

DEPARTMENT OF AGRICULTURAL STATISTICS
PALLI SIKSHA BHAVANA
VISVA-BHARATI, SRINIKETAN

Courses for B.Sc. (Hons.) Agriculture
NEP

Course Code	Course Title	Credit Hours
AST-311	Basic and Applied Agril Statistics	2+1
EC-4145	Statistical Methods	3+1
EC-4146	Design of Experiments	3+1
EC-4147	Sampling Techniques	3+1
EC-4148	Regression Analysis	3+1
EC-4149	Data Analysis	3+1

Objective

This course is meant to expose the students to various concepts of descriptive statistical methods and statistical inferential procedures which would help them in understanding the concepts involved in data collection, its presentation, analysis and interpretation. This course will also provide them with the understanding of the early development of statistics in ancient India.

Theory

Statistics as a method of learning from experience and decision making under uncertainty were practiced from the beginning of human civilization. Here the ancient foundations of statistical practices in Indian civilization with a focus on agriculture, population and economic census in villages and towns will be explored. Data collection systems found in ancient texts like Vyasa's Mahabharata, Kautilya's (321-296 BC) Arthashastra, Abul Fazal's Ain-i-Akbari (1596-1597 AD) and others which documented methods of estimation, land classification and resource planning will be discussed. A brief discussion will be made on the first use of arithmetic mean as the best representative value for a set of observations in statistical sense which can be found in ancient India's texts (Brahmagupta, 628 AD). How population data and vital statistics were maintained in ancient India as seen in Manusmriti and the administrative practices of the Chola dynasty will also be explained.

Introduction to Statistics and its Applications in Agriculture. Types of Data. Scales of measurements of Data. Summarization of Data. Classification of Data. Frequency Distribution. Methods of Classification. Definition of Grouped and Ungrouped Data. Definition of Class Interval (formula for determining the no. of class interval), Width of CI, Class Limits (Boundaries), Mid Points. Types of Frequency Distribution. Diagrammatic Presentation of Data. Bar Diagrams – Simple, Multiple, Sub-divided and Percentage Bar Diagrams. Pie-diagram. Graphical Presentation of Data – Histogram, Frequency Polygon and Ogives, Stem and leaf plot. Measures of Central Tendency. Requisites for an Ideal Measure of Central Tendency. Different Types of Measure. Arithmetic Mean– Definition, Properties, Merits, Demerits and Uses. A.M. (examples) for Grouped and Ungrouped Data. Outlier observations. Trimmed Mean. Step-deviation Method. Weighted Mean. Definition of Geometric Mean and Harmonic Mean. Relationship between A.M., G.M. and H.M. Median Definition, Merits, Demerits and Uses. Graphical Location of Median. Mode-Definition, Merits, Demerits and Uses. Graphical Location of Mode. Relationship between Mean, Median and Mode.

Measures of Dispersion. Characteristics for an Ideal Measure of Dispersion. Different Types of Measures of Dispersions. Definition of Range, Interquartile Range, Quartile Deviation and Mean Deviation. Standard Deviation- Definition, Properties. S.D. and Variance for Grouped and Ungrouped Data. Variance of Combined Series. Co-efficient of Dispersions. Co-efficient of Variation. Box plot and five-number summary statistics.

Measures of Skewness and Kurtosis. Definition of Symmetrical Distribution. Definition of Skewness, Measures of Skewness. Definition of Kurtosis. Measure of Kurtosis. Relationship between Mean, Median and Mode for Symmetrical and Skewed Distribution.

Correlation and Regression. Definition of Correlation. Scatter Diagram. Karl Pearson's Coefficient of Correlation. Types of Correlation Coefficient. Properties of Correlation Coefficient. Definition of Linear Regression. Regression Equations. Regression Coefficients. Properties of Regression Coefficients.

Introduction to Probability. Basic Terminology. Classical Probability-Definition and Limitations. Empirical Probability- Definition and Limitations. Axiomatic Probability. Addition and Multiplication Theorem (without proof). Conditional Probability. Independent Events. Simple Problems based on Probability.

Definition of Random Variable. Discrete and Continuous Random Variable. Introduction of Binomial and Poisson distributions with basic properties. Normal Distribution- Definition, Prob. Distribution, Mean and Variance. Assumptions of Normal Distribution. Normal Probability Curve.

Sampling Theory. Introduction. Definition of Population, Sample, Parameter and Statistic. Sampling Vs Complete Enumeration. Sampling Methods. Simple Random Sampling with Replacement and without Replacement. Use of Random Number Table.

Tests of Significance. Null and Alternative Hypothesis. Type I and Type II Errors. Critical Region and Level of Significance. One Tailed and Two Tailed Tests. Test Statistic. One Sample (Z and t), Two Sample independent and dependent (Z and t) test with Examples. F-test for Variance.

ANOVA and Experimental Designs. Assumptions of ANOVA. Assignable and Non assignable Factors. Analysis of One-way Classified Data. Basic Examples of Experimental Designs. Terminologies. Completely Randomized Design (CRD).

Practical

Diagrammatic and Graphical representation of data. Calculation of A.M., Median and Mode (Ungrouped and Grouped data). Calculation of S.D. and C.V. (Ungrouped and Grouped data). Correlation and Regression analysis. Application of Z and t-test (one sample, two sample independent and dependent). Analysis of variance one-way classification. CRD. Selection of random sample using simple random sampling.

Outcome

It is expected that the students will be equipped with basic statistical tools used for analyzing data sets and will be able to draw valid conclusion supported by statistical philosophy.

Suggested readings

- 1) Agriculture and Applied Statistics-I by P.K. Sahu, Kalyani Publishers.
- 2) Agriculture and Applied Statistics-II by P. K. Sahu and A. K. Das, Kalyani Publishers.
- 3) Agricultural Statistics by S.P. Singh and R.P.S. Verma, Rama Publishing House.
- 4) Basic Statistics by B. L. Agarwal, New Age International Publishers.
- 5) Fundamentals of Applied Statistics by S.C. Gupta and V. K. Kapoor, Sultan Chand and Sons.

- 6) Fundamentals of Statistics by D. N. Elhance, Kitab Mahal Publishers.
- 7) Probability and Statistical Inference by D. Bhattacharya and S. Roychowdhury, 3rd ed. U.N. Dhur and Sons
- 8) Statistics: Theory and Practice by D. Bhattacharya and S. Roychowdhury, 4th ed. U.N. Dhur and Sons

Objective

This course is meant for students who have some knowledge on basic statistical tools and techniques. The students would be exposed to those concepts of statistics that deal with modelling observed data using different probability models, how to draw a good sample from a population in order to make some valid conclusion about the population a population parameter etc.

Theory

Unit I: Descriptive statistics, Exploratory data analysis, Theory of probability, Random variables, Probability mass function. Probability density function. Mathematical Expectation, Moment generating function. Cumulant generating function.

Unit II: Probability Distribution: Binomial, Poisson, Negative Binomial, Geometric, Uniform, Normal, Exponential, Gamma and Beta.

Unit III: Sampling Distributions: Chi-Square, t and F . Introduction to theory of estimation and confidence-intervals, Parametric tests based on normal, chi-square, t and F distributions. Introduction of Non-parametric tests.

Practical

- ❖ Exploratory data analysis,
- ❖ Fitting of distributions $\sim$ Binomial, Poisson, Normal.
- ❖ Large sample tests, testing of hypothesis based on exact sampling distributions $\sim$ chi square, t and F .
- ❖ Confidence interval estimation and Correlation and regression analysis, fitting of Linear and Quadratic Model.

Outcome

Students will be well equipped to handle field level data for analysis and modelling purposes. They will learn how to draw a good sample from a population in order to draw valid inference about the population a population parameter and how to build multiple linear regression models and study correlation among them under a multivariable set-up.

Suggested Readings

1. Bhattacharya, D. and Roychowdhury, S. *Probability and Statistical Inference: Theory and Practice*. 3rd Ed. U.N. DHUR and Sons Pvt. Ltd.
2. Bhattacharya, D. and Roychowdhury, S. *Statistics: Theory and Practice*. 3rd Ed. U.N. DHUR and Sons Pvt. Ltd.
3. Goon, A.M., Gupta, M.K. and Dasgupta, B. *Fundamentals of Statistics*. Vol. I & II. The World Press Pvt. Ltd.
4. Gupta, S.C. and Kapoor, V.K. *Fundamentals of Mathematical Statistics*. Sultan Chand and Sons.
5. Hogg, R.V. McKean, J.W. and Craig, A.T. *Introduction to Mathematical Statistics*. 8th Ed. Pearson.
6. Rohatgi, V.K. and Ehsan, S. *An Introduction to Probability Theory and Mathematical Statistics*. Wiley Eastern Private Ltd.

Objective

This course is designed to give a comprehensive knowledge on how to design a study or experiment so that the results of the experiments are free from errors or biases, and then how to draw a valid conclusion using the results so obtained. In this context, laying out of different agricultural field experiments will also be covered. Designing an experiment is an integrated component of research in almost all sciences.

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. Concept of confounding.

Unit IV: Split plot and strip plot designs, missing plot techniques in randomized block and Latin square designs; transformations, Introduction of Balanced Incomplete Block Design, 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.

Outcome

The students would be exposed to various concepts of designing an experiment so as to enable them understand the science involved in planning, designing their research experiments and how to make analysis of different experimental data.

Suggested Readings

1. Cochran, W.G. and Cox, D.R. 1957. *Experimental Designs*. 2nd Ed. John Wiley & Sons.
2. Das, M.N. and Giri, N.C. *Design and Analysis of Experiments*. Wiley Eastern Ltd.
3. Dey, A. 2010. *Incomplete Block Designs*. World Scientific Publishing
4. Fisher, R.A. 1953. *Design and Analysis of Experiments*. Oliver and Boyd.
5. Gomez, K.A. and Gomez, A.A. *Statistical Procedures for Agricultural Research*. 2nd Ed. John Wiley & Sons.
6. Gupta, S.C. and Kapoor, V.K. *Fundamentals of Applied Statistics*. Sultan Chand and Sons.
7. www.drs.icar.gov.in.

Objective

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, Sampling and non-sampling errors.

Unit II: Probability and non-probability sampling, Simple random sampling with and without replacement, determination of sample size, inverse sampling.

Unit III: Stratified random sampling, Cluster sampling, Systematic sampling, Multi-stage sampling, Introduction to PPS sampling, Double Sampling.

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

Outcome

It is expected that 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 from data sets which are non-normal and are not of quantitative nature.

Suggested Readings

1. Bhattacharya, D. and Roychowdhury, S. *Probability and Statistical Inference: Theory and Practice*. 3rd Ed. U.N. DHUR and Sons Pvt. Ltd.
2. Cochran WG. 1977. *Sampling Techniques*. John Wiley.
3. Murthy MN. 1977. *Sampling Theory and Methods*. 2nd Ed. Statistical Publ. Soc., Calcutta.
4. Singh D, Singh P and Kumar P. 1982. *Handbook on Sampling Methods*. IASRI Publ.
5. Sukhatme PV, Sukhatme BV, Sukhatme S and Asok C. 1984. *Sampling Theory of Surveys with Applications*. Iowa State University Press and Indian Society of Agricultural Statistics, New Delhi.

Objective

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: Simple and multiple correlation, partial correlation, rank correlation, Simple and multiple linear regressions: Least squares fit, Properties and examples. Polynomial regression: Use of orthogonal polynomials. Analysis of multiple regression models, estimation and testing of regression parameters, sub-hypothesis testing, restricted estimation.

Unit II: Selection of variables, Stepwise and Stagewise regressions. Adequacy and validation of models. Use of dummy variables, regression with ordinal data. Introduction to non-parametric regression. Logistic regression. Parameter estimation in non-linear models.

Unit III: Regression diagnostics - non-normal errors, non-constant error variances, non independent observations, influential observations (outliers), non-linearity of the model, auto-correlation, Durbin-Watson statistic, Heteroscedasticity, multicollinearity in the data. Remedial measures - regression under non-normal errors, transformation of data, generalized least-squares, robust regression, ridge regression, Model over-fitting, model under-fitting.

Practical

- ❖ Correlation and regression analysis,
- ❖ Multiple regression fitting with three and four independent variables,
- ❖ Estimation of residuals, their applications in outlier detection, distribution of residuals,
- ❖ Test of homoscedasticity, normality and autocorrelation,
- ❖ Box-Cox transformation; Restricted estimation of parameters in the model, hypothesis testing,
- ❖ Step wise regression analysis

Outcome

It is expected that the students will develop skills to diagnostic measures such as autocorrelation, multi collinearity and heteroscedasticity. inference from data sets which are non-normal and are not of quantitative nature.

Suggested Readings

1. Belsley, D.A., Kuh, E. and Welsch, R.E. 2004. *Regression Diagnostics-Identifying Influential Data and Sources of Collinearity*. John Wiley.
2. Chatterjee, S., Hadi, A. and Price, B. 1999. *Regression Analysis by Examples*. John Wiley.
3. Draper, N. R. and Smith, H. 1998. *Applied Regression Analysis*. John Wiley.
4. McCullagh, P. and Nelder, J.A. 1999. *Generalized Linear Models*. Chapman and Hall.

5. Montgomery, D.C., Peck, E. and Vining, G. 2003. *Introduction to linear regression analysis*. John Wiley and Sons Inc., New York.
6. Rao, C.R. 1973. *Linear Statistical Inference and its applications*. John Wiley.

Objective

This course is meant for the students to let them know about the usage of various statistical packages for analyzing data. It would provide the students a hands on for analysing their research data along with some basic ideas on the statistical tools used in the research analysis with their applications.

Theory

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

Unit II: Test for normality; Testing of hypothesis: Parametric and nonparametric tests.

Unit III: Data preparation for ANOVA, Analysis of basic designs, Split and strip plot design, Factorial Experiments, multiple comparisons.

Unit IV: Correlation and regression analysis, Probit and Logit Models, Cluster analysis, Principal component analysis, 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;
- ❖ Cluster analysis and Principal component analysis

Outcome

Some expertise on analysing data using software packages will be developed. Special skills will be developed for reading of the output tables and picking up of the right numbers from the output tables for inferential purposes. At the end, Skills will be developed in writing computer programmes when it is not available in the menu driven software packages.

Suggested Readings

1. Anderson C.W. and Loynes R.M. 1987. *The Teaching of Practical Statistics*. John Wiley.
2. Bhattacharya, D. and Roychowdhury, S. *Probability and Statistical Inference: Theory and Practice*. 3rd Ed. U.N. DHUR and Sons Pvt. Ltd.
3. Chambers J.M., Cleveland W.S., Kleiner B and Tukey P.A. 1983. *Graphical Methods for Data Analysis*. Wadsworth, Belmont, California.

4. Chapman & Hall. Chatfield C. 1995.*Problem Solving: A Statistician's Guide*. Chapman & Hall.
5. Gupta, S.C. and Kapoor, V.K. *Fundamentals of Mathematical Statistics*. Sultan Chand and Sons.
6. Snell E.J. and Simpson HR. 1991. *Applied Statistics: A Handbook of GENSTAT Analyses*. Chapman and Hall.
7. Velleman PF and Hoaglin DC. 1981. *Application, Basics and Computing of Exploratory Data Analysis*. Duxbury Press.
8. Wetherill GB.1986. *Regression Analysis with Applications*. Chapman & Hall.
9. <http://freestatistics.altervista.org/en/learning.php>.
10. <http://freestatistics.altervista.org/en/stat.php>.
11. http://www.cas.lancs.ac.uk/glossary_v1.1/main.html.
12. <http://www.stat.sc.edu/~grego/courses/stat706/>.
13. www.drs.icar.gov.in.