FOUR YEAR UNDERGRADUATE (FYUG) PROGRAMME UNDER NEW EDUCATION POLICY, 2020

STATISTICS

Date of approval in Academic Council –May 30, 2024

Preface

The four-year undergraduate (FYUG) syllabus for Statistics has been framed in accordance with NEP-2020 guidelines. As per NEP-2020 recommendations, the course has been designed keeping in view multiple exit options:

(i) A certificate after successful completion of one year of study;

(ii) A diploma after successful completion of two years of study;

(iii) A 3-year UG degree after successful completion of three years of study; and

(iv) A UG Honours or UG Honours with Research after successful completion of a four years of study.

The FYUG programme in Statistics has been designed to motivate and inspire students to develop a deep interest in the subject Statistics, develop problem-solving abilities and understand statistical theory, practical and statistical tools and softwares. It includes laboratory work and practical exercises that allow students to apply theoretical concepts to real-world problems and enhance their scientific skills. It also trains the students to use various statistical packages in analysing data, obtain estimates of necessary parameters, set and test hypotheses and draw meaningful inferences thereby preparing technically proficient statisticians. It also aims to provide students with necessary knowledge and skills to design a survey instrument, select a sample scientifically from the population under study, conduct survey to collect data, represent data in various forms, analyse collected data and draw valid conclusions that could be used in decision making. The programme includes a base of theoretical work complemented by practical analysis of real-world data to give them training in applying and sharpening the skills they have acquired during the course of the programme. The course aims to disseminate academic, research, and professional development knowledge to the students.

Programme Outcomes:

Upon completion of the programme, the learners will be able to:

- Plan for a scientific investigation for an identified issue.
- Design an appropriate instrument to be used for collection of data related to the identified issue.
- Select a representative sample from the population under study and administer survey to collect relevant data.
- Organize and present data in suitable form (diagrams, graphs, measures etc.).
- Handle and analyse large datasets with computer skills and use their results and interpretations to make practical suggestions for improved decision making.
- Apply statistical methods and modelling techniques to real-world problems in both observational and designed studies.
- Employ their knowledge and expertise for the development of a research enquiry and to select the tools necessary for executing the research
- Develop independent learning strategies.
- Coordinate and work with multidisciplinary teams.
- Display a capacity for logical thinking, structured reasoning and synthesis

Structure of the Syllabus

1st Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA - 100	Introductory Statistics (Major)	3	1	4	75
	(Minor)			4	
MDC-110.....119	Any one of the available MDC Courses to be chosen			3	
AEC-120….129	Any one of the available AEC Courses to be chosen			3	
SEC-130…...139	Any one of the available SEC Courses to be chosen			3	
VAC-140	Any one of the available Vocational Education and training Course to be chosen (Minor Course)			3	

2nd Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA – 150	Introduction to Probability and Applied Statistics (Major)	3	1	4	75
	Minor			4	
MDC – 161	Fundamentals of Statistics	2	1	3	60
AEC-170…..179	Any one of the available AEC Courses to be chosen			3	
SEC-180……189	Any one of the available SEC Courses to be chosen			3	
VAC-190…...199	Any one of the available Vocational Education and training Course to be chosen (Minor Course)			3	

3rd Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA – 200	Mathematical Analysis and Numerical Methods (Major)	3	1	4	75
STA – 201	Probability Theory and Probability Distribution (Minor)	3	1	4	75
MDC 210 to 219	Any one of the available MDC Courses to be chosen			3	
AEC 220-229	Any one of the available AEC Courses to be chosen			2	
SEC 230-239	Any one of the available SEC Courses to be chosen			3	
VTC 240-249	Any one of the available Vocational Education and training Course to be chosen (Minor Course)			4	

4th Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA - 250	Distribution Theory (Major)	3	1	4	75
STA – 251	Matrix and Linear Algebra (Major)	3	1	4	75
STA - 252	Introduction to Statistical Inference (Major)	3	1	4	75
STA – 253	Introduction to Sample Survey (Major)	3	1	4	75
VTC 260-269	Any one of the available Vocational Education and training Courses to be chosen (Minor Course)			4	

5th Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA – 300	Introduction to Design of Experiments (Major)	3	1	4	75
STA – 301	Testing of Hypothesis (Major)	3	1	4	75
STA – 302	Applied Statistics(Major)	3	1	4	75
STA - 302	Theory of Attributes, Categorical data and Demand analysis(Minor)	3	1	4	75
STA-303	Internship/Apprenticeship/Community engagement and service field based learning or minor project			4	

6th Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA – 350	Non – Parametric & Sequential Procedures (Major)	3	1	4	75
STA – 351	Linear Estimation and Linear Models (Major)	3	1	4	75
STA – 352	Regression Analysis (Major)	3	1	4	75
STA – 353	Advanced Survey Sampling (Major)	3	1	4	75
VTC 260-269	Any one of the available Vocational Education and training Courses to be chosen (Minor Course)			4	

7th Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA - 400	Research Methodology & Proposal Writing (Major)	3	1	4	75
STA – 401	Applied Multivariate Analysis (Major)	3	1	4	75
STA – 402	Statistical Computing in R (Major)	3	1	4	75
STA – 403	Statistics for Sustainable Development (Minor)	3	1	4	75
STA-404	Introduction to Operations Research (Minor)	3	1	4	75

8th Semester

Course Code	Title of the Course	Credits			Total Contact hours
		Theory	Practical	Total	
STA – 450	Operations Research (Major)	3	1	4	75
STA - 451	Econometrics (Major)	3	1	4	75
STA - 452	Research Project/ Dissertation for Honors with Research only			12	
STA - 453	Biostatistics and Demography (Major#)	3	1	4	75
STA – 454	Advance Design of Experiments (Major#)	3	1	4	75
STA - 455	Time Series Analysis and Forecasting (Major#)	3	1	4	75

Assessment Approach

1. For papers having both of the theory as well as practical components, the assessment approach will be as follows:

	Internal	End Semester
Theory (Part A, 3 credits)	19	56
Theory (Part A, 2 credits)	12	38
Practical (Part B, 1 credit)	6	19
For VTC Courses	40	60

NORTH-EASTERN HILL UNIVERSITY, SHILLONG

Syllabus for Under Graduate Statistics under NEP 2020

I SEMESTER		
Paper Code	**Paper**	**Total Credit**
STA - 100	Introductory Statistics	**4**
II SEMESTER		
MDC - 161	Fundamentals of Statistics	**3**
STA - 150	Introduction to Probability and Applied Statistics	**4**
III SEMESTER		
STA - 200	Mathematical Analysis and Numerical Methods	**4**
STA - 201	Probability Theory and Probability Distribution	**4**
	IV SEMESTER	
STA - 250	Distribution Theory	**4**
STA - 251	Matrix And Linear Algebra	**4**
STA - 252	Introduction to Statistical Inference	**4**
STA - 253	Introduction to Sample Survey	**4**
	V SEMESTER	
STA - 300	Introduction to Design of Experiments	**4**
STA - 301	Testing of Hypothesis	**4**
STA - 302	Applied Statistics	**4**
STA - 303	Theory of Attributes, Categorical data and Demand analysis	**4**
STA - 304	Internship/Apprenticeship/Community engagement and service field based learning or minor project	**4**
	VI SEMESTER	
STA - 350	Non – Parametric & Sequential Procedures	**4**
STA - 351	Linear Estimation And Linear Models	**4**
STA - 352	Regression Analysis	**4**
STA - 353	Advanced Survey Sampling	**4**
	VII SEMESTER	
STA - 400	Research Methodology & Proposal Writing	**4**
STA - 401	Applied Multivariate Analysis	**4**
STA - 402	Statistical Computing in R	**4**
STA - 403	Statistics for Sustainable Development	**4**
STA - 404	Introduction to Operations Research	**4**
	VIII SEMESTER	
STA - 450	Operations Research	**4**
STA - 451	Econometrics	**4**
STA - 452*	Research Project/ Dissertation for Honours with Research only	**12**
STA - 453#	Biostatistics and Demography	**4**
STA - 454#	Advance Design Of Experiments	**4**
STA - 455#	Time Series Analysis and Forecasting	**4**

* Students securing 75% marks or more in aggregate in 6 Semester are eligible to opt for Honours with Research.

All others students must opt for UG Honours with 3 Advances Courses, viz STA-453, STA-454 and SUB-455 in lieu of STA-452.

STA -400

Name of Paper: Research Methodology & proposal writing

75 Lecture Hours **4 Credits**

Learning Objective: To impart the students a thorough knowledge on the meaning and definition of Research, various types of Research Designs, aspects of data collection and surveying techniques, simulation, and framework of research papers.

Course Outcome: On completion of this course students will know the different steps and methods used for research. To identify research problems, to frame research questions and research objective, Types of survey design, how to frame questionnaire. The students will learn different types of sampling tool used for data collection.

Part – A (Theory)

UNIT-I:

Meaning of Research, Objectives, and importance of Research, types of Research: Qualitative & Quantitative Research, Longitudinal Research, Survey & Experimental Research. Research Approaches, Significance of Research, Criteria of good research, Research Methods. Formulation of Research problems, criteria and guidelines for selecting a Research problem, Meta Analysis.

(15 hours)

UNIT-II:

Survey Methodology and Data Collection, inference and error in surveys, the target populations, sampling frames and coverage error, develop a questionnaire, validity and reliability of questionnaire, tools for data quality assessment, estimation of missing data, Ethics and data quality, scientific articles, statistical database.

(15 hours)

UNIT-III:

Simulation: Concepts and advantages, Simulation methods, Monte Carlo simulation, MCMC Principle. Measurement scales in Research. Framework for a good research article/report/publication (Guidelines and criteria), writing a research proposal/project, case study.

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

(30 hours)

Identifying the research problems, developing research ideas, formulate the research problems, literature review, Research proposals, Writing a Scientific Report, presentation of research work.

Reference Books:

1. Singh, Y. K. (2006). Fundamental of Research Methodology and Statistics. New Age International.
2. C.R. Kothari and Gaurav Garg (2018). Research Methodology: Methods and Techniques. New Age International
3. Kumar,R (2011):Research Methodology: A Step-by-Step Guide for Beginners, SAGE Publications
4. Pannerselvan, R. (2006). Research Methodology, Prentice-Hall of India Pvt., New Delhi
5. Robert, C.P and Casella, G. (2004).Monte Carlo Statistical Methods, Springer Science, New York
6. Anderson, J., Durston, B.H., Pooole, M. (1970). Thesis and Assignment Writing, Wiley Eastern Ltd., New Delhi.
7. Little, R.J.A and Rubin, D.B. (2019). Statistical Analysis with Missing Data. Wiley, New York.
8. Olkin, I. (1985). Statistical Methods for Meta Analysis, Academic Press, USA
9. Pajo, B. (2022). Introduction to research methods: A hands-on approach. Sage publications.

STA – 401: Applied Multivariate Analysis

75 Lecture Hours **4 Credits**

Learning objective: The objective of this course is to provide students with a deep understanding of advanced topics in multivariate statistical analysis and their applications in various fields. By the end of the course, students will be proficient in analyzing multivariate data, conducting hypothesis tests, and applying advanced techniques for classification, discrimination, and dimensionality reduction

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.

Part – A (Theory)

Unit-I:

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 hours)

Unit-II:

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 hours)

Unit-III:

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.

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

(15 hours)

Part – B (Practical)

Unit-IV: (Practical)

(30 hours)

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)

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.
6. Giri, N.C. (1977). Multivariate Statistical Inference. Academic Press, New York.
7. Kshirsagar, A.M. (1972). Multivariate Analysis, Marcel Dekker, New York.
8. Muirhead, R.J. (1982). Aspects of Multivariate Statistical Theory, John Wiley, New York.
9. Seber, G. A. F. (1984). Multivariate Observations, John Wiley, New York.

STA – 402: Statistical Computing in R

75 Lecture Hours **4 Credits**

Learning Objective: By the end of the course, students will be proficient in analyzing multivariate data, conducting hypothesis tests, and applying advanced techniques for classification, discrimination and dimensionality reduction.

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.

Part – A (Theory)

Unit-I:

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 hours)

Unit-II:

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 hours)

Unit-III:

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

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 hours)

Part – B (Practical)

Unit-IV: (Practical)

(30 hours)

Basic calculations using R as a calculator (e.g., arithmetic operations, exponentiation), Manipulating different data types (numeric, character, logical), creating and manipulating arrays and matrices, reading data from various sources (e.g., Excel files, database), Practicing R syntax (e.g., variable assignment, function calls), Calculating descriptive Statistics for datasets (e.g., mean, median and variance),Generating tables summarizing data(e.g., frequency

tables, cross tabulations),Exploring built-in functions for data analysis (e.g., aggregate, summary), conducting exploratory data analysis using visualizations and summary statistics, simulating probability distributions (e.g., normal distributions, binomial distribution).

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.
4. R for Data Science by Hadley Wickham and Garrett Grolemund.

Additional references

1. Chambers, J. (2008). Software for Data Analysis: Programming with R, Springer.
2. Chang, W. (2024). R Cookbook, O'Reilly.
3. Everitt, B. and Hothorn, T. (2006). A Handbook of Statistical Analyses Using R, CRC.
4. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021) Introduction to Statistical Learning with applications in R, Springer US.
5. Jones, O., Maillardet, R. and Robinson, A. (2009). Introduction to Scientific Programming and Simulation using R, CRC.
6. Wickham, R. (2014). Advanced R, CRC Press.

STA-403: Statistics for Sustainable Development

75 Lecture Hours **4 Credits**

Learning Objectives: To provide students with an in-depth understanding of sustainable development and its components. The students will understand the different types of indexes use for measuring social equity, environmental protection and SDG indicators. Human development index, quality of life index and economic index will be part of the course.

Course outcomes: On completion of the course, the students should have a comprehensive understanding of sustainable development, social equity and environmental protection. The students will know different types of law and policies for sustainable development. The students will know the types of indices used for measuring the social status, economic status and quality of life.

Part – A (Theory)

UNIT-I:

Economic Growth and Sustainability, alternative definitions and concepts of weak and strong sustainability; Social Equity and Environmental protection. UN adopted Sustainable Development Goals (SDGs), SDG indicators and Data Source,

(15 hours)

UNIT-II:

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.

(15 hours)

UNIT-III:

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. SDG1 and poverty, global Multidimensional Poverty Index (MPI) of UNDP, Composition of MPI and deprivation indicator.

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

(30 hours)

Problems based on SDGs and its components, HDI, Dimension of HDI,

Problems on Physical Quality of Life Index, Gender-equity index.

Problems based on MPI and composition of MPI.

Reference 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 Multidimensional Poverty Index 2020: Charting Pathways out of Multidimensional Poverty: Achieving the SDGs.

STA-404: Introduction to Operations Research

75 Lecture Hours **4 Credits**

Learning objectives:

This course introduces mathematical methods for formulating, solving and providing quantitative data for managerial decision making under the given constraints of resources. It also introduces decision making under deterministic and probabilistic situations. Linear programming models, transportation and assignment models, network models and decision trees will be discussed in detail. The students will also be trained to use TORA, Excel Solver and R to solve optimization problems.

Course outcomes:

On completion of the module, the students will be able to Formulate a given real-life optimization problem in to a Linear Programming Problem (LPP) with a clear objective function and constraints, Solve a given LPP using appropriate method and interpret the result, Obtain the optimal solution to a given transportation model/ assignment model and estimate the least cost for such a model.

Part – A (Theory)

UNIT-I:

Linear Programming Problems (LPP), Introduction to LPP, Formulation of LPP, Simplex Method, Use of Artificial Variables in simplex method, Degeneracy , Infeasible solutions

(15 hours)

UNIT-II:

Duality in LPP**,** Dual Simplex Method, Sensitivity analysis, Computer solution to LPP with TORA, Excel Solver and R. Introduction to Transportation Problem, TP as a case of LPP, Methods to obtain initial basic feasible solution to a TP, North West Corner Rule, Matrix Minima Method, Vogel's Approximation Method, Solution of the TP by MODI method, Degeneracy in TP, Unbalanced transportation problems and their solutions.

(15 hours)

UNIT-III:

Introduction to AP**,** AP as a complete degenerate form of TP**,** Hungarian Method for solving AP**,** Unbalanced Assignment Problems and their solutions**,** AP with restrictions

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

Computer solution to LPP with TORA, Excel Solver and R
Computer solution to TP and AP using TORA, Excel Solver and R

(30 hours)

Reading List

Essential reading

Taha, H. A. (2013). *Operations research: an introduction*. (9th ed.). Pearson Education India.

Hillier, F. S. (2021). *Introduction to operations research*. (11th ed.). Tata McGraw-Hill Education.

Vanderbei, R. J. (2020). Linear programming: foundations and extensions (Vol. 285). Springer Nature.

Additional reading

Bernard W. Taylor III. (1993). *Introduction to Management Sciences*. (4th ed.). New York: Allyn and Bacon.

Kanti Swarup, Manmohan, Gupta. (2018). *Operations Research*. (10th ed.). New Delhi: Sultan Chand and Sons Publishing Co.

Manmohan & Gupta P.K. (1987). *Operations Research and Statistical Analysis*. (3rd ed.). New Delhi: Sultan Chand and Sons Publishers.

STA – 450: OPERATIONS RESEARCH

75 Lecture Hours **4 Credits**

Learning Objectives: The objective of this course is to introduce the basic elements of Operations Research including linear programming problem, simplex method, duality theory, transportation and assignment problems

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

Part – A (Theory)

UNIT-I:

Definition and scope of operations research. Linear Programming: Introduction. Definition of general linear programming problems (LPP). Formulation and examples of LPP. Problems occurring in various fields. Graphical method of solving LPP. Slack and Surplus variables.

(15 hours)

UNIT-II:

Simplex method. Duality theory- Introduction and applications, Formulation of Primal Dual problems, weak and strong duality theorems. Role of Duality Theory in Sensitivity analysis. Dual Simplex method.

(15 hours)

UNIT-III:

Representation of transportation and assignment problems as LPP. Solution of transportation problems using initial basic feasible solution with the help of North-West Corner rule, Matrix minima and Vogel's methods. Solution of assignment problems using Hungarian method.

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

(30 hours)

1. Solving LPP by graphical method.
2. Problem based on Simplex method.
3. Problem based on Primal Dual problem.
4. Problem based on Dual Simplex method.
5. Problem based on Transportation problem: Vogel's method and North-West corner rule.
6. Solution of assignment problem by Hungarian method.

Suggested readings:

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. Gass S.I. (2003) Linear Programming: Methods and Applications.
4. Hadley, G (1962) Linear Programming, Addison- Wesley Pub.
5. Wagner, H.M. (1994). Principles of Operations Research, Prentice Hall of India, New Delhi.
6. Hillier, F.S. and Leiberman, G.J. (1962). Introduction to Operations Research, Holden-day, San Francisco.

STA -451 Econometrics

75 Lecture Hours **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 outcome: 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 results.

Part – A (Theory)

UNIT-I:

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 hours)

UNIT-II:

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

(15 hours)

UNIT-III:

Principal components regression and generalized inverse regression, concept of dummy variables and application in Regression, Distributed-lag models (DLM) specifications, estimation of parameters under various DLMs. Instrumental variables, method of restricted least squares.

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

(30 hours)

- Simple and multiple linear regression
- Model adequacy checking, residual plots
- Non-linear Regression
- Forecasting of General linear model
- Residual Analysis

Text Books

1. Anders, B. (2001), "Ridge Regression and Inverse Problems", Stockhome University, Sweden.
2. Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley, New York.
3. Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley, New York.
4. Apte, P. G. (1990), "Text Book of Econometrics", Tata McGraw Hill, ND.
5. Baltagi, B. H. (2005), "Econometrics", Springer (India), ND.
6. Bhattacharyya, G.K. and Johnson, R.A. (1977). Statistical Concepts and Methods, John Wiley, New York.
7. Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis- Forecasting and Control, Holden-day, San Francisco.
8. Brockwell, P. and Davis R.A. (2002). Introduction to Time Series and Forecasting, Springer.
9. Chatterjee, S. and Price, B. (1991). Regression Analysis by Example, John Wiley, New York.
10. Draper, N.R. and Smith, H (1998). Applied Regression Analysis, John Wiley, New York.
11. Goon, A.M., Gupta, M.K. and Disrupt, B. (2000). Fundamentals of Statistics, World Press, Kolkata.
12. Granger, C.W.J. and Newbold (1984). Forecasting Econometric Time Series, Academic Press, New York.12
13. Green, W. H. (2003), "Econometric Analysis", 5TH Ed., Prentice Hall, New Jersey.
14. Gujarati, D. (2004), "Basic Econometrics", 4TH Ed., McGraw Hill, ND.
15. Gujarati, D. (2012), "Econometrics by Example", Indian Ed.
16. Gujarati, D., Dawn, C. P. and Gunasekar, S. (2011), "Basic Econometrics", 5TH Ed., McGraw Hill education.
17. Hansen, B. E. (2016), "Econometrics", Revised, University of Wisconsin, Madison.
18. Hogg, R.V. and Tanis, E.A. (2003). Probability and Statistical Inference, Pearson Education, Delhi.
19. Johnston, J. (1984), "Econometric Methods", McGraw Hill, NY.
20. Montgomery, D. C.; Peck, E. A. and Vining G. G. (2004). Introduction to Linear Regression Analysis. John Wiley, New York.
21. Seber, G. A. F and Lee Alan J. (2003). Introduction to Linear Regression Analysis, John Wiley, New York.
22. Verbeek, M. (2004), "A Guide to Modern Econometrics", 2ND Ed., John Wiley & Sons, England.

STA -452 Research Project/ Dissertation for Honours with Research only)

75 Lecture Hours **4 Credits**

1. A Project work is intended to give students scope to apply and demonstrate statistical techniques.
2. The project works should preferably be based upon field survey. In case field surveys are not feasible, it can be based secondary data also.
3. The project report should clearly address a problem with methodology applied, hypotheses formulated and conclusions drawn.
4. The project will be supervised by a faculty member allocated by the Department preferably during the seventh semester.
5. A students has to submit a project report/dissertation to the Department and give oral presentation of the project report before a board of examiners for evaluation.

STA-453: Biostatistics and Demography

75 Lecture Hours **4 Credits**

Learning objectives: To equip students with basic understanding of the techniques and tools of biostatistics and demography. To make the students understand the various demographics events and processes that shape the population size and structure. To make students learn the tools in statistics that can be applied in health science including demography.

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.

Part – A (Theory)

UNIT-I: Theory of Demography

Meaning and Scope of Demography; distinction between demography and population studies, Components of Population change, Age-sex structure, Population pyramids, factors affecting age-sex structure of the population, factors affecting sex-ratio of the population, stages of demographic transition, sources of demographic data. Dependency ratio, population theories.

(15 hours)

UNIT-II: Elementary Biostatistics

Measuring the occurrence of disease: Measures of morbidity - prevalence and incidence rate, ratios and proportions, association between prevalence and incidence, uses of prevalence and incidence, Attributable Risk, Relative Risk, Odds Ratio, Risk Ratio along with their Standard Errors and Confidence Intervals (Point and Interval estimation). Problems with incidence and prevalence measurements; Clinical agreement: kappa statistics, Mantel-Haenszel test; intra-class correlation.

(15 hours)

UNIT-III: Epidemiology, Population Estimation and Projection

Concepts of population projections; population estimates, forecasts and projections, uses of population projections; Methods of interpolation, extrapolation using linear, exponential. Population projection based on Cohort component method, rural-urban and sub national projections. Concept of health and morbidity, Various definition of health, Epidemiology objectives, measures of Epidemiology- measures of mortality, morbidity and disability, Epidemiological (Health) Transition, basic idea of statistical methods in epidemiology. SIR model.

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

(30 hours)

Problems on construction of population pyramids, dependency ratio, sex ratio. Problems on calculation of. Problems on population projection using cohort component, rural-urban and sub national projections. Problems on population interpolation and extrapolation using linear, exponential method, Basic Statistical methods in epidemiology - Disease prevalence, incidence, attributable risk, relative risk, odds ratio, intraclass correlation, Disease Screening, Disease clustering, kappa statistics, Mantel-Haenszel test, SIR model.

Reference Books:

1. Henry S Shryock, Jacob S Siegel & Associates, The Methods & Material of Demography, U.S. Bureau of the Census, U.S. Government Printing Office, Washington D.C. - Vol I & II, 1980
2. Deepti Shyam Sunder (2019): Fundamentals of Epidemiology and Biostatistics, CBS Publishers & Distributors.
3. Pathak, K.B.& F.RAM: Techniques of Demographic Analysis, Himalaya Publishing house, Mumbay,
4. K. Park (2013): Parks's Textbook of Preventive and Social Medicine, Banarasidas Bhanot Publishers, Jabalpur.
5. Shrivastava O.S. (1995): Demography and population Studies, Vikas Publishing house private limited, 2nd edition.
6. Bhende, Asha A and Tara Kanitkar: Principles of Population Studies – 5th rev. ed. Himalaya Publishers,Delhi, 1997
7. ORGI, UNFPA (2014): Training Manuals on Demographic Techniques
8. Pagano M, Gauvreau K (2010): Principles of Biostatistics. Cengage Learning India Pvt Ltd
9. Armitage, P., Statistical Methods in Medical Research, London, Blackwell Scientific
   Publications, 1989.

STA – 454: Advanced Design of Experiments

75 Lecture Hours **4 Credits**

Learning Objectives: To develop a systematic method to determine the relationship between factors affecting a process and the output of that process. To make the student understand how to design experiments to optimize processes, products, or systems by identifying significant factors and their interactions.

Course Outcomes: By the end of the course, students should be able to familiarize with incomplete block designs and their applications and also familiarize with factorial experiments for industrial and other uses.

Part – A (Theory)

UNIT-I:

Two-way classifications with m observation per cell under fixed, random and mixed effects models. Analysis of covariance (both, one and two ways).

(15 hours)

UNIT-II:

Incomplete Block Designs: Balanced Incomplete Block Design (BIBD) – parameters, relationships among its parameters, incidence matrix and its properties, Symmetric BIBD, Resolvable BIBD, Affine Resolvable BIBD, Intra Block analysis, complimentary BIBD, Residual BIBD, Dual BIBD, Derived BIBD.

(15 hours)

UNIT-III:

Factorial experiments: 3^2 factorial experiments, design and analysis. Complete and partial confounding under 2^3 factorial experiment. Total and Partial confounding for 2^n ($n \leq 3$), 3^2 and 3^3. Factorial experiments in a single replicate.

(15 hours)

Part – B (Practical)

UNIT-IV: Practical

(30 hours)

1. Problem on ANOVA for two-way classified data with m observations per cell.
2. Intra Block analysis of a BIBD
3. Problem on factorial experiments of 3^2.
4. Analysis of a completely confounded two level factorial design in 2 blocks
5. Analysis of a completely confounded two level factorial design in 4 blocks
6. Analysis of a partially confounded two level factorial design
7. Analysis of a single replicate of a 2^n design

Suggested readings:

23. Cochran, W.G. and Cox, G.M. (1959): Experimental Design. Asia Publishing House.
24. Das, M.N. and Giri, N.C. (2017): Design and Analysis of Experiments. Wiley Eastern Ltd.
25. Goon, A.M., Gupta, M.K. and Dasgupta, B. (2016): Fundamentals of Statistics. Vol. II, 8thEdn. World Press, Kolkata.
26. Kempthorne, O. (2017): The Design and Analysis of Experiments. John Wiley.
27. Montgomery, D. C. (2012): Design and Analysis of Experiments, John Wiley.
28. Joshi, D.D. (2020), "Linear Estimation and Design of Experiments", New Age International Publishers, ND.
29. Bhuyan, K. C. (2017): Designs of Experiment and Sampling Methods, New Central Book Agency (P). Ltd

STA-455: Time Series Analysis and Forecasting

75 Lecture Hours **4 Credits**

Learning Objectives: The course provides a comprehensive understanding of time series data and various methods to handle stationary and non-stationary time series data. Students will learn different types of time series model use for estimation and forecasting.

Course Outcomes: The learning objectives for time series analysis and forecasting include understanding time series data fundamentals, learning various models, exploring data preprocessing techniques, mastering trend analysis and seasonality detection, gaining proficiency in forecasting, understanding evaluation metrics, and applying these skills to real-world datasets for informed decision-making.

Part – A (Theory)

UNIT-I:

Time series data, decomposition of series, moving average, exponential smoothing and Holt-Winters method. 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 hours)

UNIT-II:

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

(15 hours)

UNIT-III:

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

(15 hours)

Part – B (Practical)

UNIT-IV: Practical (30 Hours)

(30 hours)

1. Correlogram Analysis and Interpretation
2. Smoothing and Forecasting using simple exponential
3. Forecasts using Holt and Winter model
4. Modeling and Forecasting with pure MA/AR models
5. Modeling and Forecasting with mixed ARMA models
6. Fitting and Forecasting with ARMA models
7. Modeling seasonal data using SARIMA and Forecasting
8. Modeling volatility using ARCH
9. Residual Analysis
10. Diagnostic checking

Suggested readings:

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.