

# M.A/M.Sc. PROGRAMME IN STATISTICS



(Effective from Session 2025-26)  
(Batch: 2025-2027)



**SAMBALPUR UNIVERSITY**  
JYOTI-VIHAR, BURLA, SAMBALPUR, ODISHA-768019



**SEMESTER-WISE COURSE STRUCTURE FOR THE TWO YEARS P.G PROGRAMME IN  
UNIVERSITY P.G DEPARTMENT AND COLLEGES UNDER SAMBALPUR UNIVERSITY**

**TO BE EFFECTIVE FROM 2025-2026  
BATCH: 2025-27  
(Ref: letter No: 4873/Acd.-I Dated 21.08.2023)**

<b>For (Science/ Humanities/Social Sciences/ Commerce)</b>				
<b>Semester</b>	<b>Core Course Credit</b>	<b>Additional Course</b>	<b>Additional Course Credit</b>	<b>Total Credit</b>
First	20	<b>AECC I: Environmental Studies and Disaster Management</b>	2	22
Second	20	Inter Dept. Course (IDC) or open elective	3	23
Third	20	<b>AECC II: Industrial Field Survey</b>	2	22
Fourth (including project of 4 credit)	20	MOOCs one paper	3	23
<b>TOTAL</b>	<b>80</b>		<b>10</b>	<b>90</b>
	<b>Total credit for 2 years course = 90 Credits</b>			
	Furthermore, following non - credit course will be taken by the students			
1. Yuva Sanskar		2. N.C.C/N.S. S/Sports/Performing Arts/Yoga (Of which one has to be opted)		

## MISSION

M1	Educate society for generations by providing transformative education with deep disciplinary knowledge and concern for environment
M2	Develop problem solving, leadership and communication skill in student participants to serve the organisation of today and tomorrow
M3	Aim for the holistic development of the students by giving them value based ethical education with concern for society
M4	Foster entrepreneurial skills and mindset in the students by giving life-long learning to make them responsible citizens

## Programme Education Objectives

PEO1	Understand the nature and basic concepts of <b>Statistics</b> .
PEO2	Analyse the relationships among different concepts
PEO3	Perform procedures as laid down in the areas of study
PEO4	Apply the Basic Concepts learned to execute them

## Programme Outcomes

PO-1	<b>Critical Thinking:</b> Take informed actions after identifying the assumptions that frame our thinking and actions
PO-2	<b>Effective Communication:</b> Will be able to speak, read, write and listen clearly in person and through electronic media in English and in one Indian Language
PO-3	<b>Social Interaction (Interpersonal Relation):</b> Elicit views of others, mediate disagreements and prepared to work in team
PO-4	Entrepreneurship Capability: Demonstrate qualities to be prepared to become an entrepreneur
PO-5	Ethics: Recognize different value systems including your own, understand the moral dimensions and accept responsibility for them
PO-6	Environment and Sustainability: Understand the issues of environmental contexts and sustainable development
PO-7	Life-Long Learning: Acquire the ability to engage in independent and life-long learning in the context of socio-technological changes

**SEMESTER-WISE SKELETON OF THE TWO YEARS P.G. PROGRAMME IN UNIVERSITY P.G.  
DEPARTMENTS AND COLLEGES UNDER SAMBALPUR UNIVERSITY**

<b>P.G. IN STATISTICS</b>			
	From the Dept		
Semester	Credit		Credit
First	20	Entrepreneurship Development/ Env. Studies & Disaster Management	2
Second	20	Inter Dept. course (IDC) or Open Elective	3
Third	20	Env. Studies & Disaster Management/ Entrepreneurship Development	2
Fourth (including Project of 4 credit)	20		
Total	80	MOOCs one paper (in 2 <sup>nd</sup> or 3 <sup>rd</sup> Sem)	3
Total credit for 2 year course = 90 credits			
Furthermore, the following non-credit course will be taken by the student			
1. Yuva Sanskar	2. N.C.C./ N.S.S./Sports/ Performing Arts/Yoga (of which one has to be opted)		

- In each Semester, the Department can offer either 5 papers of 4 credits (i.e., 5x4=20 credits)
- The results of 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> Semester will be published on the basis of 20 credits core course only. The results of fourth and final semester results will incorporate taking into consideration 80 credits core course offered by the Department and 10 credit courses comprising (i) Env. Studies & Disaster Management (2 credit), (ii) Inter Dept. Course (IDC) or Open elective (3 credit), (iii) Entrepreneurship Development (2 credit) (iv) MOOCs (3 credit).
- The students will take one MOOCs Course according to his/her preference in consultation with HOD and submit the document in support of undertaking the MOOCs course to the respective Department.
- Students will apply in prescribed form their preference for NCC/NSS/Sports/Performing Art/Yoga at the beginning of the session. The consolidated list of the same will be forwarded by the HODs to the office of the Chairman, PG Council. Depending upon the number of application, maximum capacity and preference, the students will be allotted one of the above non-credit courses.

**1. Distribution of Marks in % for the theory Papers**

Theory Papers offered by the own Department		Theory Paper offered by other Department	
End Term	Mid Term	End Term	Mid Term
80%	20%	60%	40%
Mid- term will consist of (10% for 2 class tests, 5% for assignment, 5% for case study in case of mid-term of 20%; 20% for 2 class tests, 10% for assignment, 10% for case study in case of mid-term of 40%)			

2. Project work of 4<sup>th</sup> Semester will be assigned to the students (jointly or individually) at the beginning of the III Semester and will be completed in the IV semester. The distribution of the work/marks will be as follows:

Semester wise work and distribution of marks in % for Project			
III Semester (20%) Evaluation of Interim Report of the Project Work			
Background of the Problem (5%)	Review of Literature (5%)	Objectives (5%)	Methodology (5%)
IV Semester (80%) Evaluation of Final Report of the Project Work			
Project work (50%)		Viva (30%)	

3. Question Pattern and Mark Distribution of the Theory Paper

- a. For End-Term Examination of Total Marks 80 (Four Units Course)

Q.1 Twenty Questions of 1 mark each (Questions patterns should be of MCQ, Fill in the Blanks, True/False, Definition etc.)

(20 x 1 = 20)

For Unit-I, Unit-II, Unit-III and Unit-IV

Each question will be of 15 marks, and it should have alternative in each unit. The distribution of 15 marks will be decided by the paper setter.

[Suggested patterns distribution of 15 marks are 15; 8+7; 7.5+7.5; 5+10, 2+3+10, 5+5+5 etc.]

- b. For End-Term Examination of Total Marks 60 (Four Units Course)

Q.1 Twelve Questions of 1 mark each (Questions patterns should be of MCQ, Fill in the Blanks, True/False, Definition etc.)

(12 x 1 = 12)

For Unit-I, Unit-II, Unit-III and Unit-IV

Each question will be of 12 marks and there should have alternative in each unit. The distribution of 12 marks will be decided by the paper setter.

[Suggested patterns distribution of 12 marks are 12; 8+4; 6+6; 2+3+7, 4+4+4 etc.]

4. Pass percentage

For each paper pass percentage is 30% (Credit 4). For clearing the semester Overall Grade Point should be 4.5 (40%).

For IDC, Environmental Studies and Disaster Management and Entrepreneurship Development Programme the pass percentage is 30% (Credit 4).

For MOOC course the pass percentage is as per the programme guideline.

**POST GRADUATE DEPARTMENT OF STATISTICS  
AND AFFILIATED COLLEGES UNDER SAMBALPUR UNIVERSITY,  
OUTLINE OF COURSE STRUCTURE**

**MA/M.Sc. STATISTICS (Session: 2025-2027)**

<b>Subject Code</b>	<b>Title of the Course</b>	<b>End Term Marks</b>	<b>Internal Marks</b>	<b>Credit Hour</b>	<b>Maximum Marks</b>
<b>SEMESTER I</b>					
ST-C-411	Statistical Methods-I	80	20	4 CH	100
ST-C-412	Probability-I	80	20	4 CH	100
ST-C-413	Mathematical Analysis and Algebra	80	20	4 CH	100
ST-C-414	Sampling Methods	80	20	4 CH	100
ST-C-415	Data Structure and C-Programming	80	20	2 CH	100
ST-C-416	LAB-I: Laboratory using MS-Excel/ Calculator (ST-C-411&ST-C-414)	80	20	2 CH	100
ST-C-417	Env. Studies & Disaster Management	60	40	2CH	100
	Total Credit Hours for Semester	540	160	22 CH	700
<b>SEMESTER II</b>					
ST-C-421	Statistical Methods-II	80	20	4 CH	100
ST-C-422	Probability-II	80	20	4 CH	100
ST-C-423	Statistical Inference-I	80	20	4 CH	100
ST-C-424	Stochastic Modeling	80	20	4 CH	100
ST-C-425	Applied Statistics	80	20	2 CH	100
ST-C-426	LAB-II: Laboratory using SPSS/STATA(ST-C-421&ST-C-425)	80	20	2 CH	100
	Inter Disciplinary Course (IDC) or Open Elective	60	40	3CH	100
	Total Credit Hours for Semester II	540	160	20 CH	700
<b>SEMESTER III</b>					
ST-C-511	Statistical Inference-II	80	20	4 CH	100
ST-C-512	Multivariate Analysis	80	20	4 CH	100
ST-C-513	Optimization Technique-I	80	20	4 CH	100
ST-E-514(*)	Special Paper-I (anyone)	80	20	4 CH	100
ST-C-515	Time Series & Forecasting	80	20	2 CH	100
ST-C-516	LAB-III: Laboratory using R/Python (ST-C-512, ST-E-514(*)&ST-C-515)	80	20	2 CH	100
	Entrepreneurship Development	60	40	2CH	100
	Total Credit Hours for Semester III	540	160	22 CH	700
<b>SEMESTER IV</b>					
ST-C-521	Design and Analysis of Experiment	80	20	4 CH	100
ST-E-522(*)	Special Paper-II (anyone)	80	20	4 CH	100
ST-E-523(*)	Special Paper-III (anyone)	80	20	4 CH	100
ST-C-524	Project and Viva Voce	80	20	4 CH	100
ST-C-525	Research Methodology and Ethics	80	20	2 CH	100
ST-C-526	LAB-IV: Laboratory using	80	20	2 CH	100

	R/Python/STATA/SPSS/MATLAB {ST--521, ST-E-522()* , ST-E-523()* }				
	MOOCs/Alternative to MOOCs	<b>100</b>		<b>3CH</b>	<b>100</b>
	Total Credit Hours for Semester-IV	<b>580</b>	<b>120</b>	<b>23 CH</b>	<b>100</b>
<b>TOTAL (A)</b>		<b>2200</b>	<b>600</b>	<b>90 CH</b>	<b>700</b>

<b>Semester</b>	<b>Title of the course</b>	<b>Credit Hours</b>
First(Sc.)/Third(Arts)	Env. Studies & Disaster Management	<b>2 CH</b>
Second	Inter Disciplinary Course (IDC) or Open Elective	<b>3 CH</b>
First(Arts)/Third(Sc.)	Entrepreneurship Development	<b>2 CH</b>
Second or Third	MOOCs one Paper	<b>3 CH</b>
<b>Total (B)</b>		<b>10 CH</b>
<b>Grand Total (Total A+ Total B)</b>		<b>90 CH</b>

### **Distribution of Marks**

<b>AP</b>	Assignment /and Presentation	5 Marks
<b>CT-1</b>	Class test-1	5 Marks
<b>CT-2</b>	Class test-2	5 Marks
<b>PF</b>	Project/Field Work	5 marks
<b>IA</b>	Internal Assessment	20 Marks
<b>EE</b>	External Evaluation	80 Marks

## SPECIAL PAPERS

**(In 3rd Semester one Special paper has to be chosen from the following list)**

<b>Subject code ST-E-514()*</b>	<b>Title of the course</b>	<b>Credit Hours</b>
A	Econometrics	<b>4 CH</b>
B	Biostatistics& Epidemiology	<b>4 CH</b>
C	Statistical Decision Theory	<b>4 CH</b>
D	Pattern Recognition	<b>4 CH</b>

**(In 4th Semester two Special papers has to be chosen from the following list)**

<b>Subject code ST-E-522()*</b>	<b>Title of the course</b>	<b>Credit Hours</b>
E	Statistical Quality Control and Reliability	<b>4CH</b>
F	Bayesian Inference	<b>4CH</b>
G	Linear Models and Regression Analysis	<b>4CH</b>
H	Demography	<b>4CH</b>

<b>Subject code ST-E-523()*</b>	<b>Title of the course</b>	<b>Credit Hours</b>
I	Optimization-II	<b>4CH</b>
J	Actuarial Statistics	<b>4CH</b>
K	Data Warehousing and Data Mining	<b>4CH</b>
L	Advanced Stochastic Process	<b>4CH</b>

Econometrics
Biostatistics & Epidemiology
Statistical Decision Theory
Pattern Recognition
Statistical Quality Control and Reliability
Bayesian Inference
Linear Models and Regression Analysis
Demography
Optimization-II
Actuarial Statistics

ata Warehousing and Data Mining
Advanced Stochastic Process

## **M.A/M.Sc. STATISTICS COURSES OF STUDY FOR 2025-2027**

### **ACADEMIC PROGRAM**

First Semester Examination	December 2025
Second Semester Examination	June 2026
Third Semester Examination	December 2026
Fourth Semester Examination	June 2027

## PROGRAMME OUTCOME

<b>PO-1</b>	<b>Critical Thinking</b> Take informed actions after identifying the assumptions that frame our thinking and actions
<b>PO-2</b>	<b>Effective Communication</b> Will be able to speak, read, write and listen clearly in person and through electronic media in English and in one Indian Language
<b>PO-3</b>	<b>Social Interaction (Interpersonal Relation)</b> Elicit views of others, mediate disagreements and prepared to work in team
<b>PO-4</b>	<b>Entrepreneurship Capability</b> Demonstrate qualities to be prepared to become an entrepreneurship
<b>PO-5</b>	<b>Ethics</b> Recognize different value systems including your own, understand the moral dimensions and accept responsibility for them
<b>PO-6</b>	<b>Environment and Sustainability</b> Understand the issues of environmental contexts and sustainable development
<b>PO-7</b>	<b>Life-Long Learning</b> Acquire the ability to engage in independent and life-long learning in the context of socio-technological

## M.A/M.Sc. STATISTICS COURSES OF STUDY FOR 2025-2027

### MA/MSc. STATISTICS

#### SEMESTER-I

#### STATISTICAL METHODS-I

##### ST-C-4114CH

**Objective:** Statistical methods provide scientific view to conduct the survey in proper way to collect the data about specific perspective. It helps to learn how to develop regression model and apply for the specific perspective data appropriate manner. This paper presents the general theory of statistical distributions as well as the standard distributions found.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Methods-I
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

##### Unit-I

Introduction to different measures such as measures of location, dispersion and skewness, Basic concept of discrete and continuous probability distributions and their properties. Basic discrete distributions- Bernoulli, Binomial, Poisson and Uniform, Computations of their moments, means, Variances and Beta coefficients, recurrence relations between moments, M.G.F. and C.F. Basic continuous distributions: - Uniform and normal distributions their properties and applications. Computation of their M.G.F., C.F. and Moments.

##### Unit-II

Analytical Statistics: - Bivariate data, Scatter diagram, Simple correlations and linear regression. Their properties and applications. Curve fitting, linear and second-degree curve and their applications.

##### Unit-III

Random vectors, joint distribution, joint M.G.F., Independence of random variables. Multinomial distributions and its mean vector and variance and co-variance matrix, bivariate normal distributions, conditional expectation and conditional variance. Its marginal and conditional distributions.

##### Unit-IV

Sampling distributions of Statistics: Sampling distributions of sample mean, sample variance, Derivations of Chi-square, z, t and F distributions, tests of significance based on them and their applications.

##### **Books Recommended:**

1. Fundamental of Mathematical Statistics (2009)- S Chand & Sons- S.C. Gupta and V.K. Kapoor
2. Introduction to Theory and Mathematical Statistics (1988) - Wiley – V.K. Rohatgi,
3. Linear Statistical Inference and its Applications (1975) – Wiely- C.R. Rao

4. An Introduction to Theory of Statistics - Charles Griffiu Yale, G.U. and Kendall, M.G. (1953)
5. Probability and Statistics with Engineering and Computer Science Applications (2005) – Kalyani Swain, A.K.P.C.
6. John E. Freud's Mathematical Statistics with Applications(2006); (7<sup>th</sup> Edition- Pearson Education; Asia-Miller; Irwin and Miller; Marylees.
7. Applied Statistics(1999)- New Central Book Agency- Mukhapaddhyay P.

## PROBABILITY-I

**ST-C-412**

**4CH**

**Objective:** This paper helps to understand basics of set theory and probability theory which deals with uncertain occurrence situations in logical manner.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Probability-I
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Algebra of sets, Fields and Sigma fields, Limit of sequence of subsets, Sigma field generated by a class of subsets, Borel fields. Probability space, continuity of probability measure.

### Unit-II

Sample space, Probability axioms, Conditional probability, Independence of events. Bayes' theorem, Real and vector valued random variables, Distribution function, Discrete and continuous random variables, Distribution of L.V.S. Marginal and conditional distribution. Independence of random variables.

### Unit-III

Poisson theorem, Interchangeable events and their limiting properties, Expectation of a random variable. Linear properties of expectations. Conditional expectation, Moment generating function. Moment inequalities. Characteristic function and its properties.

### Unit-IV

Convergence of a sequence of random variables, Convergence in distribution, Convergence in probability, almost sure convergence and Convergence in quadratic mean and their interrelations. Monotone and dominated convergence theorem, Central limit theorem: Lindberg-Levy and Demoivre-Lapalce theorem.

### Books Recommended:

1. Bhat, B.R. (1985): Modern probability theory (Wiley).
2. Billingsley, P. (1986): Probability and measure (Wiley).
3. Feller, W. (1969): Introduction to probability theory and applications, Vol . II (Wiley)
4. Rohatgi, V.K. (1976): Introduction to theory of probability and mathematical Statistics (Wiley).
5. H.G. Tucker (1967): A graduate course in probability theory (AP)

6. Y.S. Chow and H Teicher(1979): Probability theory (Springer-Verlag),

## MATHEMATICAL ANALYSIS AND ALGEBRA

ST-C-413

4CH

**Objective:** To introduce fundamental concept of Mathematical analysis such as sequence, series of real numbers and their convergence, continuity, differentiability of real valued functions and complex analysis. It will cover the analysis and implementation of algorithms used to solve linear algebra problems in practice. Apply numerical methods to obtain approximate solutions to mathematical problems.

CO-1	Remember and understand the basic concepts/Principles of Mathematical Analysis & Algebra
CO-2	Analyze the Various Concepts to understand them through case studies
CO-3	Apply the knowledge in understanding practical problems
CO-4	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Sequences, subsequences, convergence, divergence, bounded sequences, limits superior and inferior, monotone sequences, Cauchy sequences, completeness, Series of real numbers. Heine-Borel Theorem, Bolzano-Weierstrass Theorem. Functions: limits, continuity, uniform continuity, intermediate value theorem, Differentiability; Chain rule, mean value theorem, Taylor's theorem (statement), extreme, Multivariate calculus: partial, directional and total derivatives mean value theorem, Gamma function and Beta function, Multiple integrals, change of variables, Jacobian formula.

### Unit-II

Algebra of complex numbers, operations of absolute value and conjugate, standard inequalities for absolute value, concept of analytic functions like power series, and differentiability methods, exponential and logarithmic functions, Cauchy integral formula, Holomorphic functions, Laurent Series, Singularity, calculus of residues evaluation of integration using contour. Vector Spaces, Subspaces. Linear independence, Basis, Dimensions.

### Unit-III

Algebra of matrices, Operation on matrices, Properties of matrix, Rank, row space, column space and inverse of a matrix. Cayley –Hamilton theorem, symmetric, skew-symmetric, Hermitian, skew-Hermitian, orthogonal, unitary matrices and their eigen values. Elementary operation, Echelon, normal and Hermite canonical forms, linear equations. Inner product, norm. Characteristic roots of real matrices, right and left characteristic vectors, Eigen values and Eigen vectors. Independence of characteristic vectors corresponding to distinct characteristic roots. Generalized inverse. Definition of a real quadratic form, Classification of quadratic forms.

### Unit-IV

Root finding using Newton-Raphson, Secant, Regula-Falsi methods and their convergence, Interpolation - Newton's formulae, Lagrange, Hermite, Numerical differentiation. Numerical integration -Trapezoidal, Simpson and Weddle rules, Gaussian quadrature formulae -Gauss-Laguerre, Transcendental Algebraic equations- Gauss elimination, Jacobi, Gauss Seidal methods and their convergence.

### Books Recommended:

1. Bartle G.R. & Sherbert D. R. (2000): Introduction to Real Analysis- John Wiley & Son Inc.
2. Royden (1988): Principles of Real Analysis - Macmillian.
3. Widder (1989): Advanced Calculus - Dover Publication.
4. W, Rudin (2013): Real and Complex Analysis, Tata Mc-Graw Hill.
5. E. M. Strein, R, Shakarchi (2003): Complex Analysis, Princeton University Press.
5. Rao, A.R. and Bhimasankaram,P .(2000): Linear Algebra, Hindustan Book Agency, New Delhi.
6. Scoule, S.R. (1982): Matrix Algebra Useful for Statistics, John Wiley & Sons.
7. Rao, C.R. (1995): Linear Statistical Inference and its Applications (Wiley Eastern).
8. Hohn, F .E. (1973): Elements of Matrix Algebra, McMillan.
9. M.K. Jain, S.R.K. Iyengar, R.K. Jain (1995): Numerical Methods for Scientific and Engineering Computation, Willey Eastern Ltd, New Delhi.

## SAMPLING METHODS

ST-C-414

4 CH

**Objective:** This paper helps to learn variety of probability and non-probability sampling methods depending on the type of the population for selecting a sample from it and estimation of various population parameters.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Sampling Methods
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Concept of population and sample, need for sampling, complete enumeration versus sampling, basic concepts in sampling, sampling and non-sampling error, Methodologies in sample surveys (questionnaires, sampling design and methods followed in field investigation) by NSSO. Subjective or purposive sampling, probability sampling or random sampling, simple random sampling with and without replacement, estimation of population mean, population proportions and their standard errors. Stratified random sampling, proportional and optimum allocation, comparison with simple random sampling for fixed sample size. Covariance and Variance Function, formation of strata and number of strata.

### Unit-II

Use of supplementary information for estimation, Ratio, product and regression methods of estimation, estimation of population mean, evaluation of Bias and Variance to the first order of approximation, comparison with simple random sampling. Systematic sampling (when

population size (N) is an integer multiple of sampling size (n)). Estimation of population mean and standard error of this estimate, comparison with simple random sampling.

### Unit-III

Equal size cluster sampling: estimators of population mean and total and their standard errors, comparison of cluster sampling with SRS in terms of intra-class correlation coefficient. Concept of multistage sampling and its application, two-stage sampling with equal number of second stage units, estimation of population mean and total. Double sampling in ratio and regression methods of estimation. Concept of sub-sampling.

### Unit-IV

Sampling with probability proportional to size (with and without replacement method): PPSWR/WOR methods (including Lahiri's scheme), Des Raj and Das estimators for  $n=2$ , Horvitz-Thomson's estimator and its properties, Murthy estimator. Non-sampling error with special reference to non-response problems.

### Books Recommended:

1. Des, Raj and Chandok, P. (1998): Sample Survey Theory (Narosa).
2. Sukhatme, P.V; Sukhatme, B.V. and Asok, C. (1984): Sampling Theory of Surveys with Applications, Indian Soc. of Ag. Stats., New Delhi.
3. Cochran, W.G. (1984): Sampling Technique (Wiley).
4. Swain, A.K.P.C. (2003): Finite Population Sampling - Theory and Methods, South Asian Publishers.
5. Mukhopadhyay, P. (1996): Inferential Problems in Survey Sampling, New Age International (P).
6. Chaudhuri, A. and R. Mukherjee (1988): Randomised response: theory and techniques, New York, Marcel Dekker Inc.
7. Singh D. and Chaudhary, F.S. (1986): Theory and Analysis of Sample Survey Designs, New Age International Publishers.

## DATA STRUCTURE & C-PROGRAMMING

ST-C- 415

2 CH

**Objective:** Data structure and algorithms introduces about the basic and elementary concepts of data structures. This paper involves learning of programming in C, with an emphasis on problem solving skills.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Data Structure & C-Programming
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Introduction to programming, History and importance of C, Components, basic structure programming, Data types, Constant and variables, declaration and assignment of variables, Operators and Expressions: Arithmetic, relational, logical, assignment, increment/decrement, operators, precedence of operators in arithmetic, relational and logical expression. Implicit

and explicit type conversions in expressions, library functions. Managing input and output operations: reading and printing formatted and unformatted data.

## **Unit-II**

Decision making and branching – if-else, nesting of if-else, else if ladder, switch, conditional-operator. Looping in C: for, nested for, while, do...while, jumps in and out of loops. Switch case statement, break and continue statement. Arrays: Declaration and initialization of one-dim and two-dim arrays. Character arrays and strings: Declaring and initializing string variables, reading and writing strings from Terminal (using scanf and printf).

## **Unit-III**

Overview of Functions, built in and user defined functions, recursive function, Function call by value and call by reference. Passing arrays to functions, Storage class of Variables. Arrays and its applications, Strings, Pointers, Structures and Unions, Data Files.

## **Unit-IV**

Time and space complexity of algorithm, Representation of stack and queue using arrays and its operations. Binary Tree representation, Binary Tree traversal methods, Binary search tree and its operations, Graph representation, Adjacency matrix, Depth first search, Breadth first search, Sequential and Binary Searching, Bubble sort, selection sort.

### **Books Recommended:**

1. Balguruswamy E.: Programming in ANSI C; Tata-McGraw Hill New Delhi
2. Byron S. Gottfried: Theory and Problems of Programming with C; Tata- McGraw Hill Edition (Schaum's Outline Series)
3. T Cormen, C Leiserson, R Rivest, C Stein, Introduction to Algorithms, PHI
4. E Horowitz, S Sahni, S Rajasekaran, Fundamentals of Computer Algorithms, Universities Press

### **LAB-I: LABORATORY USING MS-EXCEL/CALCULATOR**

#### **ST-C-416**

**2 CH**

**Objective:** Use of MS EXCEL/Calculator to draw graphs, diagrams, charts, classification and tabulation of data, frequency distribution, computation of summary statistics and analytical statistics.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Lab using of MS EXCEL/Calculator to draw graphs, diagrams, charts, classification and tabulation of data, frequency distribution, computation of summary statistics and analytical statistics.
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### **Problem Solving using MS EXCEL/CALCULATOR:**

Using of MS EXCEL/Calculator to draw graphs, diagrams, charts, classification and tabulation of data, frequency distribution, computation of summary statistics and analytical statistics and to analyze data from the following areas: ST-C-411&ST-C-414.

Examination: Practical (80%) and Record & Viva Voce (20%).

**MA/MSc. STATISTICS**  
**SEMESTER-II**

**STATISTICAL METHODS-II**

**ST-C-421**

**4 CH**

**Objective:** To identify appropriate sources of data and to perform basic demographic analyses using various techniques across populations. To develop a deeper understanding of the linear and non-linear regression model and its limitations. To develop scientific view to analyze the industrial data about specific perspective.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Methods-II
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

**Unit-I**

Basic discrete distributions- Negative Binomial, and Hypergeometric distributions, Computations of their moments, means, Variances and Beta coefficients, recurrence relations between moments, M.G.F. and C.F. Basic continuous distributions: - Lognormal, Beta, Gamma, and Weibull distributions their properties and applications. Computation of their M.G.F., C.F. and Moments.

**Unit-II**

Curve fitting- Polynomial, Orthogonal, Exponential, Logarithmic and Growth curves. Their applications.

Associations of attributes: - Contingency table and coefficients of contingency and their interpretations.

**Unit-III**

Theory of residues and their properties. Multiple and partial correlation coefficients. Their relationship and properties. Rank correlation coefficient and correlation ratio and their applications. Test of significance of multiple, partial and simple correlation coefficient.

**Unit-IV**

Order statistics- Distribution of range, smallest and biggest observations, distribution of  $r$ th order statistics and their functions, probability integral transformation,

**Books Recommended:**

1. Fundamental of Mathematical Statistics (2009)- S Chand & Sons- S.C. Gupta and V.K. Kapoor
2. Introduction to Theory and Mathematical Statistics (1988) - Wiley – V.K. Rohatgi,
3. Linear Statistical Inference and its Applications (1975) – Wiely- C.R. Rao
4. An Introduction to Theory of Statistics - Charles Griffiu Yale, G.U. and Kendall, M.G. (1953)
5. Probability and Statistics with Engineering and Computer Science Applications (2005) – Kalyani Swain, A.K.P.C.
6. John E. Freud's Mathematical Statistics with Applications(2006); (7<sup>th</sup> Edition- Pearson Education; Asia-Miller; Irwin and Miller; Marylees.
7. Applied Statistics(1999)- New Central Book Agency- Mukhapaddhyay P.

## PROBABILITY-II

ST-C-422

4CH

**Objective:** to provide students with the foundations of probabilistic and statistical analysis mostly used in varied applications in engineering and science like disease modeling, climate prediction and computer networks etc.

CO-1	Remember and understand the basic concepts/Principles of Probability-II
CO-2	Analyze the Various Concepts to understand them through case studies
CO-3	Apply the knowledge in understanding practical problems
CO-4	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Definition and properties of Lebegue integral, Monotone Convergence Theorem, Indefinite integral, Uniform integrability, Jensen's, Holder's, Cauchy and Schwartz, Luaponov inequalities.

### Unit-II

Levy inversion Theorem and Levy Continuity Theorem. Conditional expectation and properties, Three series theorem for independent random variables.

### Unit-III

Bernoulli's Theorem, Borel Theorem, Borel-Canteili Lemma. Convergence of Distribution Functions. Heily-Compactness Theorem, Heily-Bray Theorem.

### Unit-IV

Central Limit Theorem for binomial random variables. Law of Large Numbers and Law of the Iterated Logarithm.

#### BooksRecommended:

1. Bhat, B.R. (1985): Modern probability theory (Wiley).
2. Billingsley, P. (1986): Probability and measure (Wiley).
3. Feller, W. (1969): Introduction to probability theory and applications, Vol. II (Wiley)
4. Rohatgi, V.K. (1976): Introduction to theory of probability and mathematical Statistics (Wiley).
5. H.G. Tucker(1967): A graduate course in probability theory (AP)
6. Y.S. Chow and H Teicher(1979): Probability theory (Springer-Verlag)

## STATISTICAL INFERENCE - I

ST-C-423

4 CH

**Objective:** To derive suitable point estimators of the parameters of the distribution of a random variable and give a measure of their precision. To perform Test of Hypothesis as well as obtain MP, UMP tests. To derive suitable point estimators of the parameters of the distribution of a random variable and give a measure of their precision. To learn computational skills to implement various statistical inferential approaches.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Inference-I
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Parametric Point estimation: properties of estimators – Unbiasedness, Sufficiency, Completeness. Uniformly minimum variance unbiased estimators. Rao-Blackwell theorem. Cramer-Rao inequality. Fishers Information measure and its properties.

### Unit-II

Asymptotic properties of estimators- Consistency and Efficiency. Their relationship and properties. Some special classes of distribution admitting complete sufficient statistics. Methods of estimation- Method of maximum likelihood estimation and its properties. Methods of moments and its properties.

### Unit-III

Bayesian estimation, prior distributions, posterior distributions, loss function and risk function, Quadratic loss function and other common loss functions. Bayes and minimax estimators and their inter relationship. Properties of Bayes and minimax estimators.

### Unit-IV

Theory of Least squares. Gauss-Markov set-up, Normal equations, least squares estimators of linear parametric functions. Variances and Covariances of the estimators of linear parametric functions. Estimation of error variance.

### Books Recommended:

1. Linear Statistical Inference and its Applications (1973) - Wiley Eastern. Rao, C.R.
2. An Outline of Statistical Theory - Vol-II, World Press, Calcutta Goon, M.A., Gupta, M.K., and Dasgupta, B
3. Introduction to Theory of Probability and Mathematical Statistics (1970) - Wiley Rohatgi, V.K

## STOCHASTIC MODELING

ST-C-424

4CH

**Objective:** To learn and to understand stochastic processes predictive approach. To develop an ability to analyze and apply some basic stochastic processes for solving real life situations.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Stochastic Modelling
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Definition and classification of stochastic processes, Notion of stochastic processes, Markov chain, one step transition probabilities, Chapman-Kolmogorov equations, evaluation of higher step transition probabilities, Some examples such as gamblers ruin problem and one-dimensional random walk. Concept of absorption probabilities, Use of these to compute probability of winning the game of Gambler's Ruin Problem.

### Unit-II

Classification of states of a Markov chain, Periodicity, Recurrence, Basic limit theorems of Markov chain, Absorption probability, Criteria for recurrence.

### Unit-III

Introduction to birth process, birth and death process, death process. Expression for mean and variance of a birth process and, Applications of these processes.

### Unit-IV

Martingale-Elementary results. Brownian Motion-Definition, Continuity of paths. Branching Processes-Definition, generating function relation.

### Books Recommended:

1. Karlin, S and Taylor, H.M (1975): A First Course in Stochastic Processes. Academic Press.
2. Bhatt, B.R. (2000): Stochastic Models: Analysis and application, New Age International Publication.
3. Feller, W. (1968): Introduction to Probability and its Applications, Vol. 1, Wiley Eastern.
4. Hoel, P.G., Port, S.C. and Stone, C.J. (1972): Introduction to Stochastic Processes, Houghton Mifflin & Co.
5. Medhi, J. (1982): Stochastic Processes, Wiley Eastern.
6. Parzen, E. (1962): Stochastic Processes, Holden-Day.

## APPLIED STATISTICS

ST-C-425

2CH

**Objective:** To learn and develop scientific view to understand the time series data and its analysis.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Applied Statistics
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Index Numbers: Price relatives and quantity or volume relatives, Link and chain relatives composition of index numbers; Laspeyre's, Paasches', Marshal Edgeworth and Fisher index numbers; chain base index number, tests for index number, Construction of index numbers of wholesale and consumer prices.

### Unit- II

Statistical data:- types, sources and their methods of collection, Statistical system in India and Orissa; Functions of NSSO and CSO. Vital statistics- Rates, Ratios of Births and Deaths, Their merits, demerits, and relative comparisons.

### Unit-III

Theory and analysis of consumer's demand, law of demand, price elasticity of demand, estimation of demand curves, forms of demand functions, Engel's curve, income elasticity of demand. Analysis of income and allied distributions: Pareto distribution, graphical test, fitting of Pareto law, illustration, lognormal distribution and properties, Lorenz curve, Gini's coefficient.

### Unit-IV

Time Series Analysis: Economic time series, different components, illustration, additive and multiplicative models, determination of trend, seasonal and cyclical fluctuations. Time-series as discrete parameter stochastic process, auto covariance and autocorrelation functions and their properties. Brief study of study of the stationary processes: (1) moving average (MA), (2) auto regressive (AR).

### Books Recommended:

1. Mukhopadhyay P. (1999): Applied Statistics, New Central Book Agency Pvt. Ltd., Calcutta.
2. Goon A.M., Gupta M.K. and Dasgupta B. (1986): Fundamentals of Statistics, Vol.II, World Press, Calcutta.
3. Croxton F.E. and Cowden, D.F.: Applied General Statistics.
4. Asthana B.N. and Srivastava S.S.: Applied Statistics in India.

## LAB-II: LABORATORY USING SPSS/STATA

**ST-C- 426**

**2 CH**

**Objective:** To learn statistical techniques and their implementation using comprehensive SPSS/STATA software packages.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Laboratory on Computer using SPSS/STATA
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### **Problem Solving using SPSS/ STATA:**

Using of SPSS/ STATA software packages to gain the knowledge of software package and applications of Software for data analysis in the areas of ST-C-421 and ST-C-425.

Examination: Practical (80%) and Record & Viva Voce (20%).

## MA/MSc. STATISTICS

### SEMESTER-III

#### STATISTICAL INFERENCE-II

ST-C-511

4CH

**Objective:** To learn computational skills to implement various statistical inferential approaches. To learn types of errors, non-parametric tests. ·

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Inference-II
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### Unit-I

Basic concept of test of Hypothesis, Neyman-Pearson Lemma. Most powerful (MP) test Application of Neyman-Pearson Lemma for the test of simple hypothesis, Generalized Neyman-Pearson Lemma Uniformly most powerful (UMP) test. UMPU tests. Similar Regions, Application of Neyman- Pearson Lemma for the test of composite hypothesis.

#### Unit-II

Non-parametric tests: - One sample problem, Goodness of fit, Kolmogorov test, sign test, Run test, U statistic and its properties. Two sample problem- Kolmogorov-Smirnov test and its consistency, run test, Location problem, Wilcoxon-Mann Whitney test, median test, and their asymptotic normality.

#### Unit-III

Likelihood Ratio Test (LRT), Consistency of the test, Asymptotic distribution of LRT, Application of Likelihood Ratio Test (LRT), Comparison between Likelihood Ratio Test and Neyman- Pearson test procedure.

#### Unit-IV

Sequential probability ratio test (SPRT), procedures, Properties of SPRT, Fundamental identity of SPRT, Wald 's fundamental Identity. OC and ASN functions.

#### Books Recommended:

1. An Outline of Statistical Theory - Vol-II, World Press, Calcutta Goon, M.A., Gupta, M.K., and Dasgupta, B
2. Non-parametric Methods in Statistics - Second Edition, Marcel Dekker Gibbons, J
3. Linear Statistical Inference and its Applications (1973) - Wiley Eastern. Rao, CR

## MULTIVARIATE ANALYSIS

ST-C-512

4CH

**Objective:** To learn and develop scientific view to deal with multidimensional datasets and its uses in the analysis of research data. · To understand the extensions of univariate techniques to multivariate frameworks and learn to apply dimension reduction techniques used in the data analysis.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Multivariate Analysis
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

**Unit-I:** Multivariate normal distribution and its properties- Marginal and conditional distributions – characteristic function; Random sampling from multivariate normal distribution; Maximum likelihood estimators of parameters, distribution of sample mean vector.

**Unit-II:** Wishart matrix – its distribution and properties; distribution of sample generalized variance; Partial and Multiple correlation coefficients in the multivariate setup, Sampling distributions of Total and Multiple Correlation in the null case.

**Unit-III:** Hotelling's  $T^2$  and its sampling distribution- –Properties and Uses of  $T^2$  statistic; Mahalanobis  $D^2$  statistic and its distribution-- relation between  $T^2$  and  $D^2$ ; application of  $T^2$  statistic in test on mean vector for one and more multivariate normal population and also on equality of components of a mean vector in multivariate normal population.

**Unit-IV:** Classification problem: Standards of good classification, Classification into one of two populations (known and unknown dispersion matrix), procedure of classification based on multivariate normal distributions; Principal components, dimension reduction, canonical variates and canonical correlation— definition, use, estimation and computation.

### Books Recommended:

1. Anderson, T.W. (1984): Introduction to Multivariate Analysis, Wiley.
2. Kshirsagar, A.M. (1983): Multivariate Analysis, Marcel Dekkar.
3. Morrison, D.F. (1990): Multivariate statistical methods, McGraw Hill.
4. Rao, C .R. (1995): Linear statistical Inference and its Applications (Wiley).
5. Johnson, R.A. and Wichern, D.W. (1988): Applied Multivariate Statistical Analysis (Prentice Hall).

## OPTIMIZATION TECHNIQUE-I

### ST-C-5134CH

**Objective:** To develop the optimization techniques that will be useful in the personal and professional life.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Optimization Technique-I
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### Unit-I

Introduction to LPP, Mathematical formulation, Standard form and canonical form, Graphical solution, Simplex Method including Big-M and two phase method, Degeneracy, Solution of simultaneous equations and inversion matrix by simplex method.

#### Unit-II

Duality in Linear Programming, Duality Theorems, Dual simplex method with justifications, Transportation and Assignment algorithms.

#### Unit-III

Introduction to sensitivity analysis, variation in cost and requirement vectors, coefficient matrix and applications, Simulation, Parametric programming and revised simplex methods.

#### Unit-IV

Game Theory, two persons zero sum game, Maxmin Minimax principle, Mixed strategy, Graphical solutions, Dominance Property, Arithmetic Method and general solution.

#### Books Recommended:

1. Kambo., N.S. (1991): Mathematical Programming Tech., Affiliated
2. East-West press. Hadley, G. (1987): Linear Programming
3. Taha H. A. (1992): Operations Research, 5th Ed. (McMillan)

## SPECIAL PAPER-I (ANY ONE)

ST-E- 514

4CH

### A. ECONOMETRICS

**Objective:** Econometrics is the use of statistical and mathematical models to develop theories or test existing hypotheses in economics and to forecast future trends from historical data. It subjects real-world data to statistical trials and then compares the results against the theory being tested.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Econometrics
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### Unit-I

The Two-Variable Regression Model: Some basic ideas, the problem of estimation, CNLRM, Interval Estimation and Hypothesis testing.

#### Unit-II

Autocorrelation: Nature and Scope of Autocorrelation, OLS estimation, BLUE estimation, Consequences, Detecting Autocorrelation, Remedial Measures of Autocorrelation, Multicollinearity: Nature and Scope of Multicollinearity, Estimation in the presence of Multicollinearity, Consequences of Multicollinearity, Detecting Multicollinearity, Remedial Measures of Multicollinearity.

#### Unit-III

Heteroscedasticity: Nature and Scope of Heteroscedasticity, OLS estimation, GLS estimation, Consequences, Detection of Heteroscedasticity, Remedial Measures of Heteroscedasticity.

#### Unit-IV

Simultaneous equations models: The nature of Simultaneous equations models, Simultaneous equations bias, identification problem, restrictions on structural parameters, rank and order conditions. Estimation in simultaneous equations model, recursive systems, 2 SLS estimators.

#### Book Recommended:

1. Apte, P.G. (1990): Text books of Econometrics, Tata McGraw Hill.
2. Cramer, J.S. (1971): Empirical Econometrics, North Holland.
3. Gujarathi, D. (1979): Basic Econometrics, McGraw Hill.
4. Intrulligator, M.D. (1980): Econometric models—Techniques and applications, Prentice Hall of India.
5. Johnston, J. (1984): Econometric methods. Third edition, McGraw Hill.
6. Klein, L.R. (1962): An introduction to Econometrics, Prentice Hall of India.
7. Koutsoyiannis, A. (1979): Theory of Econometrics, Macmillan Press.
8. Malinvaud, E. (1966): Statistical methods of Econometrics, North Holland.
9. Srivastava, V.K. and Giles D.A.E. (1987): Seemingly unrelated regression equations models, Maicel Dekker.
10. Theil, H. (1982): Introduction to the theory and practice of Econometrics, John Wiley.
11. Walters, A. (1970): An introduction to Econometrics, Macmillan & Co.

## B. BIOSTATISTICS& EPIDEMIOLOGY

**Objective:** An exploration of the core principles and practice of epidemiology and biostatistics in assessing and responding to population health need. Students will learn the skills necessary to analyze and interpret data, disseminate information, and critically appraise quantitative literature.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Biostatistics & Epidemiology
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Measuring the occurrence of disease: Measures of morbidity - prevalence and incidence rate, association between prevalence and incidence, uses of prevalence and incidence, problems with incidence and prevalence measurements; Clinical agreement: kappa statistics, Mantel-Haenszel test; intra-class correlation; Surveillance.

Assessing the validity and reliability of diagnostic and screening test: Validity of screening test – sensitivity, specificity, positive predictive value and negative predictive value; Reliability; Relationship between validity and reliability; ROC curve and its applications; Overall accuracy

### Unit-II

Issues in epidemiology: Association; causation; causal inference; Errors and bias; Confounding; Controlling confounding; Measurement of interactions; Generalizability

Estimating risk: Estimating association – absolute risk, relative risk, odds ratio; Estimating potential for prevention – attributable risk; comparison of relative risk and attributable risk; Odds ratios for retrospective studies; Odds ratios approximating the prospective RR; Exact inference for odds ratio analysis of matched case-control data. Statistical process control: special and common causes of variation, Shewhart, CUSUM and EWMA charts

### Unit-III

Introduction: Definition and objectives of epidemiology; Epidemiology and clinical practice; The epidemiologic approach; Infectious disease epidemiology, occupational epidemiology, disaster epidemiology. The dynamics of disease transmission: Modes of transmission; epidemic, endemic and pandemic; Disease outbreak; Determinants of disease outbreak; Herd immunity; incubation period; outbreak investigation; epidemiological modeling

### Unit-IV

Identifying the roles of genetic and environmental factors in disease causation: Association with known genetic diseases; Age at onset; Family studies; Interaction of genetic and environmental factors

Epidemiology and public policy: Epidemiology and prevention; Population versus high- risk approaches to prevention; epidemiology and clinical medicine; Risk assessment

**Books Recommended:**

1. Altman D G: Practical Statistics for Medical Research, London: Chapman and Hall, 2006.
2. Rosner B: Fundamentals of Biostatistics, ed. 6, 2006.
3. Bonita R, Beaglehole R, Kjellstrom T: Basic Epidemiology, ed. 2. World Health Organization, 2006.
4. Gordis L: Epidemiology, ed. 3. Philadelphia, 2004.
5. Baker, D. et al.: Environmental Epidemiology: A Text Book on Study Methods and Public Health Applications, WHO/SDE/99.7, 1999.
6. Dunn G, Everitt B: Clinical Biostatistics: An Introduction to Evidence-based Medicine. Edward Arnold, 1995.

**C. STATISTICAL DECISION THEORY**

**Objective:** After completing this course, students will be able to describe the bias that affect the unaided decision making process and be capable of formulating a decision problem in terms of a matrix of alternatives, preferences and consequences, as well as defining, collecting and synthesizing the data required to make the decision. Graphical models are introduced for representing complex probability and decision models by specifying modular components. An overview is given of expected utility and decision theory. Students apply the methods to a variety of practical problems.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Decision Theory
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

**Unit I**

Decision problem and 2-person game, utility theory, loss functions, expected loss, decision rules (non-randomized and randomized), decision principles. Concept of admissibility and completeness, Bayes rules, admissibility of Bayes rules.

**Unit II**

Supporting and separating hyperplane theorems, minimax theorem of for finite parameter space, minimax estimators of Normal and Poisson means, admissibility of minimax rules.

**Unit III**

Supporting and separating hyperplane theorems, minimax theorem of for finite parameter space, minimax estimators of Normal and Poisson means, admissibility of minimax rules.

**Unit IV**

Supporting and separating hyperplane theorems, minimax theorem of for finite parameter space, minimax estimators of Normal and Poisson means, admissibility of minimax rules.

**Books Recommended:**

1. Berger, J.O. (1985): Statistical Decision Theory and Bayesian Analysis, 2nd Edition. SpringerVerlag.

2. Ferguson, T.S. (1967): Mathematical Statistics – A Decision Theoretic Approach, Academic Pres.
3. Rohatgi, V.K. (1988): An Introduction to Probability and Mathematical Statistics, Wiley Eastern, New Delhi.
4. Rao, C.R. (1973): Linear Statistical Inference and its Applications, Wiley Eastern.
5. Bernardo, J.M. and Smith, A.F.M. Bayesian Theory, John Wiley and Sons.
6. Robert, C.P.: The Bayesian Choice: A Decision Theoretic Motivation, Springer.

## D. PATTERN RECOGNITION

**Objective:** To develop the mathematical tools required for the pattern recognition.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Pattern Recognition
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Basics of Probability, Random Processes and Linear Algebra: Probability: independence of events, conditional and joint probability, Bayes' theorem; Random Processes: Stationary and nonstationary processes, Expectation, Autocorrelation, Cross-Correlation, spectra; Linear Algebra: Inner product, outer product, inverses, eigen values, eigen vectors;

### Unit-II

Bayes Decision Theory • Bayes Decision Theory: Minimum-error-rate classification, Classifiers, Discriminant functions, Decision surfaces, Normal density and discriminant functions, discrete features • Parameter Estimation Methods: Maximum-Likelihood estimation: Gaussian case; Maximum a Posteriori estimation; Bayesian estimation: Gaussian case

### Unit-III

Unsupervised learning and clustering: Criterion functions for clustering; Algorithms for clustering: K-Means, Hierarchical and other methods; Cluster validation; Gaussian mixture models; Expectation-Maximization method for parameter estimation; Maximum entropy estimation • Sequential Pattern Recognition: Hidden Markov Models (HMMs); Discrete HMMs; Continuous HMMs • Nonparametric techniques for density estimation: Parzen-window method; K-Nearest Neighbour method •

### Unit-IV

Dimensionality reduction: Fisher discriminant analysis; Principal component analysis; Factor Analysis • Linear discriminant functions: Gradient descent procedures; Perceptron; Support vector machines • Non-metric methods for pattern classification: Non-numeric data or nominal data; Decision trees: CART

### Books Recommended:

1. R.O.Duda, P.E.Hart and D.G.Stork, Pattern Classification, John Wiley, 2001
2. S.Theodoridis and K.Koutroumbas, Pattern Recognition, 4th Ed., Academic Press, 2009
3. C.M.Bishop, Pattern Recognition and Machine Learning, Springer, 2006

## TIME SERIES AND FORECASTING

ST-C-515

2 CH

**Objective:** To learn and develop scientific view to understand the time series data and its analysis. To learn stationary and non-stationary, and seasonal and nonseasonal time series models. Learn to estimate model parameters and compare different models developed for the same dataset in terms of their estimation and prediction accuracy.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Time Series and Forecasting
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Exploratory Time Series Analysis: Forecasting trend and seasonality based on smoothing. Outliers, procedure for detecting outliers; Stationary Stochastic models: weak and strong stationarity, Deseasonalising and detrending an observed time series, Auto-covariance, autocorrelation function (ACF), partial autocorrelation function (PACF) and their properties

### Unit II

Models for Time Series: Time series data, Trend, seasonality, cycles and residuals, Stationary, White noise processes, Autoregressive (AR), Moving Average (MA), Autoregressive and Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) processes, Choice of AR and MA periods

### Unit III

Spectral analysis and decomposition: Spectral analysis of weakly stationary process, Periodogram and Correlogram analysis, Spectral decomposition of weakly AR process and representation as a one-sided MA process - necessary and sufficient conditions, implication in prediction problems.

### Unit IV

Modeling Seasonal Time Series: seasonal ARIMA models, estimation and forecasting, Fitting ARIMA models with Box-Jenkins procedure, Identification, Estimation, Verification, Test for white noise, Forecasting with ARMA models.

#### Books Recommended:

1. Box, GEP and Jenkins, G.M. (1976): Time Series Analysis - Forecasting and Control, Holdenday, Sanfransico.
2. Anderson, T.W. (1971): The Statistical Analysis of Time Series. Wiley.
3. Brockwell, P.J. and Davis, RA : Time Series: Theory and Methods, 2nd Ed., Springer-Verlas.
4. Nicholas T. Thomopoulos, 1980, Applied Forecasting Methods, Prentice Hall
5. Box GEP, Jenkins G M and Reinsel GC (2004): Time Series Analysis-Forecasting and Control, Pearson Education.
6. Brockwell PJ and Davis RA (2002): Introduction to Time Series and Forecasting, Springer.
7. Montgomery D C and Johnson L A (1977): Forecasting and Time Series analysis, McGraw Hill.

### LAB- III- LABORATORY USING R/PYTHON

**ST-C-516**

**2CH**

**Objective:** To learn statistical techniques and their implementation using comprehensive R/PYTHON software packages.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of R / Python
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

**Problem Solving using R / PYTHON:**

Using of R/ PYTHON software packages to gain the knowledge of software package and applications of Software for data analysis in the areas of ST-C-512 and ST-E-514()\*&ST-C-515.

Examination: Practical (80%) and Record & Viva Voce (20%).

## MA/MSc. STATISTICS

### SEMESTER- IV

#### DESIGN AND ANALYSIS OF EXPERIMENTS

ST-C-521

4 CH

**Objective:** To learn the basic principles in the design of simple experiments. · To learn different tests for comparing pairs of treatment means, ANCOVA, factorial experiments, fractional factorial experiments, confounding, BIBD, PBIBD with solving real life examples. · To learn the applications of different designs in agriculture.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Design and Analysis of Experiments
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### **Unit-I**

Linear estimation. Theory of least squares. Gauss Markov Theorem. Normal equations, least square estimators of linear parametric functions. Variance and covariance of estimators.

Analysis of variance - fixed, random and mixed models. Analysis of variance. One-way and two-way classification with unequal and equal number of observation per cell. Basic principles of designs - CRD, RBD and Latin square and their analysis - missing plot technique. Connectedness, balance, orthogonality.

#### **Unit-II**

Factorial experiments –  $2^n$  and  $3^2$ ,  $3^3$  Presentation of main effects and interactions. Analysis. Asymmetrical factorial experiments.

#### **Unit-III**

Confounding - Total confounding of  $2^n$  design in  $2^p$  blocks, partial confounding in  $2^p$  blocks. Fractional factorial experiments and their analysis, Total and partial confounding in  $3^2$  and  $3^2$  designs. Split plot designs.

#### **Unit-IV**

Incomplete block designs - BIBD and Lattice designs - Construction and analysis. Concept of rotatable design. Control composite design. Response surface methodology.

#### **Books Recommended:**

1. Kshirasagar, A.M. (1983): Linear Models, Marcel Dekkar.
2. John, P. W.M. (1971): Linear Models, Wiley.
3. Montgomery, D.C. (2001). Design and Analysis of Experiments, Wiley.
4. Das, M.N. and Giri, N.C. ( ): Design of Experiments.

## SPECIAL PAPER-II (ANY ONE)

ST-E- 5224 CH

### E. STATISTICAL QUALITY CONTROL AND RELIABILITY

**Objective:** To learn the statistical quality control techniques used in industries such as control charts, acceptance sampling plans etc. To learn some advanced control charts, capability indices and the concept of six-sigma. To learn the reliability theory and to distinguish censored and uncensored data. To visualize and communicate time-to event data, to fit and interpret failure time model.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Statistical Quality Control & Reliability
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### Unit-I

Statistical process and product control: Quality of a product, need for quality control, basic concept of process control, process capability and product control, general theory of control charts, causes of variation in quality, control limits, sub grouping summary of out of control criteria, charts for attributes p chart, np chart, c-chart; charts for variables:  $(\bar{X}, R)$ ,  $(\bar{X}, \sigma)$  charts.

#### Unit-II

Acceptance sampling plans: classification and general properties of sampling plans by variables. Definitions of OC and ASN functions, AOQL and ATI. Single and double sampling plans and their properties, sequential sampling plan, plans for inspection by variables for one-sided and two-sided specification.

#### Unit-III

Reliability- Hazards rate, IFR and DFR. Failure time distribution, Reliability concepts and measures components and systems, coherent systems, reliability of coherent systems, cuts and paths, life distributions, survival functions, Hazard rate, Hazard function and residual life time, survival function of residual life time, mean residual life function.

#### Unit -IV

Common life distributions, Exponential, Weibull, Gamma, Pareto, Rayleigh, Log-Normal etc. Notions of aging – IFR, IFRA, NBU, DMRL and NBUE classes and their duals. Loss of memory property of the exponential distribution. Reliability estimation based on failure times in variously censored life-tests and in tests with replacement of failed items, likelihood functions based on these sampling schemes.

#### Books Recommended:

1. RE. Ballo and F. Proschan: Statistical Theory of Reliability and Life Testing. Holt, Reinhart and Wonston
2. RC. Miller: Survival Analysis. John Wiley.
3. J.L. Bain: Statistical Analysis of Reliability and Life Testing Models. Morcel Bekker.

4. Gupta and Kapoor: Applied Statistics.
5. E.L. Groant and RS. Leavenworth: Statistical Quality Control. Sixth Edn., McGraw Hill Publication.

## F. BAYESIAN INFERENCE

**Objective:** To introduce students to Bayesian inference and decision making and to provide practical experience in applications from information technology and engineering. Students will learn the fundamentals of the Bayesian theory of inference, including probability as a representation for degrees of belief, the likelihood principle, the use of Bayes Rule to revise beliefs based on evidence, conjugate prior distributions for common statistical models, and methods for approximating the posterior distribution. Modern Bayesian computational methods are introduced.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Bayesian Inference
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Subjective probability, its existence and interpretation. Prior distribution, subjective determination of prior distribution. Improper priors, non-informative (default) priors, invariant priors. Conjugate prior families, construction of conjugate families using sufficient statistics of fixed dimension, hierarchical priors and partial exchangeability. Parametric Empirical Bayes.

### Unit-II

Bayesian inference: Bayes sufficiency, summary through posterior, predictive inference.

### Unit-III

Bayesian decision theory: Bayes solutions for practical decision problems. Point estimation, credible sets, testing of hypotheses. Comparison with classical procedures.

### Unit-IV

Ideas on Bayesian robustness. Asymptotic expansion for the posterior density. Bayesian calculation, Monte-Carlo Integration and Markov chain Monte Carlo techniques (without proof).

### Books Recommended:

1. Berger, J.O.: Statistical Decision Theory and Bayesian Analysis, Springer Verlag.
2. Robert, C.P. and Casella, G.: Monte Carlo Statistical Methods, Springer Verlag.
3. Leonard, T. and Hsu, J.S.J.: Bayesian Methods, Cambridge University Press.
4. Bernardo, J.M. and Smith, A.F.M.: Bayesian Theory, John Wiley and Sons.
5. Robert, C.P.: The Bayesian Choice: A Decision Theoretic Motivation, Springer.
6. Gemerman, D.: Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Chapman Hall.
7. Box, G.P. and Tiao, G.C.: Bayesian Inference in Statistical Analysis, Addison-Wesley.

## G. LINEAR MODELS AND REGRESSION ANALYSIS

**Objective:** This course is primarily about data analysis and developing a deeper understanding of the generalized linear model. The focus is on practice, and this focus is reflected in the choice of texts and in the emphasis on applied coursework. While this course deals to some degree with the generalized linear model on a mathematical and theoretical level, its main focus is practical, the ability to use the techniques when faced with the need in practical research. Consequently the learning method combines lectures and reading with hands-on statistical programming exercises using real datasets.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Linear Models and Regression Analysis
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Gauss-Markov linear models, estimable functions, error and estimation space, normal equations and least square estimators, estimation of error variance, estimation with correlated observations, properties of least square estimators, generalized inverse of matrix and solution of normal equations, variances and covariances of least square estimators

### Unit-II

One way and two-way classifications, fixed, random and mixed effects models. Analysis of variance (two-way classification only), Multiple comparison tests due to Tukey, Scheffe and Student-Newmann-Karl.

### Unit-III

Simple linear regression, multiple, regression, fit of polynomials and use of orthogonal polynomials. Residuals and their plots as tests for departure from assumptions such as fitness of the model, normality, homogeneity of variances and detection of outliers. Remedies. Multi co-linearity, ridge regression, sub-set selection of explanatory variables, Mallows Cp Statistics

### Unit-IV

One way and two-way classifications, fixed, random and mixed effects models. Analysis of variance (two-way classification only), multiple comparison tests due to Tukey, Scheffe and Student-Newmann-Keul-Duncan.

#### Book Recommended:

1. Goon, A.M., Gupta, M.K. and Das Gupta, B. (1967): An Outline of Statistical Theory, Vol.
2. The World Press Pvt. Ltd., Calcutta. 2. Rao, C.R. (1973); Linear Statistical Inference and its Application, Wiley Eastern.
3. Graybill, I.A. (1961): An Introduction to Linear Statistical Models, Vol. 1, McGraw Hill Book Co. Inc.
4. Draper, N.R. and Smith H. (1998); Applied Regression Analysis, 3rd Ed. Wiley.
5. Weisberg, S. (1985): Applied Linear Regression, Wiley.
6. Cook, R.D. and Weisberg, S. (1982): Residuals and Inference in Regression, Chapman and Hall.
- 7.

## H. DEMOGRAPHY

**Objective:** To describe current population trends in terms of fertility, mortality and population growth and the concepts stable population.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Demography
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Sources of demographic data, Coverage and content errors in demographic data, Chandrasekharan Deming formula. Adjustment of age data, Whipples, Mayers and UN indices. Population projection methods: Component & Growth Models, Leslie Matrix, Population distribution: Lorenz curve and Gini concentration ratio, Population pyramid.

### Unit-II

Measures of fertility (period and cohort), Coales fertility index, Measures of reproduction, Calculation of PPR, Fertility models, Birth Intervals. Nuptiality rate, Net Nuptiality table, Proportion Single and Singulate. Mean age at marriage, Hajnal's method of estimating SMAM, Mean duration of fertile union.

### Unit-III

Measures of mortality, comparative mortality index, Lexis Diagram and IMR, life table functions, Construction of Reed Merrell, Greville life table, UN and Coale- Demeny model life tables, multiple decrement life table, measures of morbidity.

### Unit-IV

Measures of internal migration & international migration methods of estimation, Migration models. Stationary and stable population models, simplified example of stable population, Lotka's demonstration of conditions producing a stable population, the equations characterizing a stable Population, Identification of the intrinsic growth rate.

### Books Recommended:

1. K.B. Pathak, F. Ram, 1998, Techniques of Demographic Analysis, 2nd Edition. Himalaya Publishing House, New Delhi.
2. D.J. Bogue, 1969, Principles of Demography. John Wiley & Sons Inc.
3. C.L. Chiang, 1968, Introduction of the Stochastic Processes in Biostatistics. John Wiley New York.
4. Nathan Keyfitz, 1968, Introduction to the Mathematics of Population. Addison Wesley Publishing.
5. R. Ram Kumar, 1986, Technical Demography. Wiley Eastern Limited, New Delhi.
6. M. Spiegelman, 1955, Introduction to Demography. Harvard University Press.
7. P.R. Cox, 1950, Demography. Cambridge University Press.
8. Nathan Keyfitz, H Caswell, 2005, Applied Mathematical Demography, 3rd Edition. Springer-Verlag.
9. Heuveline, M. Guillaat, 2001, Demography: Measuring and Modeling Population Processes, Oxford, UK: Blackwell Publishers.

10. NFHS I-V International Institute for Population Sciences, Mumbai.
11. The Methods and Materials Of Demography Edited By Jacob S. Siegel David A. Swanson, Elsevier Academic Press UK

**SPECIAL PAPER-II (ANY ONE)**

**ST-E-523**

**4CH**

### I. OPTIMIZATION-II

**Objective:** To learn the mathematical formulation of complex decision-making problems and arrives at optimal or near-optimal solutions using different techniques of operations research.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Optimization-II
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

#### Unit-I

Markov process, transition matrix, transition diagram, construction of transition matrix, n-step transition prob. Equilibrium condition, Markov analysis algorithm, Network Scheduling by PERT/CPM, Sequencing Problem, Integer Programming.

#### Unit-II

Dynamic Programming, Inventory decision, cost associated with inventory, factors affecting inventories, EOQ, deterministic inventories with no shortage and shortage, inventory with uncertain demand, system of inventory control, Probabilistic inventory problems.

#### Unit-III

Decision Analysis, Goal Programming, Queuing system, operating characteristic, probability distribution, Classification of queuing models, transient and steady state, Poisson and non-Poisson Queuing System, Cost model in queuing, queuing control, queuing Theory and Inventory control.

#### Unit-IV

Formulation of Non-Linear programming, Constrained optimization with equality constraint and inequality constraint, saddle point and NLLP. Graphical solution, Kuhn- Tucker conditions, Quadratic Programming, Wolfe's and Beales' method.

#### Books Recommended:

1. Kambo., N.S. (1991): Mathematical Programming Tech., Affiliated
2. East-West press. Hadley, G. (1987): Linear Programming
3. Taha H.A. (1992): Operations Research, 5th Ed. (McMillan)
4. Operations Research: Kanti Swarup; Gupta & Mohan (S. Chand)

## J. ACTUARIAL STATISTICS

**Objective:** To strengthen your data visualization and analysis capabilities to help you sort an issue faster. This involves equipping the students with knowledge of statistical distributions, methods to summarize data, the principles of statistical inference, and regression models.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Actuarial Statistics
<b>CO-2</b>	Analyse the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Utility theory, insurance and utility theory, models for individual claims and their sums, survival function, curtate future lifetime, force of mortality. Life table and its relation with survival function, examples, assumptions for fractional ages, some analytical laws of mortality, select and ultimate tables. Multiple life functions, joint life and last survivor status, insurance and annuity benefits through multiple life functions evaluation for special mortality laws.

### Unit-II

Multiple decrement models, deterministic and random survivorship groups, associated single decrement tables, central rates of multiple decrements, net single premiums and their numerical evaluations. Distribution of aggregate claims, compound Poisson distribution and its applications. Distribution of aggregate claims, compound Poisson distribution and its applications.

### Unit-III

Principles of compound interest: Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding.

Life insurance: Insurance payable at the moment of death and at the end of the year of death-level benefit insurance, endowment insurance, deferred insurance and varying benefit insurance, recursions, commutation functions.

### Unit-IV

Life annuities: Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities. Net premiums: Continuous and discrete premiums, true monthly payment premiums, apportionable premiums, commutation functions, accumulation type benefits.

A brief outline of payment premiums and net premiums

### Book Recommended:

1. N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones and C.J. Nesbitt (1966), 'Actuarial Mathematics,' Society of Actuarial, Ithaca, Illinois, U.S.A., Second Edition (1997).
2. Spurgeon, E.T. (1972): Life Contingencies, Cambridge University Press.

3. Neill, A. (1977): Life Contingencies, Heinemann.

## K. DATA WAREHOUSING AND DATA MINING

**Objective:** This course gives an introduction to methods and theory for development of data warehouses and data analysis using data mining. Data quality and methods and techniques for preprocessing of data. Modeling and design of data warehouses. Algorithms for classification, clustering and association rule analysis. Practical use of software for data analysis.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Data Warehousing and Mining
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Overview: Data warehousing, OLAP and Data mining, classification of data mining techniques, Discovery and analysis of patterns, trends, and deviations.

### Unit-II

Data mining models: decision trees, genetic algorithms.

### Unit-III

Neural nets, clustering, Enabling data mining through data warehouse.

### Unit-IV

Data marts, multidimensional databases, Data mining applications.

### Books Recommended:

1. Pujari, A.K.: Data Mining, UBH, Bangalore
2. Data Mining-Concepts and Techniques- Jiawei Han, Micheline Kamber, Morgan Kaufmann Publishers, Elsevier, 2 Edition, 2006.
3. Introduction to Data Mining, Pang-Ning Tan, Vipin Kumar, Michael Steinbach, Pearson Education.
4. Data Warehousing Fundamentals, Punalraj Ponnaiah, Wiley Student Edition.
5. The Data Warehouse Life Cycle Toolkit- Ralph Kimball, Wiley Student Edition.
6. Data Mining, Vikaram Pudi, P Radha Krishna, Oxford University.

## L. ADVANCED STOCHASTIC PROCESS

**Objective:** to modern stochastic processes, including Brownian motion, continuous-time martingales, stochastic integration and Ito's calculus, Markov processes, stochastic differential equations, point processes and their applications. The course will include some applications but will emphasise setting up a solid theoretical foundation for the subject.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of Advanced Stochastic Process
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Introduction. Distribution of a Stochastic Process. Conditional Expectation. Stopping Times

### Unit-II

Countable State Markov Processes, Birth-Death Processes. The Markov Property and Strong Markov Property, Recurrence. Probability Transition Functions and Infinitesimal Generator. Constructive Definition. Kolmogorov's Differential Equations. Invariant Measures, Stationarity and Limit Theorems. Reversibility. Resolvents and Potential Theory.

### Unit-III

Martingales and Supermartingales, Random Walks. The Martingale Property. Maximal Inequalities. Convergence of Supermartingales and Uniformly Integrable Martingales. Reversed Martingales. Continuous Time Martingales. Supermartingales and Stopping Times Random Walks and Wald's Identity. Martingales and Markov Processes, More Potential Theory.

### Unit-IV

Brownian Motion and Diffusion Processes. Diffusions. Brownian Motion. Infinitesimal Generator and Kolmogorov's Differential Equations. Dynkin's Formula. First Passage Times, Scale and Speed Functions. Recurrence and Transience . Connections with Brownian Motion, Random Walks and Birth-Death Processes

### Book Recommended:

1. G.R. Grimmett and D.R. Stirzaker, Probability and Random Processes, 3rd ed., Oxford Univ. Press.
2. G.R. Grimmett and D.R. Stirzaker, One Thousand Exercises in Probability, Oxford Univ. Press. References: (on reserve in Evans Library)
3. R.F. Bass, Stochastic Processes, Cambridge Univ. Press.
4. R.N. Bhattacharya and E.C. Waymire, Stochastic Processes with Applications, Wiley. 5. E. Çinlar, Introduction to Stochastic Processes, Prentice-Hall.
5. M. Kijima, Markov Processes for Stochastic Modeling, Chapman & Hall.
6. T. Mikosch, Elementary Stochastic Calculus, World Scientific.
7. S.I. Resnick, A Probability Path, Birkhäuser.
8. S.I. Resnick, Adventures in Stochastic Processes, Birkhäuser.
9. D. Williams, Probability with Martingales, Cambridge Univ. Press

## PROJECT AND VIVA VOCE

ST-C- 524

4 CH

The Project Work will be taken in the final semester and spread over the whole semester. A project may be undertaken by a group of students. However, the project report shall be submitted by each member of the group separately. A project report shall clearly state the problem addressed, the methodology adopted, the assumptions and the hypotheses formulated, any previous reference to the study undertaken, statistical analyses performed and the broad conclusion drawn. The evaluation will be based on the total 100 marks assigned to the project, 60 marks will be assigned on the evaluation of the project work and 40 marks will be assigned jointly by the examiners on the oral presentation and viva – voce.

## RESEARCH METHODOLOGY AND ETHICS

ST-C- 525

2 CH

**Objective:** To understand the concept of research methodology and ethics for data management and as a cutting edge technology tool. To enable to identify data sources, processing and imparting knowledge tools to analyze sets of data to gain useful research understanding.

CO-1	Remember and understand the basic concepts/Principles of Research Methodology and Ethics
CO-2	Analyze the Various Concepts to understand them through case studies
CO-3	Apply the knowledge in understanding practical problems
CO-4	Execute/Create the Project or field assignment as per the knowledge gained in the course

### Unit-I

Science and Research, Verification Vs. Falsification, Objectivity: Facts, theory and concepts, Basic Steps for doing Research, Formulation of Research Problem, Scientific method Vs Arbitrary Method, Deductive and Inductive Reasoning, Error Analysis and Accuracy.

### Unit-II

Descriptive Statistics, Probability, Random Variables, Sampling distribution and Probability Distribution, Hypothesis Testing, Regression Analysis, Multivariate Analysis.

### Unit-III

Definition and type of research. Steps in research process: Conceptual phase, empirical phase, analytical phase. Measurements. Significance of literature review, writing scientific report, structure and components of research report, revision, writing project proposal, writing a Research Paper.

### Unit-IV

Citation counting and Impact factor, Science citation index (SCI)/ Science citation index Expanded (SCI-E), H-index, Academic Ethics and Plagiarism, Intellectual Property Rights and Patent law

**Books Recommended:**

1. C.R. Kothari, 1985, Research Methodology: Methods and Techniques, Wiley Eastern, New Delhi.
2. R.L. Dominowski, 1980, Research Methods, Prentice Hall Inc., New Jersey.
3. R.P. Mishra, 1980, Research Methodology, Handbook Concept Publishing Company, New Delhi.
4. Michael P. Marder, Research Methods for Science, Cambridge University Press, 2011.
5. Day RA, howto Write and Publish a Scientific Paper, Cambridge University Press, London, 1992.
6. . P. Oliver, Writing Your Thesis, NewDelhi: Vistaar Publications, 2004.
7. Gregory, Ethics in Research, Continuum, 2005.

**LAB-IV: LABORATORY USING STATA/SPSS/MATLAB****ST-C- 526****2 CH**

**Objective:** To understand the concept of STATA and/or SPSS and/or MATLAB for data management and as a cutting edge technology tool. To enable to identify data sources, processing and imparting knowledge tools to analyze sets of data to gain useful business and health related issues for understanding.

<b>CO-1</b>	Remember and understand the basic concepts/Principles of R/PYTHON/STATA / SPSS / MATLAB
<b>CO-2</b>	Analyze the Various Concepts to understand them through case studies
<b>CO-3</b>	Apply the knowledge in understanding practical problems
<b>CO-4</b>	Execute/Create the Project or field assignment as per the knowledge gained in the course

**Problem Solving using R/PYTHON/STATA and/or SPSS and/or MATLAB:**

Using of R/PYTHON/STATA/ SPSS/ MATLAB software packages to gain the knowledge of software package and applications of Software for data analysis in the areas of ST-C-521; ST-E-522)\* and ST-E-523)\*.

Examination: Practical (80%) and Record & Viva Voce (20%).