**Maulana Abul Kalam Azad University of Technology, WB**
**(Formerly known as West Bengal University of Technology )**
**Syllabus of B.Sc. in Statistics**
**Effective from academic session 2021-22**

## Program outcomes for BSc in Statistics

After completion of the three years bachelor program in statistics,

**PO1.** Appreciate & articulate the symmetry & patterns that underlie & govern appearance & behaviour of objects & systems in natural as well as built environments.

**PO2.** Model real-world phenomena into statistical model formulations and develop the statistical models and subsequently interpret the solutions back to the real-world as applicable recommendations.

**PO3.** Deploy a rich *portfolio* of advanced statistical techniques using contemporary software tools for the purpose of solving real-life problems.

**PO4.** Demonstrate proficiency in the knowledge of statistical principles underlying archetypical problems in real-life application domains.

**PO5.** Demonstrate efficacy in the use of computational tools for solving problems, assessing risks, and exploring opportunities in real life applications.

**PO6.** Interact with fellow professionals of their own and other fields of science, engineering, and humanities through collaborative engagement.

**PO7.** Contribute their part in addressing 'big' scientific, technical, and societal issues which demand interdisciplinary as well as transdisciplinary efforts to solve and resolve.

**PO8.** Critically read, critique, and evaluate the merit of scientific & technical documents that concern application of mathematical science.

**PO9**. Write scientific & technical documents, such as essays, research & review articles, project proposals & reports, theses, etc. in the domain of statistical sciences and analytics.

**PO10**. Acquire on their own initiative new knowledge & skills of statistics to stay current in their profession.

**Program Specific Outcome:**

PSO-1: Acquire core knowledge and applications of the basic concepts of statistics which include the major areas of Probability theory, Probability distributions, Statistical inference, Survey sampling, Time series, Statistical quality control, Designs of experiments, Mathematical methods (Linear algebra, Numerical methods, Stochastic process), Non-parametric inference and Operations research.

PSO-2: Practical exercises accompanied with theoretical studies enable students to analyse, visualise and interpret data as well as to draw valid inferences out of them. The thrust areas of such practical exercises are real life and real time applications.

PSO-3: Build strong orientation towards enhancement of research ability through research oriented projects on real data sets.

PSO-4: Understand the applications of statistical concepts in other interdisciplinary areas like Mathematics, Physics, Economics, Bioscience etc.

PSO-5: Provides a strong platform for pursuing higher studies leading to Post Graduate or Research degrees as well as successful professional careers.

**Maulana Abul Kalam Azad University of Technology, WB**
**(Formerly known as West Bengal University of Technology )**
**Syllabus of B.Sc. in Statistics**
**Effective from academic session 2021-22**

**Curriculum Structure**

| Semester-I | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sl. No.** | **Category** | **Subject Code** | **Subject Name** | **Total no of contact hours** | | | **Credits** |
| | | | | L | T | P | |
| 1 | Core Course 1 | BSTAT101 | Descriptive Statistics | 3 | 1 | 0 | 4 |
| 2 | Laboratory 1 | BSTAT191 | Laboratory for Descriptive Statistics | 0 | 0 | 4 | 2 |
| 3 | Core Course 2 | BSTAT102 | Basic Probability | 5 | 1 | 0 | 6 |
| 4 | Generic Elective 1 | BSTAT103 | Any one from GE Basket | | | | 6 |
| 5 | Ability Enhancement Compulsory Course (AECC1) (Communicative English) | BSTAT 104 | Speak English Professionally: In Person, Online & On the Phone | 2 | 0 | 0 | 2 |
| **Total of Semester-I** | | | | | | | **20** |
| **Semester-II** | | | | | | | |
| **Sl. No.** | **Category** | **Subject Code** | **Subject Name** | **Total no of contact hours** | | | **Credits** |
| | | | | L | T | P | |
| 1 | Core Course 3 | BSTAT 201 | Application of Probability in Real Life | 3 | 1 | 0 | 4 |
| 2 | Laboratory 1 | BSTAT 291 | Laboratory for Probability | 0 | 0 | 4 | 2 |
| 3 | Core Course 4 | BSTAT 202 | Mathematical Analysis | 5 | 1 | 0 | 6 |
| 4 | Generic Elective 2 | BSTAT 203 | Any one from GE Basket | | | | 6 |
| 5 | Ability Enhancement Compulsory Course (AECC 2) (Environment & Sustainability) | BSTAT 204 | Environmental Science | 2 | 0 | 0 | 2 |
| **Total of Semester-II** | | | | | | | **20** |

**Maulana Abul Kalam Azad University of Technology, WB**
**(Formerly known as West Bengal University of Technology )**
**Syllabus of B.Sc. in Statistics**
**Effective from academic session 2021-22**

| **Semester III** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sl. No.** | **Category** | **Subject Code** | **Subject Name** | **Total no of contact hours** | | | **Credits** |
| | | | | L | T | P | |
| 1 | Core Course 5 | BSTAT 301 | Sampling Distributions | 5 | 1 | 0 | 6 |
| 2 | Core Course 6 | BSTAT 302 | Statistical Inference | 3 | 1 | 0 | 4 |
| 3 | Laboratory 1 | BSTAT 392 | Laboratory for Statistical Inference | 0 | 0 | 4 | 2 |
| 4 | Core Course 7 | BSTAT 303 | Linear Algebra | 5 | 1 | 0 | 6 |
| 5 | Generic Elective 3 | BSTAT 304 | Any one from GE Basket | | | | 6 |
| 6 | Skill Enhancement Course 1 (SEC–1) | BSTAT 305A/B | Python for Data Analysis/Statistical Data Analysis using R | 2 | 0 | 0 | 2 |
| **Total of Semester-III** | | | | | | | **26** |
| **Semester IV** | | | | | | | |
| **Sl. No.** | **Category** | **Subject Code** | **Subject Name** | **Total no of contact hours** | | | **Credits** |
| | | | | L | T | P | |
| 1 | Core Course 8 | BSTAT 401 | Survey Sampling | 3 | 1 | 0 | 4 |
| 2 | Laboratory 1 | BSTAT 491 | Laboratory for Survey Sampling | 0 | 0 | 4 | 2 |
| 3 | Core Course 9 | BSTAT 402 | Statistical Quality Control | 3 | 1 | 0 | 4 |
| 4 | Laboratory 2 | BSTAT 492 | Laboratory for Statistical Quality Control | 0 | 0 | 4 | 2 |
| 5 | Core Course 10 | BSTAT 403 | Linear Model | 5 | 1 | 0 | 6 |
| 6 | Generic Elective 4 | BSTAT 404 | Any one from GE Basket | | | | 6 |
| 7 | Skill Enhancement Course 2 (SEC-2) | BSTAT 405A/B | Statistical Techniques for Research Methods/Database Management System | 2 | 0 | 0 | 2 |
| **Total of Semester-IV** | | | | | | | **26** |

**Maulana Abul Kalam Azad University of Technology, WB**
**(Formerly known as West Bengal University of Technology )**
**Syllabus of B.Sc. in Statistics**
**Effective from academic session 2021-22**

| **Semester V** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sl. No.** | **Category** | **Subject Code** | **Subject Name** | **Total no of contact hours** | | | **Credits** |
| | | | | L | T | P | |
| 1 | Core Course 11 | BSTAT 501 | Stochastic Process and Queuing Theory | 3 | 1 | 0 | 4 |
| 2 | Laboratory 1 | BSTAT 591 | Laboratory for Stochastic Process and Queuing | 0 | 0 | 4 | 2 |
| 3 | Core Course 12 | BSTAT 502 | Modern Statistical Techniques | 3 | 1 | 0 | 4 |
| 4 | Laboratory 2 | BSTAT 592 | Laboratory for Modern Statistical Techniques | 0 | 0 | 4 | 2 |
| 5 | Discipline Specific Elective 1 | BSTAT 503A/B | Time Series Analysis/Demography and Vital Statistics | 5 | 0 | 1 | 6 |
| 6 | Discipline Specific Elective 2 | BSTAT 504A/B/C | Econometrics/Financial Statistics/Actuarial Statistics | 5 | 1 | 0 | 6 |
| **Total of Semester-V** | | | | | | | **24** |

| **Semester VI** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sl. No.** | **Category** | **Subject Code** | **Subject Name** | **Total no of contact hours** | | | |
| | | | | L | T | P | |
| 1 | Core Course 13 | BSTAT 601 | Design of Experiments | 3 | 1 | 0 | 4 |
| 2 | Laboratory 1 | BSTAT 691 | Laboratory for Design of Experiments | 0 | 0 | 4 | 2 |
| 3 | Core Course 14 | BSTAT 602 | Multivariate Analysis and Nonparametric Methods | 5 | 1 | 0 | 6 |
| 4 | Discipline Specific Elective 3 | BSTAT 603A/B | Survival Analysis and Biostatistics/Operations Research | 5 | 0 | 1 | 6 |
| 5 | Research Ability Enhancement Courses (RAEC) | BSTAT 681 | Capstone Project* | 0 | 0 | 6 | 6 |
| **Total of Semester-VI** | | | | | | | **24** |

*Research work in the form of dissertation/project work preferably in interdisciplinary/multidisciplinary/trans-disciplinary areas.

**Detailed Syllabus for Semester I**

| **CC 1: BSTAT 101 Descriptive Statistics** | **Credit 4**<br>**Marks 100** |
|---|---|

**Course Outcome:**
After completion of the course the students will be able

1. To learn different data types and their proper uses along with data handling, data representation, descriptive measures and exploratory analysis,
2. To analyse using measures of central tendency, dispersion and location for interpretation of statistical data,
3. To apply the concept of Principle of least squares for curve fitting and regression lines.
4. To learn about bivariate distributions as a closure approach to real world data and study the relationships between them,
5. To select different uses of index number techniques as a basic tool for synthesising economic statistics.

**UNIT I** **10L**

Statistical Methods: Definition and scope of Statistics, concepts of statistical population and sample. Data: quantitative and qualitative, attributes, variables, scales of measurement- nominal, ordinal, interval and ratio. Presentation: tabular and graphical, including histogram and ogives, consistency and independence of data with special reference to attributes.

**UNIT II** **10L**

Measures of Central Tendency: mathematical and positional. Measures of Dispersion: range, quartile deviation, mean deviation, standard deviation, coefficient of variation, Moments, absolute moments, factorial moments, skewness and kurtosis, Sheppard's corrections.

**UNIT III** **10L**

Bivariate data: Definition, scatter diagram, simple, partial and multiple correlation (3 variables only), rank correlation. Simple linear regression, principle of least squares and fitting of polynomials and exponential curves.

**UNIT IV** **10L**

Index Numbers: Definition, construction of index numbers and problems thereof for weighted and unweighted index numbers including Laspeyre's, Paasche's, Edgeworth-Marshall and Fisher's. Chain index numbers, conversion of fixed based to chain-based index numbers and vice-versa. Consumer price index numbers.

**SUGGESTED READING:**

1. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edn. The World Press, Kolkata.
2. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
3. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co. Ltd.

| **Lab 1: BSTAT 191: Laboratory for Descriptive Statistics** | **Credit 2**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able,

1. To learn the basics of data handling, data representation, descriptive measures and exploratory analysis,
2. To apply mathematical ideas into statistical data analysis, interpretation of data,
3. To perform data analysis based on the statistical measures and interpreting the outcomes of the study,
4. To learn the interrelationship between location, scale and shape problems of data distribution and situational application of various statistical measures,
5. To understand the mechanism of statistical decision making through exploratory data analysis,
6. To collect real data from the field of enquiry and prepare complete analytical findings of inner stories of a situation.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Graphical representation of data.
2. Problems based on measures of central tendency.
3. Problems based on measures of dispersion.
4. Problems based on combined mean and variance and coefficient of variation.
5. Problems based on moments, skewness and kurtosis.
6. Fitting of polynomials, exponential curves.
7. Karl Pearson correlation coefficient.
8. Correlation coefficient for a bivariate frequency distribution.
9. Lines of regression, angle between lines and estimated values of variables.
10. Spearman rank correlation with and without ties.
11. Partial and multiple correlations.
12. Planes of regression and variances of residuals for given simple correlations.
13. Planes of regression and variances of residuals for raw data.
14. Calculate price and quantity index numbers using simple and weighted average of

price relatives.
15. To calculate the Chain Base index numbers.
16. To calculate the consumer price index number.

| **CC 2: BSTAT 102 Basic Probability** | **Credit 6**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to

1. Understand the idea of deterministic experiment and random experiment.
2. Apply the concept of probability as measure of uncertainty in practical cases
3. Understand the idea of prior, posterior and Bayes' theorem
4. Learn the idea of random variables and its applications.

**UNIT I** **25L**

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability – classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, Theorem of total probability, Bayes' theorem and its applications.

**UNIT II** **20L**

Random variables: discrete and continuous random variables, p.m.f, p.d.f. and c.d.f., illustrations and properties of random variables. Expectation and variance of random variables.

**UNIT III** **15L**

Jensen's Inequality, Cauchy Schwarz Inequality, Markov's inequality, Chebyshev's Inequality, Chernoff's Inequality, One Sided Chebyshev's Inequality.

**References**

1. Chung K.L. (1983): Elementary Probability Theory with Stochastic Process, Springer / Narosa
2. Feller W. (1968): An Introduction to Probability Theory & its Applications, John Wiley
3. Goon A.M., Gupta M.K. & Dasgupta B. (1994): An Outline of Statistical Theory (Vol-1), World Press
4. Rohatgi V.K. (1984): An Intro. to Probability Theory & Math. Statistics, John Wiley

5. Hoel P.J., Port S.C. & Stone C.J.: Introduction to Probability Theory (Vol-1), Mifflin & UBS
6. Cramer H. (1954): The Elements of Probability Theory, John Wiley
7. Parzen E. (1972): Modern Probability Theory and its Applications, John Wiley

| **AECC 1: BSTAT 104 Communicating English** | **Credit 2**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to

1. Talk in English so that he/she could express their views on any topic without any difficulty in speech,
2. Use English effectively during the entire course curriculum and enable the learner to communicate effectively and appropriately in real life situations,
3. Develop and demonstrate the speaking skills for group discussions.

**Module 1** **5L**

Small Talk & Conversational, Vocabulary; Express Yourself: Pronunciation; Elevator Speech

**Module 2** **5L**

Video Conferencing: Face to Face but Online; Group Discussion Language; Video

Conference Role Play

**Module 3** **5L**

Telephone Language; Understand and Be Understood on the Phone; Phone Role Play

**Module 4** **5L**

Get Ready for the Interview; Improve Your Pronunciation

**References:**

1. S R Inthira & amp, V Saraswathi, Enrich your English a) Communication skills b) Academic skills, CIEFL &amp, OUP
2. R.C. Sharma and K. Mohan Business Correspondence and Report Writing Tata McGraw Hill , New Delhi , 1994
3. Maxwell Nurnberg and Rosenblum Morris, All About Words- A Text Book for English for Engineers &amp, Technologists General Book Depot, New Delhi, 1995

## Detailed Syllabus for Semester II

| **CC 3: BSTAT 201 Application of Probability in Real Life** | **Credit 4**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Understand different probability models in different real-life situations,
2. Build different probability models applicable for different types of data set,
3. Learn about important probability distributions,
4. Make use of moments and moment generating functions,
5. Use data generated from different probability distributions.

**UNIT I** **15L**

Two dimensional random variables: Discrete and Continuous type, Joint, Marginal and Conditional p.m.f, p.d.f., and c.d.f., Independence of Variables, Bivariate transformations with illustrations. Wald's Equation. Applications of Bivariate random vector in real situations.

**UNIT II** **15L**

Mathematical Expectation and Generating Functions: Expectation of single and Bivariate random variables and its properties. Moments and Cumulants, Moment Generating Function, Cumulant Generating Function and Characteristic Function. Uniqueness and Inversion theorems (without proof) along with applications. Conditional Expectations.

**UNIT III** **10L**

Standard probability distributions: Binomial, Poisson, Geometric, Negative binomial, Hypergeometric, Uniform, Exponential, Normal, Cauchy, Beta and Gamma along with their properties and limiting/approximation cases.

**References:**

1. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
2. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
3. Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi

| **Lab 1: BSTAT 291 Laboratory for Application of Probability in Real Life** | **Credit 2**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Model a data set based on appropriate probability distribution,
2. Apply probability to carry out decision analysis for real life data,
3. Understand the applicability of normality assumption across different data sets.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Fitting of binomial distributions for n and p = q = 1⁄2.
2. Fitting of binomial distributions for given n and p.
3. Fitting of binomial distributions after computing mean and variance.
4. Fitting of Poisson distributions for given value of lambda.
5. Fitting of Poisson distributions after computing mean.
6. Fitting of negative binomial.
7. Fitting of suitable distribution.
8. Application problems based on binomial distribution.
9. Application problems based on Poisson distribution.
10. Application problems based on negative binomial distribution.
11. Problems based on area property of normal distribution.
12. To find the ordinate for a given area for normal distribution.
13. Application based problems using normal distribution.
14. Fitting of normal distribution when parameters are given.
15. Fitting of normal distribution when parameters are not given.
16. Gambler's ruin problem, Drunkard Step problem

| **CC 4: BSTAT 202: Mathematical Analysis** | **Credit 6**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Describe the fundamental properties of the real numbers that underpin the formal development of real analysis,

2. Demonstrate an understanding of the theory of sequences and series, continuity, differentiation and integration,
3. Apply and analyse rigorous mathematical arguments,
4. Apply procedures like interpolation, Numerical integration and solution of Differential equation in different application areas,
5. Analyse and evaluate the accuracy of common numerical methods.

**UNIT-I** **15L**

Real Analysis: Representation of real numbers as points on the line and the set of real numbers as complete ordered fields. Bounded and unbounded sets, Neighbourhoods and limit points, Supremum and infimum, Derived sets, Open and closed sets, Sequences and their convergence, Limits of some special sequences such as and Cauchy's general principle of convergence, Cauchy's first theorem on limits, Monotonic sequences, Limit superior and limit inferior of a bounded sequence.

**UNIT-II** **15L**

Infinite series, Series with positive terms and their convergence, Comparison test, D'Alembert's ratio test, Cauchy's nth root test, Raabe's test, Gauss test, Cauchy's condensation test and Integral test (Statements and Examples only). Absolute convergence of series, Leibnitz's test for the convergence of alternating series, Conditional convergence.

**UNIT-III** **15L**

Numerical Analysis: Approximations and Errors, Factorial notation, Operators: Shift, Forward, Backward, Central differences, Divided difference. Interpolation: Newton's forward, backward and divided differences interpolation, Lagrange's interpolation. Gauss and Stirling interpolation formulae.

**UNIT-IV** **15L**

Numerical integration. Trapezoidal rule, Simpson's $1/3^{rd}$ rule, $3/8^{th}$ rule, Weddle's rule with corresponding error terms, Solution of differential equations of first order: Euler's and Modified Euler's method, Runge- Kutta Method ($4^{th}$ order).

**References:**

1. Malik S.C. and Savita Arora: Mathematical Analysis, Second Edition, Wiley Eastern Limited, New Age International Limited, New Delhi, 1994.
2. Somasundram D. and Chaudhary B.: A First Course in Mathematical Analysis, Narosa Publishing House, New Delhi, 1987.
3. Gupta S.L. and Nisha Rani: Principles of Real Analysis, Vikas Publ. House Pvt. Ltd., New Delhi, 1995.
4. Apostol T.M.: Mathematical Analysis, Second Edition, Narosa Publishing House, NewDelhi, 1987.
5. Shanti Narayan: A course of Mathematical Analysis, 12th revised Edition, S. Chand & Co. (Pvt.) Ltd., New Delhi, 1987.

6. Bartle, R. G. and Sherbert, D. R., Introduction to Real Analysis (3rd Edition), John Wiley and Sons (Asia) Pte. Ltd., Singapore, 2002.
7. Jain,M. K., Iyengar, S. R. K. and Jain, R. K., Numerical methods for scientific and engineering computation, New age International Publisher, India, 2003.
8. Sastry, S.S. (2000): Introductory Methods of Numerical Analysis, 3rd edition, Prentice Hall of India Pvt. Ltd., New Delhi.
9. Gupta, A. and Bose, S.C., Introduction to Numerical Analysis, Academic Publishers, 2009.

**Lab 2: BSTAT 292: Laboratory for Introduction to Computer Programming using Python**

**Credit 2**
**Marks 100**

**Course Objectives:**

**After completion of the course the students will be able to**

1. Write, test, and debug simple Python programs.
2. Implement Python programs with conditionals and loops.
3. Use functions for structuring Python programs.
4. Read and write data from/to files in Python.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Running Code in the Interactive Shell, Input, Processing and Output, Editing, Saving and Running a Script

2. Values and Data Types, Variables, Keywords, String Literals, Escape Sequences, Operators and Operands, Expressions and Statements

3. Arithmetic, Relational Operators and Comparison Operators, Python Assignment Operators Shorthand Assignment Operators, Logical Operators or Bitwise Operators
4. String operations and indices Basic String Operations -- String Functions, Methods, Delete a string , String Multiplication and concatenation, String Formatting Operator, Python String replace() Method, Changing upper and lower case strings, Using “join” function for the string, Reversing String, Split Strings

5. If Statement, If-Elif statements, For loop, While Statement, Infinite loops, Break and Continue Statements.

| **AECC 2: BSTAT 204: Environmental Science** | **Credit 2**<br>**Marks 100** |
|---|---|

**Course Outcome:**
After completion of the course the students will be able to

1. Understand and evaluate the global scale of environmental problems,
2. Reflect critically on their roles, responsibilities, and identities as citizens, consumers and environmental actors in a complex, interconnected world,
3. Use critical thinking, problem-solving, and the methodological approaches of the social sciences, natural sciences, and humanities in environmental problem solving.

**Unit I** **5L**

Introduction to environmental studies & ecosystems: Multidisciplinary nature of environmental studies: Scope and importance; what is an ecosystem? The structure and function of the ecosystem, Energy flow in an ecosystem, food chains, food webs and ecological succession, forest ecosystem, grassland ecosystem, desert ecosystem, aquatic ecosystems.

**Unit II** **5L**

Natural resources & its management and conservation: Land resources and land use change: Land degradation, soil erosion and desertification; Deforestation: Causes and impacts, forests, biodiversity and tribal populations; Water: Use and over-exploitation of surface and ground water. Energy resources: Renewable and non-renewable energy sources, use of alternate energy sources and growing energy needs.

**Unit III** **5L**

Environmental pollution & management: Environmental pollution: types, causes, effects and controls; Air, water, soil and noise pollution, Solid waste management: Control measures of urban and industrial waste. Climate change, global warming, Environment Laws: Environment Protection. Act, Air (Prevention & Control of Pollution) Act, Water (Prevention and control of pollution) Act, Wildlife Protection Act, Forest Conservation Act; International agreements: Montreal and Kyoto protocols and Convention on Biological Diversity (CBD).

**Unit IV** **5L**

Environment & social issues: Human population growth: Impacts on environment, human health and welfare; Resettlement and rehabilitation of project affected persons; case studies; Disaster management: floods, earthquake, cyclones and landslides;

Environmental ethics: environmental conservation; environmental communication and public awareness.

**References:**

1. Carson, R. Silent Spring. Houghton Mifflin Harcourt, 2002.
2. Gadgil, M., & Guha, R. This Fissured Land: An Ecological History of India. Univ. of California Press, 1993.
3. Gleeson, B. and Low, N. (eds.). Global Ethics and Environment, London, Routledge, 1999.
4. Gleick, P. H. Water in Crisis. Pacific Institute for Studies in Dev., Environment Security. Stockholm Env. Institute, Oxford Univ. Press, 1993.
5. Groom, Martha J., Gary K. Meffe, and Carl Ronald Carroll. Principles of Conservation Biology. Sunderland: Sinauer Associates, 2012.

## Detailed Syllabus for Semester III

| **CC 5: BSTAT 301: Sampling Distribution** | **Credit 6**<br>**100 marks** |
|---|---|

**Course Outcome**

After completion of the course the students will be able to,

1. Understand the function of random variables and their distribution.
2. Find out the interrelationships of different probability distributions.
3. Make use of the transformations in solving inferential problems.
4. Apply the family of distributions in real life situations.

**UNIT I** 20 L

Definitions of random sample, parameter and statistic, sampling distribution of a statistic, sampling distribution of sample mean, standard errors of sample mean, sample variance and sample proportion. Null and alternative hypotheses, level of significance, Type I and Type II errors, their probabilities and critical region. Large sample tests, use of CLT for testing single proportion, difference of two proportions, single mean, difference of two means, standard deviation and difference of standard deviations by classical and p-value approaches.

**UNIT II** 15 L

Exact sampling distribution: Definition and derivation of p.d.f. of $\chi 2$ with n degrees of freedom (d.f.) using m.g.f., nature of p.d.f. curve for different degrees of freedom, mean, variance, m.g.f., cumulant generating function, mode, additive property and limiting form of $\chi 2$ distribution.

**UNIT III** 15 L

Exact sampling distributions: Student's and Fisher's t-distribution, Derivation of its p.d.f., nature of probability curve with different degrees of freedom, mean, variance, moments and limiting form of t distribution.
Snedecore's F-distribution: Derivation of p.d.f., nature of p.d.f. curve with different degrees of freedom, mean, variance and mode. Distribution of 1/F(n1,n2). Relationship between t, F and $\chi 2$ distributions. Test of significance and confidence Intervals based on t and F distributions.

**UNIT IV** 10 L

Order statistics and its distributions. Sampling from bivariate distribution. Sampling distribution of correlation coefficient (Null Case) and least square estimators of linear regression.

**References:**

1. Hogg, R.V. and Tanis, E.A. (2009): A Brief Course in Mathematical Statistics. Pearson Education.
2. Johnson, R.A. and Bhattacharya, G.K. (2001): Statistics-Principles and Methods, 4th Edn. John Wiley and Sons.
3. Mood, A.M., Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn. (Reprint).Tata McGraw-Hill Pub. Co. Ltd.

| **CC 6: BSTAT 302: Statistical Inference** | **Credit 4**<br>**100 marks** |
|---|---|

**Course Outcome:**
After completion the course students will be able to,

1. Explain the statistical inference and its types,
2. Understand the desirable properties of an estimator and measures of finding a good estimate from sample data for practical use.
3. Determine various optimum tests for examining a hypothesis regarding a population parameter
4. Apply the idea for carrying out research to analyse real-life data.

**UNIT I** **10 L**

Estimation: Concepts of estimation, unbiasedness, sufficiency, consistency and efficiency. Factorization theorem. Complete statistics, Minimum variance unbiased estimator (MVUE), Rao-Blackwell and Lehmann-Scheffe theorems and their applications. Cramer-Rao inequality and MVB estimators (statement and applications).

**UNIT II** **10 L**

Methods of Estimation: Method of moments, method of maximum likelihood estimation, method of minimum Chi-square, basic idea of Bayes estimators.

**UNIT III** **10 L**

Principles of test of significance: Null and alternative hypotheses (simple and composite), Type-I and Type-II errors, critical region, level of significance, size and power, best critical region, most powerful test, uniformly most powerful test, Neyman Pearson Lemma (statement and applications to construct most powerful test). Likelihood ratio test, properties of likelihood ratio tests (without proof).

**UNIT IV** **10 L**

Sequential Analysis: Sequential probability ratio test (SPRT) for simple vs simple hypotheses. Approximation of boundaries based on strength of test. Wald's fundamental identity and the derivation of operating characteristics (OC) and average sample number

(ASN) functions, examples based on Normal, Poisson, Binomial and Exponential distributions.

**References:**

1. Goon A.M., Gupta M.K.: Das Gupta. B. (2005), Fundamentals of Statistics, Vol. I, World Press, Calcutta.
2. Rohatgi V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
3. Miller, I. and Miller, M. (2002): John E. Freund's Mathematical Statistics (6th edition, low price edition), Prentice Hall of India.
4. Dudewicz, E. J., and Mishra, S. N. (1988): Modern Mathematical Statistics. John Wiley & Sons.
5. Mood A.M, Graybill F.A. and Boes D.C (1974): Introduction to the Theory of Statistics, McGraw Hill.
6. Bhat B.R, Srivenkatramana T and Rao Madhava K.S. (1997) Statistics: A Beginner's Text, Vol. I, New Age International (P) Ltd.
7. Snedecor G.W and Cochran W.G.(1967) Statistical Methods. Iowa State University Press.

| **CC 7: BSTAT 303: Linear Algebra** | **Credit: 6**<br>**100 marks** |
|---|---|

**Course Outcome:**
After completion of the course the students will be able to,

1. Make use of computational techniques and algebraic skills essential for the study of systems of linear equations, matrix algebra, vector spaces, eigenvalues and eigenvectors, orthogonality.

2. Learn the diagonalizability of a square matrix and its corresponding applications.

3. Develop knowledge and ability for visualisation, spatial reasoning, as well as geometric properties and strategies to model, solve problems and view solutions, especially in $\square^2$, $\square^3$ and extending these results to higher dimensions.

4. Determine the dimension of row or column space of a matrix with application.

**Unit I:** **15 L**

Definition of vectors, operation of vectors (angle, distance etc.). Vector spaces, Subspaces, sum of subspaces, Span of a set, Linear dependence and independence, dimension and basis, dimension theorem. Extension of basis. Orthogonal vectors, Gram-Schmidt Orthogonalization.

**Unit II:** **15 L**

Algebra of matrices. Linear transformation. Elementary matrices and their uses, theorems related to triangular, symmetric and skew symmetric matrices, Hermitian and skew Hermitian matrices, Idempotent matrices, Orthogonal matrices.

**Unit III: 15L**

Determinant of Matrices: Definition, properties and applications of determinants for 3rd and higher orders, evaluation of determinants, using transformations. Symmetric and Skew symmetric determinants, product of determinants. Use of determinants in solution to the system of linear equations. Adjoint and inverse of a matrix and related properties. Singular and non-singular matrices and their properties. Conditions for consistency, uniqueness, infinite solutions, solution sets of linear equations, linear independence, Applications of linear equations.

**Unit IV: 15L**

Rank of a matrix, row-rank, column-rank, standard theorems on ranks, rank of the sum and the product of two matrices. Partitioning of matrices and simple properties. Characteristic roots and Characteristic vectors, Properties of characteristic roots, Quadratic forms: Classification & Canonical reduction. Linear orthogonal transformation, Matrix diagonalization.

**References:**

1. Linear Algebra and its Application by Gilbert Strang, 4th edition, Cengage Learning along with MIT Lecture videos by the Author
2. Introduction to Applied Linear Algebra: Vector, Matrices and Least Squares by Stephen Boyd and Lieven Vandenberghe, 2018 edition, Cambridge University Press.
3. Matrix Computations, G. H. Golub and C. F. Van Loan, 3rd Edition, John Hopkins University Press, 1996
4. Hadley, G (2002) : Linear Algebra. Narosa Publishing House (Reprint).
5. Applied Linear Algebra by Peter J. Olver and Chehrzad Shakiban, second Edition, Springer

| **SEC 1: BSTAT 305 A: Python for Data analysis** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Understand the principles of Python and acquire skills in programming in Python.
2. Understand the use of fundamental Python libraries to perform statistical analysis.
3. Generate meaningful visualisation from data.
4. Manipulate, summarise, analyse and clean data.
5. Develop statistical models to solve problems arising in real life scenarios.

**Unit I** **5L**

**Review of Basic concepts of Python.**
Fundamentals of Data Science and role of Python as a programming tool, Analytics Life Cycle, Installing and Getting Started, Data Types and Operations, Control Flow and Functions**.**

**Unit II** **10L**

**File Handling and Data Generation:** Open, read and write data to files of different formats, Interact with raw inputs from users, Understanding Random module and generate synthetic data.
**Numpy:** Create, access, modify, and sort multidimensional Numpy arrays, Different manipulation and operations on Numpy arrays
**Pandas**: Create, access, and modify Pandas Series and Data frames, Perform arithmetic operations on Series and Data frame, load data into a Data frame.

**Unit III** **10L**

**Data Visualisation and Exploratory Data Analysis (EDA):** Visualise data using Matplotlib and Seaborn, Perform EDA and Data profiling.
**Data Processing:** Scaling and Transforming data, Encoding, Handling Missing observations.
**Model Development:** Develop a Regression model, Regression Diagnostics, Evaluating model performance.

* The curriculum is to be complemented by corresponding computer software and programming.

**References:**

1. Mark Lutz, "*Learning Python Powerful Object-Oriented Programming*", O'reilly Media, 2018, 5th Edition.
2. Wes McKinney, "*Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*", O'reilly Media, 2nd Edition, 2018.
3. Cole N Knaflic, "*Storytelling with Data: A Data Visualization Guide for Business Professionals"*, 1st Edition, John Willey & Sons, 2015.
4. John P Mueller, Luca Massaron, *"Machine Learning for Dummies"*, John Willey & Sons, 2016

| **SEC 1: BSTAT 305 B: Statistical data analysis using R** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,
1. Understand the software for ready use

2. Use the software R for basic data analytic tool
3. Apply R programing language to cross-check theoretical results
4. Adopt programing skill as fundamental industry requirement

**Unit 1** **5L**

Introduction to R: Installation, command line environment, overview of capabilities, brief mention of open source Philosophy. R as a calculator: The four basic arithmetic operations. Use of parentheses nesting up to arbitrary level. The power operation. Evaluation of simple expressions. Quotient and remainder operations for integers. Standard functions, e.g., sin, cos, exp, log.

**Unit 2** **5L**

The different types of numbers in R: Division by zero leading to Infor -Inf, NaN, NA. Variables, Creating a vector using c(), seq() and colon operator. How functions map over vectors. Functions to summarise a vector: sum, mean, sd, median etc. Extracting a subset from the vector (by index, by property). R as a graphing calculator: Introduction to plotting. Plot(), lines(), abline(). Barplot, Pie chart and Histogram. Box plot. Scatter plot and simple linear regression using lm(y~x).

**Unit 3** **5L**

Matrix operations in R: Creation. Basic operations. Extracting submatrices. Loading data from a file: read.table() and read.csv(). Mention of head=TRUE and head=FALSE. Dataframes. Mention that these are like matrices, except that different columns may be of different types.

**Unit 4** **5L**

Problems on discrete and continuous probability distributions.

*The curriculum is to be complemented by corresponding computer software and programming.

**References:**

1. Verma A. K. , R Programming , Cengage Publication, 2017
2. Garrett Grolemund, Hands on Programming with R, O'reilly Publications.
3. Hadley Wickham & Garrett Grolemund, R for data science, O'reilly Publications
4. Trevor Hastie and Rob Tibshirani, An introduction to Statistical learning with applications in R.

| **Lab 1: BSTAT 391: Laboratory for Statistical Inference** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Apply theory of point estimation for analysing sample observation.

2. Experiment with the notions of testing hypotheses in various fields e.g., biomedical, bio-science, geology, geography, behavioural science etc.
3. Determine sequential testing procedures under varying sample size.
4. Evaluate the risk associated with different inferential methods.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Unbiased estimators (including unbiased but absurd estimators)
2. Consistent estimators, efficient estimators and relative efficiency of estimators.
3. Cramer-Rao inequality and MVB estimators
4. Sufficient Estimators – Factorization Theorem, Rao-Blackwell theorem, Complete Sufficient estimators
5. Lehman-Scheffe theorem and UMVUE
6. Maximum Likelihood Estimation
7. Asymptotic distribution of maximum likelihood estimators
8. Estimation by the method of moments, minimum Chi-square
9. Type I and Type II errors
10. Most powerful critical region (NP Lemma)
11. Uniformly most powerful critical region
12. Unbiased critical region
13. Power curves
14. Likelihood ratio tests for simple null hypothesis against simple alternative hypothesis
15. Likelihood ratio tests for simple null hypothesis against composite alternative hypothesis
16. Asymptotic properties of LR tests
17. SPRT procedure
18. OC function and OC curve
19. ASN function and ASN curve

**Lab 2: BSTAT 392: Laboratory for Mathematics Methods for Computing** **Credit 2**
**Marks 100**

**Course Outcome:**
After completion of the course the students will be able to,

1. Derive unknown functions at least approximately using different interpolation formulas.
2. Analyse the accuracy for common numerical methods.
3. Find numerical solutions for different IVP and BVPs for ODE.
4. Estimate the solutions to a set of linear and nonlinear equations.

5. Build models based on the data through interpolation.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems on:

1. Applications of Lagrange's, Inverse interpolation.
2. Applications of Hermite and Spline Interpolations.
3. Approximation of a function by Chybeshev's polynomial
4. Solution of nonlinear equations using methods of bisection & Regula falsi
5. Use of Newton-Raphson and its modified version
6. Determination of Eigen value and eigen vector of a matrix
7. Finite difference method for ODE
8. Solution of IVP and BVP using finite difference methods of an ODE
9. Explicit and Implicit methods for numerical solutions of PDE

## Detailed Syllabus for Semester IV

| **CC 8: BSTAT 401: Survey Sampling** | **Credit: 4**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to

1. Understand the idea of population, sample and sampling errors.
2. Learn about the selection procedure of a sample from a target population.
3. Learn about different techniques of survey sampling.
4. Estimate different population parameters, variances of these estimates under different survey sampling techniques.
5. Learn about the sample size determination and understand the idea of sub sampling.

**UNIT I** **10L**

Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, definition and procedure of selecting a sample, estimation of population mean, total and proportion, variance of these estimators, estimates of their variances and sample size determination.

**UNIT II** **15L**

Stratified random sampling: Technique, estimation of population mean and total, variances of these estimators, proportional and optimum allocations and their comparison with SRS. Practical difficulties in allocation, estimation of gain in precision, post stratification and its performance. Systematic Sampling: Technique, estimation of population mean and total, variances of these estimators under linear systematic sampling. Comparison of systematic sampling with SRS and stratified sampling in the presence of linear trend and corrections.

**UNIT III** **15L**

Introduction to Ratio and regression methods of estimation, first approximation to the population mean and total (for SRS of large size), variances of these estimators and estimates of these variances, variances in terms of correlation coefficient for regression method of estimation and their comparison with SRS.

Cluster sampling (equal clusters only) estimation of population mean and its variance, comparison (with and without randomly formed clusters). Relative efficiency of cluster sampling with SRS in terms of intra class correlation. Concept of sub sampling.

**References:**

1. Cochran W.G. (1984): Sampling Techniques (3rd Ed.), Wiley Eastern.
2. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. Asok,C.(1984). Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics
3. Murthy M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub.Society, Calcutta.
4. Des Raj and Chandhok P. (1998): Sample Survey Theory, Narosa Publishing House.
5. Goon A.M., Gupta M.K. and Dasgupta B. (2001): Fundamentals of Statistics (Vol.2),World Press.
6. Guide to current Indian Official Statistics, Central Statistical Office, GOI, New Delhi.

| **CC 9: BSTAT 402: Statistical Quality Control** | **Credit 4**<br>**100 marks** |
|---|---|

**Course Outcome:**

On completion of the course, students will be able to:

1. Understand the general idea of quality and monitoring of industrial experiments.
2. Understand basic difference between process control and product control
3. Adapt to control chart techniques and acceptance of sampling plans.
4. Make use of six-sigma methodology.

**UNIT I** **10L**

Introduction to statistical quality control, Quality system and standards: Introduction to ISO quality standards, Quality registration, Statistical Process Control - Seven tools of SPC, chance and assignable causes of quality variation, Statistical Control Charts- Construction and Statistical basis of 3-σ Control charts, Rational Sub-grouping.

**UNIT II** **10L**

Control charts for variables: X-bar & R-chart, X-bar & s-chart. Control charts for attributes: np-chart, p-chart, c-chart and u-chart. Comparison between control charts for variables and control charts for attributes. Analysis of patterns on control chart, estimation of Process Capability.

**UNIT III** **10L**

Acceptance sampling plan: Principle of acceptance sampling plans. Single and Double sampling plan for attributes and their OC, AQL, LTPD, AOQ, AOQL, ASN, ATI functions with graphical interpretation, use and interpretation of Dodge and Romig sampling inspection plan Tables.

**UNIT IV** **10L**

Introduction to Six-Sigma: Overview of Six Sigma, Lean Manufacturing and Total Quality Management (TQM). Organisational Structure and Six Sigma training plans- Selection Criteria for Six-Sigma roles and training plans.

**References:**

1. Montgomery, D. C. (2009): Introduction to Statistical Quality Control, 6th Edition,

Wiley India Pvt. Ltd.
2. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edn. The World Press, Kolkata.
3. Mukhopadhyay, P (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied(P) Ltd.
4. Montgomery, D. C. and Runger, G.C. (2008): Applied Statistics and Probability for Engineers, 3rd Edition reprint, Wiley India Pvt. Ltd.
5. Ehrlich, B. Harris (2002): Transactional Six Sigma and Lean Servicing, 2nd Edition, St. Lucie Press.
6. Hoyle, David (1995): ISO Quality Systems Handbook, 2nd Edition, Butterworth Heinemann Publication.

| **CC 10: BSTAT 403: Linear Model** | **Credit 6**<br>**100 marks** |
|---|---|

**Course Outcome**
After completion of the course the students would be able to

1. Build statistical models under various setups.
2. Analyse data sets using appropriate models.
3. Choose among the different available models in an efficient way.

**UNIT I** **15L**
Gauss-Markov set-up: Theory of linear estimation, Estimability of linear parametric functions, Method of least squares, Gauss-Markov theorem, Estimation of error variance.

**UNIT II** **10L**
Regression analysis: Simple regression analysis, Estimation and hypothesis testing in case of simple and multiple regression models, Concept of model matrix and its use in estimation.

**UNIT III** **20L**
Analysis of variance: Definitions of fixed, random and mixed effect models, analysis of variance and covariance in one-way classified data for fixed effect models, analysis of variance and covariance in two-way classified data with one observation per cell for fixed effects model.

**UNIT IV** **15L**
Model checking: Prediction from a fitted model, Violation of usual assumptions concerning normality, Homoscedasticity and collinearity, Diagnostics using quantile-quantile plots.

**References:**

1. Weisberg, S. (2005). Applied Linear Regression (Third edition). Wiley.
2. Wu, C. F. J. And Hamada, M. (2009). Experiments, Analysis, and Parameter Design Optimization (Second edition), John Wiley.
3. Renchner, A. C. And Schaalje, G. B. (2008). Linear Models in Statistics (Second edition), John Wiley and Sons.

| **SEC 2: BSTAT 405 A: Statistical Techniques for Research Methods** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students would be able to

1. Apply statistical techniques for scientific understanding.
2. Choose the core statistical techniques through elaborated research.
3. Relate the inductive inferential logic into statistical thoughts.

**UNIT I** 5 L

Introduction: Meaning, objection and motivation in research, types of research, research approach, significance of research. Research problems: definition, selection and necessity of research problems.

**UNIT II** 5 L

Survey Methodology and Data Collection, inference and error in surveys, the target populations, sampling frames and coverage error, methods of data collection, non-response, questions and answers in surveys.

**UNIT III** 5 L

Processing, Data Analysis and Interpretation: Review of various techniques for data analysis covered in core statistics papers, techniques of interpretation, precaution in interpretation.

**UNIT IV** 5 L

Develop a questionnaire, collect survey data pertaining to a research problem with relevant examples, Interpret the results and draw inferences.

**References:**

1. Kothari, C.R. (2009): Research Methodology: Methods and Techniques, 2nd Revised Edition reprint, New Age International Publishers.
2. Kumar, R (2011): Research Methodology: A Step - by - Step Guide for Beginners, SAGE publications.

| **SEC 2: BSTAT 405 B: Database Management Systems** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome**

After completion of the course the students will be able to

1. Apply data manipulation and its processing through SQL.
2. Relate the basic concepts with the applications of database systems.
3. Apply the basics of SQL to construct queries.

**UNIT I** 5 L

Introduction: Overview of Database Management System, Introduction to Database Languages, advantages of DBMS over file processing systems.

**UNIT II** 5L

Relational Database Management System: The Relational Model, Introduction to SQL: Basic Data Types, Working with relations of RDBMS: Creating relations e.g. Bank, College Database (create table statement)

**UNIT III** 5L

Modifying relations (alter table statement), Integrity constraints over the relation like Primary Key , Foreign key, NOT NULL to the tables, advantages and disadvantages of relational Database System

**UNIT IV** 5L

Database Structure: Introduction, Levels of abstraction in DBMS, View of data, Role of Database users and administrators, Database Structure: DDL, DML, Data Manager (Database Control System).Types of Data Models Hierarchical databases, Network databases, Relational databases, Object oriented databases.

* The curriculum is to be complemented by Oracle/ SQL server relational database.

**References:**

1. Gruber, M(1990): Understanding SQL, BPB publication
2. Ramez. E, Navathe. S, (2017): Fundamentals of Database Systems, Pearson Education; Seventh edition.
3. Silberschatz. A, Korth. H and Sudarshan. S(2011) “Database System and Concepts”, 6th Edition McGraw-Hill.

| **Lab 1: BSTAT 491: Survey Sampling Laboratory** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to

1. Recognize the practical issues arising in sampling studies.
2. Analyse data from different types of sampling surveys.
3. Understand the principles underlying sampling as a procedure of making inferences about a population.
4. Understand the concepts of bias and sampling error and strategies for reducing these.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. To select a SRS with and without replacement.
2. For a population of size 5, estimate population mean, population mean square and population variance. Enumerate all possible samples of size 2 by WR and WOR and establish all properties relative to SRS.
3. For SRSWOR, estimate mean, standard error, the sample size.
4. Stratified Sampling: allocation of sample to strata by proportional and Neyman's methods Compare the efficiencies of above two methods relative to SRS.
5. Estimation of gain in precision in stratified sampling.
6. Comparison of systematic sampling with stratified sampling and SRS in the presence of a linear trend.
7. Ratio and Regression estimation: Calculate the population mean or total of the population. Calculate mean squares. Compare the efficiencies of ratio and regression estimators relative to SRS.
8. Cluster sampling: estimation of mean or total, variance of the estimate, estimate Intraclass correlation coefficient, and efficiency as compared to SRS.

| **Lab 2: BSTAT 492: Laboratory for Statistical Quality Control** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**
After completion of the course the students will be able to,

1. Understand the philosophy and basic concepts of quality improvement
2. Acquire the ability to demonstrate the methods of statistical process control.
3. Interprete the control charts for variables and attributes and use them for decision making.
4. Understand the product control by using different inspection plans.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Construction and interpretation of statistical control charts
   - X-bar & R-chart
   - X-bar & s-chart
   - np-chart
   - p-chart
   - c-chart
   - u-chart
2. Single sample inspection plan: Construction and interpretation of OC, AQL, LTPD, ASN, ATI, AOQ, AOQL curves.
3. Calculation of process capability and comparison of 3-sigma control limits with specification limits.
4. Use a case study to apply the concept of six sigma application in DMAIC: practical application.

## Detailed Syllabus for Semester V

| **CC 11: BSTAT 501: Stochastic Process and Queuing Theory** | **Credit 4**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course, the student will be able to

1. Recall the knowledge of probabilistic methods used in scientific applications.
2. Carry out derivations involving conditional probability distributions and conditional expectations.
3. Define basic concepts from the theory of Markov chains.
4. Demonstrate applicability of essential mathematical tools for handling stochastic processes.
5. Apply probabilistic and stochastic methods in modern engineering problems.
6. Classify different types of stochastic processes based on their properties.

**Module 1:** **10 L**

Stochastic Process: Introduction and applications, Stationary Process.

**Module 2:** **20 L**

Markov Chains: Definition of Markov Chain, transition probability matrix, order of Markov chain, higher transition probabilities. Chapman Kolmogorov's Theorem and Classification of states and chains, Stability of Markov system. Gambler's ruin problem: Expected duration of the game.

**Module 3:** **15 L**

Poisson Process: postulates of Poisson process, properties of Poisson process, inter-arrival time, pure birth process, pure death process, birth and death process, Yule Furry process.

**Module 4:** **15 L**

Queuing System: General concepts, steady state distribution, queuing model, M/M/1 with finite and infinite system capacity, waiting time distribution (without proof).

**Reference:**

1. Feller, William (1968): Introduction to probability Theory and Its Applications, Vol I,3rd Edition, Wiley International
2. Medhi, J. (2009): Stochastic Processes, New Age International Publishers.

3. Basu, A.K. (2005): Introduction to Stochastic Processes, Narosa Publishing.
4. Bhat, B.R.(2000): Stochastic Models: Analysis and Applications, New Age International Publishers.
5. Taha, H. (1995): Operations Research: An Introduction, Prentice- Hall India. Feller, William (1968): Introduction to probability Theory and Its Applications, Volume I,3rd Edition, Wiley International

| **CC 12: BSTAT 502: Modern Statistical Techniques** | **Credit 4**<br>**100 marks** |
|---|---|

**Course Outcome:**
After completion of the course the student would be able to

1. Learn the concept of how to learn patterns and concepts from data.
2. Understand applications of ML in various functional areas & industries.
3. To design and analyse various machine learning algorithms and techniques with a modern outlook focusing on recent advances.
4. To explore supervised and unsupervised learning paradigms of machine learning.

**Unit 1** **15L**

Frequentist inferential approach and its limitations, notions of Bayesian Paradigm Shift, comparison between classical and Bayesian approach, case studies of Bayesian Analysis.

**Unit 2** **15L**
Regression as a predictive analytics tool, Non-linear regression, Regression Tree, Classification methods.

**Unit 3** **15L**
Introduction to statistical learning, supervised and unsupervised learning, clustering methods.

**Unit 4** **15L**
Random Forest and concept of boosting, Concept of Bootstrapping , Cross Validation technique.

**References:**

1. Machine Learning: Step-by-Step Guide to Implement Machine Learning Algorithms with Python, Rudolph Russell-Online Resource, Copyright 2018

2. Machine Learning: Algorithms and Applications, M. Mohammed, M.B. Khan and E.B.M. Bashier, 2017, CRC press

| **Lab 1: BSTAT 591:Laboratory for Stochastic Processes and Queuing Theory** | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Demonstrate applications of stochastic processes including Markov processes.
2. Apply the methods for determining the stationarity of a stochastic process.
3. Demonstrate different categories of birth and death process.
4. Learn to estimate different parameters of a queue.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

**Practical :**

1. Calculation of transition probability matrix.
2. Identification of characteristics of reducible and irreducible chains.
3. Identification of types of classes.
4. Identification of ergodic transition probability matrix.
5. Stationarity of the Markov chain.
6. Computation of probabilities in case of generalisations of independent Bernoulli trials.
7. Calculation of probabilities for given birth and death rates and vice versa.
8. Calculation of probabilities for Birth and Death Process.
9. Calculation of probabilities for Yule Furry Process.
10. Computation of inter-arrival time for a Poisson process.
11. Calculation of Probability and parameters for (M/M/1) model and change in behavior of queue as N tends to infinity.
12. Calculation of generating function and expected duration for different amounts of stake.

| **Lab 2: BSTAT 592:** Laboratory for Modern Statistical Techniques | **Credit 2**<br>**100 marks** |
|---|---|

**Course Outcome:**

1. To introduce students to the basic techniques of Machine Learning.
2. To develop skills of using recent machine learning software for solving practical problems.
3. To enable the students to: state-of-the-art methods and modern programming tools for data analysis using machine learning programs and algorithms
4. To gain experience of doing independent study and research.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Linear regression, Decision trees, overfitting.
2. Bayesian Decision Making
3. Logistic Regression, Count Regression
4. Clustering: k-means, adaptive hierarchical clustering,
5. Gaussian mixture model.

*Use of any language(R or Python) is encouraged.

**References:**

1. Machine Learning: Step-by-Step Guide to Implement Machine Learning Algorithms with Python, Rudolph Russell-Online Resource, Copyright 2018
2. Machine Learning: Algorithms and Applications, M. Mohammed, M.B. Khan and E.B.M. Bashier, 2017, CRC press
3. Introduction to Machine learning with python by Andreas C. Müller and Sarah Guido, O'REILLY, 2016
4. Introduction to Machine Learning using Python, Jeeva Jose, Khanna Publishing House

**DSE 1: BSTAT 503 A: Time Series Analysis** **Credit 6**
**Marks 100**

**Course Outcome:**
On completion of the course, students will be able to:
1. Classify different components of time series and extraction of those components.
2. Explain basic time series modelling- AR (1), AR (2), MA (1) and MA (2)
3. Apply forecasting, exponential smoothing

**UNIT I** **10L**
Introduction to times series data, application of time series from various fields, Components of a times series, Decomposition of time series. Trend: Estimation of trend, method of semi averages, fitting a various mathematical curve, and growth curves.

**UNIT II** **10L**
Trend Cont.: Method of moving averages. Detrending. Effect of elimination of trend on other components of the time series. Seasonal Component: Estimation of seasonal component by Method of simple averages, Ratio to Trend,

**UNIT III** **10L**

Seasonal Component cont: Ratio to Moving Averages and Link Relative method, Deseasonalization. Cyclic Component: Harmonic Analysis. Some Special Processes: Moving-average (MA) process and Autoregressive (AR) process of orders one and two, Estimation of the parameters of AR (1) and AR (2) – Yule-Walker equations.

**UNIT IV** **10L**

Random Component: Variate difference method. Forecasting: Exponential smoothing methods, Short term forecasting methods: Brown's discounted regression, Box-Jenkins method and Bayesian forecasting. Stationary Time series: Weak stationarity, autocorrelation function and correlogram of moving average.

**SUGGESTED READING:**

1. Kendall M.G. (1976): Time Series, Charles Griffin.
2. Chatfield C. (1980): The Analysis of Time Series –An Introduction, Chapman & Hall.
3. Mukhopadhyay P. (2011): Applied Statistics, 2nd ed. Revised reprint, Books and Allied

**PRACTICAL (Using R/ Python/spreadsheet/ any software package):**

The laboratory should be carried out with the following suggested problems:

1. Fitting and plotting of modified exponential curve
2. Fitting and plotting of Gompertz curve
3. Fitting and plotting of logistic curve
4. Fitting of trend by Moving Average Method
5. Measurement of Seasonal indices Ratio-to-Trend method
6. Measurement of Seasonal indices Ratio-to-Moving Average method
7. Measurement of seasonal indices Link Relative method
8. Calculation of variance of random component by variate difference method
9. Forecasting by exponential smoothing
10. Forecasting by short term forecasting methods

**DSE 1: BSTAT 503 B: Demography and Vital Statistics** **Credit 6**
**Marks 100**

**Course Outcomes:**

On completion of the course, students will be able to:

1. Discover errors in demographic data
2. Measure of fertility and mortality.
3. Explain Life table functions and their applications.
4. Illustrate population growth and population projection.
5. Learn hands on practice on demographic data and vital statistics

**UNIT I** **10 L**

Population Theories: Coverage and content errors in demographic data, use of balancing equations and Chandrasekharan-Deming formula to check completeness of registration data. Adjustment of age data, use of Myer and UN indices, Population composition, dependency ratio.

**UNIT II** **10 L**

Introduction and sources of collecting data on vital statistics, errors in census and registration data. Measurement of population, rate and ratio of vital events. Measurements of Mortality: Crude Death Rate (CDR), Specific Death Rate (SDR), Infant Mortality, Rate (IMR) and Standardised Death Rates. Population growth models, A.P, G.P method, Logistic growth model, fitting of logistic curves by Rhode's method. Application in Indian and Chinese population.

**UNIT III** **10 L**

Stationary and Stable population, Central Mortality Rates and Force of Mortality. Life (Mortality) Tables: Assumption, description, construction of Life Tables and Uses of Life Tables.

**UNIT IV** **10L**

Abridged Life Tables; Concept and construction of abridged life tables by Reed-Merrell method, Greville's method and King's Method. Measurements of Fertility: Crude Birth Rate (CBR), General Fertility Rate (GFR), Specific Fertility Rate (SFR) and Total Fertility Rate (TFR). Measurement of Population Growth: Crude rates of natural increase, Pearl's Vital Index, Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR).

**References:**

1. Mukhopadhyay P. (1999): Applied Statistics, Books and Allied (P) Ltd.
2. Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. II, 9th Edition, World Press.
3. Biswas, S. (1988): Stochastic Processes in Demography & Application, Wiley Eastern Ltd.
4. Croxton, Fredrick E., Cowden, Dudley J. and Klein, S. (1973): Applied General Statistics, 3rd Edition. Prentice Hall of India Pvt. Ltd.
5. Keyfitz N., Beckman John A.: Demography through Problems S-Verlag New York.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**

The laboratory should be carried out with the following suggested problems:

**List of Practicals:**

1. To calculate CDR and Age Specific death rate for a given set of data
2. To find Standardised death rate by: (i) Direct method (ii) Indirect method
3. To construct a complete life table
4. To fill in the missing entries in a life table
5. To calculate probabilities of death at pivotal ages and use it construct abridged life table using (i) Reed-Merrell Method, (ii) Greville's Method and (iii) King's Method
6. To calculate CBR, GFR, SFR, TFR for a given set of data
7. To calculate Crude rate of Natural Increase and Pearle's Vital Index for a given set of data
8. Calculate GRR and NRR for a given set of data and compare them
9. Fitting of logistic curves by Rhode's method

| **DSE 2: BSTAT 504 A: Econometrics** | **Credit 6**<br>**100 marks** |
|---|---|

**Course Outcome**:
After completion of the course the students will be able to,

1. Learn the Statistical Methods in Economics.
2. Apply the basic concepts in probability theory and statistical inference in real-life cases.
3. Understand empirical economic research and to plan and execute independent research projects**.**
4. Identify the techniques for critical evaluation of empirical studies in economics.

**Unit I** **10L**

Introduction: Objective behind building econometric models, nature of econometrics, model building, role of econometrics, structural and reduced forms. General linear model (GLM). Estimation under linear restrictions.

**UNIT II** **10L**

Multicollinearity: Introduction and concepts, detection of multicollinearity, consequences, tests and solutions of multicollinearity, specification error.

**UNIT III** **10L**

Generalised least squares estimation, Aitken estimators. Autocorrelation: concept, consequences of autocorrelated disturbances, detection and solution of autocorrelation.

**UNIT IV** **10L**

Heteroscedastic disturbances: Concepts and efficiency of Aitken estimator with OLS estimator under heteroskedasticity. Consequences of heteroscedasticity. Tests and solutions of heteroscedasticity. Autoregressive and Lag models, Dummy variables, Qualitative data.

**References**:

1. Gujarati, D. and Sangeetha, S. (2007): Basic Econometrics, 4th Edition, McGraw Hill Companies.
2. Johnston, J. (1972): Econometric Methods, 2 Edition, McGraw Hill International.
3. Koutsoyiannis, A. (2004): Theory of Econometrics, 2nd Edition, Palgrave Macmillan Limited,
4. Maddala, G.S. and Lahiri, K. (2009): Introduction to Econometrics, 4 Wiley & Sons.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**
The laboratory should be carried out with the following suggested problems:

1. Problems based on estimation of General linear model
2. Testing of parameters of General linear model
3. Forecasting of General linear model
4. Problems concerning specification errors
5. Problems related to consequences of Multicollinearity
6. Diagnostics of Multicollinearity
7. Problems related to consequences of Autocorrelation (AR(I))
8. Diagnostics of Autocorrelation
9. Estimation of problems of General linear model under Autocorrelation
10. Problems related to consequences Heteroscedasticity
11. Diagnostics of Heteroscedasticity
12. Estimation of problems of General linear model under Heteroscedastic distance terms
13. Problems related to General linear model under (Aitken Estimation )
14. Problems on Autoregressive and Lag models.

**DSE 2: BSTAT 504 B: Financial Statistics** **Credit 6**
**Marks 100**

**Course Outcome:**

After completion of the course the students will be able to,

1. Adapt the fundamental statistical tools in analysing the financial data
2. Apply the hands-on experience to use statistical tools in order to make scientific decisions even in an uncertain environment.
3. Learn different stochastic models for real life situations.

**UNIT I** **10L**

Probability concepts in Finance: Real valued random variables, expectation and variance, skewness and kurtosis , conditional probabilities and expectations. Discrete Stochastic Processes, Binomial processes, General random walks, Geometric random walks, Binomial models with state dependent increments.

**UNIT II**

Tools Needed For Option Pricing: Wiener process, stochastic integration, and stochastic differential equations. Introduction to derivatives: Forward contracts, spot price, forward price, future price. “Call and Put” options, zero-coupon bonds and discount bonds.

**UNIT III**

Pricing Derivatives: Arbitrage relations and perfect financial markets, pricing futures, "Put-call parity" for European options, relationship between strike price and option price.

Stochastic Models in Finance: Discrete time process- One period binomial model.

**UNIT IV** **15L**

Stochastic Models in Finance: Continuous time process- geometric Brownian motion. Basic concepts related to: Ito's lemma, Black-Scholes differential equation, Black-Scholes formula for European options.

Basics concepts of Big data and its statistical approach.

**SUGGESTED READING:**

1. Franke, J., Hardle, W.K. And Hafner, C.M. (2011): Statistics of Financial Markets: An Introduction, 3rdEdition, Springer Publications.
2. Stanley L. S. (2012): A Course on Statistics for Finance, Chapman and Hall/CRC.

**PRACTICAL / LAB WORK (Using R/ Python/ spreadsheet/ any software package):**

The laboratory should be carried out with the following suggested problems:

1. List of Practical To verify "no arbitrage" principle
2. To verify relationship between spot price, forward price, future price
3. To price future contracts
4. To verify put-call parity for European options
5. To construct binomial trees and to evaluate options using these trees
6. To price options using black – Scholes formula.
7. To hedge portfolios using delta and gamma hedging.
8. To hedge portfolios theta hedging.
9. Pricing of call options using binomial model.
10. Computation of dividends on call options as a percentage of stock price.
11. Computation of dividends on call options as a fixed amount of money.
12. Pricing of put options using binomial model
13. Call-put parity for options following binomial models.
14. Effect of dividends on "Put options".

**DSE II: BSTAT 504 C: Actuarial Statistics** **Credit 6**
**Marks 100**

**Course Outcome:**

After completion of the course the students will be able to,

1. Identify the utility functions and their uses in insurance.
2. Adapt the individual and aggregate claims and their applications.
3. Apply Life table functions in real life applications.

**UNIT I** **10L**

Introductory Statistics and Insurance Applications: Discrete, continuous and mixed probability distributions. Insurance applications, Convolutions of random variables. Utility theory: Utility functions, expected utility criterion, types of utility function, insurance and utility theory.

**UNIT II** **10L**

Principles of Premium Calculation: Properties of premium principles, examples of premium principles. Individual risk models: models for individual claims, the sum of independent claims, approximations and their applications.

**UNIT III** **10L**

Survival Distribution and Life Tables: Uncertainty of age at death, survival function, time until-death for a person, curtate future lifetime, force of mortality, life tables with examples, deterministic survivorship group, life table characteristics, assumptions for fractional age, some analytical laws of mortality.

**UNIT IV** **10l**

Life Insurance: Models for insurance payable at the moment of death, insurance payable at the end of the year of death and their relationships. Life annuities: continuous life annuities, discrete life annuities, life annuities with periodic payments. Premiums: continuous and discrete premiums.

**SUGGESTED READING:**

1. Dickson, C. M. D. (2005): Insurance Risk And Ruin (International Series On Actuarial Science), Cambridge University Press.
2. Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A. And Nesbitt, C. J. (1997): Actuarial Mathematics, Society Of Actuaries, Itasca, Illinois, U.S.A.

**PRACTICAL / LAB WORK** (Using R/ Python/ spreadsheet/ any software package):
The laboratory should be carried out with the following suggested problems:

List of Practical:

1. Risk computation for different utility models
2. Discrete and continuous risk calculations
3. Calculation of aggregate claims for collective risks
4. Calculation of aggregate claim for individual risks
5. Computing Ruin probabilities and aggregate losses
6. Annuity and present value of contract
7. Computing premium for different insurance schemes
8. Practical based on life models and tables

## Detailed Syllabus for Semester VI

| **CC 13: BSTAT 601: Design of Experiments** | **Credit 4**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to

1. Familiar with the basic principles and methods of statistical design of experiments.
2. Understand about the different types of formal experimental designs and Incomplete Block Designs.
3. Describe the concept of variability, its causes and methods of reducing it.
4. Identify the experimental unit and recognise issues of non-independence.
5. Explain the fundamental concepts behind the output of an ANOVA (including "blocking" and "interactions").

**UNIT I** **15L**

Experimental designs: Role, historical perspective, terminology, experimental error, basic principles, uniformity trials, fertility contour maps, choice of size and shape of plots and blocks. Basic designs: Completely Randomised Design (CRD), Randomised Block Design (RBD), Latin Square Design (LSD) – layout, model and statistical analysis, relative efficiency, analysis with missing observations.

**UNIT II** **5L**

Incomplete Block Designs: Balanced Incomplete Block Design (BIBD) – parameters, relationships among its parameters, incidence matrix and its properties, Symmetric BIBD.

**UNIT III** **10L**

Factorial experiments: advantages, notations and concepts, $2^2$, $2^3 \ldots 2^n$. Factorial experiments, design and analysis, Total and Partial confounding for $2^n$ $(n \leq 5)$.

**UNIT IV** **10L**

Analysis of factorial experiments: complete and confounded experiments.

**SUGGESTED READINGS:**

1. Cochran, W.G. and Cox, G.M. (1959): Experimental Design. Asia Publishing House.
2. Das, M.N. and Giri, N.C. (1986): Design and Analysis of Experiments. Wiley Eastern Ltd.
3. Goon, A.M., Gupta, M.K. and Dasgupta, B. (2005): Fundamentals of Statistics. Vol.II, 8th Edn. World Press, Kolkata.
4. Kempthorne, O. (1965): The Design and Analysis of Experiments. John Wiley.
5. Montgomery, D. C. (2008): Design and Analysis of Experiments, John Wiley.

| **LAB 1: BSTAT 691: Laboratory for Design of Experiments** | **Credit 2**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to

1. Analyse the data from the experiments.
2. Understand the potential practical problems in a design for a particular experiment.
3. Construct good or optimal designs for a range of practical experiments.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**

The laboratory should be carried out with the following suggested problems:

1. Analysis of a CRD
2. Analysis of an RBD
3. Analysis of an LSD
4. Analysis of an RBD with one missing observation
5. Analysis of an LSD with one missing observation
6. Analysis of 2^2 and 2^3 factorial in CRD and RBD
7. Analysis of 2^2 and 2^3 factorial in LSD
8. Analysis of a completely confounded two level factorial design in 2 blocks
9. Analysis of a completely confounded two level factorial design in 4 blocks
10. Analysis of a partially confounded two level factorial design.
11. Analysis of a single replicate of a 2^n design.

| **CC14: BSTAT 602: Multivariate Analysis and Nonparametric Methods** | **Credit 6**<br>**Marks 100** |
|---|---|

**Course Outcome:**

After completion of the course the students will be able to,

1. Understand the various multivariate techniques available.
2. Understand the concept of analysing multivariate data.
3. Describe properties of bivariate and multivariate normal distribution.
4. Identify when to use a nonparametric method.
5. Understand different nonparametric methods in estimation, testing, model fitting and in analyses.

**UNIT I** **15L**

Bivariate Normal Distribution (BVN): p.d.f. of BVN, properties of BVN, marginal and conditional p.d.f. of BVN. Multivariate Data: Random Vector: Probability mass/density functions, Distribution function, Mean vector & Dispersion matrix, Marginal & Conditional distributions.

**UNIT II** **15L**

Multivariate Normal distribution and its properties. Sampling distribution for mean vector and variance- covariance matrix. Multiple and partial correlation coefficient and their properties.

**UNIT III** **10L**

Nonparametric Tests: Introduction and Concept, Test for randomness based on total number of runs, Empirical distribution function, Kolmogrov Smirnov test for one sample, Sign tests- one sample and two samples, Wilcoxon-Mann-Whitney test, Kruskal-Wallis test.

**SUGGESTED READING:**

1. Anderson, T.W. (2003): An Introduction to Multivariate Statistical Analysis, 3rdEdn., John Wiley
2. Muirhead, R.J. (1982): Aspects of Multivariate Statistical Theory, John Wiley.
3. Kshirsagar, A.M. (1972) :Multivariate Analysis, 1stEdn. Marcel Dekker.
4. Johnson, R.A. and Wichern, D.W. (2007): Applied Multivariate Analysis, 6thEdn., Pearson & Prentice Hall
5. Mukhopadhyay, P. Mathematical Statistics.
6. Gibbons, J. D. and Chakraborty, S (2003): Nonparametric Statistical Inference. 4th Edition. Marcel Dekker, CRC.

**PRACTICALS/ LAB WORK: 20L**

**(Using R/ Python/spreadsheet/ any software package):**

The laboratory should be carried out with the following suggested problems:

1. Problems based on Multiple Correlation
2. Problems based on Partial Correlation
3. Problems based on Bivariate Normal Distribution and signed test,
4. Problems based on Multivariate Normal Distribution
5. Problems based on Kolmogrov Smirnov test for one sample and Sign Tests for one sample and two samples.
6. Problems based on Wilcoxon-Mann-Whitney test.
7. Problem based on Kruskal-Wallis test.

| **DSE 3: BSTAT 603 A: Survival Analysis and Biostatistics** | **Credit 6**<br>**100 marks** |
|---|---|

**Course Outcome**

After completion of the course the students will be able to,

1. Understand different biostatistical problems.
2. Learn application of statistics in medical science.
3. Design different stages of clinical trials.
4. Find the root cause of disease.

**UNIT I** **20L**

Survival Analysis: Functions of survival times, survival distributions and their applications- exponential, gamma, Weibull, Rayleigh, lognormal, death density function for a distribution having a bath-tub shaped hazard function.
Censoring Schemes: Type I, Type II and progressive or random censoring with biological examples. Estimation of mean survival time and variance of the estimator for Type I and Type II censored data with numerical examples. Non-parametric methods: Actuarial and Kaplan-Meier methods for estimating survival function and variance of the Estimator.

**UNIT II** **20L**

Competing Risk Theory: Indices for measurement of probability of death under competing risks and their inter-relations. Estimation of probabilities of death using maximum likelihood principle and modified minimum Chi-square methods. Theory of independent and dependent risks. Bivariate normal dependent risk model.

**SUGGESTED READING:**

1. Lee, E.T. and Wang, J.W. (2003): Statistical Methods for Survival data Analysis, 3rd Edition, John Wiley and Sons.
2. Biswas, S. (2007): Applied Stochastic Processes: A Biostatistical and Population Oriented Approach, Reprinted 2ndCentral Edition, New Central Book Agency.

3. Kleinbaum, D.G. (1996): Survival Analysis, Springer.

4. Chiang, C.L. (1968): Introduction to Stochastic Processes in Biostatistics, John Wiley and Sons.

5. Indrayan, A. (2008): Medical Biostatistics, 2nd Edition Chapman and Hall/CRC.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):** **20L**

The laboratory should be carried out with the following suggested problems:

1. To estimate survival function
2. To determine death density function and hazard function
3. To identify type of censoring and to estimate survival time for type I censored data
4. To identify type of censoring and to estimate survival time for type II censored data
5. To identify type of censoring and to estimate survival time for progressively type I censored data
6. Estimation of mean survival time and variance of the estimator for type I censored data
7. Estimation of mean survival time and variance of the estimator for type II censored data
8. Estimation of mean survival time and variance of the estimator for progressively type I censored data
9. To estimate the survival function and variance of the estimator using Non-parametric methods with Actuarial methods
10. To estimate the survival function and variance of the estimator using Non-parametric methods with Kaplan-Meier method

| **DSE 3: BSTAT 603 B: Operations Research** | **Credit 6**<br>**Marks 100** |
|---|---|

**Course Outcome:**
After completion of the course, the students will be able to,

1. Develop operational research models from the actual description of the real system,
2. Understand the mathematical tools that are needed to solve optimisation problems,
3. Solve linear programming problems (LPP) e.g. Shortest route and minimal spanning tree problems, transshipment problems and assignment problems.
4. Find out the optimal strategy for two-person zero sum game with or without saddle point
5. Determine EOQ for different inventory models.

**UNIT I** **15 L**

Introduction to Operations Research, phases of O.R., model building, various types of O.R. problems, Linear Programming Problem, Mathematical formulation of the L.P.P, graphical solutions of a L.P.P., Simplex algorithm for solving L.P.P. Charne's M-technique and two phase method for solving L.P.P. involving artificial variables. Special cases of L.P.P. Concept of Duality in L.P.P: Dual simplex method. Post-optimality analysis.

**UNIT II** **15 L**

Transportation Problem: Initial solution by North West corner rule, Least cost method and Vogel's approximation method (VAM), MODI's method to find the optimal solution, Special cases of transportation problems. Assignment problem: Hungarian method to find optimal assignment, special cases of assignment problem.

**UNIT III** **15 L**

Game theory: Rectangular game, minimax-maximin principle, Nash Equilibrium, solution to rectangular game using graphical method, dominance and modified dominance property to reduce the game matrix and solution to rectangular game with mixed strategy. Networking: Shortest route and minimal spanning tree problem.

**SUGGESTED READING:**

1. Taha, H. A. (2007): Operations Research: An Introduction, 8th Edition, Prentice Hall of India.
2. Kanti Swarup, Gupta, P.K. and Manmohan (2007): Operations Research, 13th Edition, Sultan Chand and Sons.
3. Hadley, G: (2002) : Linear Programming, Narosa Publications.
4. Hillier, F.A and Lieberman, G.J. (2010): Introduction to Operations Research-Concepts and cases, 9th Edition, Tata McGraw Hill.

**PRACTICAL / LAB WORK (Using R/ Python/spreadsheet/ any software package):**

The laboratory should be carried out with the following suggested problems:

1. Linear Programming Problem
2. Simplex algorithm
3. Dual simplex algorithm
4. Transportation problem
5. Assignment Problem
6. Game Theory
7. Shortest root and Minimal Spanning Tree problem

**RAEC: BSTAT 604: Capstone Project I** **Credit 6**
**100 marks**