Bangalore University
Department of Statistics

Jnanabharathi Campus
Bengaluru – 560 056

Syllabus for
III & IV Semester Statistics Papers
Under-Graduate (UG) Program
Framed according to the National Education Policy (NEP 2020)

BANGALORE UNIVERSITY

Syllabus for III and IV Semester

in

B.Sc. and B.Sc. (Honours) Course for STATISTICS (CBCS 2021)

The Board Studies in Statistics (PG&UG) approved the syllabus for III and IV Semester of BSc, BSc(Honours) The Board of Studies consists of experts as below:

1. Prof. Parameshwar V Pandit, - Chairperson
Professor and Chairperson, Department of Statistics
Bangalore University, Bengaluru

2. Dr. Surekha B Munoli - Member
Professor, Department of Statistics
Karnataka University,Dharwad.

3. Dr Sujata Inginshetty - Member
Professor and Chairperson, Department of Statistics
Gulbarga University, Kalaburgi

4. Dr. Kala R. Nayak - Member
Registrar, Don Bosco College of Engineering,
Fatoda, Madgao, Goa

5. Dr. Satish S. Bhat - Member
Associate Professor, Department of Statistics
Yuvaraja College, Mysore

6. Dr. Suresh, R. - Member
Assistant Professor, Department of Statistics
Bangalore University, Bengaluru

7. Dr. Mallappa - Member
Assistant Professor, Department of Statistics
Bangalore University, Bengaluru

8. Dr. Mohini Bhat - Member
Assistant Professor, Department of Statistics
Christ Institute of Advanced Studies, Christ Nagar
Begur-Koppa Road, Sakkalawara Post, Bengaluru

9. Divya, V. R. - Member
Assistant Professor, Department of Statistics
St. Clariet College, MES Ring Road, Jalahalli
Bengaluru

Scheme of Instruction/ Examination

1. The subject of Statistics in this course has to be taught by MSc/MA degree holders in Statistics / Applied Statistics.
2. The theory question paper for each paper shall cover all the topics in the pertaining syllabus with proportional weightage to the number of hours of instruction prescribed.
3. The practicals are to be conducted in batches as per the University norms for the faculty of science (normally 10 students per batch per teacher).
4. Two teachers are to be assigned for each batch with not more than 20 students for giving instructions, supervision, and correction of records.
6. It is expected that each student collects and uses real life data for the practical classes.
7. Students are required to use Statistical software, run the programmes, and enclose computer outputs to the practical records in the case of computer based practicals.
8. Maximum marks for each record in the examination is 5.
9. Study tour for the students is strongly recommended to gain practical knowledge of applications of Statistics in Industries/Agriculture/Medical field.

Program Outcomes

By the end of the program the students will be able to:

1. Acquire fundamental/systematic or coherent understanding of the academic field of Statistics and its different learning areas and applications.
2. Develop and demonstrate abilities to understand major concepts in various disciplines of Statistics.
3. Demonstrate the ability to use skills in Statistics and different practicing areas for formulating and tackling Statistics related problems and identifying and applying appropriate principles and methodologies to solve a wide range of problems associated with Statistics.
4. Understand procedural knowledge that creates different types of professionals related to subject area of Statistics, including professionals engaged in government/public service and private sectors.
5. Plan and execute Statistical experiments or investigations, analyze and interpret data/information collected using appropriate methods, including the use of appropriate statistical software including programming languages, and report accurately the findings of the experiment/investigations.

6. Have a knowledge regarding use of data analytics tools like Excel and R-programming.
7. Developed ability to critically assess a standard report having graphics, probability statements.
8. Analyze, interpret the data and hence help policy makers to take a proper decision.
9. Recognize the importance of statistical modeling and computing, and the role of approximation and mathematical approaches to analyze the real problems using various statistical tools.
10. Demonstrate relevant generic skills and global competencies such as
    i. Problem-solving skills that are required to solve different types of Statistics related problems with well-defined solutions, and tackle open-ended problems, that belong to the disciplinary-area boundaries;
    ii. Investigative skills, including skills of independent thinking of Statistics-related issues and problems;
    iii. Communication skills involving the ability to listen carefully, to read texts and reference material analytically and to present information in a concise manner to different groups/audiences of technical or popular nature;
    iv. Analytical skills involving paying attention to details and ability to construct logical
    v. Arguments using correct technical language related to Statistics and ability to translate them with popular language when needed;
    vi. ICT skills;
    vii. Personal skills such as the ability to work both independently and in a group.
11. Undertake research projects by using research skills- preparation of questionnaire, conducting national sample survey, research projects using sample survey, sampling techniques.
12. Understand and apply principles of least squares to fit a model to the given data, study the association between the variables, applications of Probability Theory and Probability Distributions.

Assessment
Weightage for assessments (in percentage)

Type of Course	**Formative Assessment / IA**	**Summative Assessment**
Theory	**40**	**60**
Practical	**25**	**25 (20+5(Practical record))**
Projects	**40**	**60**
Experiential Learning (Internships, etc.)	**40**	**60**

STATISTICS

Syllabus for III and IV Semester B.Sc. with Statistics as Major and as Minor

III Semester

Course Title: **Calculus and Probability Distributions**	
Total Contact Hours: 56	Course Credits:04
Formative Assessment Marks: 40	Duration of ESA/Exam: 2 hours
Model Syllabus Authors: State level NEP-model curriculum setting committee members-Statistics	Summative Assessment Marks: 60

Number of Theory Credits	**Number of lecture hours/semester**	**Number of practical Credits**	**Number of practical hours/semester**
4	**56**	**2**	**52**

Course Objectives

To enable the students to

1. Know the concept of continuity, differentiability, integration of one and more variables.
2. Define and describe properties of Joint, Marginal and conditional distributions of variables and some key concepts of probability theory.
3. Understand different discrete, continuous and sampling distributions, properties and their applications.
4. Generate random variables from various distributions using R-code.

Course outcomes

After completion of this course the students will be able to

1. Judge continuity of a function, find integrations and solve problems of differentiability.
2. Solve problems of various analytical environments using different distributions and their properties.
3. Find sampling distributions of functions of random variables and explore their applications.

STAT 301: Calculus and Probability Distributions

Content of Theory Paper 3	56 Hrs
UNIT 1: Calculus of one and more variables	**15 Hrs**
Review of calculus of one variable: continuity, differentiability, mean value theorem and Taylor series expansion. Functions of several variables: Continuity, directional derivatives, differentials of functions of several variables, the gradient vector. The mean value theorem, a sufficient condition for the existence of the differential, partial derivatives of higher order and Taylor's formula. Applications of partial differentiation, Jacobians. Riemann integrals, integration by parts, mean value theorem. Multiple integrals and evaluation of multiple integrals by repeated integration, Mean-value theorem for multiple integrals. Sequences and Series of real numbers. convergence of sequences and series, tests for convergence of series. (Only results and applications)	
UNIT 2: Distribution of Random Variables (Two-dimensional)	**12 Hrs**
Two dimensional random variables: Joint distribution, Marginal distribution and Conditional distributions of random variables, conditional expectation, covariance, correlation and moments. Distribution of functions of random variables using m.g.f. and distribution function.	

Transformation of variable technique (one and two variables). Chebyshev's inequality- proof and its use in approximating probabilities; Statements of Weak Law of Large Numbers; Convergence in law and Central Limit theorems – De-Moivre. (Some simple examples)	
UNIT 3: Probability Distributions-II	**16 Hrs**
Discrete distributions: Rectangular, Geometric, Negative Binomial, Hypergeometric, Multinomial- definition through probability mass function, mean, variance, moments, p.g.f., m.g.f., other properties and applications. Continuous distributions: Uniform, Gamma, Exponential, Beta (type 1 and type 2), Cauchy, Weibull– definition through probability density function, mean, variance, moments, m.g.f., other properties and applications. Bivariate normal distribution- definition through probability density function, marginal and conditional distribution.	
UNIT 4: Sampling Distributions and Simulation	**13 Hrs**
Definitions of random sample, parameter and statistic, sampling distribution of sample mean, standard error of sample mean, sampling distribution of sample variance, standard error of sample variance. Exact sampling distributions: Chi square distribution- mean, variance, moments, mode, additive property. Student's and Fisher's t-distribution- mean, variance, moments and limiting form of t distribution. Snedecor's F-distribution: mean, variance and mode. Distribution of 1/F. Relationship between t, F and χ^2 distributions. Introduction to simulation. Generation of random observations from Uniform, Exponential, Normal, Binomial, Poisson distributions using R-codes.	

References

1. Andre I Khuri (2003). Advanced Calculus with Applications in Statistics, Second Edition, John Wiley & Sons.
2. Ghorpade, S. R. and Limaye, B. V. (2006). A Course in Calculus and Real Analysis, Springer
3. Gupta S.C. and V.K. Kapoor (2020), Fundamental of Mathematical Statistics, Sultan Chand and Co. 12th Edition.
4. Hogg, R. V. McKean J. W. and Craig, A. T. (2012), Introduction to Mathematical Statistics, Pearson 7th Edition.
5. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009), Probability and Statistical Inference,10th Edition, Pearson Education, New Delhi.
6. Jay Kerns, G. (2010). Introduction to Probability and Statistics using R. 1st Edition, Springer.
7. Rohatgi, V.K. and A.K. Md. Ehsanes Saleh. (2002). An Introduction to Probability Theory and Mathematical Statistics, New York, John Wiley.
8. Ross, S. M. (2014). Introduction to Probability Models. 11th Edition, Elsevier science.
9. Ross, S. M. (2012). Simulation. Academic Press.
10. Shanthi Narayana (2000), Integral Calculus, S. Chand & Co. Ltd.
11. Shanti Narayana (2000). Differential Calculus, S. Chand & Co. Ltd.
12. Verzani, J. (2002). Simple R - Using R for Introductory Statistics.

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.

2. Students are encouraged to use resources available on open sources.

Formative Assessment: Total 40 marks	
Assessment Occasion/ type	**Weightage in Marks**
Internal Test 1	15
Internal Test 2	15
Assignment/Seminar (7 marks)+Attendance(3marks)	10
Total	40

STS 302: Contents of Practical 3

Note: The first practical assignment is on R-programming. Practical assignments 2 to 10 have to be first solved manually (using scientific calculators) and executed using R-programming.

1. Demonstration of R functions for calculus, distribution of random variables, probability distributions, sampling distribution simulation.
2. Numerical differentiation and integration.
3. Bivariate Probability Distributions - Marginal and Conditional distributions,
4. Bivariate Probability Distributions - Conditional Mean, Conditional Variance, Correlation.
5. Applications of Chebyshev's inequality (For standard distributions such as Normal, Exponential, Gamma).
6. Applications of discrete probability distributions - Negative – Binomial, Geometric, Hyper geometric and discrete uniform, multinomial distributions.
7. Applications of continuous probability distributions - Exponential, Gamma, Cauchy, Weibull distributions.
8. Fitting of discrete and continuous distributions.
9. Generating random sample from discrete distributions.
10. Generating random sample from continuous distributions.

Formative Assessment: Total 25 marks	
Assessment Occasion/ type	**Weightage in Marks**
Internal Test 1	10
Internal Test 2	10
Attendance	5
Total	25

IV Semester

Course Title: **Statistical Inference-I**	
Total Contact Hours: 56	Course Credits:04
Formative Assessment Marks: 40	Duration of ESA/Exam: 2 hours
Model Syllabus Authors: State level NEP-model curriculum setting committee members-Statistics	Summative Assessment Marks: 60

Number of Theory Credits	**Number of lecture hours/semester**	**Number of practical Credits**	**Number of practical hours/semester**
4	**56**	**2**	**52**

Course Objectives

To enable the students to understand the concepts of

1. Families of distributions, order statistics and their distributions.
2. Estimation, criteria for estimators, methods of estimation, confidence interval.
3. Testing of Hypotheses and its theoretical aspects, large and small sample tests.

Course Outcomes

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

1. Carryout statistical analysis by identifying families of distributions and the use of order statistics.
2. To find estimators using different methods of estimation and compare estimators.
3. To carryout statistical inference using different tests of hypotheses under different scenarios.
4. Generate random variables and use these generated random variable for illustration of concepts studied in this course.

STS 401: Statistical Inference-I

Content of Theory Paper	56 Hrs
UNIT- 1: Point Estimation-I	**16 Hrs**
Families of distributions- location and scale families. Single parameter exponential family. Concept of order statistics, Distribution of maximum and minimum order statistics (with proof) and r^{th} order statistic (without proof). Concepts of estimator and estimate. Criteria for estimators: Unbiasedness, Consistency. Invariance property of consistent estimators. Efficiency and relative efficiency. Mean squared error as a criterion for comparing estimators. Sufficient statistics. Statement of Neyman-Factorization theorem.	
UNIT-2: Point Estimation-II	**12 Hrs**
Fisher information function. Statement of Cramer–Rao inequality and its applications. Minimum Variance Unbiased Estimator and Minimum Variance Bound Estimator. Maximum likelihood and method of moment estimation; Properties of MLE and moment estimators and examples. Method of Scoring	
UNIT- 3: Testing of Hypotheses	**18 Hrs**
Statistical hypotheses - null and alternative, Simple and composite hypotheses. Type-I and Type-II errors, test functions. Randomized and non-randomized tests. Size, level of significance, Power function, power of tests. Critical region, p- value and its interpretation. Most Powerful (MP) and UMP test. Statement of Neyman-Pearson Lemma and its applications. Likelihood ratio tests. Large and small samples tests of significance. Tests for single mean, equality of two means, single variance and equality of two variances for normal populations. Tests for proportions.	

UNIT- 4: Interval Estimation	10 Hrs
Confidence interval, confidence coefficient, shortest confidence interval. Methods of constructing confidence intervals using pivotal quantities. Construction of confidence intervals for mean, difference of two means, variance and ratio of variances, proportions, difference of two proportions and correlation coefficient.	

References

1. Chihara, L. and Hesterberg, T. (2011) Mathematical Statistics with Resampling and R. Wiley.
2. Gupta S.C. and V.K. Kapoor (2020), Fundamental of Mathematical Statistics, Sultan Chand and Co. 12th Edition.
3. Hogg, R. V. McKean J. W. and Craig, A. T. (2012), Introduction to Mathematical Statistics, Pearson 7th Edition.
4. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009), Probability and Statistical Inference, 10th Edition, Pearson Education, New Delhi.
5. Johnson, R.A. and Bhattacharyya, G.K. (2006), Statistics: Principles and methods. 5th Edition, John Wiley & Sons, New York.
6. Kale, B.K. (1999). A First Course on Parametric Inference, New Delhi, Narosa Publishing House.
7. Kendall, M.G., et. al., (1996). An Introduction to the Theory of Statistics, Universal Book Stall.
8. Rohatgi, V.K. and A.K. Md. Ehsanes Saleh. (2002). An Introduction to Probability Theory and Mathematical Statistics, New York, John Wiley.
9. Ross, S.M. (2014), Introduction to Probability and Statistics for Engineers and Scientists, 5th Edition, Academic Press.

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.

2. Students are encouraged to use resources available on open sources.

Formative Assessment: Total 40 marks	
Assessment Occasion/ type	**Weightage in Marks**
Internal Test 1	15
Internal Test 2	15
Assignment/Seminar (7 marks)+Attendance(3marks)	10
Total	40

STS 402 Practical 4

Note: The first practical assignment is on R-programming and R packages. Practical assignments 2 to 10 have to be first solved manually (using scientific calculators) and executed using R-programming.

1. Demonstration of R-functions for estimation and testing of hypotheses.
2. Point estimation of parameters and obtaining estimate of standard errors and mean square error.
3. Computing maximum likelihood estimates.
4. Computing moment estimates.
5. Interval estimation: Construction of confidence interval (large and small samples)
6. Evaluation of Probabilities of Type – I and Type – II errors and power of tests.
7. Small sample tests: Tests for mean, equality of means under normality when variance is (i) known (ii) unknown, P-values.
8. Small sample tests: single proportion and equality of two proportions, variance and equality of two variances under normality. P-values for the above tests.
9. Large sample tests: Tests for mean, equality of means when variance is (i) known (ii) unknown, under normality, variance and equality of two variances under normality. P-values for the above tests.
10. MP and UMP tests for parameters of binomial, Poisson distributions, normal and Exponential (scale parameter only) distributions and power curve.

Formative Assessment: Total 25 marks	
Assessment Occasion/ type	**Weightage in Marks**
Internal Test 1	10
Internal Test 2	10
Attendance	5
Total	25

List of Open Electives (OE)

1) **Population Studies**
2) **Survival Models**
3) **Basics of Operations Research**
4) **Quantitative Analysis Techniques**

1. Population Studies

Course Objectives

1. To enable the students to identify appropriate sources of data, perform basic demographic analysis using various techniques and ensure their comparability across populations.

2. To acquire knowledge about the construction of life table and its applications in demographic analysis.

Course Outcomes (CO)

Upon successful completion of this course the student will be able to

CO1. Study the concepts of Vital Statistics, sources of data, different measures of Fertility, Mortality and migration.

CO2. Understand the Growth rates- GRR and NRR and their interpretations.

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.

2. Students are encouraged to use resources available on open sources.

Contents

UNIT-1: Introduction and Sources of Population Data **14 hours**

History, definition, nature and scope of population Studies. Sources of population data – salient features of Census, Civil Registration System, National Sample Surveys, Demographic Surveys, relative merits and demerits of these sources. Coverage and content errors. Use of balancing equations, Chandrasekar-Deming formula to check completeness of vital registration data, use of Whipple's, Myer's and UN indices.

UNIT-2 : Fertility, Mortality **14 hours**

Basic concepts and terms used in the study of fertility. Measures of fertility- Crude Birth Rate (CBR), General Fertility Rate (GFR), Age Specific Fertility Rate (ASFR), Total Fertility Rate (TFR), Birth order statistics, Child Women ratio. Measures of reproduction- Gross Reproduction Rate (GRR)and Net Reproduction rate(NRR). Measurement of population growth rate- simple growth rate and compound growth rate.

Basic concepts and terms used in the study of mortality. Measures of mortality- Crude Death Rate (CDR), Age Specific Death Rate (ASDR), Direct and Indirect Standardized Death rates, Infant Mortality Rate (IMR), Under-five mortality Rate, Neo-natal mortality rate, Post-natal mortality rate; Maternal Mortality Rate (MMR).

UNIT-3: Life tables and Population change **14 hours**

Life tables: Components of a life table, force of mortality and expectation of life table, types of life tables. Construction of life tables using Reed-Merrell's method, Greville's method. Uses of life tables.

Basic concepts and definition of population change, migration. Types of migration- internal and international, factors affecting migration. Rates and ratios of Migration-Indirect measures of net-internal migration, national growth rate method, residual method, push-pull factors Population estimates and projections.

References

1. Barclay, G, W(1968). Techniques of Population Analysis, John Wiley and Sons, Incs. New York/London.
2. Keyfitz, H (1968). Introduction to the Mathematics of Population. Addison-Wesley Publishing Co.
3. Pathak, K.B and Ram, F (1991).Techniques of Demographic Analysis, Himalaya Publishing House.
4. Ramakumar. R (1986). Technical Demography, Wiley Eastern Ltd.
5. Srinivasan. K (1998). Basic Demographic Techniques and Applications, Sage Publication, New Delhi.
6. Wunsch G.J. & M.G. Tarmota(1978). Introduction to Demographic Analysis, Plenum Press, N.Y.

2. Survival Models

Course Objectives

1. Enable the students to construct and interpret life tables.
2. To understand the concepts of Survival analysis.
3. To study the design of clinical trials and their analysis.

Course Outcomes:

By the end of this course, the student should be able to:

CO1.Explain Life Tables, types of life tables, its functions, construction.

CO2. Describe multiple decrement life tables and their construction.

CO3. Know survival models, concepts of survival analysis, notion of ageing.

CO4. Explain key concepts in the design of clinical trials, phases, types, clinical trial protocol, analysis.

Content

Unit-1: Life Tables **14 hours**

Basic definition and notations, Types of life tables, inter – relationships between life table functions, Properties of life table functions. Construction of life tables using Reed – merrel and Greville's Method. Competing causes of failure/death, Multiple decrement life tables and their construction (with examples).

Unit-2: Survival Concepts **14 hours**

Life distributions, survival functions, failure rate, Integrated hazard function, residual life time, mean residual life time. Notion of aging: IFR, IFRA, DMRL, NBU, NBUE classes of life distributions and their dual classes. Common Life Distributions: binomial, Poisson, exponential, Weibull, gamma, Pareto and log-normal distributions.

Unit-3: Clinical Trials **14 hours**

Basics of Clinical Trials: Who can be in clinical trials? need clinical trials, Brief History of Clinical Trials, Common Terms in clinical Trials: Clinical Research, Healthy Volunteer, Inclusion/Exclusion Criteria, Informed

Consent, Patient Volunteer, Phases of Clinical Trials, Placebo, Protocol, Principal Investigator, Randomization, Single- or Double-Blind, Studies, Types of Clinical Trials. - Diagnostic trials, Natural history studies, Prevention trials, Quality of life trials, Screening trials, Treatment trials, therapeutic trials and prophylactic trials. Observational studies – Cross sectional studies, prospective studies, retrospective studies, randomized control studies. Clinical Trial Protocol and its components. Type of analyses: ITT, mITT and PP. Odds ratio, Relative risk, Sensitivity, specificity, false negative and false positive rates. Receiver operating characteristic(ROC) curve.

References

1. Deshpande, J V and Purohit, Sudha (2005). Life Time Data: Statistical Models and Methods. World Scientific.
2. Friedman, Furberg, and DeMets. (2010). Fundamentals of Clinical Trials (4th Edition). Springer, Free text available online at http://dx.doi.org/10.1007/978-1-4419-1586-3
3. Lawrence MF, Curt DF, David LD (2010), Fundamentals of clinical trials.
4. R. Ramkumar (1986), Technical Demography, Wiley Eastern, New Delhi.
5. Shryock, Henry S, Jacob S, Siegel and Associates (1964). Methods and materials of demography (condensed edition), Academic press, London.

3. Basics of Operations Research

Course Objectives

1. Students get knowledge about the scope and application of Operations Research(OR) in business and industry.
2. Exposes the students to various OR tools and models.
3. To get knowledge about various decision making through OR models.

Course Outcomes

Students will be able to

CO1- Generate mathematical models of business environment.

CO2-Analyze the business situations.

CO3-Use different solution procedures through OR models.

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.

2. Students are encouraged to use resources available on open sources.

Contents

Unit-1: Introduction to Operations Research(OR) 14 hours

Origin and growth of OR, importance of OR in managerial decision making, scope and applications of OR, models and modelling in OR. Linear programming problems(LPP): Formulation of the problem, feasible & infeasible, basic feasible solution, optimal, unbounded and multiple optimal solutions of LPP, solution by graphical method. Slack, Surplus and Artificial variables. Duality in LPP, Importance of Duality Concepts, Formulation of Dual Problem, Economic Interpretation of Duality.

Unit-2: Allocation Problems **14 hours**

Transportation problems: Formulation, methods of finding initial solution (North West Corner Rule, Least Cost Method and Vogel's Approximation Method), unbalanced transportation problems, maximization transportation problem.

Assignment problems: Formulation, methods of solution, Hungarian method, multiple optimal solutions, unbalanced problems, maximization problems.

Unit-3: Decision theory **14 hours**

Game theory: Basic concepts. Two – Person Zero Sum Game. Pure and Mixed Strategies. Maximin – Minimax principle, Games with and without saddle points. Principle of dominance.

Concepts of decision making, decision making environments, Decision making under uncertainty - Decision making under risk, decision tree analysis. Case discussion.

Concepts of network analysis, project network models, Critical Path Method, PERT.

References

1. Hillier, F S, et al. Introduction to Operations Research (9/e). Tata McGraw Hill, 2011.
2. Ravindran, A and Don T Phillips. Operations Research: Principles and Practice. John Wiley & Sons, 1987.
3. Sharma, J K. Operations Research: Theory and Applications (5/e). New Delhi: Laxmi Publications, 2013.
4. Taha, Hamdy A. Operations Research: An Introduction (9/e). Prentice Hall, 2010.
5. Vohra, N D. Quantitative Techniques for Management. Tata McGraw Hill Education, 2015.
6. Kanti Swarup, Gupta, P.K. and Man Mohan: Operations Research, Sultan Chand & Sons, New Delhi.
7. Kapoor, V.K: Operations Research, Sultan Chand & Sons, New Delhi.
8. Kapoor, V.K.: Operations Research Problems & Solutions, Sultan Chand & Sons, New Delhi.

4. Quantitative Analysis Techniques

Course Objectives

To enable the students to acquire the knowledge about

1. The concepts of correlation and regression analysis.
2. The concepts of linear programming problem and its applications.
3. The students will learn the tools of data mining.

Course Outcomes

Students will be able to

CO1. Carryout correlation and regression analysis.

CO2. Formulate and solve linear programming problems.

CO3. Use data mining tools.

Pedagogy

The course is taught using traditional chalk and talk method using problem solving through examples and exercises. Students are encouraged to use resources available on open sources.

UNIT- 1: Correlation and regression analysis **14 hours**

Correlation- Definition, Types - Simple, multiple, partial. Causation - Spurious, positive, negative, perfect and no correlation, explanation with examples. Importance of correlation analysis. Measurement of correlation- scatter diagram, Karl Pearson's coefficient of correlation, Properties of coefficient of correlation, interpretation. Spearman's coefficient of rank correlation – with and without ties, interpretation. Coefficient of determination and its interpretation.

Regression- Definition, regression lines/equations of X on Y and Y on X. Properties of regression coefficients and regression lines/equations. Principle of least squares and fitting of linear, quadratic and exponential curves. Uses of regression analysis. Comparison between correlation and regression.

UNIT-2: Linear programming problem(LPP) 18 hours

Definition and scope of Operations Research (OR). Modelling and solution. Linear Programming Problem (L.P.P): Definition, Standard forms. Formulation of LPP. Basic Solutions, degenerate and non-degenerate solutions. Graphical method of solving LPP. Criteria for unbounded, Multiple and infeasible solutions.

Transportation problem: Mathematical formulation. Existence of feasible solution. Finding initial basic feasible solution: North West Corner Rule, matrix minima method and Vogel's method. Unbalanced transportation problem.

Assignment Problem: Mathematical Formulation and Hungarian algorithm. Unbalanced assignment problem.

UNIT-3: Data Mining 10 hours

Motivations and importance of Knowledge Discovery in Databases (KDD) process - search, induction, querying, approximation and compression. Kinds of data considered for data mining, basic data mining tasks, data mining issues, Data Mining models - predictive and descriptive, interconnections between Statistics and Data Mining. Artificial Intelligence and Machine Learning. Applications of data mining.

References

1. Goon, A.M., Gupta, M.K. and Dasgupta, B. (2002). Fundamentals of Statistics, Vol. I, 8th Ed., The World Press, Kolkata.

2. Ross, S.M. (2014). Introduction to Probability and Statistics for Engineers and Scientists, 5th Edition, Academic Press.

3. Kanthi Swaroop, Manmohan and P. K. Gupta (2013). Operation Research, Sultan Chand New Delhi.

4. Mustafi, C. K. (2006). Operations Research Methods and Practice, 3/e. New Age Publication.

5. Narag. A. S. (1970). Linear Programming and Decision Making. Sultan Chand and Co.

6. Sharma, J K.(2013). Operations Research: Theory and Applications (5/e). New Delhi: Laxmi Publications.

7. Jiawei Han, Micheline Kamber (2002). Data mining concepts and Techniques, Morgan Kaufman Publishers, USA.

8. Trevor Hastie, Robert Tibshirani and Jerome Friedman (2001). The elements of Statistical learning: Data Mining, Inference and Prediction, Springer, New York.
9. Rajan Chattamvelli (2009). Data mining methods, Narosa Publishing House.