File Ref.No.Ac.A3/Faculty of Science/2023 (PART-I)

# COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

(Abstract)

Faculty of Science - Revised course structure and syllabi of M.Sc. Statistics and M.Tech Data Science and Analytics Programme effective from 2024 admission onwards - Resolution of the Academic Council - Communicated - Orders issued.

---

**ACADEMIC A SECTION**

No.CUSAT/AC(A).A3/3583/2024 Dated,KOCHI-22,22.08.2024

---

Read:-Item No. I (g)(6) of the minutes of the meeting of the Academic Council held on 30.04.2024

**ORDER**

The Academic Council considered along with the recommendations of its standing committee, the Minutes of the Faculty of Science held on 08.04.2024 and resolved to approve the revised course structure and syllabi for the following programmes:

i. M.Sc.Statistics with effect from 2024 admissions (Appendix I)
ii. M.Tech Data Science and Analytics with effect from 2024 admissions (Appendix II)

Orders are, therefore, issued accordingly.

**Dr. V. Sivanandan Achari ***
**Registrar**

To:

1. The Dean, Faculty of Science
2. Chairmen, BoS under Faculty of Science
3. The Head, Department of Statistics
4. All AR/DR Examination wing - with a request to forward to concerned sections
5. The Director,IQAC/ DoA
6. CIRM/Conference Sections
7. PS To VC/PVC;PA To Registrar/CE.

* This is a computer generated document. Hence no signature is required.

**M.Sc. (STATISTICS): ACADEMIC PROGRAM**
**(2024 Admission Onwards)**

## Objectives of the Program

The present program is intended to provide a platform for talented students to undergo higher studies in the subject as well as to train them to suit the needs of society. Apart from teaching core Statistics subjects the students can choose inter-disciplinary, intra-disciplinary and skill-based electives depending upon their interests, under the choice-based credit system. The students are also trained to handle real-life problems through practical classes and project work. As a part of the course, the students are also exposed to various statistical software's.

**Program Outcomes:**

**On successful completion of the M. Sc. Statistics program the students will be able to**

PSO.1: Understand the role of probability and statistics in solving real-life problems.
PSO.2: Acquire knowledge of modern statistical techniques relevant to today's scientific community.
PSO.3: Convince the need for systematic analysis of data in any scientific experiment.
PSO.4: Provide consultancy on experimental design and field survey.
PSO.5: Handle any statistical packages.
PSO.6: Handle real-life problems using suitable statistical tools in any discipline as well and they will be able to work in any industry which deals with data.
PSO.7: To become a professionally inclined statistics teacher/statistician/data scientist who has sound knowledge of the subject matter and specializes in knowledge discovery through statistical methods.
PSO.8: To understand basic theoretical and applied principles of statistics with adequate preparation to pursue a Doctoral (PhD) degree or enter the job force as an applied statistician.
PSO.9: To communicate key statistical concepts to non-statisticians.
PSO.10: To gain proficiency in using statistical software/utility for data analysis.

**Eligibility**

(i) Successful completion with a pass of the first six semesters of the Integrated M.Sc program of CUSAT, who have opted for Integrated M.Sc (IMSc) in Statistics.

OR

(ii) A B.Sc Degree with,

a) Mathematics as a main subject or an optional subject

AND

b) Statistics as a main subject or an optional subject with atleast 55% marks for the main and optional subjects taken together.

**Duration of the Course** : Four Semesters

**Examination** : Credit and Semester

2024 onwards : 25 (upto 15 from IMSc and all the remaining seats from CAT Rank List)

| SEMESTER - I | | | | | | |
|---|---|---|---|---|---|---|
| **Course Code** | **Title of Paper** | **Core/ Elective** | **Credits (Total Credits - 19)** | **Continuous Evaluation** | **End Semester Evaluation** | **Total Marks** |
| 24-322-0101 | Mathematical Methods for Statistics | C | 4 | 50 | 50 | 100 |
| 24-322-0102 | Probability Theory - I | C | 4 | 50 | 50 | 100 |
| 24-322-0103 | Probability Distributions | C | 4 | 50 | 50 | 100 |
| 24-322-0104 | Sampling Theory& Methods | C | 4 | 50 | 50 | 100 |
| | Elective - I | E | 3 | 50 | 50 | 100 |
| Elective - I (Choose any one) | | | | | | |
| 24-322-0105 | Data Analytics using R | E | 3 | 50 | 50 | 100 |
| 24-322-0106 | Statistical Techniques in Data Science using R | E | 3 | 50 | 50 | 100 |

| SEMESTER - II | | | | | | |
|---|---|---|---|---|---|---|
| **Course Code** | **Title of Paper** | **Core/ Elective** | **Credits (Total Credits - 19)** | **Continuous Evaluation** | **End Semester Evaluation** | **Total Marks** |
| 24-322-0201 | Statistical Inference I | C | 4 | 50 | 50 | 100 |
| 24-322-0202 | ProbabilityTheory II | C | 4 | 50 | 50 | 100 |
| 24-322-0203 | Stochastic Processes | C | 4 | 50 | 50 | 100 |
| 24-322-0204 | Practical-I & Viva Voce | C | 2 | 50 | 50 *a | 100 |
| | Elective -II | E | 3 | 50 | 50 | 100 |
| | Elective-III *b | E | 2 | - | 100 | 100 |
| Elective - II (Choose any one) | | | | | | |
| 24-322-0205 | Statistics for National Development | E | 3 | 50 | 50 | 100 |
| 24-322-0206 | Reliability Modeling and Analysis | E | 3 | 50 | 50 | 100 |

| SEMESTER - III | | | | | | |
|---|---|---|---|---|---|---|
| **Course Code** | **Title of Paper** | **Core/ Elective** | **Credits (Total Credits -21)** | **Continuous Evaluation** | **End Semester Evaluation** | **Total Marks** |
| 24-322-0301 | Statistical Inference II | C | 4 | 50 | 50 | 100 |
| 24-322-0302 | Multivariate Analysis | C | 4 | 50 | 50 | 100 |
| 24-322-0303 | Applied Regression Analysis | C | 4 | 50 | 50 | 100 |
| 24-322-0304 | Practical-II, Project & Viva Voce | C | 3 | 50 | 30 *a + 20 *c | 100 |
| | Elective -IV | E | 3 | 50 | 50 | 100 |
| | Elective-V (Inter-departmental course/MOOC course *b) | E | 3 | 50*b | 50*b | 100 |
| Elective–IV (Choose any one of the following) Either an inter-departmental course or an online course *b | | | | | | |
| 24-322-0305 | Time Series Analysis | E | 3 | 50 | 50 | 100 |
| 24-322-0306 | Topics in Mathematical Finance | E | 3 | 50 | 50 | 100 |
| 24-322-0307 | Operations Research | E | 3 | 50 | 50 | 100 |

| SEMESTER - IV | | | | | | |
|---|---|---|---|---|---|---|
| **Course Code** | **Title of Paper** | **Core/ Elective** | **Credits (Total Credits - 21)** | **Continuous Evaluation** | **End Semester Evaluation** | **Total Marks** |
| 24-322-0401 | Design and Analysis of Experiments | C | 4 | 50 | 50 | 100 |
| 24-322-0402 | Practical-III & Viva Voce*d | C | 4 | 50 | 50 *d | 100 |
| 24-322-0403 | Project | C | 4 | 50 | 50 *e | 100 |
| | Elective -VI | E | 3 | 50 | 50 | 100 |
| | Elective -VII | E | 3 | 50 | 50 | 100 |
| | Elective -VIII | E | 3 | 50 | 50 | 100 |
| Electives-VI, VII, VIII (Choose any three) | | | | | | |
| 24-322-0404 | Statistical Quality Assurance | E | 3 | 50 | 50 | 100 |
| 24-322-0405 | Actuarial Statistics | E | 3 | 50 | 50 | 100 |
| 24-322-0406 | Lifetime Data Analysis | E | 3 | 50 | 50 | 100 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 24-322-0407 | Applied Multivariate Statistical Analysis | E | 3 | 50 | 50 | 100 |
| 24-322-0408 | Statistical Forecasting | E | 3 | 50 | 50 | 100 |
| 24-322-0409 | Inference for Stochastic Processes | E | 3 | 50 | 50 | 100 |
| 24-322-0410 | Categorical Data Analysis | E | 3 | 50 | 50 | 100 |
| 24-322-0411 | Directional Data Analysis | E | 3 | 50 | 50 | 100 |
| 24-322-0412 | Statistical Decision Theory | E | 3 | 50 | 50 | 100 |
| 24-322-0413 | Theory of Entropy | E | 3 | 50 | 50 | 100 |
| 24-322-0414 | Statistics for Clinical Research | E | 3 | 50 | 50 | 100 |

***a** End semester evaluation based on viva voce.

***b** A MOOC course from SWAYAM/NPTEL/Moodle/Others will be opted for by the students with approval of the Department Council and it will have END SEMESTER EVALUATION ONLY, for 100 marks.

***c** End-semester evaluation based on project.

***d** The VivaVoce examination is to be conducted externally with at least one external examiner.

***e** End-semester evaluation will be done based on the presentation and project report.

# DETAILED SYLLABUS

## 24-322-0101: MATHEMATICAL METHODS FOR STATISTICS

| Course Out come (CO) | Cognitivelevel |
|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Demonstrate and understanding of limits and continuity of various functions | Apply |
| 2. | Evaluate the Riemann-Stieltjes integral and verify the conditions for the existence of the integrals | Evaluate |
| 3. | Distinguish between the concepts of sequence and series, and determine limits of sequences and convergence and approximate sums of series | Analysis |
| 4. | Compute the partial and total derivatives and maxima and minima of multivariable function | Apply |
| 5. | Solve systems of line are quations, diagonalize matrices and Characterize quadratic forms | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 1 | 2 | | | | | | | | |
| CO2 | 2 | | 3 | | | | 2 | 2 | | |
| CO3 | | | 3 | | | | | | 2 | |
| CO4 | | | 2 | | | | 3 | | | |
| CO5 | 2 | | | | | | 2 | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I:**
Riemann-Stieltjes Integral - definition, properties and important theorems, Eulers summation formula, integrators of bounded variation, sufficient conditions for existence of Riemann-Stieltjes integrals, Mean value theorems of Riemann-Stieltjes integrals.

**Module-II:**
Sequences and series of functions: - Pointwise and uniform convergence, properties and theorems; Multivariable Calculus- limit and continuity, total derivative, directional derivatives, differentiation of composite functions, Taylor's Theorem for a multivariable function, inverse and implicit functions, optima of a multivariable function, method of Lagrange multipliers.

**Module-III:**
Matrices:- Rank of a matrix, elementary transformations of a matrix and properties, inverse of a matrix by elementary transformations, Generalized Inverse Matrices:- Definition and existence, Solving linear equations, Moore-Penrose inverse, Symmetric matrices - properties of generalized inverse; singular value decomposition.

**Module-IV:**
Quadratic forms - classification and its characteristic properties, canonical forms, gram matrices, Characteristic roots and characteristic vectors of a matrix, nature of characteristic roots of some special types of matrices, algebraic and geometric multiplicity of a characteristic roots, Cayley-Hamilton theorem, Orthogonal and unitary reductions of quadratic forms, Spectral decomposition of a matrix, singular value decomposition.

**References:**


1. Searle S. R. and Khuri A. I. (2017). Matrix Algebra Useful for Statistics. Wiley Series in Probability and Statistics, Second Edition.
2. Khuri A.T. (1993). Advanced Calculus with Applications in Statistics, John Wiley & Sons, Inc., USA, Chapters - 3 and 7
3. Apostol T.M. (1996). Mathematical Analysis, Narosa Publshing House, New Delhi, Second Edition, Chapters - 6, 7, 9.
4. Shanti Narayan (1991). A text of book of matrices, S. Chand & Company, New Delhi, Chapters - 3, 6, 7, 10, 11.
5. Searle S.R. (1971). Linear models, John Wiley & Sons, Inc., Chapter - 1
6. Gupta S.L. and Gupta N.R. (2003) Principles of Real Analysis, Second edition, Pearson Education (Singapore) Pte. Ltd.
7. Widder D.A. (1996) Advanced Calculus, Second Edition, Prentice Hall, Inc., New Delhi.
8. Nanda S. and Saxena, V.P. (2000) Real Analysis, Allied Publishers Ltd.
9. Graybill F.A. (1969) Introduction to matrices with applications in statistics, Wadsworth Publishing Company, USA.
10. Rao C.R. (2002) Linear statistical inference and its applications, Second edition, Chapter 1b, 1c.

## 24-322-0102: PROBABILITY THEORY – I

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Identify sigma fields and Compute limits of a sequence of random variables | Apply |
| 2. | Describe properties of Probability Measure and distribution function | Remember |
| 3. | Define Expectation and moments | Understand |
| 4. | Compute Momentine qualities using Expectations | Apply |
| 5. | Concepts of Independence and its use in Multiplication properties, Zero-one laws | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | | | | | 2 | 2 | | |
| CO2 | | | 2 | | | | 2 | 2 | 3 | |
| CO3 | | | | | | | 3 | | | |
| CO4 | | 1 | | | | | 3 | | | |
| CO5 | | | | | | | | 2 | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Random variables: Algebra of sets, Fields, Sigma fields, Inverse function, Measurable functions, Random variables, Induced sigma fields, Limits of random variables.

**Module-II**

Probability: General measure space, Lebesgue measure, Lebesgue-Stieltjes measure, Counting measure and their simple properties, Discrete probability space, General probability space as normed measure space, Induced probability space, Extension of probability measures. Distribution function of a random variable, Decomposition of distribution functions, Distribution function of random vectors.

**Module-III**

Integration with respect to measure (Introduction only), Expectation and moments: Definition and properties, Moment generating functions, Moment inequalities: $C_r$-, Holder, Jenson and basic inequalities, Product spaces and Fubini's theorem (idea and statement only), Independence: Definitions, Multiplication properties, Zero-one laws.

**Module-IV**

Convergence: Modes of convergence, Convergence in probability, in distribution, in rth mean, almost sure convergence and their inter-relationships, Convergence theorem for expectation such as Monotone convergence theorem, Fatou's lemma, Dominated convergence theorem (some remarks on the corresponding theorems for general integrals with respect to measure).

**References:**


1. Billingsley P.(2017)Probability and Measure, Third Edition, Wiley, India Pvt. Ltd.
2. Bhat B.R. (2011) Modern Probability Theory, Second edition, Wiley Eastern, Chapters 1,2,3, 4,5,6,9.
3. Laha R. G. and Rohatgi V. K. (2020). Probability Theory, John Wiley and Sons, New York.
4. Rohatgi V. K. and Saleh M. (2015). An Introduction to Probability and Statistics, Third Edition, John Wiley and Sons, New Jersey.
5. Feller W.(1966)An Introduction to Probability Theory and Its Applications, Volume II, Wiley Eastern.
6. Rao C.R.(1973) Linear Statistical Inference and Its Applications, Wiley.
7. Rohatgi V.K. and A.K.E. Salah (2001) Introduction to Probability and Statistics, John Wiley and Sons.
8. Basu A.K.(1999) Measure Theory and Probability, Prentice-Hall.

## 24-322-0103: PROBABILITY DISTRIBUTIONS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Describe and employ various statistical concepts to study the discrete distributions | Apply |
| 2. | Describe and employ various statistical concepts to study the discrete distributions | Apply |
| 3. | Describe properties of bivariate continuous exponential Distributions | Understand |
| 4. | Illustrate characterization properties of the bivariate exponential | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | | | | | 2 | 3 | | |
| CO2 | 3 | | | | | | 2 | 2 | | |
| CO3 | | | | | | | 3 | 2 | | |
| CO4 | | | | | | | | 3 | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Discrete Distributions: Modified power series family - properties, moment generating functions, recurrence relations for raw, central and factorial moments, recurrence relation for cumulants, Binomial, Negative binomial, Logarithmic series and Lagrangian distributions and their properties as special cases of the results from modified power series family, hypergeometric distribution and its properties.

**Module-II**

Continuous distribution: Pearson family – identifications of the different types, Beta, Gamma, Pareto and Normal Special cases of the Pearson family and their properties. Exponential family of distributions, Compound, truncated and mixture distributions.

**Module-III**

Sampling distributions: Sampling distributions of the mean and variance from normal population, independence of mean and variance, Chi-square, students *t* and *F* distribution and their non-central forms. Order statistics and their distributions, Conditional distribution of order statistics, distribution of sample range.

**Module-IV**

Bivariate distributions: Multinominal, bivariate normal, bivariate exponential distribution of Gumbel, Marshall and Olkin and Block and Basu, Dirichlet distribution.

**References:**


1. Rohatgi V.K (1976) An introduction to Probability Theory and Mathematical Statistics, Wiley Eastern
2. Arnold B.C, Balakrishnan N and Nagaraja H.N (1992). A first course in order statistics
3. Galambos J and Kotz's (1978):Characterization of Probability distributions, Springer-Verlag.
4. Ord J.K.(1972) Families of frequency distributions Griffin
5. Johnson N.L, Kotz S and Kemp A.W (1992) Univariate discrete distributions, John Wiley.
6. Johnson N.L Kotz S and Balakrishnan N (1991) Continuous univariate distributions I & II, John Wiley.
7. Johnson N.L, Kotz S and Balakrishnan N(1995) Multivariate Distribution, John Wiley.

## 24-322-0104: SAMPLING THEORY AND METHODS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Apply various sampling procedures like SRS, Stratified, Systematic, Cluster etc., and estimate the population parameters for attributes and variables | Apply |
| 2. | Estimate population ratio, population mean and population total using ratio, difference and regression estimators | Apply |
| 3. | Explain Midzuno-Sen-Lahiri, Murthy's, Des Raj's sampling strategies Evaluate/under varying probability without replacement sampling | Apply |
| 4. | Understand various types of errors in surveys and procedures to rectify them | Understand |
| 5. | Understand quota, network and adaptive samplings; and evaluate estimator under adaptive sampling | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | 2 | 1 | | | | 1 | | |
| CO2 | | 2 | 2 | | | | | 1 | | |
| CO3 | | | 2 | | | | | | | |
| CO4 | 1 | | 1 | | | | | | | |
| CO5 | 1 | | 1 | | | | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-1**

Basic concepts:- Population, sample, sampling design, interpenetrating subsampling; Simple Random Sampling (SRS):- SRS with replacement, SRS without replacement, confidence interval, estimation of population proportion, determination of sample size, comparison between SRSWR and SRSWOR; Stratified Random Sampling:- estimation of population mean and total, optimum allocation, other types of allocation, comparison with SRS.

**Module-II**

Estimation of gain due to stratification over SRS, construction of strata, number of strata, Ratio estimator:- Bias and mean square error, estimation of variance, confidence interval, comparison with mean per unit estimator, optimum property of ratio estimator, unbiased ratio type estimator, ratio estimator in stratified random sampling; Difference estimator and Regression estimator:- Difference estimator, regression estimator, comparison of regression estimator with mean per unit and ratio estimator, regression estimator in stratified random sampling.

**Module-III**

Systematic sampling:- estimation of population mean and variance, comparison of systematic sampling with SRS and stratified random sampling, circular systematic sampling; Cluster sampling:- estimation of population mean, estimation of efficiency by a cluster sample, variance function, determination of optimum cluster size, clusters of varying sizes; Probability proportional to size with replacement sampling:- estimation of population mean and total, selection of a ppswr sample; Varying probability without replacement sampling I:- properties of a sampling design, Horvitz-Thomson estimator.

**Module-IV**

Varying probability without replacement sampling II:-Midzuno-Sen-Lahiri sampling strategy, Desraj, Murthy's; Mutistage sampling:- estimation population total with SRS sampling at both stages, multiphase sampling (outline only); Errors in surveys:- effect of unit nonresponse in the estimate, procedures for unit nonresponse; quota sampling, network sampling; Adaptive sampling:- introduction and estimators under adaptive sampling

**References:**


1. Mukhopadhay P (2009) Theory and methods of survey sampling, Second edition, PHIL earning Pvt. Ltd., New Delhi, Relevant sections of Chapters1-16.
2. Sampath S. (2001) Sampling theory and methods, Alpha Science International Ltd., India, Chapter 10.
3. Arnab R. (2017). Survey Sampling Theory and Applications. Netherlands: Elsevier Science.
4. Bansal A. (2017). Survey Sampling. United Kingdom: Alpha Science International, Limited.
5. Cochran W.G.(1999) Sampling Techniques, Third edition, John Wiley & Sons.
6. DesRaj (1976) Sampling Theory, McGraw Hill.
7. Murthy M.N.(1977) Sampling Theory and Methods, Statistical Publishing Society, Calcutta.
8. Singh D. and Chaudhary F.S. (1986) Theory and Analysis of Sample Survey Designs, Wiley Eastern.

## 24-322-0105: DATA ANALYTICS USING R

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Develop a scientific computing environment using R | Evaluate |
| 2. | Identify the use of R software to meet the given scientific objective | Analyze |
| 3. | Identify the use of various packages in R | Analyze |
| 4. | Write an efficient program using R to perform routine and specialized data manipulation /management and analysis tasks | Execute |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | 3 | | | | | | | | |
| CO2 | | 3 | | 3 | | | | | | 1 |
| CO3 | | | | | | | | | | 3 |
| CO4 | 1 | | | | | 2 | | | | 3 |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Introduction to statistical software R, Using R as a calculating environment, Arithmetic variables, Functions, Vectors, Expressions and assignments, Logical expressions, Manipulating vectors, matrices, importing of files.

Types of data, Scale of measurement, Data objects in R, Graphical summaries of data-Bar chart, Pie chart, Histogram, Box-plot, Stem and leaf plot, Frequency table, Plotting of probability distributions and sampling distributions, P-P plot, Q-Q Plot, Computations of descriptive statistics measures.

Inference from bivariate data-Scatter plot, Correlation and Regression.

**Module-II**

Basic programming, Branching with if, Looping with for, Looping with while, Vector-based programming, Program flow, Pseudo-code, Basic debugging, Programming with functions, Vectorized functions, Optional arguments and default values, Vector based programming using functions, Recursive programming, Debugging functions, Data frames, Lists, Use of apply group of functions.

**Module-III**

Simulation, Congruential generators, Seeding, Random Number Generation- Basic principles of Random number generation, Inversion method, Accept-reject method, Random number generation from Binomial, Poisson Uniform, Exponential, Cauchy and Normal, Rejection with exponential envelope, Box-Muller algorithm.

**Module-IV**

Statistical Inference Problems Using R-Estimation and confidence intervals-Point estimates of normal mean, confidence interval for normal mean with known and unknown standard deviation. Confidence interval for standard deviation. Confidence interval for proportion.

One sample t-test, two sample t-test, paired t-test, test on standard deviation (chi-square test).

**References:**


1. Jones O., Maillardet. R. and Robinson, A. (2014). Introduction to Scientific Programming and Simulation Using R. Chapman & Hall/CRC, The R Series.
2. Crawley, M, J. (2012). The R Book, 2nd Edition. John Wiley & Sons.
3. Chambers, J. M. (2008). Software for Data Analysis-Programming with R. Springer-Verlag, New York.
4. Jammalamadaka, S. R. (2007). Essential Statistics with python and R. Kendal Hunt publishing

## 24-322-0106: STATISTICAL TECHNIQUES IN DATA SCIENCE USING R

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Apply R for visualizing both quantitative and qualitative data, including techniques like Q-Q plots | Apply |
| 2. | Will be able to understand and Apply Two Sample Analysis: Matched pairs analysis, independent samples analysis | Understand and Apply |
| 3. | Have a clear understanding of the assumptions, estimation methods, inferences and diagnostics associated with simple linear regression and multiple linear regression with hands on illustrations using R | Understand and Apply |
| 4. | Execute the knowledge and techniques necessary to analyze binary and categorical response variables using the machine learning techniques | Execute |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | 2 | 2 | | | | 2 | 1 | 2 | 2 |
| CO2 | 1 | 1 | 2 | | | | 2 | 1 | 2 | 2 |
| CO3 | 2 | 2 | 2 | | | | 2 | 2 | 2 | 2 |
| CO4 | 2 | 2 | 2 | | | | 2 | 2 | 2 | 2 |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

An Introduction to R. Data Visualisation: Visualization for Quantitative Data, Visualization for Qualitative Data, Q-Q Plots, Two Sample Analysis: Matched Pairs, Independent Samples, Table Analysis, Correlation.

**Module-II**

Integrate R to study the Simple Linear Regression Model and assumptions of the Simple linear regression model. Least square estimation, Inferences on the parameters, Confidence Intervals for the regression line. Prediction Intervals. Diagnostics, Residual Analysis, Remedial transformations, Outliers and leverage points, Simple regression based on weighted least squares.

**Module-III**

Integrate R to study Multiple linear regression, Polynomial regression, Cross-validation, Estimation and inference, Use R for Regression Diagnostic of Multiple Linear Regression, Transformations, and Detecting Multicollinearity. Variable selection:

**Module-IV**

Use R to study models for binary response variables, estimation and diagnosis methods for logistic and Poisson regressions. Over dispersion Prediction and Residual Analysis. Tree-based methods, Gradient Boosting.

**References :**


1. C.J Neter J. Kutner, W. Wasserman and M.H. Nachtsheim (2005), Applied Linear Statistical Models, 4th edition. McGraw Hill/Irwin.
2. Montgomery D.C., Peck E.A. and Vining, G.G. (2012) Introduction to Regression Analysis, Fifth edition. Wiley.
3. Crawley M.J (2013), The R Book, ISBN 978-0-470-97392-9 (hardback), John Wiley & Sons, Ltd.

## SEMESTERII

## 24-322-0201: STATISTICAL INFERENCE - I

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Summarize the desirable properties of estimator of a parameter or parameters of any given distribution | Evaluate |
| 2. | Relate complete sufficient statistic, Rao-Blackwell theorem and Lehmann-Scheffe theorem | Analyze |
| 3. | Relate Cramer-Rao, Chapmann-Robbin's and Bhattacharya bounds in connection with lower bound for the variance of an unbiased estimator | Analyze |
| 4. | Compute estimator of parameter or parameters of any given distribution using method of moments, method of maximum likelihood and method of minimum variance | Apply |
| 5. | Judge MLE of parameter or parameters of any given distribution possess its invariance and large sample properties | Evaluate |
| 6. | Compare classical inference and Bayesian inference | Analyze |
| 7. | Evaluate Bayes and minimax estimator of parameter or parameters of any given distribution under given prior density and loss function | Evaluate |
| 8. | Illustrate Metropolis-Hasting algorithm, Gibbs sampler and MCMC method | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | | | | | | | | |
| CO2 | | | | | | | 2 | 3 | | |
| CO3 | | | | | | | 3 | 2 | | |
| CO4 | | | | 3 | | | 2 | 3 | | |
| CO5 | | | | | 3 | | 2 | 2 | | |
| CO6 | 2 | | | | 2 | | 2 | | | |
| CO7 | | | | | | | | 3 | | |
| CO8 | | | | | | | | 3 | | |

**Module-I**

Point estimation: Sufficiency and minimal sufficiency, Exponential family of distributions, Pitman family, Factorization criterion, Likelihood equivalence, Unbiased estimation, Completeness, Ancillary statistics and Basu's Theorem, UMVUE estimators and their characterizations, Rao-Blackwell Theorem, Lehmann-Scheffe Theorem, UMVUE estimation of parametric functions from standard distributions.

**Module-II**

Fisher information measure and its properties, Fisher information matrix, Lower bound to the variance of an unbiased estimates, Cramer-Rao, Chapman-Robbin's and Bhattacharya bounds, BLUE of parametric functions, Efficiency, Consistency, Weak and strong consistency, Marginal and joint consistent estimators, Equivariance, Pitman estimators.

**Module-III**

Methods of estimation: Methods of moments, Maximum likelihood, Minimum chi square and its modification, Least square estimation, Properties of maximum likelihood estimators, Cramer-Huzurbazar Theorem, Likelihood equation - multiple roots, Iterative methods, E.M Algorithm.

**Module-IV**

Basic elements of Bayesian Inference, Loss function, Prior distribution, Bayes Theorem, Posterior distributions, Bayes risk, Bayes principle, Bayes estimators, Minimax estimators, Metropolis-Hastings algorithm, Gibbs sampler, MCMC method.

**References:**


1. E.L.Lehmann (1998) Theory of Point Estimation, John Wiley and Sons.
2. V. K. Rohatgi and A.K.L.Saleh (2015). An Introduction to Probability and Mathematical Statistics, Third Edition, Wiley.
3. B.K.Kale (2005) A First Course in Parametric Inference, Alpha Science International.
4. Robert C.P.and Casella G (2013) Monte Carlo Statistical Methods, Springer Verlag.
5. Rao C.R.(2009) Linear Statistical Inference and its Applications, Second Edition, Wiley.
6. Casella, Gand Berge R.L (2002) Statistical Inference, Second Edition, Thompson-Duxbury Press.
7. Mukhopadhay P.(1996) Mathematical Statistics, New Central Book Agency Pvt. Ltd.
8. Rajagopalan, M. and Dhanavanthan, P. (2012). Statistical Inference, PHI Learning Pvt. Ltd.

## 24-322-0202: PROBABILITY THEORY–II

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Employ the Inversion formula, Uniqueness theorem | Apply |
| 2. | Illustrate Convergence of distribution function characteristic functions and moments | Apply |
| 3. | Define Convergence of a series of in dependent random variables | Understand |
| 4. | Describe different forms of Central limit theorems | Understand |
| 5. | Define Conditional expectation and conditional probability | Understand |
| 6. | Demonstrate the Randon-Nikodym Theorem and its applications | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | | | | | | 2 | 2 | | |
| CO2 | 2 | | | | | 3 | 1 | 2 | | |
| CO3 | 2 | | | | | 3 | 2 | | | |
| CO4 | 2 | | | | | 2 | 3 | 1 | | |
| CO5 | 3 | | | | | 2 | | | | |
| CO6 | | | | | | | 2 | 3 | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Characteristic functions**:** Definition and simple properties, Inversion formula, Uniqueness theorem, Characteristic function and moments, Bochner's Theorem (Statement only), Convergence of distribution function: Weak convergence, Convergence of distribution functions and characteristic functions, Convergence of moments.

**Module-II**

Laws of Large Numbers**:** Convergence of series of independent random variables, Kolmogorov's inequality, Three series theorem, Weak law of large numbers (Kninchine's and Kolmogorov's), Kolmogorov's strong law of large numbers, Glivenko-Cantelli theorem, Kolmogorov's law of iterated logarithms (without proof).

**Module-III**

Limit Theorems**:** Central limit theorems for i.i.d random variables, Lindberg-Levy and Liaponov's CLT, Lindberg-Feller CLT, Infinitely divisible distributions--definition, elementary properties and examples, Canonical representation (without proof).

**Module-IV**

Conditioning: Conditional expectation and its properties, Conditional probabilities, Randon-Nikodym Theorem (Statement only) and its applications. Martingales, Submartingales, Martingale convergence theorem, Decomposition of submaritingales.

**References:**


1. Bhat B.R. (2011) Modern Probability Theory, Second edition, Wiley Eastern, Chapters 7,8,10, 11, 12.
2. Laha R.G. and Rohatgi V.K. (2020) Probability Theory, John Wiley and Sons, New York, Relevant sections of Chapters 2,4, 6.:
3. Billingsley P.(2017) Probability and Measure, Third edition, John Wiley, India Pvt. Ltd.
4. Feller W. (1976) An Introduction to Probability Theory and its Applications, Volume II Wiley Eastern.
5. Hoffmann-Jorgensen J.(1994) Probability with a view towards Statistics, Chapman & Hall.
6. Loeve M.(1977) Probability Theory, Volume I, Fourth edition, Springer-Verlag
7. Loeve M.(1978) Probability Theory, Volume II, Fourth edition, Springer-Verlag.
8. Rohatgi V.K. and Saleh A.K.E. (2015) An Introduction to Probability and Statistics, John Wiley & Sons, New York
9. Resnich S.I.(2005). A Probability Path. Birhauser, Springe

**24-322-0203: STOCHASTIC PROCESSES**

| | **Course Outcome (CO)** | **Cognitivelevel** |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand the classifications of random processes and concepts such as strict stationarity, wide-sense stationary and ergodicity | Understand |
| 2. | Classify the states of a Markov chain and apply the ergodic theorem for finding limiting distributions on states | Understand |
| 3. | Understand and apply Poisson, birth-death, renewal processes and Brownian motion | Apply |
| 4. | Describe and use the recurrence relation for generation sizes in a Branching Process and determine the probability of ultimate extinction | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | | | | | 3 | 3 | | |
| CO2 | 2 | | | | | | 3 | 3 | 1 | |
| CO3 | 3 | 2 | | | | 3 | 2 | 3 | | |
| CO4 | 2 | 2 | | | | 2 | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Markov Chains: Definition, Examples and classification, Discrete renewal equation and basic limit theorem, Absorption probabilities, Criteria for recurrence.

**Module-II**

Continuous time Markov chains, Examples, General pure birth process, Poisson process, Birth and death process, Finite state continuous time Markov chains, Applications to queuing models.

**Module-III**

Galton-Watson branching processes, generating function, Extinction probabilities, Continuous time branching processes, Extinction probabilities, Branching processes with general variable lifetime.

**Module-IV**

Renewal equation, Renewal theorem, Applications, Generalizations and variations of renewal processes, Applications of renewal theory, Brownian motion.

**References:**


1. Karlin .S. and Taylor H.M. (1975) A First Course in Stochastic Processes, Second edition, Academic Press, Relevant sections of Chapters 1, 2,3, 4, 5and 8.
2. Medhi J. (2020). Stochastic Processes, Fifth Edition, New Age International
3. Bhat B.R.(2002).Stochastic Processes, Second edition, New Age Publication.
4. Cinlar E.(2013).Introduction to Stochastic Processes, Dover Publications.
5. V. G. Kulkarni. (2010). Introduction to Modelling and Analysis of Stochastic Systems, Second Edition, CRC Press.
6. S. M. Ross.(2008).Stochastic Processes, Second edition, John Wiley

## 24-322-0204: PRACTICAL - I & VIVA VOCE

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Apply the different sampling methods for designing and selecting a sample from a population | Apply |
| 2. | Apply the methods of generating random numbers from different probability distributions and its goodness-of-fit using R software | Apply |
| 3. | Formulate and solve problems which involve setting up stochastic models | Evaluate |
| 4. | Understand the notion of parametric models, point and interval estimation of the parameters of those models using real data | Understand |
| 5. | Apply topics related to the Elective in Semester II using real data sets and interpretation of the results | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | | 2 | 3 | | 2 | | | | |
| CO2 | 2 | | | | | 2 | 3 | 2 | | |
| CO3 | 3 | | 2 | | | | 3 | | | |
| CO4 | 2 | | | | | 3 | 2 | 2 | | |
| CO5 | 3 | | | | | | 3 | 2 | | |

**Practicals based on topics covered in**

24-322- 0104 : Sampling Theory and methods; A sample survey to be executed
24-322- 0201 : Statistical Inference I
24-322- 0203 : Stochastic Processes
24-322- XXXX : Elective II

## ELECTIVE II

## 24-322-0205: STATISTICS FOR NATIONAL DEVELOPMENT

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Explain the concept of economic development, growth in per capita income and distributive justice | Understand |
| 2. | Define the indices of development like Human development index etc. | Understand |
| 3. | Estimate national income through income and expenditure approaches | Apply |
| 4. | Measure inequality in incomes, and measure poverty through measures of incidence and intensity combined | Analyze |
| 5. | Define components of Time series Remember Determine the trend, analyze seasonal Fluctuations, construct seasonal indices Measure cyclical movement | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | 1 | 2 | | | 3 | | | | |
| CO2 | | 2 | 2 | | | 1 | | 1 | 3 | |
| CO3 | | | 3 | | | 3 | | | 3 | |
| CO4 | 2 | | 2 | | | | | | 2 | |
| CO5 | 2 | 2 | 1 | | | 3 | | 2 | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Demographic methods:- Sources of demographic data - census, register, adhoc survey, hospital records, demographic profiles of Indian census; Measurement of mortality and life tables - crude, death rates, infant mortality rates, death date by cause, standardized death rate; Complete life tables – its main features, mortality rate and probability of dying, use of survival tables; Measurement of fertility - crude birth rate, general fertility rate, total fertility rate, gross reproduction rate, net reproduction rate; Population growth in developing and developed countries; Population projection using Leslie metric; Labour force projection.

**Module-II**

Economic statistics:- Index number - its definition, price relatives and quantity or volume relatives, link and chain relatives, consumer price index; Demand analysis - static laws of demand and supply, price elasticity of demand, analysis of income and allied size distribution - Pareto distribution, graphical test, fitting of Pareto's law, log normal distribution and its properties, Lorenz curve and estimation of elasticity; Gini coefficient.

**Module-III**

Economic development, growth in per capita income and distributive justice, indices of development; Human Development Index, Estimation of national income - product approach, income approach and expenditure approach; Measuring inequality in incomes, poverty measurement - measures of incidence and intensity combined; Time Series:-components of time series, determination of trend, analysis of seasonal fluctuations, construction of seasonal indices, measurement of cyclic movement, random component in time series, smoothing methods.

**Module-IV**

Introduction to Indian and International Statistical System - role, function activities of Central and State Statistical Organizations; Organization of large scale sample surveys; Role of National sample survey organization; General and special data dissemination systems; Principal publications containing such statistics on the topics - population, agriculture industry, trade, price, labour and employment transport and communications, and finance; Educational and Psychological statistics:-Scaling individual test items, scaling of scores on a test, different types of scores and scaling, scaling of ranking and rating in terms of normal curve, Reliability of test scores, Rulon and Kuder Richardson methods, Reliability of a test, validity, comparison between reliability and validity, Intelligence coefficient.

**References:**


1. Basic Statistics Relating to Indian Economy(CSO), 1990-Current Indian Statistics
2. CoxPR(1957) Demography, Cambridge University Press
3. Croxton F E and Crowder D J(1967) Applied General statistics, Prentice-HallIndia.
4. Guide to current Indian Official Statistics CSO, Govt. of India ,New Delhi
5. Guide to official Statistics(CSO)-1990
6. Kendall M.G. and Stuart A.(1966).The Advanced Theory of Statistics, Charles Griffin
7. Keyfitz N.(1977) Applied Mathematical Demography-Springer Verlag
8. Mukhopadhyay, P Applied Statistics, Books and Allied (P) Ltd
9. Pollard A H, Yusuf F and Pollard G.N.(1998) Demographic Techniques
10. Saluja M.P, Indian Official Statistical Systems, Statistics Publishing Society, Calcutta
11. Sen A.(1997): Poverty and inequality
12. Statistical System in Indian (CSO)1995
13. UNESCO: Principles for Vital Statistics system, SeriesM-12

## ELECTIVE II

## 24-322-0206: RELIABILITY MODELING AND ANALYSIS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand the various concepts and different notions of ageing used in Reliability analysis and their inter relations | Describe |
| 2. | Identify the various aspects like monotonic failure rates, Bath tub and upside down bathtub shaped failure rates and other related measures for various lifetime distributions | Evaluate |
| 3. | Understand and discover the system reliability using the concept of structure functions | Understand |
| 4. | Understand the concepts like positive dependency and various measures of dependence viz-RCSI, LCSD, PF2, WPQD and their inter relations | Evaluate |
| 5. | Estimate the reliability function for complete and censored samples through the maximum likelihood estimation | Evaluate |
| 6. | Estimate the reliability function for complete and censored samples through Uniformly minimum variance unbiased estimation | Evaluate |
| 7. | Estimate the reliability function for complete and censored samples through the Bayesian Estimation | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | | | | | | 2 | 3 | | |
| CO2 | 3 | | 2 | | | | 2 | 3 | 2 | |
| CO3 | | | 2 | | | | 1 | 3 | 1 | |
| CO4 | 1 | | 1 | 2 | | | 2 | 2 | 1 | |
| CO5 | 2 | | 2 | | | | 2 | 3 | 1 | |
| CO6 | 2 | | 2 | | | 3 | 2 | 3 | 1 | |
| CO7 | 2 | | 2 | | | 3 | 2 | 3 | 1 | |
| CO8 | | | | | | 2 | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Structure functions, Coherent Systems, Basic concepts in reliability: Failure rate, mean, variance and percentile residual life, identities connecting them; Notions of ageing - IFR, IFRA, NBU, NBUE, DMRL,HNBUE, NBUC etc and their mutual implications; TTT transforms and characterization of ageing classes.

**Module-II**

Structure function, Coherent systems, Reliability systems with dependent components:-Parallel and series systems, k out of n systems, ageing properties with dependent and independents components, concepts and measures of dependence in reliability - RCSI, LCSD, WPQD.

**Module-III**

Non monotonic failure rates and mean residual life functions, Study of lifetime models viz. exponential,Weibull, lognormal, generalized Pareto, gamma with reference to basic concepts and ageing characteristics: Bivariate Exponential distributions, Marshall-Olkin, Bath tub and upside down bath tub failure rate distributions.

**Module-IV**

Reliability estimation using MLE - exponential, Weibull and gamma distributions based on censored and non censored samples; UMVUE estimation of reliability function; Bayesian reliability estimation of exponential and Weibull models.

**References**:


1. Lai C.D and Xie M. (2006): Stochastic ageing and dependence in reliability (Relevant topics) Springer.
2. Sinha S K (1986) Reliability and Life Testing, Wiley Eastern.
3. Barlow R.E. and Proschan F. (1975) Statistical Theory of Reliability and Life Testing, Holt, Reinhart and Winston.
4. Marshall A.W. and Olkin I. (2007) Life Distributions, Springer
5. Galambos J and Kotz S. (1978) Characterization of Probability distributions, Springer
6. Lawless J.F.(2003) Statistical Models and Methods for Life Data, Wiley

## SEMESTER III

## 24-322-0301: STATISTICAL INFERENCE – II

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Summarize the testing problem in statistical testing problem | Evaluate |
| 2. | Evaluate MP and UMP tests corresponding to any given testing problem | Evaluate |
| 3. | Relate confidence interval estimation and testing of hypothesis | Analyze |
| 4. | Compute shortest confidence interval for parameter/s of any given distribution using different methods | Apply |
| 5. | Formulate LR test corresponding to any given testing problem | Evaluate |
| 6. | Construct SPRT corresponding to any given testing problem | Evaluate |
| 7. | Distinguish non-parametric confidence interval and boots trap Confidence intervals | Analyze |
| 8. | Examine the non-parametric alternatives for each parametric tests | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | | | | | 3 | | | |
| CO2 | | | | | | | 3 | 3 | | |
| CO3 | | | | | | 2 | 3 | 2 | | |
| CO4 | | | | | | | 2 | 3 | | |
| CO5 | | | | | | | | 3 | | |
| CO6 | | | | | | | | 3 | | |
| CO7 | 3 | | | | | 2 | | | | |
| CO8 | 2 | | | | | | | 3 | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Tests of hypotheses, Formulation of problem, Null and alternative hypotheses, Size of a test, Two kinds of errors, Simple and composite hypotheses, Randomized and non-randomized tests, Power of a test, Most powerful test, Neyman-Pearson lemma and its generalization, Monotone likelihood ratio property, UMP tests, Unbiased tests and UMPU tests with examples., Multiple hypothesis testing, False discovery rate.

**Module-II**

Confidence interval estimation, Relationship between confidence interval estimation and testing of hypothesis, UMA and UMAU confidence intervals, Shortest confidence intervals, Construction of confidence intervals using pivots, Large sample confidence interval based on maximum likelihood estimator, central limit theorem and Chebyshev's inequality, Bayesian credible regions.

**Module-III**

Likelihood ratio tests and their properties, Testing mean and variance of a normal population, Testing equality of means and variances of two normal populations, Sequential probability ratio tests, Construction of sequential probability ratio tests with examples.

.

**Module-IV**

Non-parametric inference: Goodness of fit tests- Chi square test and Kolmogorov Smirnov test for one and two sample problems, Sign test, Signed rank test, Wald-Wolfowitz run test, Median test, Man-Whitney U-test, Non-parametric confidence intervals, Bootstrapping confidence intervals, P-P Plot and Q-Q plot, Kendall's tau, Tests for independence and homogeneity.

.

**References:**


1. Lehmann E.L.(1998) Testing Statistical Hypothesis, John Wiley.
2. Wald A.(2013) Sequential Analysis, Doves
3. Gibbons J.K. (1971) Non-Parametric Statistical Inference, Mc Graw Hill
4. Rohatgi V.K. and Saleh A.K.E. (2011) An Introduction to Probability and Statistics, John Wiley and Sons.
5. Kale B.K. (2005) A First Course in Parametric Inference, Alpha Science International.
6. Rao C.R. (2009) Linear Statistical Inference and its Applications, Second Edition, Wiley.
7. Casella, Gand Berger R.L(2002) Statistical Inference, Second Edition, Thompson-Duxbury Press.
8. Rajagopalan M and Dhanavanthan P.(2012). Statistical Inference, PHI Learning Pvt. Ltd.
9. Dixit U.J.(2016).Examples in Parametric Inference with R Springer.

## 24-322-0302: MULTIVARIATE ANALYSIS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|
| **After completion of this course the student should be able to:** | | |
| 1. | Describe random vectors and their properties | Understand |
| 2. | Discuss the multivariate normal distribution and its properties | Understand |
| 3. | Understand the concept of Wishart distribution, distributions of simple, partial and multiple correlations and $T^2$ and $D^2$ statistics | Understand |
| 4. | Identify various classification methods for multivariate data and cluster analysis | Analyze |
| 5. | Explain principal component analysis and factor analysis | Evaluate |
| 6. | Identify canonical variables and quantify canonical correlation | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | 2 | 2 | | | | | | | |
| CO2 | | | 2 | | | 2 | | | | |
| CO3 | | | 3 | | | 2 | | | | |
| CO4 | | | | | | 2 | | | | |
| CO5 | | | 2 | | | 2 | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Multivariate data, preliminary analysis, notion of multivariate distributions, multivariate normal distribution, marginal and conditional distributions, characteristic function, estimation of mean vector and covariance matrix. Wishart distribution and its properties, distribution of simple, partial and multiple correlations based on samples from normal population.

**Module-II**

Hotelling's$T^2$ and Mahalanobis $D^2$ statistics, properties of $T^2$ and $D^2$, multivariate Fisher-Behren's problem. Testing independence of sets of variates, testing equality of covariance matrices and means, Sphericity tests, testing the hypothesis that a covariance matrix equal to given matrix, Mean and covariance equal to a given vector and given matrix.

**Module-III**

**Classification problem** - standards of good classification, procedures of classification into one of two populations with known probability distributions, classification into one of two known multivariate normal populations, classification into one of several populations. **Clustering of Observations**: Hierarchical clustering for continuous and categorical data - different choices of proximity measures, Agglomerative and Divisive algorithms, K-means clustering optimum choice of the number of clusters

**Module-IV**

**Principal component analysis**- definition, properties and ML estimation; canonical variables, canonical correlation. **Factor Analysis**: The orthogonal factor model, Estimation of factor loading, Factor rotation, Estimation of Factor scores, Interpretation of Factor Analysis. Multidimensional Scaling.

**References:**


1. Anderson T.W.: An Introduction to Multivariate Statistical Analysis, John Wiley
2. Brian Everitt and Torsten Hothorn( 2011) An Introduction to Applied Multivariate Analysis with R, Springer.
3. Johnson R. A. and Wichern, D.W.(2001): Applied Multivariate Statistical Analysis, Prentice Hall of India.
4. Rao C. R.: Linear Statistical Inference and its Applications, John Wiley
5. Rencher A. C. (2002): Methods of Multivariate Analysis, 2nd Ed., John Wiley & Sons.

## 24-322-0303: APPLIED REGRESSION ANALYSIS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Identify a linear and non linear regression problem | Apply |
| 2. | Model a data using an appropriate Regression model | Analyze |
| 3. | Identify and interpret a regression model | Understand |
| 4. | Examine model diagnostics | Analyze |
| 5. | Identify a Nonparametric Regression problem | Analyze |
| 6. | Apply Non Parametric Regression techniques | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | 2 | | | 2 | | | | | |
| CO2 | | | 2 | | 2 | | | | | |
| CO3 | | | | | | 2 | | 2 | | |
| CO4 | | 2 | | | | | 3 | | | |
| CO5 | | | | | | | 2 | | | |
| CO6 | | | | | | 2 | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Simple Linear Regression Model, Multiple linear regression model, Least squares estimation, Gauss Markov Theorem, Properties of the estimates, Distribution Theory, Maximum likelihood estimation, Hypothesis testing - likelihood ratio test, F-test; Confidence intervals.: Bonferroni-t-intervals, max modulus t intervals, Scheffes's method, Estimation with linear restrictions, Generalised least squares.(12+4+4hrs)

**Module-II**

Residual analysis, Departures from underlying assumptions, Effect of outliers, Collinearity, Non-constant variance and serial correlation, Departures from normality, Diagnostics and remedies.

**Module-III**

Polynomial regression in one and several variables, Orthogonal polynomials, Indicator variable Subset selection of explanatory variables, stepwise regression and Mallows Introduction to non-parametric regression.

**Module-IV**

Introduction to nonlinear regression, Least squares in the nonlinear case and estimation of parameters, Models for binary response variables, estimation and diagnosis methods for logistic and Poisson regressions. Prediction and residual analysis, Generalized Linear Models – estimation and diagnostics.

**References:**


1. Montgomery D.C., Peck E.A. and Vining G.G.(2001) Introduction to Regression Analysis, Third edition. Wiley.Chapter2, 3,
2. Seber A.F. and Lee A.J.(2003) Linear Regression Analysis, John Wiley, Relevant sections from chapters 3,4, 5,
3. Searle S.R.(1971) Linear models, John Wiley & Sons, Inc.
4. N. Draperand H. Smith (1986) Applied Regression Analysis–John Wiley & Sons.
5. Fox J. (1984) Linear Statistical Models and Related methods, John Wiley, Chapter 5.
6. Christensen R.(2001) Advanced Linear Modeling, Chapter7.
7. B. Abraham and Ledotter J.(1983) Statistical Methods fo-statistics, or Forecasting, John Wiley & Sons

## 24-322-0304: PRACTICAL - II, PROJECT & VIVA VOCE

| Course Outcome (CO) | Cognitivelevel |
|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand various tools using a programming software | Understand |
| 2. | Apply different statistical testing problems using real datasets and interpretation of the results | Analyze |
| 3. | Apply different multivariate techniques using real datasets and interpretation of the results | Analyze |
| 4. | Apply different regression techniques using real data sets and interpretation of the results | Evaluate |
| 5. | Apply topics related to the Elective in the Semester III using real data sets and interpretation of the results | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | | | | 3 | 3 | | | 2 | 3 |
| CO2 | 1 | 2 | | | 3 | 3 | 2 | | 2 | 3 |
| CO3 | | | | | 3 | 3 | 2 | | 2 | 3 |
| CO4 | | | | | 3 | 3 | 2 | | 2 | 3 |
| CO5 | | | | | 3 | 3 | 2 | | 2 | 3 |

**Practicals based on topics covered in**

24-322-0301; Statistical Inference II
24-322-0302: Multivariate Analysis
24-322-0303: Applied Regression Analysis
24-322-**xxxx**: Elective III

## ELECTIVE – IV

## 24-322-0305: TIME SERIES ANALYSIS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Define time series in time and frequency domain | Remember |
| 2. | Assess the stationarity of time series and its decomposition | Evaluate |
| 3. | Identify suitable ARMA models for the stationary component of the given time series | Analyze |
| 4. | Estimate the parameters of the identified models | Analyze |
| 5. | Discuss the validity of the model by residual analysis | Understand |
| 6. | Prediction by MMSE methods | Evaluate |
| 7. | Analyze Spectral density and periodogram | Analyze |
| 8. | Analyze time series in a state space setup | Analyze |
| 9. | Compute Smooth and filter by Kalmanalgorithm | Apply |
| 10. | Identify a model for the given time series | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | 3 | | | 3 | | 1 | 2 | | 1 |
| CO2 | 3 | 2 | | | 1 | | | 2 | | |
| CO3 | | | | | | | 2 | 3 | | |
| CO4 | 2 | 1 | | | 3 | | 2 | | | 1 |
| CO5 | 3 | | | | | | | 2 | | 1 |
| CO6 | | | | | | | | 3 | | |
| CO7 | | | | | | | | 2 | | |
| CO8 | | | | | | | | 3 | | |
| CO9 | | | | | | | | 2 | | |
| CO10 | 2 | | | | | | | | | |

**Module-I**

Characteristics of time series: Time series as a discrete parameter stochastic process, Auto-correlation function (ACF) and cross correlations, Stationary time series, Estimation of autocorrelations. Classical regression in time series context, exploratory data analysis, smoothing methods for time series. Wold representation of linear stationary processes.

**Module-II**

Linear time series models: Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models. Forecasting and estimation of ARMA models. Seasonal ARIMA models, Residual analysis and diagnostic checking. ARCH and GARCH models (Basic definition only). Non Gaussian time series (Basic definition only)

**Module-III**

Spectral analysis: Time series in frequency domain, spectral density, periodogram and discrete Fourier transforms, estimation of spectral density, multiple series and cross spectra, linear filters.

**Module-IV**

State space models: Filtering, smoothing and forecasting using state space models, Kalman smoother, Maximum likelihood estimation, Missing data modifications.

**References:**


1. Shumway R. H and Stoffer, D. S. (2006). Time series Analysis and its Applications, Springer.
2. Chatfield C. (2004). The Analysis of Time Series - An Introduction, Sixth edition, Chapman and Hall.
3. Box G. E. P. Jenkins, G. M. and Reinsel G. C. (1994). Time Series Analysis: Forecasting and Control, Pearson Education.
4. Brock well P.J and Davis R.A. (2006). Tim32e Series: Theory and Methods, 2ndedn Springer- Verlag
5. Abraham B. and Ledolter, J.C. (1983), Statistical Methods for Forecasting, Wiles 3. Anderson T.W (1971). Statistical Analysis of Time Series. Wiley.
6. Fuller W.A. (1978). Introduction to Statistical Time Series, Jolin Wiley
7. Kendall M.G. (1978). Time Series, Charler Graffin.
8. Tanaka K. (1996). Time Series Analysis, Wiley Series

## 24-322-0306: TOPICS IN MATHEMATICAL FINANCE

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|
| **After completion of this course the student should be able to** | | |
| 1. | Define the terms: interest rate, options, pay-off, arbitrage, Brownian geometric motion, mean reversion etc. | Remember |
| 2. | Describe and prove arbitrage theorem, Black Scholes theorem | Evaluate |
| 3. | Distinguish call and put options | Understand |
| 4. | Analyze portfolios viautility functions | Analyze |
| 5. | Apply CAPM | Apply |
| 6. | Assess the value at risk | Evaluate |
| 7. | Describe exotics by simulation | Understand |
| 8. | Employ and fit AR models for log prices | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | | | | | | | 2 | | |
| CO2 | | | | | | | | 3 | | |
| CO3 | | | | | | | 2 | 3 | | |
| CO4 | | | | | | | 3 | 3 | | |
| CO5 | | | | | | | | 3 | | |
| CO6 | | | | | | | 3 | 3 | | |
| CO7 | | | | | 3 | | | | | |
| CO8 | 2 | | | | | 3 | | 2 | | |

**Module-I**

Interest rate and Present value analysis, rate of return, Continuously varying interest rate. Options, Pricing contracts via arbitrage, Arbitrage theorem, single and multi-period binomial model.

**Module-II**

Geometric Brownian motion, The Black-Scholes formula, Properties of the Black-Scholes option cost, the delta hedging arbitrage strategy, Derivatives, Call options on dividend-paying securities, Pricing American put options.

**Module-III**

Adding jumps to geometric Brownian motion, Estimating the volatility parameter, Valuing investments by expected utility, The portfolio selection problem, Value at risk and conditional value at risk, The Capital Assets Pricing Model.

**Module-IV**

Exotic Options, Barrier options, Asean and look back options, Pricing exotic options by simulation, Pricing with nonlinear payoffs, Approximation via multi period binomial models, Crude oil data, Autoregressive moving average models for returns, Mean reversion.

**References:**


1. Sheldon M. Ross (2003). An Elementary Introduction to Mathematical Finance.
2. Cambridge University Press.
3. A.N. Shiryave (1999). Mathematical Finance, Theory and Practice, World Scientific.
4. David Ruper (2004). Statistics and Finance-An Introduction, Springer International Edition.
5. Fima C.Klebener (1997). Introduction to Mathematical Finance, World Scientific
6. John C.Hull (2008).Options, Futures and other derivatives, Pears on Education India.

## 24-322-0307: OPERATIONS RESEARCH

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Examine the properties of linear programming problem | Analyze |
| 2. | Solve different types of LPP | Apply |
| 3. | Solve LPP using duality | Apply |
| 4. | Employ transportation and assignment problems | Apply |
| 5. | Solvenon-linear programming problems | Apply |
| 6. | Explain quadratic and convex programming problems | Understand |
| 7. | Examine deterministic and probabilistic inventory models | Analyze |
| 8. | Employ inventory models in real situations | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | | 2 | | | 1 | | | | |
| CO2 | 2 | | 2 | | | | | | | |
| CO3 | 2 | | 2 | | | | | | | |
| CO4 | 3 | | | | | 1 | | | | |
| CO5 | 1 | | 1 | | | 1 | | | | |
| CO6 | 1 | | 1 | | | 1 | | | | |
| CO7 | 1 | | 2 | | | 1 | | | | |
| CO8 | | | 2 | | | | | | | |

**Module-I**

Linear programming:- convex sets and associated theorems, graphical method, definition of linear programming problem, properties of a solution to the linear programming problem, generating extreme-point solutions, simplex computational procedure, artificial variables technique - big M method, two phase method; Revised simplex method.

**Module-II**

Duality problems of linear programming:-unsymmetric primal-dual problems, symmetric primal-dual problems, Degeneracy and anticycling procedures:- perturbation techniques. Transportation problems:- general transportation problem, Finding initial basic feasible solution, test for optimality, degeneracy in transportation problem, unbalanced transportation problem, maximization transportation problem, Assignment problem:- mathematical formulation of the problem, the assignment method (Hungarian method).

**Module-III**

Non-linear programming problem (NLPP):- general non-linear programming problem, Constrained optimization with equality constraints - necessary conditions for a generalized NLPP, sufficient conditions for a general NLPP with one constraint, sufficient conditions for a general problem with $m(<n)$ constraints, Constrained optimization with inequality constraints - Kuhn-Tucker conditions for general NLPP with $m(<n)$ constraints, quadratic programming problem, convex programming problems.

**Module-IV**

Inventory models:- Deterministic inventory models - general inventory model, Static economic-order quantity (EOQ) models - classic EOQ model, EOQ with price breaks, multi-item EOQ with storage limitation, Probabilistic inventory models:- Continuous review models - "probabilitized" EOQ model, probabilistic EOQ model, Single-period models - No setup model (Newsvendor model), setup model (*s*-*S* policy).

**References:**


1. Gass S.I.(1985) Linear Programming-methods and applications, Fifth edition, McGraw Hill, USA, Chapters 2-7.
2. Kanti Swarup, Gupta,P.K. and ManMohan (2001) Operations Research, Ninth edition, Sultan Chand & Sons, Chapters 3,4, 10, 11 & 24.
3. Taha H.A. (2007) Operations Research- An introduction, Eighth edition, Prentice-Hall of India Ltd., Chapters 11, 14 &15.
4. Ravindran A, Philips D.T and Soleberg J.J.(1997) Operation Research-Principles and Practice, John Wiley & Sons.
5. Sinha S.M.(2006) Mathematical programming theory and methods, Elsevier, a division of Reed Elsevier India Pvt. Ltd., New Delhi.
6. Paneerselvam R.(2008) Operations Research, Second edition, Prentice Hall of India Pvt. Ltd., New Delhi.

## SEMESTER IV

## 24-322-0401: DESIGN AND ANALYSIS OF EXPERIMENTS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand the basic principles and guidelines of Design of experiments | Apply |
| 2. | Design and analyze CRDRBD, LSD and Greaco LSD | Apply |
| 3. | Apply in complete block designs in designing experiments and analyze them | Analyze |
| 4. | Understand and apply the factorial designs and its various versions | Apply |
| 5. | Apply Response surface methodology understanding various aspects involved in it | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | 2 | 3 | 3 | | | 3 | 3 | 3 | |
| CO2 | 3 | 3 | 3 | 3 | | | 3 | 3 | 3 | |
| CO3 | 3 | 3 | 3 | 3 | | | 3 | 3 | 3 | |
| CO4 | 3 | 3 | 3 | 3 | | | 3 | 3 | 3 | |
| CO5 | 3 | 3 | 3 | 3 | | | 3 | 3 | 3 | |

**Module-I**

Randomization, Replication and local control, One way and two way classifications with equal and unequal number of observations per cell with and without interaction, Fixed effects and Random effects model. Model adequacy checking, CRD, RBD and Latin Square designs, Analysis of co-variance for completely randomized and randomized block designs. Analysis of experiments with missing observations.

**Module-II**

Incomplete Block Designs: Balanced Incomplete Block designs, Construction of BIB Designs, Analysis with recovery of inter-block information and intra-block information. Partially balanced incomplete block designs, Analysis of partially balanced incomplete block designs with two associate classes, Lattice designs.

**Module-III**

$2^n$ Factorial experiments. Analysis of $2^n$ factorial experiments. Total confounding of $2^n$ designs in $2^P$ blocks. Partial confounding in $2^P$ blocks. Fractional factorial designs, Resolution of a design, $3^n$ factorial designs. Split plot design and strip plot design (outline only).

**Module-IV**

Response surface designs - orthogonality, rotatability blocking and analysis - Method of Steepest accent, Models, properties and Analysis.

**References:**


1. Montgomery D.C.(2001)) Design and Analysis of Experiments, John Wiley.
2. Das M N and Giri N.C.(1979) Design and Analysis of Experiments, second edition, Wiley.
3. Hinkleman and Kempthrone,C.(1994) Design and Analysis of Experiments-I, John Wiley.
4. Joshi D.D.(1987) Linear Estimation and Design of Experiments, Wiley Eastern.
5. Chakrabarti M.C.(1964) Design of experiments, ISI, Calcutta

## 24-322-0402: PRACTICAL – III &VIVA VOCE

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand the various computational techniques using R | Understand |
| 2. | Develop programming skill to meet the given Scientific objective | Analyze |
| 3. | Apply different DoE techniques using real data sets and interpretation of the results | Apply |
| 4. | Apply topics related to the Elective I in the Semester IV using real data sets and interpretation of the results | Apply |
| 5. | Apply topics related to the Elective II in the Semester IV using real data sets and interpretation of the results | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | | | 2 | 3 | 2 | | | | 3 |
| CO2 | | 2 | 2 | | 3 | 2 | | | | 3 |
| CO3 | | | | 3 | 2 | 3 | | | | 3 |
| CO4 | | | 2 | | 1 | 3 | 2 | | | 3 |
| CO5 | | | 2 | | 3 | 3 | 3 | | | 3 |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Practical based on topics covered in**

24-322-0401: Design and Analysis of Experiments.
Elective IV, Elective V, Elective VI

**ELECTIVE VI, VII, VIII (Choose any three)**

## 24-322-0404: STATISTICAL QUALITY ASSURANCE

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|
| **After completion of this course the student should be able to:** | | |
| 1. | Apply different statistical quality control techniques including various types sampling plans for attributes and measure the performance of these plans | Apply |
| 2. | Explain and design various types of control charts, design control charts and distinguish between them | Apply |
| 3. | Explain acceptance sampling by variables, Sampling Plans for a single and double specification limits with known and unknown variance | Analyze |
| 4. | Sampling plans with double specification limits | Apply |
| 5. | Compare sampling plans by variables and attributes and Continuous sampling plans I, II & III | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 1 | | | | | 2 | | | 2 | 2 |
| CO2 | 2 | 1 | 2 | | | 1 | | | 2 | 1 |
| CO3 | | 2 | 2 | | | 1 | | | 2 | |
| CO4 | | 2 | 1 | | | | 2 | | 2 | |
| CO5 | | | 2 | | | 2 | 2 | | 2 | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Quality and quality assurance, Methods of quality assurance, Introduction to TQM and ISO 9000 standards, statistical quality control: Acceptance sampling for attributes, Single sampling, Double sampling, Multiple and sequential sampling plans, Measuring the performance of these plans.

**Module-II**

Control charts, Basic ideas, designing of control charts for the number of non-conformities and fraction non-conformities, mean charts, Median charts, Extreme value charts, R-charts, and S-charts, ARL, Economic design of Shewarts control charts.

**Module-III**

Acceptance sampling by variables, Sampling plans for a single specification limit with known and unknown variance, Sampling plans with double specification limits, Comparison of sampling plans by variable and attributes, Continuous sampling plans I, II and III.

**Module-IV**

Process capability studies, Statistical aspect of six sigma philosophy, Lean concepts, Control charts with memory - CUSUM charts, EWMA-mean charts, OC and ARL for control charts, The Taguchi Method: The Taguchi philosophy of Quality, Loss functions, SN ratios, Performance measures, Experimental design in Taguchi Methods: Orthogonal arrays and linear graph, Estimation of effects, Parameter Design.

**References:**


1. Montgomery R.C.(1985).Introduction to Statistical Quality Control, Fourth Edition, Wiley.
2. Mittag H.J.& amp; Rinne H.(1993) Statistical Methods for Quality Assurance, Chapman & amp; Hall, Chapters 1, 3 and 4,15
3. The ISO 9000 book, Second Edition, Rabbit J T and Bergle P A Quality resources,Chapter-I
4. Schilling E.G.(1982)Acceptance Sampling in Quality Control, Marcel Dekker.
5. Amitava Mitra-Fundamentals of Quality Control and Improvement–Pearson Education Asia 2001 – Chapter 12 (relevant parts)
6. Duncan A.J.(1986) Quality control and Industrial Statistics.
7. Grant E.L. and Leaven Worth R.S.(1980) Statistical Quality Control, Mc Graw Hill.
8. Chin-KneiCho (1987) Quality Programming, John Wiley.

## 24-322-0405: ACTUARIAL STATISTICS

| Course Outcome (CO) | Cognitivelevel |
|---|---|
| **After completion of this course the student should be able to:** | |
| 1. Understand the concept of risk and lifetime tables | Understand |
| 2. Evaluate life insurance products, compound interest and discount factor, benefit payable at the moment of death, benefit payable at the end of year of death | Evaluate |
| 3. Understand and evaluate annuities and premiums | Understand and Evaluate |
| 4. Understand concepts related to Reserves - fully continuous reserves, fully discrete reserves; Multiple life contracts - joint life status | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 2 | 2 | 2 | | | 2 | 2 | | 2 | |
| CO2 | 2 | 2 | 1 | | | 2 | 2 | | 2 | |
| CO3 | 2 | 2 | 2 | | | 2 | 2 | | 2 | |
| CO4 | 2 | 2 | 2 | | | 2 | 2 | | 2 | |

**Module-I**

Insurance business - introduction, insurance companies as business organizations, concept of risk - failure lifetime distributions and life tables, future lifetime random variable, curtate future lifetime, life tables.

**Module-II**

Actuarial present values or benefit in life insurance products, compound interest and discount factor, benefit payable at the moment of death, benefit payable at the end of year of death – relation between these quantities.

**Module-III**

Annuities - annuities certain, continuous life annuities, discrete life annuities, life annuities with monthly payments, premiums- Loss at issue random variable, fully continuous premiums, fully discrete premiums.

**Module-IV**

Reserves - fully continuous reserves, fully discrete reserves; Multiple life contracts - joint life status, last survivor status.

**References:**


1. Desmukh S. R. (2009). Actuarial Statistics - An Introduction Using R, 3rd Edition. Universities Press (India) Private Ltd., Hyderabad.
2. Promislow, S. D. (2006). Fundamentals of Actuarial Mathematics, John Wiley, Chapters 2-11 and 14.
3. Dickson, C. M. D. (2016). Insurance Risk and Ruin, 2nd Edition. Cambridge University Press, United Kingdom.
4. Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A. And Nesbitt, C. J. (1997). Actuarial Mathematics, 2nd Edition. Society Of Actuaries, Illinois.

## 24-322-0406: LIFETIME DATA ANALYSIS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|
| **After completion of this course the student should be able to:** | | |
| 1. | Understand the basic concepts and ideas of survival analysis | Understand |
| 2. | Examine the properties and methods for standard survival time distributions | Analyze |
| 3. | Estimate survival functions using parametric and non-parametric methods | Evaluate |
| 4. | Apply and interpret semi-parametric and parametric regression models for survival data | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | 2 | | | | 3 | 1 | 2 | 2 | |
| CO2 | 2 | 2 | | | | 3 | 1 | 2 | 2 | |
| CO3 | 2 | 1 | 2 | | | 3 | 3 | 2 | 1 | |
| CO4 | 1 | 1 | 1 | | | 3 | 3 | 2 | 1 | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Basic Quantities and Models - Survival function, Hazard function, Mean residual life function, Common Parametric Models for Survival Data; Censoring and Truncation - Right Censoring, Left or Interval Censoring, Truncation, Likelihood Construction for Censored and Truncated Data, Counting Processes.

**Module-II**

Nonparametric Estimation of Basic Quantities for Right Censored and Left Censored Data - Estimators of the Survival and Cumulative Hazard Functions for Right Censored Data, Point wise Confidence Intervals for the Survival Function (without derivation), Estimators of the Survival Function for Left-Truncated and Right-Truncated Data; Estimation of the Survival Function for Left, Estimating the Hazard Function, Hypothesis Testing - One-Sample Tests, Tests for Two or More Samples.

**Module-III**

Semi-parametric Proportional Hazards Regression with Fixed Covariates - Coding Covariates, Partial Likelihoods for Distinct-Event Time Data, Partial Likelihoods when Ties are present, Model Building using the Proportional Hazards Model, Estimation for the Survival Function; Regression Diagnostics - Cox-Snell Residuals for assessing the fit of a Cox Model, Graphical Checks of the Proportional Hazards Assumption, Deviance Residuals.

**Module-IV**

Inference for Parametric Regression Models - Exponential, Weibull and Log Logistics; Multiple Modes of Failure – Basic Characteristics and Model Specification, Likelihood Function Formulation, Nonparametric Methods.

**References:**


1. Klein J.P. and Moeschberger M.L.(2003) Survival Analysis-Techniques for censored and truncated data, Second Edition, Springer-Verlag, New York,
2. Lawless J.F (2003) Statistical Models and Methods for Lifetime Data, Second Edition, John Wiley & Sons, Relevant Sections of the Chapters9.
3. Kalbfleisch J.D and Prentice, R.L. (2002) The Statistical Analysis of Failure Time Data, Second Edition, John Wiley &Sons Inc.
4. Hosmer Jr.D.W and Lemeshow S(1999)Applied Survival Analysis-Regression Modelling of Time to Event Data, John Wiley &Sons. Inc.
5. Nelson.W(2003) Applied Life Data Analysis.
6. Miller R.G.(1981) Survival Analysis, John Wiley.
7. Deshpande J.V. and Purohit S. G. (2006). Lifetime Data: Statistical Models and Methods. World Scientific.

## 24-322-0407: APPLIED MULTIVARIATE STATISTICAL ANALYSIS

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Distinguish multivariate data and its preliminary analysis | Understand |
| 2. | Examine properties of principal component analysis | Analyze |
| 3. | Apply PCA and canonical variates to real data | Apply |
| 4. | Analyze factor model | Understand |
| 5. | Illustrate the factor analysis | Apply |
| 6. | Outline different clustering and similarity techniques | Understand |
| 7. | Apply various clustering and similarity techniques | Apply |
| 8. | Infer multivariate data using MANOVA | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | 2 | | | | | 2 | | | |
| CO2 | | | 2 | | | | 2 | | | |
| CO3 | | | | | | 2 | | | | |
| CO4 | | | | | | 2 | | | | |
| CO5 | | | 2 | | | | | | | |
| CO6 | | | 2 | | | 2 | | | | |
| CO7 | | | | | | 2 | | | | |
| CO8 | | | 2 | | | 2 | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Multivariate Data, Types and preliminary methods of analysis, Principal components Analysis: - population principal components, summarizing sample variation by principal components, graphing the principal components, large sample inference, monitoring quality with principal components; Canonical correlation analysis: - canonical variates and canonical correlations, interpreting the population canonical variables, the sample canonical variates and sample canonical correlations.

**Module-II**

Factor analysis: - orthogonal factor model; methods of estimation, factor rotation, factor scores, perspectives and a strategy for factor analysis.

**Module-III**

Cluster analysis: - similarity measures, hierarchical clustering methods, non-hierarchical clustering methods; Distance methods: - multidimensional scaling, correspondence analysis.

**Module-IV**

Comparison of several multivariate population means (one-way MANOVA), simultaneous confidence intervals for treatment effects, two-way multivariate analysis of variance; profile analysis; Repeated measures designs and growth curves, path analysis.

**References:**


1. Johnson, R.A. and Wichern, D.W. (2007) Applied Multivariate Statistical Analysis, PHI Learning Private Ltd, New Delhi, Sixth edition, Relevant sections from Chapters 1, 6, 8, 9, 10 & 12.
2. DillonW.R.and Goldstein M (1984) Multivariate Analysis, John Wiley, Relevant sections from Chapter12.
3. Seber G.A.F. (1983) Multivariate Observations, Wiley.
4. Tabachnick B.G. and Fidell L.S.(2018) Using multivariate statistics, Sixth edition, Pearson India Education Services Pvt. Ltd, India.
5. Gnandesikan R., Methods of Statistical Data Analysis of Multivariate Observations, Wiley.
6. Jambu, MandLebeauxM.O., Cluster Analysis and Data Analysis.
7. Lebart, Lmorinean A. and Warwick K.M., Multivariate Descriptive Statistical Analysis, JohnWiley.
8. Davison, Multidimensional Scaling, John Wiley.
9. Morrison D.F., Multivariate Statistical Methods, Mc Graw Hill.
10. Rencher A.C.(1995) Methods of Multivariate Analysis, John Wiley.

## 24-322-0408: STATISTICAL FORECASTING

| | Course Outcome(CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Define various types offorecasts and their performance measures | Remember |
| 2. | Compute forecasts based on regression models | Apply |
| 3. | Compute forecast by simple and double exponential smoothing | Apply |
| 4. | Understand algorithms of Holt-Winters methods for forecasting | Understand |
| 5. | Estimate and forecast seasonal time series | Evaluate |
| 6. | Describe explicit forms of stationary and non-stationary time series models and derive the forecasts | Evaluate |
| 7. | Describe forecast formula for state-space models | Evaluate |
| 8. | Smoothing and filtering by Kalmanfilters | Analyze |
| 9. | Choose a model and construct forecast formula for a given time series | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | | | | | | 3 | 2 | | |
| CO2 | 2 | 3 | | | | 2 | | 2 | | 2 |
| CO3 | | | | | | | 3 | 3 | | |
| CO4 | | | | | | | | 3 | | |
| CO5 | | | | | | | 2 | 2 | | |
| CO6 | 2 | | | | | | | 3 | | |
| CO7 | | | | | | | 2 | 2 | | |
| CO8 | | | | | | | | 3 | | |
| CO9 | | | | | | | | 3 | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Review of linear regression and time series models. Forecasting in constant mean model, Locally constant mean model and simple exponential smoothing. Regression models with time as independent variable, Discounted least squares and general exponential smoothing. Locally constant linear trend model and double exponential smoothing, Prediction intervals.

**Module-II**

Seasonal time series, Globally constant seasonal models, Locally constant seasonal models, Winters' seasonal forecast procedures (additive and multiplicative), Seasonal adjustment procedures.

**Module-III**

Forecasts based on stationary ARMA and non-stationary ARIMA models. Transfer function models and forecasting.

**Module-IV**

State-space models- Filtering, smoothing and forecasting. Kalman filter. Vector ARMA models and Forecasting.

**References:**


1. Abraham B and Ledolter J (2005) Statistical Methods for Forecasting, John Wiley and Sons, NewYork.
2. Shumway R. HandStoffer, D.S.(2006). Timeseries Analysis and its Applications. Springer.
3. Montgomery D.C., Jennings C.L. and Kulachi M.(2011). Introduction to Time series analysis and Forecasting, Wiley Inter science.
4. Pankratz A. (2009) Forecasting with univariate Box-Jenkins models, John Wiley Sons, NewYork
5. Makridakis S, Wheel wright S.C., Rob J. Hyndman.(2005). Forecasting Methods and Applications, Third Edition, John Wiley and Sons
6. Box,G.E.P. Jenkins, G.M. and Reinsel G.C. (2015).Time Series Analysis: Forecasting and Control, Fourth Edition, Wiley.
7. Brockwell P.J. and Davis R.A. (2013) Introduction to Time Series and Forecasting, second edition, Forth Edition, Springer.

## 24-322-0409: INFERENCE FOR STOCHASTIC PROCESSES

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Define relevant optimality criteria for inference in stochastic processes | Remember |
| 2. | Choose suitable method of estimation and test procedure for given process | Evaluate |
| 3. | Compute estimates for relevant parametric functions for Markov chains in discrete and continuous time space | Apply |
| 4. | Produce the asymptotic properties of the estimators for such processes | Apply |
| 5. | Compute the estimates and test statistics for continuous time Markov processes such as Poisson processes, birth-death processes, etc. | Apply |
| 6. | Give examples for processes satisfying the regularity conditions under which estimators and test functions behave well | Understand |
| 7. | Identify a suitable stochastic model for the given situation | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 1 | 2 | | | | 2 | 2 | | 2 | |
| CO2 | 1 | 2 | | | | 1 | 2 | | 2 | |
| CO3 | | 1 | | | | 2 | | | 2 | |
| CO4 | 2 | | | | | | 2 | | 1 | |
| CO5 | | 2 | 1 | | | 2 | 2 | | 1 | |
| CO6 | 1 | 1 | 1 | | | 2 | | | 2 | |
| CO7 | 2 | 2 | 2 | | | 2 | | | 1 | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Brief review of basic principles of methods of statistical inference, Inference for the Galton-Watson process, The Markov branching process, Estimation and prediction in Auto regressive process.

**Module-II**

Inference in discrete Markov chains: Maximum likelihood estimation, Asymptotic properties of estimators, Asymptotic distribution of serial correlation, Tests of hypothesis tests of independence based on serial correlation Bayesian analysis, Inference for an absorbing chain Inverse likelihood estimation of states, Macro model, grouped Markov chains, Estimation in countable state-space Markov chain.

**Module-III**

Inference in continuous time Markov chains: Inference in finite Markov chains, queuing models, pure birth and death process, Homogeneous and non-homogeneous Poisson processes, Inference for renewal process in relation to reliability applications.

**Module-IV**

Large sample theory for discrete parameter stochastic process, Estimation, Consistency, Asymptotic normality, Efficiency, Robustness, Maximum likelihood estimation for some optimal asymptotic tests.

**References:**


1. Basava, I.V. and Prakasa Rao, B.L.S. (1980) Statistical Inference for Stochastic Processes Academic Press Chapters 1-7.
2. Billingsley,P. (1961) Statistical Inference for Markov Processes, University of Chicago Press.
3. Chung K.L. (1967) Markov Chain with Stationary Transition Probabilities 2nd edition, Springer-Varlag
4. Karr A.R.(1991) Point Processes and Their Statistical Inference, Marcel Dickker
5. Keiding N. (1974) Estimation in the Birth Process, Biometrika,61,71-80.
6. Keiding N. (1975) Maximum Likelihood Estimation in the Birth and Death Process, Annals of Statistics, 3,363-372.
7. Rajarshi M.B. (2013). Statistical Inference for Discrete time Stochastic Processes. Springer.

## 24-322- 0410 Categorical Data Analysis

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand a categorical data | Understand |
| 2. | Able to perform table analysis | Apply |
| 3. | Analyze categorical data using appropriate statistical models | Analyze |
| 4. | Able to use R software to effectively implement the above | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | 3 | 3 | | | 3 | 3 | 3 | 3 | |
| CO2 | 3 | 3 | 3 | | | 3 | 3 | 3 | 3 | |
| CO3 | 3 | 3 | 3 | | | 3 | 3 | 3 | 3 | |
| CO4 | | | | | | | | | | 3 |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-1**

Introduction to Categorical Data-Categorical Data, Discrete Distributions and Related Inference Problems, Statistical Inference with Categorical Data, Classes of Models for Discrete Data Analysis of Two-way Tables- Analyzing 2 2 Tables, Analyzing *IJ T*ables, Test of Independence for Ordinal Variables, Graphs for Two-way Tables

**Module-II**

Analysis of Multi-way Tables.- Describing Multi-way Contingency Tables, On Partial and Marginal Tables, Analysis of *K*2x2 Tables, Types of Independence for Three-way Tables, Graphs for Multi-way Contingency Tables

**Module-III**

Log-Linear Models-Log-Linear Models for Two-way Tables, On Inference and Fit of Log-Linear Models, Log-Linear Models for Three-way Contingency Tables, Hierarchical Log-Linear Models for Multi-way Tables, Maximum Likelihood Estimation for Log-Linear Models, Model Fit and Selection, Graphical Models, Collapsibility in Multi-way Tables

**Module-IV**

Generalized Linear Models and Extensions- The Generalized Linear Model (GLM), Log-Linear Model: Member of the GEM Family, Inference for GLMs, Software for GLMS, Independence for Incomplete Tables, Models for Joint and Marginal Distributions. Association Models-Basic Association Models for Two-way Tables, Maximum Likelihood Estimation for Association Models, Association Model Selection, Features of Association Models, Association Models of Higher Order: The RC(M) Model, Software Applications for Association Models, Association Models for Multi-way Tables.

**References:**


1. Kateri M. (2014). *Contingency Table Analysis: Methods and Implementation using R,* Springer.
2. Agresti A. (2013). *Categorical Data Analysis*, III Edition, John Wiley, New York.

## 24-322-0411 Directional Data Analysis

| Course Outcome (CO) | | Cognitivelevel |
|---|---|---|
| **After completion of this course the student should be able to:** | | |
| 1. | Understand a circular data and its corresponding distribution theory | Understand |
| 2. | Analyze the various statistical summary statistics associated with circular data | Analyze |
| 3. | Be able to estimate parameter of a circular distribution | Execute |
| 4. | Understand non parametric methods associated analysis of a circular data | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | 3 | 3 | | | 2 | 3 | 3 | 3 | |
| CO2 | 2 | 2 | 2 | | | 2 | 2 | 2 | 2 | |
| CO3 | 2 | 2 | 2 | | | 2 | 2 | 2 | 2 | |
| CO4 | 2 | 2 | 2 | | | 2 | 2 | 2 | 2 | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Graphical representation of data, Frequency distribution, Measures of location, circular variance and concentration, Correction for mean grouping, Measures of skewness and kurtosis.

**Module-II**

Circular models, Distribution theory, Independence, Convolution, Moments, Distributions of an arc, Mixtures, Lattice distributions, Wrapped normal, Cauchy, Poisson distributions, Von Mises, Fisher distribution characteristic functions, Polar distributions, Isotropic random walk on the circle.

**Module-III**

Point estimation, Cramer Rao type bound, Sufficiency, Methods of estimation, Testing hypothesis from parametric models, Neyman- Pearson and likelihood ratio principles.

**Module-IV**

Non-parametric methods: Tests for randomness, Goodness of fit, Rayleigh's test, Durand and Greenwood's test, Range test, Kuper's test, Watson's test, Uniform score tests, Runs test, Rank sum test, Tests for dispersion.

### References:


1. Mardia K.V. (1972). Statistics of Directional data, Academic Press.
2. Jammalamadaka S. R. & Sengupta A. (2001). Topics in circular statistics (Vol. 5), World Scientific
3. Batschelet E. (1981). Circular Statistics in Biology, Academic Press.
4. Watson G.S (1983). Statistics on Spheres, Wiley.

## 24-322-0412 STATISTICAL DECISION THEORY

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand the concept of decision rules | Understand |
| 2. | Understand and apply Bayesian concepts | Understand |
| 3. | Able to execute prior elicitation | Execute |
| 4. | Understand concepts of Game theory | Understand |
| 5. | Execute a complete Bayesian decision theoretic problems | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | 3 | 3 | 3 | | | 2 | 3 | 3 | 3 | |
| CO2 | 2 | 1 | 1 | | | 1 | 1 | 1 | 1 | |
| CO3 | 3 | 3 | 3 | | | 2 | 3 | 3 | 3 | |
| CO4 | 3 | 3 | 3 | | | 2 | 3 | 3 | 3 | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Basic elements of a decision problem, Randomized and non-randomized decision rules, Estimation and testing of hypothesis as decision problems, Bayes approach to inference and decision, Loss functions, Prior and posterior distributions, Prior - Posterior analysis for Bernoulli, Poisson and normal processes, Decision principles and Baye's risk.

**Module-II**

Utility theory, Axioms, Construction of utility functions, Sufficiency, Equivalence of Classical and Bayesian sufficiency, Complete and essentially complete classes of decision rules.

**Module-III**

Minimax analysis, Basic elements of game theory, General techniques of solving games, Finite games, Supporting and separating hyperplane theorems, Minimax theorem, Minimax estimation for normal and Poisson means.

**Module-IV**

Admissibility of Bayes and minimax rules, General theorems on admissibility, Robustness of Bayes rules, Invariant decision rules, Location parameter problems, Confidence and credible sets.

**References:**


1. James O. Berger (1980). Statistical Decision Theory and Bayesian Analysis, Springer Verlag.
2. M.H.DeGroot(1970).Optimal Statistical Decisions, John Wiley.
3. H. Raiffa and R. Schlaifer (2000). Applied Statistical Decision Theory, Wiley Classics.
4. Zellener (1971). An Introduction to Bayesian inference in Econometrics, Wiley.
5. Hayes J, Gand Winkler R I(1976). Probability, Statistics and Decision, Dower.
6. Anthony O' Hagan(1994). Kendall's Advanced theory of Statistics vol.2B, Bayesian Inference, John Wiley

## 24-322-0413 THEORY OF ENTROPY

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Describe measures of additive entropy and its properties | Understand |
| 2. | Illustrate Conditional entropy relative entropy and mutual Information | Analyze |
| 3. | Describe Renyi entropy, Conditional Renyi entropy | Understand |
| 4. | Discuss non-additive entropy and its properties | Understand |
| 5. | Describe maximum entropy principle | Evaluate |
| 6. | Compute various inequalities in information theory | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | 2 | | | | | | 2 | | |
| CO2 | | | 2 | | | | 2 | | | |
| CO3 | | 2 | | | 2 | | | | | |
| CO4 | | 2 | | | 2 | | | | | |
| CO5 | | | 3 | | | | 2 | | | |
| CO6 | | 3 | | | 2 | | | | | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Measures of Additive Entropy- Statistical Preliminaries, The Shannon entropy, Fadeev characterization, Tverberg characterization, Lee Characterization, Properties of Shannon entropy, Entropy rate, Conditional entropy and information, Chain rules for entropy, relative entropy and mutual information, Jensen's inequality and its consequences, the log sum inequality and its applications, The Renyi entropy and its characterizations, Conditional Renyi entropy.

**Module-II**

Measures of non-additive entropy- Polynomial entropy and their characterizations, non-polynomial algebraic entropy and characterizations, Transcendental entropy and characterizations, Tsallis entropy- definition and properties.

**Module-III**

Maximum entropy distributions, examples, Entropy based parameter estimation, Entropy as a criterion for goodness fit, dependence of entropy on sample size. Comparison of other parameter estimation methods with entropy method.

**Module-IV**

Inequalities in Information Theory- Basic inequalities of information theory, Differential entropy, Bounds on entropy and relative entropy, Inequalities for types, Entropy rates of subsets, Entropy and Fisher information, The entropy power inequality and the Brunn Minkowski inequality, Inequalities for determinants, Inequalities for ratios of determinants

**References:**


1. Behara M (1990) Additive and non-additive Measures of entropy, Wiley Eastern Limited.
2. Jagat Narain Kapur(1989)Maximum-entropy Models in Science and Engineering, John Wiley & Sons.
3. Robert M Gray (1990) Entropy and Information theory, Springer-Verlag
4. Thomas M.Cover and Joy A. Thomas (1991) Elements of Information Theory, John Wiley & Sons, Inc.
5. Tsallis C.(2023) Introduction to Non-extensive Statistical Mechanics: Approaching a Com-plex World, 2nded.;Springer: NewYork, NY,USA.
6. Vijay P Singh (1998) Entropy based Parameteres timation in hydrology, Kluwer Academic Publishers

## 24-322-0414 STATISTICS FOR CLINICAL RESEARCH

| | Course Outcome (CO) | Cognitivelevel |
|---|---|---|

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Understand the role of Statistics in Clinical Research | Understand |
| 2. | Understand modern statistical techniques relevant for today's scientific community | Understand |
| 3. | Identify the need for systematic analysis of data in any scientific experiment using statistical packages such as SAS Analysis | Analyze |
| 4. | Practice consultancy on experimental design and field survey | Apply |
| 5. | Solve real life problems using suitable statistical tools in numerous clinical domains | Apply |
| 6. | Understand basic theoretical and applied principles of statistics with adequate preparation to pursue a higher degree or enter job force as an applied Biostatistician | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 | PSO6 | PSO7 | PSO8 | PSO9 | PSO10 |
|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | | 1 | | | 1 | | | | | |
| CO2 | | 2 | | | 1 | | | | | 1 |
| CO3 | | | | 1 | | | | | | |
| CO4 | | | | 2 | | | | | 1 | |
| CO5 | | | | | | | | | 2 | |
| CO6 | | 1 | | | | | | | 1 | |

1– Slightly/Low, 2 – Moderate/Medium, 3-Substantial/High

**Module-I**

Drug Discovery and Development Duration: 1 week (4 hours) Introduction to clinical research, history, different phases, clinical data flow and ICH GCP E6 guidelines. Introduction to study documents (Protocol, SAP, CRF) and different activities related to Bios department. Different types of clinical study and study designs.

**Module-II**

SAS Programming Duration: 3 weeks (12 hours) Introduction to SAS, PDV, SAS statements and general rules of programming. Combining Datasets (Set and Merge), functions and controlling outputs. Data manipulations, various SAS procedures, introduction to SQL and Macros. V

**Module-III**

CDISC SDTM & ADaM Implementation Duration : 3.5 weeks (14 hours) Introduction to SDTM/ADaM Package, overview, and fundamentals of SDTM/ADaM. Introduction to SDTM/ADaM domains/classes assumptions and types. SDTM/ADaM Specification Development, Pinnacle 21 and submission package.

**Module-IV**

TLF Report Generation &Advance Biostatistics using SAS Duration: 4 weeks (17hours) Creation of Table/Listing/Figures (TLFs) using REPORT procedure. Introduction to Sensitivity & Tipping point analyses, statistical approaches to handle missing data& common imputation approaches. KM survival analysis using oncology examples and CSR case study using ICH GCP E3 guidelines.

**References:**


1. Overview of Clinical Research Industry https://www.centerwatch.com/clinical-trials/overview https://www.grandviewresearch.com/industry-analysis/global-clinical trials market#:~:text=Report%20Overview,countries%20is%20fueling% 20 market % 20 growth
2. Overview of Biological Sciences
   https://www.youtube.com/watch?v=URUJD5NEXC8 (Human Cell Structure)
   https://www.youtube.com/watch?v=e_1utfWwdD4 (Human Body System)
   https://www.youtube.com/watch?v=1vaEVcMfa1E (Cellular Death)
   https://www.youtube.com/watch?v=xlvLQtNF-Lk (Cell Multiplication)
   https://www.youtube.com/watch?v=fSEFXl2XQpc (Immune System)
   https://www.youtube.com/watch?v=SgVVGs7nqZA (Diabetes and its types)
   https://www.youtube.com/watch?v=46Xh7OFkkCE (Oncology Introduction)
3. SAS Programming 1 : Free online course with certificate https://support.sas.com/training/outlines/pg1.html#s1=2
4. ICH (International Conference on Harmonization): · E3 (Structure of Clinical Study Report) · E6 (Good Clinical Practice - GCP) · E9 (Statistical Principle for Clinical Trials) Guidelines
5. The Little SAS Book: A Primer, Fifth Edition
6. CDISC SDTM IG V3.3

# M.TECH. IN DATA SCIENCE AND ANALYTICS

A program offered by

**The Department of Statistics**

Under the Faculty of Science

DEPARTMENT OF STATISTICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
Kochi 682 022

# M.TECH. IN DATA SCIENCE AND ANALYTICS

*Under Faculty of Science*

## Objectives of the Programme:

Data scientists and professionals with data analytic skill are in great demand. The objective of this programme is to train graduates from different disciplines to become Data Scientists who can handle data analytics efficiently. Data Analytics is the application of structured statistical techniques on collected data in order to detect the underlying pattern as well as make predictions. The programme aims at learning Data Science via a comprehensive course curriculum covering Statistical tools and techniques, key programming languages such as R, Python, Machine learning algorithms and more.

Students are trained on the simultaneous application of statistics, computer programming, operations research and optimization techniques to analyze data patterns which help in effective decision making. The course also covers a broad spectrum of analytics and application knowledge areas required for overall professional development. The first two semesters are devoted to classroom teaching and laboratory practicals. In the third and fourth semesters, the candidates will be focusing on advanced level elective courses and they will be sent to undertake project work in industries of their choice.

## Program Outcomes:

On successful completion of 'M.Tech in Data Science and Analytics' program the students will be able to

1. Become Data science professionals.
2. Acquire knowledge in modern statistical and technology tools required for handling data analytics in a wide variety of application domains.
3. Critically investigate the prevailing complex problem scenarios in industries and arrive at optimal solutions by applying the acquired theoretical and practical knowledge.
4. Handle advanced level machine learning algorithms to solve complex problems and master the relevant statistical packages for interpreting the results.
5. Undertake teaching/research careers in Data Science/Statistics and allied areas and be involved in the process of knowledge discovery and effective communication of the same.

## Eligibility Criteria:

Pass in M.Sc (Statistics/Mathematics/Computer Science)/MCA or equivalent to any of this approved by this University with at least 60% marks or CGPA 6.5/10 or equivalent grade

OR

Pass in B.Tech in any branch with atleast 60% marks or CGPA 6.5/10 or equivalent grade

| | | |
|---|---|---|
| **Duration of the Programme** | : | Four Semesters |
| **Examination** | : | Credit and Semester |
| **Intake** | : | 18 |
| **Mode of admission** | : | Based on Admission Test |

# COURSE STRUCTURE (with effect from 2024 ADMISSION onwards)

## SEMESTER I

| Sl. No. | Course Code | Name of the Paper | Core/ Elective | Credits | CE Marks | ES Marks | Total Marks |
|---|---|---|---|---|---|---|---|
| 1 | 24-478-0101 | Mathematical Methods for Data Science | C | 3 | 50 | 50 | 100 |
| 2 | 24-478-0102 | Probability and Statistical Inference | C | 4 | 50 | 50 | 100 |
| 3 | 24-478-0103 | Data Structures and Algorithms | C | 3 | 50 | 50 | 100 |
| 4 | 24-478-0104 | Python Programming-Practical I | C | 2 | 50 | 50 | 100 |
| 5 | | Elective I | E | 3 | 50 | 50 | 100 |
| 6 | | Elective II | E | 3 | 50 | 50 | 100 |

Minimum Credits: 18 (Core: 12, Elective: 6)

### List of Electives for Semester I

| Sl. No. | Course Code | Name of the Paper |
|---|---|---|
| 1 | 24-478-0105 | Systems and Decision Analytics |
| 2 | 24-478-0106 | Data Warehousing and Data Mining |
| 3 | 24-478-0107 | Data Analysis and Visualization using Python |
| 4 | 24-478-0108 | Operations and Supply Chain Management |
| 5 | 24-478-0109 | System Reliability and Risk Analysis |
| 6 | 24-478-0110 | Computational Thinking |

## SEMESTER II

| Sl. No | Course Code | Name of the Paper | Core/ Elective | Credits | CE Marks | ES Marks | Total Marks |
|---|---|---|---|---|---|---|---|
| 1 | 24-478-0201 | Simulation Modelling and Analysis | C | 3 | 50 | 50 | 100 |
| 2 | 24-478-0202 | Machine Learning | C | 4 | 50 | 50 | 100 |
| 3 | 24-478-0203 | Multivariate Analysis and Statistical Techniques for Data Mining | C | 3 | 50 | 50 | 100 |
| 4 | 24-478-0204 | R/R-Studio Programming-Practical II | C | 2 | 50 | 50 | 100 |
| 5 | | Elective III | E | 3 | - | 100 | 100 |
| 6 | | Elective IV | E | 3 | 50 | 50 | 100 |

Minimum Credits: 18 (Core: 12, Elective: 6)

### Elective -III for Semester II

A MOOC course, from SWAYAM/NPTEL/Moodle/Others, will be opted for by students with the approval of the Department Council and will have the end semester evaluation ONLY, for 100 marks.

**List of Courses for Electives IV**

| Sl. No. | Course Code | Name of the Paper |
|---|---|---|
| 1 | 24-478-0205 | Optimization Techniques |
| 2 | 24-478-0206 | Design of Experiments (Integrated with R) |
| 3 | 24-478-0207 | Artificial Intelligence and Deep Learning |
| 4 | 24-478-0208 | Natural Language Processing |
| 5 | 24-478-0209 | Financial Risk Analytics and Management |
| 6 | 24-478-0210 | Marketing and HR Analytics |
| 7 | 24-478-0211 | Bioinformatics |
| 8 | 24-478-0212 | Big Data Technology |

## SEMESTER III

| Sl. No. | Course Code | Name of the Paper | Core/ Elective | Credits | CE Marks | ES Marks | Total Marks |
|---|---|---|---|---|---|---|---|
| 1 | 24-478-0301 | Project on Data Analytics in Industry | C | 15 | 150 | 150 | 300 |
| 2 | | Elective V | E | 3 | 50 | 50 | 100 |

Minimum Credits: 18 (Core: 15, Elective: 3)

**List of Electives for Semester III**

| Sl. No. | Course Code | Name of the Paper |
|---|---|---|
| 1 | 24-478-0302 | Business Analytics |
| 2 | 24-478-0303 | Statistical Forecasting Methods |
| 3 | 24-478-0304 | Quality Management and Six Sigma |
| 4 | 24-478-0305 | Applied Longitudinal Data Analysis |
| 5 | 24-478-0306 | Lifetime Studies in Data Science (Integrated with R). |
| 6 | 24-478-0307 | Bayesian Computing &Analysis |

## SEMESTER IV

| Sl. No | Course code | Name of the Paper | Core/ Elective | Credits | CE Marks | ES Marks | Total Marks |
|---|---|---|---|---|---|---|---|
| 1 | 24-478-0401 | Project Dissertation Evaluation and Viva | C | 18 | 200 | 200 | 400 |

Minimum Credits: 18 (Core: 18)

*Additional electives from Industry/Institutions can be offered during third and fourth semesters with the approval of Department Council and University.

# DETAILED SYLLABUS

## 24-478-0101: MATHEMATICAL METHODS FOR DATA SCIENCE

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the applications of vectors and matrices in data science | Understand |
| 2. | Represent the problems in the form of a system of linear equations and how to construct machine learning algorithm | Evaluate |
| 3. | Identify the applications of calculus in optimization methods such as least squares, machine learning, etc. | Apply |
| 4. | Formulate an optimization problem | Apply |
| 5. | Prepare algorithm to solve optimization problems | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | 2 | | 2 | 2 |
| CO2 | 3 | | 2 | 3 | 3 |
| CO3 | | 2 | | 2 | |
| CO4 | 3 | 2 | | 2 | |
| CO5 | 1 | 2 | 2 | | 3 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Basic properties of matrix and vectors: linear transformation, rank, determinant, inner and outer products, matrix multiplication rule and various algorithms, matrix inverse.
Special matrices: square matrix, identity matrix, triangular matrix, idea about sparse and dense matrix, unit vectors, symmetric matrix, Hermitian, skew-Hermitian and unitary matrices, Generalized inverses.

**Module 2:**

Matrix factorization concept/LU decomposition, Gaussian/Gauss-Jordan elimination, solving Ax=b linear system of equation,vector space,basis, span, orthogonality, orthonormality, linear least square, Eigen values, eigenvectors, diagonalization, singular value decomposition, Quadratic forms ,Spectral decomposition.

**Module 3:**

Functions of a single variable, limit, continuity, differentiability, Mean value theorems, Product and chain rule, Taylor's series, infinite series summation/integration. Fundamental and mean value theorems of integral calculus, evaluation of definite and improper integrals, Beta and gamma functions, Functions of multiple variables, limit, continuity, partial derivatives, Basics of ordinary and partial differential equations, numerical integration.

**Module 4:**

Sets, subsets, power sets, counting functions, combinatorics, countability, Basics of optimization techniques: problem formulation, Maxima, minima, convex function, global solution, Linear programming, simplex algorithm, Integer programming, Constraint programming, Randomized optimization techniques: hill climbing, simulated annealing, genetic algorithms.

**Reference Books:**

1. Gilbert Strang (2014) .Linear Algebra and Its Applications, 4th Edition, Cengage Learning India Pvt. Ltd.

2. Marc Peter Deisenroth, A.Aldo Faisal and Cheng Soon Ong (2020). Mathematics for Machine Learning .Cambridge University Press.

3. AvrimBlum, John Hopcroft, Ravindran Kannan; Foundations of Data Science, 2018
https://www.cs.cornell.edu/jeh/book.pdf

# 24-478-0102: PROBABILITY AND STATISTICAL INFERENCE

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Describe sample space and events | Understand |
| 2. | Understand axiomatic probability and know how to model real-world problems using it | Understand |
| 3. | Understand Discrete and Continuous Random Variables, their Distributions, their properties, moments and correlations | Understand |
| 4. | Understand the limiting behavior of large amounts of data by limit theorems | Understand |
| 5. | Demonstrate conceptual understanding of sampling distributions and the central limit theorem | Apply |
| 6. | Estimate parameters of distributions using maximum Likelihood and Other estimation techniques | Evaluate |
| 7. | Assess the properties of estimators | Evaluate |
| 8. | Construct confidence intervals for point estimates | Evaluate |
| 9. | Test hypotheses about different parameters of distributions of populations using samples of data | Analysis |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | | | | | 1 |
| CO2 | | | | | 1 |
| CO3 | 2 | | | | |
| CO4 | 2 | | | | 2 |
| CO5 | | 3 | | | |
| CO6 | 1 | 2 | | | |
| CO7 | | | | | |
| CO8 | | | | | 2 |
| CO9 | 2 | 2 | | | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Sample space, Probability axioms, Probability on finite sample spaces, conditional probability and Bayes theorem, independent of events, random variables, probability distribution of a random variable, discrete and continuous random variables, functions of random variables, moments of a distribution function, generating functions, Discrete distributions: Binomial distribution,negative binomial distribution, geometric distribution, hypergeometric distribution, negative hypergeometric distribution, Poisson distribution, Continuous distribution: Uniform distribution, exponential distribution, gamma distribution, beta distribution, Normal distribution

**Module 2:**

Mode of convergence, weak law of large numbers, strong law of large numbers, central limit theorems, random sampling, sample characteristics and their distributions, Chi-square, t-and F-distributions, large sample theory.

**Module 3:**

Concept of sufficiency, minimal sufficiency, Neyman factorization criterion, Fisher information, exponential families, Maximum likelihood estimators, method of moment estimators, percentile estimators, least squares estimators, minimum mean squares estimators, uniformly minimum variance unbiased estimators, Rao-Blackwell theorem, Cramer-Rao lower bound, different examples.

**Module 4:**

Statistical Hypotheses-simple and composite, statistical tests, critical regions, Type-I and Type-II errors, size and power of a test, Neyman Pearson lemma and its different applications. Most powerful test, uniformly most powerful test, unbiased test and uniformly most unbiased test, Likelihood ratio test, Interval estimation, confidence intervals, construction of confidence intervals, shortest expected length confidence interval, most accurate one-side d confidence interval and its relation to UMP test.

**Reference Books:**

1. V.K. Rohatgi and A.K .Md. Saleh (2001) An introduction to probability and Statistics, Wiley India.
2. George Casella and Roger L Berger (2002) Statistical Inference, Thomson Learning.
3. Efron B. and Hastie T (2017). Computer Age Statistical Inference-Algorithms, Evidence and Data Science, Cambridge University Press.
4. Avrim Blum, John Hopcroft, Ravindran Kannan; Foundations of Data Science, 2018 https://www.cs.cornell.edu/jeh/book.pdf

# 24-478-0103: DATA STRUCTURES AND ALGORITHMS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Basic ability to analyze algorithms and to determine algorithm correctness | Analyze |
| 2. | Choose appropriate data structures and algorithms, understand the ADT/libraries, and use it to design algorithms for a Specific problem | Evaluate |
| 3. | Understand the necessary mathematical abstraction to solve problems | Analyze |
| 4. | Comprehend and select algorithm design approaches in a problem specific manner | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | | | | 2 | |
| CO2 | | | 2 | | |
| CO3 | 2 | | 2 | | |
| CO4 | | | 1 | | 1 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to Basic Data Structures: Importance and need of good data structures and algorithms, Linked lists, Queues, Heaps, Hash tables, Binary search trees.

**Module 2:**

Advanced Data Structures: Red-Black Trees, B-trees, Fibonacci heaps, Data Structures for Disjoint Sets. Design Strategies: Divide-and-conquer, Dynamic Programming, Greedy Method.

**Module 3:**

Internal and External Sorting algorithms: Linear Search, Binary Search, Bubble Sort, Insertion Sort, Shell Sort, Quick Sort, Heap Sort, Merge Sort, Counting Sort, Radix Sort. Advanced String-Matching Algorithms: The naive string-matching algorithm, Rabin-Karp, String matching with finite automaton, Knuth-Morris-Pratt algorithm.

**Module 4:**

Graph Algorithms: Elementary graph algorithms, Minimum spanning trees, shortest path algorithms: single source and all pair, Max flow problem and its solutions, Graph coloring problem and its solutions, Bio-inspired algorithms: Swarm Intelligence, Ant Colony Optimization, and recent trends in data structures.

**Reference Books:**


1. Cormen T.H., Leiserson, C.E., Rivest R.L. and Stein C (2010). Introduction to Algorithms 3rded. MIT Press.
2. S.Sridhar, Design and Analysis of Algorithms, Oxford University Press India, 2014
3. Aho A.V., Hopcroft J.E and Ullman J.D. Data Structures and Algorithms. India: Pearson Education,2009
4. Horowitz E., Sahni S. and Rajasekaran S. Fundamentals of Computer Algorithms, Galgotia Publications, 2010.
5. Weiss M.A, Data Structures and Algorithm Analysis in C++, India: Pearson, Third Edition, 2014.
6. Horowitz, Sahni and Anderson-Freed, Fundamentals of Data Structures in C, Universities Press, 2008

# 24-478-0104: PYTHON PROGRAMMING – PRACTICAL I

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Define python environment | Understand |
| 2. | Explain python programming language | Evaluate |
| 3. | Develop a scientific computing environment using python | Evaluate |
| 4. | Identify the use of python software to meet the given Scientific objective | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | | 1 | | | |
| CO2 | | 3 | | 3 | 3 |
| CO3 | | 3 | | 3 | 3 |
| CO4 | | 1 | | 1 | 3 |
| CO5 | | | | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction: The Process of Computational Problem Solving, Python Programming Language, Python Data Types: Expressions, Variables and Assignments, Strings, List, Objects and Classes, Python Standard Library, Imperative Programming: Python programs, Execution Control Structures, User-Defined functions, Python Variables and Assignments, Parameter Passing. Text Files: Strings, Formatted Output, Files, Errors and Exception Handling, Execution and Control Structures: if Statement, for Loop, Two Dimensional lists, while Loop, More Loop Patterns, Additional Iteration Control Statements, Containers and Randomness: Functions: Definition, Call, Arguments, Scope rules and Name resolution, Lambda functions Built-in data types for data collections: Lists, Tuple, set and dictionary: Basic Operations, Iteration, Indexing, Slicing and Matrices; Dictionaries: Basic dictionary operations Error Handling and debugging.

**Module 2:**

Python Modules; Python file handling: Import, Write/create files, Delete files; Scientific Programming with Python: 1. Numpy Basics – Creating array, indexing, slicing, data types, Copy and view, iteration, Split, search, sort, filter; 2. Pandas Basics: Introduction to Pandas and its data structures (Series, Data Frame), Reading and writing data with Pandas, Data cleaning and preprocessing with Panda, Checking Null values, remove duplicates, correlations, plotting, 3. Plotting with matplotlib and seaborn.

**Reference Books:**


1. McKinney W. (2013). Python for Data Analysis. India: O'Reilly Media, Incorporated.
2. Charles Dierbach. (2012). Introduction to computer science using Python a computational problem-solving focus, John-Wiley & Sons.
3. McKinney W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython. United States: O'Reilly Media.
4. Perkovie L. (2011). Introduction to computing using python: An Application development focus. Wiley Publishing.
5. McKinney W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy and IPython. "O Reilly Media, Inc."

# 24-478-0105: SYSTEMS AND DECISION ANALYTICS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the basic concepts of systems and system thinking | Understand |
| 2. | Apply diagrammatic aids to capture problem situation | Apply |
| 3. | Identify elements for an optimization model | Evaluate |
| 4. | Use the basics tools for system studies such as system dynamics and simulations | Apply |
| 5. | Evaluate decision problems and perform utility or risk analysis | Evaluate |
| 6. | Develop decisions aids for informed decisions using hard/soft OR methodology | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | | 3 | | 2 |
| CO2 | | 3 | | | |
| CO3 | | | 2 | 2 | |
| CO4 | 2 | | | 2 | |
| CO5 | | | | 1 | |
| CO6 | | 1 | | | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Overview of decision-making process; role of analytics in decision making; Systems thinking, efficiency and effectiveness, unplanned and counter intuitive outcomes, reductionist and cause-and-effect thinking, black boxes approach, hierarchy of systems, feedback loops, control of systems, mind maps, cognitive mapping, causal loop diagrams, influence diagrams and other system diagrams.

**Module 2:**

Soft systems thinking, Checkland's soft systems methodology; Hard OR paradigm, problem scoping and modelling phase, the implementation phase, code of ethics
Decision making over time: planning horizon, production planning problem.

**Module 3:**

Prescriptive Analytics: Linear Optimization: Identifying elements for an Optimization Model - Translating Model -Solving Linear Optimization Models -Graphical Interpretation of Linear Optimization - Solving Models with General Integer Variables; Simulation and system dynamics: the structure of simulation models, computer simulation packages, basics of system dynamics.

**Module 4:**

Decision Analysis - Formulating Problems - Strategies for Minimizing and Maximizing Objectives - Conflicting Goals - Strategies with Known Probabilities - Average Payoff and Expected Value - Decision Trees and Monte Carlo Simulation - Value of Information - Sample Information - Bayes's Rule - Utility - Multicriteria Decision Making - Analytic Hierarchy Process.

**Reference Books:**

1. H.G. Daellenbach and D.C. McNickle, Management Science, Decision Making Through Systems Thinking, Palgrave Macmillan, 2005

2. F.S.Hillier and J.Lieberman, Introduction to Operations Research, Tenth Edition, McGraw-Hill, New York, 2015

3. Law.A.M, Simulation Modeling and Analysis, Fifth edition, McGraw-Hill New York, 2015
4. J.R.Evans, Business Analytics, Pearson Education; Second edition, 2017.
5. K.C. James, Systems. Models and Decision Making, Amazon Publishing, 2023

# 24-478-0106: DATA WAREHOUSING AND DATA MINING

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the functionality of the various data mining and data warehousing components | Understand |
| 2. | Examine the types of the data to be mined and apply pre-processing methods on raw data | Evaluate |
| 3. | Expose various real-world data mining applications | Apply |
| 4. | Apply the functionality of the various data warehousing components | Apply |
| 5. | Develop an understanding of the data mining techniques and issues | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | | | 3 | 3 |
| CO2 | 2 | | | 3 | 3 |
| CO3 | | | 3 | | |
| CO4 | 3 | | | | 3 |
| CO5 | 3 | | | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Data Warehouse: Introduction to Data Ware House, Differences between operational data base systems and data Ware House, Data Ware House characteristics, Data Ware House Architecture and its components, Extraction-Transformation Loading, Logical (Multidimensional), Data Modeling, Schema Design, star and snow-Flake Schema, Fact Constellation, Fact Table, Fully Addictive, Semi-Addictive, Non-Addictive Measures; Fact Less-Facts, Dimension Table characteristics; Fact-Less-Facts, Dimension Table characteristics; OLAP cube, OLAP Operations, OLAP Server Architecture-ROLAP, MOLAP and HOLAP.

**Module 2:**

Introduction to Data Mining: Introduction, What is Data Mining, Definition, KDD, Challenges, Data Mining Tasks, Data Pre-processing - Data Cleaning, Missing Data, Dimensionality Reduction, Feature Subset Selection, Discretization and Binaryzation, Data Transformation; Measures of similarity and dissimilarity-Basics.

**Module 3:**

Association Rules: Problem Definition, Frequent Item Set Generation, The Apriori Principle, Support and Confidence Measures, Association Rule Generation, Apriori Algorithm, The Partition Algorithms, FP-Growth Algorithms, Compact Representation of Frequent Item Set-Maximal Frequent Item Set, Closed Frequent Item Set.

**Module 4:**

Classification: Problem definition, General Approaches to solving a classification problem, Evaluation of Classifiers, Classification techniques, Decision trees, Naïve-Bayes Classifier, K-nearest neighbor classification-Clustering: Partitioning clustering K-Means Algorithm, Hierarchical Clustering-Algorithm- Agglomerative Methods and Divisive Methods, Basic Agglomerative Hierarchical Clustering Algorithm

**Reference Books:**

1. Jiawei Han and Micheline Kamber, Data Mining-Concepts and Techniques, Morgan Kaufmann Publishers, Elsevier, 3 Edition, 2012.

2. Pang-NingTan, Vipin Kumar, Michael Steinbanch Introduction to Data Mining, 2nd edition, Pearson Education, 2018.

3. Parteek Bhatia, Data Mining and Data Warehousing: Principles and Practical Techniques, Cambridge University Press, 2019.

4. Alex Berson and Stephen Smith, Data Warehousing, Data Mining & OLAP. McGraw Hill Education, 2017

# 24-478-0107: DATA ANALYSIS AND VISUALIZATION USING PYTHON

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Write an efficient program using python to perform routine and specialized data manipulation/management and analysis tasks | Evaluate |
| 2. | Develop student's knowledge in the area of Data Science with emphasis on predictions using associated statistical techniques and software tools | Evaluate |
| 3. | Develop various quantitative and classification predictive models Based on various regression and decision tree methods | Evaluate |
| 4. | Apply specific statistical methods applicable to predictive analytics to identify new trends and patterns, uncover relationships, create forecasts, predict likelihood sand test predictive hypothesis | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | | | | |
| CO2 | | 3 | | | 3 |
| CO3 | | | | 3 | |
| CO4 | | | 3 | 3 | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to Data Acquisition-Applications-Process-Data Extraction-Data Cleaning and Annotation-Data Integration-Data Reduction-Data Transformation-Visualization-Introduction-Terminology-Basic Charts and Plots-Multivariate Data Visualization- Data Visualization Techniques-Pixel-Oriented Visualization Techniques-Geometric Projection Visualization Techniques-Icon-Based Visualization Techniques-Hierarchical Visualization Techniques-Visualizing Complex Data and Relations.

**Module 2:**

Data Visualization Tools-Rank Analysis Tools- Trend Analysis Tools-Multivariate Analysis Tools-Distribution Analysis Tools-Correlation Analysis Tools-Geographical Analysis Tools.

**Module 3:**

Regression model building framework: Problem definition, Data pre-processing; Model building; Diagnostics and validation Simple Linear Regression: Coefficient of determination, Significance tests, Residual Analysis, Confidence and Prediction Intervals.

**Module 4:**

Multiple Linear Regression: Coefficient of multiple coefficients of determination, Interpretation of regression coefficients, Categorical variables, Heteroscedasticity, Multi-collinearity, outliers, Auto regression and transformation of variables, Regression model building.

**Reference Books:**


1. Andy Kirk, Data Visualization a Handbook for Data Driven Design, Sage Publications, 2016
2. Philipp K.Janert, Gnuplotin Action, Understanding Data with Graphs, Manning Publications, 2010
3. Alberto Cordoba, "Understanding the Predictive Analytics Lifecycle", Wiley, 2014.

4. Eric Siegel, Thomas H. Davenport, “Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die”, Wiley, 2013.

5. James R Evans, “Business Analytics-Methods, Models and Decisions”, Pearson 2013.

6. R. N.Prasad, Seema Acharya, “Fundamentals of Business Analytics”, Wiley, 2015.

7. Perkovie,L.(2011).Introduction to computing using python: An Application development focus. Wiley Publishing.

8. McKinney W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy and IPython. “OReilly Media, Inc.”

# 24-478-0108: OPERATIONS AND SUPPLY CHAIN MANAGEMENT

**After completion of this course the student should be able to:**

| | | |
|---|---|---|
| 1. | Describe operations management, objectives and Identify strategic approaches for competitive advantage | Understand |
| 2. | Compute single-factor productivity and multi fact or productivity And identify the critical variables in enhancing productivity | Evaluate |
| 3. | Understand the strategic importance of forecasting and Select appropriate method for forecasting in supply chains | Evaluate |
| 4. | Describe basics of supply chain, performance and processes | Understand |
| 5. | Formulate and Solve basic inventory order related problems in supply chain | Evaluate |
| 6. | Understand and apply analytical tools for supply chain decisions | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | | 1 | 2 | | 1 |
| CO2 | | | 2 | | |
| CO3 | 3 | 2 | | 1 | 2 |
| CO4 | | | 1 | | |
| CO5 | | | 1 | | |
| CO6 | 2 | 1 | 2 | | 1 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Operations, Logistics and Supply Chain Management: Definitions and Objectives- operations strategy and competitiveness-Basic concepts of operations management-types of manufacturing systems and their characteristics - Basics of Product and Process Design -Goods and Services. The Importance of Project Management - Project Planning, Scheduling and Controlling-PERT and CPM

**Module 2:**

Forecasting: The Strategic Importance of Forecasting - Steps in the Forecasting – Forecasting Approaches - Time-Series Forecasting -Associative Forecasting Methods: Regression and Correlation Analysis - Monitoring and Controlling Forecasts - Forecasting in the Service Sector. Process Analysis and Design - Process Redesign. The Strategic Importance of Location - Methods of Evaluating Location Alternatives - Job Design - Ergonomics and the Work Environment-Methods Analysis.

**Module 3:**

Understanding supply chain, supply chain performance - supply chain drivers and obstacles -Supply chain micro and macro processes - Push and pull systems - Aggregate planning in supply chain -planning supply and demand - managing predictable variability - Economic Order Quantity Models, Reorder Point Models - Manufacturing Planning and Control Systems.

**Module 4:**

Supply Chain Management Analytics: features of supply chain analytics - Techniques for Evaluating Supply Chains - Evaluating Disaster Risk in the Supply Chain - Managing the Bullwhip Effect - Supplier Selection Analysis - Transportation Mode Analysis, Warehouse Storage-Stochastic Inventory Models.

**Reference Books:**

1. Jay Heizer, Barry Render and Chuck Munson, Operations management; sustainability and supply chain management, Twelfth Edition, Pearson Education,Inc,2017.

2. Cecil C. Bozarth and Robert B. Handfield, Introduction to operations and supply chain management, Fifth edition, NY: Pearson, 2019.

3. Iman Rahimi et al., Big data analytics in supply chain management: theory and applications, CRC-Press,2021.

4. Sunil Chopra, Peter Meindl and Dharam Vir Kalra, Supply Chain Management Strategy, Planning and Operation, Sixth Edition, Pearson India, 2016.

5. Dmitry Ivanov, Alexander Tsipoulanidis and Jörn Schönberger, Global Supply Chain and Operations Management: A Decision-Oriented Introduction to the Creation of Value, Second Edition, Springer, 2019

# 24-478-0109: SYSTEM RELIABILITY AND RISK ANALYSIS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the concept of reliability, reliability data and model selection | Understand |
| 2. | Examine classical non-parametric distribution estimation and Bayesian estimation procedure | Understand |
| 3. | Understand system reliability analysis | Understand |
| 4. | Distinguish different methods of evaluation of reliability | Analyze |
| 5. | Understand accelerated life testing problems and its popular reliability models | Understand |
| 6. | Examine AL model for time-dependent stress and PH model | Analyze |
| 7. | Understand basic concepts in risk analysis | Understand |
| 8. | Relate different risk assessment methods | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | 2 | 2 | | 2 |
| CO2 | | | | 2 | |
| CO3 | | 2 | 2 | 2 | 1 |
| CO4 | | 2 | | | |
| CO5 | | 1 | 2 | | |
| CO6 | | | 2 | | |
| CO7 | 2 | 2 | | | 2 |
| CO8 | | 1 | | | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Component Reliability Analysis: Concept of reliability, Reliability data and model selection, Censored data, Classical non-parametric distribution estimation, Bayesian estimation procedure Estimation of the parameter of exponential distribution.

**Module 2:**

System Reliability Analysis: Reliability block diagram method, Failure tree and success tree methods, Event tree method, Master logistic diagram, Failure Mode and Effect Analysis.

**Module 3:**

Accelerated Life (AL) testing: Basic AL notions, Some popular AL reliability models, AL data analysis, AL model for time-dependent stress, Exploratory data analysis for time-dependent stress, PH model data analysis.

**Module 4:**

Risk Analysis: Determination of risk values, Quantitative risk assessment, probabilistic risk assessment, A simple fire protection risk analysis, precursor-based risk analysis.

**Reference Books:**

1. Modarres, M., Kaminskiy,M.P. and Krivtsov,V.(2017). Reliability Engineering and Risk Analysis–A Practical Guide, Third Edition, C RC Press, Boca Raton, FL.

2. Rao S.S.(1992)Reliability-based Design, McGraw-Hill, Inc., USA.

3. Gnedenko,B., Pavlov,I. and Ushakov,I. (1999). Statistical Reliability Engineering, John Wiley & Sons, Inc.,USA.

4. Zio,E. (2007). An introduction to basics of reliability and risk analysis, World Scientific Publishing Co.Pvt Ltd., Singapore

5. Birolini.A.(2007). Reliability engineering–Theory and Practice, Fifth Edition, Springer-Verlag, Berlin.

# 24-478-0201: SIMULATION MODELLING AND ANALYSIS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand how simulation works, application areas and the types of software that are available for simulation modelling | Understand |
| 2. | Develop a conceptual model for a given problem | Evaluate |
| 3. | Apply general modelling principles of simulation | Apply |
| 4. | Create credible models of systems | Create |
| 5. | Evaluate model's models validation using appropriate methods | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | 2 | 2 | | 3 |
| CO2 | 2 | 1 | 2 | | 2 |
| CO3 | 2 | 2 | | | 2 |
| CO4 | | 2 | | 2 | 2 |
| CO5 | | 1 | | 2 | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Simulation: Definition, areas of application, System: discrete and continuous Systems, Model of System, Common types of simulation models and applications, Steps in a simulation study. General principles of discrete event–Simulation, Event scheduling/time advance algorithms, World views, Simulation examples: single channel queues newspaper selling problem, reliability problem, lead-time demand.

**Module 2:**

Random number generation, Properties of random numbers, Techniques of generation of pseudo–random numbers, Test for random numbers, Random variate generation: Inverse transform technique, Convolution method, Acceptance–rejection technique. Queuing Models, Long run measures of performance, Steady state models: M/G/1, M/M/1/N/$\infty$,M/M/c/$\infty$/$\infty$ and M/M/c/K/K.

**Module 3:**

Simulation of manufacturing and material handling systems: Modeling of manufacturing system, Material handling systems, Goals and performance measurement, Modeling of downtimes and failures, Trace driven models; Features of Simulation languages: Promodel –Extend - Auto Mod – Taylor II – Witness, Simul8– AIM – Arena, Basic introduction to agent-based simulation and applications.

**Module 4:**

Input modeling, Verification, Calibration and validation, Face validity, Validation of model assumption, Validating input-output. Analysis of simulation data: Output analysis for terminating simulations, Output analysis for steady state simulations.

**Reference Books:**


1. Jerry Banks et.al.: Discrete–Event System Simulation, Fifth Edition, Prentice Hall ,2009.
2. Law A.M, Simulation Modeling and Analysis, Fifth edition, McGraw Hill New York, 2015.
3. Robinson S: Simulation, The Practice of Model Development and Use, Red Globe Press; Second edition, 2014.
4. Gordon G, System Simulation, Second Edition, Prentice Hall, 1978

# 24-478-0202: MACHINE LEARNING

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understanding of the basic concepts of Machine Learning | Understand |
| 2. | Solve the problems using various machine learning techniques | Apply |
| 3. | Describe the underlying mathematical relationships within across Machine Learning algorithms and the paradigms of supervised and unsupervised learning | Evaluate |
| 4. | Identify machine learning techniques suitable for a given problem | Analysis |
| 5. | Understand neural networks, including MLPs, activation functions and back propagation. Gain basic knowledge of CNNs for image recognition and RNNs | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | 2 | | 3 | 3 |
| CO2 | 3 | | 2 | 3 | 3 |
| CO3 | 2 | | | | 3 |
| CO4 | 3 | | 2 | 3 | 3 |
| CO5 | 3 | | | | 3 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Review of linear algebra: Vectors, Matrices and Eigen values; Probability and Optimization: Random Variables, Discrete and Continuous distributions, Gradient, Hessian, Least Squares, Genetic Algorithms.

**Module 2:**

Machine Learning: Introduction, Basic Definitions, Types of Learning, Handling Multicollinearity, Residual Analysis, Supervised Learning- Linear Regression, Logistic Regression; Lasso and Ridge Regularization; Bias and Variance; Overfitting and Underfitting.

**Module 3:**

Supervised Learning: Classification-, Decision Tree, Native Bayes, K-Nearest Neighborhood, Support Vector Machine; Evaluation Matrices for Classifier models.
Dimensionality Reduction: Recursive Feature Elimination, Principal Component Analysis

**Module 4:**

Reinforcement Learning: Introduction, Elements of Reinforcement Learning, Q Learning, Nondeterministic Rewards and Actions, Temporal Difference Learning Overview of Neural Networks: Multilayer Perceptrons (MLPs) and feed forward networks - Activation functions and their role in neural network computation - Training neural networks: Back propagation algorithm - Basic introduction to Convolutional Neural Networks (CNNs) for image recognition and Recurrent Neural Networks (RNNs).

**Reference Books:**

1. Kroese, D. P., Botev, Z., Taimre, T., Vaisman, R. (2019). Data Science and Machine Learning: Mathematical and Statistical Methods. United States: CRC Press.

2. Watt, J., Borhani, R., Katsaggelos, A. K. (2020). Machine Learning Refined: Foundations, Algorithms, and Applications. United States: Cambridge University Press.

3. Deisenroth, M. P., Faisal, A. A., Ong, C. S. (2020). Mathematics for Machine Learning. United Kingdom: Cambridge University Press.

4. Tom M. Mitchell (1997). Machine Learning, New York, N Y: Mc Grow-Hill.

5. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Switzerland: Springer New York.

6. Langley P. (1995). Elements of Machine Learning, Morgan Kaufmann.

7. James K.C.(2024)Regression Modelling with Classical and Statistical Learning Methods: An Easy Guide for Data Scientists, Business Analysts and Engineers using Python. Amazon Kindle Publishing, .

8. Charu C. Aggarwal (2018). Neural Networks and Deep Learning: A Textbook, Springer

# 24-478-0203: MULTI VARIATE ANALYSIS AND STATISTICAL TECHNIQUES FOR DATA MINING

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand basic concepts of data mining | Understand |
| 2. | Understand different methods of data pre-processing, data cleaning and data reduction | Understand |
| 3. | Apply principal component analysis | Apply |
| 4. | Perform factor analysis and canonical correlation | Evaluate |
| 5. | Examine classification problems | Evaluate |
| 6. | Perform MANOVA | Evaluate |
| 7. | Apply different tests of multivariate populations and associated problems | Apply |
| 8. | Apply cluster analysis and different methods of it | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | | 1 | | 2 |
| CO2 | 2 | | | | 2 |
| CO3 | 2 | 2 | | 2 | 2 |
| CO4 | 2 | 2 | | 2 | 2 |
| CO5 | | | | 2 | 2 |
| CO6 | | 2 | | 1 | 2 |
| CO7 | | 2 | | | |
| CO8 | | 2 | | 2 | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to data mining; data types for Data mining, Data mining functionalities -Concept/class description: characterization and discrimination, Association analysis, Classification and prediction, Clustering analysis, Evolution and deviation analysis, Data Pre-processing, Data cleaning, Data integration and transformation, Data reduction, Discretization and concept hierarchy generation.

**Module 2:**

Dimension Reduction methods: Profile Analysis and the associated tests, Principal Component Analysis-Method of extraction-properties, the associated tests, Factor Analysis-Orthogonal Model-Estimation of factor loadings, Canonical variates and canonical correlation, use, estimation and computation.

**Module 3:**

Classification problems: Discriminant Analysis-Bayes' procedure, Classification into one of the two populations (Normal distribution only), Classification into several populations(Normal distribution only), Fishers linear discriminant function and its associated tests, Cluster Analysis: proximity measures, Hierarchical and non-hierarchical methods.

**Module 4:**

Multivariate General linear models- MANOVA (one way and two way), Wilk's $\Lambda$, Rau's U, Pillai's trace, Hotelling-Lawley trace, Roy's Maximum Root Statistics (Concepts only), Tests-Independence of sets of variables, Equality of dispersion matrices and Sphericity test.

**Reference Books:**

1. JiaweiHan, Micheline Kamber, JianPei (2012): Data Mining: Concepts and Techniques, 3rd Edition, Morgan Kaufmann Publishers, Elsevier Inc., USA.

2. Johnson R.A. and Wichern D.W. (2008) Applied Multivariate Statistical Analysis, 6th Edition, Pearson Education.

3. Rencher A.C.(2012) Methods of Multivariate Analysis. 3rd Edition, John Wiley.

4. Johnson D.E.(1998). Applied Multivariate Methods for Data Analysts, Duxbury Press, USA-An International Thomson Publishing Company.

5. Morrison,F (2003). Multivariate Statistical Methods, Brooks/Cole, 4thRevised Edition, McGrawHill Book Company.

6. Srivastava M.S. and KhatriC.G.(2002): Methods of Multivariate Statistics, John Wiley & Sons, NewYork.

7. Anderson T.W.(2010). An Introduction to Multivariate Statistical Analysis, 3rd Edition, JohnWiley.

8. Seber G.F.(2004) Multivariate Observations, John Wiley.

# 24-478-0204: R/R-STUDIO PROGRAMMING – PRACTICAL II

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understanding of basic commands in R programming | Understand |
| 2. | Practice different problems in regression analysis using R programme | Apply |
| 3. | Explain different optimization problems using R programme | Evaluate |
| 4. | Apply the R programming from a statistical perspective | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | 2 | 2 | | 2 |
| CO2 | 2 | | 3 | 3 | 2 |
| CO3 | 2 | | 2 | 3 | 2 |
| CO4 | 2 | 3 | | | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to R: basic commands, graphics, indexing data, loading data; Regression: linear regression, test of significance, residual analysis, polynomial regression, qualitative predictor, logistic-regression

**Module 2:**

Resampling methods: cross validation, bootstrap; Subset selection: best subset selection, forward and backward stepwise selection, choosing among models using the validation; Markov chain Monte Carlo.

**Module 3:**

Optimization in R: Common R Packages for Linear, Quadratic and Non-linear optimization, Built-in Optimization functions, Linear Programming in R-lpsolve

**Module 4:**

Quadratic Programming: quadprog, Non-Linear Optimization:One-dimensiona1:Golden Section Search; Multidimensional: Gradient-based, Hessian based, Non-gradient based

**Reference Books:**

1. G. James, D. Witten,T. Hastie and R.Tibshirani (2013). An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics), Springer.
2. W John Braun, Duncan J (2008). A First Course in Statistical Programming with R, Murdoch, Cambridge University, Press.

# 24-478-0205: OPTIMIZATION TECHNIQUES

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the basic concepts of linear programming | Understand |
| 2. | Distinguish different methods of solving nonlinear programming problem | Analyze |
| 3. | Understand genetic algorithms and associated properties | Understand |
| 4. | Distinguish different types of advanced genetic algorithms and applications | Analyze |
| 5. | Understand differential evolution and its principle | Understand |
| 6. | Distinguish different algorithms of differential evolution | Analyze |
| 7. | Understand particle swarm optimization methods | Understand |
| 8. | Understand protein folding problems and protein structure analysis | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | 3 | 3 | | 3 |
| CO2 | 2 | 2 | 2 | 2 | |
| CO3 | 1 | 1 | | | |
| CO4 | 1 | 1 | | | |
| CO5 | | 1 | | | |
| CO6 | | 1 | | | |
| CO7 | 1 | 1 | | | |
| CO8 | | 1 | | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to optimization: formulation of optimization problems-Review of classical methods-Linear programming-Nonlinear programming-Constraint optimality criteria-constrained optimization-Population based optimization techniques.

**Module 2:**

Genetic Algorithm-Introduction: Working principle-Representation-selection-fitness assignment reproduction-cross over-mutation-constraint handling-advanced genetic algorithms-Applications-Artificial Immune Algorithm-Introduction-Clonal selection algorithm- Negative selection algorithm-Immune network algorithms-Dendritic cell algorithms

**Module 3:**

Differential Evolution: Introduction-Working principles-parameter selection-advanced algorithms in Differential evolution–Biogeography-Based Optimization-Introduction-Working Principles-Algorithmic variations

**Module 4:**

Particle Swarm Optimization: Introduction-Working principles-Parameter selection- Neighborhoods and Topologies-Convergence-Artificial Bee Colony Algorithm-Introduction-Working principles-Applications Cuckoo search based algorithm-Introduction- Working principles- Random walks and the step size-Modified cuckoo search, Hybrid Algorithms: Concepts- divide and conquer- decrease and conquer-HPABC-HBABC-HDABCHGABC-Shuffled Frog Leaping Algorithm—Working principles-Parameters-Grenade Explosion Algorithm-Working principle-Applications.

**Reference Books:**

1. Rao S.S. (2019).Engineering optimization: theory and practice. John Wiley & Sons.
2. Venkata Rao R.(2016). Teaching Learning Based Optimization Algorithm: And Its Engineering Applications, 1e, Springer.
3. Simon D. (2013). Evolutionary optimization algorithms. John Wiley & Sons.
4. Yang X.S.(2010). Engineering optimization: An Introduction with Metaheuristic Applications. John Wiley & Sons.

# 24-478-0206: DESIGN OF EXPERIMENTS (INTEGRATED WITH R)

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Illustrate the statistical tool-Analysis of Variance | Apply |
| 2. | Describe the different experimental designs | Understand |
| 3. | Describe concepts of Taguchi experiments and orthogonal array | Understand |
| 4. | Apply design of experiments to product design, Taguchi concept of quality, Reliability enhancement techniques and accelerated life testing | Apply |
| 5. | Be able to execute using R | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 1 | 3 | | 1 | |
| CO2 | 2 | 3 | 2 | | 2 |
| CO3 | 1 | 2 | 2 | | 2 |
| CO4 | 1 | 2 | 2 | | 2 |
| CO5 | 1 | | | 2 | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Analysis of variance: General Linear Model, One-way analysis of variance, Two way analysis of variance, Interaction, Confidence intervals in one-way and two-way analysis of variance, Multiple treatment comparisons, Random ordering of the experimental sequence, CRD,RBD.

**Module 2:**

Factorial experiments: Simple and main effects, Factorial applied to randomized block designs, Taguchi experiments, Generalization of orthogonal arrays, Dealing with interactions, Confounding, Other designs. Response surface Methodology, First and second order models.

**Module 3:**

Application of design of experiments, Product design, Managing the uncontrollable factors, Type of uncontrollable factors, Use of outer arrays for concurrent engineering, Application of design of experiments to reliability assurance.

**Module 4:**

Accelerated Life Testing, Environmental and operating stress, Interpreting data from accelerated tests, Developing an integrated reliability test program, Reliability improvements with design of experiments.

**Reference Books:**


1. D.C. Montgomery (1997): Design and Analysis of Experiments, John Wiley and Sons, New York.
2. Lawson John (2015): Design and Analysis of Experiments with R; CRC Press.
3. Bloyd W. Condra (1993): Reliability Improvements with Design of Experiments, Marcel Dekker.
4. R. M Bethea and R. Rhireheart (1991): Applied Engineering Statistics, Marcel Dekker.
5. Alain Villemeur (1992): Reliability Availability, Maintainability and Safety Assessment Vol.2, John Wiley ands ons.
6. E. Daniel (1976): Application of Statistics to Industrial Experimentation, John Wiley and Sons.
7. T.J Lorenzen and V.L. Anderson (1993): Design of Experiment, Marcel Dekker.

# 24-478-0207: ARTIFICIAL INTELLIGENCE AND DEEP LEARNING

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Discuss the history, current applications, future challenges and the controversies in artificial intelligence | Understand |
| 2. | Apply principle of AI in the design of an agent and model its actions | Apply |
| 3. | Design a heuristic algorithm for search problems | Apply |
| 4. | Analyze and represent the fact using logic for a given scenario | Apply |
| 5. | Represent uncertainty using probabilistic models | Evaluate |
| 6. | Describe the basics of learning and deep learning | Understand |
| 7. | Apply principles of Convolution neural network to practical problems | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | 1 | | | 1 |
| CO2 | 3 | 3 | | 2 | |
| CO3 | 3 | 2 | | 2 | 2 |
| CO4 | | | 2 | 1 | |
| CO5 | | 1 | | | 1 |
| CO6 | 3 | 3 | | 2 | 3 |
| CO7 | 3 | 3 | 3 | 3 | 3 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction: AI:, Brief history-Agents and rationality, task environments, agent architecture Types- Search and Knowledge representation - Search spaces, Uninformed and informed search.

**Module 2:**

Techniques: Hill climbing, simulated annealing, genetic algorithms-Logic based representations (PL, FoL) and inference, Prolog - Rule based representations, forward and backward chaining, matching algorithms - Probabilistic reasoning and uncertainty - Bayes networks and reasoning with them.

**Module 3:**

Learning: Uncertainty and methods to handle it - Forms of learning - Statistical methods: naive-Bayes, nearest neighbors, kernel, neural network models, noise and over fitting, Decision trees, inductive learning.

**Module 4:**

Introduction to Convolution Neural Networks and Recurrent Neural Networks – Deep Learning, Autoencoders and unsupervised learning - Stacked auto-encoders and semi-supervised learning -Regularization – Dropout and Batch normalization.

**Reference Books:**


1. Russel S. and Norvig P. (2015), Artificial Intelligence: A Modern Approach, 3rd Edition, Pearson Education Limited, 2016
2. Goodfellow, I., Bengio,Y. and Courville, A., Deep Learning, The MIT Press, 2016.
3. DavidPoole, Alan Mackworth, Artificial Intelligence: Foundations of Computational Agents, Cambridge University Press, 2010.
4. Keith Frankish, William M. Ramsey (eds) The Cambridge Handbook of Artificial Intelligence, Cambridge University Press, 2014.

# 24-478-0208: NATURAL LANGUAGE PROCESSING

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Describe the basics of language processing technologies for Processing the text | Understand |
| 2. | Acquire knowledge on text data analytics using language models | Evaluate |
| 3. | Process the text data at syntactic and semantic level | Evaluate |
| 4. | Extract key information from text data | Apply |
| 5. | Analyze the text content to provide predictions related to a specific domain using language models | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | 1 | | 2 | 3 |
| CO2 | 1 | | | 3 | 3 |
| CO3 | 2 | | 1 | 3 | 3 |
| CO4 | 3 | | | 1 | 3 |
| CO5 | 2 | | 3 | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to NLP, Knowledge Acquisition, Regular expression (RE) and Text Processing, Word Tokenization, Word Normalization and Word Stemming, Sentence Segmentation, Edit Distance, Word Alignment Problem and Statistical Machine Translation (MT),Word Alignment Problem, Parallel Corpora, Decoding, Evaluation, Statistical MT, Modern MT Systems

**Module 2:**

Language Modeling, Introduction to N-grams, Estimating N-gram Probabilities, Evaluation and Perplexity, Generalization and Zeros, Interpolation, Good-Turing Smoothing, Kneser- Ney Smoothing, Spelling Correction, Noisy-Channel Model for Spelling

**Module 3:**

Text Classification and Sentiment Analysis, Naïve Bayes Classifier, Precision, Recall and the F measure, Text Classification, Evaluation, Sentiment Analysis - Baseline Algorithm ,Learning Sentiment Lexicons and Discriminative Classifier - Maximum Entropy Classifier, Generative vs. Discriminative Model Making features from text, Feature-based Linear Classifier, Problem of Over counting evidence, Named Entity Recognition (NER) and Maximum Entropy Sequence Model.

**Module 4:**

Introduction to Information Extraction, NER and Evaluation of NER, Sequence Models for NER,Maximum Entropy Sequence Model, Relation extraction by using patterns, Supervised, Semi-supervised and Unsupervised Relation Extraction, Advanced Maximum Entropy Models, Parts of Speech (POS) Tagging , Sequence Models for POS Tagging, Parsing, Syntactic Parsing - Constituency vs Dependency, Context Free Grammar (CFG) and PCFG, Grammar Transforms, CKY Parsing, Lexicalized Parsing, Lexicalization and PCFGs, Charniak' Model, Unlexicalized PCFGs, Latent Variable PCFGs, Context Sensitive Grammar(CSG).

**Reference Books:**

1. Jurafsky D. and Martin J.H. Speech and language processing.Vol.3.Prentice Hall, 2014.

2. JanŽižka, František Dařena and Arnošt Svoboda, Text mining with machine learning : principles and techniques, CRC Press, 2019.

3. Bird S., Klein E. & Loper, Natural language processing with Python: analyzing text witht the natural language toolkit. O'Reilly Media, Inc,2009..

4. Manning C.D., Manning, C.D. & Schütze, H Foundations of statistical natural language processing. MIT press, 1999.

# 24-478-0209: FINANCIAL RISK ANALYTICS AND MANAGEMENT

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Describe the financial risk management process and identify key factors that affect interest rates, exchange rates and commodity prices | Evaluate |
| 2. | Understand and evaluate ways to manage the different risks pertaining to stock market and its instruments | Apply |
| 3. | Identify the various risks faced by an organization such as credit risk, operational risk, and systemic risk | Evaluate |
| 4. | Develop an organizational profile to support risk management policy and evaluate opportunities to develop or refine a risk management policy | Evaluate |
| 5. | Identify and solve legal issues that impact financial and other risk affecting business | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 1 | | 2 | | 2 |
| CO2 | 1 | 2 | 2 | | 2 |
| CO3 | 1 | | 2 | | 2 |
| CO4 | 1 | | 2 | | 2 |
| CO5 | 1 | | 2 | | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to Risk-Understanding Risk-Nature of Risk, Source of Risk, Need for risk management, Benefits of Risk Management, Risk Management approaches. Risk Classification- credit risk, market risk, operational risk and other risk

**Module 2:**

Risk Measurements-Measurement of Risk–credit risk measurement, market risk measurement, interest rate risk measurement, Asset liability management, measurement of operational risk

**Module 3:**

Risk Management- Risk management- Managing credit risk, managing operational risk, managing market risk, insurance

**Module 4:**

Risk in Instruments -Tools for risk management – Derivatives, combinations of derivative instruments - Neutral and volatile strategies - credit derivatives - credit ratings- Regulation and Other Issues: Other issues in risk management – Regulatory framework, Basel committee, legal issues, accounting issues, tax issues, MIS and reporting, integrated risk management

**Reference Books:**

1. Dun, Bradstreet, Financial Risk Management, TMH, 2006.
2. Jimmy Skoglund and Wei Chen, Financial risk management: applications in market, credit, asset and liability management and firmwide risk, Wiley finance series, 2015

3. McNeil A.J., Frey, R. and Embrechts, P. Quantitative Risk Management, Princeton University Press, Oxford, 2005.

4. Richard Apostolik and Christopher Donohue, Foundations of Financial Risk: An Over view of Financial Risk and Risk-Based Regulation, Wiley, 2015

5. JohnCHull, Risk management and Financial Institutions, Pearson, 2015.

6. Aswath Damodharan, Strategic Risk Taking, Pearson, 2008.

# 24-478-0210: MARKETING AND HR ANALYTICS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Interpret the benefits and objectives of marketing analytics | Evaluate |
| 2. | Apply metrics-driven techniques to improve marketing decisions | Apply |
| 3. | Design appropriate hands-on computer models and metrics | Apply |
| 4. | Identify necessary skills to carry out the personnel roles in the domain of HR | Evaluate |
| 5. | Apply HR analytics to improve organizational performance by providing better insights on human resources data | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 1 | 1 | 1 | | 1 |
| CO2 | 1 | | 1 | | 1 |
| CO3 | 1 | | 3 | | 1 |
| CO4 | 1 | 1 | | | |
| CO5 | 1 | | 2 | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Introduction to Marketing Analytics – Models & Metrics – Market Insight – Terminologies – Data Sources - Market sizing - PESTLE analysis – Introduction to forecasting tools.

**Module 2:**

Market Segmentation - Segment Targeting and Positioning- Competitor - identification and information gathering – Competitive analysis – Competitive -actions - Cluster Analysis –Classification Trees-Conjoint Analysis, Products, Attributes, and Levels, Full Profile Conjoint Analysis- Discrete Choice Analysis- Random Utility Theory.

**Module 3:**

Introduction to HR Analytics: Overview of HR Process, HR as an expense, the analytics and prediction Strategic Human capital measures, business analysis and rational action – Benefits of Analytics in Improving HR Process, Intersection of people and profits -Technology Used, SWOT Analysis of HR analytics.

**Module 4:**

Employee Engagement Measurement Process: Attracting, motivating and retaining people Organization Gap and Alignment Analytics - Recruitment Analytics and On Boarding Analytics - Staffing Analytics - Performance & Skill Gap Analytics - Compensation & Benefit Analytics Training & Learning Analytics - Promotion and Succession Planning Analytics Compliance Analytics Attrition & Retention Analytics - Identification of Key Business Objectives Conducting HR Practice

**Reference Books:**


1. Wayne L.Winston, Marketing Analytics: Data-Driven Techniques with Microsoft® Excel, Published by John Wiley & Sons, 2014
2. Jac FITZ-ENZ, The New HR Analytics: Predicting the Economic Value of Your Company's Human Capital Investments, American Management Association, 2010.
3. JohnW.Boudreau, Beyond HR: The New Science of Human Capital, Harvard Business School Press, 2007.
4. Stephan Sorger, Marketing Analytics: Strategic Models and Metrics, 1st Edition, Create Space Independent Publishing Platform, ISBN: 1481900307, 2013.

# 24-478-0211: BIOINFORMATICS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand features of molecular biology | Understand |
| 2. | Distinguish the genomic-genetic synthesis, translation-transcription-protein synthesis | Analyze |
| 3. | Understand pair wise sequence alignment, PAM and BLOSSUM matrices | Understand |
| 4. | Distinguish different dynamic programming methods | Analyze |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 1 | 2 | 3 | | 1 |
| CO2 | | 2 | 2 | | 1 |
| CO3 | | 2 | | 2 | 1 |
| CO4 | | 2 | | | 1 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Molecular Biology and Bioinformatics: Introduction to molecular biology- Nucleic acids-DNA-RNA Proteins-Gene-Genome-Genetic synthesis -Translation-Transcription-Protein synthesis-Chromosomes-Maps and sequences-Human genome project.

**Module 2:**

Sequence alignment and database search: Pair-wise sequence alignment- Substitution matrices -PAM and BLOSSUM matrices, Dot plots - Local and global alignment theory -Dynamic programming methods -FASTA and BLAST algorithms - database search using BLAST and FASTA - Similarity & distance -Similarity scores - Weight matrices - Heuristic method - Hidden Markov Models and their application in sequence analysis.

**Module 3:**

Phylogenetic trees: Introduction -Dendrogram construction/Molecular Phylogenetics/Tree definitions/Optimality criteria/Distance matrix methods and maximum parsimony/Multiple/ sequence alignments- tree alignments, star alignments, pattern in pair wise alignment Genetic algorithm.

**Module 4:**

DNA Micro-arrays and Gene Expression- Gene profiling- DNA Microarray technology-Gene regulatory network-Heuristic Algorithms for GRN- S-system model – Computational methods for pathways and system biology-metabolic pathways- genetic pathways- signaling pathways, Molecular Structure Prediction- RNA secondary structure prediction-Protein Folding problems-Protein threading- Protein structure analysis.

**Reference Books:**

1. Ghosh,Z.,&Mallick,B.(2008). Bioinformatics: Principles and Applications.Rastogi S. C., Rastogi, P., & Mendiratta, N. (2008). Bioinformatics Methods And Applications: Genomics Proteomics And Drug Discovery 3RdEd. PHI Learning Pvt. Ltd..

2. Jones N.C., Pevzner,P.A. & Pevzner,P. (2004). An introduction to Bioinformatics Algorithms. MIT press.

# 24-478-0212: BIG DATA TECHNOLOGY

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Process data in Big Data platform and explore the big data analytics Techniques for business applications | Analyze |
| 2. | Analyze Map Reduce technologies in big data analytics | Analyze |
| 3. | Learn to build and maintain reliable, scalable, distributed systems with Apache Hadoop | Evaluate |
| 4. | Design efficient algorithms for stream data mining on big data platform | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | | | 1 | 3 |
| CO2 | 2 | | | 3 | 3 |
| CO3 | 2 | | 1 | 3 | 3 |
| CO4 | 3 | | | 3 | 3 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

NoSQL Database: NoSQL Databases – Schema less Models, Increasing Flexibility for Data Manipulation-Key Value Stores, Document Stores, Tabular Stores, Object Data Stores –Graph Databases, Big Data: Evolution of Big data, Best Practices for Big data Analytics - Big data characteristics - Big Data Use Cases, Characteristics of Big Data Applications, Big data for twitter, Big data for E-Commerce blogs.

**Module 2:**

Big Data Modelling, Map Reduce algorithm. Hadoop Introduction: Apache Hadoop & Hadoop Ecosystem – Moving Data in and out of Hadoop – Understanding inputs and outputs of MapReduce-Data Serialization, Hadoop Architecture.

**Module 3:**

HDFS performance and tuning, Pig: Introduction to PIG, Execution Modes of Pig, Grunt, Pig Latin, User Defined Functions, Data Processing operators. Hive : Hive Shell, Hive Services, HiveQL, Tables, Querying Data and User Defined Functions. Hbase: HBasics, Concepts, Clients, Example, Spark

**Module 4:**

Mining Data Streams: Introduction to Streams Concepts, Stream Data Model and Architecture - Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream –Real time Analytics Platform (RTAP) applications, Case Studies, Real Time Sentiment Analysis-Stock Market Predictions.

**Reference Books:**


1. Jure Leskovec, Anand Rajaramanand Jeffrey David Ullman, Mining of Massive Datasets, Cambridge University Press, 2014.
2. Arshdeep Bahga & VijayMadisetti,BigDataAnalytics:AHands-OnApproach,2019
3. Tom White, Hadoop: The Definitive Guide, 4th edition O'Reily Publications, 2015
4. Jules S.Damji, Brooke Wenig, TathagataDas,andDennyLee,LearningSpark:Lightning-FastDataAnalytics, O-ReillyPublications,2020

5. David Loshin, Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL and Graph, 2013.

6. Bart Baesens, Analytics in a Big Data World: The Essential Guide to Data Science and its Applications", Wiley Publishers, 2015.

7. Dietmar Jannach, Markus Zanker, Alexander Felfernig and Gerhard Friedrich, Recommender Systems: An Introduction, Cambridge University Press, 2010.

# 20-478-0301: BUSINESS ANALYTICS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand and explore problems in business | Evaluate |
| 2. | Summarize the evolution of business analytics and explain the concepts of business intelligence, operations research and management science, and decision support systems | Evaluate |
| 3. | Explain the purpose of regression analysis and provide examples in business | Understand |
| 4. | Apply a systematic approach to build good regression models | Apply |
| 5. | Apply the appropriate forecasting technique in a given business situation | Apply |
| 6. | Apply association rule mining, k- Nearest Neighbors, discriminant analysis, and logistic regression in business applications such as Market basket analysis. | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 3 | | 3 | | 2 |
| CO2 | | | 2 | | 2 |
| CO3 | | | | 2 | 2 |
| CO4 | 3 | 3 | | 2 | 2 |
| CO5 | 3 | 3 | | 2 | 2 |
| CO6 | 3 | | | 2 | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Business Analytics: Evolution of Business Analytics, Impacts and Challenges - Big Data - Metrics and Data Classification - Data Reliability and Validity - Models in Business Analytics - Decision Models - Model, Assumptions - Uncertainty and Risk - Prescriptive Decision Models - Problem Solving with Analytics - Recognizing a Problem - Defining the Problem - Structuring the Problem - Analyzing the Problem - Interpreting Results and implementation.

**Module 2:**

Modeling Relationships and Trends in Data- Regression applications in industry and business- Least-Squares Regression - Regression Analysis of Variance-Testing Hypotheses for Regression Coefficients –Confidence Intervals for Regression Coefficients- Residual Analysis and Regression Assumptions- Checking Assumptions- Multiple Linear Regression, Polynomial regression-diagnostics, transformations and model improvements - Regression with categorical independent variables.

**Module 3:**

Forecasting Techniques - Qualitative and Judgmental Forecasting - Forecasting Models for Stationary Time Series - Moving Average Models-Error Metrics and Forecast Accuracy - Exponential Smoothing Models Regression - Based Forecasting for Time Series with a Linear Trend -Forecasting Time Series with Seasonality - Regression - Based Seasonal Forecasting Models - Holt-Winters Forecasting for Seasonal Time Series - Holt-Winters Models for Forecasting Time Series with Seasonality and Trend - ARMA and ARIMA Modeling and Forecasting - selecting appropriate Time - Series based Forecasting Models - Regression Forecasting with Causal Variables.

**Module 4:**

Introduction to Data Mining - The Scope of Data Mining -Data Exploration and Reduction- Sampling - Data Visualization- Dirty Data - Cluster Analysis -Classification - Measuring Classification Performance-Using Training and Validation Data-Classifying New Data- k-Nearest Neighbors (k-NN) - Discriminant Analysis -Logistic Regression-Association Rule Mining-Cause-and-Effect Modeling

**Reference Books:**

1. Galit Shmueli et al, Data Mining For Business Analytics: Concepts, Techniques, and Applications in R John Wiley & Son, 2018
2. J.R.Evans, Business Analytics, Pearson Education; Second edition, 2017
3. Jeffrey D Camm, Essentials of Business Analytics; South Western, 2015
4. S.Christian Albright and Wayne L. Winston, Business Analytics: Data Analysis & Decision Making, Cengage Learning; Sixth Edition, 2017.
5. K.C James, Regression Modelling with Classical and Statistical Learning Methods: An Easy Guide for Data Scientists, Business Analysts and Engineers using Python. Amazon Kindle Publishing, 2024.
6. Hanke, J. E.,& Wichern, D. W. Business Forecasting: Pearson New International Edition. Pearson Higher Ed, 2013.

# 24-478-0303: STATISTICAL FORECASTING METHODS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Choose a suitable model for the given time series | Evaluate |
| 2. | Define forecasts under various Optimization criteria such as MMSE ,MAP, MAPE etc. | Remember |
| 3. | Prediction by suitable smoothing methods | Evaluate |
| 4. | Compute forecasts using regression and time series models | Apply |
| 5. | Evaluate the forecasts using suitable measures | Evaluate |
| 6. | Prediction using logistic regression, intervention models and neural networks | Evaluate |
| 7. | Interpret the model performance based on residuals | Evaluate |
| 8. | Choose forecast methods for a given data | Evaluate |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | 2 | | | 3 |
| CO2 | 2 | 2 | | | 2 |
| CO3 | | 2 | | 2 | 2 |
| CO4 | 1 | 2 | | 2 | 2 |
| CO5 | 2 | 2 | | 2 | |
| CO6 | 2 | 2 | | 2 | 2 |
| CO7 | | 2 | | | 2 |
| CO8 | 2 | 1 | | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Need for forecasting, Basic forecasting tools, Time series and cross-sectional data, Graphical and numerical summaries, Measuring forecast accuracy, Transformations and adjustments. Time series decomposition: principle of decomposition, moving averages. Exponential smoothing: single, double, Holt's and Holt – Winters' smoothing methods.

**Module 2:**

Simple regression: model and properties, Inference and forecasting with simple regression. Multiple linear regression: Method of analysis and examples. Variable selection methods, Multicollinearity, multiple linear regression and forecasting, Regression with correlated errors, Durbin–Watson test.

**Module 3:**

Box-Jenkins Methodology for forecasting: Tests for independence and stationarity, methods of removing non-stationary, ARMA and ARIMA models for time series: Identification, estimation and diagnosis methods. Forecasting with ARIMA models.

**Module 4:**

Modelling and forecasting of Regression models with ARIMA errors: Dynamic regression models, Intervention analysis, non-linear models: logistic regression, Neural network forecasting.

**Reference Books:**

1. S.Makridakis, S.C.Wheelwright and R.J.Hyndman (2005): Forecasting Methods and Applications. 3rd Edition. John Wiley and Sons, NewYork.

2. B.Abraham and J.Ledolter (2006): Introduction to Regression Modeling Thomson, Canada

3. B.Abraham and J. Ledolter (1983).Statistical Methods for Forecasing. Wiley, NewYork.

4. D.C.Montgomery, E.A.Peckand G.G.Vining (2003): Introduction to Linear Regression Analysis.Wiley

# 24-478-0304: QUALITY MANAGEMENT AND SIX SIGMA

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the basic ideas of quality and quality management and the necessity for quality improvement for organizational excellence | Understand |
| 2. | Understand various quality costs and see how quality Improvement efforts could reduce costs in industries | Analyze |
| 3. | Evaluate current level of quality in industries | Evaluate |
| 4. | Know Project Management Using the DMAIC and DMADV Models | Apply |
| 5. | Understand and apply various analytical tools of quality, Six Sigma and quality management systems | Apply |
| 6. | Understand and apply appropriate statistical tools for quality control | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 2 | 2 | | | 1 |
| CO2 | | 1 | 3 | | 1 |
| CO3 | | 1 | | | 1 |
| CO4 | 2 | 2 | 2 | | 1 |
| CO5 | 2 | 2 | 2 | | 1 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Basics of Quality and Total Quality Management – Scope and Objectives - Models and Frameworks for TQM-Analysis of Quality Costs–Quality Improvement and Cost Reduction. Strategic Quality Management – Quality Policy – Training for Quality –Quality Teams – Developing Quality Culture – TQM & JIT- Quality System: ISO 9000/Q S9000 And ISO 14000.

**Module 2:**

Introduction to Six Sigma-Six Sigma as a Statistical Measure-The DMAIC Process–Basics of Lean concepts, Strategic Planning and Deployment, Project Selection, Lean Six Sigma Team and Project Management - Design for Six Sigma (DFSS) - Phases of DFSS - Process of DFSS.

**Module 3:**

Analytics Tools: Basic concepts of SWOT, Affinity Diagram, SIPOC, VOC, CTQ, Pareto Diagram, Prioritization Matrix, Monte Carlo Analysis, Gauge R&R Methods, Cause-and-Effect Analysis, Quality function deployment (QFD), Bench Marking, FMEA, DOE, VSM, Regression, RSM.

**Module 4:**

Quality Control: Basic concepts of Acceptance sampling, Quality using SPC Charting-Common Causes and Assignable Causes- Rational Subgroups, Control Limits, Control charts for Attributes and Variables - Cumulative Sum Chart - EWMA Chart - Process Capability Studies.

**Reference Books:**

1. Erick C. Jones, Quality Management for Organizations Using Lean Six Sigma Techniques,CRCPress,2014
2. Theodore T. Allen, Introduction to Engineering Statistics and Lean Six Sigma Statistical Quality Control and Design of Experiments and Systems, Third Edition, Springer-Verlag London Ltd.,2019

3. K. Muralidharan, Six Sigma for Organizational Excellence A Statistical Approach, Springer India,2015

4. Amitava Mitra, Fundamentals of Quality Control and Improvement Fourth Edition, John Wiley& Sons Inc.,2016

5. Juran,J. M and Gryna,F.M: Quality Planning and Analysis for Enterprise Quality, 5th Edition, McGraw Hill, 2007.

# 24-478-0305: APPLIED LONGITUDINAL ANALYSIS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand features of longitudinal data | Understand |
| 2. | Understand the descriptive methods of analysis and modelling mean and covariance | Apply |
| 3. | Understand linear mixed effects models | Understand |
| 4. | Distinguish fixed effects versus random effects | Analyze |
| 5. | Apply residual analysis and diagnostics | Apply |
| 6. | Distinguish GLM and generalized linear mixed effects models | Analyze |
| 7. | Examine missing data and dropout | Analyze |
| 8. | Understand multilevel data | Understand |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | | | 2 | 2 | 2 |
| CO2 | | | | | 2 |
| CO3 | | 1 | | | 2 |
| CO4 | | 1 | 1 | 1 | 2 |
| CO5 | 3 | | | 2 | 2 |
| CO6 | 3 | | | 3 | 2 |
| CO7 | 3 | | | 3 | 2 |
| CO8 | 3 | | | 3 | 2 |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Longitudinal Data-Introduction, Defining features of Longitudinal data, Linear models for Longitudinal Data-Simple descriptive methods of analysis, modelling the mean, modelling the covariance, estimation, missing data, modelling the.

**Module 2:**

Linear mixed effects models-random effects covariance structure, two-stage random effects formulation, fixed effects versus random effects models, residual analyses and diagnostics-residuals, transformed residuals, aggregating residuals, semi-variogram.

**Module 3:**

Salient features of GLM, ordinal regression models, overdispersion, marginal models, generalized estimating equations, generalized linear mixed effects models, incorporating random effects in GLM, interpretation of regression parameters.

**Module 4:**

Missing data and dropout- multiple imputation, inverse probability weighted methods, repeated measures and related designs, multiple source data, Multilevel data, multilevel linear models, multilevel GLM.

**Reference Books:**

1. Fitzmaurice, G.M., Laird, N.M. and Ware, J.H. (2011). Applied Longitudinal Analysis, 2nd Edition, John Wiley and Sons, New Jersey.

2. Diggle, P.J., Heagerty, P., Yiang, K-Y., Zeger, S.L. (2002). Analysis of longitudinal data. Oxford University Press.

3. Fress, E.W. (2004). Longitudinal and Panel Data-Analysis and Applications in the Social Sciences, Cambridge University Press, UK.

4. Hedeker.D. and Gibbons, R.D. (2006). Longitudinal Data Analysis, John Wiley and Sons, New Jersey.

# 24-478-0306: LIFETIME STUDIES IN DATA SCIENCE (INTEGRATED WITH R)

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand the basic concepts and ideas of survival analysis | Understand |
| 2. | Examine the properties and methods for standard survival time distributions | Analysis |
| 3. | Estimate survival functions using parametric and non-parametric methods | Evaluate |
| 4. | Apply and interpret semi-parametric and parametric regression models for survival data | Apply |
| 5. | Execute the methodologies using R | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 1 | 3 | 2 | | 1 |
| CO2 | 2 | 2 | 2 | 1 | |
| CO3 | 2 | 3 | 1 | 1 | |
| CO4 | 2 | 3 | 1 | 2 | |
| CO5 | 2 | 3 | | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Basic Quantities and Models - Survival function, Hazard function, Mean residual life function, Common Parametric Models for Survival Data; Censoring and Truncation - Right Censoring, Left or Interval Censoring, Truncation, Likelihood Construction for Censored and Truncated Data, Ageing

**Module 2:**

Some Parametric Families of Probability Distributions, Parametric Analysis of Survival Data, Non-parametric Estimation of Basic Quantities for Right Censored and Left Censored Data - Estimators of the Survival and Cumulative Hazard Functions for Right Censored Data, Pointwise Confidence Intervals for the Survival Function (without derivation), Estimators of the Survival Function for Left-Truncated and Right-Truncated Data; Estimation of the Survival Function for Left censored data.

**Module 3:**

Test for exponentiality, Two sample parametric problem.

**Module 4:**

Proportional Hazards Model: A Method of Regression- Coding Covariates, Partial Likelihoods for Distinct-Event Time Data, Partial Likelihoods when Ties are present, Model Building using the Proportional Hazards Model, Estimation for the Survival Function; Test for Constant of Proportionality in PH Model Analysis of Competing Risks, Repairable Systems

**Reference Books:**

1. J V Deshpande and Sudha Purohit (2005), Lifetime data: statistical models and Methods, World Scientific, Chapters 1-10.

   Reference Books:

2. Klein J.P. and Moeschberger M.L. (2003) Survival Analysis - Techniques for censored and truncated data, Second Edition, Springer-Verlag, New York.

3. Lawless J.F (2003) Statistical Models and Methods for Lifetime Data, Second Editon, John Wiley & Sons, Relevant Sections of the Chapters 9.

4. Kalbfleisch J.D and Prentice, R.L. (2002) The Statistical Analysis of Failure Time Data, Second Edition, John Wiley & Sons Inc.

5. Hosmer Jr. D.W and Lemeshow S (1999) Applied Survival Analysis - Regression Modelling of Time to event Data, John Wiley & Sons. Inc.

6. Nelson. W (2003) Applied Life Data Analysis.

7. Miller, R.G. (1981) Survival Analysis, John Wiley.

# 24-478-0307: BAYESIAN COMPUTING AND ANALYSIS

**After completion of this course the student should be able to:**

| | Course Outcome (CO) | Cognitive Level |
|---|---|---|
| 1. | Understand basic concepts in Bayesian statistics including Bayes rule, prior, posterior distributions | Understand |
| 2. | Describe different Bayesian models and its uses | Evaluate |
| 3. | Explain different computation and simulation techniques in Bayesian statistics | Evaluate |
| 4. | Apply Bayesian inference approaches to scientific and real-world problems | Apply |

| | PSO1 | PSO2 | PSO3 | PSO4 | PSO5 |
|---|---|---|---|---|---|
| CO1 | 1 | 1 | | | 1 |
| CO2 | 1 | 1 | 1 | | 1 |
| CO3 | 1 | 1 | 2 | | 1 |
| CO4 | | | 2 | | |

1-Slightly/Low, 2-Moderate/Medium, 3-Substantial/High

**Module 1:**

Subjective notions of probability, Bayes Theorem and applications. Likelihood, Prior and posterior

**Module 2:**

Loss function, Bayes Rule, priors and other priors, Sensitivity Analysis, Posterior Convergence.

**Module 3:**

Bayesian Models: Poisson, Binomial, Univariate Gaussian, Multivariate Gaussian. Hierarchical Bayesian Model, Classification with Bayesian Logistic Regression, Discriminant Analysis.

**Module 4:**

Bayesian Computation and simulation techniques: (integrated with R Jags and Stan. Estimation of Posterior Mode with Optimization, Estimation of Posterior Mean and other summary with Monte Carlo Simulation, Accept-Rejection Sampling, Importance Sampling, Markov Chain and Monte Carlo, Metropolis-Hastings, Hamiltonial Monte Carlo

**Reference Books:**

1. James O. Berger (1985): Statistical Decision Theory and Bayesian Analysis, 2'nd Edition, Springer-Verlag New York.
2. Kruschke: (2014) Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan, Academic Press.
3. Jim Albert: (2007) Bayesian Computation with R, Springer, New York.