NORTH-EASTERN HILL UNIVERSITY, SHILLONG

Syllabus for M.Sc. Statistics under NEP 2020

Paper Code	Paper	Total Credit
FIRST SEMESTER		
STA-CC-500	Probability Theory & Stochastic Processes	4
STA-CC 501	Distribution Theory & Matrix Analysis	4
STA-DSEC-502	Statistical Computing using R	4
STA-DSEC-503	Practical – I	4
STA-GEC-504	Statistical Analysis of Data for Natural Sciences	4
	Total	**20**
SECOND SEMESTER		
STA-CC-505	Sample Survey and Estimation theory	4
STA-CC-506	Linear Models & Design of Experiment	4
STA-DSEC-507	Statistics for Sustainable Development	4
STA-DSEC-508	Applied Multivariate Analysis	4
STA-RM-509	Research Methodology & Proposal Writing	4
STA-SEC-510	Practical - II	4
	Total	**24**
THIRD SEMESTER		
STA-CC-600	Statistical Inference	4
STA-CC-601	Regression Analysis	4
STA-CC-602	Time Series Analysis and Forecasting	4
STA-DSEC-603 **STA-DSEC-604**	1. Demography and Population Dynamics/Operations Research/MOOCs Courses 2. Reliability Theory/Econometrics/MOOCs Courses	4+4=8
STA-SEC-605	Practical - III	4
	Total	**24**
FOURTH SEMESTER		
STA-DSEC-606	Planning, Execution and Dissertation	16
	Presentation and Viva-Voce	4
	Total	**20**

Abbreviations: ***CC-Core Course, DSEC-Discipline Specific Elective Course, SEC-Skill Enhancement Course, RM-Research Methodology, RP-Research Project.***

Note:

a) One Credit Course is equivalent to 15 hours of teaching for theory papers and 30 hours of practical activities for Practical Papers.

b) Up to 40% of the total credits of DSEC courses can be achieve through MOOCS (SWAYAM) platform.

c) 25% of all courses will be Internal

d) Paper no. STA-RP-606 shall preferably be based on Field Study

e) The duration of examination for a 4 Credit Theory course is of 3 hours and 4 Credit Practical Course is of 5 hours with 15 marks for Viva-voce.

f) The Course 'Statistical Computing using R' and Practical papers will introduce Statistical packages like Minitab, SPSS, SYSTAT etc.

Objectives of the Programme:

1. To inculcate and develop aptitude towards applying statistical tools at various data generating fields related to real life problems.
2. To train students to handle large data sets and carry out data analysis and decision making using softwares and programming languages.
3. To teach a wide range of statistical skills, including problem-solving, project work, field study and presentation so as to enable students to take prominent role in a wide spectrum of scientific research, planning process and employment.

Outcomes of the Programme:

On successful completion of the course a student is expected to:

1. Gain sound knowledge in theoretical and practical aspects of Statistical science.
2. Describe and make understand 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 improved decision making.
4. Get wide range of job opportunities in industries as well as in government sectors.

I SEMESTER

STA-CC-500

Probability Theory & Stochastic Processes

Full Marks: 100 **4 Credits**

Learning objectives: The course aims to provide students with a comprehensive understanding of probability theory and an introduction to Stochastic Processes, their properties, and applications.

Course outcomes: After completion of course, students are expected to develop a comprehensive understanding of probability theory and the basic understanding of stochastic processes. Learners will be able to apply the tools to various real life situations, and will possess the ability to interpret and communicate the results effectively.

Unit 1 Classes of sets, fields, sigma-fields, minimal sigma-field, Borel sigma -field in R^k, sequence of sets and their limits. Measure, Probability measure and its properties, Caratheodory extension theorem (statement only).Measurable functions as limit of simple functions, Random variables, sequence of random variables.

(15 Contact hours)

Unit 2 Convergence of sequence of random variables**,** almost sure convergence, convergence in probability (and in measure), convergence in distribution. Integration of a measurable function with respect to a measure. Expectation and moments. (Statements of) Monotone convergence theorem, Fatou's lemma, and Dominated convergence theorem (and discussion).

(15 Contact hours)

Unit 3 Probability inequalities (Markov, Tchebyshef, Jensen), WLLN. Independence, Borel- Cantelli Lemma, Kolmogorov zero-one law and Borel zero-one law. Kolmogorov's Strong Law of Large numbers for iid sequences. Polya's theorem and Levy's continuity theorem (statement only), De Moivre-Laplace Central Limit Theorem (CLT), Lindeberg-Levy's CLT, Liapounov CLT and applications.

(15 Contact hours)

Unit 4 Introduction to stochastic processes; classification of stochastic processes according to state space and time domain, Countable state Markov chains (MCs), Chapman-Kolmogorov and spectral decomposition theorems, calculation of n-step transition probability and its limit, classification of states; transient and persistent MC, stationary distribution, random walk and gambler's ruin problem; applications in social, biological and physical sciences. Poisson process, birth and death process; applications to queues and storage problems.

(15 Contact hours)

Text Books

1. Parthasarthy, K. R. (1980). Introduction to Probability Theory and Measure. McMillan, India.
2. Ross, S.M. (2005). A First Course in Probability, 7th Edition, Prentice Hall, New Jersey.
3. Ross, S. M. and Erol, A. P. (2007): A Second Course in Probability, www. Probability Book Store. Com, Boston, USA.
4. Singh, B. M. (2002). Measure, Probability and Stochastic Processes, South Asian Publishers, New Delhi.
5. Ross, S. M. (2001), "Probability Models for Computer science", Academic Press.
6. Dudley, R. M. (2002), "Real Analysis and Probability", 2ND Ed., Cambridge.
7. Robert, B. A., Birnbaum, Z. W. and Lucacs, E. (1972), "Real Analysis and Probability", Academic Press.
8. Adke, S. R. and Manjunath, S. M. (1984): An Introduction to Finite Markov Processes, Wiley Eastern.
9. Feller, W. (1968): Introduction to Probability and its Applications, Vol.1, Wiley Eastern, New Delhi
10. Medhi, J. (2013), "Stochastic Processes", 3RD Ed., Wiley Eastern.
11. Minh, D.L. (2000) Applied Probability Models, Duxbury Press.
12. Ross, S.M. (1996). Stochastic Processes, John Wiley and Sons.
13. Ross, S.M. (2007): Introduction to Probability Models (IXth edition). Elsevier, USA.

Additional References

1. Billingsley, P. (1986). Probability and Measure. John Wiley & Sons, New York.
2. Feller, W. (1985). Introduction to Probability theory and its Applications. (Vol. 1&2). Wiley Eastern, New Delhi.
3. Kingman, J F C and Taylor, S. J. (1966). Introduction to Measure and Probability. Cambridge University Press.
4. Natarajan, A. M. and Tamilarasi. A. (2003). Probability, Random Process and Queuing Theory, New Age International Publishers,
5. New Delhi.
6. Rao, C. R. (1995). Linear Statistical Inference and its Applications. Wiley Eastern, New Delhi.
7. Guttorp, P.(1995) Stochastic Modelling for Scientific Data, Springer.
8. Karlin, S. and Taylor, H. M. (1975): A First Course in Stochastic Processes, Vol.1, Academic Press.
9. Parzen, E. (1962): Stochastic Processes, Holden- Day. San - Francisco

STA-DSEC-501

Distribution Theory & Matrix Analysis

Full Marks: 100 **4 Credits**

Learning objectives: The course aims to provide students with a comprehensive understanding of continuous probability distributions, their properties, and applications starting from random variables, marginal probability, joint probability and conditional probability. Bivariate probability distributions and their applications also have been discussed. The concept of sampling distributions with their characteristics included here. Order statistics which is very useful concept in statistical sciences and have a wide range of applications in distribution is also included here. Moreover, some fundamental concepts of matrix algebra, to solve systems of linear equations and their applications in statistics will be discussed here.

Course outcomes: By course completion, students should have a comprehensive understanding of probability distributions along with their characteristics and will be able to apply them to various situations, and possess the skills to interpret and communicate the results effectively. They will have an in-depth grasp of the fundamentals, characteristics, and practical applications of random variables. They should confidently apply sampling distribution theory in practical scenarios. After completing this course, students will demonstrate an understanding of fundamental concepts of matrix algebra, application of matrix algebra in Statistics

Unit 1 Random variable, pmf, pdf**,** Joint, marginal and conditional distributions. Computations of probability, expectations and variances by conditioning, generating functions (m.g.f and p.g.f) of random variables, Bivariate and Multivariate p.g.f, their properties and applications. Some continuous distributions (Exponential, Gamma, Beta, Cauchy, Pareto, Weibull, lognormal), Bivariate normal and bivariate exponential distributions and their properties, multinomial distribution.

(15 Contact hours)

Unit 2 Functions of random variables and their distributions using Jacobian and other tools, convolution and compound distributions; idea of truncated and mixture distributions; concept of weighted distribution and length biased distributions; sampling distributions from normal population, central and non - central Chi-square, t- and F-distributions.

(15 Contact hours)

Unit 3 Order statistics and their distributions and properties. Joint, marginal and conditional distributions of order statistics. Pdf of sample median, range, quasi range, moments and recurrence relations, modal equations.

(15 Contact hours)

Unit 4 Non-singular matrices and their inversion, ranks, row and column rank of a matrix, partitioned matrices, G - inverse. Systems of homogeneous and non-homogeneous linear equations, their consistency and maximal linearly independent solutions, minimal and characteristic polynomials of a square matrix, Characteristic roots and vectors and their extraction, similarity and diagonalization of square matrices. Real quadratic forms and their value classes, canonical reductions and simultaneous reducibility of quadratic forms.

(15 Contact hours)

Text Books

1. Hogg, R.V. and Craig, A.L. (1978). Introduction to Mathematical Statistics, McMillan, New York.
2. Mood, A.M., Grabyl, F.A. and Boes, D.C. (2001). Introduction to Theory of Statistics, Tata McGraw Hill, New Delhi.
3. Ross, S. M. (2004) Introduction to Probability and Statistics for Engineers and Scientist, Third Edition, Elsevier Academic Press, USA.
4. Rohatgi V.K. and Saleh, A.K.Md. E. (2001). An Introduction to Probability and Statistics (Second Edition), John Wiley and Sons (Asia), Singapore.
5. Lay, David C. (1997). Linear Algebra and its Applications, Addison-Wesley,
6. Singh, B.M. (2008). Introductory Linear Algebra, South Asian Publishers Pvt. Ltd., New Delhi.
7. Searle, S.R. (1982). Matrix Algebra useful for Statistics, John Wiley & Sons, New York.
8. Kolman, B. and Hill, D. R. (2010). Elementary Linear Algebra with Applications, 10^{TH} Ed., Pearson.

Additional References

1. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley, New York.
2. Johnson, S and Kotz, S. (1970). Continuous univariate Distributions I and II John Wiley, New York.
3. Pitman, J. (1993). Probability, Narosa Publishing House, New Delhi.
4. David, W.S.(2003). Order Statistics. (Second Edition). John Wiley and Sons, New York.
5. Ferguson, T.S. (1996). A Course on Large Sample Theory. Chapman and Hall, London.
6. Johnson, N.L. and Kotz, S.(1970). Continuous Univariate Distributions – 2, John Wiley and Sons, New York.
7. Gentle, James E. (2005). Matrix Algebra: Theory, Computations and Applications in Statistics, Springer Text in Statistics, Springer-Verlag, New York.
8. Banerjee, S. and Roy, A.(2014). Linear Algebra and Matrix Analysis, CRC
9. Rao, A.R. and Bhimasankaram, P. (1992). Linear Algebra, Tata McGraw-Hill, New Delhi

STA-DSEC - 502

Statistical Computing in R

Full Marks: 100 **4 Credits**

Learning Objectives: This course aims to instill the proficiency in R programming language and to provide with a comprehensive understanding of statistical computing using R which is a powerful and flexible programming language for data analysis. Through hands-on experience, students will learn to navigate the R environment, manage data structures, and perform essential data manipulation and visualization tasks. The course covers fundamental programming concepts such as loops, functions, and flow control, along with advanced topics like statistical modeling, probability distributions, and simulations.

Course Outcomes: Upon completing this course, students will have developed the skills to proficiently use R for statistical computing and data analysis. They will be able to efficiently handle various data types, import and manipulate data, and create detailed visualizations using R's graphical capabilities. Students will gain a solid understanding of programming concepts, enabling them to write custom functions and control the logic and flow of their analysis. Additionally, they will be able to implement statistical models, conduct simulations, and apply inferential techniques to analyze real-world datasets. Overall, students will emerge with the practical knowledge and confidence to use R for statistical analysis and decision-making in academic and professional settings.

Unit 1 Introduction to R environment and R Studio, using the help facility. R as a calculator, data types and data structures, simple manipulations of data, vectors and vector arithmetic, objects, their modes and attributes, ordered and unordered factors, arrays and matrices, lists and data frames.

(15 Contact hours)

Unit 2 Reading data into R from various sources, merging data files. Basics of R syntax, sub setting, loops and conditional execution, writing R functions, logic and flow control, iterations.

(15 Contact hours)

Unit 3 R graphics: low and high level, lattice and ggplot2, descriptive statistics and tables. Built-in functions, exploratory analysis, probability distributions and simulations

(15 Contact hours)

Unit 4 Linear and generalized linear models: lm and glm functions, statistical inference, contingency tables, chi-square goodness of fit, least squares, maximum likelihood, non-linear optimization, resampling.

(15 Contact hours)

Text Books

1. Dalgaard, P. (2000). Introductory Statistics with R, Springer.
2. Dennis, B. (2013). The R Student Companion, Taylor and Francis.
3. Crawley, M J. (2013). The R Book, John Wiley and Sons.

Additional references

1. Chambers, J. (2008). Software for Data Analysis: Programming with R, Springer.
2. Jones, O., Maillardet, R. and Robinson, A.(2009). Introduction to Scientific Programming and Simulation using R, CRC.
3. Everitt, B. and Hothorn, T.(2006). A Handbook of Statistical Analyses Using R, CRC.

STA-DSEC-503

Practical-I

Full Marks: 100 **4 Credits**

Learning Objectives: Practical applications of several continuous distributions and matrix algebra will be discussed in this paper.

Course outcomes: From this paper students will be able to focus on real life situations of distribution theory and matrix algebra using software R.

Problems based on following topics:

Section A: Distribution Theory

- Stem and Leaf, Box and Whisker's plots
- Empirical Distribution plots
- Fitting of some standard distributions; goodness of fit test, p-p and q-q plots
- Plotting of density and distribution functions for exponential family with varying location, scale, and shape parameters
- Sample generation from chi-square, t, F and lognormal distributions using standard normal variates; comparison of histogram of the generated data and the corresponding density plot
- Fitting of Pareto distribution
- Fitting of Weibull distribution
- Fitting of lognormal distribution
- Multiple and partial correlation

(60 Contact hours)

Section B: Matrics Algebra

- Determinant: pivotal consideration method
- Solutions of a system of linear equations: Gauss elimination method
- Matrix inversion – Gauss Jordon elimination method
- Computation of G – inverse
- Characteristic roots and vectors by power method/singular value decomposition
- Triangular reduction of a positive definite matrix
- QR decomposition of a non-singular matrix
- Spectral decomposition of a real symmetric matrix
- Canonical reduction of quadratic forms Solutions of a system of linear equations and least squares
- Matrix inversion and computation of G - inverse
- Eigen analysis of matrices
- Spectral decomposition of a real symmetric matrix
- Solution to non-linear equation Newton-Raplson/Steepest descent

(60 Contact hours)

Text Books

1. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley, New York.
2. Goon, A.M., Gupta, M.K. and Disrupt, B. (2000). Fundamentals of Statistics, Vol. I, World Press, Collate.
3. Hogg, R.V. and Tanis, E.A. (2003). Probability and Statistical Inference, Pearson Education, Delhi.
4. Rohatgi V.K. and Saleh, A.K.Md. E. (2001). An Introduction to Probability and Statistics (Second Edition), John Wiley and

1. Sons (Asia), Singapore.

5. Rao, A. R. and Bhimasankaran, P. (1992), Linear Algebra, Tata McGraw Hill, New Delhi.
6. Searle, S. R.(1982). Matrix Algebra useful for Statistics, John Wiley & Sons, Inc., New York.
7. Thisted, R. A. (1988). Elements of Statistical Computing. Chapman and Hall.
8. Dalgaard, P. (2000). Introductory Statistics with R, Springer.
9. Dennis, B. (2013). The R Student Companion, Taylor and Francis Golub Gene H. and Loan C.F. Van (1996). Matrix

2. Computations (John Hopkins Studies in Mathematical Science) 3rd Edition; John Hopkins University Press, USA.

10. Fieller, N.(2016). Basics of Matrix Algebra for Statistics with R, CRC.
11. Gentle, James E (2005). Matrix Algebra: Theory, Computations and Applications in Statistics, Springer, New York.
3. 12 Givens, G. H. and Hoefting, J. A. (2005). Computational Statistics, John Wiley & Sons, New York.

Reference Books

1. Wasserman, L. (2004). All of Statistics: A Concise Course in Statistical Inference, Springer Science Business Media, Inc.,

1. New York.

2. Golub Gene H. and Loan C.F. Van (1996). Matrix Computations (John Hopkins Studies in Mathematical Science) 3rd Edition;
3. John Hopkins University Press, USA.

STA-GEC- 504

Statistical Analysis of Data for Natural Sciences

Full Marks: 100 **4 Credits**

Learning objectives: The objective of this module is to introduce the basics of statistics, information about variables in a data set and emphasis the potential relationship between variables and descriptive statistical techniques that would help the students in understanding the importance and need of statistics. It would help them in understanding the concepts involved in data presentation, analysis, and interpretation using suitable statistical data and measures.

Course outcomes: On completion of the module, students will be able to define the basic concepts and scope of statistics, explain the descriptive statistical techniques and its scope in describing the real world problems, present data in a meaningful manner after the proper arrangement using Spreadsheet Package, interpret data using appropriate statistical measures using statistical functions.

Unit 1 Data types and scales of measurements, variables and their descriptive measures, bivariate and multivariate data, Correlation and regression, Spearman's Rank Correlation.

(15 Contact hours)

Unit 2 Probability and probability models: binomial, Poisson, geometric, uniform and normal distributions, their properties and applications.

(15 Contact hours)

Unit 3 Population and sample, types of sampling, parameter and statistic, sampling distribution of a statistic, point and interval estimation, tests of hypothesis, large sample tests.

(15 Contact hours)

Unit 4 Sampling distributions chi-square, t and F and their applications, test of association, goodness of fit, equality of means, variance ratio, analysis of variance for one and two factors with applications.

(15 Contact hours)

Text Books

1. Bevan, A. (2013). Statistical Data Analysis for Physical Sciences, Cambridge University Press, Cambridge.
2. McCleery,R.H., Watt, T.A., and Hart, T. (2007), Introduction to Statistics for Biology, Routledge, India.
3. Heumann, C., Schomaker, M., and Shalabh (2016). Introduction to Statistics and Data Analysis, Springer nature, Switzerland, AG.
4. Ott, L. and Longnecker, M. (2015). An Introduction to Statistical Methods and Data Analysis, 7th Edition, Cengage Learning, India.

II SEMESTER

STA-CC 505

Sample Survey and Estimation Theory

Full Marks: 100 **4 Credits**

Learning Objectives: To introduce students to the concepts and principles of random sample based inference on characteristics of finite and infinite populations. To make them understand equal and unequal sampling methods and to extend these basic sampling methods towards handling complex survey designs in finite population. To provide students with the knowledge and skills to design sample surveys effectively and efficiently. To help students understand model based estimation theory and methods.

Course Outcomes: After successful competition of the course, the learners are expected to be skilled in taking decision on the population characteristics of infinite and finite populations. The learners may be able to undertake large scale design based sample surveys as well be able to apply knowledge on model based sampling in planning and efficient handling of day to day situations involving uncertainty.

Unit 1 Simple random sampling WR & SRSWOR - estimation based on distinct units in SRSWR. Systematic sampling (linear, circular, population with linear trend), domain estimation in SRS. Ratio and regression estimators based on double sampling and SRS. Stratified sampling, allocation problem and construction of strata; Cluster and Multi-stage sampling under stratification.

(15 Contact hours)

Unit 2 Unequal probability sampling; pps wr and wor methods (including Lahiri's scheme) and related estimators of a finite population mean. Hansen – Hurwitz and Desraj estimators for a general sample size and Murthy's estimator for a sample of size 2. Horvitz – Thompson Estimator (HTE).

(15 Contact hours)

Unit 3 Point Estimation, joint distribution of a sample and induced sampling distribution of a statistic; examples form standard discrete and continuous models. Likelihood function, sufficiency, Neyman factorizability criterion, likelihood equivalence, minimal sufficient statistic, exponential family, invariance property of sufficiency, Fisher information for one and several parameters model. Interval estimation; confidence level.

(15 Contact hours)

Unit 4 Methods of estimation: maximum likelihood method, method of moments, method of minimum chi-square, method of scoring; choice of estimators based on unbiasedness, minimum variance, mean squared error, minimum variance unbiased estimators.

(15 Contact hours)

Text Books

1. Cochran, W.G. (1997). Sampling Techniques, Wiley Eastern, New Delhi.
2. Mukhopadhyay, P. (1998). Theory and Methods of Survey Sampling, Prentice Hall of India, New Delhi.
3. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). Sampling Theory of Surveys with Applications, Iowa State University Press and IARS.
4. Casella. G and Berger R.L. (1990) Statistical Inference, Wordsworth and Brooks, California.
1. 5.Hogg, R.V. and Craig, A.T. (2002). Introduction to Mathematical statistics, Pearson Education, Delhi.
5. Kale, B.K. (1999). A First Course on Parametric Inferences, Narosa Publishing House, New Delhi.
6. Rohatgi V. (1998). An Introduction to Probability and Mathematical Statistics. Wiley Eastern.

Additional References

.

1. Chaudhury, A. and Mukerjee, R. (1988). Randomized Response: Theory and Techniques, Marcel Decker, New York.
2. Murthy, M.N. (1977). Sampling Theory and Methods, Statistical Publishing Society, Kolkata.
3. Raj, D. and Chandhok, P. (1998). Sampling Theory. Narosa Publishing House, New Delhi.
4. Mukhopadhyay, Nitis. (2000). Probability and statistical inference, CRC Press.
5. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley, New York.
6. Lehman, E.L. (1986). Testing of Hypothesis, John Wiley, Singapore.
7. Lehman, E.L. (1996). Theory of Point Estimation, John Wiley, Singapore
8. Rao, C.R. (1995), Linear Statistical Inference, Wiley Eastern, New Delhi.

STA-CC-506

Linear Models & Design of Experiment

Full Marks: 100 **4 Credits**

Learning Objectives: This module is meant to expose the students to the basic principles of design of experiments. The students would be provided with mathematical background of various basic designs involving one-way and two-way elimination of heterogeneity and their properties. This module also prepares students for undertaking research in this area so that it helps them in applications of this important subject to other sciences. The module will be delivered through the use of R and Excel.

Course outcomes: On completion of the module, students will be able to apply basic principles of design of experiments to get optimal information from scarce resources, apply completely randomized design (CRD), randomized complete block design (RCBD) and Latin square design (LSD) for treatment comparison, analyze the uses of mutually orthogonal Latin squares and Graeco Latin squares, construct balanced incomplete block (BIB) designs using various methods of construction, impute the values using missing plot techniques in analysis of experimental designs and apply multiple comparison procedures to find the differences between treatments.

Unit 1 Gauss-Markov setup, full rank linear model, estimability of parameters, least-squares (LS) estimators, Gauss-Markov theorem, variance and covariance of LS estimators, estimation of error variance, correlated observations, LS estimations with restrictions on parameters, simultaneous estimates of linear parametric functions. Tests of hypotheses, linear models with restricted hypothesis, confidence intervals.

(15 Contact hours)

Unit 2 Introduction to random effects linear models and estimation of variance components, mixed effects models and ANOVA. Concept of various experimental designs, basic principles of experimental designs, CRD, RBD and LSD and their analyses and uses, multiple range tests.

(15 Contact hours)

Unit 3 General factorial experiments - main and interaction effects, factorial experiments in CRD and RBD; principles of confounding and partial confounding for symmetric factorials, introduction to fractional factorial, split plot and split block designs.

(15 Contact hours)

Unit 4 Block designs, general properties, BIB designs with constructions and analysis; connectedness and balancing, intra block analysis, recovery of inter block information, Split plot and strip plot design and their analysis.

(15 Contact hours)

Text Books

1. Rencher, A.C., Schaalje, B. G. (2007). Linear Models in Statistics, 2nd Edition, Wiley.
2. .Zimmerman, D. L. (2020). Linear model theory: with examples and exercises, Springer
3. Searle, S. R. (1971). Linear Models. Wiley, New York.
4. Montgomery, C.D. (2013). Design and Analysis of Experiments, John Wiley, New York.
5. Cochran, W.G. and Cox, G.M. (1959). Experimental Designs, Asia Publishing House, Singapore.
6. Das, M.N. and Giri, N. (1979). Design and Analysis of Experiments, Wiley Eastern, New Delhi.
7. Giri, N. (1986). Analysis of Variance, South Asian Publishers.
8. Joshi, D.D. (1987). Linear Estimation and Design of Experiments, Wiley Eastern, New Delhi.

Additional References

1. Chatterjee, S, and Hadi, Ali S. (2013). Regression Analysis by Example, John Wiley. New York.
2. Dean, Angela and Voss, Daniel (1999). Design and Analysis of Experiments, Springer-Verlag, New York.
3. Dey, Aloke (1986). Theory of Block Designs, Wiley Eastern, New Delhi.
4. Pearce, S.C. (1984). Design of Experiments, John Wiley, New York.
5. Searle, S.R. Casella, G. and McGulloch, C.E. (1992). Variance Components, John Wiley, New York.
6. Dey, Aloke (2010). Incomplete Block Designs, World Scientific, Singapore.
7. Raghavarao, D. and Laxmi, P. (2005) “Block Designs”, World Scientific, Singapore.

STA-DSEC-507

Statistics for Sustainable Development

Full Marks: 100 **4 Credits**

Learning Objectives: This course aims to equip students with the statistical tools and knowledge necessary to understand and contribute to sustainable development. It explores the role of data in assessing and achieving sustainable goals, particularly focusing on how statistical methods can drive evidence-based policy making and monitor progress toward global initiatives like the United Nations' Sustainable Development Goals (SDGs). Students will delve into various sustainability concepts, human development metrics, and multidimensional poverty indices, learning how to interpret and evaluate these in the context of social equity, environmental protection, and economic growth. Through theoretical exploration and case studies, the course will foster a practical understanding of how data is used to measure progress and address gaps in achieving sustainable development at both national and global levels.

Course outcomes:

Upon successful completion of this course, students will have a comprehensive understanding of the key statistical concepts and methodologies used in sustainable development analysis. They will be able to critically assess and interpret sustainability metrics such as the Human Development Index (HDI), the Multidimensional Poverty Index (MPI), and various SDG indicators. Students will be adept at using data to inform evidence-based policymaking, addressing data gaps, and monitoring progress toward the United Nations' Sustainable Development Goals (SDGs). Additionally, they will gain the ability to analyze case studies, conduct disaggregated assessments of poverty and inequality, and evaluate the role of national and international statistical systems in supporting sustainable development efforts. Overall, students will be equipped with the skills to contribute to sustainability discussions and apply statistical tools to real-world policy challenges at local, national, and global levels.

Unit 1 Economic Growth and Sustainable Development, alternative definitions and concepts of sustainability; Social Equity and Environmental protection. UN adopted Sustainable Development Goals (SDGs), SDG indicators and Data Sources.

(15 Contact hours)

Unit 2 Role of data and statistics in evidence based policy making, National and International Statistical systems – their roles for capacity building, existing data gaps for goals, monitoring and evaluation of SDGs and targets, development and implementation of the SDGs from a range of theoretical, policy and practical perspectives. Country level adoption within an international, domestic and global context.

(15 Contact hours)

Unit 3 Concept of Human Development, human development index (HDI) and its dimensions and compositions, HDI and its link with SDGs, Physical Quality of Life Index, Gender- equity Index,

(15 Contact hours)

UNIT 4 SDG1 and poverty, global Multidimensional Poverty Index (MPI) of UNDP, Composition of MPI and deprivation indicators, Measurement and trends in MPI, disaggregated analysis of MPI and assessment of progress towards SDG1 – examples and case studies.

(15 Contact hours)

Text Books

1. Jennifer A. Elliot, (2013). An Introduction to Sustainable Development, 4th Ed, Routledge, New York.
2. Willies, K., (2011). Theories and Practices of Development, 2nd ed., Routledge, New York.
3. Hartmut, B., (1999). Indicators for sustainable development: Theory, Method and applications, IISD, Canada.
4. UNDP & OPHI(2020). Global Multidimesional Poverty Index 2020: Charting Pathways out of Multidimensional Poverty : Achieving the SDGs, https://ophi.org.uk/multidimensional-poverty-index/.

Additional References

1. United Nations(2015). Transforming Our World: The 2030 Agenda for Sustainable Development, https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981.
2. United Nations (2020). The Sustainable Development Goals Report 2020. https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.
3. General Economics Division (GED), Bangladesh Planning Commission (2017), Data Gap Analysis of Sustainable Development Goals (SDGs): Bangladesh Perspective, Government of the People's Republic of Bangladesh.
4. Kusters, C.S.L. and Batjes, K. with Wigboldus, S., Brouwers, J. and Baguma, S.D. (2017) Managing for Sustainable Development Impact: An Integrated Approach to Planning, Monitoring and Evaluation, Wageningen: Wageningen Centre for Development Innovation, Wageningen University & Research, and Rugby, UK: Practical Action Publishing, http://dx.doi.org/10.3362/9781780449807.

STA-DSEC-508

Applied Multivariate Analysis

Full Marks: 100 4 **Credits**

Learning objectives: The objective of this course is to provide students understanding about various extensions of bivariate probability distributions with applications. Some advanced topics in multivariate statistical analysis and their applications in various fields will be discussed here. By the end of the course, students will be proficient in analysing multivariate data, conducting hypothesis tests, and applying advanced techniques for classification, discrimination, cluster analysis, Principal component analysis and multivariate analysis of variance.

Course outcomes: Students will have acquired advanced skills in multivariate statistical analysis, enabling them to analyze complex datasets, make informed decisions, and contribute to research or practical applications in their respective fields.

Unit 1 Distribution of linear and quadratic forms in normal variables, expectations, variances and covariances, characteristic functions, independence of quadratic forms, conditions for a quadratic form to be distributed as chi-square and non-central chi-square, decomposition of quadratic forms.

(15 Contact hours)

Unit 2 MLEs of the parameters of multivariate normal distribution and their sampling distributions, Wishart distribution and its properties, tests of hypothesis about the mean vector of a multinormal population, Hotelling's T^2 - statistic, its distribution and applications.

(15 Contact hours)

Unit 3 Classification and discrimination for two known populations: Bayes', minimax and likelihood ratio procedures, Mahalonobis D^2 - statistic and its application, sample discriminant function and discrimination based on Fisher's method, cluster Analysis and evaluation of clusters.

(15 Contact hours)

Unit 4 Introduction to principal component analysis, canonical correlation analysis, factor analysis, MANOVA and its applications (sans derivation of the distribution of Wilk's λ).

(15 Contact hours)

Text Books

1. Anderson, T.W. (1983). An Introduction to Multivariate Statistical Analysis, Wiley Eastern, New Delhi.

2. Johnson, R. and Wychern, D.W. (2002). Applied Multivariate Statistical Analysis, Pearson Education, Delhi.

3. Rao, C.R. (1995), Linear Statistical Inference, Wiley Eastern, New Delhi.

4. Sharma, S. (1996). Applied Multivariate Techniques. John Wiley, New York.

5. Singh, B.M. (2002). Multivariate Statistical Analysis, South Asian Publishers, New Delhi.

Additional References

1. Giri, N.C. (1977). Multivariate Statistical Inference. Academic Press, New York.

2. Kshirsagar, A.M. (1972). Multivariate Analysis, Marcel Dekker, New York.

3. Muirhead, R.J. (1982). Aspects of Multivariate Statistical Theory, John Wiley, New York.

4. Seber, G. A. F. (1984). Multivariate Observations, John Wiley, New York.

STA-RM-509

Research Methodology & Proposal Writing

Full Marks: 100 **4 Credits**

Learning Objectives: The course mainly focuses on research methodology in statistics. This course will provide basic knowledge on research methods and statistical tools. Several types of research will be studied here. Finally various steps to write a research proposal will be discussed here.

Course outcomes: By course completion, students should have a comprehensive understanding on statistical tools and techniques in Research and will be able to write a research proposal.

Unit 1 Simulation methods, Monte Carlo integration, importance sampling, MCMC – Metropolis Hestings, Random Walk MH, Gibbs sampling, Resampling methods - Bootstrap, Jackknife, cross-validation, Permutation tests, Nonparametric curve estimation, EM algorithm..

(15 Contact hours)

Unit 2 Tools for Data Quality Assessment, detection of outliers and robust statistics, extreme value techniques, missing value estimation, imputation and related methods, Principles of data quality assurance and framework in surveys, Ethics and data quality, Paradata and their use in data quality improvement, paradata analytics and data quality metrics, indicators to measure data quality.

(15 Contact hours)

Unit 3 Scientific thinking and research, characteristics of research, classification of research, comparison and examples of basic, applied and action research, focus group research method, Synthesis of research findings and meta-analysis, techniques to combine statistics to common metric/effect size measures.

(15 Contact hours)

Unit 4 Formulation of research problem, the source of problem, criteria and guidelines of selection of a problem, literature review and different approaches, guidelines to write research report/article, writing a research proposal/project, case study.

(15 Contact hours)

Text Books

1. Wasserman, L. (2004). All of Statistics – A Concise Course in Statistical Inference. Springer Science + Buisness Media, New York.
2. Agarwal, C.C.(2017). Outlier Analysis. Springer Cham, New York.
3. Good, P.I. (2006). Resampling Methods A Practical Guide to Data Analysis, Barkhauser Boston, MA, USA.
4. Little, R.J.A and Rubin, D.B. (2019). Statistical Analysis with Missing Data. Wiley, New York.
5. Olkin, I. (1985). Statistical Methods for Meta Analysis, Academic Press, USA.

Additional References

1. ICMR-National Institute of Medical StatisticsMS. (2021). National Guidelines for Data Quality in Surveys. New Delhi: ICMR-NIMS.

STA-SEC-510

Practical – II

Full Marks: 100 **4 Credits**

Learning Objectives: All practical applications of sample survey, multivariate data using multivariate statistical techniques will be done here with the help of Software R

Course outcomes: After completion of these practical students will be able to see real life situations with sample survey and multivariate data

Section A: Sample Survey and Estimation Theory

- Simple random sampling – all possible samples
- Estimation using SRSWR and SRSWOR
- Estimation using SRSWR based on distinct units
- PPSWR & PPSWOR sampling – selection and estimation
- Stratified Sampling – estimation, sample allocation and construction of strata
- Estimation in linear and circular systematic sampling
- Estimation in cluster Sampling – equal and unequal cluster size
- Estimation in two-stage sampling – equal and unequal size units
- Estimation in double Sampling – ratio and regression estimator
- Estimation in double Sampling – for stratification
- Plotting of likelihood function and finding m.l.e. of parameters – using numerical methods
- Estimation by the method of scoring
- Estimation by the method of minimum Chi – square
- Estimation by the method of moments
- Estimation by the method of scoring for Cauchy and Chi – square distribution.

(40 Contact hours)

Section B Linear Model and Design of Experiment

- Single factor design
- Multiple Range Test
- Randomized Complete Block Design (RCBD)
- Latin Square Design (LSD)
- Balanced Incomplete Block Design (BIBD)
- 2^n factorial experiments, n = 3, 4
- 3^n factorial experiments, n = 2, 3
- Complete confounding
- Partial confounding
- Split-plot design

(40 Contact hours)

Section C: Applied Multivariate Analysis

- Estimation of mean and dispersion matrix
- Application of Hotelling's T^2 – statistic for single and two sample problems
- Discrimination between two multivariate normal populations with unknown parameters and common dispersion matrix
- Application of D^2 - Statistic
- Extraction of clusters
- Extraction of principal components and summarization of sample variations
- Canonical correlation analysis
- Factor analysis
- MANOVA (one way)

(40 Contact hours)

III SEMESTER

STA-CC-600

Statistical Inference

Full Marks: 100 **4 Credits**

Learning objectives: This course covers hypothesis testing for various scenarios. Students will learn to formulate hypotheses and to apply optimal tests. Students will be able to evaluate tests and interpret results accurately. Additionally, students will be able to explore the application of testing of hypothesis. Moreover various non parametric test also will be discussed here. Bayesian inference also included here to study prior and posterior distributions.

Course outcomes: After completing this course, students will understand the basic principles and methods of hypothesis testing, Non parametric inference and a Bayesian inference.

Unit 1 Tests of Hypotheses: concepts of critical regions, test functions, two kinds of errors, size function, power function, level of significance, MP test, Neyman-Pearson Lemma, likelihood ratio test, Monotone likelihood ratio property, UMP and UMPU tests, Karlin-Rubin theorem, Wald's SPRT with pre-specified errors of two kinds, OC and ASN functions.

(15 Contact hours)

Unit 2 One sample location problem, sign test and signed rank test, one and two sample Kolmogorov Smirnov tests, Two sample location problems. Wilcoxon-Mann-Whitney test, normal score test, ARE of various tests based on linear rank statistics. Kruskal-Wallis K sample test, one and two sample U statistics, asymptotic distribution of U statistics.

(15 Contact hours)

Unit 3 Basic concepts of decision theory; inference problems viewed as decision problem, Problem of classification, minimax approach and Bayes' approach, structure of Bayes' rule, complete class of rules, construction of minimax rule.

(15 Contact hours)

Unit 4 Concepts and evaluation of subjective probability of an event; subjective prior distribution of a parameter. Bayes' theorem and computation of posterior distribution; natural conjugate family of prior for a model, loss function, Bayes' risk, Bayesian estimation of parameters of binomial, Poisson, normal and exponential distributions.

(15 Contact hours)

Text Books

1. Bansal, A. K. (2007): Bayesian Parametric Inference, Narosa Publishing House, New Delhi.
2. Casella. G and Berger, R.L. (1990) Statistical Inference, Wordsworth and Brooks, California.
3. Ferguson, T.S. (1996). Mathematical Statistics- A Decision theory approach, Academic press, London.
4. Gibbons, J.D. (1985). Non-parametric Statistical Inference, Marcel Dekker, New York.
5. Kale, B.K. (1999). A first Course on Parametric Inference, Narosa Publishing House, New Delhi.
6. Rohatgi V. (1988). An Introduction to Probability and Mathematical Statistics, Wiley Eastern, New Delhi
7. Lehmann, E. L. (2015), "Theory of Point Estimation", 2ND Ed., Springer.

Additional References

1. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics. John Wiley and Sons, New York.
2. Mukhopadhyay, N. (2000). Probability and Statistical Inference, CRC, London
3. Mukhopadhyay, N. (2006). Introductory Statistical Inference, CRC, London

STA-CC-601

Regression Analysis

Full Marks: 100 **4 Credits**

Learning Objectives: The objective of this course is to equip students with a comprehensive understanding on the fundamental concepts of regression analysis. Students will learn regression models and their applications in analysing relationships between variables, the assumptions underlying regression analysis and checks the validity of regression assumptions, detect influential observations, and identify potential problems such as multicollinearity.

Course Outcomes: After completing this course, students should be equipped with the knowledge and skills necessary to conduct regression analysis effectively, critically evaluate regression models and apply them to solve practical problems in various fields.

Unit 1 Introduction to one-way random effects linear models and estimation of variance components, mixed effects models. Linear regression: Simple and Multiple - estimation of parameters, hypothesis testing and interval estimation, prediction of new observations; prediction of new observations. Polynomial regression and orthogonal polynomials.

(15 Contact hours)

Unit 2 Model adequacy checking, residual analysis, residual plots, normal probability plots, detection of outliers, transformations to improve normality and stabilizing variance, influential observations. Problems of multicollinearity. Power transformations for dependent and independent variables. Subset selection of explanatory variables, Mallow's Cp statistic. Introduction to model selection, AIC, BIC.

(15 Contact hours)

Unit 3 Nonlinear regression models: nonlinear LS, maximum likelihood estimation, transformations to linearize models. Exponential family of distributions, Inference in Exponential Family of distributions. Generalized Linear models (GzLMs), structure for the class of GzLM, Likelihood equations.

(15 Contact hours)

Unit-4 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 measure, analysis of deviance, Lack-of-Fit tests in Logistic regression. Introduction to regression for count and categorical data, MLE for Poisson regression and applications. Concept of over-dispersion in binary and count data regression. Idea of regression models for dependent data.

(15 Contact hours)

Text Books

1. Draper, N. R. and Smith, H. (2011). Applied Regression Analysis, John Wiley, New York.
2. Montgomery, D. C.; Peck, E. A. and Vining G. G. (2006). Introduction to Linear Regression Analysis. John Wiley, New York.
3. Seber, G. A. F and Lee Alan J. (2014). Linear Regression Analysis, John Wiley, New York.
4. Cook, R. D. and Weisberg, S. (1982). Residual and influence in Regression. Chapman and Hall.
5. Draper, N. R. and Smith, H. (1998). Applied Regression Analysis, John Wiley, New York.
6. Searle, S. R. (1971). Linear Models. Wiley, New York.
7. Seber, G. A. F. and Wild, C. J. (2003). Non-linear Regression, Wiley.
8. Dobson, A.J. and Barnett, A.G. (2008). Introduction to Generalized Linear Models, Third

Edition, Chapman and Hall/CRC. London.

9. Hosmer, D.W. and Lemeshow, S. (2000). Applied Logistic Regression, Second Edition.

Wiley, New York.

Additional References

1. Chatterjee, S, and Hadi, Ali S. (2013). Regression Analysis by Example, John Wiley. New York.
2. Mc Cullagh, P. and J.A. Nelder (1989): Generalized Linear Models, 2nd Edition Chapman & Hall, London.
3. Myers, R.H., Montgomery, D.C and Vining, G.G. (2002). Generalized Linear Models with Applications in Engineering and the Sciences, John Wiley & Sons.
4. Molenberghs, G and Verbeke, G. (2005): Models for Discrete Longitudinal Data, Springer-Verlag; New York.
5. Diggle, Liang, K.-Y. and Zeger, S.L. (1994): Longitudinal Data Analysis, Oxford University Press, Oxford.

STA-CC-602

Time Series Analysis and Forecasting

Full Marks: 100 **4 Credits**

Learning objectives:

The objective of this course is to provide students with a solid foundation in time series analysis and forecasting techniques, equipping them to model, analyze, and predict time-dependent data effectively. By covering the most basic and essential concepts and using hands-on practice the course indents to instill the proficiency of analyzing and forecasting business and economic time series data.

Course outcomes: By the end of the course, students will gain the ability to understand time series patterns, specify and estimate model parameters, perform diagnostic checks, and produce accurate forecasts, preparing them for real-world data analysis in fields such as economics, finance, and environmental studies.

Unit 1 Time series data, graphical display and exploratory analysis; trend, seasonal and cyclical components; decomposition of series, moving average, exponential smoothing and Holt-Winters method

(15 Contact hours)

Unit 2 Stationary time Series, basic time series models: white noise, random walk, AR, MA and ARMA models, Box-Jenkins correlogram analysis, ACF and PACF, choice of AR and MA orders.

(15 Contact hours)

Unit 3 Non-stationary time series, ARIMA models, deterministic and stochastic trends, introduction to SARIMA and ARCH models.

(15 Contact hours)

Unit 4 Model specification, estimation of ARIMA model parameters. Forecasting using exponential smoothing and Box – Jenkins model, Residual analysis and diagnostic checking.

(15 Contact hours)

Text Books

1. Brockwell, P. and Davis R.A.(2002). Introduction to Time Series and Forecasting, Springer.
2. Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley, New York.
3. Box, G.E.P., Jenkins, G.M. and Reinsel, G.C (2004). Time Series Analysis- Forecasting and Control, Pearson Education, Singapore.
4. Makridakis, S.G., Wheelwright, S.C. and Hyndman, R.J. (2005), Forecasting Methods and Applications, John Wiley and Sons.
5. Montgomery, D.C.,Jennings, C.L. and Kulachi, M.(2015). Introduction to Time Series Analysis and Forecasting, John Wiley and Sons.

Additional References

1. Enders, W. (2004). Applied Econometric Time Series, John Wiley and Sons.
2. Brokwel, P.J and Davis. R.A (1987). Time Series: Theory and Methods, Springer - Verlag, New York.
3. Fuller, W.A. (1976). Introduction to Statistical Time Series, John Wiley, New York.
4. Granger, C.W.J. and Newbold (1984). Forecasting Econometric Time Series, Academic Press, New York.
5. Shumway, R. H. and Stoffer, David S. (2006) Time Series Analysis and Its Applications: With R Examples. Springer-Verlag.

STA-DSEC-603

1. Demography and Population Dynamics

Full Marks: 100 4 **Credits**

Learning objectives: The course introduces the subject of Demography to the Lerners. Students will become familiar with basic concepts and sources of data in Demography and also will be able to comprehend the processes and events in Demography and their interactions. The also provides a higher level of understanding of population sciences including an in-depth knowledge of the linkages between population and various dimensions of socio-economic development and health. This course also provides a comprehensive idea to conduct further research in various aspects of population and development.

Course outcomes: At the end of the course, the students would be able to distinguish between demography and population studies, understand the population distribution by age and sex, and understand the importance of demographic transition, sex ratio and factor affecting population change. Will be able to calculate various demographic rates and ratios and epidemiological measures of mortality and morbidity using practical data. They are expected to develop ability to model population on important characteristics.

Unit 1 Introduction to Population: Meaning of Population, Size, structure, distribution of population, the structure of demographic rates. Age-sex pyramids. Demographic data: Census, Registration system, Indian SRS, and surveys. NFHS- 2, 3 & 4. Evaluation of Quality of demographic data: Chandrasekaran-Deming formula, accuracy of data on sex and age: Whipple's, Myer's and UN indices.

(15 Contact hours)

Unit 2 Mortality: concepts and rates, measures of infant mortality rate. Force of mortality, mortality laws - Gompertz and Makeham. Life table and its construction: complete and abridged, Greville's, Reed-Merrel's and Chiang's methods

(15 Contact hours)

Unit 3 Fertility and Reproduction: Period and cohort measures. P/F ratio and own children method, reproductive measures. Nuptiality rates. Gross and net nuptiality tables, internal and international migration: concept and rates, uses of place of birth and duration of residence data.

(15 Contact hours)

Unit 4 Theory of stable population model (one sex), quasi and stationary population. Lotka's stable population model. The equations characterizing a stable population, the effect of changes in fertility and mortality on age structure, growth rates, birth rates and death rates. Momentum of population growth. Population projection: Mathematical curves viz., growth curves, modified exponential, logistic curves and its properties, and their fitting, component method and matrix method of population projection.

(15 Contact hours)

Text Books

1. Samuel H. Preston (2001). Demography: measuring and modeling population processes. Blackwell.
2. Shryock, H.S. (1976).The methods and Materials of Demography. Academic Press, New York.
3. Lundquist, J. H. and Anderton, D. L. (2014). Demography: The Study of Human Population, Waveland Press, Inc.
4. Preston, S. and Heuvline, P. (2000). Demography: Measuring and Modeling Population Process, Blackwell publishers Ltd.

Additional References

1. International Institute of Population Sciences (1995). National Family Health Survey, 1992-93, Mumbai, IIPS.
2. International Institute of Population Sciences (2002). National Family Health Survey, 1998-99, Mumbai, IIPS
3. International Institute for Population Science (IIPS) and Macro International (2007). National Family Health Survey (NFHS - 3) 2005-06, Mumbai, IIPS.
4. Keyfitz, N. and Caswell, H. (2005). Applied Mathematical Demography, Third Edition, Springer, New York.
5. Krishnan Namboodiri. (1996). A primer of Population Dynamics. Plenum.
6. Mishra, B.D (1995). An Introduction to The Study of Population. South Asian Publications, New Delhi.
7. Ram Kumar (1986).Technical Demography. Wiley Eastern, New Delhi.
8. International Institute for Population Science (IIPS) and Macro International (2007). National Family Health Survey (NFHS - 4) 2015-16, Mumbai, IIPS.

STA-DSEC-603

2. Operations Research

Full Marks: 100 **4 Credits**

Learning Objectives: The objective of this course is to review the basic elements of Operations Research including linear programming problem, simplex method, duality theory, and transportation and assignment problems. The course is intended to elaborate upon Inventory problems and Queueing theory together with applications.

Course Outcomes: After completing this course, students will be able to understand the Operations Research tools and use the methods and techniques for effective decision-making in real life applications.

Unit 1 Definition and scope of operations research; phases in operations research, models and their solutions, decision-making under uncertainty and risk, use of different criteria, Review of linear programming (LP) problems - duality theorem.

(15 Contact hours)

Unit 2 Transportation and assignment problems; sensitivity analysis; non-linear programming; Kuhn Tucker conditions, Wolfe's and Beale's algorithms for solving quadratic programming problems.

(15 Contact hours)

Unit 3 Analytical structure of inventory problems; Economic order quantity (EOQ) formula of Harris, its sensitivity analysis and extension allowing quantity discounts and shortages, multi-item inventory subject to constraints, P and Q- systems with constant and random lead items.

(15 Contact hours)

Unit 4 Queueing models; specifications and effectiveness measures, steady-state solutions of M/M/1 and M/M/c models with associated distributions of queue length and waiting time, M/G/1 queue and Pollazcek-Khinchin result, steady state solutions of M/Ek/1, simulation.

(15 Contact hours)

Text Books

1. Kanti, S., Gupta, P.K. and Singh, M.M. (1995). Operations Research, Sultan Chand & Sons, New Delhi.
2. Taha, H.A. (1982). Operational Research: An Introduction, Macmillan, New York.
3. Wagner, H.M. (1994). Principles of Operations Research, Prentice Hall of India, New Delhi.

Additional Reference

1. Hillier, F.S. and Leiberman, G.J. (1962). Introduction to Operations Research, Holden-day, San Francisco.

STA-DSEC-604

3. Reliability Theory

Full Marks: 100 **4 Credits**

Learning Objectives: The objective of the course is to impart reliability theory concepts into the students. To make them understand the failure time distributions of various types and their modelling.

Course outcomes: After successful completion of the course the students thoroughly account for industrial applications of different methods in reliability theory. They are expected to effectively model the life time data under constant and variable failure rates, under censored conditions and up to bivariate shock models.

Unit 1 Reliability concepts and measures; components and systems, coherent systems; reliability coherent systems, cuts and paths, modular decomposition, bounds and system reliability; structural and reliability importance components.

(15 Contact hours)

Unit 2 Failure time distributions; reliability function, hazard rate, common failure time distributions; exponential, Weibull, gamma etc., estimation of parameters and tests in these models, notions of ageing; increasing failure rate (IFR).

(15 Contact hours)

Unit 3 Increasing failure rate average (IFRA), not better than used (NBU), decreasing mean residual life (DMRL) and not better than used in expectation (NBUE), classes and their duals; loss of memory property of the exponential distribution; closures of these classes under formation of coherent systems, convolutions and mixtures.

(15 Contact hours)

Unit 4 Univariate shock models and the distributions arising out of them; bivariate shock models, common bivariate exponential distributions and their properties, reliability estimation based on failure times in censored life tests and in tests with replacement of failed items; stress-strength reliability and its estimation.

(15 Contact hours)

Text Books

1. Bain L.J. (1991) Statistical Analysis of Reliability and Life Testing Models, Marcel Dekker, New York.
2. Barlow R.E. and Proschan, F. (1985). Statistical Theory of Reliability and Life Testing, Holt Rinehart and Winston, New York.

Additional References

1. Lawless, J.F. (1982). Statistical Models and Methods of Life Time Data, John Wiley, New York.
2. Nelson, W. (1982). Applied Life Data Analysis, John Wiley, New York.

STA-DSEC-604

4. Econometrics

Full Marks: 100 **4 Credits**

Learning objectives: This course provides a comprehensive understanding of the techniques of econometrics and knowledge of simple and multiple linear regressions, model diagnostic, logistic regression, panel data models and time series analysis. The goal of the course is to introduce the students to the various modeling techniques under regression analysis that would help them in decision making process.

Course outcomes: At the end of this course students will be able to conduct independent econometric and statistical analysis of data, demonstrate their understanding of applied econometric analysis models/ methods with respect to choice of model, estimation method and interpretation of

UNIT 1 Definition of Econometrics, Single equation linear model, types of Econometrics, Measurement scales of variables, population & sample regression functions, linearity, definition and specification of general linear model with assumptions, Stochastic specification of PRF; techniques of ordinary least squares (OLS) and Generalized least squares (GLS), difference between OLS and GLS.

(15 Contact hours)

UNIT 2 Problems of heteroscedasticity, autocorrelation and multicollinearity - their consequences and diagnosis - VIF, tolerance, eigenvalues, condition index, condition number and remedies; ridge regression with applications.

(15 Contact hours)

UNIT 3 Principal components regression and generalized inverse regression, concept of dummy variables and application in Regression, probit and logit regressions.

(15 Contact hours)

UNIT 4 Distributed-lag models (DLM) specifications, estimation of parameters under various DLMs. Instrumental variables, method of restricted least squares.

(15 Contact hours)

Text Books

1. Anders, B. (2001), "Ridge Regression and Inverse Problems", Stockhome University, Sweden.
2. Apte, P. G. (1990), "Text Book of Econometrics", Tata McGraw Hill, ND.
3. Baltagi, B. H. (2005), "Econometrics", Springer (India), ND.
4. Gujarati, D. (2004), "Basic Econometrics", 4TH Ed., McGraw Hill, ND.
5. Gujarati, D. (2012), "Econometrics by Example", Indian Ed.
6. Gujarati, D., Dawn, C. P. and Gunasekar, S. (2011), "Basic Econometrics", 5TH Ed., McGraw Hill education.
7. Green, W. H. (2003), "Econometric Analysis", 5TH Ed., Prentice Hall, New Jersey.
8. Hansen, B. E. (2016), "Econometrics", Revised, University of Wisconsin, Madison.
9. Johnston, J. (1984), "Econometric Methods", McGraw Hill, NY.
10. Verbeek, M. (2004), "A Guide to Modern Econometrics", 2ND Ed., John Wiley & Sons, England.

Additional Books

1. Intrilligator, M. D. (1980), "Econometric Models – Techniques and applications", Prentice Hall of India, ND.
2. Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H. and Lee, T. C. (1985), "The Theory and Practice of Econometrics", John Wiley and Sons, NY.
3. Koutsoyiannis, A. (1977), "Theory of Econometrics", Macmillan, London.
4. Maddala, G. S. (1977), "Econometrics", McGraw Hill Book Company, Singapore.
5. Malinvaud, E. (1970), "Statistical Methods of Econometrics", North Holland, Amsterdam.
6. Schmidt, P. (1976), "Econometrics", Marcel Dekker, NY.
7. Theil, H. (1971), "Principles of Econometrics", John Wiley and Sons, NY.

STA-SEC-605

Practical – III

Full Marks: 100 **4 Credits**

Learning Objectives: All practical applications of statistical inference, regression and time-series will be done here with the help of Software R

Course outcomes: After completion of these practical students will be able to use statistical inference, regression and time-series in real life situations.

Problems based on following topics:

Section A: Statistical Inference

- One and two sample Sign and Signed rank tests
- One and two sample Kolmogorov-Smirnov tests & p – p plot
- Wilcoxon-Mann-Whitney test
- Normal Score Test
- Kruskal-Wallis K Sample Test
- Minimax estimation
- Hypothesis testing: UMP, UMPU tests, its power function and plotting
- Bayesian estimation of risk under different priors and losses
- SPRT

(40 Contact hours)

Section B: Regression Analysis

- Simple linear regression
- Multiple linear regression
- Model adequacy checking, residual plots
- Transformations to improve normality and stabilize variance
- Polynomial Regression
- Logistic regression
- Non-linear Regression
- Poisson Regression

(40 Contact hours)

Section C: Time Series Analysis and Forecasting

- Correlogram Analysis and Interpretation
- Smoothing and Forecasting using simple exponential
- Forecasts using Holt and Winter model
- Modeling and Forecasting with pure MA/AR models
- Modeling and Forecasting with mixed ARMA models
- Fitting and Forecasting with ARMA models
- Modeling seasonal data using SARIMA and Forecasting
- Modeling volatility using ARCH
- Residual Analysis
- Diagnostic checking

(40 Contact hours)

Text Books

1. Bhattacharyya, G.K. and Johnson, R.A. (1977). Statistical Concepts and Methods, John Wiley, New York.
2. Goon, A.M., Gupta, M.K. and Disrupt, B. (2000). Fundamentals of Statistics, World Press, Kolkata.
3. Hogg, R.V. and Tanis, E.A. (2003). Probability and Statistical Inference, Pearson Education, Delhi.
4. Chatterjee, S. and Price, B. (1991). Regression Analysis by Example, John Wiley, New York.
5. Draper, N.R. and Smith, H (1998). Applied Regression Analysis, John Wiley, New York.
6. Montgomery, D. C.; Peck, E. A. and Vining G. G. (2004). Introduction to Linear Regression Analysis. John Wiley, New York.
7. Seber, G. A. F and Lee Alan J. (2003). Introduction to Linear Regression Analysis, John Wiley, New York.
8. Brockwell, P. and Davis R.A. (2002). Introduction to Time Series and Forecasting, Springer.
9. Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley, New York.
10. Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley, New York.
11. Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis- Forecasting and Control, Holden-day, San Francisco.
12. Granger, C.W.J. and Newbold (1984). Forecasting Econometric Time Series, Academic Press, New York.12

IV SEMESTER

STA-DSEC-606

1. **Planning, Execution and Dissertation** **16 Credits**

2. **Presentation and Viva-Voce** **4 Credits**

Learning Objectives:

Students are expected to carry out a Dissertation within a time frame under a supervisor's guidance culminating in a final Dissertation. The Dissertation aims to provide practice and experience to integrate and apply knowledge and skills gained throughout the statistics academic discourse on a real-time project. A platform to use theoretical foundations of the programme will better prepare students to handle statistical tools undertakings in a real work environment.

Course outcomes: On completion of the project, students will be able to identify a suitable topic for a project, formulate and articulate questions, hypotheses, and objectives, carry out a literature search and summarize the state of the art, articulate suitable methodology and effectively and professionally engage with primary or secondary data sources and apply robust statistical techniques to solve real-world problems.

Note: This paper shall be based on some real life application of Statistical analysis of data. The learners are expected to be able to formulate a problem situation, collect relevant data from field, analyze data using available Statistical softwares, interpret the results and prepare a dissertation.