COTTON UNIVERSITY

DEPARTMENT OF STATISTICS

Ph.D. PROGRAMME

(Courses effective from Academic year 2018-2019)

SYLLABUS OF COURSE WORK TO BE OFFERED

COTTON UNIVERSITY

DEPARTMENT OF STATISTICS

Ph. D. Course Work

Statistics Syllabus

DISTRIBUTION OF PAPERS

Paper Code	Paper Name	Credit (L+T+P)
STA 2501C	Statistical Techniques for Data Analysis	(4+0+0)
STA 2502 SP1	Multivariate Analysis	(4+0+0)
STA 2502 SP2	Regression Analysis	(4+0+0)
STA 2502 SP3	Economics and Social Statistics	(4+0+0)
STA 2502 SP4	Statistical Techniques in Population Studies	(4+0+0)
STA 2502 SP5	Biostatistics	(4+0+0)
STA 2502 SP6	Statistical techniques in Medical Research	(4+0+0)
STA 2502 SP7	Environmental Statistics	(4+0+0)
STA 2502 SP8	Stochastic Processes	(4+0+0)

Paper: STA 2501C

Statistical Techniques for Data Analysis

Credit: 4 (4+0+0)

UNIT I (Lectures: 16)

Sample survey: Definition of related terms with example from different fields of study, Principles of sample survey, Different steps in a large scale sample survey, size of sample, Judgement and random sampling with examples. Concept of different sampling techniques, namely, Simple random sampling with and without replacement, Stratified random sampling, Systematic sampling, Multistage sampling, Cluster sampling, Double sampling, Probability proportional to size sampling – with examples (derivation of different estimates are not required).

UNIT II (Lectures: 16)

Testing of Hypothesis: Parameter and statistic, Sampling distribution, Standard error, Concept of hypothesis, Type I and Type II errors, Level of significance, Applications of Chi-square, t and F statistics, Large sample tests and Confidence interval.

UNIT III (Lectures: 16)

Analysis of Variance: Concept of Analysis of Variance (ANOVA), Assumptions and effects of violation of the assumptions., Concept of Fixed effects and Mixed linear models, ANOVA for one way and two way classified data with m (> 1) observations per cell (for fixed effects only).

UNIT IV (Lectures: 16)

Nonparametric Tests: Introduction and concept, Test for randomness based on total number of runs, Empirical distribution function, Kolmogrov-Smirnov test for one and two samples, Sign test-one sample and two samples, Wilkoxon-Mann-Whitney test, Kruskal Wallis test, Wilkoxon signed rank test, mediun test, Spearman rank test (without derivation for all the tests)

SUGGESTED READING:

1. M. N. Das and N. C. Giri, Design and Analysis of Experiments, New Age International Ltd., 2008.
2. P. Mukhopadhaya, Applied Statistics, New Central Book Agency, 2014.
3. J. D. Gibbons, Non Parametric Statistical Inference, Marcel Dekker Inc., 1985.
4. A. M. Goon, M. K. Gupta and B. Dasgupta, Fundamental of Statistics, Vol. II, World Press, 1999.
5. S. C. Gupta and V. K. Kapoor, Fundamentals of Applied Statistics, S. Chand & Sons, 2018.

STA 2502 SP1
Multivariate Analysis
Credit: 4 (4+0+0)

UNIT I (Lectures: 16)

Classification of Data Analytic Methods: Dependence methods and Independence methods, Some Applications of Multivariate tests.

Multivariate Analysis of Variance (MANOVA): Two-Group MANOVA, Multiple-Group MANOVA.

UNIT II (Lectures: 16)

Principal Component Analysis (PCA): Analytical Approach, How to perform PCA, Issues relating to the use of PCA.

UNIT III (Lectures: 16)

Factor Analysis: Factor Analysis techniques, How to perform Factor Analysis, Exploratory versus Confirmatory Factor Analysis, Factor Analysis versus PCA.

UNIT IV (Lectures: 16)

Discriminant Analysis: Analytical approach to Discriminant Analysis, Regression approach to Discriminant Analysis, Stepwise Discriminant Analysis, External validation of the Discriminant function.

SUGGESTED READING:

1. T. W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley, 2003.
2. N. C. Giri, Multivariate Statistical Inference, Academic Press, 1977.
3. W. K. Hardle and L. Simar, Applied Multivariate Statistical Analysis, Pearson Education India, 2015.
4. R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis, Prentice Hall, 2015.
5. R. J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, 1982.
6. C. R. Rao, Linear Statistical Inference and its Application, Wiley Eastern Ltd., 1991.
7. M. S. Srivastava and C. G. Khatri, An introduction to Multivariate Statistics, North-Holland, 1979.
8. S. Sharma, Applied Multivariate Techniques, John Willey & Sons, 1996.

STA 2502 SP2
Regression Analysis
Credit: 4 (4+0+0)

UNIT I (Lectures: 20)

Multiple Linear Regression (MLR) model , Methodology for fitting models, Statistical inference , ANOVA , Diagnostic measures of model fit, Variable selection and model building, Testing in multiple linear regression, Regression without intercept terms, use of R^2 & R^2 adjusted, Residual Analysis.

UNIT II (Lectures: 12)

Nonlinear regression models, Estimation of parameters and Statistical Inferences in nonlinear regression, Interpretation of parameters relation with categorical data analysis.

UNIT III (Lectures: 12)

Generalized Linear model: introduction and as link functions to Poisson, binomial, inverse binomial, inverse Gaussian, gamma.

UNIT IV (Lectures: 20)

Logistic regression models: Logistic regression model for dichotomous data with single and multiple explanatory variables, ML estimation, large sample tests about parameters, Goodness-of-Fit tests (Concept of deviance), analysis of deviance, Lack-of-Fit tests in Logistic regression.

Log linear models for contingency tables: interpretation of parameters, ML estimation of parameters.

SUGGESTED READING:

1. S. Chatterjee and B. Price, Regression Analysis by Example, John Wiley and Sons, 1995.
2. N. R. Draper and H. Smith, Applied Regression Analysis, John Wiley, 1998.
3. J. Fox, Applied Regression Analysis & Generalized Linear Models, Sage, 2016.
4. D. W. Hosmer and S. Lemeshow, Applied Logistic Regression, John Wiley, 1989.
5. T. Z. Keith, Multiple Regression and Beyond, Routledge, 2019.
6. P. McCullagh and J. A. Nelder, Generalized Linear Models, Chapman & Hall, 1989.
7. D. C. Montgomery, A. E. Peck and G. G. Vining, Introduction to Linear Regression Analysis, Wiley, 2003.
8. M. Kunter, Neter, W. Wasserman, J. Neter and C. Nachtsheim, Applied Linear Statistical Models, McGraw-Hill Education, 2004.
9. D. A. Ratkowsky, Nonlinear Regression Modelling, Marcel Dekker, 1983.
10. G. E. F. Seber and C. J. Wild, Nonlinear Regression, John Wiley, 1989.

STA 2502 SP4
Statistical Techniques in Population Studies
Credit: 4 (4+0+0)

UNIT I (Lectures: 20)

Regression Analysis: Simple and multiple linear regression, Polynomial regression and orthogonal polynomials, Test of significance and confidence intervals for parameters. Residuals and their analysis for test of departure from the assumptions such as fitness of model, normality, homogeneity of variances, detection of outliers, Influential observations, Power transformation of dependent and independent variables. Logistic regression, multinomial logistic regression.

UNIT II (Lectures: 14)

Nonlinear Regression Analysis: Nonlinear regression, parameter estimation, confidence intervals and testing of hypothesis, transformations to linear forms.

UNIT III (Lectures: 16)

Multivariate Analysis: Multivariate normal distribution, assessing normality, Comparisons of several multivariate means, multivariate linear regression models; principal components, factor analysis; discriminant analysis. Multivariate Analysis of Variance (MANOVA)

UNIT IV (Lectures: 14)

Simultaneous equation systems: Introduction, identification problems of simultaneous equation models, different methods of solving these equations.

SUGGESTED READING:

1. T. W. Anderson, An Introduction to Multivariate Statistical Analysis, Wiley, 1984.
2. K. C. Bhuyan, Multivariate Analysis and its Apllications, New Central Book Agency Ltd., 2005.
3. B. L. Bowerman and R. T. O'Connell, Linear Statistical Models: An Applied Approach, PWS-KENT Pub., 1990.
4. N. R. Draper and H. Smith, Applied Regression Analysis, John Wiley and Sons (Asia) Pvt. Ltd., Series in Probability and Statistics, 2003.
5. R. Gnanadesikan, Methods for Statistical Data Analysis of Multivariate Observations, John Wiley, 1997.
6. G. M. K. Madnani, Introduction to Econometrics- Principles and Application, Oxford & IBH Publishing Co. Pvt. Ltd, 1996.

7. D. C. Montgomery, E. A. Peck and G. G. Vining, Introduction to Linear Regression Analysis, John Wiley, 2003.
8. A. A. Sen and M. Srivastava, Regression Analysis – Theory, Methods & Applications, Springer-Verlag, 1990.
9. G. A. F. Seber and C. J. Wild, Nonlinear Regression, John Wiley & Sons, 1989.
10. H. D. Vinod and A. Ullah, Recent advances in regression methods, Marcel Dekker, 1981.

STA 2502 SP5

Bio-Statistics

Credit: 4 (4+0+0)

UNIT I (Lectures: 20)

Introduction to survival data, time and event. Functions of survival time, survival distributions and their applications viz. Exponential, Gamma, Weibull, Rayleigh, Lognormal, death density function for a distribution having bath-tub shape hazard function. Censored data: different types of censoring viz. right (type I, type II and type III), left and interval censoring with real life examples. Estimation of mean survival time and variance of the estimator for type I and type II censored data with numerical examples. Non-parametric methods for estimating survival function and variance of the estimator and confidence intervals viz. Acturial and Kaplan–Meier methods. Methods for comparing two survival distributions: parametric methods viz. Likelihood Ratio test, Cox's F-test and non parametric methods viz. Log Rank test, Wilcoxon test. Semi parametric model for survival data: Cox Proprtional hazard model.

UNIT II (Lectures: 20)

Analysis of Epidemiologic and Clinical Data: Studying association between a disease and a

characteristic: (a) Types of studies in Epidemiology and Clinical Research (i) Prospective study (ii)Retrospective study (iii) Cross-sectional data, (b) Dichotomous Response and Dichotomous Risk Factor: 2X2 Tables (c) Expressing relationship between a risk factor and a disease (d) Inference for relative risk and odds ratio for 2X2 table, Sensitivity, specificity and predictivities.Competing risk theory: Indices for measure-ment of probability of death under competing risks and their inter-relations. Estimation of probabilities of death under competing risks by maximum likelihood and modified minimum Chi-square methods. Theory of independent and dependent risks. Bivariate normal dependent risk model. Conditional death density functions.

UNIT III (Lectures: 14)

Stochastic epidemic models: Simple and general epidemic models (by use of random variable technique). Basic biological concepts in genetics, Mendels law, Hardy- Weinberg equilibirium, random mating, distribution of allele frequency (dominant/co-dominant cases), Approach to

equilibirium for X-linked genes, natural selection, mutation, genetic drift, equilibirium when both natural selection and mutation are operative, detection and estimation of linkage in heredity.

UNIT IV (Lectures: 10)

Planning and design of clinical trials, Phase I, II, and III trials. Consideration in planning a clinical trial, designs for comparative trials. Sample size determination in fixed sample designs.

SUGGESTED READING:

1. S. Biswas, Applied Stochastic Processes. A Biostatistical and Population Oriented Approach, Wiley Eastern Ltd, 1995.
2. D. Collett, Modelling Survival Data in Medical Research, Chapman & Hall/CRC, 2003.
3. D. R. Cox and D. Oakes, Analysis of Survival Data, Chapman and Hall, 1984.
4. R. C. E. Johnson, Probability Models and Statistical Methods in Genetics, John Wiley & Sons, 1971.
5. W. J. Ewens, Mathematics of Population Genetics, Springer Verlag, 1979.
6. W. J. Ewens and G. R. Grant, Statistical methods in Bio informatics: An Introduction, Springer, 2001.
7. L. M. Friedman, C. Furburg and D. L. DeMets, Fundamentals of Clinical Trials, Springer Verlag, 1998.
8. A. J. Gross and V. A. Clark, V. Survival Distribution: Reliability Applications in Biomedical Sciences, John Wiley & Sons, 1975.
9. A. Indrayan, Medical Biostatistics, Second Edition, Chapman & Hall/CRC, 2008.
10. E. T. Lee, Statistical Methods for Survival Data Analysis, John Wiley & Sons, 1992.
11. C. C. Li, First Course of Population Genetics, Boxwood Press, 1976.

STA 2502 SP6
Statistical methods in Medical Research
Credit: 4 (4+0+0)

UNIT I (Lectures: 16)
The role of cohort studies in Cancer Epidemiology: Historical perspectives, present significance and specific strengths of cohort studies, limitations of cohort studies, Implementation and interpretation of cohort studies.

UNIT II (Lectures: 18)
Rates and rate standardization: Calculation of age specific and calendar specific rates, interpretations of different rates, comparative and proportional measures of incidence and mortality, comparison of standardized mortality ratios, comparisons of internally standardized mortality ratios, proportional and dose-response analyses.

UNIT III (Lectures: 16)
Fitting of models to cancer data: Additive and multiplicative models for rates, proportional mortality analyses, general models of relative risk, fitting relative and excess risk models to grouped data, fundamentals of continuous data analysis, likelihood inference, non-parametric estimation.

UNIT IV (Lectures: 12)
Multivariate analysis: Multivariate linear regression models, principal component, factor analysis, correspondence analysis.

SUGGESTED READING:

1. K. C. Bhuyan, Multivariate Analysis and its Applications, New Central Book Agency Ltd., 2005.
2. N. E. Breslow and N. E. Day, Statistical methods in Cancer Research, International Agency for Research in Cancer (IARC) Scientific Publications No 82, 1987.
3. R. Gnanadesikan, Methods for Statistical Data Analysis of Multivariate Observations, John Wiley, 1997.

4. M. Greenacre, Theory and Application of Correspondence Analysis, Academic Press Inc. Limited, London, 1993.
5. M. Greenacre, Corresponding Analysis in Practice (3rd Edition), CRC Press, Chapman & Hall, 2017
6. R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis, Pearson Prentice Hall, 2007.

STA 2502 SP7
Environmental Statistics
Credit: 4 (4+0+0)

UNIT I (Lectures: 10)
The concept of environmental statistics, Objective of environment statistics, Scope of environment statistics, Sources of environment statistics.

UNIT I (Lectures: 18)
Frequency Analysis of Extreme Events: Order Statistics- Definitions and distributions, Functions of order statistics, expected value and variance of order statistics; Extreme Value Distributions-Basic concepts of extreme value theory, Gumbel distribution, Fr´echet distribution, Weibull distribution as an extreme value model, Generalized extreme value distribution, Generalized Pareto distribution; Practical study and empirical models.

UNIT I (Lectures: 16)
Time-Series Methods: Space and Time Effects, Estimating and eliminating trend, periodicities, Stationary Time series, Time-Domain Models and Methods, Frequency-Domain Models and Methods, Point Processes

UNIT I (Lectures: 18)
Principal Component Analysis (PCA) and Factor Analysis (FA): Conditioning the data for PCA and FA; Scree plot, The biplot, Mapping of principal components; Choice of factor analysis method-Choice of rotation method, Number of factors extracted, Selection of elements for factor analysis, graphical representation of the results of factor analysis.

SUGGESTED READING:

1. P. U. Berthouex, Statistics for Environmental Engineers, Lewis Publ., 1994.
2. R. Wayne, Environmental Statistics and Data Analysis, CRC Press. 1995.
3. C. Reimann, P. Filzmoser, R. G. Robert and R. Dutter, R., Statistical Data Analysis Explained, John Wiley & Sons Ltd, 2008.
4. N. T. Kottegoda and R. Rosso, Applied Statistics for Civil and Environmental Engineers, Wiley-Blackwell, 2008.
5. V. Barnett, Environmental Statistics: Methods and Applications, John Wiley & Sons, Ltd, 2004.
6. B. F. J. Manly, Statistics for Environmental Science and Management, Chapman and Hall/CRC, 2000.