

MASTER OF SCIENCE(M. SC.)
IN
STATISTICS



Division of Statistics & Computer Science,
Faculty of Basic Sciences,
Sher-e-Kashmir University of Agricultural Sciences & Technology - Jammu.
Main Campus, Chatha, Jammu (J&K)- 180009.

M. Sc. (Statistics)

Description of Courses (Course Content):

STAT 501

PROBABILITY THEORY

2+0

SEM I

Theory

UNIT I

Basic concepts of probability. Elements of measure theory: class of sets, field, sigma field, minimal sigma field, Borel sigma field in \mathbb{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, Jensen, 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.

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*. 4th 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.

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.

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 nonrandomized 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 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.

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, 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.
- Kemphorne, 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.

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.

UNIT V

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

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.

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.

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

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.

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.

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 R² 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*. 3rd Ed. John Wiley.

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

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

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

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.

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.

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. 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.

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.

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.

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.

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.

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.

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.

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 nonparametric 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.

Supporting Courses (Course Content):

STAT 551

MATHEMATICAL METHODS – I

3+0

SEM I

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, 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. <i>Elements of Real Analysis</i> .	John Wiley.
Chatterjee SK.	1970. <i>Mathematical Analysis</i> .	Oxford & IBH.
Gibson GA.	1954. <i>Advanced Calculus</i> .	Macmillan.
Henrice P.	1964. <i>Elements of Numerical Analysis</i> .	John Wiley.
Hildebrand FB.	1956. <i>Introduction to Numerical Analysis</i> .	Tata McGraw Hill.
Priestley HA.	1985. <i>Complex Analysis</i> .	Clarenton Press.
Rudin W.	1985. <i>Principles of Mathematical Analysis</i> .	McGraw Hill.
Sauer T.	2006. <i>Numerical Analysis with CD-Rom</i> .	Addison Wesley.
Scarborough JB.	1976. <i>Numerical Mathematical Analysis</i> .	Oxford & IBH.
Stewart J.	2007. <i>Calculus</i> .	Thompson.
Thomas GB Jr. & Finney RL.	1996. <i>Calculus</i> .	9 th Ed. Pearson Edu.

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, Gramian 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.

Service Courses (For M.Sc. and Ph.D. programs of other disciplines)

Description of Course (Course Content)

STAT 553 MATHEMATICAL METHODS FOR APPLIED SCIENCES 2+0 SEM I

Theory

UNIT I

Variables and functions; limit and continuity. Specific functions. Differentiation: theorems of differentiation, differentiation of logarithmic, trigonometric, exponential and inverse functions, function of a function, derivative of higher order, partial derivatives. Application of derivatives in agricultural research; determination of points of inflexion, maxima and minima in optimization, etc.

UNIT II

Integration as a reverse process of differentiation, methods of integration, reduction formulae, definite integral; Applications of integration in agricultural research with special reference to economics and genetics, engineering, etc.

UNIT III

Vectors and vector spaces, Matrices, notations and operations, laws of matrix algebra; transpose and inverse of matrix; Eigen values and eigen vectors. Determinants - evaluation and properties of determinants, application of determinants and matrices in solution of equation for economic analysis.

UNIT IV

Set theory-set operations, finite and infinite sets, operations of set, function defined in terms of sets.

Suggested Readings

- Harville DA. 1997. *Matrix Algebra from a Statistician's Perspective*. Springer.
Hohn FE. 1973. *Elementary Matrix Algebra*. Macmillan.
Searle SR. 1982. *Matrix Algebra Useful for Statistics*. John Wiley.
Stewart J. 2007. *Calculus*. Thompson.
Thomas GB. Jr. & Finney RL. 1996. *Calculus*. 9th Ed. Pearson Edu.

STAT 554 STATISTICAL METHODS FOR APPLIED SCIENCES 3+1 SEM I

Theory

UNIT I

Descriptive statistics. Exploratory data analysis; Theory of probability. Random variable and mathematical expectation.

UNIT II

Discrete and continuous probability distributions: Binomial, Poisson, Negative Binomial, Normal distributions, and their applications. Concept of sampling distribution: chi-

square, t and F distributions. Tests of significance based on Normal, chi-square, t and F distributions. Large sample theory.

UNIT III

Introduction to theory of estimation and confidence-intervals. Correlation and regression. Simple and multiple linear regression model, estimation of parameters, predicted values and residuals, correlation, partial correlation coefficient, multiple correlation coefficient, rank correlation, test of significance of correlation coefficient and regression coefficients. Coefficient of determination. Polynomial regression models and their fitting. by least squares method

UNIT IV

Non-parametric tests - sign, Wilcoxon, Mann-Whitney U-test, Wald Wolfowitz run test, Run test for the randomness of a sequence. Median test, Kruskal- Wallis test, Friedman two-way ANOVA by ranks. Kendall's coefficient of concordance.

UNIT V

Introduction to multivariate analytical tools- Hotelling's T^2 Tests of hypothesis about the mean vector of a multinormal population. Classificatory problems and discriminant function, D^2 -statistic and its applications; Cluster analysis, principal component analysis.

Practical

Exploratory data analysis, Descriptive statistics; Fitting of distributions ~ Binomial, Poisson, Negative Binomial, Normal; Large sample tests, testing of hypothesis based on exact sampling distributions ~ chi square, t and F ; Confidence interval estimation and point estimation of parameters of binomial, Poisson and Normal distribution; Correlation and regression analysis, fitting of orthogonal polynomial regression; applications of dimensionality reduction and discriminant function analysis; Nonparametric tests.

Suggested Readings

- Anderson TW. 1958. *An Introduction to Multivariate Statistical Analysis*. John Wiley.
Dillon WR & Goldstein M. 1984. *Multivariate Analysis - Methods and Applications*. John Wiley.
Goon AM, Gupta MK & Dasgupta B. 1977. *An Outline of Statistical Theory*. Vol. I. The World Press.
Goon AM, Gupta MK & Dasgupta B. 1983. *Fundamentals of Statistics*. Vol. I. The World Press.
Hoel PG. 1971. *Introduction to Mathematical Statistics*. John Wiley.
Hogg RV & Craig TT. 1978. *Introduction to Mathematical Statistics*. Macmillan.
Morrison DF. 1976. *Multivariate Statistical Methods*. McGraw Hill.
Siegel S, Johan N & Casellan Jr. 1956. *Non-parametric Tests for Behavior Sciences*. John Wiley.

STAT 555 EXPERIMENTAL DESIGNS

2+1

SEM II

Theory

UNIT I

Need for designing of experiments, characteristics of a good design. Basic principles of designs- randomization, replication and local control.

UNIT II

Uniformity trials, size and shape of plots and blocks; Analysis of variance; Completely randomized design, randomized block design and Latin square design.

UNIT III

Factorial experiments, (symmetrical as well as asymmetrical). orthogonality and partitioning of degrees of freedom, Confounding in symmetrical factorial experiments, Factorial experiments with control treatment.

UNIT IV

Split plot and strip plot designs; Analysis of covariance and missing plot techniques in randomized block and Latin square designs; Transformations, crossover designs, balanced incomplete block design, resolvable designs and their applications ~ Lattice design, - concepts, randomization procedure, analysis and interpretation of results. Response surfaces.

UNIT V

Bioassays- direct and indirect, indirect assays based on quantal dose response, parallel line and slope ratio assays potency estimation.

Practical

Uniformity trial data analysis, formation of plots and blocks, Fairfield Smith Law; Analysis of data obtained from CRD, RBD, LSD; Analysis of factorial experiments without and with confounding; Analysis with missing data; Split plot and strip plot designs; Transformation of data; Analysis of BIBD; Fitting of response surfaces.

Suggested Readings

- Cochran WG & Cox GM. 1957. *Experimental Designs*. 2nd Ed. John Wiley.
Dean AM & Voss D. 1999. *Design and Analysis of Experiments*. Springer.
Federer WT. 1985. *Experimental Designs*. MacMillan.
Fisher RA. 1953. *Design and Analysis of Experiments*. Oliver & Boyd.
Nigam AK & Gupta VK. 1979. *Handbook on Analysis of Agricultural Experiments*. IASRI Publ.
Pearce SC. 1983. *The Agricultural Field Experiment: A Statistical Examination of Theory and Practice*. John Wiley.

STAT 556 SAMPLING TECHNIQUES

2+1

SEM II

Theory

UNIT I

Concept of sampling, sample survey vs complete enumeration, planning of sample survey, sampling from a finite population.

UNIT II

Simple random sampling, sampling for proportion, determination of sample size; inverse sampling, Stratified sampling.

UNIT III

Cluster sampling, PPS sampling, Multi-stage sampling, double sampling, systematic sampling; Use of auxiliary information at estimation as well as selection stages.

UNIT IV

Ratio and regression estimators. Construction and analysis of survey designs, sampling and non-sampling errors; Preparation of questionnaire Non-sampling errors.

Practical

Random sampling ~ use of random number tables, concepts of unbiasedness, variance, etc.; simple random sampling, determination of sample size; Exercises on inverse sampling, stratified sampling, cluster sampling and systematic sampling; Estimation using ratio and regression estimators; Estimation using multistage design, double sampling and PPS sampling.

Suggested Readings

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

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

Singh D, Singh P & Kumar P. 1982. *Handbook on Sampling Methods*. IASRI Publ.

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.

STAT 557 APPLIED REGRESSION ANALYSIS

2+1

SEM I

Theory

UNIT I

Introduction to correlation analysis and its measures; Correlation from grouped data, Biserial correlation, Rank correlation; Testing of population correlation coefficients; Multiple and partial correlation coefficients and their testing.

UNIT II

Problem of correlated errors; Auto correlation; Durbin Watson Statistics; Removal of auto correlation by transformation; Analysis of collinear data; Detection and correction of multicollinearity; Regression analysis; Method of least squares for curve fitting; Testing of regression coefficients; Multiple and partial regressions.

UNIT III

Examining the multiple regression equation; Concept of weighted least squares; regression equation on grouped data; Various methods of selecting the best regression equation; regression approach applied to analysis of variance in one way classification.

UNIT IV

Heteroscedastic models, Concept of nonlinear regression and fitting of quadratic, exponential and power curves; Economic and optimal dose, Orthogonal polynomial.

Practical

Correlation coefficient, various types of correlation coefficients, partial and multiple, testing of hypotheses; Multiple linear regression analysis, partial regression coefficients, testing of hypotheses, residuals and their applications in outlier detection; Handling of correlated errors, multicollinearity; Fitting of quadratic, exponential and power curves, fitting of orthogonal polynomials.

Suggested Readings

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

Ezekiel M. 1963. *Methods of Correlation and Regression Analysis*. John Wiley.

Kleinbaum DG, Kupper LL, Muller KE & Nizam A. 1998. *Applied Regression Analysis and Multivariable Methods*. Duxbury Press.

Koutsoyiannis A. 1978. *Theory of Econometrics*. MacMillan.

Kutner MH, Nachtsheim C J & Neter J. 2004. *Applied Linear Regression Models*. 4th Ed. with Student CD. McGraw Hill.

STAT 558 DATA ANALYSIS USING STATISTICAL PACKAGES 2+1 SEM II

Theory

UNIT I

Use of Software packages for: Summarization and tabulation of data; Descriptive statistics; Graphical representation of data, Exploratory data analysis.

UNIT II

Fitting and testing the goodness of fit of discrete and continuous probability distributions; Testing of hypothesis based on large sample test statistics; Testing of hypothesis using chi-square, t and F statistics.

UNIT III

Concept of analysis of variance and covariance of data for single factor, multi-factor, one-way and multi-classified experiments, contrast analysis, multiple comparisons, Analyzing crossed and nested classified designs.

UNIT IV

Analysis of mixed models; Estimation of variance components; Testing the significance of contrasts; Correlation and regression including multiple regression.

UNIT V

Discriminant function; Factor analysis; Principal component analysis; Analysis of time series data, Fitting of non-linear models; Time series data

Practical

Use of software packages for summarization and tabulation of data, obtaining descriptive statistics, graphical representation of data. Robust Estimation, Testing linearity and normality assumption, Estimation of trimmed means etc., Cross tabulation of data including its statistics, cell display and table format and means for different sub-classifications; Fitting and testing the goodness of fit of probability distributions; Testing the hypothesis for one sample t -test, two sample t -test, paired t -test, test for large samples - Chi-squares test, F test, One way analysis of variance, contrast and its testing, pairwise comparisons; Multiway classified analysis of variance - cross-classification, nested classification, factorial set up, fixed effect

models, random effect models, mixed effect models, estimation of variance components; Bivariate and partial correlation, Distances - to obtain a distance matrix, dissimilarity measures, similarity measures; Linear regression, Multiple regression, Regression plots, Variable selection, Regression statistics, Fitting of growth models - curve estimation models, examination of residuals; Discriminant analysis - fitting of discriminant functions, identification of important variables, Factor analysis. Principal component analysis - obtaining principal component, spectral composition; Analysis of time series data .

Suggested Readings

- Anderson CW & Loynes RM. 1987. *The Teaching of Practical Statistics*. John Wiley.
- Atkinson AC. 1985. *Plots Transformations and Regression*. Oxford University Press.
- Chambers JM, Cleveland WS, Kleiner B & Tukey PA. 1983. *Graphical Methods for Data Analysis*. Wadsworth, Belmont, California.
- Chatfield C & Collins AJ. 1980. *Introduction to Multivariate Analysis*. Chapman & Hall.
- Chatfield C. 1983. *Statistics for Technology*. 3rd Ed. Chapman & Hall.
- Chatfield C. 1995. *Problem Solving: A Statistician's Guide*. Chapman & Hall.
- Cleveland WS. 1985. *The Elements of Graphing Data*. Wadsworth, Belmont, California.
- Ehrenberg ASC. 1982. *A Primer in Data Reduction*. John Wiley.
- Erickson BH & Nosanchuk TA. 1992. *Understanding Data*. 2nd Ed. Open University Press, Milton Keynes.
- Snell EJ & Simpson HR. 1991. *Applied Statistics: A Handbook of GENSTAT Analyses*. Chapman & Hall.
- Sprent P. 1993. *Applied Non-parametric Statistical Methods*. 2nd Ed. Chapman & Hall.
- Tufte ER. 1983. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Conn.
- Velleman PF & Hoaglin DC. 1981. *Application, Basics and Computing of Exploratory Data Analysis*. Duxbury Press.
- Weisberg S. 1985. *Applied Linear Regression*. John Wiley.
- Wetherill GB. 1982. *Elementary Statistical Methods*. Chapman & Hall.
- Wetherill GB. 1986. *Regression Analysis with Applications*. Chapman & Hall.