University of Lucknow
Master of BIOSTATISTICS Programme
Regulations 2020

1. Applicability

These regulations shall apply to the Master in Biostatistics Programme from the session 2020-21.

2. Minimum Eligibility for admission

A three/four-year Bachelor's degree or equivalent degree awarded by a University or Institute established as per law and recognised as equivalent by this University with minimum 45% marks or equivalent grade with Mathematics as one of the subject at Intermediate Higher Secondary level shall constitute the minimum requirement for admission to the Master in Biostatistics programme. (any other additional requirement may also be specified)

3. Programme Objectives

Preamble: M. A. / M. Sc. Biostatistics programme is of minimum 96 credits spread over four semesters. The programme emphasizes both theory and applications of Biostatistics in biological/medical sciences and is structured to provide knowledge and skills in depth necessary for the employability of students in industry, other organizations, as well as in academics. The program has some unique features such as independent projects, number of elective courses, extensive computer training of statistical computations including standard software packages such as STATA, R, and SPSS. The department has the academic autonomy and it has been utilized to add the new and need based elective courses. The independent project work is one of the important components of this program. In all semesters I, II, III & IV some courses are compulsory and others are elective. The syllabus has been framed to have a good balance of theory, methods and applications of Biostatistics.

It is possible for the students to study basic courses from other disciplines such as health sciences, life sciences, physical sciences, social sciences, bioinformatics ,biomathematics, analytical softwares etc in place of electives.

4. Programme Outcomes

The programme focuses on the integration of epidemiology, biostatistics and public health. It involves diverse topics like Bio statistical Inference, Survival Analysis, Clinical Trials, Statistical Genetics, etc. where students learn fundamental and advanced concepts in demography, survey sampling, epidemiological aspects in infectious and chronic diseases, public health issues in nutrition and reproductive health. Advanced mathematical & computational skills are employed through Software to analyse disease pathogenesis, transmission and control, whereby, original research projects can also be undertaken in these areas.

- To cover the basic biostatistical concept and theory needed by practicing as a biostatisticians.
- To furnish students to teach themselves new skills in what is a fast developing

area under discussion.

- To enable students to turn a problem described in terms into something that can be tackled by a biostatistical analysis.
- Teach, and provide the opportunities to learn, a core of advanced biostatistical methods, together with a range of more specialized options in biostatistics.

5. Specific Programme Outcomes

At the ***Department of Statistics, University of Lucknow,*** we provide theoretical foundations that will motivate and prepare the students to take up theoretical and applied research in the field of Biostatistics. Students have a variety of rewarding job options available to them in avenues like Pharmaceutics, Medical Statistics, and Consultancy etc.

- The student's can handle and analyze biological/medical/clinical databases with computer skills.
- The students may able to describe complex biostatistical ideas to non-biostatisticians and to present the results of their analyses in written, oral forms and can make practical suggestions for fulfilling the objectives of the study such as assessment, monitoring, evaluation, improvement etc

- The students may get wide range of opportunities of Biostatistics in Health industry sector as well as in government/private sector.

- The students will get wide range of biostatistical skills, including problem-solving, project work and presentation; they may enable to take prominent roles in a wide spectrum of employment and research.

Department of Statistics

University of Lucknow

Syllabus for M.A./M.Sc. Biostatistics Programme

(Proposed to be implemented from July 2020)

Course No.	Name of the Course	Credits	Remark
	Semester I		
BSTATCC-101	Descriptive Statistics	04	Core Course
BSTATCC -102	Elements Of Demography	04	Core Course
BSTATCC -103	Probability Theory And Distributions	04	Core Course
BSTATCC -104	Research Methodology	04	Core Course
BSTATCC -105	Practical	04	Core Course
BSTATVC-101	Sustainable Development Goals	04	Value added course (Credited)
	Semester Total	24	
	Semester II		
BSTATCC -201	Sampling Theory and Design of Experiments	04	Core Course
BSTATCC -202	Biostatistical Inference	04	Core Course
BSTATCC -203	Epidemiology	04	Core Course
BSTATCC -204	Applied Regression Analysis and Generalized Linear Models	04	Core Course
BSTATCC -205	Applied Multivariate Analysis	04	Core Course
BSTATCC -206	Practical	04	Core Course
BSTATVNC-201	Environmental Statistics	00	Value added course (Non Credited)
	Semester Total	24	
	Semester III		
BSTATCC -301	Survival Analysis	04	Core Course / MOOC
BSTATCC -302	Practical	04	Core Course
BSTATEL -301A	Statistical Genetics	04	Elective
BSTATEL -301B	Qualitative Data Analysis		
BSTATEL-302A	Data Analysis Using SPSS	04	Elective
BSTATEL-302B	Official Statistics		
BSTATIN-301	Summer Internship	04	Summer Internship
BSTATIER-301	Disaster Risk Management	04	Interdepartmental Course
	Semester Total	24	
	Semester IV		
BSTATCC -401	Clinical Trials	04	Core Course
BSTATEL -401A	Computer Intensive Statistical Techniques	04	Elective
BSTATEL -401A	Gender Statistics		
BSTATEL-402A	Data Management Health Informatics	04	Elective
BSTATEL-402B	Ethics, Integrity and Aptitude		
BSTATMT-401	Master Thesis	08	Master Thesis
BSTATIRA-401	Data Analysis Using STATA	04	Intradepartmental Course
	Semester Total	24	
	GRAND TOTAL	96	

SEMESTER - I

BSTATCC-101: DESCRIPTIVE STATISTICS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcomes:

The objective of this course is to make the students aware of theproperties and applications of descriptive statistics.

Course Specific Outcomes:

On successful completion of the course the student will be able to:

1. Find joint, marginal and conditional probability distributions of descriptive statistics in the continuous and discrete cases.
2. Find the distribution of sample range and other systematic statistics in case of sampling from an arbitrary continuous population and, in particular, from some specific continuous distributions such as uniform and exponential.
3. Learn how to obtain distribution-free confidence intervals for population quantile and distribution-free tolerance intervals for population distributions based on order statistics.
4. Understand the distribution-free bounds for moments of order statistics and of the range.
5. Find the approximations to moments of order statistics in terms of quantile function and its derivatives.
6. Derive the recurrence relations and identities for moments of order statistics drawn from an arbitrary population (discrete or continuous), as well as from some specific distributions.

UNIT I

Statistical Data, Types of Data: attributes and variables, discrete & continuous data, Primary data, Secondary data, Different types of scales- nominal, ordinal, ratio and interval. Presentation of data: Construction of tables with one or more factors of classification.

UNIT II

Diagrammatic and graphical representation of data: Pictorial representation, Bar chart, Pie Chart, histogram, frequency polygon, frequency curve and ogives. Stem and leaf chart. Box Plot Central tendency and its measures: Mean, Median and Mode

UNIT III

Dispersion and relative dispersion and their measures: Range, Mean deviation, Standard deviation, Quartile deviation, coefficient of range, coefficient of variation, coefficient of mean deviation and coefficient of quartile deviation, Skewness and Kurtosis and their measures.

UNIT IV

Bivariate Data: Scatter diagram. Correlation and its measures: Karl Pearson's correlation coefficient and its properties, Spearman's rank correlation coefficient Kendall's τ, Intra-class correlation coefficient, correlation for bivariate frequency distribution, Correlation ratio, Multiple and partial correlation coefficients.

UNIT V

Attributes, contingency tables, concept of homogeneity and independence of attributes, Measures of association: Yule's coefficient, coefficient of colligation, Chi-Square coefficient, Karl Pearson's coefficient of contingency, Tschuprow's coefficient, Phi coefficient, Cramer's V, relative risk, odds ratio, Measures of agreement- Kappa and weighted kappa

REFERENCES

1. Alan Agresti: Categorical Data analysis; John wiley and Sons, New York, USA.
2. Bhatt B R, Srivenkatramana T and Rao Madhva K S (1996): Statistics: A Beginner's Text, Vol 1, and New Age International (P) Ltd.
3. Goon A M, Gupta M K, Das Gupta B. (1991): Fundamentals of Statistics. Vol 1, World Press, Calcutta.

ADDITIONAL REFERENCES

1. Anderson T W and Sclove S L (1978): An Introduction to Statistical Analysis of Data, Houghton Mffin\Co.
2. Snedecor G W and Cochran W G (1967): Statistical Methods. Iowa State University Press.
3. Spiegel, M R (1967): Theory and Problems of Statistics, Schaum's Publishing Series.

BSTATCC -102: ELEMENTS OF DEMOGRAPHY

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcomes:

The main objective of this course is to describe current population trends, in terms of fertility, mortality and population growth and the concepts of Statistical Quality Control, Quality Assurance and Performance Analysis.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Identify principle sources of demographic data and assess their strengths and weaknesses.
2. Discuss the demographic significance of age structures and the implications of variations in age structure.
3. Construct and interpret single-decrement life tables.
4. Specify and calculate the principal demographic measures, and standardize these measures for comparison and interpretation.
5. Identify the components of population change, including the effects of changing birth, death and migration rates, and demonstrate their influences on age structure.

UNIT I

Introduction to Demography: Source of Demography Data: Census, vital events, registration, survey, extent of under registration, Chandrasekhar Deming Index Mortality,

UNIT II

Measurements: Crude and specific rates, direct and indirect methods of standardization, Determinants of mortality

UNIT III

Fertility: Crude and specific rates, gross reproduction rates, net reproduction rates, parity progression ratio, child-women ratio, Determinants of fertility, fertility differentials,

UNIT IV

Life table- construction and uses, Abridged Life table-construction and uses, Concept of model life table

UNIT V

Migration: Measures of migration, Balancing equation, survival ratio method, Selectivity and differential. Factors affecting Migration: Push and pull Factors.

REFERENCES

1. Ram Kumar, R. (1986): Technical demography, Wiley Eastern Ltd, New Delhi,
2. Patnaik, K.B., Ram F.1991.Techniques of Demographic Analysis, Himalaya Publishing House, New Delhi.
3. Premi, M.K. Ramanamma, A and V. Bambawale; 1983.An Introduction to social Demography, Vikas Publishing Pvt. Ltd., New Delhi.
4. Rao. P.S.S., Jesudian G., Richard J,1983.An Introduction to Biostatistics 2nd Ed.,Department of Biostatistics, Christian Medical College, Vellore,
5. Bartholomew, D J (1982): Stochastic Models for Social Processes, John Wilew.
6. Benjamin B (1969): Demographic Analysis, George, Allen and Unwin.
7. Chiang, C L (1968): Introduction to Stochastic Processes in Biostatistics, John Wiley.
8. Cox P R (1970): Demography, Cambridge University Press.

9. Keyfitz N (1977): Applied Mathematical Demography, Springer Verlag.
10. Siegelman M (1969): Introduction to Demographic Analysis. Harward University Press.
11. Wolfenden H H (1954): Population Statistics and Their Compilation; American Actuarial Society.

BSTATCC -103: PROBABILITY THEORY AND DISTRIBUTIONS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The aim of this course is to provide a thorough theoretical grounding in probability theory and different type of probability distributions and their applications.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Formulate the mathematical/statistical models for real data sets arising in various fields in order to analyze with respect to various useful characteristics of the populations.
2. Understand how to use probability distributions in real life problems.

UNIT I

Random experiment, trial, sample point and sample space, events, operations of events, concepts of equally likely, mutually exclusive and exhaustive events. Definition of probability: Classical, relative frequency and axiomatic approaches, Discrete probability space, properties of probability under set theoretic approach Independence of events, Conditional probability, total and compound probability theorems, Bayes theorem and its applications

UNIT II

Random variables – discrete and continuous, probability mass function (pmf) and probability density function (pdf), Cumulative distribution function (cdf) and their properties. Joint pmf / pdf of several random variables, Marginal and conditional distributions of functions of discrete and continuous random variables, Independence of random variables, Expectation of a random variable and its properties, Moment Generating Function (m.g.f) and Characteristic Function. Inversion Formula

UNIT III

Special discrete distributions: discrete uniform, binomial, Poisson, geometric, negative binomial, hyper-geometric, Uniform, properties of these distributions.

UNIT IV

Special continuous distributions: continuous uniform, normal distribution, exponential distribution, Gamma distribution, Beta distribution, Bivariate Normal distribution, properties of these distributions.

UNIT V

Chebyshev's inequality, Weak law of large numbers and Central Limit Theorem for a sequence of independently and identically distributed random variables and their applications. Law of large numbers and central theorem, Review of Sampling Distributions, χ^2, t and F distributions with reference to Biostatistical aspects.

REFERENCES

1. Mood A. M., Graybill F A and Boes D. C. (1974): Introduction to the Theory of Statistics, Mcgraw Hill.
2. Chung K. L. (1979): Elementary Probability Theory with Stochastic Processes, Springer International Student Ed.
3. David Stirzaker (1994): Elementary Probability, Cambridge University Press.
4. Feller, W. (1968): An Introduction to Probability Theory and Its Applications, Wiley.

5. Mukhopadhyay, P. (1996): Mathematical Statistics, New Central Book Agency, Kolkatta.
6. Parzen E (1960): Modern Probability Theory and Its Applications, Wiley Eastern.
7. Pitman, Jim (1993): Probability, Narosa Publishing House.
8. Meyer, P.: Introductory Probability and Statistical Applications.

BSTATCC -104: RESEARCH METHODOLOGY

(4 Credits – 4 hours of T/L per week)

Course Outcome:

The main objectives of this course are:

1. To learn statistical techniques useful for research work.
2. To understand the qualitative and quantitative methods used biostatistical research.

Course Specific Outcomes:

After completing this course, the students will be able to:

1. Know different types of data produced in their area of study.
2. Create, manage, visualize, and summarize datasets.
3. Use and understand the inferential procedures.
4. Apply suitable sampling design.
5. Understand and apply basic designs.
6. Apply suitable statistical techniques to analyse the data and interpret the results.

UNIT I

Research Methods: Scientific Research: The scientific method and problem solving. Characteristics to the scientific approach. Purpose of scientific approach. Research as Decision- making process: Research alternatives, role of research methods in business and industry. Limitations of research. Major steps in the research process: Literature review, theoretical contexts, research problem, research hypothesis.

UNIT II

Research Designs: exploratory research studies in case of descriptive research and studies, causal research studies, Experimental, Quasi-Experimental and non-experimental research. Surveys and evolution research, Retrospective (case- control), Prospective (cohort and case-cohort) and case studies, historical research. Sampling techniques and sample size determination,

UNIT III

Techniques of control, Internal and external validity, characteristic of good design and time dimension. Measurement and Data collection: Primary and Secondary, Self report, observational, physiological measure. Projective techniques: Recodes and available data, Questionnaires and interview schedules, designing self- Report instruments, scales, response bias.

UNIT IV

Unstructured observational and structural observational methods, observational sampling, evolution, Errors of measurement, reliability, validity and other criteria for assessing measures. Research Report: Data preparation and preliminary Data Analysis.

UNIT-V

Statistical analysis, model building and decision making, the context, style of research report. Types of research documents, writing and formatting of report, presentation of critique.

REFERENCES

1. Cenise F. Polit and and applications, J.B. Bemadette P. Hungler (1984) Essential of Nursing Research Methods Lippinott Company, U.K.
2. Carol T. Bush (1985): Nursing Research, Reston Publishing C. Reston,
3. Bhattacharya, D. K. (2003): Research Methodology, Excel Books, New Delhi

ADDITIONAL REFERENCES

1. Cheryl B. Setter (1984): Nursing Research in a Service Setting Reston Publishing Co. Reston.
2. Louise H. Kidder (1981): Research Methods in Social Relations. IV Edition, Holt Rinehart and Winston, New York.
3. Pauling V. Young (1968): Scientific Social surveys and research Pretence- Hall, New Delhi.

BSTATCC -105: CORE VI : PRACTICAL

(4 Credits – 6 hours of T/L per week)

Note on Practical: Each practical session should correspond to two teaching hours.

List of Practical Experiments

1. Graphical representation of data by Histogram, Frequency polygons, frequency curves and Ogives, stem & Leaf Plot, Box Plot.
2. Calculation of measures of location.
3. Calculation of measures of dispersion.
4. Calculation of moments, measures of skewness and measures of Kurtosis.
5. Fitting of curves by method of least squares.
6. Determination of regression lines and calculation of correlation coefficient – grouped and ungrouped data.
7. Calculation of correlation ratios and rank correlation coefficients.
8. Calculation of multiple and partial correlation coefficients for three variables
9. Calculation of measures of association in contingency tables.
10. Direct and Indirect Methods of Standardization.
11. Construction of Life Table and Abridged Life Table.
12. Calculation of measures of mortality and fertility.
13. Population estimation using exponential and modified exponential function; Logistic, Makeham and Gompertz curves.
14. Estimation of net migration using survival Ratios.

And

Any other relevant practicals based on Papers BSTATCC101, BSTATCC102, BSTATCC103 and BSTATCC104 & BSTATCC105.

Practical work should be done on statistical packages or using high level languages. The purpose of this part is to use a statistical package such as MS- Excel /SPSS /S+/ R/ MINITAB/ etc. to carry out statistical procedures already known to students based on theory papers taught in semester I. No new statistical methods should be presented but interesting data can be analyzed using known methods on the package. Topic should include graphics, descriptive statistics, and representation of multivariate data.

BSTATVC-101: Credited: SUSTAINABLE DEVELOPMENT GOALS

(4 Credits – 4 hours of T/L per week)

Course Outcome:

The course aims to sensitize the students towards indicators and measures of sustainable development, give knowledge of and capacity to analyse policies and practices for sustainable development of different sectors, management strategies for water, waste, energy etc., and understand the components and techniques of EIA .It examine the relationship between SDGs, community problems and current sustainable and social solutions to serve as a starting point for developing new solutions that might serve as the business or social cases for new start-ups in health, sustainability or social ventures

Course Specific Outcomes:

After completing this course, students will:

1. Gain insight into the need, indicators and measures of sustainable development along with challenges and responses
2. Understand and critically analyze policies and practices regarding various sectors- energy, air, water, waste, agriculture, etc.

UNIT I

Concept of sustainability, framework, principles, dimensions of sustainability - social, economic and technological; changing perspective, promoting sustainable development, Need, indicators and measures for sustainable development.

UNIT II

Challenges to Sustainable Development- agriculture, population, food security, public health, education, natural resources, climate change, Responses to sustainable development: Public Policy (Community Participation and Participatory Specific), gender and human rights, technology etc.

UNIT III

SDG Index, Collection ,analysis and reporting of data, methodologies for index, computation of Index scores for different countries, MoSPI's National Indicator framework.

UNIT IV

Critical evaluation of government initiatives on sustainable development to produce a policy brief or a policy discussion paper. Focus areas may include, but are not limited to, policy analysis (context, history, objectives, framework and the process of policy formulation), linkages of the identified policy with other goals (SDGs etc.), implementation challenges (internal resistance, technical capacity, institutional and political perspectives, information etc.), outcomes and recommendations for potential modification of the policies. Case Studies.

UNIT V

SDG's and India's commitment, Biodiversity Conservation and Management, SDG index score for India, Highlighting data gaps related across SDGs for India to develop its statistical systems at the national and State levels.

REFERENCES

1. Scruton, Roger. (2012). How to Think Seriously about the Planet: The Case for an Environmental Conservatism, Oxford, Oxford University Press.
2. United Nations (2015), Transforming our world: the 2030 Agenda for Sustainable Development[Resolution adopted by the General Assembly on 25 September 2015], 70/1, New York, NY, UN General Assembly.
3. Sachs, Jeffrey D. (2015),The Age of Sustainable Development. New York, NY, Columbia University Press.

SEMESTER – II

BSTATCC -201: SAMPLING THEORY AND DESIGN OF EXPERIMENT

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The main objective of this course is to learn techniques in survey sampling and design of experiment with practical applications in biological and medical sector.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Understand the distinctive features of sampling schemes and its related estimation problems
2. Learn about various approaches (design based and model-based) to estimate admissible parameters; with and without replacement sampling scheme, sampling with varying probability of selection.
3. Learn about the methods of post–stratification (stratified sampling) and controlled sampling and also double sampling procedure with unequal probability of selection.
4. Learn about the applications of sampling methods; systematic, stratified and cluster sampling.
5. Understand the cluster and two stage sampling with varying sizes of clusters/first stage units.
6. Understand the super population approach to estimation.
7. Learn about the randomized response techniques.
8. To understand basic concepts of design of experiment and factorial experiments and apply them in biological fields.

Note:

For all unit I to IV: Concepts are to be discussed with simple proofs where needed or without proofs. Emphasis is to be given on the applications of these concepts in biological fields.

UNIT I

Concepts of population and sample, census and sample surveys, Basic concepts in sampling and designing of a large scale surveys, steps involved in sample survey, Types of sampling: sample and the probability sample; simple random sampling with and without replacement, Probability proportional to size sampling.

UNIT II

Systematic sampling, Stratified sampling, Cluster sampling, Multistage and Multiphase sampling, Ratio and Regression method of estimation, Double sampling. Estimation of sample size for clinical experiments, sources of error in surveys.

UNIT III

Estimation of mean proportion and standard error in Non-probability samplings: the convenience sampling, Judgment sampling, Quota sampling, interpenetrating sub-sample, Snowball sampling.

UNIT IV

Fundamental principles of design of experiments. Basic deigns - CRD, RBD and LSD. Their layout and analyses. Orthogonality of classification in two-way lay-outs, advantages of

orthogonality relation, simple illustrations. Fixed, mixed and random effect models; Analysis of co-variance, missing plot techniques -general theory and application, Crossover design,

UNIT V

Factorial experiments: 2^n, 3^2 factorial experiments, illustrations, main effects and interactions, confounding and illustrations. Application areas: Response surface experiments; first order designs and orthogonal designs.

REFERENCES

1. Murthy M N: (1967): Sampling theory and methods Statistical Publishing Society, Kolkatta.
2. Cochran W G: (1994): Sampling Techniques 3rd Edition, Wiley Eastern.
3. Des Raj and Chandok (1997): Sampling Design, Tata Mc Graw Hill.
4. Goon A M, Gupta, M K and Das Gupta, B (1986): Fundamental of Statistics; Vol II, World Press, Kolkatta.
5. Mukhopadhyay, Parimal (1996): Theory and Methods of Survey Sampling. Prentice Hall.
6. Sukhatme, Sukhatme, Sukhatme&Asok: Sampling Theory of Surveys with applications.

BSTATCC -202: BIOSTATISTICAL INFERENCE

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

To make students aware of estimation (point, as well as, interval) and testing (simple, as well as, composite hypotheses) procedures.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Apply various estimation and testing procedures to deal with real life problems.
2. Understand Fisher Information, Lower bounds to variance of estimators, MVUE.
3. Understand Neyman-Pearson fundamental lemma, UMP test, Interval estimation and confidence interval.

Note: For all unit I to IV: Concepts are to be discussed with simple proofs where needed or without proofs. Emphasis is to be given on the applications of these concepts in biological fields.

UNIT I

Concept of Statistical Inference: Parametric models, parameters, problem of inference Estimation: properties of point estimations, Minimum variance unbiased estimator, best linear unbiased estimator, Methods of point estimation: Maximum Likelihood estimators, Method of moments, Method of minimum variance, Method of Chi-square, Method of modified minimum- chisquare, Method of least squares, Illustration of Cramer – Rao inequality, Bhatacharya inequality, Rao – Blackwell inequality

UNIT II

Testing Hypothesis: Critical region and level o significance, test of a simple hypothesis against simple alternative, composite hypothesis, Neyman-Pearson test of hypothesis, UMP test, UMP unbiased test, Likelihood ratio test

UNIT III

Parametric tests based upon t , χ^2 and F distributions: Test for the mean of normal population, paired t-test, independent sample t-test, χ^2-test for the variance of normal population, χ^2- test for goodness of fit χ^2-test for association, F-test for equality of variances, Large sample tests (z-test) for means and proportions, test of correlation coefficient, One way ANOVA

UNIT IV

Non-Parametric tests: One sample sign test, Wilcoxon signed rank test for one sample and paired sample, Wilcoxon rank sum test, Median test, Mann-Whitney U test, Kolmogorov-Smirnov test for goodness of fit, Kolmogorov-Smirnov two sample test, Run test for randomness, Run test for equality of two distributions, Kruskal-Wallis test.

UNIT V

Interval estimation: Confidence interval for mean, variance of normal distribution, proportions, Correlation and regression coefficients, Confidence interval of mean and variance for small samples, Sample size: Its requirement and calculation, Sequential Analysis and Sequential probability ratio test.

REFERENCES

1. Kale B K (1999): A first Course on Parametric Inference, Narosa Publishing House.
2. Lehmann E L (1986): Theory of Point Estimation (Student Edition)
3. Lehman E L (1986): Testing Statistical Hypotheses (Student Edition)

4. Armitage P., Berry G. (1990).: Medical Research: Blackwell Scientific Publications
5. Hogg &Graig: Mathematical Statistics.
6. Rao C.R. (1973): Linear Statistical inference and its applications 2nd Ed. John Wiley & Sons, Inc.

BSTATCC -203: EPIDEMIOLOGY

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

Epidemiology is one area of Biostatistics that concerns itself with theapplication of statistical methods to medical, biological, and health relatedproblems.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Tackle the challenges associated with the study design and data analysis conducted in the health sciences.
2. Use and understand the principal numeric and graphical techniques to display and summarize medical and health related data.
3. Understand the basic principles of probability and how they relate to biostatistics.
4. Studying the relationship between a vector of covariates x and the rate of occurrence of specific types of failure
5. Evaluate, from simple datasets, evidence for linkage disequilibrium and disease associations using basic association tests
6. Discuss the wider issues involved in applying association tests to whole genomes.
7. Formulate and analyze stochastic epidemic models for specific purposes
8. Understand modelling epidemics with a structured underlying population.
9. Understand How to handle censored data, estimation of Kaplan-Meier survivor curves, the Log-Rank test for testing differences between survival curves, and Cox' regression model for estimating and testing effects of covariates (and interpretation of the statistical results).
10. Understand the introduction to basic statistics for clinical trials.

UNIT I

Definition, agent, host and environment, mode of transmission, incubation period, spectrum of disease, herd immunity, classification of cause of death, measure of mortality, studies of mortality, Measure of morbidity, morbidity surveys, issues and problems. Risk, cause and bias, Observational studies: retrospective, cross sectional and prospective studies.

UNIT II

Clinical trials: Methods of randomization, ethical issue, cross over trails, Sequential and group sequential trails, Interim analysis, multiple testing and stopping rules. Equivalence trails

UNIT III

Clinical Epidemiology: Definition, reliability, validity, sensitivity, specificity, predictive values, Likelihood ratio test, selection and interpretation of diagnostic test, Deciding on the best therapy, ROC curves, multiple and parallel test

UNIT IV

Screening for diseases, Critical approach, and Meta-analysis, Epidemiological Models- Epidemometric studies- Deterministic epidemic models: simple, General Recurrent- Stochastic epidemic models

UNIT V

Chain binomial and branching processes- Spatial models, Applications of Time series analysis in epidemiology simple descriptive techniques for detecting seasonal, cyclical, secular and random variations- Transformations-Trend analysis, Auto correlations, Auto regression, Forecasting.

REFERENCES

1. K J Rothman and S Greenland (ed.): Modern Epidemiology, Lippincott-Raven.
2. S. Selvin (1996): Statistical Analysis of Epidemiology data, Oxford University Press.
3. D Mcneil (1996): Epidemiological Research Methods, Wiley and sons.
4. J F Jekel, J G Elmore, D L Katz (1996): Epidemiology, Biostatistics and Preventing Medicine, WB Saunders Co.
5. Lilenfeld, A.M. &Lilenfeld, D.C (1980): Foundations epidemiology, II Ed., Oxford Univ. Press, New York.
1. Fletcher, R.H., Fletcher, S.W. and Wagner, E.H. 1982: Clinical Epidemiology- the essentials II Ed.
6. Harold A Kahn, 1989: Statistical Methods in Epidemiology. Oxford Univ. Press, New York.
7. David G, Kleinbaum, Lawrance L. Kupper and Hal Morgenstern 1982: Epidemiologic Research, Van Nostrand, USA.
8. Chatfield, C 1984 (Chapters 1, 2, 3, 4, 5, 6, 7): The Analysis of Time Series- An Introduction III Ed, Chapman & Hall, London.
9. Bailey, N.T.J, 1967 (Chapters 1, 2 and 9): The Mathematical Approach to Biology and Medicine. John Wiley.

ADDITIONAL REFERENCES

1. David L Sackett, 1967: Clinical Epidemiology, Little Brown & Co. USA,
2. Brian Mac Mohan and Thomas, F. Pugh: Epidemiology–Principles and Methods, Little Brown & Co. USA

3. Moolgavkar, S.H. &Prentice R L (Editors): Modern Statistical Methods in Chronic Disease Epidemiology. John Wiley & Sons

BSTATCC -204: APPLIED REGRESSION ANALYSIS & GENERALIZED LINEAR MODELS

(4 Credits – 4 hours of Theory teaching per week)

Course Outcome:

The main objective of this course is to provide students the ability to learn and use linear and non-linear models for normal data, and generalized linear models for normal and non-normal responses.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Use linear and Non-linear models, apply data transformations, and appreciate the need and uses of generalized linear models.
2. Use logistic and Poisson regression models.
3. Understand deviance, analysis of deviance, Lack-of-Fit tests in Logistic and Poisson regression, and the concept of over dispersion.
4. Use Log linear models for contingency tables, and likelihood ratio tests for various hypotheses including independence, marginal and conditional independence, and partial association.
5. Understand graphical and non-graphical models.
6. Use the concepts of Generalized Linear Models in real life problems.
7. Understand and apply Quasi likelihood.

UNIT I

Linear Models of full rank and not of full rank: Assumptions, Estimation, Analysis of variance table Dummy variables and its use in one way and two way classified data (One way and Two way ANOVA), Random and Mixed effect models, Analysis of covariance.

UNIT II

Regression diagnostic: Residuals and residual plots, Tests for departure from the assumptions of homoscedaticity, Multi-collinearity, Normality and autocorrelation,

Selection of best regression equation: Forward selection method, Backward selection method, Stepwise regression

UNIT III

Logistic and Poisson regression: logit model for dichotomous data with single and multiple explanatory variables, estimation of parameters, large sample tests about parameters, goodness of fit, Multinomial and ordinal logistic regression

UNIT IV

Log linear models for two and more dimensional contingency tables: interpretation of parameters, marginal and conditional independence, ML estimation of parameters, tests for various hypotheses including independence

UNIT V

Nonparametric regression and generalized linear models: interpolating and smoothing splines for simple regression, use of cross- validation, applications to logistic and Poisson regression, introduction to additive models and generalized additive models

REFERENCES

1. D. W. Hosmer and S. Lemeshow (2000). Applied logistic regression, second edition, Wiley,New York.

2. A. Agresti (1990). Categorical data analysis. Wiley, New York.
3. R. Christensen (1997). Log linear models and logistic regression, second edition, Springer,New York.
4. P. McCullagh and J. A. Nelder (1999): Generalized linear models, second edition, chapmanand Hall, New York.
5. P. J. Green and B. W. Silverman (1994): Nonparametric regression and generalized linearmodels, chapman and Hall, New York.
6. T. J. Hastie and R. J. Tibshirani (1999): Generalized Additive models, second edition, Chapmanand Hall, New York.
7. Draper N R and Smith H (1988): Applied Regression Analysis, 3rd edition, Wiley, New York.
8. Bates D M and Watts D G (1988): Non-linear Regression Analysis and its Application, Wiley,New York.
9. Cooks R D and Weisberg, S (1982): Residuals and Inference in Regression, Chapman andHall, London.

ADDITIONAL REFERENCES

1. D. A. Belsley, E. Kuh, and R. E. Welsch (1980): Regression diagnostics, Wiley.
2. P. McCullagh and J. A. Nelder (1999): Generalized linear models, Third edition, Chapman andHall, New York.
3. G. E. F. Seber and C. J. Wild (1989): Nonlinear regression. Wiley.
4. J. S. Simonoff (1996): Smoothing methods in statistics, Springer.

BSTATCC -205: APPLIED MULTIVATIATE ANALYSIS

(4 Credits – 4 hours of Theory teaching per week)

Course Outcome:

The main objective of this course is to introduce students to the analysis of observations on several correlated random variables for a number of individuals. Such analysis becomes necessary in Anthropology, Psychology, Biology, Medicine, Education, Agriculture and Economics when one deals with several variables simultaneously.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Account for important theorems and concepts in multivariate analysis.
2. Summarize and interpret multivariate data.
3. Appreciate the range of multivariate techniques available.
4. Understand the link between multivariate techniques and corresponding univariate techniques.
5. Conduct statistical inference about multivariate means including hypothesis testing, confidence region calculation, etc.
6. Use multivariate techniques appropriately, and draw appropriate conclusions.
7. Analyse multivariate data using the SPSS statistical software package.

UNIT-I

Multivariate data and their diagrammatic representation, Multivariate Normal Distribution, Its properties, Maximum likelihood estimators of parameters

UNIT-II

Maximum likelihood estimators of total, partial and multiple correlations, Tests of significance of total, partial and multiple correlations, Multinomial distribution and its properties.

UNIT III

Tests of significance of total, partial and multiple correlations. Hotelling's T^2 statistic and its use in testing- the significance of 1.mean vector from N ($\mu, \sum$) with unknown dispersion matrix $\sum$, 2. Equality of mean vectors from two Multivariate Normal Populations having same unknown dispersion matrix $\sum$,

UNIT IV

Problem of Classification and Discriminant analysis, Mahalnobis D^2 statistic and its use in Discriminant analysis. Principal Component analysis, Factor analysis: objectives, uses, limitations,

UNIT V

Canonical variates and canonical correlations and their uses, Cluster analysis and its uses, Multivaraite analysis of variance

REFERENCES

1. Morrison D F (1976): Multivariate Statistical methods, 2nd Edition, McGraw Hill.
2. B Everitt: Graphical Methods for Multivariate Analysis.
3. B Everitt: Cluster Analysis
4. Dellon W and Gold Stein M: Multivariate Analysis: Methods and Applications.
5. R A Johnson and D W Wichern (1988): Applied Multivariate Statistical Analysis.

6. A E Maxwell (1977): Multivariate Analysis in Behavioural Research.
7.Gorsuch, Richard L. (1983) Factor Analysis. Hillsdale, NJ: Erlbaum
9. Morrison, Donald F. (1990) Multivariate Statistical Methods. New York: McGraw-Hill.

ADDITIONAL REFERENCES

1. Armitage, P and Berry G. 1987, Statistical methods in Medical Research (chapter 10), Blackwell Scientific Publications, USA,
2. Anderson T.W.An Introduction to Multivariate Statistical Analysis. (Chapter 6, 11, 14), Wiley Eastern, New York, 1984.
3. Jae-On Kim and Charles W. Mueller 1986.Factor analysis: Statistical Methods and Practical issues, Sage Univ. paper 14, Sage publications, New Delhi,
4 .M.S. Aldenderfer and Roger K. Blissfield 1984.Cluster Analysis. Sage Univ. paper 44, Sage publications, New Delhi,
5. W.D. Berry and Stanley Feldman 1985, multiple regressions in practice Sage Univ. Paper 50, Sage Publications, New Delhi,
6. W.R. Klecka. 1980.Discriminate analysis Sage Univ. Paper 19, Sage Publications, New Delhi,
7. Chatfield C, and Collins, A.J. 1979. Introduction to Multivariate analysis. (Chapters, 4, 5, 6, 7) Chapman and Hall, New York, Arthertine Inc.
8. Johnson, R.A, and Wichern, W.D., 1988, Applied Multivariate Analysis. Prentice Hall, International Inc., USA

BSTATCC -206: PRACTICAL II

(4 Credits – 6 hours of T/L per week)

Note on Practicals: Each practical session should correspond to two teaching hours.

List of Practical Experiments

1. Fitting of Binomial, Poisson and Normal distributions to observed data and testing of goodness of fit.
2. Testing of independence of attributes in m x n contingency table and calculation of measures of association.
3. t – test for (i) $\mu = \mu_0$ (ii) $\mu_1 = \mu_2$ (iii) $\alpha = \alpha_0$ (iv) $\beta = \beta_0$ (v) $\rho = 0$ (vi) $\rho_{12..3} = 0$
4. F-test for (i) $\sigma_1^2 = \sigma_2^2$ (ii) $\rho_{1.23} = 0$
5. Fisher's Z-transformation and its use in testing (i) $\rho_1 = \rho_2$ (ii) $\rho_1 = \rho_2 = \ldots = \rho_k$ (iii) $\rho = \rho_0$
6. Calculation of power curve for the test of $\mu = \mu_0$ against $\mu \neq \mu_0$ for a normal distribution with known variance.
7. Large sample tests.
8. Drawing a simple random sample with the help of table of random numbers.
9. Estimation of population means and variance in simple random sampling.
10. Stratified Random Sampling-Estimation of mean and standard error-proportional allocation, Optimum allocation, Estimation of gain due to stratification.
11. Systematic Sampling
12. Estimation of sampling size in different sampling techniques.
13. Ratio and Regression estimation methods- Estimation mean, total and S.E.
14. Cluster Sampling- Equal and unequal cluster sizes.
15. Double sampling using regression and ratio estimates and double sampling for stratification.
16. Two stage and three stage sampling schemes.
17. Sequential probability ratio tests for parameters of binomial, Normal and Exponential distributions. (One sided alternatives with given β and).
18. ASN & OC functions for SPRT
19. Non- parametric tests. Test, Kolmogorov- Smirnov test, Median test, Wald run test, Mann Whitney U-test, multiple range test.

And

Any other relevant practical based on Papers BSTATCC202, BSTATCC204 and BSTATCC205.

Practical work should be done on statistical packages or using high level languages. The purpose of this part is to use a statistical package such as MS- Excel /SPSS /S+/ R/ MINITAB/ etc. to carry out statistical procedures already known to students based on theory papers taught in semester II. No new statistical methods should be presented but interesting data can be analyzed using known methods on the package.

BSTATVNC-201: (Non credited) ENVIRONMEMTAL STATISTICS

Course Outcome:

This course is to provide a systematic advanced treatment of areas of current interest in the statistical theory and methods for environmental data. Learn, discuss and apply statistical methods to important problems in environmental sciences, with a focus on understanding and quantifying change in environmental sciences or problems of this nature.

Course Specific Outcomes:

This paper introduces the students to various methods used in the collection of data and analysis for environmental studies. Simulation model are increasingly used to investigate the complexity of environmental processors. The paper introduces the students to the basics of modelling along with the application of remote sensing and GIS in different aspects of environmental studies.

Topics:

Data collection, survey and processing including social and cultural parameters. Statistics: Normal and binomial distribution. Hypothesis testing, t and chi square tests. Correlation and regression. Introductory analysis of variance. Multi-criteria analysis.

Modeling: Types of models: Mechanistic, economic, simulation etc. Fundamentals of building a model. Treatment of 2 or 3 environmental related models: Eutrophication model, global climate change model, wildlife habitat suitability model, air pollution model, ground water pollution model.

Remote Sensing: definition, principles, satellites and sensors. Aerial photography to Satellite Remote Sensing. Digital Image processing and image interpretation. GPS principles and applications.

Geographic Information System: concepts, database generation and analysis. Applications of Remote Sensing-GIS for Environmental Studies. Case studies, such as, (a) Land use / land cover change, Forest degradation, Urban sprawling. (b) Mining Hazards / Impacts, and (c) Forest Fire / Coal Fire Monitoring.

Applications of Biostatistics for Environmental Studies. Quantitative and qualitative research, Research question, Experimental design, Exercises on choice of statistical tools, presentation and interpretation of statistical results. Case studies based on contemporary research in ecology and environment.

REFERENCES:

1. Burrough, P.A., McDonnell, R.A. and Lloyd, C.D. (2015). Principles of Geographical Information Systems. Oxford University Press.
2. Dillon, J. and Wals, A. E. (2016). On the dangers of blurring methods, methodologies and ideologies in environmental education research. In Towards a Convergence Between Science and Environmental Education (pp. 113-124). Routledge.
3. Freund, R.J., Wilson, W.J. (2003). Statistical Methods. Academic Press.
4. Lillesand, T.M. and Kiefer. R.W. (1999). Remote sensing and image interpretation (4thed). Wiley.
5. Nieuwenhuijsen, M. J. (ed.) (2015). Exposure Assessment in Environmental Epidemiology. Oxford University Press, USA.
6. Wainwright, J. and Mulligan, M. eds. (2005). Environmental Modelling: Finding Simplicity in Complexity. John Wiley & Sons.

SEMESTER III

BSTATCC -301: SURVIVAL ANALYSIS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

Survival Analysis is one area of Applied Statistics that concerns itself with the application of statistical methods to medical, biological, epidemiological and health related problems.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Tackle the challenges associated with the study design and data analysis conducted in the health sciences.
2. Use and understand the principal numeric and graphical techniques to display and summarize medical and health related data.

UNIT I

Concepts of time, Order and random Censoring, likelihood in these cases, Life distributions- Exponential gamma, Weibull, Lognormal, Pareto, Linear Failure rate

UNIT II

Parametric inference (Point estimation, Confidence Intervals, Scores, LR, MLE tests (Rao – willks –Wald)) for these distributions.Life tables, failure rate, mean residual life and elementary properties

UNIT III

Ageing classes and their properties, Bathtub Failure rate. Estimation of survival function – Actuarial Estimator, Kaplan-Meier Estimator, Estimation under the assumption of IFR/DFR.

UNIT IV

Tests of exponentiality against non-parametric classes- total time on test, Despande test. Two sample problem- Gehan test, Log rank test, Mantel-Haenzel test, Tarone – ware tests.

UNIT V

Semi-parametric regression for failure rate- Cox's proportional hazards model with one and several covariates, Rank test for regression coefficients, Competing risk model, parametric and non-parametric inference for this model

REFERENCES

1. Cox D R and Oakes D (1984): Analysis of Survival Data, Chapman and Hall, New York.
2. Gross A J and Clark V A (1975): Survival Distribution: Reliability applications in the Biomedical Sciences, John Wiley and sons.
3. Elandt- Johson, R E Johnson N L , Survival Models and Data Analysis, John Wiley and sons.
4. Miller, R G (1981): Survival Analysis, John Wiley.
5. Kalbfleisch J D and Prentice R L (1980), The Statistical of failure Time Data, John Wiley.

OR

MOOC (such as SWAYAM, etc.)/ Project/Internship (4 Credits)

The candidates may opt any course from MOOC of the UGC/other websites and credit to be obtained is to be submitted in the office of CoE.

BSTATCC -302: PRACTICAL

(4 Credits – 6 hours of T/L per week)

List of Practical Experiments

1. Analysis of variance-Two way classification. Missing plot technique-R.B.D / L.S.D. Mixed up plots technique.
2. Analysis of variance in one-way and two-way classification (with and without interaction terms).
3. Analysis of a Latin square design.
4. Analysis of variance in RBD and LS design with one or two missing observations.
5. Analysis of Covariance. Two way classification- concommitment variable.
6. Factorial Experiments 2^3, 3^2 , and 2x3x4 types.
7. Confounding (partial/ Full) in factorial Experiment- 2^2, 3^2
8. OLS estimation and prediction in GLM.
9. GLS estimation and prediction.
10. Tests for autocorrelation. BLUS procedure.
11. Measurement of trend and seasonal factors
12. Experiments based on survival analysis
13. Design of Epidemiologic Studies
14. Analysis and inference of Epidemiological Data.

And

Any other relevant practicals based on Papers BSTATCC201 and BSTATCC301.

Practical work should be done on statistical packages or using high level languages. The purpose of this part is to use a statistical package such as MS- Excel /S+/ R/ MINITAB/ SPSS etc. to carry out statistical procedures already known to students based on theory papers taught in semester III. No new statistical methods should be presented but interesting data can be analyzed using known methods on the package

BSTATEL -301A: STATISTICAL GENETICS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

1. Acquire deep understanding of Mendelian genetics and its applications; population genetics, human genetics and microbial genetics. Learns about applied genetics and gene mapping methods.
2. Study descriptive statistics and applies in academic and research fields.
3. Develop practical understanding of research methodology and applies in research (formulation and defining a research problem, different types of research design techniques involved, Literature survey, execution of research plan and scientific documentation

Course Specific Outcomes:

After completing this course, the students will be able to:

1. Understand Population Genetics, Computational Biology, Statistical Genomics.
2. Study gene action, estimate genetic parameters and genetic merit, genetic progress and other understand related statistical methods in genetics.

UNIT I

Basic biological concepts in genetics (relevant to this course), Mendel's law, Hardy Weinberg equilibrium, Mating tables, estimation of allele frequency (dominant /co-dominant cases)

UNIT II

Approach to equilibrium for x-linked gone, natural selection, mutation, genetic drift, equilibrium when both natural selection and mutation are operative

UNIT III

Non- random Mating, inbreeding, phenotypic assortative mating Analysis of family data (a) relative pair data, 1, T, 0 matrices, identity by descent (b) Family data – estimation of segregation ratio under ascertainment bias, (c) pedigree data: Elston- Stewart algorithm for calculation of likelihoods.

UNIT IV

Linkage, Estimation of recombination fraction, inheritance of quantitative traits, Models and estimation of parameters, Sequence similarity, homology and alignment

UNIT V

Algorithms for (a) pair wise sequence alignment, (b) multiple sequence alignment, construction of phylogenetic trees, UPGMA, neighbour joining, parsimony and maximum likelihood algorithms.

REFERENCES

1. C. C. Li (1976): First course on population genetics. Boxwood Press, California
2. W. J. Ewens (1979): Mathematical population genetics. SprigerVerlog.
3. T. Nagylaki (1992): Introduction to theoretical population genetics. SprigerVerlog.
4. R. Durbin, S. R. Eddy, A. Krogh, G. Mitchison (1998): Biological sequential Analysis: Probabilistic models of proteins and nucleic acids.
5. Emery A.E.G. (1986): Methodology in Medical Genetics Anintroduction to statistical methods. Churchill Livingstone, New York.
6. Crow J.F. and Kimura M. (1970): An Introduction to population Genetic Theory. Harper and Row New York, (Selected Chapters).
7. Curt Stern and W.H. Freeman (1960): Principles of Human Genetics, company, San Francisco.

ADDITIONAL REFERENCES

1. R. C. Elandt- Johnson (1975) Probability models and statistical Methods in genetics. John Wiley.
2. Cavalli – Sforza, LL. W.H. Freeman and Bodmer W.F. (1971): The Genetics of Human population company, San Francisco (Selected Chapters).
3. Walter J. Burdette. (1962): Methodology in Human Genetics. Holden Day Inc., San Francisco.
4. Schull W.J. and Neel J.V. (1965): The Effects of Inbreeding of Japanese Children, Harper & Row, New York.

BSTATEL -301B: QUALITATIVE DATA ANALYSIS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This course equips students with the necessary skills to analyse and interpret qualitative data to answer research and policy questions

Course Specific Outcomes:

1. Upon successful completion, students will have the knowledge and skills to:
2. Understand a range of qualitative analysis approaches that are commonly employed across disciplines;
3. Critically evaluate the advantages and disadvantages of a variety of qualitative analysis methods and select appropriate methods for application;
4. Develop and apply skills in thematic coding techniques;
5. Apply skills in qualitative data analysis using appropriate data management software; and
6. Assemble and present the results of qualitative research analyses in written and oral formats.

UNIT I

Categorical response variables: Nominal, ordinal, interval. Probability structure for contingency tables: joint, marginal and conditional probabilities, sensitivity and specificity, independence. Comparing proportions in 2x2 Tables: difference of proportions, relative risk. Odds Ratio: definitions and properties of odds ratio with examples, inference for odds ratio and log odds ratio, relationship between odds ratio and relative risk. Chi-square tests of independence: Pearson statistic, likelihood ratio statistic, tests of independence, partitioning Chi-squared.

UNIT II

Testing independence for ordinal data: linear trend alternative to independence, extra power with ordinal test, choice of score, trend tests for Ix2 and 2xJ tables, nominal-ordinal tables. Exact inference for small samples: Fisher's exact test for 2x2 table, p-values and conservatism for actual P(Type I error), small sample confidence interval for odds ratio. Association in three-way table: partial tables, conditional versus marginal associations, Simpson's paradox, conditional and marginal odds ratios, conditional independence versus marginal independence, homogeneous associations.

UNIT III

Models for binary response variables: logit, log linear, linear probability and logistic regression models. Logit models for categorical data, probit and extreme value models, models with log-log link, model diagnostics. Fitting logit models, conditional logistic regression, exact trend test. Log-linear models for two dimensions - independence model, saturated model and models for cell probabilities.Item Response Theory, Rasch Model.

UNIT IV

Loglinear models for two-way and three-way tables: log-linear model of independence for two-way table, saturated model for two-way tables, log-linear models for three-way tables. Inference for loglinear models: Chi-squared goodness of fit tests, log-linear cell residuals, tests about conditional associations, confidence intervals for conditional odds ratios, three factor interactions, large samples and statistical versus practical significance.Fitting Log-linear models. Strategies in model selection, analysis of residuals, Cochran-Mantel-Haenszel test.

UNIT V

Models for Matched Pairs: Comparing Dependent Proportions, Conditional Logistic Regression for Binary Matched Pairs, Marginal Models for Square Contingency Tables, Symmetry, Quasi-symmetry, and Quasiindependence, Analyzing Repeated Categorical Response Data: Comparing Marginal Distributions: Multiple Responses, Marginal Modeling: Maximum Likelihood Approach, Marginal Modeling: Generalized Estimating Equations Approach, Quasi-likelihood and Its GEE Multivariate Extension, Markov Chains: Transitional Modeling.

REFERENCES

1. Agresti, A. (2010): Analysis of ordinal categorical data, Wiley.
2. Agresti, A. (2013): Categorical Data Analysis, Third Edition, Wiley.
3. Bilder, C. R. and Loughin, T.M. (2013): Analysis of Categorical Data with R, CRC Press.
4. Bowerman, O. (2000): Linear Statistical models.
5. Congdon, P. (2005): Bayesian Models for Categorical Data, Willey.
6. Kleinbaum, D. G. (1994): Logistic Regression, Springer Verlag.
7. Sutradhar, B. C. (2014): Longitudinal Categorical Data Analysis, Springer.
8. Upton, G.J.G. (2017): Categorical Data Analysis by Example, Wiley.

BSTATEL-302A: DATA ANALYSIS USING SPSS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

a) To train students in quantitative analysis of data

b) To formulate an optimization problem

c) To make students well versed with charts & graphs and their interpretation in business for decision making

d) To make a meaningful interpretation from data in business decision making.

Course Specific Outcomes:

On completion of this course, the students will be able

1. To deal with statistical and mathematical tools for precision in decision-making situation
2. To understand the optimality of decision making while dealing operations
3. To give a meaning interpretation to the data set to solve the business problem
4. To know the concept of probability to deal with business problem under uncertain condition
5. To develop the skill in statistical tools like correlation and regression to find the relationship between variables

UNIT – I

Data files, Distributed Analysis Mode, Data Editor, Working with Multiple Data Sources, Variable View-Data Preparation, Data Transformation : File Handling, File Transformation, Working with Command Syntax.

UNIT – II

Analyzing Data, Code Book, Frequencies, Descriptives, Explore, Cross Tabs, Summarize, Means, OLAP Curves, Tables and IGRAPH, Command Syntax, Syntax Commuter.

UNIT – III

Parametric Tests (t-tests), One way ANOVA, Different Non Parametric Tests, Bivariate Correlation, Partial Correlation, Association and its different measure

UNIT - IV

Linear Regression, Ordinal Regression Partial Least Square Regression. General Linear Models and its use in fitting ANOVA and ANCOVA models, Logistic regression

UNIT – V

Nearest Neighbor Analysis, Discriminant Analysis, Principal component analysis, Factor Analysis, Choosing a Procedures for Clustering, Two Step Cluster Analysis, K-Means Cluster Analysis,

REFERENCES

1. Margan G A: SPSS for Introductory Statistics; Uses and Interpretation.
2. Practical Work Book by Bristol Information Services: Introduction to SPSS for Windows.

BSTATEL-302B: OFFICIAL STATISTICS

Course Outcome:

This paper will help to know about different dimensions and issues related to our country through data and images such as graphs and visuals. It will help to provide basic information for decision making, evaluation related to administrative issues and policy making.

Course Specific Outcomes:

This course will provide students with an opportunity to develop:

1. Skill to visualize status of country.
2. To enhance basic information for decision making, evaluation, monitoring etc.
3. It will help in good policy and decision making.

UNIT-I

National and International official statistical system Official Statistics: (a) Need, Uses, Users, Reliability, Relevance, Limitations, Transparency, its visibility (b) Compilation, Collection, Processing, Analysis and Dissemination, Agencies Involved, Methods.

UNIT-II

National Statistical Organization: Vision and Mission, NSSO and CSO; roles and responsibilities; Important activities, Publications etc.

UNIT-III

National Statistical Commission: Need, Constitution, its role, functions etc.; Legal Acts/ Provisions/ Support for Official Statistics; Important Acts
Index Numbers: Different Types, Need, Data Collection Mechanism, Periodicity, Agencies Involved, Uses.

UNIT-IV

Sector Wise Statistics: Agriculture, Health, Education, Women and Child etc. Important Surveys & Census, Indicators, Agencies and Usages etc.

UNIT-V

National Accounts: Definition, Basic Concepts; issues; the Strategy, Collection of Data and Release. Population Census: Need, Data Collected, Periodicity, Methods of data collection, dissemination, Agencies involved, Misc.: Socio Economic Indicators, Gender Awareness/Statistics, Important Surveys and Censuses

REFERENCES

1. Goon A.M., Gupta M.K. and Dasgupta B. (2008): Fundamentals of Statistics (Vol.2), World Press.
2. Guide to current Indian Official Statistics, Central Statistical Office, GOI, and New Delhi.
3. C.S.O. (1984) : Statistical System in India
4. Mudgett B.D. (1951): Index Numbers, John Wiley
5. Allen R.G.D. (1975): Index Numbers in Theory and Practice, Macmillan
6. Mukhopadhyay P. (1999): Applied Statistics

BSTATIN-301 Summer Internship

Candidates/Students are required to attend an Internship Program during Summer Break. The Credits shall be provided by the Departmental Committee on submission of the report of internship.

BSTATIER-301 DISASTER RISK MANAGEMENT

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The course is intended to provide a general concept in the dimensions of disasters caused by nature beyond the human control as well as the disasters and environmental hazards induced by human activities with emphasis on disaster preparedness, response and recovery in global, national and regional level scenario of disaster management.

Course specific Outcomes:

The students will be able to differentiate the types of disaster, causes and their impact on environment and society. They will be able to assess vulnerability and various methods of risk reduction measures in Indian context and apply the skills in the interest of greater good of all.

UNIT-I

Understanding the Concepts and definitions of Disaster, Hazard, Vulnerability, Risk, Capacity – Disaster and Development, and disaster management, Geological Disasters (earthquakes, landslides, tsunami, mining); Hydro-Meteorological Disasters (floods, cyclones, lightning, thunder-storms, hail storms, avalanches, droughts, cold and heat waves) Biological Disasters (epidemics, pest attacks, forest fire); Technological Disasters (chemical, industrial, radiological, nuclear) and Manmade Disasters (building collapse, rural and urban fire, road and rail accidents, nuclear, radiological, chemicals and biological disasters) Global Disaster Trends – Emerging Risks of Disasters – Climate Change and Urban Disasters

UNIT-II

Risk : Its concept and analysis, Risk Reduction, Vulnerability : Its concept and analysis, Strategic Development for Vulnerability Reduction

UNIT-III

Disaster Preparedness: Concept and Nature, Disaster Preparedness Plan, Prediction, Early Warnings and Safety Measures of Disaster, Role of Information, Education, Communication, and Training, Role of Government, International and NGO Bodies, Role of IT in Disaster Preparedness and Disaster Management.

UNIT-IV

Introduction to Disaster Response, Disaster Response Plan, Communication, Participation, and Activation of Emergency Preparedness Plan, Search, Rescue, Evacuation and Logistic Management, Psychological Response and Management (Trauma, Stress, Rumor and Panic), Relief and Recovery, Medical Health Response to Different Disasters

UNIT-V

Reconstruction and Rehabilitation as a Means of Development, Damage Assessment, Post Disaster effects and Remedial Measures, Creation of Long-term Job Opportunities and Livelihood Options, Disaster Resistant House Construction, Sanitation and Hygiene, Education and Awareness, Dealing with Victims' Psychology, Long-term Counter Disaster Planning, Role of Educational Institution.

REFERENCES

1. Mrinalini Pandey: Disaster Management; Wiley India Pvt. Ltd.
2. Tushar Bhattacharya: Disaster Science and Management; McGraw Hill Education (India) Pvt. Ltd.

3. Jagbir Singh: Disaster Management : Future Challenges and Opportunities; K W Publishers Pvt. Ltd.
4. J. P. Singhal: Disaster Management; Laxmi Publications
5. Shailesh Shukla, Shamna Hussain: Biodiversity, Environment and Disaster Management; Unique Publications
6. C. K. Rajan, NavalePandharinath; Earth and Atmospheric Disaster Management : Nature and Manmade; B S Publication
7. Natural Disasters, David Alexander, Kluwer Academic London, 1999
8. Management of Natural Disasters in developing countries, H.N. Srivastava & G.D. Gupta, Daya Publishers
9. National Disaster Management Policy, 2009, GoI
10. Coppola D P, 2007. Introduction to International Disaster Management, Elsevier Science (B/H), London.
11. KapurAnu Vulnerable India: A Geographical Study of Disasters, IIAS and Sage Publishers, New Delhi, 2010.
12. Gupta Anil K, Sreeja S. Nair. Environmental Knowledge for Disaster Risk Management, NIDM, New Delhi, 2011

SEMESTER IV

BSTATCC -401: CLINICAL TRIALS

(4 Credits – 4 hours of Theory Teaching per week)

Course Ooutcome:

To have an understanding of current research methodology, collection and interpretation of data, ability to carry out research projects on clinical and epidemiological aspects, a working knowledge on current databases, automated data retrieval systems, referencing and skill in writing scientific papers.

Course Specific Outcomes:

After successful completion of this course, student will be able to:

1. Apply statistical techniques to solve clinical and research problems.
2. Gain knowledge on the different techniques and bio instruments that can be used for the research.

Carry out advanced research in specialized areas and transmit their knowledge to the society.

UNIT I

Introduction to clinical trials: the need and ethics of clinical trials, bias, and random error in clinical studies, conduct of clinical trials, overview of Phase I – IV trials, multi-center trials. Protocol: Inclusion/ exclusion criterion

UNIT II

Primary and secondary responses, Monitoring, Analytical methods, single and double blinding. Data management: Data definitions, case report forms, database design, data collection systems for good clinical practice.

UNIT III

Design of clinical trials: Parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, review of factorial designs, objectives and endpoints of clinical trials, design of phase I trials, design of single-stage and multi-stage Phase II trials,

UNIT IV

design and monitoring of Phase III trials with sequential stopping, design of bioequivalence trials. Non-inferiority trials. Superiority trial, Reporting and analysis: Analysis of categorical outcomes from Phase I – III trials, analysis of survival data from clinical trials.

UNIT V

Multi centric trials, Surrogate endpoints: Selection and design of trials with surrogate endpoints, analysis of surrogate endpoint data, Multiple end points and decisions, Meta-analysis of clinical trials.

REFERENCES

1. S. Piantadosi (1997): Clinical Trials, A Methodologic Perspective, Wiley and sons.
2. C. Jennison and B W Turnbull (1999): Group Sequential Methods with Applications to Clinical Trials, CRC Press.
3. L M Friedman, C Furburg, D L Demets (1998): Fundamentals of Clinical Trials, Springer Verlag.
4. J L Fleiss (1989): The Design and Analysis of Clinical Experiments, Wiley and sons.
5. E Marubeni and M G Valsecchi (1994): Analyzing Survival Data from Clinical Trials and Observational Studies, Wiley and sons.
6. Armitage, P. and Berry, G. (1994) Statistical Methods in Medical Research, 3rd Ed. Blackwell, Oxford.

7. Breslow, N.E. and Day, N.E. (1987) Statistical methods in cancer research. Volume II - the design and analysis of cohort studies IARC, Lyon.
8. Easterbrook. P.J., Berlin, J.A., Gopalan, R., and Mathews, D.R. (1991) Publication bias in clinical research. Lancet 337 867-72.
9. Glasziou, P.P. and Mackerras, D.E.M. (1993) Vitamin A supplementation in infectious disease: a meta-analysis. British Medical Journal 306 366-70.
10. Pocock, S.J. and Hughes, M.D. (1990) Estimation issues in clinical trials and overviews. Statistics in Medicine 9 657-71.
11. Thompson, S.G. (1993), Controversies in meta-analysis: the case of the trials of serum cholesterol reduction. Statistical methods in medical research 2 173-92.

BSTATEL -401A: COMPUTER INTENSIVE STATISTICAL TECHNIQUES

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

1. To inculcate and develop aptitude to apply statistical tools at a number of data generating fields in real life problems.
2. To train students to handle large data sets and carry out data analysis using software and programming language.
3. To teach a wide range of statistical skills, including problem-solving, project work and presentation so as enable students to take prominent roles in a wide spectrum of employment and research.

Course specific Outcomes:

On successful completion of the course a student will be able to:

1. Gain sound knowledge in theoretical and practical aspects of Statistics.
2. Describe complex statistical ideas to non-statisticians.
3. Handle and analyse large databases with computer skills and use their results and interpretations to make practical suggestions for improvement.

UNIT I

Stochastic simulation: generating random observations, simulating multivariate distributions, simulating stochastic processes, such as simple queues. Variance reduction: importance sampling for integration.

UNIT II

Monte Carlo simulation methods, Gibbs sampling for multivariate simulation, simulated annealing for optimization.

UNIT-III

Simulation based testing: simulating test statistics and power functions, permutation tests.

UNIT IV

Bootstrap methods: resampling paradigms, bias and standard errors, confidence intervals, bootstrapping in regression.

UNIT V

Jackknife and cross validation: jackknife in sample surveys cross validation for tuning parameters.

REFERENCES

1. G. S. Fishman (1969): Monte Carlo concepts, algorithms, and applications. Springer.
2. R. Y. Rubinstein (1981): Simulation and the Monte Carlo method. Wiley.
3. M. A. Tanner (1996): Tools for Statistical inference, Third edition. Springer.
4. B. Efron and R. J. Tibshirani (1993): An introduction to the Bootstrap. Chapman and Hall.
5. J. Shao and D. Tu (1995): The jackknife and the bootstrap. Springer Varlag.
6. McCullagh P and Nelder J A (1989): Generalized Linear Models, 2nd edition, Chapman and Hall, London.
7. Searle S R (1987) Linear Models for unbalanced Data, Wiley, New York.
8. Seber, G A and Wild, G J (1989): Nonlinear Regression, Wiley, New York.

BSTATEL -401A: GENDER STATISTICS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This course will introduce basic concepts relating to gender and provide logical understanding of gender roles. It will provide insight on gender disparities within family, economy, education and political system.

Course specific Outcomes:

The student will be able to acquire comprehensive knowledge regarding gender perspective in every sphere of life. They will be able to understand gender specific developmental need and work for better gender equity in all spheres with clarity of concepts.

UNIT-I

Sex and Gender, Male and Female Gaze and Objectivity, Social construction of Femininity and Masculinity, Gender roles: Biological vs cultural determinism, Private vs public dichotomy, Existential foundation of gender-power relations- A Paradigm Shift : from statistics to Gender statistics: Objectives and goals of Gender Statistics, Gender and Family - Gender role socialization and formation of identity, Psychoanalysis, social constructionist and discursive analysis of gender, Gender and Economy-Gender auditing, and budgeting in local governance.

UNIT-II

Definition of violence and gender based violence, Classification of gender based violence- Violence in Family : Domestic violence: physical, sexual, psychological and verbal etc., Forms of violence against women in family, Community and State, Violence by law enforcing agency, Response to Violence: National Human Rights Commission– National Commission for Women– Law Enforcing Agencies: All Women's Police Stations, Vigilance Cells, Legal Aid– Cells Judiciary: Family Courts/Mahila Courts– Women and Children Helplines, Roles of INGOs and NGOs in context of gender violence, Women Empowerment, National policies and programs for Women Empowerment , Engendering development policies

UNIT-III

Gender division of labour and asymmetric role structure, Gender Inequality in Labor Market - Gendered jobs and Social Inequality, Sex Segregation at Work Place, Gender Marginalisation and Glass Ceiling , Gender disparity: Global and Indian scenario

Theoretical perspectives of Gender and Development,Measures of Gender Development indicators and Human development indicators, Gendered impact of globalization, Gender specific consequences of environmental degradation, Role of women in Sustainable Development, Gender dimensions of poverty, Role of Health and Education in Gender Development, Gender Stereotyping in Media, Portrayal of Gender in Print Media, Gender and Electronic Media,Gender Stereotyping- Commercialization and Objectification

UNIT-IV

Gender Equality and Equity, Gender Inequality, Practical needs and Strategic, Gender as Analytical Category, Objectives of Gender Analysis, Gender Analysis Framework, Gender Roles Framework, Gender Planning Framework, Gender Analysis Matrix, Gender Analysis Process - Collecting Disaggregated Data, Assessing Gender Division of Labour and Decision making Pattern, Understanding Complexity of Gender Relation, Assessing Counterpart/ Partner Capacity for Gender Sensitive Plan, Assessing Potential of Program/ Project to Empower Women,Developing Gender Sensitive Indicators, Addressing Gender Issues in Project Cycle, Gender Analysis Tools- Problem Wall, Activity Calendar, FGD, Pair Wise Rank, Knowledge Mapping, PRA Techniques, Constraints and Opportunity Mapping,

Gender Analysis in Various Contexts,Programme Planning and Projects Designing, Gender Analysis in Sectoral Programmes – Case studies.

UNIT- V

Definition and Strategies of Gender Mainstreaming, Steps in Gender Mainstreaming - Collecting Disaggregated Data and Developing Analytical Framework, Developing Gender Equality Action Plan, Gender Sensitive Indicators, Tools and Techniques for Gender Action plan- Analytical Tools, Statistics, Surveys, Research reports, Forecasts, Checklists, Guidelines, Gender– Impact Assessment methods : Educational Tools, Awareness raising, Training, Follow up Action and Consultative Tools, Gender Mainstreaming in National/ State Policy.

Gender representation and Discrimination in Indian Polity, Political Participation of Women in Pre- and Post Independent India, Gender perspectives of Voting Behaviour and Electoral Process, Political Participation of Women: Opportunities and constraints. Indian philosophy on gender issues and policies.

REFERENCES

1. Cornell R W (1995) Gender. Cambridge, Polity Press.
2. Holmes M (2007) What is Gender. New Delhi, Sage Publications.
3. Jackson S and Scott S (2002) Gender: A Sociological Reader. New York:
4. Routledge. Kesseler S J And Mckenna W (1978) Gender: An Ethnomethodological Approach. Chicago: University of Chicago Press.
5. Kimmel S Michael (2004) The Gendered Society; Reader. Oxford: Oxford University Press.
6. Lipman- Blumen J (1984) Gender Roles and Power. New Jersey: Prentice Hall.
7. Oakley A (1985) Sex, Gender and Society. London: Temple Smith. Stanley L and Wise S. (1983) Breakling out Again: Feminist Methodology and Epistemology. London: Routledge.
8. Datta, R. and Kornberg, J. (eds.) (2002) Women in Developing Countries, Assessing Strategies for Empowerment. London: Lynne Rienner Publishers.
9. Eade, D. (ed.) (1999) Development with Women: Selected essays from Development in Practice. Great Britain: Oxfam.
10. Hunt, J, 2004. 'Introduction to gender analysis concepts and steps', Development Bulletin, no. 64, pp. 100-106

BSTATEL-402A: DATA MANAGEMENT AND HEALTH INFORMATICS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The main objective of this paper is to introduce some computational and data management techniques to extract information, visualization and learn their application in health domain.

Course Specific Outcomes:

After successful completion of this course, student will be:

1. Equipped with different theoretical methods and practicable techniques to achieve the objectives.
2. Enhanced with the basic concepts of statistical theories besides developing their ability to handle real world problems with large scale data.

UNIT I

Principles and practice or Health Data Management- Manual, electrical, electronic devices-strategies, Data Base Management (using Fox Pro or Visual Basic or any available dbms package) : Creating structure and command file, interacting with the user commands, accepts, input, wait, get. DO WHILE, IF ELSE….ENDIF, DO CASE. ENDCASE. Debugging techniques. Procedure files, public memory variables. Techniques for report and form. Linking database, adding and deleting data to linked to database.

UNIT II

Essentials of C++: Functions and parameters, classes, constructions, input/output, control statements such as if-self, switch, for, while and do-while, pointers and references,

UNIT III

Dynamic allocation, processing of linked lists, array and character strings, and libraries. Introduction to program analysis: simple testing and debugging.

UNIT IV

Introduction to available Statistical packages as SPSS / R / S Plus, Use any one of relevant package for Defining data, data transformation, selective, weighting and ordering cases, Restructuring files

UNIT V

Frequencies, Descriptive, Cross tabs and mean procedures. t-test correlation, ANOVA and Non-parametric procedures. Discriminant function and principal component procedure.

Note: Student Are Expected To Interpret The Output.

REFERENCES

1. Byron S, Gottfried (1986): Theory and Problems of programming with BASIC, Mc Graw Hill Co. New York.
2. Marija J. Norusis (1988): Statistical Package for Social Sciences, SPSS Inc., USA,

ADDITIONAL REFERENCES

1. Thomas WM. Madron (1985):Using Microcomputers in Research, Sage Univ.52 Sage C. Neal Tates and Robert Publications, New Delhi.G. Brookshire
2. Philip A. Schrodt. (1984):Microcomputer Methods for social scientists. Sage Univ. paper 40.Sage Publications, New Delhi.

3. R Decker and S. Hirshfield (1998): The object Concept: AnIntroduction to Computer Programming using C++, PWS Publishing.
4. S B Lippmann and J. Lajoie (1998): C++ Primer. 3rd edition, Addition_Wesley.
5. Brian Croritt and Sophia-RabeHesketh (2006): Analysing Medical Data using SPSS.
6. W J Savitch: Problem Solving with C++: The Object of Programming. 3rd edition, Addition-Wesley Longman.

BSTATEL-402B: ETHICS, INTERGRITY AND APTITUDE

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This paper includes questions to develop the student's attitude and approach to issues relating to integrity, probity in public and social life and his/her approach to various issues and conflicts faced by him/her while dealing with society. Case study approach may be utilized to inculcate these values and appropriate positive aptitude for in depth understanding.

Course specific Outcome:

Understanding and clarity of the human values, ethics and integrity in society and tackle various situations in life with positive attitude for common good.

UNIT- I

Ethics and Human Interface- Essence, determinants and consequences of Ethics in human actions; dimensions of ethics; ethics in private and public relationships.

Human Values - lessons from the lives and teachings of great leaders, reformers and administrators; role of family, society and educational institutions in inculcating values.

UNIT- II

Attitude- content, structure, function; its influence and relation with thought and behaviour; moral and political attitudes; social influence and persuasion.

Aptitude and foundational values for integrity, impartiality and non-partisanship, objectivity, dedication to social service, empathy, tolerance and compassion towards the weaker-sections.

UNIT- III

Emotional intelligence-concepts, and their utilities and application in academic, corporate sector and public services. Contributions of moral thinkers and philosophers from India and world. Values and Ethics in academic and pubic administration- Status and problems; ethical concerns and dilemmas in government and private institutions; laws, rules, regulations and conscience as sources of ethical guidance; strengthening of ethical and moral values in academic life.

UNIT- IV

Probity in Education and Research- Concept of Intellectual Property Rights (IPR) , Philosophical basis of teaching and probity; Information sharing and transparency in Education and Research, Codes of Ethics, Codes of Conduct, Citizen's Charters, Work culture, Quality of service delivery, challenges of corruption.

UNIT- V

Global Issues: Globalization and MNCs –Cross Culture Issues – Business Ethics – Media Ethics – Environmental Ethics – Endangering Lives – Bio Ethics – Computer Ethics – War Ethics – Research Ethics -Intellectual Property Rights.

REFERENCES

1. P.D.Sharma, Ethics, integrity and Aptitude, Rawat Publications, Jaipur.
2. G. Subba Rao and P.N. Roy Chowdhury, Ethics, Integrity and Aptitude, Access Publishing.
3. *Nanda Kishore Reddy and Santosh Ajmera*, Ethics, Integrity and Aptitude, Mcgraw Hill Education.
4. Professional Ethics by R. Subramaniam – Oxford Publications, New Delhi.
5. Ethics in Engineering by Mike W. Martin and Roland Schinzinger – Tata McGraw-Hill – 2003.
6. Professional Ethics and Morals by Prof.A.R.Aryasri, DharanikotaSuyodhana – Maruthi Publications.

7. Engineering Ethics by Harris, Pritchard, and Rabins, Cengage Specific, New Delhi.
8. Human Values & Professional Ethics by S. B. Gogate, Vikas Publishing House Pvt. Ltd., Noida.
9. Engineering Ethics & Human Values by M.Govindarajan, S.Natarajan and V.S.SenthilKumar-PHI Specific Pvt. Ltd – 2009.
10. Professional Ethics and Human Values by A. Alavudeen, R.Kalil Rahman and M. Jayakumaran – University Science Press.
11. Professional Ethics and Human Values by Prof.D.R.Kiran-Tata McGraw-Hill – 2013

BSTATMT-401 Master Thesis

(Credits 08)

BSTATIRA-401 DATA ANALYSIS USING STATA

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

1. To inculcate and develop aptitude to apply statistical tools at a number of data generating fields in real life problems.
2. To train students to handle large data sets and carry out data analysis using software and programming language.
3. To teach a wide range of statistical skills, including problem-solving, project work and presentation so as enable students to take prominent roles in a wide spectrum of employment and research.

Course Specific Outcomes:

On successful completion of the course a student will be able to:

1. Gain sound knowledge in theoretical and practical aspects of Statistics.
2. Describe complex statistical ideas to non-statisticians.
3. Handle and analyse large databases with computer skills and use their results and interpretations to make practical suggestions for improvement.

UNIT I

STATA Environment: STATA windows- command window and result window, Review window, variable window, data browser, data editor, do-file editor, viewer and Help. Files-log files, STATA data files; initial setup- memory allocation and setting setup file and other system parameters; functions, operators and expressions in STATA;

UNIT II

STATA command syntax; working with STATA: basic unit of data, loading and saving data, working with data, value and variable labels, some elementary commands-commands for loading or importing and saving data in main memory, commands related to data manipulation, commands related to tabulation, command related to combining data, commands for reshaping/ re-structuring the datasets and commands which having `replace` options. other important commands- sorting, dealing with the variables and observations, and dealing with missing data.

UNIT III

Role of log files, concept of immediate commands, other immediate handy options-sample size and power estimation. Exploratory data analysis. Frequencies analysis, cross tabulations, descriptive Statistics, Three-way crosstabs, creation and editing of basic graphs. formatting graphs. Advanced Graphs-Scatter plots, Histograms, Catplot, Bars etc.

UNIT IV

Analysis of continuous and binary outcomes, chi-squared test, t-test, one-way ANOVA, correlation, rank correlation, Simple linear regression, univariate and multivariate logistic regression analysis, some epidemiological tests using STATA, Multiple regression.

UNIT V

Preparing data for advance statistical analysis- dealing with dates, setting time and time series variables. Advanced analysis- analysis of longitudinal data in STATA, Survival Analysis in STATA, Principal component analysis and Time series analysis

REFERENCES

1. A handbook of statistical analysis using STATA. Sophia Rabe- Hesketh, Brain Everitt.
2. Categorical data analysis. Alan Agresti.
3. Data Analysis with STATA. Prasad Kothari.
4. Using STATA for Quantitative Analysis. Kyle C. Longest