Course Name: Mathematics for Applied Sciences Course No.: STAT – 501 Credit Hour: 2 (2+0)

Objective

This course is meant for students who do not have sufficient background of Mathematics. The students would be exposed to elementary mathematics that would prepare them to study their main courses that involve knowledge of Mathematics. The students would get an exposure to Linear Algebra, differentiation, integration and differential equations etc.

Theory

UNIT I

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

UNIT II

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, Solutions of Linear Equations.

UNIT III

Variables and functions, limits and continuity of specific functions. Differentiation: theorems of differentiation, differentiation of logarithmic, trigonometric, exponential and inverse functions, Differentiation of function of a function, derivatives of higher order, partial derivatives. Application of derivatives, determination of points of inflexion, maxima and minima.

UNIT IV

Integration, methods of integration, reduction formulae, definite and indefinite integral, Applications of integration in Agriculture, Differential Equations.

Course Title: Statistical Methods for Applied Sciences Course No.: STAT – 502 Credit Hour – 4(3+1)

Objective

This course is meant for students who do not have sufficient backgroundof Statistical Methods. The students would be exposed to concepts of statistical methods and statistical inference that would help them in understanding the importance of statistics. It would also help them in understanding the concepts involved in data presentation, analysis and interpretation. The students would get an exposure to presentation of data, probability distributions, parameter estimation, tests of significance, regression and multivariate analytical techniques.

Theory

UNIT I

Box-plot, 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 distribution, Beta and Gamma 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.

UNIT III

Introduction to theory of estimation and confidence-intervals, Simple and multiple correlation coefficient, partial correlation, rank correlation, Simple and multiple linear regression model, test of significance of correlation coefficient and regression coefficients, Coefficient of determination, Fitting of quadratic models.

UNIT IV

Non-parametric tests - sign, Wilcoxon, Mann-Whitney U-test, Run test for the randomness of a sequence. Median test.

UNIT V

Introduction to ANOVA: One way and Two Way, Introduction to Sampling Techniques, Introduction to Multivariate Analysis, Transformation of Data.

Practical

Exploratory data analysis, 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 Correlation and regression analysis, fitting of Linear and Quadratic Model; Non- parametric tests. ANOVA: One way, Two Way, SRS

Course Title: Experimental Designs Course No.: STAT – 511 Credit Hour: 3(2+1)

Objective

This course is meant for students of agricultural and animal sciences other than Agricultural Statistics. Designing an experiment is an integrated component of research in almost all sciences. The students would be exposed to concepts of Design of Experiments so as to enable them to understand the concepts involved in planning, designing their experiments and analysis of experimental data.

Theory

UNIT I

Need for designing of experiments, characteristics of a good design. Basic principles of designsrandomization, 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. Concept of confounding.

UNIT IV

Split plot and strip plot designs, analysis of covariance and missing plot techniques in randomized block and Latin square designs; Transformations, Balanced Incomplete Block Design, resolvable designs and their applications, Lattice design, alpha design - concepts, randomization procedure, analysis and interpretation of results. Response surfaces. Combined analysis.

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; Analysis with missing data; Split plot and strip plot designs.

Course Title: Basic Sampling Techniques Course No.: STAT – 512 Credit Hour - 2+1

Objective

This course is meant for students of agricultural and animal sciences other than Statistics. The students would be exposed to elementary sampling techniques. It would help them in understanding the concepts involved in planning and designing their surveys, presentation of survey data analysis of survey data and presentation of results. This course would be especially important to the students of social sciences.

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 with and without replacement, sampling for proportion, determination of sample size, inverse sampling, Stratified sampling.

UNIT III

Cluster sampling, Multi-stage sampling, systematic sampling; Introduction to PPS sampling, UNIT IV

Use of auxiliary information at estimation, Ratio product and regression estimators. Double Sampling, sampling and non-sampling errors.

Practical

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

Course Name: Applied Regression Analysis Course No.: STAT 521 Credit Hours: 3(2+1)

Objective

This course is meant for students of all disciplines including agricultural and animal sciences. The students would be exposed to the concepts of correlation and regression. Emphasis will be laid on diagnostic measures such as autocorrelation, multi collinearity and heteroscedasticity. This course would prepare students to handle their data for analysis and interpretation.

Theory

UNIT I

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

UNIT II

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

UNIT III

Diagnostic of multiple regression equation; Concept of weighted least squares; regression equation on grouped data; Various methods of selecting the best regression equation.

UNIT IV

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, multi collinearity; Fitting of quadratic, exponential and power curves, fitting oforthogonal polynomials.



Course Title: Data Analysis Using Statistical Packages Course No.: STAT 522 Credit Hours: 3(2+1)

Objective

This course is meant for exposing the students in the usage of various statistical packages for analysis of data. It would provide the students a hands on experience in the analysis of their research data. This course is useful to all disciplines.

Theory

UNIT I

Introduction to various statistical packages: Excel, R, SAS, SPSS. Data Preparation; Descriptive statistics; Graphical representation of data, Exploratory data analysis.

UNIT II

Test for normality; Testing of hypothesis using chi-square, t and F statistics and Z-test.

UNIT III

Data preparation for ANOVA and ANCOVA, Factorial Experiments, contrast analysis, multiple comparisons, Analyzing crossed and nested classified designs.

UNIT IV

Analysis of mixed models; Estimation of variance components; Correlation and regression analysis, Probit, Logit and Tobit Models.

UNIT V

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

Practical

Use of software packages for summarization and tabulation of data, obtaining descriptive statistics, graphical representation of data. 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, Designs for Factorial Experiments, fixed effect models, random effect models, mixed effect models, estimation of variance components, Linear regression, Multiple regression, Regression plots, Discriminant analysis - fitting of discriminant functions, identification of important variables, Factor analysis. Principal component analysis - obtaining principal component.

Course Title: Probability Theory Course No.: STAT – 552 Credit Hour: 2(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 caseonly). 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. Chebyshev, Markov, Cauchy-Schwartz, Sequence of random variables and modes of convergence (convergence in distribution in probability, almost surely, and quadratic mean) and their interrelations.

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 and simple applications.

Course Title: Statistical Methods Course No. : STAT – 553 Credit Hour: 3(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). 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. Categorical data analysis, Association between attributes. Variance Stabilizing Transformations.

UNIT V

Order statistics, distribution of rth order statistics, joint distribution of several order statistics and their functions, marginal distributions of order statistics.

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.

Course Title: Statistical Inference Course No.: STAT – 562 Credit Hour: 3(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: 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 their properties. Interval estimation-Confidence level, shortest length CI. CI for the parameters of Normal, Exponential, Binomial and Poisson distributions.

UNIT III

Fundamentals 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, asymptotic properties of LR tests with applications (including homogeneity of means and variances). Relation between confidence interval estimation and testing of hypothesis.

UNIT IV

Sequential Probability ratio test, Properties of SPRT. 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. 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.

Course Title: Design of Experiments Course No.: STAT – 563 Credit Hour: 3(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 24 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 andblocks, Randomization procedure.

UNIT III

Basic designs - completely randomized design, randomized complete block design and Latin square design; Construction of 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 BIB designs. Partially balanced incomplete block designs with two associate classes - properties, analysis and construction, Lattice designs, alpha designs, cyclic designs, augmented designs.

UNIT V

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

UNIT VI

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; 2ⁿ, 3ⁿ 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.

Course Title: Sampling Techniques Course No.: STAT – 564; Credit Hour: 3(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 enumeration, probability sampling, samplespace, sampling design, sampling strategy; Determination of sample size; Confidence-interval; Simple random sampling, Estimation of population proportion, Stratified random sampling, Proportional allocation and optimal allocation, Inverse sampling.

UNIT II

Ratio, Product 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. Unbiased ratio type estimators

UNIT III

Non-sampling errors – sources and classification, Non-response in surveys, Randomized response techniques, Response errors/Measurement error – interpenetrating sub-sampling.

UNIT IV

PPS Sampling with and without replacement, 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.

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.

Course Title: Statistical Genetics Course No.: STAT – 565 Credit Hour: 3(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.

Course Title: Multivariate Analysis Course No.: STAT – 571 Credit Hour: 3(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 datareduction 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 T² and Mahalanobis D² statistics. Null distribution of Hotelling's T². Rao's U statistics and its distribution. Wilks' λ criterion and its properties. Concepts of discriminant analysis, computation of linear discriminant function, classification between k (≥ 2) multivariate normal populationsbased on LDF and MahalanobisD².

UNIT III

Principal Component Analysis, factor analysis. Canonical variables and canonical correlations. Cluster analysis: similarities and dissimilarities of qualitative and quantitative characteristics, Hierarchical clustering. Single, Complete and Average linkage methods. K-means cluster analysis.

UNIT IV

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

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.

Course Title: Regression Analysis Course No.: STAT – 572 Credit Hour: 2(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; residual analysis ~ Studentized residuals, applications of residuals in detecting outliers, identification of influential observations. Lack of fit, Pure error. Test of normality, test of linearity, Testing homoscedasticity and normality of errors, Durbin-Watson test. Test of goodness of fit for the model evaluation and validation. Concept of multi-collinearity

UNIT III

Weighted least squares method: Properties, and examples. Box-Cox family of transformations. Use of dummy variables, Over fitting and under fitting of model, Selection of variables: Forward selection, Backward elimination. Stepwise and Stage wise regressions.

UNIT IV

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.

Course Title: Statistical Computing Course No.: STAT – 573 Credit Hour: 2(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, Use of Software packages like, SAS, SPSS or "R: The R Project for Statistical Computing". Data analysis principles and practice, Summarization and tabulation of data, Exploratory data analysis; Graphical representation of data. Statistical Distributions: Fitting and testing the goodness of fit of discrete and continuous probability distributions;

UNIT II

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

UNIT III

Numerical linear algebra, numerical optimization, graphical techniques, numerical approximations, Time Series Analysis.

UNIT IV

Analysis of mixed models; Estimation of variance components, Analysis of Covariance, Fitting of non-linear model, Discriminant function; Principal component analysis. 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; 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;mixed effect models, estimation of variance components; Categorical data analysis, dissimilarity measures, similarity measures;analysis of discrete data, analysis of binary data; Numerical algorithms; Spatial modeling, cohort studies; Clinical trials, analysis of survival data; Handling missing data.Analysis of time series data - fitting of ARIMA models.