Presidency University, Kolkata
04 Years Bachelor Programme under CHOICE BASED CREDIT SYSTEM for
B.Sc. Honours with Research in Statistics
(Total Credits : 194)
Effective from 2023 – 2024 Academic Session

Department of Statistics
(Faculty of Natural and Mathematical Sciences)
Presidency University
Previously Hindu College (1817 – 1855),
Presidency College (1855 – 2010)
86/1, College Street, Kolkata – 700 073
West Bengal, India

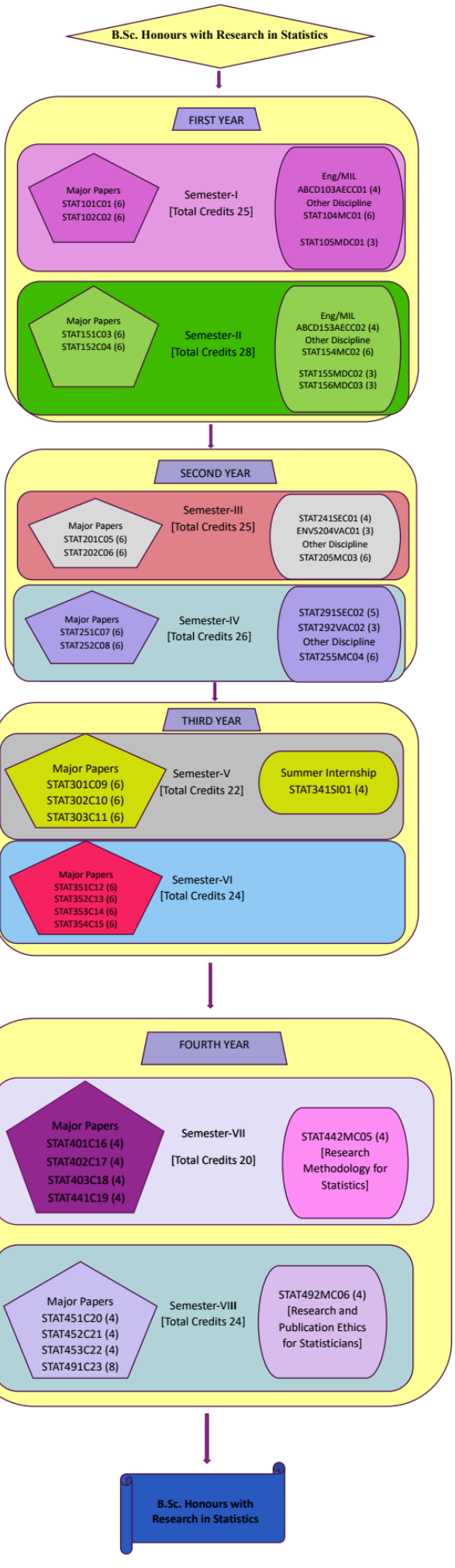
CONTENT

1.1 Academic Sessions :

Odd Semesters : One, Three, Five, Seven

Even Semesters : Two, Four, Six, Eight

1.2 Flow-chart :


B.Sc. Honours with Research in Statistics
FIRST YEAR
Major Papers
STAT101C01 (6)
STAT102C02 (6)
Semester-I
[Total Credits 25]
Eng/MIL
ABCD103AECC01 (4)
Other Discipline
STAT104MC01 (6)
STAT105MDC01 (3)
Major Papers
STAT151C03 (6)
STAT152C04 (6)
Semester-II
[Total Credits 28]
Eng/MIL
ABCD153AECC02 (4)
Other Discipline
STAT154MC02 (6)
STAT155MDC02 (3)
STAT156MDC03 (3)
SECOND YEAR
Major Papers
STAT201C05 (6)
STAT202C06 (6)
Semester-III
[Total Credits 25]
STAT241SEC01 (4)
ENVS204VAC01 (3)
Other Discipline
STAT205MC03 (6)
Major Papers
STAT251C07 (6)
STAT252C08 (6)
Semester-IV
[Total Credits 26]
STAT291SEC02 (5)
STAT292VAC02 (3)
Other Discipline
STAT255MC04 (6)
THIRD YEAR
Major Papers
STAT301C09 (6)
STAT302C10 (6)
STAT303C11 (6)
Semester-V
[Total Credits 22]
Summer Internship
STAT341SI01 (4)
Major Papers
STAT351C12 (6)
STAT352C13 (6)
STAT353C14 (6)
STAT354C15 (6)
Semester-VI
[Total Credits 24]
FOURTH YEAR
Major Papers
STAT401C16 (4)
STAT402C17 (4)
STAT403C18 (4)
STAT441C19 (4)
Semester-VII
[Total Credits 20]
STAT442MC05 (4)
[Research Methodology for Statistics]
Major Papers
STAT451C20 (4)
STAT452C21 (4)
STAT453C22 (4)
STAT491C23 (8)
Semester-VIII
[Total Credits 24]
STAT492MC06 (4)
[Research and Publication Ethics for Statisticians]
B.Sc. Honours with Research in Statistics

COURSE DISTRIBUTION (Type-II)

8-Semester Bachelor (Honours with Research) Programme
under CHOICE BASED CREDIT SYSTEM for

B. Sc. Honours with Research in STATISTICS
194 Credits (4 Years)

The distribution of papers in the under-graduate program will be as follows:

SEM	MAJOR COURSE (C)	ABILITY ENHANCEMENT COMPULSORY COURSE (AECC)	SKILL ENHANCEMENT COURSE (SEC)	VALUE ADDED COURSE (VAC)	MINOR COURSE (MC)	MULTI DISCIPLINARY COURSE (MDC)
I	STAT101C01 (Major) Descriptive Statistics 6 Credits STAT102C02 (Major) Probability and Probability Distributions I 6 Credits	103AECC01 English Communication /MIL			STAT104MC01 Statistical Methods 6 Credits	STAT105MDC01 Statistics for All 3 Credits
II	STAT151C03 (Major) Linear Algebra 6 Credits STAT152C04 (Major) Probability and Probability Distributions II 6 Credits	153AECC02 English Communication/MIL			STAT154MC02 Introductory Probability 6 Credits	STAT155MDC02 Statistics for Bioscience 3 Credits STAT156MDC03 Statistics for Practitioners 3 Credits
III	STAT201C05 (Major) Mathematical Analysis and Calculus 6 Credits STAT202C06 (Major) Inference I 6 Credits	SEC01 is a Sessional Paper	STAT241SEC01 (Major) Statistical Computing Using R and Excel 4 Credits	ENVS204VAC01 Environmental Science 3 Credits	STAT205MC03 Basics of Statistical Inference 6 Credits	
IV	STAT251C07 (Major) Inference II 6 Credits STAT252C08 (Major) Survey Sampling and Indian Official Statistics 6 Credits	SEC02 and VAC02 are Sessional Papers	STAT291SEC02 (Major) Statistical Computing Using C 5 Credits	STAT292VAC02 Data Analysis with Python 3 Credits	STAT255MC04 Applied Statistics 6 Credits	
V	STAT301C09 (Major) Linear Models and ANOVA 6 Credits STAT302C10 (Major) Multivariate Analysis		STAT341SI01 Summer Internship 4 Credits			

	6 Credits STAT303C11 (Major) Optional Paper 6 Credits					
VI	STAT351C12 (Major) Advanced Statistical Methods 6 Credits STAT352C13 (Major) Design of Experiments 6 Credits STAT353C14 (Major) Inference III 6 Credits STAT354C15 (Major) Optional Paper 6 Credits					
VII	STAT441C16 (Major) Advanced Regression Analysis 4 credits STAT402C17 (Major) Stochastic Processes and Queuing Theory 4 Credits STAT403C18 (Major) Optional Paper 4 Credits STAT441C19 (Major) (Project/ Dissertation) 4 Credits			MC05 is a Sessional Paper	STAT442MC05 Research Methodology for Statistics 4 Credits	
VIII	STAT451C20 (Major) Time Series and Spatial Data Analysis 4 Credits			MC06 is a Sessional Paper	STAT492MC06 Research and Publication Ethics for Statisticians	

	STAT452C21 (Major) Optional Paper 4 Credits STAT453C22 (Major) Optional Paper 4 credits STAT491C23 (Major) (Project/ Dissertation) 8 Credits				4 Credits	

List of Optional Papers for **STAT303C11** of Semester V, and **STAT354C15** of Semester VI (**Credits 6**):

1) Operations Research and Statistical Quality Control **A**
2) Modern Survey Sampling **B**
3) Advanced Mathematical Analysis **C**
4) Actuarial Science and Demography **D**

List of Optional Papers for **STAT403C18** of Semester VII, and **STAT452C21** and **STAT453C22** of Semester VIII (**Credits 4**) :

1) Survival Analysis and Clinical Trials **A**
2) Statistics in Finance **B**
3) Graph Theory and its Applications **C**
4) Game Theory **D**
5) Data Mining **E**
6) Advanced Design of Experiments **F**

N.B. :- (i) The Optional Paper(s) which is/are to be offered in a particular semester will be fixed by the Departmental Academic Committee and will be announced before the commencement of the semester. No Course will be offered in both the semesters V and VI of an academic year, and nor in the semesters VII and VIII of an academic year.

(ii) Students will have to choose one Paper as the Major Course STAT303C11 in the Semester V and one Paper as the Major Course STAT354C15 in the Semester VI.

(iii) The students will have to opt for one Paper from the list as Major Course STAT403C18 in Semester VII and two Papers as Major Courses STAT452C21 and STAT453C22 in the Semester VIII.

(iv) No student is allowed to opt for the same Paper more than once.

(v) The list of Optional Papers is tentative and is subject to change.

Skill Enhancement (SEC) Courses/Papers

Semester	Paper Code	Paper Title	Credits
SEM 3	STAT241SEC01	Statistical Computing using R and Excel	4

SEM 4	STAT291SEC02	Statistical Computing using C	5

Minor Elective (MC) Courses/Papers

[The following 4 MC Courses/Papers are to be selected by the students of other Departments/Disciplines]

1. **STAT104MC01** : Statistical Methods 6 Credit

2. **STAT154MC02** : Introductory Probability 6 Credit

3. **STAT205MC03** : Basics of Statistical Inference 6 Credit

4. **STAT255MC04** : Applied Statistics 6 Credit

The following MC Courses/Papers are to be taken by the students of Department of Statistics

5. **STAT442MC05** (for B.Sc. Honours with Research) : Research Methodology for Statistics 4 Credit

6. **STAT441MC05** (for B.Sc. Honours) : Statistical Consulting 4 Credit

7. **STAT492MC06** (for B.Sc. Honours with Research) : Research and Publication Ethics for Statisticians 4 Credit

8. **STAT491MC06** (for B.Sc. Honours) : Seminar Presentation 4 Credit

Value Added (VAC) Courses/Papers

A VAC Course/Paper would be offered in SEM 4.

1. **STAT292VAC02** : Data Analysis with Python 3 Credit

Multidisciplinary Courses (MDC) Courses/Papers

The following 3 MDC Courses/Papers are to be taken by students of other Departments. The difficulty levels of these MDC Courses/Papers are lower compared to the MC Courses/Papers offered for other Departments.

1. **STAT105MDC01** : Statistics for All 3 Credit

2. **STAT155MDC02** : Statistics For Bioscience 3 Credit

3. **STAT156MDC03** : Statistics For Practitioners 3 Credit

Summer Internship

1. **STAT341SI01** An in-house or external summer internship which can be either academic or industrial in nature, would be offered in SEM 5. It would carry 4 Credit and would provide hands on training to students on data analysis.

N.B :- The lecture hours in all the papers include both theory and practical/ tutorial classes.
Use of softwares such as MS-EXCEL / SPSS / R / PYTHON or similar others depends on the availability of faculty and resources for all the core practical courses.

2. Paper wise credit distribution

For 1st to 6th Semester:

Type of Course	Credit per paper (If Practical involved with theory)	Credit per paper (If Tutorial involved with theory)
MAJOR **COURSE (C)**	Theory: **4 credits** Practical : **2 Credits** Total credit per paper **6**	Theory: **5 credits** Tutorial : **1 Credits** Total credit per paper **6**
MINOR COURSE (MC)	Theory: **4 credits** Practical : **2 Credits** Total credit per paper **6**	Theory: **5 credits** Tutorial : **1 Credits** Total credit per paper **6**

Type of Course	Credit per paper
ABILITY ENHANCEMENT COMPULSORY COURSE (AECC)	Theory: **4 credits** Total credit per paper **4**
SKILL ENHANCEMENT COURSE (SEC)	Theory: **4 credits** and **5 credits** Total credit per paper **4** and **5**
MULTI DISCIPLINARY COURSE (MDC)	Theory: **3 credits** Total credit per paper **3**
VALUE ADDED COURSE (VAC)	Theory: **3 credits** Total credit per paper **3**
SUMMER INTERNSHIP	Total credit per paper **4**

For 7th and 8th Semester:

Type of Course	Credit per paper
MAJOR	Theory: **4 credits** Total credit per paper **4**
MINOR COURSE (MC)	Total credit per paper **4**
PROJECT or **DISSERTATION**	Total credit per paper **4 or 8** * ***May be maximum of 12**

Total Course Credits is 194

3. Some Definitions

Programme	A range of learning experiences offered to students in a formal manner over a period of one-to-four years leading to certificates/ diplomas/ degrees. Examples: BA (Economics) BSc (Physics). All possible formal degree Programmes are identified by UGC
Programme Options	A range of courses offered to students to choose at various levels leading to degrees/ diplomas/ certificates.
Programme Outcomes (POs)	Programme Outcomes (POs) are what knowledge, skills and attitudes a graduate should have at the time of graduation. While no agency has formally defined the POs of General Higher Education 3-year degree Programmes in India, POs of all professional Programmes in engineering and other areas are identified at national level by the concerned accrediting agency. POs are not specific to a discipline.
Programme Specific Outcomes (PSOs) in B.Sc. (Hons.) Statistics	The student graduating with the Degree B.Sc. (Honours) Statistics should be able to 1. use skills in Statistics and its related areas of technology for collection data and tackling Statistics related problems and identifying and applying appropriate principles and methodologies to solve a wide range of problems associated with Statistics. 2. acquire (i) a fundamental/systematic or coherent understanding of the academic field of Statistics, its different learning areas and applications in Medical Statistics, Actuarial Statistics, Agricultural Statistics, Geo-Statistics, Financial Statistics, Population Statistics, Financial Econometrics, Clinical Trials and Epidemiology, Queuing Theory, Stochastic Processes, etc., (ii) procedural knowledge that creates different types of professionals related to the disciplinary/subject area of Statistics, including professionals engaged in research and development, teaching and government/public service; (iii) skills in areas related to one's specialization within the disciplinary/subject area of Statistics and current and emerging developments in the field of Statistics. 3. Recognize the importance of statistical modelling, simulation and computing, and the role of approximation and mathematical approaches to analyze the real world problems. 4. Plan and execute Statistics related experiments or investigations, analyze and interpret data/ information collected using appropriate methods, including the use of appropriate software such as programming languages and purpose-written packages, and report accurately the findings of the experiment/investigations while relating the conclusions/findings to relevant theories of Statistics. 5. 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 investigation of Statistics related issues and problems; (iii) communication skills involving the ability to listen carefully, to read texts and research papers analytically and to present complex information in a concise manner to different groups/audiences of technical or popular nature; (iv) analytical skills involving paying attention to detail and ability to construct logical arguments using correct technical language related to Statistics and ability to translate them with popular language when needed; (v) ICT skills; (vi) personal skills such as the ability to work both independently and in a group. 6. Demonstrate professional behaviour such as (i) being objective, unbiased and truthful in all aspects of work and avoiding unethical, irrational behaviour such as fabricating, falsifying or misrepresenting data or committing plagiarism; (ii) the ability to identify the potential ethical issues in work-related situations; (iii) appreciation of intellectual property, environmental and sustainability issues; and (iv) promoting safe learning and working environment.
Course	A course is a unit of 2 to 6 credits in a formal program. A 3-credit course will have three classroom sessions of one-hour duration during each week for the entire semester. Example: Program: BA Economics; Course: Kerala Economy; Credits: 3:0:1
Course Outcomes (COs)	COs are statements that describe what students should be able to do at the end of a course. They can be 6±2 for courses with 2 to 4 credits, and 8±2 for courses with 5 to 6 credits.

4.

Aims and Objectives:
The aims of this programme are to build upon the basic knowledge of Statistics, demonstrate the ability to use skills in Statistics and its related areas of technology for formulating and tackling related problems and identifying and applying appropriate principles and methodologies to solve a wide range of problems associated with Statistics.

Program Outcomes
On completion of this course, students should
A. gain fundamental knowledge on various aspects of
(i) fundamental, systematic and coherent understanding of the academic field of Statistics, its different learning areas and applications in Medical Statistics, Actuarial Statistics, Agricultural Statistics, Geo-Statistics, Financial Statistics, Population Statistics, Financial Econometrics, Clinical Trials and Epidemiology, Queuing Theory, Stochastic Processes, etc.,
(ii) procedural knowledge that creates different types of professionals related to the disciplinary/subject area of Statistics, including professionals engaged in research and development, teaching and government/public service;
(iii) skills in areas related to one's specialization area within the disciplinary/subject area of Statistics and current and emerging developments in the field of Statistics.
(iv) 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;
(v) investigative skills, including skills of independent investigation of Statistics related issues and problems;
(iii) communication skills involving the ability to listen carefully, to read texts and research papers analytically and to present complex information in a concise manner to different groups/audiences of technical or popular nature;
(vi) analytical skills involving paying attention to detail and ability to construct logical arguments using correct technical language related to Statistics and ability to translate them with popular language when needed;
(vii) ICT skills;
(viii) personal skills such as the ability to work both independently and in a group.
B. demonstrate professional behaviour such as
(ix) being objective, unbiased and truthful in all aspects of work and avoiding unethical, irrational behaviour such as fabricating, falsifying or misrepresenting data or committing plagiarism;
(x) the ability to identify the potential ethical issues in work-related situations;
(xi) appreciation of intellectual property, environmental and sustainability issues; and
(xii) promoting safe learning and working environment.

Program Specific Outcomes
The students should be able to make rational decisions about their career in fields such as academics, recognize the importance of statistical modelling simulation and computing, and the role of approximation and mathematical approaches to analyse the real-world problems.

Teaching-learning Process
Teaching will include lectures (online or offline), hands-on training, laboratory dissertation. Doubt-clearing classes/sessions are arranges in each semester. Teachers with expertise in a certain field will teach that module by having a proper idea of
(i) the curriculum, assessing learning needs, and establishing specific learning objectives. Teachers will be in continuous interaction with the students so that the various teaching and learning strategies can be implemented, while maintaining the students' motivation and curiosity about the subjects. Special care will be taken for underperforming students to make them feel confident about the subject.
(ii) plan and execute Statistics related experiments or investigations, analyse and interpret data/information collected using appropriate methods, including the use of appropriate software such as programming languages and purpose-written packages, and report accurately the findings of the experiment/investigations while relating the conclusions/findings to relevant theories of Statistics.

Teaching and Mode of Assessment
Evaluations will be in two parts – internal assessment and final assessment/examination. Both time-bound written and oral examinations will be held. The presentations and interaction during presentations will be evaluated in an objective manner. Class tests, quizzes and group discussion will be conducted for continuous assessment. Regular performance for the laboratory courses will also be assessed in an objective manner.

5.

Detailed Syllabus of CBCS Courses in B.Sc. (Honours) Statistics

Major/Core Papers in Statistics Honours

Semester	**ONE**
Paper Number	**STAT101C01**
Paper Title	**Descriptive Statistics**
No. of Credits	**6**
No. of Classes	Theory: 4 Practical : 4
Course Learning Outcomes	Students will acquire (a) knowledge on brief history of the subject Statistics and its interplay with other disciplines. (b) knowledge of various types of data, handling the data and of its scope and importance in various areas such as Medical, Engineering, Agricultural and Social Sciences etc. (c) skill to organise data, graphical displays and evaluation of summary measures such as measures of central tendency and dispersion etc. (d) knowledge of data reflecting quality characteristics including concepts of independence and association between two attributes.
Syllabus	**Unit 1** Introduction: Nature of Statistics, Uses of Statistics, Statistics in relation to other disciplines, Abuses of Statistics. Concepts of population and sample, complete enumeration and sampling, observational study and experimental study. Types of Data : Primary data and secondary data, quantitative and qualitative data, cross-sectional and time-series data, discrete and continuous data. Different types of scales: Nominal, ordinal, interval and ratio. Collection and Scrutiny of Data : designing a questionnaire and a schedule, checking its consistency, Missing data tools: nature and sources of missingness, imputation based measures, Secondary data – its major sources. Presentation of data: Construction of Tables with one or more factors of classification, diagrammatic representations, frequency distributions and cumulative frequency distributions and their graphical representations, stem and leaf displays. **30L**
	Unit 2 Measures of Central Tendency: Mean, Median, Mode. Measures of Dispersion: range, mean-deviation, standard deviation, coefficient of variation, Gini's Coefficient, Lorenz Curve. Moments, Sheppard's corrections (without proof), skewness and kurtosis, Quantiles and measures based on them, Liapunov's inequality and other inequalities related to measures of skewness and kurtosis. Box Plot, Outlier Detection. Quantile-Quantile Plot. **40L**
	Unit 3 Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, Correlation Index, Correlation Ratio. Intra-class correlation coefficient. Spearman's Rank correlation and Kendall's Tau (including tie cases). **30L**
	Unit 4 Analysis of Categorical Data: Contingency table, association of attributes and different measures, odds ratio, relative risk, Pearson's measure, Goodman-Kruskal's Gamma **28L**
List of Practicals	Graphical representation of data.

	Problems based on measures of central tendency. Problems based on measures of dispersion. Problems based on combined mean and variance and coefficient of variation. Problems based on moments, skewness and kurtosis. Fitting of polynomials, exponential curves. Karl Pearson correlation coefficient. Correlation coefficient for a bivariate frequency distribution. Lines of regression, angle between lines and estimated values of variables. Correlation ratio and correlation index. Rank correlation with and without ties. Computation of intra class correlation coefficient Problems on measures of association.
Reading/Reference Lists	Freedman, Pisani, Purves: Statistics. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I, 8th Edn. The World Press, Kolkata. Yule G.U. and Kendall M.G. : An Introduction to the Theory of Statistics. Snedecor & Cochran : Statistical Methods (6th ed). Croxton F.E., Cowden D.J. & Klein : Applied General Statistics. Moore, D.S & Notz. W.I.: Statistics – Concepts and Controversies. Siegel, A.F. & Morgan, C.J.: Statistics and Data Analysis – An Introduction. Wallis F.E. & Roberts H.V. : Statistics- a new approach. Lewis-Beck M.S. (ed.) : Regression Analysis. A. Agresti : Analysis of Ordinal Categorical Data.

Semester	**ONE**
Paper Number	**STAT102C02**
Paper Title	**Probability and Probability Distributions I**
No. of Credits	**6**
No. of Classes	Theory : 5 Tutorial : 1
Course Learning Outcomes	Students will acquire (a) ability to distinguish between random and non-random experiments. (b) knowledge to conceptualise the probabilities of events including frequentist and axiomatic approach, notion of conditional probability including the concept of Bayes' Theorem, independence of events. (c) knowledge related to concept of discrete and continuous random variables, their probability distributions including expectation and moments, generating functions. (d) knowledge of bivariate probability distributions, sum-law and product-law of expectation.
Syllabus	**Unit 1**

	Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability – classical, statistical. Limitations of Classical definition. Probability of union and intersection of events, Probability of occurrence of exactly m and at least m events out of n events, Examples based on classical approach and repeated trials, Kolmogorov's Axiomatic definition. **30L**
	Unit 2 Conditional Probability, laws of addition and multiplication, theorem of total probability, Bayes' theorem and its applications, independent events. **15L**
	Unit 3 Random variables, distribution function and properties, p.m.f., p.d.f., illustrations and properties of random variables. Mathematical Expectation and properties. Probability generating function. Moments, Dispersion, Skewness, Kurtosis and Quantiles. Cauchy-Swartz Inequality, inequalities related to moments and measures of skewness and kurtosis. Moment generating function, Cumulant generating function and Characteristic function. Uniqueness and inversion theorems (without proof) along with applications. Gambler's ruin problem. **40L**
	Unit 4 Two dimensional random variables: discrete type, joint, marginal and conditional p.m.f and c.d.f., statement of properties of c.d.f, independence of variables, Sum-law and Product-law of expectation, trinomial distribution. **11L**
List of Practicals	Tutorial Only
Reading/Reference Lists	S.M. Ross : A First Course in Probability. Feller W.: An Introduction to Probability Theory & its Applications. Anirban DasGupta: Fundamentals of Probability- A First Course. K.L. Chung : Elementary Probability Theory with Stochastic Process. Rohatgi V.K. (1984): An Intro. to Probability Theory & Math. Statistics. Chandra T.K. & Chatterjee D. : A First Course in Probability. Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol-1). Hoel P.J., Port S.C. & Stone C.J.: Introduction to Probability Theory (Vol-1). Cramer H. : The Elements of Probability Theory. Parzen E. : Modern Probability Theory and its Applications. Uspensky J.V. : Introduction to Mathematical Probability. Cacoullos T. : Exercises in Probability. Pitman J. : Probability. Stirzaker D. : Elementary Probability. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia. Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi.

Semester	**TWO**
Paper Number	**STAT151C03**
Paper Title	**Linear Algebra**
No. of Credits	**6**
No. of classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) vector space, linear dependence and independence of vectors, spanning vector space, projection of vector. (b) matrices, trace, Determinant, Adjoint and inverse of a matrix, product of determinants, related results. (c) theory of equations, generalised inverse of matrix; quadratic forms, linear transformations. (d) characteristic roots, characteristic vectors, and different related methods. (e) Inner Product and Norm. (f) applications of Linear Algebra in Statistics as the foundation to the courses like Multivariate Analysis and Linear Models.
Syllabus	**Unit 1** Real vectors (generalization of co-ordinates), Angle and Norm of vectors, Orthogonality and Gram-Schmidt Orthogonalization Process. Axiomatic Approach and examples. Subspaces, intersection and sum of subspaces. Span of a set, Linear dependence and independence, dimension and basis, dimension theorem. Direct Sum and Complement subspace, Orthogonal Projection of a vector. **30L**
	Unit 2 Algebra of matrices, theorems related to triangular, symmetric and skew symmetric matrices, idempotent matrices, orthogonal matrices, singular and non-singular matrices and their properties. Trace of a matrix, Determinant, Adjoint and inverse of a matrix and related properties. Product of determinants, inverse of a matrix. Rank of a matrix, row-rank, column-rank, standard theorems on ranks, rank of the sum and the product of two matrices. Rank factorization and Sylvester's Inequality. Partitioning of matrices and determinant and inverse of partitioned matrices. Elementary transformations, Echelon form and Normal form. **35L**
	Unit 3 System of homogeneous and non-homogeneous linear equations, Projection Matrix and application to least square method. Generalized inverse, Moore-Penrose inverse. Quadratic forms: Classification & canonical reduction. Linear transformations. **30L**
	Unit 4 Characteristic roots and Characteristic vector, Properties of characteristic roots (symmetric and general matrices). Diagonalization of matrices, Spectral Decomposition, and Singular value decomposition. Power method, Cayley Hamilton theorem, Extrema of Quadratic forms. General concepts of Inner Product and Norm (Brief discussion), Applications of Linear Algebra in Statistics.

	33L
List of Practical	Linear independence and dependence. Orthogonality and Gram-Schmidt Orthogonalization Process. Basis and Dimension. Basis of sum, intersection and complement of subspaces. Projection of vectors on a subspace. Determinant of a matrix. Inverse of matrix. Rank and Rank factorization of matrices. Elementary transformations. Solutions of system of linear equations. Finding g-inverse of a matrix. Problems on quadratic forms. Problems relayed to characteristic roots and vectors. Power method of finding characteristic roots. Problems related to linear transformations.
Reading/Reference Lists	Hadley G. : Linear Algebra. Rao A.R. & Bhimasankaram P. : Linear Algebra. Searle S.R. : Matrix Algebra – useful for Statistics. Rao C.R. : Linear Statistical Inference & its Applications. Hoffman K. & Kunze R. : Linear Algebra. Goon A.M. : Vectors and Matrices.

Semester	**TWO**
Paper Number	**STAT152C04**
Paper Title	**Probability and Probability Distributions II**
No. of Credits	**6**
No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	This is an advanced level course designed to provide knowledge to the students on (a) some discrete and continuous probability distributions and their properties. (b) some probability inequalities. (c) some scaling methods. (d) bivariate normal distribution with its properties. (e) intuitive explanation of which of these theoretical models are applicable for what kind of datasets. (f) law of large numbers, Central Limit theorems.
Syllabus	**Unit 1** Standard discrete probability distributions: Binomial, Poisson, geometric, negative binomial, hypergeometric, uniform. Standard continuous probability distributions: uniform, normal, exponential, Cauchy, beta, gamma, lognormal, logistic, double exponential and Pareto along with their properties and limiting/approximation cases. Truncated distributions. **40L**

	Unit 2 Probability Inequalities (Univariate Cases) : Markov's & Chebyshev's (one- and two- sided) inequalities, Jensen's Inequality, Holder's Inequality, Minkowski's Inequality, Cr Inequality etc. Scaling methods : Z, Percentile, Thurstone, Equivalent scaling procedures. **23L**
	Unit 3 Review of Bivariate c.d.f and p.d.f. and generating functions in continuous case. Marginal and Conditional distributions, Independence, Conditional Expectation, Correlation and Regression. Theorems on sum and product of expectations of random variables. Bivariate Normal Distribution (BVN): p.d.f., properties, marginal and conditional distribution. **35L**
	Unit 4 Limit laws: Sequence of random variables, convergence in probability, almost sure convergence, convergence in mean square and convergence in distribution and their interrelations, W.L.L.N., S.L.L.N and their applications, De-Moivre Laplace Limit theorem, Statement of Central Limit Theorem (C.L.T.) for i.i.d. variates, applications of C.L.T. **30L**
List of Practicals	Fitting of binomial distribution for given n and p. Fitting of binomial distribution after computing mean and variance. Fitting of Poisson distribution for given value of lambda. Fitting of Poisson distribution after computing mean. Fitting of negative binomial. Fitting of suitable distribution. Application problem based on binomial distribution. Application problem based on Poisson distribution. Application problem based on negative binomial distribution. Problems based on are property of normal distribution. To find the ordinate for a given area for normal distribution. Application based problems using normal distribution. Fitting of normal distribution when parameters are given . Fitting of normal distribution when parameters are not given. Fitting of some other continuous distributions. Scaling of scores. Fitting of truncated distributions.
Reading/Reference Lists	S.M. Ross : A First Course in Probability. Feller W.: An Introduction to Probability Theory & its Applications. Anirban DasGupta: Fundamentals of Probability- A First Course. K.L. Chung : Elementary Probability Theory with Stochastic Process. Rohatgi V.K. (1984): An Intro. to Probability Theory & Math. Statistics. Chandra T.K. & Chatterjee D. : A First Course in Probability. Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol-1). Hoel P.J., Port S.C.&Stone C.J.: Introduction to Probability Theory (Vol-1). Cramer H. : The Elements of Probability Theory. Parzen E. : Modern Probability Theory and its Applications. Uspensky J.V. : Introduction to Mathematical Probability. Cacoullos T. : Exercises in Probability. Pitman J. : Probability. Stirzaker D. : Elementary Probability. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia. Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH

	Publishing, New Delhi.

Semester	**THREE**
Paper Number	**STAT201C05**
Paper Title	**Mathematical Analysis and Calculus**
No. of Credits	**6**
No. of Classes	Theory : 5 Tutorial : 1
Course Learning Outcomes	At the end of the course, a student will have knowledge on 1. representation of real numbers, identifying sequences of real numbers and their properties; function and its properties. 2. series of real numbers and different tests to study their convergence/divergence. 3. limit and its properties including related theorems. 4. Riemann Integration , improper integration. 5. multiple integration and, transformation of variables and jacobian of transformation.
Syllabus	**Unit 1** Representation of real numbers as points on a line, Algebraic, Field Structure, Order Structure and Completeness properties of **R** (Concepts only) , Archimedean Property , Bounded and unbounded sets, neighbourhood of a point, Supremum and infimum, Topological properties of real line. Functions, Countable, Uncountable sets and Uncountability of **R**. Sequences and their convergence, Subsequences, monotonic sequences, bounded sequences, squeeze theorem Concept of limsup and liminf. Infinite series, positive termed series and their convergence, Comparison test, ratio test and root test. Absolute convergence of series, Leibnitz's test for the convergence of alternating series, Conditional convergence, Rearrangement and Riemann's Theorem (Statement only). **31L**
	Unit 2 Review of limit, Concepts of o and O. Continuity and Uniform Continuity and boundedness of a function. Differentiability. Darboux Theorem, Rolle's and Lagrange's Mean Value theorems. Taylor's theorem with Lagrange's and Cauchy's form of remainder (without proof). Taylor's and Maclaurin series expansions of sinx, cosx, $e^x e^x$, $(1+x)^n(1+x)^n$, log(1+x). Maxima and Minima of Functions. Successive Differentiation. **25L**
	Unit 3 Riemann Integration of Real valued Functions. Fundamental Theorem of Integral Calculus. Improper Integral, Convergence of Integrals, Simple tests. Beta and Gamma functions: properties and relationship between them. **25L**
	Unit 4 Lagrange Multiplier Technique. Double Integral (intuitive-graphical approach), Multiple Integration, change of order of integration, transformation of variables and Jacobians (statement of relevant theorems and their uses). **15L**
List of Practicals	Tutorials only

Reading/Reference Lists	R G Bartle, Sherbert D R.: Introduction to Real Analysis. Apostol, T.M. : Mathematical Analysis. Malik, S.C. & Arora, S. : Mathematical Analysis. Kumaresan, S:A Basic Course in Real Analysis. Chakraborty, Arnab (2014): Real Analysis, volumes 1,2,3, second edition. Sarat Book House.

Semester	**THREE**
Paper Number	**STAT202C06**
Paper Title	**Inference I**
No. of Credits	**6**
No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	A student will be able to understand (a) The notion of sampling distribution of a statistic. (b) Derivation of some exact sampling distributions of statistics like Chi-square, t, F etc. (c) Sampling distributions of statistics related to samples from bivariate normal distribution(s). (d) The notion of order statistics and related sampling distributions. (e) The importance of sampling distributions in Statistical Inference. The basics of Testing of Hypotheses and confidence intervals. (f) The basic principle underlying tests of significance with applications.
Syllabus	**Unit 1** Functions of Random Vectors (univariate distributions): Jacobian, Polar transformations and Orthogonal Transformations. Derivation of the sampling distribution of sample mean and variance for a normal population, standard errors of sample mean, sample variance and sample proportion. Exact sampling distribution: Definition and derivation of p.d.f. of χ^2 with n degrees of freedom (d.f.), nature of p.d.f. curve for different degrees of freedom, mean, variance, m.g.f., additive property of χ^2 distribution. Exact sampling distributions: Student's and Fisher's t-distributions, Derivation of its p.d.f., nature of probability curve with different degrees of freedom, mean, variance and limiting form of t distribution. Snedecor's F-distribution: Derivation of p.d.f., nature of p.d.f. curve with different degrees of freedom, mean, variance. Distribution of 1/F (n1, n2). Relationship between t, F and χ^2 distributions. **50L**
	Unit 2 Sampling distribution based on BVN: Distribution of sample correlation coefficient in the null case, regression coefficients and other related results with non-stochastic covariate. Order Statistics: Introduction, distribution of the rth order statistic, smallest and largest order statistics. Joint distribution of rth and sth order statistics, distribution of sample median and sample range. **25L**
	Unit 3

	Problems of Statistical Inference: Population & parameter, random sample & statistic, Point and Interval Estimation, Confidence level, Testing of Hypothesis, Null and alternative hypotheses, level of significance, Type I and Type II errors, their probabilities and critical region. **15L**
	Unit 4 Exact tests and confidence intervals: classical and p-value approaches related to Binomial proportion(s), Poisson mean(s), Univariate Normal mean(s), standard deviation(s). Standard tests related to Bivariate normal parameters. **38L**
List of Practicals	Testing of significance for single proportion and difference of two proportions. Testing of significance for single Poisson mean and difference of means of two independent Poisson distributions. Testing of significance and confidence intervals for single mean and difference of two means and paired tests. Testing if the population variance has a specific value and its confidence intervals. Testing of significance and confidence intervals of correlation coefficient. Testing of equality of population variances for two independent normal populations and related confidence intervals. Testing of ratio of variances for bivariate normal population and related confidence interval. Tests related to regression and related confidence intervals.
Reading/Reference Lists	Rohatgi, V. K. and Saleh, A.K.M.E. (2015): An Introduction to Probability and Statistics, Third Edn,Wiley, NJ. Mukhopadhay, N.: Probability and Statistical Inference. Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol-1). Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I, 8th Edn. The World Press, Kolkata. Mood, A.M., Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd. Casella, G. and Berger R.L. (2002).: Statistical Inference, 2nd Edn, Thomson Learning. Bhattacharya GK & Johnson R. A. : Concepts & Methods of Statistics.

Semester	**FOUR**
Paper Number	**STAT251C07**
Paper Title	**Inference II**
No. of Credits	**6**
No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	A student will be able to understand (a) The notion of sufficiency and completeness. (b) How to find an UMVUE using Cramer-Rao inequality and Lehmann-Scheffe theorem. (c) The derivation of the cutoff value of a test. (d) The calculation of power of a test. (e) How to find an MP test using the NP Lemma. (f) The calculation of shortest length confidence intervals. (g) The decision theoretic foundation of mathematical statistics.

Syllabus	**Unit 1:** Estimation: Concepts of estimation, unbiasedness, sufficiency, consistency and efficiency. Factorization theorem. Complete statistic, Ancillary statistic, Basu's Theorem and Applications, Minimum variance unbiased estimator (MVUE), Necessary and Sufficient condition for UMVUE, Rao-Blackwell and Lehmann-Scheffe theorems and their applications. Cramer-Rao inequality and MVB estimators. Methods of Estimation: Method of moments, method of maximum likelihood estimation, method of least square, method of minimum Chi-square and statements of their properties. **54L**
	Unit 2: Concept of test function and randomized test, Review of level of significance, power and power curve. Most powerful test, uniformly most powerful test, Neyman- Pearson Lemma (statement and proof of sufficiency part only) and its applications to construct uniformly most powerful test, unbiased test (definition only), Monotone likelihood ratio, Likelihood ratio test, properties of likelihood ratio tests (without proof). **44L**
	Unit 3: Confidence intervals, Confidence set, Shortest length confidence interval, Concepts of Uniformly Most Accurate (UMA) confidence sets, relationship with tests of hypotheses. **20L**
	Unit 4: Elements of decision theory: concepts of action space, loss function, risk, Linear ordering principle and Restriction principle, Idea of Bayes rule and Minimax rule. **10L**
List of Practicals	Maximum Likelihood Estimation Estimation by the method of moments, minimum Chi-square Most powerful critical region (NP Lemma) Uniformly most powerful critical region Unbiased critical region Power curves Confidence intervals
Reading/Reference Lists	Rohatgi V.K. (1984): An Intro. to Probability Theory & Math. Statistics Mukhopadhay, N.: Probability and Statistical Inference Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol-2) Casella, G. and Berger R.L. (2002).: Statistical Inference, 2ndEdn. Thomson Learning Kale, B.K.: A first course in parametric inference, Narosa. Bickel, P.J., Doksum, K.A.: Mathematical Statistics: Basic Ideas and Selected Topics, Volume 1

Semester	**FOUR**
Paper Number	**STAT252C08**
Paper Title	**Survey Sampling and Indian Official Statistics**
No. of Credits	**6**

No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) Population and sample, probability sampling. (b) Different sampling schemes and situations where these are applicable. (c) introducing auxiliary variable in the improvement of estimation procedures under certain situations. (d) sources of official statistics, mechanisms of collection of official data in India under MoSPI. (e) national income. (f) price index numbers, quantity index numbers, measuring formulae, their tests for criteria of index numbers, interpretation and uses.
Syllabus	**Unit 1** Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, definition and procedure of selecting a sample, estimates of: population mean, total and proportion, variances of these estimates, estimates of their variances and sample size determination. Probability proportion to size sampling, inclusion probabilities, Horvitz-Thompson estimator. **32L**
	Unit 2 Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. Practical difficulties in allocation, estimation of gain in precision, post stratification and its performance. Systematic Sampling: Technique, estimates of population mean and total, variances of these estimates (N=n x k). Comparison of systematic sampling with SRS and stratified sampling in the presence of linear trend and corrections. **25L**
	Unit 3 Introduction to Ratio and regression methods of estimation, estimation of the population mean and total (for SRS of large size), MSE of these estimates and estimates of these variances, MSE in terms of correlation coefficient for regression method of estimation and their comparison with SRS. Cluster sampling (equal clusters only) estimation of population mean and its variance, comparison (with and without randomly formed clusters). Concept of sub sampling. Two-stage Sampling, Estimation of Population mean and variance of the estimate, Randomized Response Technique: Warner Model. **32L**
	Unit 4 An outline of present official statistical system in India, Methods of collection of official statistics, their reliability and limitations. Role of Ministry of Statistics & Program Implementation (MoSPI), Central Statistical Office (CSO), National Sample Survey Office (NSSO), Registered General Office and National Statistical Commission. Government of India's Principal publications containing data on the topics such as Agriculture, price, population, industry, finance and employment. National Income: Basic idea and a brief description of income, expenditure and production approaches. Index Numbers: Price, Quantity and Value indices. Price Index Numbers: Construction, Formulae, Uses – base shifting and splicing, Limitations, Tests for index numbers, Various formulae and their comparisons, Chain-Index Numbers. Consumer Price Index, Wholesale Price Index and Index of Industrial Production. **39L**
List of Practicals	Selection of an SRS with and without replacement from finite populations, theoretical

	populations and given geometrical shapes. For a population of size 5, estimation of population mean, population mean square and population variance. Enumeration of all possible samples of size 2 by WR and WOR and establish all properties relative to SRS. For SRSWOR, estimation of mean, standard error and the sample size. Stratified Sampling: allocation of sample to strata by Proportional and Neyman's methods. Comparing the efficiencies of above two methods relative to SRS. Estimation of gain in precision in stratified sampling. Comparison of systematic sampling with stratified sampling and SRS in the presence of a linear trend. Ratio and Regression estimation: Calculation of the population mean or total of the population. Calculate mean squares. Comparing of the efficiencies of ratio and regression estimators relative to SRS. Cluster sampling: estimation of mean or total, variance of the estimate, estimate of intra-class correlation coefficient, efficiency as compared to SRS. Two-stage sampling. Calculation of price and quantity index numbers using simple and weighted average of price relatives. Calculation of the Chain Base index numbers. Problems on cost of living index numbers.
Reading/Reference Lists	Cochran, W.G. (1984): Sampling Techniques (3rd Ed.), Wiley Eastern. Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. Asok, C. (1984). Sampling Theories of Survey With Application, IOWA State University Press and Indian Society of Agricultural Statistics. Murthy, M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta. Des Raj and Chandhok P. (1998): Sample Survey Theory, Narosa Publishing House. Goon A.M., Gupta M.K. and Dasgupta B. (2008): Fundamentals of Statistics, Vol-II, World Press. Guide to current Indian Official Statistics, Central Statistical Office, GOI, and New Delhi. http://mospi.nic.in/ Gupta, S. C. and Kapoor, V. K. : Applied Statistics Mudgett, B. D.: Index Numbers Allen, R. G. D. : Index Numbers in Economic Theory and Practice Yule, G. U. and Kendall, M : Introduction to the Theory of Statistics Paul S. Levy and Stanley Lemeshow (2013): Sampling of populations : methods and applications, 4th Edition (Wiley).

Semester	**FIVE**
Paper Number	**STAT301C09**
Paper Title	**Linear Models and ANOVA**
No. of Credits	**6**
No. of Classes	Theory : 4

	Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) Gauss-Markov set-up, its identification and use in estimation of parameters. (b) ANOVA models to test the effects of different factors and their interactions. (c) testing problems related to regression models. (d) the use of concomitant variables under ANOCOVA models. (e) Regression Diagnostics including quantile-quantile plots.
Syllabus	**Unit 1** General objectives of model building: inference and prediction, difference between parametric and non-parametric approaches. Review of simple linear regression and its properties (including sampling distributions), confidence interval for mean and prediction interval. Extension of linear regression: polynomial regression, multiple regression. General linear model: Introduction and Examples, Use of dummy variables, different types of linear models and their parameter interpretations. **35L**
	Unit 2 Gauss-Markov set-up: Theory of linear estimation, Estimability of linear parametric functions, Method of least squares, Gauss-Markov theorem, Estimation of error variance. Fundamental Theorems on least squares (statements only), General Linear Hypothesis: Testing and confidence interval. **25L**
	Unit 3 Analysis of variance: Definitions of fixed-, random- and mixed-effects models, analysis of variance and covariance in one-way classified data for fixed-effects models, analysis of variance and covariance (with one concomitant variable) in two-way classified data with equal number of observations per cell for fixed-effects models. Analysis of variance for random-effects models. **48L**
	Unit 4 Regression analysis: Estimation and hypothesis testing in case of simple and multiple regression models. Tests for parallelism and identity, linearity of simple regression. **20L**
List of Practicals	*Some practical problems are to be done preferably by using R/ statistical packages.* Estimability when X is a full rank matrix and not a full rank matrix. Simple linear regression. Multiple regression. Tests for linear hypothesis. Analysis of variance of one-way classified data. Analysis of variance of a two-way classified data with one observation per cell. Analysis of variance of a two-way classified data with equal number of observations per cell. Analysis of covariance of a one-way classified data with one concomitant variable. Analysis of covariance of a two-way classified data with one concomitant variable. Hypothesis testing in case of simple and multiple regression models and related tests. Fitting of linear model using R/ statistical package.
Reading/Reference Lists	Kutner, Nachtsheim, Neter: Applied Linear Regression Models Faraway, J.J.: Linear Models with R Goon, A.M., Gupta, M.K., and Dasgupta, B. (2002), Fundamental of Statistics, Volume 1 & 2, 8th Edn. The World Press, Kolkata.

	Scheffe, H, Linear Models. Rao, C.R., Linear Statistical Inference. Stapleton, J. H.: Linear Statistical Models. Mukhopadhyay, P. (2011): Applied Statistics, 2 nd edition revised reprint, Books and Allied(P) Ltd. Sengupta D. and Jammalamadaka, S. R.: Linear Models, An Integrated Approach. Hocking, R. R.: Methods and Applications of Linear Models. Wu, C. F. J. and Hamada, M. (2009). Experiments, Analysis and Parameter Design Optimization (Second edition), John Wiley. Renchner, A.C. and Schaalje, G.B. (2008). Linear Models in Statistics (Second edition), John Wiley and Sons. S. R. Jammalamadaka & D. Sengupta (2020): Linear Models and Regression with R : An integrated approach (World Scientific)

Semester	**FIVE**
Paper Number	**STAT302C10**
Paper Title	**Multivariate Analysis**
No. of Credits	**6**
No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) Multivariate data and measures of correlation, regression; random vectors and multivariate probability distribution. (b) Multinomial and Multivariate Normal distributions along with their properties, ellipsoid of concentration. (c) Multivariate Normal distribution- application of multivariate techniques in regression analysis. (d) Inferential aspects related to multivariate normal data.
Syllabus	**Unit 1** Multivariate data – multiple regression, multiple correlation and partial correlation – their properties and related results. Random Vector: Probability mass/density functions, Distribution function, Mean vector & Dispersion matrix, Marginal & Conditional distributions. **30L**
	Unit 2 Multinomial Distribution, Multivariate Normal distribution and its properties Marginal and Conditional Distributions, Ellipsoid of Concentration. **35L**
	Unit 3 Sampling distribution for sample mean vector and sample variance-covariance matrix, Null distribution of the sample multiple and partial correlation correlation coefficients

	and their applications. **33L**
	Unit 4 Methods of estimation of the parameters of some useful multivariate distributions, concepts of the test procedures related to the parameters of multivariate normal distributions. **30L**
List of Practicals	*Some practical problems are to be done preferably by using R/ statistical packages.* Multiple Correlation and Regression. Partial Correlation. Checking for multivariate normal distribution Test related to the parameters assuming multivariate normal distribution Test of the nullity of population multiple correlation and population partial correlation in the context of multiple linear regression of normal response Test of the partial regression coefficients in the context of multiple linear regression of normal response.
Reading/Reference Lists	Anderson, T.W. (2003): An Introduction to Multivariate Statistical Analysis, 3rdEdn., John Wiley. Muirhead, R.J. (1982): Aspects of Multivariate Statistical Theory, John Wiley. Kshirsagar, A.M. (1972): Multivariate Analysis, 1stEdn. Marcel Dekker. Johnson, R.A. And Wichern, D.W. (2007): Applied Multivariate Analysis, 6thEdn., Pearson & Prentice Hall. Mukhopadhyay, P.: Mathematical Statistics.

STAT303C11 : Optional Paper

Semester	**SIX**
Paper Number	**STAT351C12**
Paper Title	**Advanced Statistical Methods**
No. of Credits	**6**
No. of classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) Visualisation of multivariate data and checking for normality (b) Applied multivariate techniques (c) Identifying causal effects and structures from data. (d) Implementing diagnostics measures through regression analysis.
Syllabus	**Unit 1** Introduction to multivariate data and applications, Visualization of multivariate data and checking for normality, curse of dimensionality with examples. Principal components and uses, Canonical variables and correlations, Orthogonal Factor model: estimation of factor loadings and scores. **30L**

	Unit 2 Supervised and unsupervised learnings, Cluster Analysis: Hierarchical methods, K-means and related methods, Evaluation of cluster algorithms, model based clustering. Classification using linear, logistic and nearest-neighbour approaches, Decision Theoretic formulation: Bayes, minimax and likelihood ratio approaches, discrimination between multivariate normal populations, Fisher's LDF. **30L**
	Unit 3 MANOVA, Multidimensional Scaling, Introduction to Causal Inference: Graphical Causal Models, Identifying Causal Effects and Causal structures. **30L**
	Unit 4 Regression Diagnostics: Outlier detection, checking for heteroscedasticity and autocorrelation, multicollinearity, tests for normality of errors. **38L**
List of Practical	*The entire practical are to be done preferably by using R/ statistical packages.* Visualisation of multivariate data. Checking for normality Finding principal components and their variances. Finding canonical correlations and variables. Finding factor loadings and scores. Identifying clusters in the data using hierarchical methods' Finding clusters using partitioning methods Finding LDF for classification among populations Implementing MANOVA Identifying causal effects and structures from data. Outlier detection and influential observations. Testing for homoskedasticity and autocorrelation. Testing for multicollinearity Checking normality of errors.
Reading/Reference Lists	R. A. Johnson & D. W. Wichern : Applied Multivariate Statistical Analysis. Léopold Simar & Wolfgang Härdle : Applied Multivariate Statistical Analysis. K. V. Mardia, J. T. Kent, J. M. Bibby : Multivariate Analysis T.W. Anderson : An Introduction to Multivariate Analysis. G.A.F. Seber : Multivariate Observations. Draper,N.R. and Smith,H.:Applied_Regression Analysis. Johnston, J. and DiNardo, J.: Econometric Methods Weisberg, S. :Applied linear regression.

Semester	**SIX**
Paper Number	**STAT352C13**

Paper Title	**Design of Experiments**
No. of Credits	**6**
No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) the basics - Randomization, Replication as essential principles and Local Control as a desirable principle in statistical designing of experiments. (b) carry out one-way and two-way Analysis of Variance. (c) construction of standard designs – CRD, RBD and LSD and apply ANOVA techniques to analyse the data produced thereof. (d) Comparing relative efficiency of one design with respect to another. (e) Analysis of data in case of missing observation. (f) Incomplete Block Designs. (g) Construction un-confounded and confounded Factorial Designs, and analyse the data produced. (h) Construction of Fractional Factorials by creating aliases.
Syllabus	**Unit 1** Role, historical perspective. Terminologies: Experimental error, Basic principles, Uniformity trials, Fertility contour maps, Choice of size and shape of plots and blocks. Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – Layout, Model and Analysis, Relative Efficiencies, Analysis with one missing observation. **48L**
	Unit 2 Balanced Incomplete Block Design (BIBD) – parameters, relationships among its parameters, incidence matrix and its properties, Symmetric BIBD, resolvable BIBD, Affine Resolvable BIBD, Intrablock Analysis, Complementary BIBD, Residual BIBD, Dual BIBD, Derived BIBD. **25L**
	Unit 3 Advantages, Notations and Concepts of 2^n factorial experiments – their design and analysis. Total and Partial confounding for 2^n($n \leq 5$) factorial experiments. **30L**
	Unit 4 Construction of one-half and one-quarter fractions of 2^n (n≤5) factorial experiments, Alias structure, Resolution of a design, Brief idea of Response Surface Methodology. **25L**
List of Practicals	Analysis of CRD. Analysis of an RBD. Analysis of an LSD. Analysis of an RBD with one missing observation. Analysis of an LSD with one missing observation. Intra Block analysis of a BIBD. Analysis of 2^2 and 2^3 factorial in CRD and RBD. Analysis of 2^2 and 2^3 factorial in LSD. Analysis of a completely confounded two level factorial design in 2 blocks. Analysis of a completely confounded two level factorial design in 4 blocks. Analysis of a partially confounded two level factorial design. Analysis of a single replicate of a 2^n design. Analysis of a fraction of 2^n factorial design. Analysis of a design with Response Surface Methodology.

Reading/Reference Lists	Goon, A.M., Gupta, M.K. and Dasgupta, B. (2005): Fundamentals of Statistics. Vol. II, 8thEdn. World Press, Kolkata. Mukhopadhyay, P. : Applied Statistics. Cochran, W.G. and Cox, G.M. (1959): Experimental Design. Asia Publishing House. Dey, A. (1986) : Theory of Block Designs, Wiley Eastern Limited. Montgomery, D. C. (2008): Design and Analysis of Experiments, John Wiley. Das, M.N. and Giri, N.C. (1986): Design and Analysis of Experiments. Wiley Eastern Ltd. Kempthorne, O. (1965): The Design and Analysis of Experiments. John Wiley.

Semester	**SIX**
Paper Number	**STAT353C14**
Paper Title	**Inference III**
No. of Credits	**6**
No. of classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) Asymptotic theory for large sample size (b) Large-sample tests related to proportion and mean (c) Applications of chi-squared tests (d) Nonparametric tests (e) Resampling techniques
Syllabus	**Unit 1** Delta Method, Derivation and uses of large sample standard error of sample moments, Standard deviation, Coefficient of Variation, b1 & b2 measures, Correlation coefficient. Asymptotic distribution of sample quantiles. Transformations of Statistics to stabilize variance: derivation and uses of $\sin^{-1}$, square root. Uses of logarithmic and z-transformations. **48L**
	Unit 2 Large sample tests for binomial proportions, Poisson means (single and two independent samples cases) and correlation coefficients. Large Sample distribution of Pearsonian $\chi 2$ – statistic and its uses. **25L**
	Unit 3 Empirical distribution function, Glivenko-Cantelli Lemma, Density Estimation, Nonparametric Tests: Introduction and Concept, Test for randomness based on total number of runs, Empirical distribution function, Kolmogrov Smirnov test, Sign tests and Signed Rank tests, Wilcoxon-Mann-Whitney test, Median test, Kruskal-Wallis test, Non-parametric confidence interval, tolerance and prediction limits. **30L**
	Unit 4 Likelihood methods for incomplete data problems: EM Algorithm, Introduction to Jackknife,

	Jackknife estimate of bias and standard error, Tukey's pseudo values, Introduction to Bootstrap: Parametric and Non-parametric Bootstrap, Boostrap estimate of standard error and bias, Comparison between Jackknife and Bootstrap, Bootstrap methods in Regression, Bootstrap Consistency measures. **25L**
List of Practical	Testing of significance and confidence intervals for single proportion and difference of two proportions using CLT. Testing of significance and confidence intervals for single Poisson mean and difference of two Poisson means using CLT. Testing of significance and confidence intervals concerning sample standard deviation, coefficient of variation and correlation coefficient (both single sample and two sample cases). Testing of significance and confidence intervals using variance stabilizing transformations. Determination of the minimum sample size required to achieve normality by sample proportion, mean and standard deviation. Tests for goodness of fit, independence and homogeneity using Pearsonian chi-square statistic.Test for randomness based on total number of runs. Kolmogorov Smirnov test for one sample. Sign test. Signed Rank test. Wilcoxon-Mann Whitney U-test. Kruskal-Wallis test. Non-parametric confidence intervals. Non-parametric tests using R/statistical packages. Finding bias and variance of a statistic using Jackknife Finding bias and variance of a statistic using Bootstrap Finding confidence intervals using Bootstrap Bootstrap in regression
Reading/Reference Lists	Rohatgi V.K. (1984): An Intro. to Probability Theory & Math. Statistics Mukhopadhay, N.: Probability and Statistical Inference Goon A.M., Gupta M.K. & Dasgupta B.: An Outline of Statistical Theory (Vol 2) Casella , G. and Berger R.L. (2002).: Statistical Inference, 2ndEdn. Thomson Learning Kale, B.K.: A first course in parametric inference, Narosa. Bickel, P.J., Doksum, K.A.: Mathematical Statistics: Basic Ideas and Selected Topics, Vol 1 Gibbons, J. D. and Chakraborty, S (2003): Nonparametric Statistical Inference. 4th Edition. Marcel Dekker, CRC. Wasserman, L.: All of Nonparametric Statistics, Springer Davison, A. C. and D. V. Hinkley (1997). Bootstrap Methods and their Applications B.Efron : The Jackknife, the Bootstrap and other Sampling Plans

STAT354C15 : Optional Paper

Semester	**SEVEN**
Paper Number	**STAT401C16**
Paper Title	**Advanced Regression Analysis**
No. of Credits	**4**
No. of classes	Theory : 3 Practical : 2
Course Learning Outcomes	Students will acquire knowledge of (a) Fitting a GLM to a given dataset. (b) In sample and extra sample error and different approaches of regression model selection. (c) Different nonparametric regression techniques. (d) Ridge regression and LASSO, and their computations. (e) PC regression and PLS regression. (f) Nonparametric regression for many covariates using GAM.
Syllabus	**Unit 1** Generalized linear models: Components of a GLM, Goodness of fit – deviance, Residuals, Maximum likelihood estimation. Binary data and Count data: ungrouped and grouped. Polytomous data Overdispersion, Quasi-likelihood. Models with constant coefficient of variation, joint modeling of mean and variance. **35L**
	Unit 2 Supervised learning and its decision theoretic formulation, Regression as a supervised learning method, Concepts of in sample error and extra sample error, different approaches of model selection : Adjusted R2, AIC, Mallow's Cp, Splitting the dataset, Cross-validation. Bias and Variance Tradeoff. General concepts of Smoothing, Linear Smoothers and Non-parametric Regression Techniques: K-NN, Kernel Regression, Splines. **21L**
	Unit 3 High dimensional Regression techniques: Variable selection methods: Best Subset selection, Forward Selection, Backward selection. Shrinkage methods: Ridge Regression and LASSO, Computation of LASSO: Least Angle Regression Technique, Soft-thresholding and co-ordinate descent approach, Uniqueness of LASSO solutions, LASSO for generalized linear models, Elastic net. Methods of derived inputs: Principal Component Regression and Partial Least Squares. **16L**
	Unit 4 Additive models, Classification and Regression Trees. **8L**
List of Practical	Fitting GLM to a given dataset. Regression model selection using different approaches of model selection. Nonparametric kernel regression. Spline smoothing for regression problems.

	Computation of LASSO using LARS. Principal Component Regression and Partial Least Squares. GAM fitting to datasets with many covariates.
Reading/Reference Lists	McCullagh, P and Nelder, A.J. : Generalized Linear Models. James G, Witten D, Hastie T,Tibshirani R : Introduction to Statistical Learning with Applications in R Hastie,T, Tibshirani R., Friedman J: Elements of Statistical Learning. Faraway, J.J.: Extending the Linear Model with R Breiman, Leo et. al. : Classification and Regression Trees. Agresti, A.: Categorical Data Analysis. Györfi, László, et. al.: A Distribution-Free Theory of Nonparametric Regression. Simonoff, Jeffrey S. : Smoothing Methods in Statistics.

Semester	**SEVEN**
Paper Number	**STAT402C17**
Paper Title	**Stochastic Processes and Queuing Theory**
No. of Credits	**4**
No. of classes	Theory : 3 Tutorial : 1
Course Learning Outcomes	Students will acquire knowledge of (a) Idea of stochastic processes. (b) Markov chains including the notion of transition probability matrix. (c) various other stochastic processes such as generalised Bernoulli process, Poisson process, birth and death processes. (d) queuing theory, finite and infinite system capacity, waiting time distribution. (e) application of these processes in real life problems.
Syllabus	**Unit 1** Stochastic Processes: Introduction, An overview of its applications. **4L**
	Unit 2 Markov Chains: Definition of Markov Chain, Examples including 2-state chain, random walk, etc., Transition probability matrix, order of a Markov chain, Markov chain as graphs, Classification of states of a Markov Chain, Stationary distribution, Limiting distribution, period of a Markov Chain, Convergence theorem. **30L**
	Unit 3 Poisson Process: postulates of Poisson process, properties of Poisson process, inter-arrival time, pure birth process, Yule Furry process, birth and death process, pure death process. **15L**
	Unit 4 Queuing System: General concept, steady state distribution, queuing model, M/M/1 with finite

	and infinite system capacity, waiting time distribution (without proof). **15L**
List of Practical	Calculation of transition probability matrix. Identification of characteristics of reducible and irreducible chains. Identification of types of classes. Identification of ergodic transition probability matrix Stationarity of Markov chain. Computation of probabilities in case of generalizations of independent Bernoulli trials. Calculation of probabilities for given birth and death rates and vice versa. Calculation of probabilities for Birth and Death Process. Calculation of probabilities for Yule Furry Process. Computation of inter-arrival time for a Poisson process. Calculation of Probability and parameters for (M/M/1) model and change in behaviour of queue as N tends to infinity.
Reading/Reference Lists	P. G. Hoel, S. C. Port and C. J. Stone: Introduction to Stochastic Processes. Medhi, J. (2009): Stochastic Processes, New Age International Publishers. S. Karlin and H.M.Taylor: A first course in stochastic process. S. Ross: Stochastic Process. J. G. Kemeny, J. L. Snell and A. W. Knapp: Finite Markov Chains. Bhat,B.R.(2000): Stochastic Models: Analysis and Applications, New Age International Publishers. Taha, H. (1995): Operations Research: An Introduction, Prentice- Hall India. Feller, William (1968): Introduction to probability Theory and Its Applications, Vol I, 3rd Edition, Wiley International. R. N. Bhattacharya and E. Waymire: Stochastic Process and Applications.

STAT403C18 : Optional Paper

Semester	**SEVEN**
Paper Number	**STAT441C19 (for B.Sc. Honours with Research)**
Paper Title	**Project/ Dissertation**
No. of Credits	**4**
No. of classes	**4 Contact Hours**
Course Learning Outcomes	Students will acquire knowledge of (a) analysing and interpreting and taking appropriate decisions in solving real life problems using statistical tools. (b) use different Statistical packages for graphical interface, data analysis and interpretation. (c) write a systematic Statistical report.
Syllabus	Literature review, collection of primary/secondary data, simulation of synthetic data, synopsis of potential topics for the project, seminar presentation and evaluation.

Semester	**EIGHT**

Paper Number	**STAT451C20**
Paper Title	**Time Series and Spatial Data Analysis**
No. of Credits	**4**
No. of Classes	Theory : 3 Practical : 2
Course Learning Outcomes	Students will acquire knowledge of (a) time series data, its applications to various fields and components of time series. (b) fitting and plotting of various growth curves such as modified exponential, Gompertz and logistic curve. (c) fitting of trend by Moving Average method. (d) measurement of Seasonal Indices by Ratio-to-Trend, Ratio-to-Moving Average and Link Relative methods. (e) calculation of variance of random component by variate-component method. (f) forecasting by exponential smoothing and short term forecasting methods such as Box Jenkins Method and Bayesian forecasting. (g) weak stationarity, autocorrelation and correlogram. (h) calculation of summary measures associated with spatial data.
Syllabus	**Unit 1** Introduction to time series data. Time series data arising from various fields. Components of a time series (trend, seasonal, cyclic and irregular component). Classical decomposition of time series. Estimation of trend: free hand curve fitting, mathematical curve fitting, method of moving averages. Linear filters and their application in estimating trend and trend removal. Estimation of seasonal component by Method of simple averages, ratio to trend method and ratio to moving averages. Correlogram and its interpretation. **20L**
	Unit 2 Weakly and strongly stationary stochastic process on time and their interrelationships, Gaussian process on time. WN, MA, AR and ARMA process, existence and uniqueness of stationary solution, causality, invertibility, solution of recurrence relation for finding their ACVF and ACF. **15L**
	Unit 3 Estimation in MA, AR and ARMA model, Yule-Walker, MOM and conditional least squares approach, PACF and order selection of MA, AR and ARMA process. Forecasting – Simple Exponential Smoothing, Holt's and Holt-Winters' method. Box-Jenkins approach, ARIMA models. Projection approach for optimal prediction/forecast. Forecast interval and density forecasting. Unit root test for detection of stationarity. **25L**
	Unit 4 Introduction to spatial data – some examples, classifications. Graphical display and summary statistics associated with point-referenced, lattice and spatial point process data, variogram analysis. Modeling of point-referenced spatial data and different methods of kriging. **20L**
List of Practicals	*Some practical problems are to be done preferably by using R/ statistical packages.*

	Plotting a real-life time series, and detecting various features (trend, periodic behaviours etc.). Suggested data sets: Sun spot data Dollar-Rupee exchange rates Stock market data Fitting and plotting of mathematical curves: modified exponential curve Gompertz curve Fitting of trend by Moving Average Method. Plotting detrended series. Measurement of Seasonal indices Ratio-to-Moving Average method. Plotting ACF and PACF of a given time series. Using Yule-Walker equation to fit AR (1) and AR (2) models to real life data. Forecasting by short term forecasting methods. Forecasting by exponential smoothing. Problems related to spatial data
Reading/Reference Lists	Gun, Gupta and Dasgupta (2002) Fundamentals of Statistics Vol II, World Press. Cooray TMJA(2008) Applied Time Series, Analysis and forecasting, Narosa Publishing house. Chatfield C (2004) Analysis of Time Series, Chapman & Hall. Cryer, J.D. and Chan, K-S: Time Series Analysis with applications in R. P.Brockwell & R.A.Davis : Introduction to time series and forecasting. Mukhopadhyay P. : Applied Statistics. B Ripley : Spatial Statistics N. Cressie : Statistics for Spatial Data N. Cressie & C. Wikle : Statistics for Spatio-temporal Data Sudipto Banerjee, B. P. Carlin & A. E. Gelfand : Hierarchical Modeling and Analysis for Spatial Data

STAT452C21 : Optional Paper

STAT453C22 : Optional Paper

Semester	**EIGHT**
Paper Number	**STAT491C23 (for B.Sc. Honours with Research)**
Paper Title	**Project/ Dissertation**
No. of Credits	**8**
Marks distribution	Internal 40% + External 60%
No. of classes	**8 Contact Hours**
Course Learning Outcomes	Students will acquire knowledge of (a) analysing and interpreting and taking appropriate decisions in solving real life problems using statistical tools.

	(b) use different Statistical packages for graphical interface, data analysis and interpretation. (c) write a systematic Statistical report.
Syllabus	Plan of the work, complete proposal of the project, review on the proposed project topic, seminar presentation and evaluation. Progression of project through methodology, progression of analysis and computation, preparation of complete project report, seminar presentation and evaluation.

The Optional Paper(s) which is/are to be offered in a particular semester will be fixed by the Departmental Academic Committee and will be announced before the commencement of the semester. No Course will be offered in both the semesters V and VI of an academic year, and nor in the semesters VII and VIII of an academic year.

Details of the Optional Papers for STAT303C11 of Semester V, and STAT354C15 of Semester VI:

Semester	**FIVE / SIX**
Paper Number	**Optional Paper (STAT303C11A / STAT354C15A)**
Paper Title	**Operations Research and Statistical Quality Control**
No. of Credits	**6**
No. of classes	Theory : 4 Practical : 4
Course Learning Outcomes	Students will acquire knowledge of (a) tackling various types of OR problems and mathematical formulation of linear programming problem (LPP). (b) graphical and simplex method of solving LPP for finding degenerate, unbounded, alternate and infeasible solutions. (c) concept of Artificial Variables. (d) concept of duality in LPP (e) Transportation Problem: Initial solution and the optimal solution, special cases. (f) assignment problem: Hungarian method to find optimal assignment, special cases of assignment problem. (g) Hungarian Method for solving assignment problems. (h) non-linear programming with different types of problems and solution methods. (i) application of statistics in industry. (j) various phases of SQC, quality of the manufactured items. (k) SQC approaches namely Control Charts in Process Control and Sampling Inspection Techniques in Product Control; measure of process capability.
Syllabus	**Unit 1** Introduction to Operations Research. Linear Programming Problem, Mathematical formulation of the L.P.P, graphical solution of a L.P.P. Simplex method for solving L.P.P.

	Artificial Variables - Charnes M method, Two-Phase Method. Introduction to the Concept of Duality in L.P.P. Transportation Problem: Methods of finding IBFS, MODI's method to find the optimal solution. Assignment problem: Hungarian method to find optimal assignment. **40L**
	Unit 2 Introduction to nonlinear programming. Unconstrained problems, problems with inequality and equality constraints, Fritz John and Karush-Kuhn-Tucker conditions. **16L**
	Unit 3 Definition, dimensions of quality. Quality system and standards: Introduction to ISO quality standards, Quality registration. Statistical Process Control - Seven tools of SPC, chance and assignable Causes of quality variation. Statistical Control Charts- Statistical basis and construction of 3-σ Control charts, Rational Sub-grouping and different control charts. Comparison between control charts for variables and control charts for attributes. Analysis of patterns on control chart. Estimation of process capability. **20L**
	Unit 4 Principle of acceptance sampling plans. Single and Double sampling plan their OC, AQL, LTPD, AOQ, AOQL, ASN, ATI functions with graphical interpretation, use and interpretation of Dodge and Romig's sampling inspection plan tables. **20L**
List of Practical(Using TORA/WINQSB/LINGO)	Mathematical formulation of L.P.P and solving the problem using graphical method, Simplex technique Charnes Big M method and Two-Phase Method. Degenerate solution Unbounded solution Alternate solution Infeasible solution Allocation problem using Transportation model. Allocation problem using Assignment model. Applications of KKT conditions. Construction and Interpretation of statistical control charts. X-bar & R chart, X-bar & s-chart, np- chart, p-chart, c-chart, u- chart. Single sample inspection plan: Construction and interpretation of OC, AQL, LTPD, ASN, ATI, AOQ, AOQL curves. Calculation of process capability and comparison of 3-sigma control limits with specification limits.
Reading/Reference Lists	Taha, H. A. (2007): Operations Research: An Introduction, 8 Hall of India. Hadley, G: (2002) : Linear Programming, Narosa Publications. Hillier, F.A and Lieberman, G.J. (2010): Introduction to Operations Research- Concepts and cases, 9th Edition, Tata McGraw Hill. Bazaraa M.S. et.al., Linear programming and Network flows. Bazaraa M.S. et.al., Non-linear programming: Theory and Algorithms. A.J.Duncan Quality Control and Industrial Statistics, 5th edition, McGraw-Hill Education. Montgomery, D.C. (2009): Introduction to Statistical Quality control, 6th edition, Wiley India, Pvt Ltd. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol 2, 8th edition, The world Press, Kolkata. Mukhopadhyay, P. (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied(P) Ltd.

Semester	**FIVE / SIX**
Paper Number	**Optional Paper (STAT303C11B / STAT354C15B)**
Paper Title	**Modern Survey Sampling**
No. of Credits	**6**
No. of Classes	Theory : 4 Practical : 4
Course Learning Outcomes	At the end of the course, a student will have knowledge on (a) notions of modern sampling design and unequal probability sampling (b) Horvitz-Thompson estimator (c) Systematic sampling using varying size-values, interpenetrating subsampling (d) The technique of probability proportional to aggregate size sampling (e) Single-stage and two-stage cluster sampling with unequal number of second stage units (f) Double sampling for stratification (g) Ranked set sampling and its uses (h) Comparison between different sampling approaches.
Syllabus	**Unit 1** Probability sampling from a finite population – notions of sampling design, sampling scheme, inclusion probabilities, Horvitz-Thompson estimator of a population total, basic sampling schemes, unequal probability sampling with and without replacement, systematic sampling, related estimators of population total/mean, their variances and variance estimators – mean per distinct unit in simple random with replacement sampling. Hansen-Hurwitz estimator in unequal probability sampling with replacement. Des Raj and Murthy's estimator (for sample of size two) in unequal probability sampling without replacement, unbiased ratio estimators – probability proportional to aggregate size sampling, Hartley-Ross **55L**
	Unit 2 Sampling and sub-sampling of clusters. Two-stage sampling with unequal number of second stage units and simple random sampling without replacement/unequal probability sampling with replacement at first stage, ratio estimation in two-stage sampling. **30L**
	Unit 3 Double sampling for stratification. Double sampling ratio and regression estimators. Sampling on successive occasions. **25L**
	Unit 4 Ranked Set Sampling : idea of the technique, framing unbiased estimator of population mean/total – its unbiasedness and sampling variance, estimation of the sampling variance. Comparison with SRS. **18L**
List of Practicals	Probability proportion to size with replacement sampling Probability proportion to size without replacement sampling Mean per distinct units in SRSWR Cluster sampling Two-stage sampling Ratio and Regression Estimation PPAS Sampling Double sampling Ranked set sampling
Reading/Reference Lists	Cochran, W. G. : Sampling Techniques (1977) Raj, D. & Chandhok, P. : Sample Survey Theory

	Murthy, M. N. : Sampling : Theory and Methods (1968) Chaudhuri, A. : Modern Survey Sampling (2014) Hedayat, A.S. and Sinha, B.K. (1991). Design and Inference in Finite Population Sampling, Wiley Inter-Science. Mukhopadhyay, P. (1996). Inferential Problems in Survey Sampling, New Age International (P) Ltd. 6. Muhopadhyay, P. (2007). Survey Sampling, Nerosa Publishing House, New Delhi. Chaudhuri, A. : Randomized Response and Indirect Questioning in Surveys (2011) Chen, Z., Bai, Z. and Sinha, B.K. : Ranked Set Sampling – Theory and Applications, Springer Bouza-Herrera, C.N. and Al-Omari, A.I.F. : Ranked Set Sampling -65 Years Improving the Accuracy in Data Gathering, Academic Press Paul S. Levy and Stanley Lemeshow (2013): Sampling of populations : methods and applications , 4th Edition (Wiley).

Semester	**FIVE / SIX**
Paper Number	**Optional Paper (STAT303C11C / STAT354C15C)**
Paper Title	**Advanced Mathematical Analysis**
No. of Credits	**6**
No. of classes	Theory : 5 Tutorial : 1
Course Learning Outcomes	Students will acquire knowledge of (a) Sequence and Series of functions. (b) Multivariable Calculus (c) Metric spaces. (d) Group theory and its properties, ring, field.
Syllabus	**Unit 1** Sequence and series of functions: Pointwise & Uniform convergence. Simple tests, Consequences of Uniformly Convergence.Power series. **15L**
	Unit 2 Metric Spaces: Definition, Examples. Some Topological Aspects. Convergent sequences and Cauchy sequences and the notion of Completeness. Compactness, First and Second Countability, Separability and Lindelof properties, Connectedness, Continuity and Uniform Continuity. **35L**
	Unit 3 Partial Derivative, Directional Derivative, and Total Derivative. Maxima and Minima of Multivariable functions. Multivariate Taylor Series Expansion. **15L**
	Unit 4 Introduction to Group Theory: Definition, Elementary properties using definition, integral powers of elements, Subgroups, Cyclic group, Groups of Permutations. Definition of Ring, Special types of Rings: Integral Domain, Field, elementary results. **31L**

List of Practical	Only Tutorials.
Reading/Reference Lists	W. Rudin : Principles of Mathematical Analysis. G.F. Simmons: Introduction to Topology and Modern Analysis. S. Kumaresan : Topology of Metric Spaces. S. Shirali and H.L.Vasudeva : Metric Spaces. A. Chakraborty : Metric Space. Tom M Apostol: Mathematical Analysis. I.N. Herstein : Topics in Algebra. Sen, Ghosh, Mukhopadhay : Topics in Abstract Algebra. M. Artin : Algebra.

Semester	**FIVE / SIX**
Paper Number	**Optional Paper (STAT303C11D / STAT354C15D)**
Paper Title	**Demography and Actuarial Statistics**
No. of Credits	**6**
No. of classes	Theory : 4 Tutorial : 4
Course Learning Outcomes	Students will acquire knowledge of (a) Death rates and birth rates (b) Life table and its characteristics (c) Concept of population growth (d) Basic idea of actuarial modelling (e) Concepts of probability related to insurance (f) Idea of premium calculation (g) Basic idea of life insurance and life annuities
Syllabus	**Unit 1:** Introduction and sources of collecting data on vital statistics, errors in census and registration data. Measurement of population, rate and ratio of vital events. Measurements of Mortality: Crude Death Rate (CDR), Specific Death Rate (SDR), Infant Mortality, Rate (IMR) and Standardized Death Rates. Stationary and Stable population, Central Mortality Rates and Force of Mortality. Life (Mortality) Tables: Assumption, description, construction of Life Tables and Uses of Life Tables. Measurements of Fertility: Crude Birth Rate (CBR), General Fertility Rate (GFR), Specific Fertility Rate (SFR) and Total Fertility Rate (TFR). **20L**
	Unit 2: Measurement of Population Growth Theory: Crude rates of natural increase, Pearl's Vital Index, Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR). Population

	Estimation, Projection and Forecasting: Use of A.P. and G.P. methods for population estimates, Fitting of Logistic curve for population forecasting using Rhode's method. **15L**
	Unit 3: Probability in Insurance: Loss distributions with and without risk sharing, compound distributions and their applications in risk modelling, Basic idea of Copulas, Introduction to extreme value theory and applications. Theory of Interests: Principles of compound interest. Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding. Utility theory: Utility functions, expected utility criterion, types of utility function, insurance and utility theory. Principles of Premium Calculation: Properties of premium principles, examples of premium principles. Individual risk models: models for individual claims, the sum of independent claims, approximations and their applications. 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. Multiple decrement models, deterministic and random survivorship groups, associated single decrement tables, central rates of multiple decrement, net single premiums and their numerical evaluations. Distribution of aggregate claims, compound Poisson distribution and its applications. **48L**
	Unit 4: 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. Life annuities: Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities-immediate and apportionable annuities-due. Net premiums: Continuous and discrete premiums, true monthly payment premiums, apportionable premiums, commutation functions, accumulation type benefits. Payment premiums, apportionable premiums, commutation functions, accumulation type benefits. Net premium reserves: Continuous and discrete net premium reserve, reserves on a semicontinuous basis, reserves based on true monthly premiums, reserves on an apportionable or discounted continuous basis, reserves at fractional durations, allocations of loss to policy years, recursive formulas and differential equations for reserves, commutation functions. Some practical considerations: Premiums that include expenses-general expenses types of expenses, per policy expenses. Claim amount distributions, approximating the individual model, stop-loss insurance. **45L**
List of Practical	Topic related practicals.
Reading/Reference Lists	Dickson, C. M. D. (2005): Insurance Risk And Ruin (International Series On Actuarial Science), Cambridge University Press. Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A. And Nesbitt,C. J. (1997): Actuarial Mathematics, Society Of Actuaries, Itasca, Illinois,U.S.A. Atkinson, M.E. and Dickson, D.C.M. (2000). An Introduction to Actuarial Studies, Elgar Publishing. Bedford, T. and Cooke, R. (2001). Probabilistic risk analysis, Cambridge. Actuarial Models (The Mathematics of Insurance) by Vladimir I. Rotar, Chapman and Hall (CRC). Fundamentals of Actuarial Mathematics by S. David Promislow , Wiley.

	Neill, A. (1977). Life Contingencies, Heineman. Philip, M. et. al (1999). Modern Actuarial Theory and Practice, Chapman and Hall. Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1998). Stochastic Processes for Insurance and Finance, Wiley.

Details of the Optional Papers for **STAT403C18** of Semester VII, and **STAT452C21, STAT453C22** of Semester VIII:

Semester	**SEVEN / EIGHT**
Paper Number	**Optional Paper (STAT403C18A/STAT452C21A/STAT453C22A)**
Paper Title	**Survival Analysis and Clinical Trials**
No. of Credits	**4**
No. of Classes	Theory : 3 Practical : 2
Course Learning Outcomes	At the end of the course, a student will have knowledge on (a)
Syllabus	**Unit 1** Survival Analysis: Functions of survival times, survival distributions and their applications, exponential, gamma, Weibull, Rayleigh, lognormal, death density function for a distribution having bath-tub shaped hazard function. Censoring Schemes: Type I, Type II and progressive or random censoring with biological examples. Estimation of mean survival time and variance of the estimator for Type I and Type II censored data with numerical examples. Non-parametric methods: Actuarial and Kaplan-Meier methods for estimating survival function and variance of the Estimator. **30L**
	Unit 2 Competing Risk Theory: Indices for measurement of probability of death under competing risks and their inter-relations. Estimation of probabilities of death using maximum likelihood principle and modified minimum Chi-square methods. Theory of independent and dependent risks. Bivariate normal dependent risk model. **15L**
	Unit 3 What is clinical trial? Different phases; Major steps of executing a controlled clinical trial; Type of control groups; Blinding; Bias; Ethics of randomization. Determination of trial size; Randomized clinical trial; Balancing treatment assignments; Complete and restricted randomization; Random allocation rule; Truncated binomial design, Efron's biased coin design. **20L**
	Unit 4 Discussion on covariate-adaptive and response-adaptive randomization procedures with examples. **15L**
List of Practicals	*The entire practical are to be done preferably by using R/ statistical packages.* To estimate survival function To determine death density function and hazard function To identify type of censoring and to estimate survival time for type I censored data

	To identify type of censoring and to estimate survival time for type II censored data To identify type of censoring and to estimate survival time for progressively type I censored data Estimation of mean survival time and variance of the estimator for type I censored data Estimation of mean survival time and variance of the estimator for type II censored data Estimation of mean survival time and variance of the estimator for progressively type I censored data To estimate the survival function and variance of the estimator using Nonparametric methods with Actuarial methods To estimate the survival function and variance of the estimator using Nonparametric methods with Kaplan-Meier method To estimate Crude probability of death To estimate Net-type I probability of death To estimate Net-type II probability of death To estimate partially crude probability of death To simulate the random sequence of treatment assignments. To plot the probability of imbalance. To simulate the treatment allocation ratio.
Reading/Reference Lists	Lee, E.T. and Wang, J.W. (2003): Statistical Methods for Survival data Analysis, 3rd Edition, John Wiley and Sons. Biswas, S. (2007): Applied Stochastic Processes: A Biostatistical and Population Oriented Approach, Reprinted 2ndCentral Edition, New Central Book Agency. Chiang, C.L. (1968): Introduction to Stochastic Processes in Bio Statistics JohnWiley and Sons. Indrayan, A. (2008): Medical Biostatistics, 2nd Edition Chapman and Hall/CRC. Rosenberger and Lachin: Randomized Clinical Trials: Theory and Practice Ding-Geng (Din) Chen and Karl E. Peace: Clinical Trial Data Analysis Using R

Semester	**SEVEN / EIGHT**
Paper Number	**Optional Paper (STAT403C18B/STAT452C21B/STAT453C22B)**
Paper Title	**Statistics in Finance**
No. of Credits	**4**
No. of Classes	Theory : 3 Practical : 2
Course Learning Outcomes	At the end of the course, a student will have knowledge on (a) Derivative Markets (b) Random walks, Brownian Motion and Martingales. (c) Option Pricing. (d) Portfolio Optimization.
Syllabus	**Unit 1** Introduction to derivatives: Forward contracts, spot price, forward price, future price. Options, zero-coupon bonds and discount bonds.

	15L
	Unit 2 Introduction to Random walk, Brownian Motion and Martingales. **25L**
	Unit 3 Option pricing: One-step and Two-step Binomial Models, General Binomial Tree Model, Black-Scholes formula, Implied Volatility, Properties of the European Call Option, Put Options, Put-call parity for European options **20L**
	Unit 4 Portfolio Optimization: Efficient frontier and Tangency portfolio - Efficient portfolio with N risky assets and one risk-free asset, Notion of VaR, Capital Asset Pricing Model, Capital Market Line (CML) . Security Market Line (SML) - Security Characteristic Line (SCL), Testing for CAPM. **20L**
List of Practicals	Problems based on Derivatives Finding European Call/Put Option using One-Step and Two-Step Binomial Tree Models. Finding European Call/Put Option using Black Scholes Formula Finding an Efficient Portfolio of N risky assets and one risk free asset.
Reading/Reference Lists	Introduces Quantitative Finance – Paul Wilmott Options, Futures and other derivatives – John C Hull. An Elementary Introduction to Statistics in Finance – S M Ross Statistics in Finance – David Ruppert Statistics of Financial Markets – J Franker, C M Hafner Stochastic Processes-S.M.Ross

Semester	**SEVEN / EIGHT**
Paper Number	**Optional Paper (STAT403C18C/STAT452C21C/STAT453C22C)**
Paper Title	**Graph Theory and its Applications**
No. of Credits	**4**
No. of Classes	Theory : 3 Tutorial : 1
Course Learning Outcomes	At the end of the course, a student will have knowledge on (a) Basic notions and real-world applications of Graph Theory. (b) Spanning Trees and their applications (c) The Four-Color Problem. (d) Applications in Game theory, Network Flows and Markov Processes.
Syllabus	**Unit 1** Introduction to Graph theory and its applications. Definitions of Graph, Finite and Infinite Graphs and other elementary notions. Connectedness, Euler Graphs, Hamiltonian Paths and Circuits, The Travelling Salesman Problem.

	20L
	Unit 2 Trees, Spanning Trees, Cutsets, Connectivity and Separability, Network Flows. **16L**
	Unit 3 Matrix Representation of Graphs, An Application to a Switching Network, Coloring, Covering and Partitioning, The Four-Color Problem. **16L**
	Unit 4 Graph theory in Operations Research and Markov Processes. **12L**
List of Practicals	Tutorials only
Reading/Reference Lists	Narsingh Deo, Graph Theory, with Applications to Engineering and Computer Science. ORE, O., Theory of Graphs, American Mathematical Society, Providence, R.I., 1962. BECKENBACH, E. F., "Network Flow Problems," Chapter 12 in Applied Combinatorial Mathematics (E. F. Beckenbach, ed.), John Wiley & Sons, Inc., New York, 1964. IRI, M., Network Flow, Transportation and Scheduling, Academic Press, Inc., New York, 1969. BERGE, C., The Theory of Graphs and Its Applications, John Wiley & Sons, Inc., New York, 1962. BUSACKER, R. G., and T. L. SAATY, Finite Graphs and Networks: An Introduction with Applications, McGraw-Hill Book Company, New York, 1965. HARARY, F., Graph Theory, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1969.

Semester	**SEVEN / EIGHT**
Paper Number	**Optional Paper (STAT403C18D/STAT452C21D/STAT453C22D)**
Paper Title	**Game Theory**
No. of Credits	**4**
No. of Classes	Theory : 3 Practical : 1
Course Learning Outcomes	At the end of the course, a student will have knowledge on (a) basic notions of game theory (b) Mixed Strategies and the value of the game. (c) Symmetric Games (d) Graphical Solution of 2 x m and m x 2 games (e) Solving a Game problem using Simplex method (f) Two Person Nonzero sum games (g) Cooperative Games (h) Applications of Game theory

Syllabus	**Unit 1** Introduction, some applications and examples, payoffs, strategies, pure strategies, a saddle point in pure strategies, The von Neumann minimax theorem, Kakutani's theorem (Statement only). Mixed Strategies, Expected payoff, a saddle point in mixed strategies, the value of the game. Dominated Strategies, Solving 2 x 2 games graphically, Graphical solution of 2 x m and m x 2 games, best response strategies. **24L**
	Unit 2 Invertible Matrix Games, Completely Mixed Games, Symmetric Games, Matrix Games and Linear Programming, setting up for the Linear program, Duality theorem (Statement only), A Direct formulation without transforming. A Game theory model of Economic growth. **20L**
	Unit 3 Two-Person Nonzero Sum Games, Nash Equilibrium, Prisoner's Dilemma, The Arms Race, 2 x 2 Bimatrix Games, Interior Mixed Nash points by Calculus, Nonlinear Programming Method for Nonzero Sum Two-Person Games, Choosing among Several Nash Equilibria. **20L**
	Unit 4 Introduction to Cooperative Games, The Coalitions and Characteristic Functions, The Nucleolus, The Shapley Value, Bargaining. **16L**
List of Practicals	Problems based on game matrix. Graphical solution to mx2 / 2xn rectangular game. Mixed strategy. Transformation of a Game problem to an LPP. Solution of a Game by Simplex Method. Solution of a Game by Matrix Method. Problems on Two-Person Nonzero Sum Games
Reading/Reference Lists	E.N.Barron , Game theory: An Introduction, Wiley, 2007. Martin J. Osborne, An Introduction to Game Theory, Oxford University Press, New Delhi, 2004. Stef Tijs, Introduction to Game Theory, Hindustan Book Agency. Steven Tadelis, Game Theory: An Introduction, Princeton University Press.2013. Hamdy A Taha: Operations Research, Pearson India Education Ltd. J G Chakravorty, P R Ghosh: Linear Programming and Game theory

Semester	**SEVEN / EIGHT**
Paper Number	**Optional Paper (STAT403C18E/STAT452C21E/STAT453C22E)**
Paper Title	**Data Mining**
No. of Credits	**4**

No. of classes	Theory : 3 Practical : 2
Course Learning Outcomes	
Syllabus	**Unit 1:** Introduction to data mining: Overview and potential applications, Data wrangling using softwares, Recommender systems, Information Retrieval. Limitations of data mining. **10L**
	Unit 2: Basic concepts of convex functions and convex optimization (proofs excluded), Gradient descent method: Algorithm and rate of convergence, choice of step sizes. Concept of subgradients with examples, Proximal gradient descent method, stochastic gradient descent method. Concept of duality, KKT conditions, Co-ordinate descent method. **20L**
	Unit 3: Brief review of basic regression and classification techniques, Issues with Classification and Regression Trees. Bagging and Random forest, Boosting: ADABOOST and connection with forward stepwise modelling, Gradient Boosting and applications, general idea of ensemble learning. Review of linear classification, concept of margin, support vector machines, random kitchen sinks. Artificial neural networks: basic idea, fitting by backpropagation algorithm and other relevant issues. **20L**
	Unit 4: Deep feed-forward networks, Regularization for deep learning, Convolution networks, Recurrent neural networks and variants, Concept of modules, Reinforcement learning, Deep graphical models. Concept of transfer learning, Deep unsupervised learning: Autoencoders, Variational Autoencoders, Generative Adverserial networks, Boltzman machines. Theory of deep learning. **30L**
List of Practical	Topic related practicals
Reading/Reference Lists	D. J. Hand, Heikki Mannila and Padhric Smyth, Principles of Data Mining. Cambridge, Massachusetts: MIT Press, 2001 Richard A. Berk, Statistical Learning from a Regression Perspective (3rd edition, New York: Springer, 2020) Jerome H. Friedman, Robert Tibshirani, and Trevor Hastie, Elements of Statistical Learning, Springer Gareth James et. al., An Introduction to Statistical Learning with Applications in R, Springer. Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT PressAston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, Dive Into Deep Learning, Lieven Vandenberghe, Stephen Boyd, and Stephen P. Boyd, Convex Optimization, Cambridge University Press.

Semester	**SEVEN / EIGHT**
Paper Number	**Optional Paper (STAT403C18F/STAT452C21F/STAT453C22F)**
Paper Title	**Advanced Design of Experiments**
No. of Credits	**4**
No. of Classes	Theory : 3 Practical : 2
Course Learning Outcomes	At the end of the course, a student will have knowledge on (a) notions of advances design of experiments (b) balanced incomplete block designs. (c) construction of PBIBD (d) idea and technique of fractional designs (e) arrays and their constructions (f) response surface designs
Syllabus	**Unit 1** Incomplete block designs : PBIBD : Properties, analysis. Dual and linked block designs. **15L**
	Unit 2 Construction of BIBD, PBIB designs. Lattice Designs. **30L**
	Unit 3 Symmetrical factorial experiments, construction and analyses of confounded designs. Fractional factorials. Orthogonal arrays and balanced arrays. **20L**
	Unit 4 Response surface methodology – orthogonality, rotatability and blocking. **15L**
List of Practicals	Analysis of data arising from partially balanced incomplete block designs lattice design fractional design cross over design Construction of BIBD PBIBD symmetric and asymmetric arrays
Reading/Reference Lists	Chakrabarti, M.C. (1962). Mathematics of Design and Analysis of experiments, Asia Publishing House. Cornell, John A. (2002). Experiments with Mixtures, John Wiley & Sons. Das, M. N. and Giri, N. C. (1986). Design and Analysis of Experiments, Wiley Eastern Limited. Dey, A. (1986). Theory of Block Designs, John Wiley & Sons. Dey, A. and Mukerjee, R. (1999). Fractional Factorial Plans, John Wiley & Sons. Hedayat, A. S., Sloane, N. J.A. and Stufken, J. (1999). Orthogonal Arrays: Theory and

	Applications, Springer. Hinkelmann, K. and Kempthorne, O. (2005). Design and Analysis of Experiments, Vol. 2: Advanced Experimental Design, John Wiley & Sons. Jones, B. and Kenward, M.G. (2003). Design and Analysis of Cross-over Trials. Chapman & Hall/CRC Press. Montgomery, D. C. (2005). Design and Analysis of Experiments, Sixth Edition, John Wiley & Sons. Myers, R. H. and Montgomery, D. C. (2002). Response Surface Methodology: Process and Product Optimization using Designed Experiments, John Wiley & Sons. Raghavarao, D. (1970). Construction and Combinatorial Problems in Design of Experiments, John Wiley & Sons. Wu, C. F. J. and Hamada, M. (2000). Experiments: Planning, Analysis and Parameter Design Optimization, John Wiley & Sons.

Skill Enhancement Elective Courses (SEC) Papers in Statistics Honours

Semester	**THREE**
Paper Number	**STAT241SEC01**
Paper Title	**Statistical Computing using R and Excel**
No. of Credits	**4**
No. of Classes	Sessional : 5
Course Learning Outcomes	Students will acquire knowledge of (a) use of Excel, diagrammatic representation, summary measures, logical commands like IF, AND, NOT, OR. (b) basic objects of R (c) functional programming using R language (d) data wrangling and visualisation using R (e) basic statistical methods using R (f) Inverse transform method for simulating various probability distributions and stochastic models. (g) data base management system with special emphasis on significance of topic to the statisticians. (h) Entity relationship, Relational, Hierarchical and Network Models. (i) Generation of reports using Latex: Suggested Editors – Lyx/ Kile/ Texnic-center.
Syllabus	*This course will review and expand upon core topics in statistics and probability, particularly by initiating the beneficiaries of the course to use of spreadsheet and R.* **Unit 1:** Use of Excel: Creating grouped frequency distribution, different diagrammatic representations, Data manipulation : Subsetting a data, sorting, searching and creating new variables, Basic summary measures, Linear regression. Logical commands: IF, AND, NOT, OR etc. Generation of reports using Latex: Suggested Editors – Lyx/ Kile/ Texnic-center. Use of R package Knitr/ Markdown to produce reports, Case study using any inbuilt or external dataset to understand and apply the statistical techniques discussed in R/Excel and prepare a report. **30L**
	Unit 2: Introduction to open source philosophy, Introduction to R, Use of R as a calculator: Arithmetic operations, Use of parentheses, power operation, Quotient and remainder operations. Standard

	inbuilt functions, Different types of numbers in R: Division by zero leading to Inf or -Inf. NaN. NA, Use of R scripts, Concept of R libraries. Variables in R. Basic data structures: vectors, matrices, lists, arrays and data frames in R. Basic plotting in R: plot(), histogram, barplot, boxplot, points, lines, segments, arrows, paste inserting mathematical symbols in a plot, pie diagram, setting graphical parameters from par(), Basic summary statistics, Usual tests of significance and confidence intervals, Linear regression: Estimation, finding predicted values, plotting the regression line on scatterplot. Standard functions, e.g., sin, cos, exp, log., Different types of numbers in R: Division by zero leading to Infor -Inf. NaN. NA. Use of R scripts, R libraries: what is an r library?, how to load and use a library how to get help- documentation and vignettes. Some useful inbuilt functions: getwd(), setwd(), source(). **16L**
	Unit 3: Use of if and ifelse, Loops in R, writing functions in R, setting default values of arguments of a function. Debugging and testing, checking compatibility of arguments in function and print error/warning messages. Advance data manipulation in R: Reading and writing non-R formats. Importing data from the Web, Selective access to data, applying the same function to all parts of a data object. Transforming the data, merging dataframes, reshaping dataframes from wide to long or long to wide.Using "apply" family of functions, Split-apply-combine technique in R, Use of plyr and dplyr functions, Use of Tidyverse, Basic concepts of relational databases; how a database is like an R dataframe. The client/server model. The structured query language (SQL) and queries; SELECT and JOIN. R/SQL translations. Accessing databases through R. **44L**
	Unit 4: Generation of random samples from univariate discrete and continuous probability distributions, cdf inversion method, box-muller transformation. Simulation of random variables from mixture distribution, simulating bivariate normal random variable (using conditional approach), Acceptance rejection sampling, Simulating random experiments like coin tossing, rolling of a die, card shuffling to illustrate probabilities of different events, Monte Carlo integration and variance reduction techniques, Approximating the expectation of a given function of a random variable using simulation. Graphical demonstration of the Law of Large Numbers and Central limit theorem. Using simulation to compute the level of significance, power, critical value and p-value of tests. **38L**
List of Practical	Topic related practicals
Reading/Reference Lists	The R Cookbook, by Paul Teetor. The R Graphics Cookbook, by Winston Chang. Data Manipulation with R, by Phil Spector. The R Inferno, by Patrick Burns (freely available at http://www.burns-stat.com/pages/ Tutor/ R_inferno.pdf) Simple R, by John Verzani (freely available at https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf). Quick R (freely available at https://www.statmethods.net/).

Semester	**FOUR**
Paper Number	**STAT291SEC02**

Paper Title	**Statistical Computing Using C**
No. of Credits	**5**
No. of Classes	Sessional : 6
Course Learning Outcomes	Students will acquire knowledge of (a) the basic structure of C programming language with different data types. (b) Operations, loop structures and their uses. (c) conditional statements, arrays and their uses. (d) Functions, multi-function program, processor, pointer. (e) the basic ways of handling file in C, usage of C programming in some selected areas of Statistics. (f) Preliminaries of numerical analysis – interpolation, numerical integration and solution of transcendental equations. (g) Use of C programmes to perform basic statistical analysis on several real datasets.
Syllabus	**Unit 1** Components, basic structure programming, character set, C/C++ tokens, Keywords and Identifiers and execution of a C/C++ program. Data types: Basic data types, Enumerated data types, derived data types. Constants and variables: declaration and assignment of variables, Symbolic Constants, overflow and underflow of data. 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. Decision making and branching - if…else, nesting of if…else, else if ladder, switch, conditional (?) operator. Looping in C/C++: for, nested for, while, do…while, and jumps in and out of loops. 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 only). **30L**
	Unit 2 User-defined functions: A multi-function program using user-defined functions, definition of functions, return values and their types, function prototypes and calls. Category of Functions : no arguments and no return values, arguments but no return values , arguments with return values, no arguments but returns a value, functions that return multiple values. Recursion function. Passing arrays to functions, Storage class of Variables. Pointers: Declaration and initialization of pointer variables, accessing the address of a variable, accessing a variable through its pointer, pointer expressions, pointer increments/decrement and scale factor. Pointers and arrays, arrays of pointers, pointers as function arguments, functions returning pointers Structure: Definition and declaring, initialization, accessing structure members, copying and comparison of structure variables, array of structures, structure pointers. Dynamic memory allocation functions: malloc, calloc and free. Pre-processors: Macro substitution, macro with argument File inclusion in C/C++: Defining and opening a file (only r, w and a modes), closing a file, I/O operations on files-fscanf and fprintf functions. Introduction to statistical libraries. **30L**
	Unit 3 Drawing of random sample from standard univariate discrete and continuous distributions, cdf inversion method, box-muller transformation, polar transformation. Drawing of random samples from mixture distribution and bivariate normal (conditional distribution approach).

	Acceptance rejection sampling. Monte Carlo Integration, Variance Reduction techniques. **20L**
	Unit 4 Numerical Analysis: Approximate Numbers, Significant Figures, Rounding of Numbers, Absolute Error and Relative Error. Errors in arithmetic computations. Polynomials and Difference Tables. Approximation of functions and Weierstrass Theorem (statement). Lagrange and Newton formulae for Interpolation. Trapezoidal and Simpson's 1/3 Rules for approximations of definite integrals. Approximate solutions of Numerical Equations by Fixed-point Iteration and Newton-Raphson methods. Conditions of convergence. **33L**
List of Practicals	Roots of a quadratic equation (with imaginary roots also). Sorting of an array and hence finding median. Mean, Median and Mode of a Grouped Frequency Data. Variance and coefficient of variation of a Grouped Frequency Data. Preparing a frequency table. Value of n factorial using recursion. Random number generation from uniform, exponential, calculate sample mean and variance and compare with population parameters. Matrix addition, subtraction, multiplication, Transpose and Trace. Fitting of Binomial, Poisson distribution. Compute ranks and then calculate rank correlation (without tied ranks). Fitting of lines of regression. Numerical methods: Solving one-variable equations using Newton-Raphson method. Trapezoidal rule for numerical integration. Solving a linear system of equation. Generation of random samples from standard discrete and continuous distributions. Generation of random samples from mixture distributions. Generation of random samples from bivariate normal distribution. General of random samples by acceptance rejection method. Monte Carlo integration and related techniques.
Reading/Reference Lists	Kernighan, B.W. and Ritchie, D.(1988): C Programming Language,2ndEdition, Prentice Hall. Balagurusamy, E. (2011): Programming in ANSI C, 6th Edition Tata McGraw Hill. Ross, S: Simulation. Scarborough, J.B. (1966): Numerical Mathematical Analysis. Oxford and IBH Publishing. Mollah, S. A. : Numerical Analysis & Computational Procedures. Atkinson K. : Elementary Numerical Analysis. Sastry S.S.: Introductory Methods of Numerical Analysis. Hildebrand F.B. : Introduction to Numerical Analysis.

Details of Minor Elective Papers (MC)

Semester	**ONE**
Paper Number	**STAT104MC01**
Paper Title	**Statistical Methods**
No. of Credits	**6**
No. of classes	Theory : 4 Practical : 4
Course Learning	This course will make the students conversant with

Outcomes	(a) various techniques used in summarization, presentation and analysis of different types of Statistical data. (b) various summary measures of central tendency, dispersion, moments, skewness and 47 kurtosis. (c) simple and rank correlation, Partial and Multiple correlation coefficients. (d) fitting of linear and quadratic regressions using principle of least squares. (e) measures of association for 2×2 and r×s contingency tables.
Syllabus	**Unit 1** Introduction: Definition and scope of Statistics, concepts of statistical population and sample. Data: quantitative and qualitative, attributes, variables, scales of measurement - nominal, ordinal, interval and ratio. Presentation: tabular and graphic, including histogram and ogives. **35L**
	Unit 2 Measures of Central Tendency: mathematical and positional. Measures of Dispersion: range, quartile deviation, mean deviation, standard deviation, coefficient of variation, moments, skewness and kurtosis. **40L**
	Unit 3 Bivariate data: Definition, scatter diagram, simple, partial and multiple correlation (3 variables only), rank correlation. Simple linear regression, principle of least squares and fitting of polynomials and exponential curves. **30L**
	Unit 4 Theory of attributes, consistency of data, independence and association of attributes, measures of association and contingency. **23L**
List of Practical	Graphical representation of data Problems based on measures of central tendency Problems based on measures of dispersion Problems based on combined mean and variance and coefficient of variation Problems based on moments, skewness and kurtosis Fitting of polynomials, exponential curves Karl Pearson correlation coefficient Partial and multiple correlations Spearman rank correlation with and without ties. Correlation coefficient for a bivariate frequency distribution Lines of regression, angle between lines and estimated values of variables. Checking consistency of data and finding association among attributes.
Reading/Reference Lists	Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I& II, 8th Edn. The World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I and II, Tata McGraw HillPub. Co. Ltd. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co. Ltd.

Semester	**TWO**

Paper Number	**STAT154MC02**
Paper Title	**Introductory Probability**
No. of Credits	**6**
No. of classes	Theory: 4 Practical: 4
Course Learning Outcomes	This course will make the students conversant with (a) probability theory and Statistical modelling of outcomes of real life random experiments through various Statistical distributions. (b) writing of sample space, events and algebra of events and finding Probability of events. (c) conditional Probability and applications of Theorem of total probability and Bayes' theorem. (d) discrete and continuous Random Variables, Probability mass function (p.m.f.) and Probability density function (p.d.f.), Cumulative distribution function (c.d.f.). (e) Expectation, variance, moments and moment generating function. (f) problem solving pertaining to binomial, Poisson, geometric, uniform, normal and exponential distributions. (g) fitting of Binomial, Poisson and Normal distributions. (h) various aspects as outlined above through practical assignments.
Syllabus	**Unit 1** Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability – classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications. **38L**
	Unit 2 Random Variables: Discrete and continuous random variables, p.m.f., p.d.f. ,c.d.f. Illustrations of random variables and its properties. Expectation, variance, moments and moment generating function. **30L**
	Unit 3 Convergence in probability, almost sure convergence, Chebyshev's inequality, weak law of large numbers, De-Moivre Laplace and Lindeberg-Levy Central Limit Theorem (C.L.T.). **20L**
	Unit 4 Standard probability distributions: Binomial, Poisson, geometric, negative binomial, hypergeometric, uniform, normal, exponential, beta, gamma. **40L**
List of Practical	Fitting of binomial distributions for n and p = q = ½ given Fitting of binomial distributions for n and p given Fitting of binomial distributions computing mean and variance Fitting of Poisson distributions for given value of lambda

	Fitting of Poisson distributions after computing mean Application problems based on binomial distribution Application problems based on Poisson distribution Problems based on area property of normal distribution To find the ordinate for a given area for normal distribution Application based problems using normal distribution Fitting of normal distribution when parameters are given Fitting of normal distribution when parameters are not given
Reading/Reference Lists	Goon A.M., Gupta M.K. and Dasgupta B. (2002): An Outline of Statistical Theory, Vol. I , The World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I and II, Tata McGraw Hill Pub. Co. Ltd. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia. Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delh

Semester	**THREE**
Paper Number	**STAT205MC03**
Paper Title	**Basics of Statistical Inference**
No. of Credits	**6**
No. of classes	Theory: 4 Practical: 4
Course Learning Outcomes	This course will make the students conversant with (a) estimation of parameters using moments and maximum likelihood. (b) exact tests and confidence intervals related to univariate and bivariate normal populations, types critical regions of a given test-statistic. (c) ideas of large sample distributions and large sample tests related to sample proportion, sample mean, sample correlation etc. and applications. (d) nonparametric methods of inference including (i) ordinary sign test and Wilcoxon signed-rank test for single-sample and paired sample data, (ii) Mann-Whitney test for two-sample problem, (iii) Kolmogorov-Smironov Approach. (e) Pearsonian chi-square tests and applications. (f) an outline of the methods of ANOVA under one-way and two-way classified data. (g) necessity for test for normality and Levin's test for homoscedasticity. (h) Kruskal-Wallis test.
Syllabus	**Unit 1**

	Estimation of population mean, confidence intervals for the parameters of a normal distribution (one sample and two sample problems). The basic idea of significance test. Null and alternative hypothesis. Type I & Type II errors, level of significance, concept of p-value. Tests of hypotheses for the parameters of a normal distribution (one sample and two sample problems). **40L**
	Unit 2 Categorical data: Tests of proportions, tests of association and goodness-of-fit using Chi-square test, Yates' correction. **25L**
	Unit 3 Tests for the significance of correlation coefficient. Sign test for median, Sign test for symmetry, Wilcoxon two-sample test. **20L**
	Unit 4 Analysis of variance, one-way and two-way classification. Brief exposure of three basic principles of design of experiments, treatment, plot and block. Analysis of completely randomized design randomized complete block design. **43L**
List of Practical	Estimators of population mean. Confidence interval for the parameters of a normal distribution (one sample and two sample problems). Tests of hypotheses for the parameters of a normal distribution (one sample and two sample problems). Chi-square test of proportions. Chi-square tests of association. Chi-square test of goodness-of-fit. Test for correlation coefficient. Sign test for median. Sign test for symmetry. Wilcoxon two-sample test. Analysis of Variance of a one way classified data Analysis of Variance of a two way classified data. Analysis of a CRD. Analysis of an RBD.
Reading/Reference Lists	Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I& II, 8th Edn. The World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I and II, Tata McGraw Hill Pub. Co. Ltd. Gupta, S. C. and Kapoor, V.K. (2008): Fundamentals Of Mathematical Statistics, Sultan Chand & Sons Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia. Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi

Semester	**FOUR**
Paper Number	**STAT255MC04**
Paper Title	**Applied Statistics**
No. of Credits	**6**
No. of classes	Theory : 4 Practical : 4
Course Learning Outcomes	This course will make the students conversant with (a) Ideas of trend, Fitting of linear, quadratic trend, exponential curve and plotting of trend values and comparing with given data graphically. (b) Ideas and measurement of seasonal indices (c) Construction of price and quantity index numbers by and comparison and interpretation. (d) Construction of wholesale price index numbers (e) Construction and interpretation of X bar & R-chart (f) Construction and interpretation p-chart (fixed sample size) and c-chart (g) Computation of measures of mortality and completion of life table (h) Computation of measures of fertility and population growth
Syllabus	**Unit 1** Economic Time Series: Components of time series, Decomposition of time seriesAdditive and multiplicative model with their merits and demerits, Illustrations of time series. Measurement of trend by method of free-hand curve, method of semi-averages and method of least squares (linear, quadratic and modified exponential). Measurement of seasonal variations by method of ratio to trend. **30L**
	Unit 2 Index numbers: Definition, Criteria for a good index number, different types of index numbers. Construction of index numbers of prices and quantities, consumer price index number. Uses and limitations of index numbers. **26L**
	Unit 3 Statistical Quality Control: Importance of statistical methods in industrial research and practice. Determination of tolerance limits. Causes of variations in quality: chance and assignable. General theory of control charts, process & product control, Control charts for variables: X- bar and R-charts. Control charts for attributes: p- and c-charts. **36L**
	Unit 4 Demographic Methods: Introduction, measurement of population, rates and ratios of vital events. Measurement of mortality: CDR, SDR (w.r.t. Age and sex), IMR, Standardized death rates. Life (mortality) tables: definition of its main functions and uses. Measurement of fertility and reproduction: CBR, GFR, and TFR. Measurement of population growth: GRR, NRR. **36L**
List of Practical	Measurement of trend: Fitting of linear, quadratic trend, exponential curve and plotting of trend values and comparing with given data graphically. Measurement of seasonal indices by Ratio-to-trend method and plotting of trend values and

	comparing with given data graphically. Construction of price and quantity index numbers by Laspeyres' formula, Paasche's formula, Marshall-Edgeworth's formula, Fisher's Formula - Comparison and interpretation. Construction of wholesale price index number, fixed base index number and consumer price index number with interpretation Construction and interpretation of X bar & R-chart Construction and interpretation p-chart (fixed sample size) and c-chart Computation of measures of mortality Completion of life table Computation of measures of fertility and population growth
Reading/Reference Lists	Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd. Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis, PHI. Hardle W. and Simar, L. Applied Multivariate Statistical Analysis. Kutner, M.H. et.al., Applied Linear Statistical Models. Belsley D.A. et.al., Regression Diagnostics. Draper N.R. and Smith, H. Applied Regression Analysis.

Semester	**SEVEN**
Paper Number	**STAT442MC05 (for B.Sc. Honours with Research)**
Paper Title	**Research Methodology for Statistics**
No. of Credits	**4**
No. of Classes	Theory : 3 Practical : 1
Course Learning Outcomes	Students will acquire knowledge of (a) providing scientific approaches to develop the domain of human knowledge largely through empirical studies. (b) understanding basic concepts and aspects related to research, data collection, analyses and interpretation. (c) preparing and finalizing research report on some real life situations.
Syllabus	**Unit 1:** Foundations of Research: Meaning, Objectives, Motivation, Utility. Concept of theory, empiricism, deductive and inductive theory. Characteristics of scientific method – Understanding the language of research – Concept, Construct, Definition, Variable. Research Process. Problem Identification & Formulation – Research Question – Investigation Question – Measurement Issues – Hypothesis – Qualities of a good Hypothesis –Null Hypothesis & Alternative Hypothesis. Hypothesis Testing – Logic & Importance. **24L**
	Unit 2: Research Design: Concept and Importance in Research – Features of a good research design – Exploratory Research Design – concept, types and uses, Descriptive Research Designs –

	concept, types and uses. Experimental Design: Concept of Independent & Dependent variables. Qualitative and Quantitative Research: Qualitative research – Quantitative research – Concept of measurement, causality, generalization, replication. Merging the two approaches. **24L**
	Unit 3: Preparing the text: Title, list of authors, abstract, introduction, materials and methodologies, results, acknowledgements, bibliography, cross-citation, cross-reference. Design effective tables, graphs. Submitting the manuscript: Rights and Permissions, Cover letter, Review process, publishing process. Writing review paper, book chapter, preparing a poster, conference report, preparing thesis, grant proposals and progress report. **40L**
	Unit 4: Review and summarize some published statistical works in journals. The students may work in groups and the list of publications may include some popular, foundational and classical works on the discipline. Each group need to prepare a report based in own language on the journal paper assigned to them by the instructor. **40L**
List of Practical	Topic related practicals.
Reading/Reference Lists	Robert A. Day and Barbara Gastel, How to Write and Publish a Scientific Paper M. Alley, The Craft of Scientific Writing Research Methodology – C.R.Kothari Wayne C. Booth, Gregoy G. Colomb and Joseph M. Williams, The Craft of Research (3rd edition, Chicago: University of Chicago Press, 2008. G. D. Gopen and J. A. Swan, "The Science of Scientific Writing", American Scientist **78** (1990): 550--558

Semester	**EIGHT**
Paper Number	**STAT492MC06 (for B.Sc. Honours with Research)**
Paper Title	**Research and Publication Ethics for Statisticians**
No. of Credits	**4**
No. of classes	Sessional
Course Learning Outcomes	
Syllabus	**Unit 1** Philosophy and ethics: Introduction: definition, nature and Scope, Concept, Branches, Ethics: definition, moral philosophy, nature of moral judgements and reaction. Scientific conduct: Ethics with respect to science and research, Intellectual honesty and research integrity, Scientific misconducts: Falsification, Fabrication, and Plagiarism (FFP), Redundant publications: duplicate and overlapping publications, salami slicing, Selective

	reporting and misrepresentation of data. **24L**
	Unit 2 Publication ethics: Definition, introduction and importance, Best practices /Standards setting initiatives and guidelines: COPE. WAME, etc., Conflicts of interest, Publication misconduct: definition, concept, problems that lead to unethical behavior and vice versa, types, Violation of publication ethics, authorship and contributorship, Identification of publication misconduct, complaints and appeals, Predatory publishers and journals. **24L**
	Unit 3 Open access publishing: Open access publications and initiatives, SHEERPA/RoMEO online resource to check publisher copyright & Self – archiving policies, Software tool to identify predatory publications developed by SPPU, Journal finder /Journal suggestion tools viz.JANE., Elsevier journal Finder, Springer Journal Suggester, etc. Publication misconduct: Group Discussions: Subject specific ethical issues, FFP, authorship, Conflicts of interest, Complaints and appeals: examples and fraud from India and abroad, Software tools Use of plagiarism software like Turnitin, Urkund and other open-source software tools. Databases and research metrics: Databases: Indexing, Citation: Web of Science, Scopus, etc., Research Metrics: Impact Factor of Journal as per Journal Citation Report, SNIP, SJR, IPP, Cite Score, Metrics: h-index, g index, i10 index, altmetrics. **32L**
	Unit 4 Professional Ethics for Statisticians: The ASA Ethical Guidelines for Statistical Practice, ACM Code of Ethics, The Data Science Ethics Checklist (DSEC) and Data Ethics Framework (DEFW). Data privacy: How to ask sensitive questions?, Warner's model, Different approaches of anonymity, Examples, Differential Privacy and applications. Selective inference and its impact on replicability of scientific results. **16L**
Reading/Reference Lists	Bird, A.(2006). Philosophy of Science. Routledge Maclntyre, Alasdair (1967) A Short History of Ethics. London P. Chaddah, (2018) Ethics in Competitive Research: Do not get Scooped; do not get Plagiarized. Ethical Practice of Statistics and Data Science, Rochelle E. Tractenberg, Ethics International Press Ltd, UK The Ethical Agorithm: The Science of Socially Aware Algorithm Design, Michael Kearns and Aaron Roth. D. J. Finney, "Ethical aspects of statistical practice", *Biometrics* **47** (1991): 331—339 Ethical Guidelines for Statistical Practice: The First 60 Years and Beyond Shelley Hurwitz and John S. Gardenier The American Statistician Vol. 66, No. 2. Ethics in statistical practice and communication: Five recommendations, Andrew Gelman (2018) Selective Inference: The Silent Killer of Replicability, Yoav Benjamini (2020). On Being a Scientist: A Guide to responsible conduct in Research: Third Edition, National Academies Press. Indian National Science Academy (INSA), Ethics in Science Education, Research and Governance (2019)

Value Added Course (VAC) Paper in Statistics Honours

Semester	**FOUR**
Paper Number	**STAT292VAC02**

Paper Title	**Data Analysis with Python**
No. of Credits	**3**
No. of classes	Sessional
Course Learning Outcomes	Students will be able to do (a) Programming in Python (b) Data input and output in Python (c) Implementing basic statistical methodologies in Python
Syllabus	**Unit 1:** Introduction to Python, Python installation, Scripting and Interacting, Setting up and customization of Jupyter. Variables and operators, reading data from keyboard, Data types - numbers, lists, dictionary, sets. Decision making Loops - for, while; Loop. control statements - break, continue, pass. Concept of functions, modules and packages. Functional programming: Iterators, List comprehensions and Generator expressions, Generators, functools, operator, itertools, functional modules. **12L**
	Unit 2: Introduction to Numpy: Matrices and Vectors, N- dimensional arrays, N-dimensional matrices, indexing and slicing, descriptive statistics, random number generators. Introduction to Pandas - Pandas Installation, data structures - series, data frame, panels. Reading from csv files, loc(), iloc() function, descriptive statistics, inserting columns into Data Frame, deleting columns from Data Frame, concatenating Data Frame, writing back to csv files, reading from excel files Merging or joining Data Frame Group by and aggregate functions Connecting Python to MySQL **12L**
	Unit 3: Introduction to SciPy: Matrix Algebra, Numerical Integration, Optimization, Statistical measures: Probability Distributions, Tests of hypothesis. Introduction to Statsmodels: Regression, Tests on categorical data. **12L**
	Unit 4: Graphics with Matplotlib, Data Visualization with pandas, Seaborn, Bokeh, Plotly. **12L**
List of Practical	Topic related practicals
Reading/Reference Lists	Python for Data Analysis by Wes McKinney, O'Reilly. Statistics and Data Visualisation with Python by Jesus Rogel-Salazar, CRC Press. Introduction to Statistics with Python by Thomas Haslwanter, Springer. Python for Probability, Statistics, and Machine Learning by José Unpingco, Springer. Think Stats: Exploratory Data Analysis by Allen Downey, O'Reilly.

Details of Multidisciplinary Courses (MDC)

Semester	**ONE**
Paper Number	**STAT105MDC01**

Paper Title	**Statistics for All**
No. of Credits	**3**
No. of classes	Theory : 2 Practical : 2
Course Learning Outcomes	This course will make the students conversant with (a) Understanding of statistical population, variables, collection of sample data on different types of variables and exploring the fundamental properties of data through summary measures. (b) Understanding paired data and multivariate (three-variable only for demonstration) data along with different types of association-measures and exploring cause-effect relation through the study of linear regression. (c) The basic concepts of statistical inference through some standard testing problems. (d) Handling comparatively little advanced but very practical analytical methods like ANOVA in both parametric and non-parametric scenario, survey-sampling approach, binary data regression analysis.
	Objective of the course is not to include any derivation of the theory but to demonstrate the theory by as many data-driven applications as possible supported by the use of software if required.
Syllabus	**Unit 1**. Understanding univariate data: Variable, notion of population and sample, different types of data, methods of collecting primary and secondary data, presentation of data, summary measures on data with central tendency (primarily arithmetic mean, median, mode), dispersion (primarily range, quartile deviation, standard deviation, coefficient of variation), ideas of skewness and kurtosis (through diagrams only). **16L**
	Unit 2. Understanding bivariate and multivariate data: Paired data and ideas (without mathematical details) of different measures of associations, primarily Pearson's correlation coefficient, Spearman's Rank correlation, measures of association of attributes through contingency table, two-variable linear regression and multiple (three-variable only) linear regression (without derivation of the regression coefficients' formulae). **16L**
	Unit 3. Statistical Inference (testing of hypothesis): Basic idea of normal population (primarily graphically, derivation of the properties excluded) Concepts of hypotheses, knowledge on test statistic and decision making in terms of critical value and *p*-value for some standard testing problems like test of mean based on single (normal) sample, test on comparing means based on two-sample, and paired sample data, etc. **16L**
	Unit 4. Miscellaneous discussion: Applications of one-way and two-way ANOVA (without derivation and details) assuming normality, Kruskal-Wallis test (without derivation and details), sample size determination, estimate of population mean and variability for finite population, idea and application of logistic regression for binary response data. **16L**
List of Practical	Meaures of mean, median, mode, range, QD, SD, CV for univariate data case Fitting of linear regression on bivariate and on three-variable multivariate data, measures of Pearson's correlation coefficients, Spearman's Rank correlation, measures

	of association of attributes through contingency table Tests of means for single sample, two-sample, and paired sample data on normal response using *p*-value approach Applications of ANOVA and Kruskal-Wallis test Sample size determination, estimate of population mean and variability for finite population Fitting of logistic regression for binary response data
Reading/Reference Lists	Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd. Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis, PHI. Hardle W. and Simar, L. Applied Multivariate Statistical Analysis. Kutner, M.H. et.al., Applied Linear Statistical Models. Belsley D.A. et.al., Regression Diagnostics. Draper N.R. and Smith, H. Applied Regression Analysis.

Semester	**TWO**
Paper Number	**STAT155MDC02**
Paper Title	**Statistics for Bioscience**
No. of Credits	**3**
No. of classes	Theory : 2 Practical : 2
Course Learning Outcomes	This course will make the students conversant with (a) Random variables and their probability distributions. (b) Understanding of statistical population, variables, collection of sample data on different types of variables and exploring the fundamental properties of data through summary measures, graphical presentations including Box-plot. (c) Understanding paired data and multivariate (three-variable only for demonstration) data along with different types of association-measures and exploring cause-effect relation through the study of linear, and nonlinear regressions including exponential, binary and count data regressions. (d) The basic concepts of statistical inference through methods of maximum likelihood, and some standard testing problems, handling comparatively little advanced but very practical analytical methods like ANOVA in both parametric and non-parametric scenario, standard non-parametric test procedures.
	Objective of the course is not to include any derivation of the theory but to demonstrate the theory by as many data-driven applications as possible supported by the use of software if required.
Syllabus	**Unit 1**. About probability and probability distributions: Definitions of probability, Theorem of total

	probability (statement only), conditional probability and Bayes Theorem (statement only), random variable and its probability distribution, probability models for frequently used discrete and continuous random variables, primarily Bernoulli, Poisson, Normal, Exponential etc. **20L**
	Unit 2. About univariate data: Variable, notion of population and sample, different types of data, Study design, graphical representation of data including Box plot, features of frequency distribution, summary measures of central tendency including quartiles, dispersion, ideas (primarily graphical) of skewness and kurtosis, outliers and extremes. **16L**
	Unit 3. About paired and multivariate data: Association and independence of the attributes through contingency table, correlations and linear regression in bivariate and multivariate set up (derivations excluded), exponential regression, applications of binary and count data distributions in biostatistical modelings. **12L**
	Unit 4. About statistical inference: Likelihood function and maximum likelihood estimate, $\chi^2-, t-, F-$tests as standard hypothesis testing problems under normality assumption, Sign-test, Mann-Whitney test, one-way and two-way ANOVA with normality assumption, Kruskal-Wallis test. **16L**
List of Practical	Measures of mean, median, mode, range, QD, SD, CV for univariate data case Fitting of linear and exponential regression on bivariate and on three-variable multivariate data, measures of Pearson's correlation coefficients, Spearman's Rank correlation, measures of association of attributes through contingency table Tests of means and variances for single sample, two-sample, and paired sample data on normal response using *p*-value approach, test of nullity of correlation for paired data with bivariate normality assumption Applications of ANOVA and Kruskal-Wallis test Fitting of regression for binary and count response data
Reading/Reference Lists	Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd. Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis, PHI. Hardle W. and Simar, L. Applied Multivariate Statistical Analysis. Kutner, M.H. et.al., Applied Linear Statistical Models. Belsley D.A. et.al., Regression Diagnostics. Draper N.R. and Smith, H. Applied Regression Analysis. Daniel, W.W. : Biostatistics – *A Foundation for Analysis in the Health Sciences*, Wiley Student Publication, 7th Edn.

Semester	**TWO**

Paper Number	**STAT156MDC03**
Paper Title	**Statistics for Practitioners**
No. of Credits	**3**
No. of classes	Theory : 2 Practical : 2
Course Learning Outcomes	This course will make the students conversant with (a) Random variables and their probability distributions including the idea of normal, χ^2, t and *F* distributions. (b) Understanding of statistical population, variables, collection of sample data on different types of variables and exploring the fundamental properties of data through summary measures, graphical presentations including Box-plot. (c) Understanding paired data and multivariate (three-variable only for demonstration) data along with different types of association-measures and exploring cause-effect relation through the study of linear, and nonlinear regressions including exponential, binary and count data regressions, goodness of fit measures, tackling multicollinearity through PCA and Factor analysis. (d) Tests of means and variances for single sample, two-sample, and paired sample data on normal response using *p*-value approach, test of nullity of correlation for paired data with bivariate normality assumption, handling comparatively little advanced but very practical analytical methods like ANOVA in both parametric and non-parametric scenario, standard non-parametric test procedures including test for normality.
	Objective of the course is not to include any derivation of the theory but to demonstrate the theory by as many data-driven applications as possible supported by the use of software if required.
Syllabus	**Unit 1**. Probability and probability distributions: Definitions, random variables and its probability distributions, probability models for discrete and continuous random variables, recognition of normal, χ^2, t and *F* distributions (without derivations and details). **8L**
	Unit 2. Univariate data and related measures: variables, classification of data, diagrammatic presentation of data, summary measures like mean, trimmed mean, median, mode, quartile deviation, standard deviation, coefficient of variation, idea (mainly graphical) of skewness and kurtosis, Box plot. **12L**
	Unit 3. Vector data and related measures: understanding bivariate and multivariate data, measures of correlation for numerical data, fitting of linear and polynomial regressions by least square method for numerical response data, concepts on R^2, adjusted $-R^2$criteria for goodness of fitted regression models, measures of association of attributes, Spearman's and Kendal's rank correlation, binary and count data regression model, applications of PCA and Factor Analysis as methods to tackle multicollinearity problem in regression. **30L**
	Unit 4. Basic inferential methods: Basic testing problems using normal-, $\chi^2-, t-$ and *F*- statistics, applications of ANOVA technique for one-way and two-way classified data, non-parametric methods of inference using Sign test, Mann-Whitney test, Kruskal-Wallis test, test of normality of the data. **14L**

List of Practical	Measures of mean, median, mode, range, QD, SD, CV for univariate data case Fitting of linear and polynomial regression on bivariate and on three-variable multivariate data, measures of Pearson's correlation coefficients, Spearman's and Kendall's Rank correlation, measures of association of attributes through contingency table Applications of PCA and Factor Analysis to tackle multicollinearity problem Tests of means and variances for single sample, two-sample, and paired sample data on normal response using *p*-value approach, test of nullity of correlation for paired data with bivariate normality assumption, test of normality Applications of ANOVA and Kruskal-Wallis test Fitting of regression for binary and count response data
Reading/Reference Lists	Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata. Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd. Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis, PHI. Hardle W. and Simar, L. Applied Multivariate Statistical Analysis. Kutner, M.H. et.al., Applied Linear Statistical Models. Belsley D.A. et.al., Regression Diagnostics. Draper N.R. and Smith, H. Applied Regression Analysis.

Details of Summer Internship

Semester	**FIVE**
Paper Number	**STAT341SI01**
Paper Title	**Summer Internship**
No. of Credits	**4**

Semesterwise Marks Distribution and Credit Allocation for 8 Semester Bachelor (Hons.) Programme under CBCS in B.Sc. Honours with Research, Department of Statistics, Presidency University, Kolkata

Sem	Course Type	Course Code	Course Name	Type of Paper	Credit				Marks		
					Theory	Practical	Tutorial	Total	Midterm	Endterm	Total
I	Major	STAT101C01	Descriptive Statistics	T	4	2		6	30	70	100
	Major	STAT102C02	Probability and Probability Distributions I	T	5		1	6	30	70	100
	AECC	103AECC01	English Communication/MIL		4			4			50/100
	Minor	STAT104MC01	Statistical Methods	T	4	2		6	30	70	100
	MDC	STAT105MDC01	Statistics for All	T	2	1		3	30	70	100
								25			500
II	Major	STAT151C03	Linear Algebra	T	4	2		6	30	70	100
	Major	STAT152C04	Probability and Probability Distributions II	T	4	2		6	30	70	100
	AECC	153AECC02	English Communication/MIL		4			4			50/100
	Minor	STAT154MC02	Introductory Probability	T	4	2		6	30	70	100
	MDC	STAT155MDC02	Statistics for Bioscience	T	2	1		3	30	70	100
	MDC	STAT156MDC03	Statistics for Practitioners	T	2	1		3	30	70	100
								28			600*
III	Major	STAT201C05	Mathematical Analysis and Calculus	T	5		1	6	30	70	100
	Major	STAT202C06	Inference I	T	4	2		6	30	70	100
	SEC (Major)	STAT241SEC01	Statistical Computing Using R and Excel	S				4		100	100
	VAC	ENVS204VAC01	Environmental Science		3			3			50/100
	Minor	STAT205MC03	Basics of Statistical Inference	T	4	2		6	30	70	100
								25			500
IV	Major	STAT251C07	Inference II	T	4	2		6	30	70	100
	Major	STAT252C08	Survey Sampling and Indian Official Statistics	T	4	2		6	30	70	100
	SEC (Major)	STAT291SEC02	Computing Using C	S				5		100	100
	VAC	STAT292VAC02	Data Analysis with Python	S				3		100	100
	Minor	STAT255MC04	Applied Statistics	T	4	2		6	30	70	100
								26			500*
V	Major	STAT301C09	Linear Models and ANOVA	T	4	2		6	30	70	

											100
	Major	STAT302C10	Multivariate Analysis	T	4	2		6	30	70	100
	Major	STAT303C11	Optional Paper	T				6	30	70	100
	Summer Internship	STAT341SI01		S				4		100	100
								22			400
VI	Major	STAT351C12	Advanced Statistical Methods	T	4	2		6	30	70	100
	Major	STAT352C13	Design of Experiments	T	4	2		6	30	70	100
	Major	STAT353C14	Inference III	T	4	2		6	30	70	100
	Major	STAT354C15	Optional Paper	T				6	30	70	100
								24			400
VII	Major	STAT401C16	Advanced Regression Analysis	T	3	1		4	30	70	100
	Major	STAT402C17	Stochastic Processes and Queuing Theory	T	3		1	4	30	70	100
	Major	STAT403C18	Optional Paper	T				4	30	70	100
	Major	STAT441C19	Project/ Dissertation	S				4		100	100
	Minor	**STAT442MC05**	**Research Methodology for Statistics**	**T**	**3**	1		**4**	**30**	**70**	**100**
								20			500
VIII	Major	STAT451C20	Time Series and Spatial Data Analysis	T	3	1		4	30	70	100
	Major	STAT452C21	Optional Paper	T				4	30	70	100
	Major	STAT453C22	Optional Paper	T				4	30	70	100
	Major	**STAT491C23**	**Project/ Dissertation**	S				**8**		**200**	**200**
	Minor	**STAT492MC06**	**Research and Publication Ethics for Statisticians**	S				**4**		**100**	**100**
								24			600

The courses marked with colour would be taken by students of other departments. T stands for taught papers and S stands for sessional papers.

***Total may vary depending on the marks of noncore papers.**