

Course Title with Credit load Ph.D. in Agricultural Statistics

Course Code	Course Title	Credit Hours	Semester
*STAT 601	Advanced Data Analytics	1+2	I
*STAT 602	Simulation Techniques	1+1	I
*STAT 603	Linear Models	2+0	I
*STAT 604	Advanced Statistical Methods	2+1	I
*STAT 611	Bayesian Inference	2+0	II
STAT 691	Seminar I	0+1	I
STAT 692	Seminar II	0+1	II
STAT 699	Research	0+75	II-VI
STAT 605	Modeling Techniques for Forecasting	2+1	I
STAT 606	Stochastic Processes	2+0	I
STAT 607	Survival Analysis	2+0	I
STAT 608	Spatial Statistics	1+1	I
STAT 612	Advanced Design of Experiments	2+1	II
STAT 613	Advanced Sampling Techniques	2+1	II
STAT 614	Advanced Statistical Genetics	2+1	II
STAT 615	Advanced Time Series Analysis	2+0	II
STAT 616	Advanced Bioinformatics	2+0	II
STAT 617	Advanced Econometrics	2+0	II
STAT 618	Recent Advances in the Field of Specialization	1+0	II

*Core Courses

I. **Course Title** : **Advanced Data Analytics**

II. **Course Code** : **STAT 601**

III. **Credit Hours** : **1+2**

IV. **Aim of the course**

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.

V. **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.

VI. **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 Reading

- Everitt B.S. and Dunn G. 1991. *Advanced Multivariate Data Analysis*. 2nd Ed. Arnold.
- Geisser S. 1993. *Predictive Inference: An Introduction*. Chapman & Hall.
- Gentle J.E., Härdle W and Mori Y. 2004. *Handbook of Computational Statistics- Concepts and Methods*. Springer.
- Han J and Kamber M. 2000. *Data Mining: Concepts and Techniques*. Morgan.
- Hastie T, Tibshirani R and Friedman R. 2017. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer. 2nd Ed.
- Kennedy W.J. and Gentle J.E. 1980. *Statistical Computing*. Marcel Dekker.
- Miller R.G. Jr. 1986. *Beyond ANOVA, Basics of Applied Statistics*. John Wiley.
- Rajaraman V. 1993. *Computer Oriented Numerical Methods*. Prentice-Hall.
- Robert C.P. and Casella G. 2004. *Monte Carlo Statistical Methods*. 2nd Ed. Springer.
- Ross S. 2000. *Introduction to Probability Models*. Academic Press.
- Simonoff J.S. 1996. *Smoothing Methods in Statistics*. Springer.
- Thisted R.A. 1988. *Elements of Statistical Computing*. Chapman & Hall.
- Venables W.N. and Ripley B.D. 1999. *Modern Applied Statistics With S-Plus*. 3rd Ed. Springer.
- Free Statistical Softwares: <http://freestatistics.altervista.org/en/stat.php>.
- Design Resources Server: www.drs.icar.gov.in.

I. **Course Title** : Simulation Techniques

II. **Course Code** : STAT 602

III. **Credit Hours** : 1+1

IV. **Aim of the course**

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.

V. **Theory**

Unit I

Uses and purposes of simulation; Classification of models. Generation and testing of random numbers, Review of simulation methods; Implementation of simulation methods - for Discrete and continuous probability distribution, sampling 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

Simulation of generalized linear models and time series models, Simulated data sets to be analyzed using popular computer software packages.

Unit IV

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

VI. **Practical**

- Simulation from various probability models;
- Resampling methods, jackknife and the bootstrap;

- Randomization tests;
- Simulating multivariate distributions, MCMC methods and Gibbs sampler;
- Simulated data sets to be analyzed using popular computer software packages;
- Markov Chain, Monte Carlo, Gibbs' sampling
- Reversible- jump MCMC and multi-grid methods.

VII. Suggested Reading

- Averill M.L. 2017. *Simulation, Modeling and Analysis*. Tata McGraw Hill.
- Balakrishnan N, Melas V.B. and 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 B.L. and Scharge L.E. 1987. *A Guide to Simulation*. Springer. Davison A.C. and Hinkley D.V. 2003. *Bootstrap Methods and their Application*. Cambridge Univ. Press.
- Gamerman D, Lopes H.F. and Lopes H.F. 2006. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. CRC Press.
- Gardner F.M. and Baker J.D. 1997. *Simulation Techniques Set*. John Wiley. Gentle J.E. 2005. *Random Number Generation and Monte Carlo Methods*. Springer.
- Janacek G and Louise S. 1993. *Time Series: Forecasting, Simulation, Applications*. Ellis Horwood Series in Mathematics and its Applications.
- Kleijnen J and Groenendaal W.V. 1992. *Simulation: A Statistical Perspective*. John Wiley.
- Kleijnen J. 1974 (Part I), 1975 (Part II). *Statistical Techniques in Simulation*. Marcel Dekker.
- Law A and Kelton D. 2000. *Simulation Modeling and Analysis*. McGraw Hill.
- Press W.H., Flannery B.P., Tenkolsky S.A. and Vetterling W.T. 1986. *Numerical Recipes*. Cambridge Univ.Press.
- Ripley B.D. 1987. *Stochastic Simulation*. John Wiley. Ross SM. 1997. *Simulation*. John Wiley.

I. **Course Title** : **Linear Models**

II. **Course Code** : **STAT 603**

III. **Credit Hours** : **2+0**

IV. Aim of the course

The students would be exposed to the theory of linear models, estimation of variance components for unbalanced data and advanced techniques for analysis of data in agriculture.

V. Theory

Unit I

General Gauss Markoff set up, Gauss-Markoff's theorem, Aitken's transformation. Theory of linear estimation, test of hypothesis in linear models. Analysis of variance, partitioning of degrees of freedom. Restricted least squares. Special cases of one and two way classifications (including disproportionate cell frequencies and interaction, cross and nested classifications).

Unit II

Analysis of covariance. Variance components models, estimation of variance components from unbalanced data. Unified theory of least-squares, MINQUE, MIVQUE. Mixed

models. LAR, LASSO.

VI. Suggested Reading

- Bapat, R.B. 2012. *Linear Algebra and Linear Models*. Springer-Verlag.
- Graybill, F. A. 1976. *Theory and Application of the Linear Model*. Duxbury, North Scituate.
- Joshi, D.D. 1987. *Linear Estimation and Design of Experiments*. Wiley Eastern.
- Rao, C. R. 2001. *Linear Inference and its Application*. Wiley Eastern.
- Searle, S. R. 1998. *Variance Components*. John Wiley.
- Searle, S.R. 1971. *Linear Models*. John Wiley.
- Seber, G.A. F. 1996. *The Linear Hypothesis: A General Theory*. Griffin, Charles and Co. Ltd.
- Sheffe, H. 1999. *Analysis of Variance*. John Wiley.

I. **Course Title** : **Advanced Statistical Methods**

II. **Course Code** : **STAT 604**

III. **Credit Hours** : **2+1**

IV. Aim of the course

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.

V. Theory

Unit I

Truncated and compound distributions. Fitting of orthogonal polynomials. Pearsonian curves. Categorical data analysis - loglinear models, Association between attributes. Variance stabilizing transformations.

Unit II

Sampling distribution of correlation coefficient, regression coefficient, correlation ratio, intra class correlation coefficient.

Unit III

Non-central t, F^2 and F distributions. Distribution of quadratic forms. Cochran's theorem. Tests for normality. Large sample tests. Tests of significance based on t, F^2 and F distributions. 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.

Unit IV

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

VI. Practical

- Fitting of truncated distribution,
- Fitting of Pearsonian curves,
- Analysis of association between attributes, categorical data.
- Fitting of non-central t, F^2 and F distributions.
- Computation of Tests of significance based on t, F^2 and F distributions.
- Order statistics.

VII. Suggested Reading

- Chatterjee S, Hadi A and Price B. 2013. *Regression Analysis by Examples*. 5th Ed. John Wiley.

- Draper N.R. and Smith H. 1998. *Applied Regression Analysis*. 3rd Ed. John Wiley.
- Rao C.R. 2009. *Linear Statistical Inference and its Applications*. 2nd Ed. John Wiley.
- Searle S.R, Casella G and McCulloch C.E. 1992. *Variance Components*. John Wiley.
- Searle S.R. 1971. *Linear Models*. John Wiley.

I. **Course Title** : **Modeling Techniques for Forecasting**

II. **Course Code** : **STAT 605**

III. **Credit Hours** : **2+1**

IV. **Aim of the course**

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.

V. **Theory**

Unit I

Empirical and mechanistic models. Nonlinear growth models: 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 Parameterization.

Unit III

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

Unit IV

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

Unit V

Forecasting techniques with special reference to agriculture. Forecast based on time series data: exponential smoothing, Box – Jenkins approach and non-linear models. Forecast models using weather parameters, crop-weather relationships and their use in yield forecast. Forecast using plant characters.

Unit VI

Forecast surveys, between-year models (regression model, Markov chain probability model and group method of data handling) and within-year models. Agrometeorological models: climatic water balance model and crop yield assessment. Forewarning of crop pests and diseases. Application of remote sensing techniques in forecasting. Use of ANN in forecasting.

VI. **Practical**

- Fitting of mechanistic non-linear models;
- Application of Schaefer and Fox non-linear models;
- Fitting of compartmental models. Fitting of forecast models using weather parameters.
- Time series analysis: plots, decomposition, stationarity tests, exponential smoothing.
 - Univariate Box – Jenkins ARIMA models and seasonal ARIMA models.
 - Forecast models using plant characters,
- Agrometeorological models for crop forecasting, Markov chain models and ANN models.

VII. Suggested Reading

- Draper, N.R. and Smith, H. 1998. *Applied Regression Analysis*. 3rd Ed. John Wiley.
- Efromovich S. 1999. *Nonparametric Curve Estimation*. Springer.
- Fan, J. and Yao, Q. 2003. *Nonlinear Time Series-Nonparametric and Parametric Methods*. Springer.
- France, J. and Thornley, J.H.M. 1984. *Mathematical Models in Agriculture*. Butterworths.
- Harvey, A.C. 1996. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge Univ. Press.
- Makridakis, S., Wheelwright, S.C. and Hyndman, R.J. 1998. *Forecasting: Methods and Applications*. John Wiley.
- Pankratz, A. 1983. *Forecasting with Univariate Box Jenkins Models: Concepts and Cases*. John Wiley.
- Thornley, J. and France J. 2006. *Mathematical Models in Agriculture: Quantitative Methods for the Plant, Animal and Ecological Sciences* (Cabi) 2nd Ed.

I. **Course Title** : **Stochastic Processes**

II. **Course Code** : **STAT 606**

III. **Credit Hours** : **2+0**

IV. Aim of the course

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.

V. 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 queueing model M/M/1: steady state behaviours. Birth and death process in queueing theory- Multi channel models. Network 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.

VI. Suggested Reading

- Adke SR and 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 and Prakasa Rao BLS. 1980. *Statistical Inference for Stochastic Processes*. Academic Press.
- Bharucha-Reid AT. 2012. *Elements of the Theory of Markov Processes and their Applications*. McGraw Hill.
- Bhat BR. 2000. *Stochastic Models; Analysis and Applications*. New Age.
- Draper NR and Smith H. 1981. *Applied Regression Analysis*. Wiley Eastern. France J & Thornley JHM. 1984. *Mathematical Models in Agriculture*. Butterworths.
- Lawler GF. 2006. *Introduction to Stochastic Processes*. Chapman & Hall. 2nd Ed.
- Medhi J. 2001. *Stochastic Processes*. 2nd Ed. Wiley Eastern.
- Prakasa Rao BLS and 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 and Wild CJ. 1989. *Non-linear Regression*. John Wiley.

I. **Course Title** : **Survival Analysis**

II. **Course Code** : **STAT 607**

III. **Credit Hours** : **2+0**

IV. Aim of the course

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.

V. 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.l.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.

Unit V

Interval Censoring Competing Risk Theory: Indices for measurement of probability of death under competing risks and their inter-relations. Concept of COX regression Stochastic Epidemic Models: Simple epidemic models, general epidemic model definition and concept (without derivation). Duration of an epidemic.

VI. Suggested Reading

- Anderson B. 1990. *Methodological Errors in Medical Research*. Blackwell.
- Armitage P and Berry G. 1987. *Statistical Methods in Medical Research*. Blackwell.
- Biswas, S. 2007. *Applied Stochastic Processes: A Biostatistical and Population Oriented Approach*, 2nd Ed., New Central Book Agency.
- Collett D. 2014. *Modeling Survival Data in Medical Research*. Chapman & Hall. 3rd Ed.
- Cox D.R. and Oakes D. 1984. *Analysis of Survival Data*. Chapman & Hall.
- Elandt-Johnson R.C. and Johnson N.L. 1980. *Survival Models and Data Analysis*. John Wiley.
- Everitt B.S. and Dunn G. 1998. *Statistical Analysis of Medical Data*. Arnold.
- Hosmer D.W. Jr. and Lemeshow S. 1999. *Applied Survival Analysis: Regression Modeling or Time to Event*. John Wiley.
- Indrayan, A. 2008. *Medical Biostatistics*, 2nd Ed. Chapman and Hall/CRC.
- Lee E.T. 1980. *Statistical Methods for Survival Data Analysis*. Lifetime Learning Publ.
- Kalbfleisch J.D. and Prentice. R.L. 2002. *The Statistical Analysis of Failure Time Data*. John Wiley.

- Klein J.P. and Moeschberger M.L. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.
- Kleinbaum D.G. and Klein M. 2002. *Logistic Regression*. Springer.
- Kleinbaum D.G. and Klein M. 2005. *Survival Analysis*. Springer.

I. **Course Title** : **Spatial Statistics**

II. **Course Code** : **STAT 608**

III. **Credit Hours** : **1+1**

IV. **Aim of the course**

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

V. **Theory**

Unit I

Spatial Analysis and types of spatial data; Visualizing Spatial Data – Exploratory data Analysis.

Unit II

Spatial Relationship- Random forest, spatially autocorrelated data, weight matrix, measures of spatial Auto-correlation – Moran's I & Geary's C; Measuring of autocorrelation of spatially continuous data.

Unit III

Spatial Sampling – Methods and procedures, Statistical Analysis of Spatial Point Process – homogenous Poisson Process, Spatial interpolation – non-statistical methods; Variogram modelling; Spatial Prediction – Simple Kriging, Co-kriging;

Unit IV

Modelling Areal data – Autoregressive and spatial regression models and model diagnostics. Examples of Spatial Data analysis in Agriculture– Disease Mapping; Incorporating spatial effects in Agricultural Field experiments

VI. **Practical**

- Spatial Data – Import, export;
- Spatial Classes in R;
- Visualizing Spatial Data;
- Spatial Auto-correlation;
- Spatial Sampling, Spatial Interpolation, Spatial Autoregressive Models, Spatial Regression Model

VII. **Suggested Reading**

- Cressie, N.A.C. 1993. *Statistics for Spatial Data*. Revised Edition. JohnWiley
- Richard E.P. 2018. *Spatial Data Analysis in Ecology and Agriculture Using R*, 2nd Ed.
- Roger S. Bivand, E Pebesma J. and Rubio B.G. 2008. *Applied Spatial Data Analysis using R*. Springer-Verlog.

I. **Course Title** : **Bayesian Inference**

II. **Course Code** : **STAT 611**

III. **Credit Hours** : **2+0**

IV. **Aim of the course**

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.

V. Theory**Unit I**

Introduction and history and criticism of Bayesian Approach; Subjective interpretation of Probability, Review of Bayes Theorem, Sufficiency, Likelihood Principle.

Unit II

Subjective Prior distribution of a parameter; Posterior Distribution of parameters using Bayes Theorem

Unit III

Informative and non-informative priors for Location and scale; Conjugate families – Discrete and Continuous and interpretation of Hyper-parameters of conjugates.

Unit IV

Non-informative, improper and invariant priors for location and scale and in general settings.

Unit V

Bayesian Point Estimation – squared error loss, absolute error loss etc. Bayesian Interval Estimation – Credible Interval, interpretation and comparison with frequentist confidence Intervals

Unit VI

Bayesian Hypothesis Testing - Specification of the appropriate form of the prior distribution for a Bayesian testing of hypothesis problem. Prior odds, Posterior odds, Bayes factor for various types of testing hypothesis problems

Unit VII

Bayesian Prediction; Numerical and Monte-Carlo Integrations

Unit VIII

Applications of Bayesian Inference - Bayesian Data Analysis

VI. Suggested Reading

- Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag.
- Box, G.P. and Tiao, G.C. 1992. *Bayesian Inference in Statistical Analysis*, Addison – Wesley
- Pilon C.D. 2015. *Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference* (Addison-Wesley Data and Analytics)

I. Course Title : Advanced Design of Experiments

II. Course Code : STAT 612

III. Credit Hours : 2+1

IV. Aim of the course

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.

V. 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. Structurally Incomplete block designs

Unit II

General properties and analysis of two-way heterogeneity designs, Youden type designs,

generalized Youden designs, Pseudo Youden 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, Resolution plans, Regular and irregular fractions.

Unit IV

Response surface designs - Symmetrical and asymmetrical factorials, Response optimization and slope estimation, Blocking, Canonical analysis and ridge analysis, CCD, Box-Jenkins, 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.

VI. Practical

Analysis of block designs, Analysis of Latin square type designs, group divisible designs, triangular designs, lattice designs. Analysis of fractional replications of factorial experiments, analysis of asymmetrical factorials and block designs with factorial structure. Analysis of second order response surface designs.

VII. Suggested Reading

- Chakraborti M.C. 1962. *Mathematics of Design and Analysis of Experiments*. Asia Publ.House.
- Dean A.M. and Voss D. 1999. *Design and Analysis of Experiments*.
- pringer. Dey A and 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 A.S., Sloane N.J.A. and Stufken J. 1999. *Orthogonal Arrays: Theory and Applications*. Springer.
- John J.A. and Quenouille M.H. 1977. *Experiments: Design and Analysis*. Charles and Griffin.
- Khuri A.I. and Cornell J.A. 1996. *Response Surface Designs and Analysis*. 2nd Ed. Marcel Dekker.
- Montgomery D.C. 2005. *Design and Analysis of Experiments*. John Wiley.
- Ogawa J. 1974. *Statistical Theory of the Analysis of Experimental Designs*. Marcel Dekker.
- Parsad R, Gupta V.K., Batra P.K., Satpati S.K. and Biswas P. 2007. *Monograph on a-designs*. IASRI, New Delhi.
- Raghavarao D. 1971. *Construction and Combinatorial Problems in Design of Experiments*. John Wiley.
- Shah K.R. and Sinha B.K. 1989. *Theory of Optimal Designs. Lecture notes in Statistics*. Vol. 54. Springer.
- Sharma M.K. 2012. *Design and Analysis of Experiments*. Kindle Ed. 1st Ed.
- Street A.P. and Street D.J. 1987. *Combinatorics of Experimental Designs*. Oxford Science Publ.
- Design Resources Server: www.drs.icar.gov.in.

I. Course Title : **Advanced Sampling Techniques**

II. Course Code : **STAT 613**

III. Credit Hours : 2+1**IV. Aim of the course**

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.

V. Theory**Unit I**

Optimum Stratification, two-way stratification, collapsed strata, controlled selection, Use of combinatorics in controlled selection, Systematic sampling in two dimensions. Sampling with varying probabilities without replacement, Horvitz – Thompson estimator

Unit II

Variance estimation in complex surveys. Taylor's series linearization, balanced repeated replication, Jackknife and bootstrap methods. Ordered and unordered estimators, Sampling strategies, Midzuno-Sen, Rao-Hartley-Cochran, PPS Sampling: procedures such as Brewer, Durbin and Sampford,

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

Post-stratified estimator, imperfect frames, multiple frames, randomized response techniques. 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. Longitudinal survey.

VI. Practical

- Sampling with varying probability,
- Ordered and un-ordered estimators,
- Sampling strategies due to Horvitz-Thompson, Midzuno-Sen, Rao-Hartley-Cochran and PPS sampling: procedures such as Brewer, Durbin and Sampford, etc.
- Imperfect frames, Randomized response technique.
- Small area estimation.

V. Suggested Reading

- Berger J.O. 1993. *Statistical Decision Theory and Bayesian Analysis*. Springer.
- Bolfarine H and Zacks S. 1992. *Prediction Theory for Finite Population Sampling*. Springer.
- Cassel C.M., Sarndal C.E and Wretman J.H. 1977. *Foundations of Inference in Survey Sampling*. John Wiley.
- Des Raj and Chandhok P. 1998. *Sample Survey Theory*. Narosa Publ.
- House. Ghosh M and Meeden G. 1997. *Bayesian Method for Finite Population Sampling. Monograph on Statistics and Applied Probability*. Chapman and Hall.
- Mukhopadhyay P. 1998. *Theory and Methods of Survey Sampling*. Prentice Hall of India.
- Rao J.N.K. 2003. *Small Area Estimation*. John Wiley.
- Sarndal C.E., Swensson B and Wretman J.H. 1992. *Model Assisted Survey Sampling*. Springer.

- I. Course Title** : **Advanced Statistical Genetics**
II. Course Code : **STAT 614**
III. Credit Hours : **2+1**

IV. Aim of the course

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.

V. 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

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.

VI. Practical

- Hardy-Weinberg law,
- Estimation of genetic load and random genetic drift.
- Effect of finite population size.
- Estimation of path coefficients.
- Detection and estimation of multiple allelism in continuous variation, sex-linked genes, maternal effects.
- Analysis of $G \times E$ interaction, measurement of stability and adaptability.
- Analysis of data of diallel and partial diallel crosses.

VII. Suggested Reading

- Crow J.F. and Kimura M. 1970. *An Introduction of Population Genetics Theory*. Harper & Row.
- Ewens W.J. 1979. *Mathematical Population Genetics*. Springer.

- Falconer D.S. 1985. *Introduction to Quantitative Genetics*. ELBL.
- Fisher R.A. 1949. *The Theory of Inbreeding*. Oliver & Boyd.
- Fisher R.A. 1958. *The Genetical Theory of Natural Selection*. Dover Publ.
- Haldane J.B.S. 1932. *The Causes of Evolution*. Harper & Bros.
- Kempthorne O. 1957. *An Introduction to Genetic Statistics*. The Iowa State Univ. Press.
- Lerner I.M. 1950. *Population Genetics and Animal Improvement*. Cambridge Univ. Press.
- Lerner I.M. 1958. *The Genetic Theory of Selection*. John Wiley.
- Li C.C. 1982. *Population Genetics*. The University of Chicago Press.
- Mather K and Jinks J.L. 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.
- Nielsen R, Montgomery S. 2013. *An Introduction to Population Genetics: Theory and Applications* 1st Ed.

I. Course Title : Advanced Time Series Analysis

II. Course Code : STAT 615

III. Credit Hours : 2+0

IV. Aim of the course

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.

V. 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, Vector Error Correction Model (VECM).

Unit II

Volatility: The class of ARCH and GARCH models; Extensions of GARCH models: TGARCH, IGARCH, PGARCH, EGARCH, GJR-GARCH, ARCH and GARCH model with-t distributed error; ARCD (Auto-Regressive Conditional Density), Multivariate GARCH model: estimation, analysis and forecasting, stochastic volatility.

Unit III

Structural time-series modelling: State space models, Kalman filter, Local level model, Local linear trend model, Seasonal models, Cyclical models. Threshold and Functional coefficient autoregressive models, Structural Break in time series.

Unit IV

Fuzzy time series models, Artificial Neural Network (ANN) methodology, Support vector machines, Wavelets for time series analysis, combinations of time series models.

VI. Suggested Reading

- Box G.E.P., Jenkins G.M. and Reinsel G.C. 2015. *Time Series Analysis: Forecasting and Control*. 5th Ed. John Wiley.

- Brockwell P.J. and Davis R.A. 1991. *Time Series: Theory and Methods*. 2nd Ed. Springer.
- Chatfield C. 2004. *The Analysis of Time Series: An Introduction*. 6th Ed. Chapman & Hall/ CRC.
- Johnston J. 1984. *Econometric Methods*. McGraw Hill.
- Singh, P. 2016. *Applications of Soft Computing in Time Series Forecasting: Simulation and Modeling Techniques*. Springer International Publishing AG
- Tong H. 1995. *Nonlinear Time Series: A Dynamical System Approach*. Oxford Univ. Press.
- Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*. Springer- Verlag, New York.
- Percival, D.B. and Walden, A.T. 2000. *Wavelet Methods for Time-Series Analysis*. Cambridge University Press, U.K.

I. Course Title : Advanced Bioinformatics

II. Course Code : STAT 616

III. Credit Hours : 2+1

IV. Aim of the course

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.

V. Theory

Unit I

EM algorithm and other statistical methods to discover common motifs in bio sequences. Concepts in phylogeny. Gene prediction based on codons, Decision trees, Clustering Techniques, 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/ RNAseq, 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, simulated annealing, MC with absorbing States). Bayesian techniques and use of Gibbs Sampling. Advanced topics in design and analysis of DNA microarray experiments.

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.

VI. Practical

- Genomic databases and analysis of high-throughput data sets, BLAST and related sequence comparison methods.
- Statistical methods to discover common motifs in bio sequences.
- Multiple alignment and database search using motif models, clustalw, classificatory analysis, neural networks, genetic algorithms, pattern recognition,
- Hidden Markov models.
- Computational analysis of protein sequence.
- Expression profiling by microarray/ gene chip, proteomics.
- Modelling and prediction of structure of proteins.
- Bayesian techniques and use of Gibbs Sampling.
- Analysis of DNA microarray experiments.
- Analysis of one DNA sequence, multiple DNA or protein sequences.
- Computationally intensive methods, multiple hypothesis testing,
- Phylogenetic tree estimation, Analysis of SNPs and haplotypes.

VII. Suggested Reading

- Baldi P and Brunak S. 2001. *Bioinformatics: The Machine Learning Approach*. MIT Press.
- Baxevanis AD and Francis BF. (Eds.). 2004. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. John Wiley.
- Duda RO, Hart PE and Stork DG. 1999. *Pattern Classification*. John Wiley.
- Ewens WJ and Grant GR. 2001. *Statistical Methods in Bioinformatics*. Springer.
- Jones NC and Pevzner PA. 2004. *Introduction to Bioinformatics Algorithms*. The MIT Press.
- Koskinen T. 2001. *Hidden Markov Models for Bioinformatics*. Kluwer.
- Krane DE and 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 and Seefeld K. 2005. *R for Bioinformatics*. O'Reilly and Associates.
- Percus JK. 2001. *Mathematics of Genome Analysis*. Cambridge Univ. Press.
- Sorensen D and Gianola D. 2002. *Likelihood, Bayesian and MCMC Methods in Genetics*. Springer.
- Tisdall J.D. 2001. *Mastering Perl for Bioinformatics*. O'Reilly & Associates.

- Wang J.T.L., Zaki M.J., Toivonen H.T.T. and Shasha D. 2004. *Data Mining in Bioinformatics*. Springer.
- Wu C.H. and McLarty J.W. 2000. *Neural Networks and Genome Informatics*. Elsevier.
- Wunschiers R. 2004. *Computational Biology Unix/Linux, Data Processing and Programming*. Springer.
- Yang M.C.C. 2000. *Introduction to Statistical Methods in Modern Genetics*. Taylor & Francis.

I. Course Title : Advanced Econometrics

II. Course Code : STAT 617

III. Credit Hours : 1+1

IV. Aim of the course

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

V. 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 timeseries. Co integration analysis.

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].

VI. Practical

Fitting of equation with serially correlated errors, ordinary least-squares and generalized least squares methods of estimation. Non-stationary multivariate time series analysis. Fitting of 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 Fitting of ARCH model.

VII. Suggested Reading

- Banerjee A, Dolado J, Galbraith J and Hendry D.F. 1993. *Co-integration, Error Correction, and the Econometric Analysis of Nonstationary Data*. Oxford Univ. Press.
- Bauwens L, Lubrano M. and Richard J.F. 1999. *Bayesian Inference in Dynamics of Econometric Models*. Oxford Univ. Press.

- Carlin B.P. and Louis T.A. 2008. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall.
- Gilks W.R., Richardson S and Spiegelhalter D. 1996. *MCMC in Practice*. Chapman & Hall.
- Greenberg E. 2012. *Introduction to Bayesian Econometrics*. Cambridge Univ. Press.
- Hamilton J.D. 1994. *Time Series Analysis*. Princeton Univ. Press.
- Judge G.G., Griffith W.E., Hill R.C., Lee C.H. and Lutkepohl H. 1985. *The Theory and Practice of Econometrics*. 2nd Ed. JohnWiley.
- Koop G, Poirier D and 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 R.S. and Rubinfeld D.L. 1981. *Econometric Models and Economic Forecasts*. McGraw Hill.

I. Course Title : Recent Advances in the Field of Specialization

II. Course Code : STAT 618

III. Credit Hours : 1+0

IV. Aim of the course

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

V. 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.

VI. Suggested Reading

Recent journals related to the research works

67th BoS (Faculty of Agriculture)