II

Computing techniques, numerical methods, simulation and general implementation of biostatistical analysis techniques with emphasis on data applications. Analysis of survival time data using parametric and non-parametric models, hypothesis testing, and methods for analyzing censored (partially observed) data with covariates. Topics include marginal estimation of a survival function, estimation of a generalized multivariate linear regression model (allowing missing covariates and/or outcomes).

UNIT III

Proportional Hazard model: Methods of estimation, estimation of survival functions, time-dependent covariates, estimation of a multiplicative intensity model (such as Cox proportional hazards model) and estimation of causal parameters assuming marginal structural models.

UNIT IV

General theory for developing locally efficient estimators of the parameters of interest in censored data models. Rank tests with censored data. Computing techniques, numerical methods, simulation and general implementation of biostatistical analysis techniques with emphasis on data applications.

UNIT V

Newton, scoring, and EM algorithms for maximization; smoothing methods; bootstrapping; trees and neural networks; clustering; isotonic regression; Markov chain Monte Carlo methods.

Suggested Readings

Biswas S. 1995. *Applied Stochastic Processes. A Biostatistical and Population Oriented Approach*. Wiley Eastern Ltd.

Collett D. 2003. *Modeling Survival Data in Medical Research*. Chapman & Hall.

Cox DR & Oakes D. 1984. *Analysis of Survival Data*. Chapman & Hall.

Hosmer DW Jr. & Lemeshow S. 1999. *Applied Survival Analysis: Regression Modeling or Time to Event*. John Wiley.

Klein JP & Moeschberger ML. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.

Kleinbaum DG & Klein M 2005. *Survival Analysis. A Self Learning Text*. Springer.

Kleinbaum DG & Klein M. 2005. *Logistic Regression*. 2nd Ed. Springer.

Lee ET. 1992. *Statistical Methods for Survival Data Analysis*. John Wiley.

Miller RG. 1981. *Survival Analysis.* John Wiley.

Therneau TM & Grambsch PM. 2000. *Modeling Survival Data: Extending the Cox Model*. Springer.

STAT 577 STATISTICAL ECOLOGY 2+0

Objective

This course is meant for exposing the students to the importance and use of statistical methods in collections of ecological data, species-abundance relations, community classification and community interpretation.

Theory

UNIT I

Ecological data, Ecological sampling; Spatial pattern analysis: Distribution methods, Quadrant-variance methods, Distance methods.

UNIT II

Species-abundance relations: Distribution models, Diversity indices; Species affinity: Niche-overlap indices, interspecific association, interspecific covariation.

UNIT III

Community classification: Resemblance functions, Association analysis, Cluster analysis; Community Ordination: Polar Ordination, Principal Component Analysis, Correspondence analysis, Nonlinear ordination.

UNIT IV

Community interpretation: Classification Interpretation and Ordination Interpretation.

Suggested Readings

Pielou EC. 1970. *An introduction to Mathematical Ecology*. John Wiley.

Reynolds JF & Ludwig JA. 1988. *Statistical Ecology: A Primer on Methods and Computing*. John Wiley.

Young LJ, Young JH & Young J. 1998. *Statistical Ecology: A Population Perspective*. Kluwer.

STAT 601 ADVANCED STATISTICAL COMPUTING 2+1

Objective

This is an advanced course in Statistical Computing that aims at describing some advanced level topics in this area of research with a very strong potential of applications. This course also prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject to agricultural sciences and use of statistical packages.

Theory

UNIT I

Measures of association. Structural models for discrete data in two or more dimensions.

Estimation in complete tables. Goodness of fit, choice of a model. Generalized Linear Model for discrete data, Poisson and Logistic regression models. Log-linear models.

UNIT II

Elements of inference for cross-classification tables. Models for nominal and ordinal response.

UNIT III

Computational problems and techniques for robust linear regression, nonlinear and generalized linear regression problem, tree-structured regression and classification, cluster analysis, smoothing and function estimation, robust multivariate analysis.

UNIT IV

Analysis of incomplete data: EM algorithm, single and multiple imputations. Markov Chain, Monte Carlo and annealing techniques, Neural Networks, Association Rules and learning algorithms.

UNIT V

Linear mixed effects models, generalized linear models for correlated data (including generalized estimating equations), computational issues and methods for fitting models, and dropout or other missing data.

UNIT VI

Multivariate tests of linear hypotheses, multiple comparisons, confidence regions, prediction intervals, statistical power, transformations and diagnostics, growth curve models, dose-response models.

Practical

Analysis of qualitative data; Generalized linear for correlated data; Generalized linear models for discrete data; Robust methods of estimation and testing of non-normal data; Robust multivariate analysis; Cluster analysis; Analysis of Incomplete data; Classification and prediction using artificial neural networks; Markov Chain; Analysis of data having random effects using Linear mixed effects models; Analysis of data with missing observations; Applications of multiple comparison procedures; Building Simultaneous confidence intervals; Fitting of growth curve models to growth data; Fitting of dose-response curves and estimation of parameters.

Suggested Readings

Everitt BS & Dunn G. 1991. *Advanced Multivariate Data Analysis*. 2nd Ed. Arnold.

Geisser S. 1993. *Predictive Inference: An Introduction*. Chapman & Hall.

Gentle JE, Härdle W & Mori Y. 2004. *Handbook of Computational Statistics -Concepts and Methods*. Springer.

Han J & Kamber M. 2000. *Data Mining: Concepts and Techniques*. Morgan.

Hastie T, Tibshirani R & Friedman R. 2001. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer.

Kennedy WJ & Gentle JE. 1980. *Statistical Computing*. Marcel Dekker.

Miller RG Jr. 1986. *Beyond ANOVA, Basics of Applied Statistics*. John Wiley.

Rajaraman V. 1993. *Computer Oriented Numerical Methods*. Prentice-Hall.

Robert CP & Casella G. 2004. *Monte Carlo Statistical Methods*. 2nd Ed. Springer.

Ross S. 2000. *Introduction to Probability Models*. Academic Press.

Simonoff JS. 1996. *Smoothing Methods in Statistics*. Springer.

Thisted RA. 1988. *Elements of Statistical Computing*. Chapman & Hall.

Venables WN & Ripley BD. 1999. *Modern Applied Statistics With S-Plus*. 3rd Ed. Springer.

Free Statistical Softwares: http://freestatistics.altervista.org/en/stat.php.

Design Resources Server: www.iasri.res.in.

SAS Online Doc 9.1.3:
http://support.sas.com/onlinedoc/913/docMainpage.jsp

STAT 602 SIMULATION TECHNIQUES 1+1

Objective

This course is meant for students who have a good knowledge in Statistical Inference and Statistical Computing. This course would prepare students for undertaking research in the area of simulation techniques and their applications to agricultural sciences.

Theory

UNIT I

Review of simulation methods; Implementation of simulation methods - for various probability models, and resampling methods: theory and application of the jackknife and the bootstrap.

UNIT II

Randomization tests, analysis using computer software packages. Simulating multivariate distributions, MCMC methods and Gibbs sampler.

UNIT III

Correlograms, periodograms, fast Fourier transforms, power spectra, cross-spectra, coherences, ARMA and transfer-function models, spectral-domain regression. Simulated data sets to be analyzed using popular computer software packages

UNIT IV

Stochastic simulation: Markov Chain, Monte Carlo, Gibbs' sampling, Hastings-Metropolis algorithms, critical slowing-down and remedies, auxiliary variables, simulated tempering, reversible- jump MCMC and multi-grid methods.

Practical

Simulation from various probability models; Resampling methods, jackknife and the bootstrap; Randomization tests; Simulating multivariate distributions, MCMC methods and Gibbs sampler; Correlograms, periodograms, fast Fourier transforms, power spectra, cross-spectra, coherences; ARMA and transfer-function models, spectral-domain regression; Simulated data sets to be analyzed using popular computer software packages; Markov Chain, Monte Carlo, Gibbs' sampling; Reversible- jump MCMC and multi-grid methods.

Suggested Readings

Averill ML, Kelton D. 2005. *Simulation, Modeling and Analysis*. Tata McGraw Hill.

Balakrishnan N, Melas VB & Ermakov S. (Ed.). 2000. *Advances in Stochastic Simulation Methods*. Basel-Birkhauser.

Banks J. (Ed.). 1998. *Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice*. John Wiley.

Brately P, Fox BL & Scharge LE. 1987. *A Guide to Simulation*. Springer.

Davison AC & Hinkley DV. 2003. *Bootstrap Methods and their Application*. Cambridge Univ. Press.

Gamerman D, Lopes HF & Lopes HF. 2006. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. CRC Press.

Gardner FM & Baker JD. 1997. *Simulation Techniques Set*. John Wiley.

Gentle JE. 2005. *Random Number Generation and Monte Carlo Methods*. Springer.

Janacek G & Louise S. 1993. *Time Series: Forecasting, Simulation, Applications*. Ellis Horwood Series in Mathematics and Its Applications.

Kleijnen J & Groenendaal WV. 1992. *Simulation: A Statistical Perspective*. John Wiley.

Kleijnen J. 1974 (Part I), 1975 (Part II). *Statistical Techniques in Simulation*. Marcel Dekker.

Law A & Kelton D. 2000. *Simulation Modeling and Analysis*. McGraw Hill.

Press WH, Flannery BP, Tenkolsky SA & Vetterling WT. 1986. *Numerical Recipes*. Cambridge Univ. Press.

Ripley BD. 1987. *Stochastic Simulation*. John Wiley.

Ross SM. 1997. *Simulation*. John Wiley.

STAT 611 ADVANCED STATISTICAL METHODS 2+0

Objective

This is an advanced course in Statistical Methods that aims at describing some advanced level topics in this area of research with a very strong potential of applications. This course also prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject to agricultural sciences.

Theory

UNIT I

Ridge regression: Basic form, Use as a selection procedure. Robust regression: Least absolute deviations regression, M-estimators, Least median of squares regression. Nonparametric regression.

UNIT II

Introduction to the theory and applications of generalized linear models, fixed effects, random effects and mixed effects models, estimation of variance components from unbalanced data. Unified theory of least - squares, MINQUE, MIVQUE, REML.

UNIT III

Quasi-likelihoods, and generalized estimating equations - logistic regression, over-dispersion, Poisson regression, log-linear models, conditional likelihoods, generalized mixed models, and regression diagnostics. Theory of statistical methods for analyzing categorical data by means of linear models; multifactor and multi-response situations; interpretation of interactions.

UNIT IV

Fitting of a generalized linear model, mixed model and variance components estimation, MINQUE, MIVQUE, REML.

UNIT V

Fitting of Logistic regression, Poisson regression, ridge regression, robust regression, non-parametric regression.

Suggested Readings

Chatterjee S, Hadi A & Price B.1999. *Regression Analysis by Examples*. John Wiley.

Draper NR & Smith H. 1998. *Applied Regression Analysis*. 3 Ed.rdJohn Wiley.

Rao CR. 1965. *Linear Statistical Inference and its Applications*. 2nd Ed. John Wiley.

Searle SR, Casella G & McCulloch CE. 1992. *Variance Components*. John Wiley.

Searle SR. 1971. *Linear Models*. John Wiley.

STAT 612 ADVANCED STATISTICAL INFERENCE 3+0

Objective

This course aims at describing the advanced level topics in statistical methods and statistical inference. This course would prepare students to have a strong base in basic statistics that would help them in undertake basic and applied research in Statistics.

Theory

UNIT I

Robust estimation and robust tests, Robustness, M-estimates. L-estimates, asymptotic techniques, Bayesian inference. Detection and handling of outliers in statistical data.

UNIT II

Loglinear models, saturated models, hierarchical models, Analysis of multi - dimensional contingency tables. Non-parametric maximum likelihood estimation.

UNIT III

Density Estimation: Density Estimation in the Exploration and Presentation of Data. Survey of existing methods. The Kernel method for Univariate Data: Rosenblatts naïve estimator, its bias and variance. Consistency of general Kernel estimators, MSE and IMSE. Asymptotic

normality of Kernel estimates of density. Estimation of distribution by method of kernels.

UNIT IV

Consistency and asymptotic normality (CAN) of real and vector parameters. Invariance of consistency under continuous transformation. Invariance of CAN estimators under differentiable transformations, generation of CAN estimators using central limit theorem. Exponential class of densities and multinomial distribution, Cramer-Huzurbazar theorem, method of scoring.

UNIT V

Efficiency: asymptotic relative efficiency and Pitman's theorem. Concepts and examples of Bahadur efficiency and Hodges-Lehmanns efficiency with examples. The concepts of Rao's second order efficiency and Hodges-Lehmann's Deficiency with examples. Rank tests, permutation tests, asymptotic theory of rank tests under null and alternative (contiguous) hypotheses.

UNIT VI

Inference on Markov Chains: Maximum likelihood estimation and testing of Transition Probability Matrix of a Markov Chain, testing for order of a Markov chain, estimation of functions of transition probabilities.

UNIT VII

Concept of loss, risk and decision functions, admissible and optimal decision functions, a-priori and posteriori distributions, conjugate families. Bayes and Minimax decision rules and some basic results on them. Estimation and testing viewed as cases of decision problems. Bayes and Minimax decision functions with applications to estimation with quadratic loss function. Concept of Bayesian sequential analysis. Bayes sequential decision rule. The SPRT as a Bayes procedure. Minimax sequential procedure.

UNIT VIII

U-Statistics: definitions of estimable parametric function, kernel, symmetric kernel and U-statistics. Variance and covariance of U-statistics. Hoeffding's decomposition of U-statistics -examples. U-statistics based on sampling from finite populations and weighted U-statistics with examples. Some convergence results on U-statistics. Asymptotic normality of U-statistics with examples.

UNIT IX

Resampling Plans : Estimation of standard and biased deviation of point estimator by the Jackknife, the Bootstrap, the Infinitesimal Jackknife, the Delta and the Influence function methods. Estimation of excess error in regression by cross validation, the Jackknife and Bootstrap methods. Nonparametric confidence interval for the median by the Percentile method.

Suggested Readings

Casela G & Berger RL. 2001. *Statistical Inference*. Duxbury Thompson Learning.

Daniel W.1990. *Applied Nonparametric Statistics*. Houghton Mifflin, Boston.

DeGroot MH. 1970. *Optimal Statistical Decisions*. McGraw Hill.

Efron B & Tibshirani RJ. 1993. *An Introduction to Bootstrap*. Chapman Hall/CRC.
Ferguson TS. 1967. *Mathematical Statistics, A Decision Theoretic Approach*. Academic Press.
Gibbons JD & Chakraborty S. 1992. *Non-parametric Statistical Inference*. Marcel Dekker.
Gray HL & Schucany WR.1972. *The Generalized Jackknife Statistics*. Marcel Dekker.
Kale BK.1999. *A First Course on Parametric Inference*. Narosa Publ.
Prakasa Rao BLS. 1983. *Nonparametric Functional Estimation*. Academic Press.
Rao CR.1965. *Linear Statistical Inference and its Applications*. 2 Ed. nd John Wiley.
Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.
Silvey SD. 1975. *Statistical Inference*. Chapman & Hall.
Tapia RA & Thompson JR. 1978. *Nonparametric Probability Density Estimation*. Johns Hopkins Univ. Press.
Tiku ML, TanWY & Balakrishnana N. 1986. *Robust Inference*. Marcel Dekker.
Wald A. 2004. *Sequential Analysis*. Dover Publ.
Wasserman L. 2006. *All of Nonparametric Statistics*. Springer.

STAT 613 ADVANCED DESIGN OF EXPERIMENTS 2+0

Objective

This is an advanced course in Design of Experiments that aims at describing some advanced level topics for students who wish to pursue research in Design of Experiments. This course prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject to agricultural sciences.

Theory

UNIT I

General properties and analysis of block designs. Balancing criteria. *m*-associate PBIB designs, and their association schemes including lattice designs - properties and construction, Designs for test treatment - control(s) comparisons; Nested block designs, Mating designs.

UNIT II

General properties and analysis of two-way heterogeneity designs, Youden type designs, generalized Youden designs, Pseudo Youden designs. Structurally Incomplete block designs, Designs for two sets of treatments.

UNIT III

Balanced factorial experiments - characterization and analysis (symmetrical and asymmetrical factorials). Factorial experiments with extra treatment(s). Orthogonal arrays, Mixed orthogonal arrays, balanced arrays, Fractional replication, Regular and irregular fractions.

UNIT IV

Response surface designs - Symmetrical and asymmetrical factorials, Response optimization and slope estimation, Blocking. Canonical analysis and ridge analysis. Experiments with mixtures: design and analysis. Experiments with qualitative cum quantitative factors.

UNIT V

Optimality criteria and optimality of designs, robustness of designs against loss of data, outliers, etc. Diagnostics in design of experiments.

Suggested Readings

Chakraborti MC. 1962. *Mathematics of Design and Analysis of Experiments*. Asia Publ. House.

Dean AM & Voss D. 1999. *Design and Analysis of Experiments*. Springer.

Dey A & Mukerjee R. 1999. *Fractional Factorial Plans*. John Wiley.

Dey A 1986. *Theory of Block Designs*. Wiley Eastern.

Hall M Jr. 1986. *Combinatorial Theory*. John Wiley.

Hedayat AS, Sloane NJA & Stufken J. 1999. *Orthogonal Arrays: Theory and Applications*. Springer.

John JA & Quenouille MH. 1977. *Experiments: Design and Analysis*. Charles & Griffin.

Khuri AI & Cornell JA. 1996. *Response Surface Designs and Analysis*. 2nd Ed. Marcel Dekker.

Montgomery DC. 2005. *Design and Analysis of Experiments*. John Wiley.

Ogawa J. 1974. *Statistical Theory of the Analysis of Experimental Designs*. Marcel Dekker.

Parsad R, Gupta VK, Batra PK, Satpati SK & Biswas P. 2007. *Monograph on α-designs*. IASRI, New Delhi.

Raghavarao D. 1971. *Construction and Combinatorial Problems in Design of Experiments*. John Wiley.

Shah KR & Sinha BK. 1989. *Theory of Optimal Designs. Lecture notes in Statistics*. Vol. 54. Springer.

Street AP & Street DJ. 1987. *Combinatorics of Experimental Designs*. Oxford Science Publ.

Design Resources Server: www.iasri.res.in.

STAT 614 ADVANCED SAMPLING TECHNIQUES 2+0

Objective

This is an advanced course in Sampling Techniques that aims at describing some advanced level topics for students who wish to pursue research in Sampling Techniques. This course prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject to Statistical System in the country.

Theory

UNIT I

Controlled selection. Two way stratification, collapsed strata. Systematic sampling in two dimensions. Use of combinatorics in controlled selection. Integration of surveys - Lahiri and Keyfitz's procedures.

UNIT II

Variance estimation in complex surveys. Taylor's series linearisation, balanced repeated replication, Jackknife and bootstrap methods.

UNIT III

Unified theory of sampling from finite populations. UMV - Non-existence theorem and existence theorem under restricted conditions. Concept of sufficiency and likelihood in survey sampling. Admissibility and hyper-admissibility.

UNIT IV

Inference under super population models - concept of designs and model unbiasedness, prediction approach. Regression analysis and categorical data analysis with data from complex surveys. Domain estimation. Small area estimation.

UNIT V

Stochastic parameter models, Bayes' linear predictor, Bayesian models with multi-stage sampling. Measurement error and small area estimation, Time series approach in survey sampling. Dynamic Bayesian prediction, Kalman filter, Empirical and Hierarchical Bayes predictors, Robust linear prediction, Bayesian robustness.

Suggested Readings

Berger JO. 1993. *Statistical Decision Theory and Bayesian Analysis*. Springer.

Bolfarine H & Zacks S. 1992. *Prediction Theory for Finite Population Sampling*. Springer.

Cassel CM, Sarndal CE & Wretman JH. 1977. *Foundations of Inference in Survey Sampling*. John Wiley.

Des Raj & Chandhok P. 1998. *Sample Survey Theory*. Narosa Publ. House.

Ghosh M & Meeden G. 1997. *Bayesian Method for Finite Population Sampling*. *Monograph on Statistics and Applied Probability*. Chapman & Hall.

Mukhopadhyay P. 1998. *Theory and Methods of Survey Sampling*. Prentice Hall of India.

Rao JNK. 2003. *Small Area Estimation*. John Wiley.

Sarndal CE, Swensson B & Wretman JH. 1992. *Model Assisted Survey Sampling*. Springer.

STAT 615 ADVANCED STATISTICAL GENETICS 2+0

Objective

This is an advanced course in Statistical Genetics that aims at describing some advanced level topics for students who wish to pursue research in Statistical Genetics. This course prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject in plant and animal breeding.

Theory

UNIT I

Hardy-Weinberg law with multiple allelic systems, auto-tetraploids and self-sterility alleles. Complex cases of selection with two or more loci.

UNIT II

Different approaches to study inbreeding process, methods of path co-efficient, probability and generation matrix. Fisher's approach to inbreeding. Stochastic process of gene frequency change, transition matrix approach using finite Markov chains, diffusion approximation, Steady decay and distribution of gene frequency, Probability of fixation of a gene, Conditional process - Markov chains and diffusion approaches, Distribution of time until fixation, random fluctuations in selection intensity, stationary distribution of gene frequency. Effective population size.

UNIT III

Prediction and estimation of genetic merit. Best linear unbiased prediction, Use of mixed model methodology in analysis of animal and plant breeding experiments. Newer reproductive technology and its effect in genetic evaluation of individual merit. Estimation of genetic parameters - problems relating to computational aspects of genetic variance components, parameter estimation in variance component models for binary response data.

UNIT IV

Identification of genes with large effects, Use of molecular markers (RFLP, PCR-AFLP, RAPD and SSR), Gene mapping and Quantitative trait loci. Molecular manipulation for genetic variability.

UNIT V

Survival analysis and concept of censored observation in animal breeding. Phylogeny and analysis of molecular variance.

Suggested Readings

Crow JF & Kimura M. 1970. *An Introduction of Population Genetics Theory*. Harper & Row.

Ewens WJ. 1979. *Mathematical Population Genetics.* Springer.

Falconer DS. 1985. *Introduction to Quantitative Genetics*. ELBL.

Fisher RA. 1949. *The Theory of Inbreeding*. Oliver & Boyd.

Fisher RA. 1958. *The Genetical Theory of Natural Selection.* Dover Publ.

Haldane JBS. 1932. *The Causes of Evolution*. Harper & Bros.

Kempthorne O. 1957. *An Introduction to Genetic Statistics*. The Iowa State Univ. Press.

Lerner IM. 1950. *Population Genetics and Animal Improvement.* Cambridge Univ. Press.

Lerner IM. 1958. *The Genetic Theory of Selection*. John Wiley.

Li CC. 1982. *Population Genetics.* The University of Chicago Press.

Mather K & Jinks JL. 1982. *Biometrical Genetics*. Chapman & Hall.

Mather K. 1951. *The Measurement of Linkage in Heredity*. Methuen.

Nagilaki T. 1992. *Introduction to Theoretical Population Genetics*. Springer.

Narain P. 1990. *Statistical Genetics*. Wiley Eastern.

STAT 616 STATISTICAL MODELING 1+1

Objective

This is an advanced course in Statistical Methods that aims at describing some advanced level topics in this area of research with a very strong potential of applications. This course also prepares students for undertaking research in the area of empirical and mechanistic models and nonlinear estimation and the replications in different disciplines of agricultural sciences.

Theory

UNIT I

Empirical and mechanistic models. Nonlinear growth models like monomolecular, logistic, Gompertz, Richards. Applications in agriculture and fisheries.

UNIT II

Nonlinear estimation: Least squares for nonlinear models, Methods for estimation of parameters like Linearization, Steepest, and Levenberg-Marquardt's Reparameterization.

UNIT III

Two-species systems. Lotka-Volterra, Leslie-Gower and Holling-Tanner non-linear prey-predator models. Volterra's principle and its applications. Gause competition model.

UNIT IV

Compartmental modelling - First and second order input-output systems, Dynamics of a multivariable system.

Practical

Fitting of mechanistic non-linear models; Application of Schaefer and Fox non-linear models; Fitting of compartmental models.

Suggested Readings

Draper NR & Smith H. 1998. *Applied Regression Analysis*. 3 Ed.rdJohn Wiley.

Efromovich S. 1999. *Nonparametric Curve Estimation*. Springer.

Fan J & Yao Q. 2003. *Nonlinear Time Series-Nonparametric and Parametric Methods*. Springer.

France J & Thornley JHM. 1984. *Mathematical Models in Agriculture*. Butterworths.

Harvey AC. 1996. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge Univ. Press.

Ratkowsky DA. 1983. *Nonlinear Regression Modelling: A Unified Practical Approach*. Marcel Dekker.

Ratkowsky DA. 1990. *Handbook of Nonlinear Regression Models*. Marcel Dekker.

Seber GAF & Wild CJ. 1989. *Non-linear Regression*. John Wiley.

Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.

STAT 617 ADVANCED TIME SERIES ANALYSIS 2+0

Objective

This is an advanced course in Time Series Analysis that aims at describing some advanced level topics in this area of research with a very strong potential of applications. This course also prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject to agricultural sciences.

Theory

UNIT I

Multivariate time series: modelling the mean, stationary VAR models: properties, estimation, analysis and forecasting, VAR models with elements of nonlinearity, Non-stationary multivariate time series: spurious regression, co-integration, common trends.

UNIT II

Volatility: Modelling the variance, The class of ARCH models: properties, estimation, analysis and forecasting, stochastic volatility, realized volatility, Extensions: IGARCH, ARCH-t, ARCD, Multivariate GARCH, Time-varying risk and ARCH-in-mean.

UNIT III

Structural time-series modelling: State space models, Kalman filter. Local level model, Local linear trend model, Seasonal models, Cyclical models. Nonlinear time-series models: Parametric and nonparametric approaches. Autoregressive conditional heteroscedastic model and its extensions. Threshold and Functional coefficient autoregressive models.

UNIT IV

Non-linear programming, Kuhn-Tucker sufficient conditions, Elements of multiple objective programming, Dynamic Programming, Optimal control theory - Pontryagin's maximum principle, Time-optimal control problems.

Suggested Readings

Box GEP, Jenkins GM & Reinsel GC. 2008. *Time Series Analysis: Forecasting and Control*. 3rd Ed. John Wiley.

Brockwell PJ & Davis RA. 1991. *Time Series: Theory and Methods*. 2nd Ed. Springer.

Chatfield C. 2004. *The Analysis of Time Series: An Introduction*. 6th Ed. Chapman & Hall/CRC.

Tong H. 1995. *Nonlinear Time Series: A Dynamical System Approach*. Oxford Univ. Press.

STAT 618 STOCHASTIC PROCESSES 2+0

Objective

This is a course on Stochastic Processes that aims at describing some advanced level topics in this area of research with a very strong potential of applications. This course also prepares students for undertaking research in this area. This also helps prepare students for applications of this important subject to agricultural sciences.

Theory

UNIT I

Introduction to stochastic process - classification according to state space and time domain. Finite and countable state Markov chains; time-homogeneity; Chapman-Kolmogorov equations, marginal distribution and finite dimensional distributions. Classification of Markov chain. Canonical form of transition probability matrix of a Markov chain. Fundamental matrix; probabilities of absorption from transient states into recurrent classes in a finite Markov chain, mean time for absorption. Ergodic state and Ergodic chain. Stationary distribution of a Markov chain, existence and evaluation of stationary distribution. Random walk and gamblers ruin problem.

UNIT II

Discrete state continuous time Markov process: Kolmogorov difference - differential equations. Birth and death process, pure birth process (Yule-Fury process). Immigration-Emigration process. Linear growth process, pure death process.

UNIT III

Renewal process: renewal process when time is discrete and continuous. Renewal function and renewal density. Statements of Elementary renewal theorem and Key renewal theorem.

UNIT IV

Stochastic process in biological sciences: Markov models in population genetics, compartmental analysis. Simple deterministic and stochastic epidemic model. General epidemic models-Karmack and McKendrick's threshold theorem. Recurrent epidemics.

UNIT V

Elements of queueing process; the queuing model M/M/1: steady state behaviors. Birth and death process in queuing theory- Multi channel models. Net work of Markovian queuing system.

UNIT VI

Branching process: Galton-Watson branching process. Mean and variance of size of nth generation, probability of ultimate extinction of a branching process. Fundamental theorem of branching process and applications.

UNIT VII

Wiener process- Wiener process as a limit of random walk. First passage time for Wiener process. Kolmogorov backward and forward diffusion equations and their applications.

Suggested Readings

Adke SR & Manjunath SM. 1984. *Finite Markov Processes*. John Wiley.

Bailey NTJ. 1964. *Elements of Stochastic Processes with Applications to the Natural Sciences*. Wiley Eastern.

Bartlett MS. 1955. *Introduction to Stochastic Processes*. Cambridge Univ. Press.

Basawa IV & Prakasa Rao BLS. 1980. *Statistical Inference for Stochastic Processes*. Academic Press.

Bharucha-Reid AT. 1960. *Elements of the Theory of Markov Processes and their Applications*. McGraw Hill.

Bhat BR. 2000. *Stochastic Models; Analysis and Applications*. New Age.

Cox DR & Miller HD. 1965. *The Theory of Stochastic Processes*. Methuen.

Draper NR & Smith H. 1981. *Applied Regression Analysis*. Wiley Eastern.

France J & Thornley JHM. 1984. *Mathematical Models in Agriculture*. Butterworths.

Karlin S & Taylor H.M. 1975. *A First Course in Stochastic Processes*. Vol. I. Academic Press.

Lawler GF. 1995. *Introduction to Stochastic Processes*. Chapman & Hall.

Medhi J. 2001. *Stochastic Processes*. 2nd Ed. Wiley Eastern.

Parzen E. 1962. *Stochastic Processes*. Holden-Day.

Prabhu NU. 1965. *Stochastic Processes*. Macmillan.

Prakasa Rao BLS & Bhat BR.1996. *Stochastic Processes and Statistical Inference*. New Age.

Ratkowsky DA. 1983. *Nonlinear Regression Modelling: a Unified Practical Approach*. Marcel Dekker.

Ratkowsky DA. 1990. *Handbook of Nonlinear Regression Models*. Marcel Dekker.

Seber GAF & Wild CJ. 1989. *Non-linear Regression*. John Wiley.

STAT 619 SURVIVAL ANALYSIS 2+0

Objective

The course deals with the study of demographic profiles and survival times. In-depth statistical properties and analysis is an important component of this course.

Theory

UNIT I

Measures of Mortality and Morbidity: Ratios and proportions, rates of continuous process, rates of repetitive events ,crude birth rate, Mortality measures used in vital statistics relationships between crude and age specific rates, standardized mortality ratios ,evaluation of person-year of exposed to risk in long term studies, prevalence and incidence of a disease, relative risk and odds ratio.

Survival Distribution: Survival functions, hazard rate, hazard function, review of survival distributions: exponential, Weibull, Gamma, Rayleigh, Pareto, Lognormal~ IFR and TFRA, Gompertz and Makeham. Gompertz and logistic distributions. Parametric (m.l.e) estimation. Types of Censoring: Type I, Type II, random and other types of censoring, right and left truncated distributions. Expectation and variance of future life time, series and parallel system of failures.

Life Tables: Fundamental and construction.

UNIT II

Complete Mortality data, Estimation of Survival Function : Empirical survival function , estimation of survival function from grouped mortality data, joint distribution of the number of deaths, distribution of the estimation P_i covariance of estimate, estimation of curves of deaths and central death rate and force of mortality rate .

Incomplete Mortality data (non-parametric models): Actuarial method, m.1.e method, moment and reduced sample method of estimation and their comparison. Product limit (Kaplan-Meier) method and cumulative hazard function (CHF) of estimation of survival function.

UNIT III

Fitting Parametric Survival Distribution : Special form of survival function cumulative hazard function (CHF) plots, Nelson's method of ungrouped data, construction of the likelihood function for survival data, least squares fitting, fitting a Gompertz distribution to grouped data.

Some tests of Goodness of fit: Graphical, Kolmogorov-Smirnov statistics for complete, censored and truncated data, Chi-Square test and Anderson-Darling A^2-statistics.

Comparison of Mortality Experiences: Comparison of two life tables, some distribution- free methods (two samples) for ungrouped data, Two samples Kolmogorov-Smirnov test, Wilcoxon test for complete data and modified Wilcoxon test for incomplete data .Gilbert and Gehan's test, mean and variance of Wilcoxon statistics, generalization of Gehan's test. Testing for Consistent Differences in Mortality : Mantel-Haenszel and log rank test. Generalized Mantel-Haenszel test (k-sample).

UNIT IV

Concomitant Variables: General parametric model for hazard function with observed concomitant variables. Additive and multiplicative models of hazard rate functions. Estimating multiplicative models, selection of

concomitant variables. Logistic linear model, Concomitant Variable regarded as random variable. Age of onset distributions: Models of onset distributions and their estimation.

Gompertz distribution, parallel system and Weibull distribution, Fatal short models of failure. Two component series system.

Suggested Readings

Anderson B. 1990. *Methodological Errors in Medical Research*. Blackwell.

Armitage P & Berry G. 1987. *Statistical Methods in Medical Research*. Blackwell.

Collett D. 2003. *Modeling Survival Data in Medical Research*. Chapman & Hall.

Cox DR & Oakes D. 1984. *Analysis of Survival Data*. Chapman & Hall.

Elandt-Johnson RC & Johnson NL. 1980. *Survival Models and Data Analysis*. John Wiley.

Everitt BS & Dunn G. 1998. *Statistical Analysis of Medical Data*. Arnold.

Hosmer DW Jr. & Lemeshow S. 1999. *Applied Survival Analysis: Regression Modeling or Time to Event*. John Wiley.

Kalbfleisch JD & Prentice. RL 2002. *The Statistical Analysis of Failure Time Data*. John Wiley.

Klein JP & Moeschberger ML. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.

Kleinbaum DG & Klein M. 2002. *Logistic Regression*. Springer.

Kleinbaum DG & Klein M. 2005. *Survival Analysis*. Springer.

Lawless JF. 2003. *Statistical Models and Methods for Lifetime Data*. 2nd Ed. John Wiley.

Lee ET. 1980. *Statistical Methods for Survival Data Analysis*. Lifetime Learning Publ.

STAT 620 ADVANCED BIOINFORMATICS 2+0

Objective

This is a course on Bioinformatics that aims at exposing the students to some advanced statistical and computational techniques related to bioinformatics. This course would prepare the students in understanding bioinformatics principles and their applications.

Theory

UNIT I

Genomic databases and analysis of high-throughput data sets, sequence annotation, ESTs, SNPs. BLAST and related sequence comparison methods. EM algorithm and other statistical methods to discover common motifs in biosequences. Multiple alignment and database search using motif models, ClustalW and others. Concepts in phylogeny. Gene prediction based on codons, Decision trees, Classificatory analysis, Neural Networks, Genetic algorithms, Pattern recognition, Hidden Markov models.

UNIT II

Computational analysis of protein sequence, structure and function. Expression profiling by microarray/gene chip, proteomics etc., Multiple alignment of protein sequences, Modelling and prediction of structure of proteins, Designer proteins, Drug designing.

UNIT III

Analysis of one DNA sequence (Modeling signals in DNA; Analysis of patterns; Overlaps and Generalizations), Analysis of multiple DNA or protein sequences (Alignment algorithms - Gapped global comparisons and Dynamic programming; use of linear gap models; protein sequences and substitution matrices - BLOSUM, PAM; Multiple sequences), BLAST (Comparison of two aligned sequences - Parameter calculation; Choice of a score; Bounds for P-value; Normalized and Bit scores, Karlin - Altschul sum statistic; comparison of two unaligned sequences; Minimum significance Lengths).

UNIT IV

Markov chains (MC with no absorbing states; Higher order Markov dependence; patterns in sequences; Markov chain Monte Carlo - Hastings-Metropolis algorithm, Gibbs sampling, Simulated Annealing; MC with absorbing States, Continuous-Time Markov chains) Hidden Markov Models (Forward and Backward algorithm; Viterbi algorithms; Estimation algorithm;

UNIT V

Modeling protein families; Multiple sequence alignments; Pfam; Gene finding), Computationally intensive methods (Classical estimation methods; Bootstrap estimation and Confidence Intervals; Hypothesis testing; Multiple Hypothesis testing), Evolutionary models (Models of Nucleotide substitution; Discrete time models - The Jukes-Cantor Model, The Kimura Model, The Felsenstein Model; Continuous-time models),

UNIT VI

Phylogenetic tree estimation (Distances; Tree reconstruction - Ultrametric and Neighbor-Joining cases; Surrogate distances; Tree reconstruction; Parsimony and Maximum Likelihood; Modeling, Estimation and Hypothesis Testing;) Neural Networks (Universal Approximation Properties; Priors and Likelihoods, Learning Algorithms - Backpropagation; Sequence encoding and output interpretation; Prediction of Protein Secondary Structure; Prediction of Signal Peptides and their cleavage sites; Application for DNA and RNA Nucleotide Sequences), Analysis of SNPs and Haplotypes.

Suggested Readings

Baldi P & Brunak S. 2001. *Bioinformatics: The Machine Learning Approach*. MIT Press.

Baxevanis AD & Francis BF. (Eds.). 2004. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. John Wiley.

Duda RO, Hart PE & Stork DG. 1999. *Pattern Classification*. John Wiley.

Ewens WJ & Grant GR. 2001. *Statistical Methods in Bioinformatics*. Springer.

Jones NC & Pevzner PA. 2004. *Introduction to Bioinformatics Algorithims*. The MIT Press.

Koskinen T. 2001. *Hidden Markov Models for Bioinformatics*. Kluwer.

Krane DE & Raymer ML. 2002. *Fundamental Concepts of Bio-informatics*. Benjamin / Cummings.

Krawetz SA & Womble DD. 2003. *Introduction to Bioinformatics: A Theoretical and Practical Approach*. Humana Press.

Lesk AM. 2002. *Introduction to Bio-informatics*. Oxford Univ. Press.

Linder E & Seefeld K. 2005. *R for Bioinformatics*. O'Reilly & Associates.
Percus JK. 2001. *Mathematics of Genome Analysis*. Cambridge Univ. Press.
Sorensen D & Gianola D. 2002. *Likelihood, Bayesian and MCMC Methods in Genetics*. Springer.
Tisdall JD. 2001. *Mastering Perl for Bioinformatics*. O'Reilly & Associates.
Wang JTL, Zaki MJ, Toivonen HTT & Shasha D. 2004. *Data Mining in Bioinformatics*. Springer.
Wu CH & McLarty JW. 2000. *Neural Networks and Genome Informatics*. Elsevier.
Wunschiers R. 2004. *Computational Biology Unix/Linux, Data Processing and Programming*. Springer.
Yang MCC. 2000. *Introduction to Statistical Methods in Modern Genetics*. Taylor & Francis.

STAT 621 ADVANCED ECONOMETRICS 2+0

Objective

This is a course on Econometrics aims at exposing the students to some advanced level econometric methods and their applications to agricultural situations.

Theory

UNIT I

Quantile regression, binary quantile regression, extreme values, copula, loss functions, Point and interval forecasting, unconditional and conditional forecasting, forecasting with serially correlated errors, bootstrap: asymptotic expansion, bootstrap consistency, asymptotic refinement, recent developments for dependent time series

UNIT II

Multivariate time series: modelling the mean, stationary VAR models: properties, estimation, analysis and forecasting, VAR models with elements of nonlinearity, Non-stationary multivariate time series: spurious regression, co-integration, common trends; Volatility: Modelling the variance, The class of ARCH models: properties, estimation, analysis and forecasting, stochastic volatility, realized volatility.

UNIT III

Basic Concepts of Bayesian Inference, Probability and Inference, Posterior Distributions and Inference, Prior Distributions. The Bayesian linear model and autoregressive (AR) processes; Model selection with marginal likelihoods and fractional priors, Comparison of Bayesian Methods with Classical approaches, Bayes risk and their applications, and Sample Selection Monte Carlo integration, importance sampling and Gibbs sampling, The Regression Model with General Error Covariance Matrix, Qualitative Choice Models, Bayesian information criterion (BIC), Markov Chain Monte Carlo (MCMC) Model Composition and stochastic search variable selection, BUGS [Bayesian Inference Using Gibbs Sampling] , BUCC [Bayesian Analysis, Computation and Communication].

- Technometrics

Suggested Readings

Banerjee A, Dolado J, Galbraith J & Hendry DF. 1993. *Co-integration, Error Correction, and the Econometric Analysis of Nonstationary Data*. Oxford Univ. Press.

Bauwens L, Lubrano M & Richard JF. 1999. *Bayesian Inference in Dynamics of Econometric Models.* Oxford Univ. Press.

Carlin BP & Louis TA. 1996. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall.

Gilks WR, Richardson S & Spiegelhalter D. 1996. *MCMC in Practice*. Chapman & Hall.

Greenberg E. 2008. *Introduction to Bayesian Econometrics.* Cambridge Univ. Press.

Hamilton JD. 1994. *Time Series Analysis*. Princeton Univ. Press.

Judge GG, Griffith WE, Hill RC, Lee CH & Lutkepohl H. 1985. *The Theory and Practice of Econometrics*. 2nd Ed. John Wiley.

Koop G, Poirier D & Tobias J. 2007. *Bayesian Econometric Methods.* Cambridge Univ. Press.

Koop G. 2003. *Bayesian Econometrics*. John Wiley.

Lancaster A. 2004. *An Introduction to Modern Bayesian Econometrics*. Blackwell.

Pindyck RS & Rubinfeld DL. 1981. *Econometric Models and Economic Forecasts.* McGraw Hill.

STAT 651 RECENT ADVANCES IN THE FIELD OF SPECIALIZATION 1+0

Objective

To familiarize the students with the recent advances in the areas of their specialization to prepare them for undertaking research.

Theory

Recent advances in the field of specialization - sample surveys / design of experiments /statistical genetics / statistical modeling / econometrics / statistical inference, etc. will be covered by various speakers from the University / Institute as well as from outside the University / Institute in the form of seminar talks.

List of Journals

Agricultural Statistics

- American Statistician
- Annals of Institute of Statistical Mathematics
- Annals of Statistics
- Australian and New Zealand Journal of Statistics
- Biometrical Journal
- Biometrics
- Biometrika
- Bulletin of Calcutta Statistical Association
- Canadian Journal of Statistics
- Communication in Statistics (Simulation & Computation)
- Communication in Statistics (Theory & and Methods)
- Experimental Agriculture
- Institute of Mathematical Statistics Bulletin (IMSB)
- Journal of American Statistical Association
- Journal of Applied Statistics
- Journal of the Indian Society of Agricultural Statistics
- Journal of the International Statistical Review
- Journal of Statistical Planning and Inference
- Journal of Statistical Theory and Practice
- Journal of Statistics, Computer and Applications
- Journal of Royal Statistical Society, Series A
- Journal of Royal Statistical Society, Series B
- Journal of Royal Statistical Society, Series C
- Metrika
- Metron
- Scandinavian Journal of Statistics (Theory & Applied)
- Sankhya
- Statistica
- Statistical Science
- Statistics and Probability Letters
- Technometrics

Computer Application

- ACM Transactions on Knowledge Discovery from Data
- Applied Intelligence - The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies
- Computational Statistics & Data Analysis, Elsevier Inc.
- Computers and Electronics in Agriculture, Elsevier Inc.
- Data Mining and Knowledge Discovery: An International Journal (DMKD)
- Expert Systems with Applications, Elsevier Inc.
- IEEE Transactions on Knowledge and Data Engineering
- IEEE Transactions on Neural Networks

- IEEE Transactions on Pattern Analysis and Machine Intelligence
- International Journal of Computing and Information Sciences
- International Journal of Information and Management Sciences
- International Journal of Information Technology
- Journal of Artificial Intelligence Research
- Journal of Combinatorics, Information and System Sciences
- Journal of Computer Sciences and Technology
- Journal of Computer Society of India
- Journal of Indian Society of Agricultural Statistics
- Journal of Intelligent Information Systems - Integrating Artificial Intelligence and Database Technologies
- Journal of Machine Learning Research
- Journal of Statistics, Computer and Applications
- Journal of Systems and Software
- Journal of Theoretical and Applied Information Technology
- Knowledge and Information Systems: An International Journal (KAIS)
- Lecture Notes in Computer Science, Springer Verlag.
- Machine Learning
- Transactions on Rough Set

e-Resources

- Design Resources Server. *Indian Agricultural Statistics Research Institute(ICAR), New Delhi 110 012, India*. www.iasri.res.in/design.
- Design Resources: www.designtheory.org
- Free Encyclopedia on Design of Experiments
- http://en.wikipedia.org/wiki/Design_of_experiments
- Statistics Glossary http://www.cas.lancs.ac.uk/glossary_v1.1/main.html.
- Electronic Statistics Text Book: http://www.statsoft.com/textbook/stathome.html.
- Hadamard Matrices http://www.research.att.com/~njas/hadamard;
- Hadamard Matrices http://www.uow.edu.au/~jennie/WILLIAMSON/williamson.html.
- Course on Experimental design: http://www.stat.sc.edu/~grego/courses/stat706/.
- Learning Statistics: http://freestatistics.altervista.org/en/learning.php.
- Free Statistical Softwares: http://freestatistics.altervista.org/en/stat.php.
- Statistics Glossary http://www.cas.lancs.ac.uk/glossary_v1.1/main.html.
- Statistical Calculators: http://www.graphpad.com/quickcalcs/index.cfm
- SAS Online Doc 9.1.3: http://support.sas.com/onlinedoc/913/docMainpage.jsp

Suggested Broad Topics for Research

Agricultural Statistics

- Design and analysis of multi-response experiments
- Design and analysis of micro-array experiments
- Design and analysis of experiments for precision agriculture
- Design and analysis of agroforestry experiments
- Bayesian designing of experiments, Bayesian optimality and Bayesian analysis of experimental data
- Computer aided search of efficient experimental designs for various experimental settings
- Fractional factorials including search designs, supersaturated designs, computer experiments, etc.
- Statistical techniques in bioinformatics, biotechnology, microbiology, genomics, etc.
- Optimality aspects and robustness of designs against several disturbances under various experimental settings (single factor, multi-factor, nested classifications, etc.)
- Small area estimation
- Computer intensive techniques in sample surveys
- Analysis of survey data, regression analysis, categorical data analysis, analysis of complex survey data
- Assessment and impact survey methodologies, valuation of natural resources, its degradation, depletion, etc.
- Linear and non-linear modeling of biological and economical phenomena
- Non-linear time series modeling
- Non-linear stochastic modeling
- Forecast models for both temporal and spatial data
- Innovative applications of resampling techniques
- Applications of remote sensing, GIS, ANN etc. in modeling various phenomena
- Econometric models for risk, uncertainty, insurance, market analysis, technical efficiency, policy planning, etc.
- Statistical studies on value addition to crop produce

Computer Application

- Web solutions in agriculture
- Decision Support/Expert Systems/Information Management Systems in Agriculture
- Software for Statistical Data Analysis
- Modelling and Simulation of Agricultural Systems
- Application Software for GIS and Remote Sensing
- Office Automation and Management System