Khandesh Education, Society's

PRATAP COLLEGE (AUTONOMOUS), AMALNER

An "Autonomous College" Affiliated to
KBC North Maharashtra University, Jalgaon

"A+"Grade NAAC Re-Accredited (3rd cycle)

SYLLABUS

FOR

M.Sc. STATISTICS

(Semester I and II)

Under New Education Policy(NEP)
[w.e.f. Academic Year 2023-2024]

Summary of Distribution of Credits under NEP Scheme for M.Sc. (Statistics)

Sr. No.	Course Type	Sem I	Sem II	Total
01	DSC	14	14	28
02	DSE	04	04	08
03	RM	04	--	04
04	OJT/FP	--	04	04
05	RP	--	--	--
	Total Credits	**22**	**22**	**44**

Subject Type	DSC	DSE	RM	OJT/ FP	RP	Total
Credits	**28**	**08**	**04**	**04**	**--**	**44**

Total Credits = 44

KES's Pratap College (Autonomous) Amalner

SYLLABUS FOR
First Year of M.Sc. STATISTICS
(Semester-I and Semester-II)
With effect from July 2023

Syllabus Structure

Semester-I

Course Type	BOS Code	Title of the Course	Contact hours/week			Distribution of Marks for Examination						Credits
						Internal		External		Total		
			Th(L)	Pr	Total	Th	Pr	Th	Pr	Th	Pr	
DSC	ST-MJ-501	Distribution Theory	04	--	04	40	--	60	--	100	--	04
DSC	ST-MJ-502	Real Analysis	02	--	02	20	--	30	--	50	--	02
DSC	ST-MJ-503	Sampling theory and Statistics for National development	04	--	04	40	--	60	--	100	--	04
DSC	ST-MJP-501	Practical-I	--	08	08	--	40	--	60	--	100	04
DSE	ST-EC-521	R programing and Numerical Methods	04	--	04	40	--	60	--	100	--	04
	ST-EC-522	C++ programming										
Research	ST-RM-541	Research Methodology	04	--	04	40	--	60	--	100	--	04

Semester-II

Course Type	Course Code	Title of the Course	Contact hours/week			Distribution of Marks for Examination						Credits
						Internal		External		Total		
			Th(L)	Pr	Total	Th	Pr	Th	Pr	Th	Pr	
DSC	ST-MJ-551	Probability Theory	04	--	04	40	--	60	--	100	--	04
DSC	ST-MJ-552	Linear Algebra	02	--	02	20	--	30	--	50	--	02
DSC	ST-MJ-553	Linear Models and Regression Analysis	04	--	04	40	--	60	--	100	--	04
DSC	ST-MJP-554	Practical-II	--	08	08	--	40	--	60	--	100	04
DSE	ST-EC-571	Python Programming	04	--	04	40	--	60	--	100	--	04
	ST-EC-572	SAS Programming										
FP/OJT/RP	ST-OJT-591	On job training: Internship	04	--	04	40	--	60	--	100	--	04

Program Objectives for M.Sc. Program:

1. To impart the profound theoretical and practical knowledge of the specific science discipline along with the fundamental core concepts
2. To train the students to employ modern techniques, tools, methodologies, equipment, hardware/software etc. to perform objective oriented scientific and planned experiments
3. To groom the students for all-round development and mould them in a trained workforce to provide teaching-learning, research, business, professional supports in the various science disciplines
4. To make the student to develop the ability to think analytically, independently and draw logical conclusions to solve real-life problems.
5. To utilize the skills and knowledge gained through the subject to deal with real life situations and problems related to society, environment, research and development etc.

SEMESTER-I
ST-MJ-501 DISTRIBUTION THEORY

Course Objectives (CObs):

- To demonstrate an understanding of the basic concepts on probability and random variables.
- To formulate mathematical concepts on probability and probability distributions with practical applications.
- To determine marginal and conditional distributions from bivariate distributions. And to study related properties such as independence etc.
- To study functions of random variables and to determine their distributions using various techniques. To study general theory of sampling distributions of statistics and order statistics.

- **Brief review of basic distribution theory: (5 Marks)**
  - Random experiment and its sample space, events. (1L)
  - Probability axioms. (1L)
  - Random variables, Discrete random variables, Continuous random variables. (1L)
  - P.d.f., p.m.f., c.d.f. of random variables. (1L)
  - M.g.f., p.g.f., c.g.f., characteristic function of random variables. (1L)
  - Moments: raw moments, Central moments, Factorial moments. (1L)

- **Standard discrete and continuous distributions: (8 Marks)**
  - Bernoulli, Binomial, Geometric, Negative Binomial, Poisson, Hypergeometric distributions. (2L)
  - Exponential, Normal, Gamma, Beta, Uniform, Chi-square, Lognormal, Weibull, Cauchy distributions. (2L)
  - M.g.f, p.g.f., c.g.f., charactseristic function, Moments of above distributions. (2L)
  - Properties of above distributions. (2L)

- **Joint, Marginal and Conditional distributions: (10 Marks)**
  - Concept of random vectors, Joint, Marginal and conditional distributions Variance-covariance matrix. (1L)
  - Joint p.m.f. of discrete random variables, Joint p.d.f. of continuous random variables.(1L)
    - Marginal and conditional density using joint density. (1L)
    - Conditional expectation and variance. (1L)
    - Independence of random variables. (1L)
  - Bivariate normal distribution; Joint p.d.f. Marginal p.d.f.s, Conditional p.d.f., Joint m.g.f., Some properties. (2L)
    - Bivariate exponential distribution: joint p.d.f., Marginal p.d.f.s, properties. (1L)
  - Multivariate normal distribution: joint p.d.f., Marginal p.d.f., Conditional p.d.f., Joint m.g.f. (2L)
  - Multinomial distribution: joint p.m.f., Marginal p.m.f., Conditional p.m.f., Joint m.g.f. (2L)

- **Functions of random variables and their distributions: (10 Marks)**
  - Function of random variables. (1L)
  - Joint density of functions of random variables using Jacobian of transformation.(3L)
  - Convolution of random variables. (1L)

- **Compound, Truncated and Mixture Distributions: (3 Marks)**
  - Concept, applications, examples and problems. (3L)

- **Correlation: (3 Marks)**
  - Multiple and Partial Correlation. (2L)

- **Sampling Distributions: (6 Marks)**
  - Sampling distribution of statistics from univariate normal random samples. (2L)
  - Non-central Chi-square, t and F-distributions and their properties. (5L)

- **Quadratic forms under Normality: (6 Marks)**
  - Distribution of linear and quadratic forms in i.i.d. Standard normal variables (Technique based on m.g.f.). (2L)
  - Independence of two linear forms, Independence of two quadratic forms and independence of linear form and quadratic form. (2L)
  - Fisher Cochran's theorem. (2L)
- **Order Statistics: (9 Marks)**
  - Distribution of r^{th} order statistics, Joint distribution of several order statistics and their functions. (4L)
  - Distribution of function of order statistics. (2L)
  - Extreme values and their asymptotic distributions (statement only) with applications. (2L)
  - Distribution of spacings, normalized spacings with illustration to exponential case. (3L)

REFERENCES

1. Rohatgi V.K. and Ehsanes Saleh A. K. MD. (2003). An Introduction to Probability Theory and Mathematical Statistics, (Wiley Eastern, 2nd Ed.).
2. Hogg, R.V. and Craig, A.T. (1978). Introduction to Mathematical Statistics, (5th Ed. Pearsons Education).
3. Hogg, R.V. and Tanis E.(2002) An. Probability and Statistical Inference (6th Ed. Pearsons Education).
4. Rao, C.R. (2002). Linear Statistical Inference and Its Applications, (2nd Ed, Wiley Eastern).
5. Dudewicz, E. J. and Mishra, S. N. (1988). Modern Mathematical Statistics, (Wiley & Sons).
6. Pitman, J. (1993). Probability, (Narosa Publishing House).
7. Johnson, S. and Kotz, (1972). Distributions in Statistics, (Vol..I, II and III, Houghton and

Miffin).
8. Cramer H. (1946). Mathematical Methods of Statistics, (Princeton).
9. DasGupta, Anirban. *Fundamentals of probability: a first course*. Springer Science & Business Media, 2010.
10. David, Herbert Aron, and Haikady Navada Nagaraja. "Order statistics." *Encyclopedia of Statistical Sciences* (2004).
11. Johnson, Norman Lloyd, Samuel Kotz, and Narayanaswamy Balakrishnan. "Continuous univariate distributions." *Journal of the Royal Statistical Society-Series A Statistics in Society* 159.2 (1996): 343.

ST-MJ-502 REAL ANALYSIS

Course Objectives (CObs):

- To learn the concepts of basic topological objects such as open sets, closed sets, compact sets
- To understand the concept of convergence of sequence, series, functions and sequences of functions
- To work comfortably with continuous, differentiable and Riemann integrable functions.

- **The Real Number System: (6 Marks)**

  - Introduction, The field axioms, the order axioms, Geometric representation of real numbers, Intervals, Integers, The unique factorization theorem for integers. (statement only) (1L)
  - Rational numbers, Irrational numbers, Upper bounds, Lower bounds, Least upper bound, Greatest lower bound of the sets of real numbers. (2L)
  - The completeness axiom, some properties of the supremum and infimum, Archimedean property of the real number system. (2L)
  - Absolute values and the triangle inequality, The Cauchy-Schwarz inequality, Plus and minus infinity and the extended real number system R^*. (1L)

- **Basic Notions of Set Theory: (4 Marks)**

  - Ordered pairs, Cartesian product of two sets, Relations and functions. One-to-one functions and inverses, Composite functions. (2L)
  - Similar (equinumerous) sets, Finite and infinite sets, Countable and uncountable sets, Uncountability of real number system. (2L)

- **Sequences and Series of Real Numbers: (10 Marks)**

  - Introduction and examples of sequences of real number. (1L)
  - Convergence of sequences, limit of a sequence, limit superior and limit inferior of a real-valued sequences, Monotone sequences of real numbers. (2L)
  - Cauchy sequences and related results (only statements) (1L)
  - Infinite series, Alternating series. (1L)
  - Convergence of Series, Absolute and conditional convergence. (1L)
  - Comparison test and limit comparison test (1L)
  - The power series and its convergence, Cauchy Hadmard Theorem.(without proof) (1L)
  - Ratio test and Root test, Cauchy integral test, Dirichlet's test, Abel's test. (2L)

- **Limit and Continuity: (5 Marks)**

  - Limits of functions. (1L)
  - Continuous functions. (1L)
  - Uniform continuity. (1L)
  - Discontinuities. (1L)
  - Continuity and compactness, Monotone function and discontinuities. (1L)

- **Sequences of Functions: (5 Marks)**

  - Introduction and examples of sequences of real-valued functions. (1L)
  - Pointwise convergence of sequences of functions. (1L)
  - Definition of uniform convergence, Uniform convergence and continuity. (2L)
  - Power series and radius of convergence. (1L)

REFERENCES

1. Apostol, T. M. (1985). Mathematical Analysis, (Narosa, Indian Ed.).
2. Courant, R. and John, F. (1965). Introduction to Calculus and Analysis, (Wiley).
3. Miller, K. S. (1957). Advanced Real Calculus, (Harper, New York).
4. Rudin, Walter (1976). Principles of Mathematical Analysis, (McGraw Hill).
5. Malik, S. C. (2005). Principles of Real Analysis, (New Age Inter-national (P) Ltd.).
6. Bartle, R. G. (1976). Elements of Real Analysis, (Wiley).

ST-MJ-503 SAMPLING THEORY AND STATISTICS FOR NATIONAL DEVELOPMENT

Course Objectives (CObs):

- To provide knowledge and training of sample surveys, methods of estimations of population parameters under different sampling schemes.
- To make students aware of National Economy and National Indicators of Economy and teach them role of statistics in National Developments.

- **Sample Surveys:**
  - **Preliminaries: (6 Marks)**
    - Objectives of sample survey, planning for sample survey. (1L)
    - Basic issue related to estimation [biased and unbiased estimator, mean square error (MSE)] and confidence interval (2L)
    - Concept of sampling distribution of statistic (2L)
    - Sampling and non-sampling errors (1L)
  - **Basic methods of sample selection from finite population. (10 Marks)**
    - Simple random sampling with replacement, Simple random sampling without replacement, Systematic sampling and related results on estimation of population total, mean and proportion. (5L)
    - Stratified sampling: Formation of strata and number of strata, Allocation problems and estimation problems. (5L)
  - **Unequal Probability Sampling Designs: (8 Marks)**
    - Inclusion probabilities, Horwitz-Thompson estimator and its properties. (3L)
    - PPSWR, PPSWOR methods (including Lahiri's scheme) and related estimators of a finite population mean (Hansen-Horwitz and Desraj estimators for a general sample size and Murthy's estimator for a sample of size 2). (5L)
    - Midzuno sampling design, πps design. (3L)
    - Use of supplementary information for estimation, Ratio and Regression estimators based on SRSWOR method of sampling. Their properties and MSEs. (5 Marks,5L)
    - The Jackknife technique. (2 Marks,2L)
    - Cluster sampling, Estimator of population mean and its properties. (4 Marks,3L)
    - Two-stage sampling with equal number of second stage units. (2 Marks,2L)
    - Double sampling and its uses in ratio and regression estimation. (3 Marks,3L)
    - Randomized response technique, Warner's model; related and unrelated questionnaire methods. (4 Marks, 3L)

- **Statistics for National Development:**
  - **Economic Development: (6 Marks)**

- Growth in per capita income and distributive justice. (1L)
- Indices of development. (1L)
- Human Development indexes. (1L)
- Estimation of national income-product approach, income approach and expenditure approach. (2L)

- Population growth in developing and developed countries, Population projection using Leslie matrix, Labour force projection. (2 Marks, 2L)
- Measuring inequality in incomes, Lorenz curve, Gini coefficient, Theil's measure. (2 Marks, 2L)
- **Poverty measurement: (6 Marks)**
  - Different issues related to poverty. (2L)
  - Measures of incidence and intensity. (2L)
  - Combined measures e.g. Indices due to Kakwani, Sen etc. (2L)

REFERENCES

Sampling Methods:

1. Cochran, W.G. (1984). Sampling Techniques, (Wiley).
2. Des Raj and Chandok (1999). Sample Survey Theory, (Narosa).
3. Sukhatme, P.V, Sukhatme, B.V and Ashok C. (1984). Sampling Theory of Surveys with Applications, (Iowa State University Press & IARS).
4. Mukhopadhay P. (2002). Theory and Method of Sample Survey, (Chapman and Hall)

Statistics for National Development:

1. CSO. National Accounts Statistics- Sources and Health.
2. Sen, A. (1997). Poverty and Inequality.
3. Datt R., Sundharam, K. P. M. (Revised edition). Indian Economy, (Sultan Chand & company Ltd.)

ST-MJP-504 PRACTICALS-I

Course Objectives (CObs):

- To understand various statistical tools used in presentation of data.
- To make students aware of sample survey data and its analysis.
- To generate random samples from various probability distributions

A. Introduction to different Statistical Software Packages (9 Hrs, 6 Marks)

1. Classification, tabulation and frequency tables.
2. Bar graphs, histogram.
3. Stem-and- Leaf plots, Box plots.
4. Summary statistics.
5. Two-way tables and plots.
6. Scatter diagram correlation coefficient.

B. Practicals based on the Sampling Theory and Statistics for National Development. (Using software packages) (30 Hours, 20 Marks)

1. Model Sampling and Estimation
   - Drawing simple random samples from a given finite population using SRSWR and SRSWOR.
   - Estimating the population total, mean and proportion using the sample drawn.
   - Estimating the variance of the estimator obtained above using the sample drawn.
   - Confidence interval for population total, mean and proportion.
   - Comparison of two estimators.
   - Minimum sample size requirement.
2. Stratified Random Sampling
   - estimation of population total and mean with S.E.
   - Various kinds of allocations
   - Post stratification.
3. Using Auxiliary Information
   - Ratio method of estimation
   - Regression method of estimation.
4. H-T estimator and PPS, π PS designs
5. Double Sampling.
6. Systematic Sampling
7. Cluster Sampling
8. Two stage sampling
9. Randomized Response Technique
10. Estimation of national income, Income inequality, Poverty measurement.

C. Practical based on Distribution Theory. (Using software packages) (6Hrs , 4 Marks)

1. Generating random samples from discrete, continuous and mixture distributions
2. Fitting of standard distributions and tests for goodness of fit.

D. Practicals based on R Programming and Numerical Methods (30 Hrs, 20 Marks)

1. Install and configuration of R programming environment, Basic language elements and data structures, Data input/output, Data storage formats, Subsetting objects, Functions, Loop functions, Graphics and visualization, Statistical simulation
2. Writing R programs to calculate different summary statistics (mean median, mode, variance, standard deviation, order statistics, range and quantiles) based on the given n observations.
3. R Programs to compute and plot p.m.f.'s and c.d.f.'s of standard probability distributions. (Binomial, Poisson, Geometric, Hyper Geometric, Negative Binomial)
4. Drawing random samples from standard distributions (Binomial, Poisson, Geometric, Exponential, Normal, Gamma, Beta, Discrete, Mixture), preparing frequency distribution of given data.
5. Drawing a random sample of size n using SRSWR and SRSWOR.
6. Calculation of double integrals, limits of functions, computing integrals by statistical methods, computing expectations of complicated functions.
7. Calculation of regression and correlation coefficients, ANOVA for one-way and two way models, Analysis of 2 x 2 contingency table, calculation of p-value for standard normal distribution (for given Z value), box Plot, pie charts, histograms, dot plots, density plots, mean and variance of estimates, calculation of empirical power & level of significance.
8. To locate, install and load R packages, development of personalized functions and R-GUI using R-shiney.
9. Programs based on the numerical methods: Bisection method, Newton-Raphson Method, Numerical Integration by Trapezoidal rule Simpson's rules.

E. Assignment Problem to be solved by students.

1. Preparing frequency distribution of given data.
2. Calculation of p-value for standard Normal distribution (for given Z value)
3. Calculation of regression and correlation coefficients.
4. Sketching p.d.f of the given distribution for various parameters.(Using software)

ST-EC-521 R PROGRAMMING AND NUMERICAL METHODS

Course Objectives (CObs):

- To review the core topics in probability and statistics through the study and practice of data analysis and graphical interpretation using 'R'.
- Students will learn R-reporting and developing own R code/packages/ Apps.

Unit I (12L, 12M)

- Introduction to R-A programming language and environment for data analysis and graphics.
- Syntax of R expressions: Vectors and assignment, vector arithmetic, generating regular sequence, logical vector, character vectors, Index vectors; selecting and modifying subsets of data set
- Data objects: Basic data objects, matrices, partition of matrices, arrays, lists, factors and ordered factors, creating and using these objects; Functions- Elementary functions and summary functions, applying functions to subsets of data.
- Data frames: Benefits of data frames, creating data frames, combining data frames, Adding new classes of variables to data frames; Data frame attributes.

Unit II (12L, 12M)

- Importing data files: import.data function, read.table function; Exporting data: export.data function, cat, write, and write.table functions; Outputting results - sink function, formatting output - options, and format functions; Exporting graphs - export.graph function.
- Graphics in R: creating graphs using plot function, box plot, histogram, line plot, steam and leaf plot, pie chart, bar chart, multiple plot layout, plot titles, formatting plot axes; 3-D plots: Contour plots, perspective plots, and image plots: Visualizing the multivariate data: Scatter plot matrices, Star plots, Faces
- Interactively adding information of plot - Identifying the plotted points, adding trend lines to current scatter plot, adding new data to current plot, adding text and legend
- Loops and conditional statements: Control Statements; if statement, if else Statement. Looping statement; for loop, repeat, while loop

Unit III (12L, 12M)

- Developing simple programs in R for data analysis tasks, saving programs, executing stored programs, defining a new binary operator, assignment within function, more advanced examples, object oriented programme. Creating function libraries- library function, attaching and detaching the libraries, R packages and data sets.
- Performing data analysis tasks: Reading data with scan function, Exploring data using graphical tools, computing descriptive statistics, one sample tests, two sample tests, Goodness of fit tests, vector and matrix computation, Defining Statistical Models: Introduction for defining models, Generic functions for extracting model information.

Unit IV (6L, 6M)

- Probability and Distributions: Random sampling and combinatory, obtaining density, cumulative density and quantile values for discrete and continuous distributions,

  generating samples from discrete and continuous distributions, plotting density and cumulative density curves, Q-Q plot.

Unit V (18L, 18M)

- Errors in Numerical Calculations: Introduction, Errors and their Analysis, general error formula, error in series approximation.
- Iterative methods: Introduction, bisection, Newton-Raphson method for finding roots of polynomial equation.
- Interpolation: Newtons methods, Langrangian method, direct method
- Solution of Simultaneous Algebraic Equations: Introduction, Direct method, Matrix Inversion Method, Jacobi iteration method, Gauss elimination method, Pivoting, Gauss-Seidel iterative method, Gauss Jordan method, Eigen value Problem.
- Numerical Integration: Introduction, Simpson's 1/3 rule, Trapezoidal rule, Quadrature rule, Simpson's 3/8 rule, Errors in integration formulae.

REFERENCES

1. Peter Dalgaard (2002). Statistics and computing: Introductory Statistics with R (Springer).
2. Purohit, G.S., Gore, S.D. and Deshmikh, S.R. (2008). Statistics Using R (Narosa Publishing House)
3. Maindonald, J. and Braum, J. (2007). Data Analysis and Graphics Using R: An example-based approach (2nd Ed. Cambridge Series in Statistical and Probabilistic Mathematics)
4. Hey-Jahans, C. (2012). An R Companion to Linear Statistical Models (CRC Press)
5. Gardener, M. (2012). Beginning R: The Statistical Programming Language (Wiley & Sons)
6. Acharya, S. (2018). Data Analytics using R (McGraw Hill Education)
7. Wickham, H. and Grolemund, G. (2017). R for Data Science (O'Reilly Media)
8. Lander, J.P. (2017). R for Everyone: Advanced Analytics and Graphics (Addison-Wesley Professional)
9. Kabacoff, R.I. (2015). R in Action: Data Analysis and Graphics with R (2nd Ed. Manning Publications)
10. V. Rajaraman (1993). Computer Oriented Numerical Methods, (3rd Ed. Prentice-Hall)

ST-EC-522 C++ PROGRAMMING

1. Principles of Object-Oriented Programming (4 lectures, 4M)

- Introduction to object-oriented programming
- Encapsulation, inheritance, and polymorphism
- Objects, classes, and methods

2. Beginning with C++ (4 lectures, 4M)

- Introduction to C++ programming language
- Structure of a C++ program
- Compiling and executing C++ programs

3. Tokens, Expressions, and Control Structures (8 lectures, 8M)

- Data types and variables
- Operators and expressions
- Decision-making statements (if, else, switch)
- Looping statements (for, while, do-while)

4. Functions in C++ (4 lectures, 4M)

- Defining and calling functions
- Function parameters and return types
- Function overloading

5. Classes and Objects (8 lectures, 8M)

- Introduction to classes and objects
- Data members and member functions
- Access specifiers (public, private, protected)
- Class constructors and destructors

6. Constructors and Destructors (8 lectures, 8M)

- Introduction to Constructors and Destructors
- Default Constructors
- Parameterized Constructors
- Copy Constructors
- Constructor Overloading
- Destructors
- Destructor Implementation
- Memory Management and Constructors/Destructors

7. Operator Overloading and Type Conversions (8 lectures)

- Operator overloading for built-in and user-defined types
- Overloading unary and binary operators

- Type conversions and type casting

8. Inheritance: Extending Classes (4 lectures)

- Introduction to inheritance
- Base and derived classes
- Single and multiple inheritance

9. Pointers, Virtual Functions, and Polymorphism (4 lectures)

- Pointers and references in C++
- Virtual functions and dynamic binding
- Polymorphism and function overriding

10. Managing Console I/O Operations (4 lectures)

- Input and output streams
- Reading and writing data from/to the console
- Formatting console output

11. Working with Files (including linking to databases) (4 lectures)

- File handling in C++
- Reading from and writing to files
- Linking to databases for data storage

REFERENCES

1. Balagurusamy, (2006). Object-Oriented Programming with C++, (Ed. Tata McGrawHill).
2. Gottfried. Programming in C++, (Schaum's Outline Series).
3. K. R. Venugopal, Rajkumar, J.Ravishankar. Mastering C++.
4. V. Rajaraman (1993). Computer Oriented Numerical Methods, (3rd Ed. Prentice-Hall)
5. W. H. Press, S. A. Teukolsky, W.T. Vellering and B.P.Flannery (1993). Numerical Recipesin C, (2nd Ed. Cambridge University Press).
6. R.A. Thisted (1988). Elements of Statistical Computing, (Chapman and Hall).Ross, S. (2005). Introduction to Probability Models, (6th Ed. Academic Press).

ST-RM-541 Research Methodology

Course Objectives (CObs):

The course will help to ensure that the scholars are able to:

- Demonstrate knowledge pertinent to the role of research in all disciplines of science, humanities, commerce and management and interdisciplinary.
- Critically analyze and interpret the results of scientific outcome
- Evaluate its limits and possibilities with respect to knowledge in the research and its implementation.
- Improve the teaching learning process with the help of research

Unit-I Basics of research and research types (12 Marks, 12L)

Research objectives, Scientific research, Importance of research methodology, Good Research Practices (GRP),Types of research and methods, surveys, case studies, Experiments and field studies, etc.

Unit-II Review of Literature and research (12 Marks, 12L)

Basic Concept and its need, literature search, types of literature review, review of research, sources, synthesis process, planning of review and documentation

Unit-III Research design (12 Marks, 12L)

Process, identification and Formulation of problem, Hypothesis, Tools of research.

Unit-IV Data analysis and interpretation (12 Marks, 12L)

Editing, Coding, Transcription, Tabulation, Introduction to Analytical/Statistical software (SPSS, MINITAB, MATLAB) and presentation of data (Graphical)

Unit-V Statistical analysis and report writing (12 Marks, 12L)

Measures of central tendency, dispersion and Association/Relationship, Variance, Regression and Correlation analysis, Hypothesis testing and Test of significance, Research paper and thesis writing

References:

1. Kothari, C.R. (2014) Research Methodology: Methods and Techniques, 2nd edn., New Age International Publishers, New Delhi.
2. Best, J. W. and Kahn, J. V. (2006) Research in Education, 10th edn., Pearson Publication, New Delhi.
3. Koul, L. (2019) Methodology of Educational Research, 5th edn., Vikas Publ., New Delhi (ISBN- 9789353386368)
4. Garrett, H.E. (2005) Statistics in Psychology and Education, Paragon International Publishers, New Delhi
5. Kambadur, M., Ghosh, A. and Singhvi A. K. (2019) Ethics in Science Education, Research and Governance, Indian National Science Academy, New Delhi (ISBN: 978-81-939482-1-7)
6. Best Practice Guidelines on Publishing Ethics(2014), A Publisher's Perspective, 2nd edn., John Wiley & Sons, Ltd.
7. Chaddah, P. (2018) Ethics in Competitive Research: Do not get Scooped; Do not get Plagiarized, ISBN-978-9387480865
8. National Academy of Sciences, National Academy of Engineering and Institute of Medicine (2009) On Being a Scientist: A Guide to Responsible Conduct in Research,

3rd edn., National Academies Press.
9. Beall, J. (2012) Predatory Publishers Corrupting Open Access, Nature, 489(7415):179.
10. Bird, A. (2006) Philosophy of Science, Routledge.
11. Software Manuals SPSS, MINITAB, MATLAB
12. UGC syllabus for Research and Publication Ethics, D.O. no. F.1-1/2018 (Journal/CARE), Dec., 2019

SEMESTER-II

ST-MJ-551 PROBABILITY THEORY

Course Objectives (CObs):

- To teach students the measure theoretical concepts of Probability theory.
- To describe concept of moment inequalities, convergence of sequence of random variables and their applications as WLLN and CLT.

- **Sets and Classes of Events: (6 Marks)**
  - Random experiment, Sample space and events. (1L)
  - Algebra of sets. (1L)
  - Sequence of sets, limit supremum and limit infimum of sequence of sets. (2L)
  - Classes of sets, Sigma-fields (σ-fields), Minimal fields, Minimal σ-field, Parition. (3L)
  - Borel fields in R^1 and R^k , Monotone field. (2L)

- **Random Variables: (6 Marks)**
  - Point function and set function, Inverse function. (2L)
  - Measurable function, Borel function, induced σ-field, Function of a function, Borel function of measurable function. (2L)
  - Real and vector-valued random variable. (2L)
  - σ-field induced by a sequence of random variables. (1L)
  - Limits of Random variable. (2L)

- **Measure and Probability Measure: (6 Marks)**
  - Measure (Definition and simple properties). (1L)
  - Probability measure, Properties of a measure. (1L)
  - Probability space (finite, countable) Continuity of a probability measure. (1L)
  - Extension of probability measure, Caratheodory Extension theorem(without proof) (1L)
  - Probability space induced by r.v. X , Distribution of Borel functions of r.v. (1L)
  - Other measures: Generalized Probability measure, Conditional Probability measure, Counting measure, Lebesgue measure. (2L)

- **Distribution Functions: (6 Marks)**
  - Distribution functions of a r.v. and its properties. (1L)
  - Jordan decomposition theorem, Mixture of distribution functions. (2L)
  - Distribution functions of vector valued r.v.s. (1L)
  - Empirical distribution functions. (1L)

- **Expectation and Moments: (8 Marks)**
  - Integration of measurable function with respect to a measure. (1L)
  - Expectation of a r.v. (Definition for simple, Nonnegative and arbitrary r.v.), Properties of expectation, Expectation of Complex r.v. (3L)

- Moments, Moment generating function. (1L)
- Moment inequalities: C-inequality, Holder inequality, Schwarz's inequality, Minkowski's inequality, Jensen's inequality, Basic inequality, Markov inequality, Chebyshev's inequality. (3L)

- **Convergence of Sequence of Random variables: (8 Marks)**

  - Convergence in distribution, Convergence in probability, Almost sure convergence and convergence in quadratic mean and their inter-relations. (5L)
  - Monotone convergence theorem, Fatou's Lemma, Dominated convergence theorem. (3L)

- **Characteristic function: (4 Marks)**

  - Definition and simple properties, Some inequalities. (2L)
  - Uniqueness theorem and Levy's continuity thereon (Statements only). (1L)

- **Independence: (5 Marks)**

  - Independence of two events, Independence of n>2 events, sequence of independent events, independent classes of events, independence of r.v.s, Borel zero-one law. (4L)

- **Law of large numbers: (5 Marks)**
  - Weak laws of large numbers (WLLN), Khintchine's WLLN, Kolmogorov's strong lawof large number (Statement only) and their applications (4L)

- **Central limit theorem (CLT): (6 Marks)**

  - CLT for a sequence of independent r.v.s. under Lindeberg's condition, CLT for i.i.d. r.v.s. and its applications. (3L)

REFERENCES

1. Ash, Robert. (1972). Real Analysis and Probability, (Academic Press).
2. Bhat, B.R. (1999). Modern Probability Theory, (3rd Ed. New Age Inter-national (P) Ltd. Pub)
3. Billingsley, P. (1986). Probability and Measure, (Wiley).
4. Basu, A. K. (1999). Measure Theory and Probability (Prentice Hall of India).
5. Feller, W. (1969). Introduction to Probability and its applications Vol.II, (Wiley Easter Ltd.)
6. Loeve, M. (1978). Probability Theory, (4th Ed. Springer Verlag).
7. Gnedenko, B.V. (1988). Probability Theory, (Mir. Pub.).

ST-MJ-552 LINEAR ALGEBRA

Course Objectives (CObs):

- To understand the vector spaces and subspaces and important concepts of vector spaces such as independence, basis, dimensions, orthogonality etc.
- To understand the link between linear transformation, its properties and matrices, matrix-operations; the spaces associated with matrices and interrelation between them, rank-factorization, eigen and spectral analysis of a matrix.
- To determine the existence and uniqueness of the solution of a linear system of equations, solution set and determining all possible solutions using generalized inverse, Quadratic forms and their definiteness categories, maxima/minima etc

- **Vector Spaces (VS) : (8 Marks)**
  - Binary operations and its properties (1L)
  - Definition of Vector Space and examples (1L)
  - Definition of Subspace and examples. Results on Intersection and union of subspaces (2L)
  - Definition of Linear dependence and linear independence of set of vectors (1L)
  - Steinitz exchange theorem (statement only), Maximal linearly independent set, minimal generating sets. (1L)
  - Definitions of Basis of Vector Spaces and Dimension, relation between dimensions of subspaces; one of which is subset of other.(only statements) (1L)
  - Sum of two sets, modular law (only statement), examples on modular law (1L)

- **Algebra of Matrices: (8 Marks)**
  - Definition of Linear transformation and examples. (1L)
  - Addition, Scalar multiple and composition of linear transformation (only statements). Representation of linear transformation into matrix form w.r.t canonical basis and given basis. (1L)
  - Row and column spaces, Rank of a matrix, Left inverse, Right inverse and inverse of a matrix, properties of inverse, Upper bound for rank of product of matrices, (2L)
  - Null space of a matrix, rank of null space of a matrix with rank of a matrix, Nullity of matrix. Rank nullity Theorem (statement only), examples (1L)
  - Lower bound for rank of product of two matrices, Rank of sum of matrices. (1L)
  - Partitioned matrix, , Determinant of a matrix, Determinant of partitioned matrix, (1L)
  - Generalized inverse of a matrix and its properties (only statements of properties without proofs) (1L)

- **Inner Product and Orthogonally: (6 Marks)**

  - VS with inner product, Normed vector spaces, Cauchy-Schawrz inequality, (statement only) Orthogonality and linear independence. (2L)
  - Orthonormal basis, Expression of any vector in VS as a linear combination of elements of orthonomal basis. (1L)
  - Gram-Schmidt orthogonalization process, Extension of any orthogonal set to orthonormal basis of VS, Examples. (2L)
  - Orthogonal and unitary matrices and their properties. (1L)

- **Eigen Values: (4 Marks)**

  - Characteristic polynomial and characteristic equation of a matrix, Characteristic roots, their properties. (properties without proofs) (1L)
  - Eigen values and eigen vectors, Eigen spaces, Geometric and algebraic multiplicity of an eigen value, Relation between the 2 multiplicities, Simple and regular eigen values, Properties of eigen values. (only statements) (2L)
  - Cayley-Hamilton theorem (without proof) and minimal polynomial, Singular values and singular vectors. (1L)

- **Quadratic Forms (QF): (4 Marks)**
  - Real QF, Classification, Rank and signature, reduction of any QF to diagonal form. (2L)
  - Definiteness of a matrix, equivalence of nonnegative definite matrix and variance-covariance matrix. (1L)
  - Extrema of QF, Maxima and Minima of ratio of two QF. (only statements) (1L)

REFERENCES

1. Graybill, F.A.(1983). Matrices with Applications in Statistics (2nd Ed. Wadsworth)
2. Rao, A.R. and Bhimasankaram, P. (2000). Linear Algebra. (Hindustan Book Agency).
3. Rao, C.R. (2002). Linear Statistical Inference and its Applications. (2nd ed. John Wiley and Sons)
4. Searle, S. R. (1982). Matrix Algebra Useful for Statistics. (John Wiley and Sons Inc.).
5. Bellman, R.(1970). Introduction to Matrix Analysis, (2nd ed.Tata McGraw Hill).
6. Biswas, S.(1984). Topics in Algebra of Matrices, (Academic Publictions).
7. Hadley, G.(1987). Linear Algebra, (Narosa Publishing House).
8. Halmos, P.R.(1958). Finite-dimensional Vector Spaces, (2nd ed. D.Van Nostrand Company, Inc.).
9. Hoffman, K. and Kunze, R. (1971). Linear Algebra, (2nd Ed.Prentice Hall, Inc.)
10. Rao, C.R. and Mitra, S.K. (1971). Generlized Inverse of Matrices and its Applications, (John Wiley and Sons Inc.).

ST-MJ-553 LINEAR MODELS AND REGRESSION ANALYSIS

Course Objectives (CObs):

- To provide the theoretical foundation for Linear models.
- To familiar with principles of multiple linear regression and non-linear regression models.
- To study concept of generalized linear models.
- To develop and validate models on the basis of collected data.

- **General Linear Model: (15 Marks)**

  - Gauss-Markov set up, Least square estimation, Normal equations, Consistency of system of normal equations and their solution. (3L)
  - Estimability of linear parametric function, necessary and sufficient condition for estimability, Best Linear Unbiased Estimator (BLUE). (2L)
  - Gauss-Markov theorem, Variances and covariances of BLUE's. (2L)
  - Estimation space, Error space, their ranks, Orthogonality of estimation space and error space. (2L)
  - Simultaneous estimates of linear parametric function, Estimation of error variance, Estimation with correlated observations. (3L)
  - Least square estimates with restriction on parameters, Method of generalized least squares. (3L)

- **Interval Estimation and Test of Hypothesis: (15 Marks)**

  - Under the normality assumption, Distribution of error sum of squares, Regression sum of squares and distribution of BLUE's, their independence. (2L)
  - Distribution of conditional error sum of squares, Distribution of sum of squares due to null hypothesis. (3L)
  - Test of hypothesis for one or more than one estimable linear parametric function, Test of hypothesis of equality of all estimable functions to zero, Testing of sub hypothesis for full rank model, Power of F-test. (3L)
  - Simultaneous confidence interval for n linearly independent estimable parametric functions. (2L)
  - One way and two way classified data, multiple comparison tests due to Tukey-Scheffe. (4L)

- **Regression Analysis: (30 Marks)**

  - Simple and multiple linear regression in Gauss-Markov set up. Estimation of regression coefficients, Regression analysis of variance, Fitted values and residuals. (4L)
  - Polynomial regression, Orthogonal polynomials, Response analysis using orthogonal polynomials. (3L)
  - Residuals and their plots as tests for departure from assumptions such as fitness of the model, Normality, Homogeneity of variances and detection of outliers. (2L)
  - Remedial measures and validation, Multi-collinearity, Ridge regression, Robust regression principal component regression subset selection of explanatory variables, Mallows Cp statistic. (7L)
  - Introduction to non-linear regression models, Least square estimation in non-linear regression, Model building and diagnostics. (4L)

- Generalized Linear model: Link functions required for dependent variable following distributions like Poisson, binomial, inverse binomial, inverse Gaussian, gamma.

  Logistic Regression: Logit transform, ML estimation, Test of hypotheses, Wald test, LR test, score test. (7L)

REFERENCES

1. Cook, R.D. and Weisberg, S.(1982). Residual and Influence in Regression, (Chapman & Hall).
2. Draper, N.R.and Smith, H. (1998). Applied Regression Analysis, (3rd Ed.Wiley).
3. Gunst, R.F.and Mason, R.L. (1980). Regression Analysis and Its Applications- A Data Oriented Approach, (Marcel and Dekker).
4. Montegomery D.C,Peck, E.A and Vining G.G(2003). Introduction to Linear Regression Analysis, (3rd Ed. Wiley)
5. Rao,C.R. (2002). Linear Statistical Inference and its Applications, (2nd Ed.Wiley).
6. Weisberg, S. (1985). Applied Linear Regression., (Wiley).

ST-MJP-554 PRACTICALS-II

Course Objectives (CObs):

- Introduce the basic operations and numeric computations in MATLAB software, exercise the understanding of algebraic concepts through computation of inverse, generalized inverse, definiteness; eigen and spectral analysis, orthogonalization of a matrix, etc
- To deal and process the multivariate data, use R and Matlab to perform exploratory analysis of such a data, apply the multivariate statistical techniques to solve the objective specific problems
- To understand and apply regression model building techniques to various data sets.
- To study sampling distribution of estimators, plotting of likelihood functions, power functions.

A. Practicals based on Linear Algebra. (Using software packages) (15 Hours, 10 Marks)

1. Checking linear dependence/independence of set of vectors using system of linear equations.
2. Getting vectors in row/column space and null space of the given matrix.
3. Computation of inverse of a given matrix.
   - Natural inverse.
   - G-inverse, left and right inverse
4. Computing higher order powers of a given matrix using spectral decomposition
5. To obtain rank factorization of given non-null matrix.
6. Gram-Schmidt orthonormalization, forming an orthogonal matrix of specified order using Gram-Schmidt orthogonalization, forming an orthogonal matrix containing a specified vector as a row/column of the matrix.
7. Checking and demonstrating the definiteness of the given matrix, getting vectors from eigen-space, algebraic and geometric multiplicity of an eigen value etc.
8. Demonstration of occurrence of maxima and minima of
   - Quadratic forms over normed vectors.
   - Ratio of two quadratic forms over normed vectors.
9. Verification of Cayley-Hamilton theorem

B. Practicals based on Linear Models and Regression Analysis. (23 Hours, 16 Marks)

1. Linear Estimation.
2. Analysis of CRD, RBD, LSD.
3. Test of hypotheses for one and more than one linear parametric functions.
4. Multiple Regression:
   - Estimation of regression coefficient, Fitting of multiple linear regression.
   - Testing of hypothesis concerning regression coefficient.
   - Testing of significance of association between the dependent and independent variables.
   - Lack of fit test, Extra sum of squares principle.
5. Orthogonal Polynomials: Fitting of orthogonal polynomials.
6. Residual Analysis.
7. Non-linear regression.
8. Logistic Regression.

C. Practicals based on Python Programming (22 Hours, 14 Marks)

Concerned teacher is expected to design the practical according to the syllabus.

ST-EC-571 PYTHON PROGRAMMING

Course Objectives (CObs):

- To understand why Python is a useful scripting language for developers.
- To learn how to use lists, tuples, and dictionaries and to design and program Python applications.
- To learn how to use import, export, indexing and slicing to access data i.e data handling techniques.
- To learn how to write conditions, loops, decision functions and pass arguments.

- Structure of a Python Program, Elements of Python (4L,4M)

- Introduction to Python: Python Interpreter, Using Python as calculator, Python shell, Indentation. Atoms, Identifiers and keywords, Literals, Strings, Operators (Arithmetic operator, Relational operator, Logical or Boolean operator, Assignment, Operator, Ternary operator, Bit wise operator, Increment or Decrement operator). (8L,8M)

- Creating Python Programs: Input and Output Statements, Control statements (Branching, Looping, Conditional Statement, Exit function, Difference between break, continue and pass.), Defining Functions, default arguments. (6L,6M)

- Structures: Numbers, Strings, Lists, Tuples, Dictionary, Date & Time, Modules, Defining Functions, Exit function, default arguments. (10L,10M)

- Introduction to Advanced Python: Objects and Classes, Inheritance, Regular Expressions, Event Driven Programming, GUI Programming. (12L,12M)

- Basic Statistics in Python with NumPy, PyDev and Jupyter Notebook development environments, Pandas data analysis library, including reading and writing of CSV files, Matplotlib 2D plotting library, Git and GitHub (20L,20M)

REFERENCES

1. VanderPlas, Jake. *Python data science handbook: essential tools for working with data*. " O'Reilly Media, Inc.", 2016.
2. Boschetti, Alberto, and Luca Massaron. *Python data science essentials*. Packt Publishing Ltd, 2015.
3. Lutz, Mark. *Learning python: Powerful object-oriented programming*. " O'Reilly Media, Inc.", 2013.
4. McKinney, Wes. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.", 2012.

ST-EC-572 SAS Programming

Course Objectives (CObs):

- To learn SAS Studio and Enterprise Guide programming environment
- To deal with various types of data and importing and exporting data sets
- Apply various statistical methods using SAS

1. Introduction to SAS (6 lectures, 6M)

- Overview of SAS software and its applications
- SAS interface and basic functionalities
- SAS data sets and data libraries
- Introduction to SAS programming

2. Accessing and Describing Data (6 lectures, 6M)

- Importing data into SAS
- Reading and writing data sets
- Exploring and summarizing data
- Basic descriptive statistics using SAS

3. Creating Data Structures (6 lectures, 6M)

- Creating and modifying variables in SAS
- Working with data formats and informats
- Combining and merging data sets
- Sorting and indexing data

4. Managing and Formatting Data (6 lectures, 6M)

- Data manipulation using SAS functions and operators
- Subsetting and filtering data
- Transforming and recoding data
- Applying conditional logic to data

5. Statistical Testing (12 lectures, 12M)

- Introduction to hypothesis testing
- Performing t-tests and chi-square tests in SAS
- Analysis of variance (ANOVA) using SAS
- Introduction to non-parametric tests in SAS

6. Statistical Analysis and Modelling (12 lectures, 12M)

- Simple linear regression analysis in SAS
- Introduction to multivariate analysis in SAS
- Time series analysis using SAS

- Survival analysis with SAS

7. Exporting Data and Generating Reports (8 lectures, 8M)

- Exporting SAS data sets to other file formats (e.g., Excel, CSV)
- Generating basic reports using SAS procedures
- Customizing reports with formatting and graphics
- Introduction to SAS Output Delivery System (ODS)

8. Handling Errors (4 lectures, 4M)

- Debugging SAS programs
- Error handling and data validation techniques
- Understanding SAS log messages
- Troubleshooting common errors in SAS

REFERENCES

1. Cody, Ron. *An introduction to SAS university edition*. SAS Institute, 2018.
2. Cody, Ron. *Learning SAS by example: a programmer's guide*. SAS Institute, 2018.
3. Delwiche and Slaughter: *The Little SAS Book*, 5th Edition.
4. Cody and Smith: *Applied Statistics and the SAS Programming Language*, 5th Edition.
5. McDaniel, Stephen, and Chris Hemedinger. *SAS for Dummies*. John Wiley & Sons, 2007.

ST-OJT-591 On job training: (Internship/Apprentionship)

Course Objectives:	
1	To offer the opportunity for the young students to acquire on job the skills, knowledge, attitudes,and perceptions along with the experience needed to constitute a professional identity.
2	To provide means to immerse students in actual supervised professional experiences.
3	To give an insight into the working of the real organizations.
4	To gain deeper understanding in specific functional areas.
5	To appreciate the linkages among different functions and departments.
6	To help the students in exploring career opportunities in their areas of interest.

- On Job Training (OJT) will be arranged by Department at some manufacturing/service industry or govt organizations. Students may also search such places for training. 15-20 days period after first semester will be given for this training purpose. This OJT will be evaluated by host institute and Department jointly.

  • Field Projects will be generally based on some survey conducted by students (may be allowed in groups). 15-20 days period after first semester will be given for this training purpose. These projects will be guided by faculty members of the Department and evaluated by Department by two examiners.

  • Project Guide: Teachers from the Department of Statistics and/or personnel from organization where student is going to visit for field work or training. Each project group will be guided by concerned teacher (guide) for one hour per week throughout the semester.

  • Fieldwork: Students will be given 15-20 days period during respective end of semester for their industrial work/data collection/survey or any other fieldwork involved in the project.

  • Project Topic: Students in consultation with the guide will decide Project Topic/Area. Project work may be carried out in a group of students depending upon the depth of fieldwork/problem involved.

  • Project report: Project report based on the on job training/field project should be submitted in typed form with binding by the end of semester. Project viva will be scheduled.

  • Project evaluation: Project evaluation will be based on

  ▪ Evaluation by Internal/External Guide/Mentor (40 Marks)

  ▪ External Examination (60 Marks)

    - Project report (20 Marks)
    - Presentation by student or group of students. (20 Marks)
    - Viva voce (20 marks)
    - Two examiners will evaluate project work.

List of Equivalent Courses

Course Code	Title of the Course under old Syllabus (w.e.f. 2022-23)	Course Code	Title of Equivalent Course under New Syllabus (w.e.f. 2021-22)
ST-111	Real Analysis	ST-MJ-502	Real Analysis
ST-112	Linear Algebra	ST-MJ-552	Linear Algebra
ST-113	Sample Survey and Statistics for National Development	ST-MJ-503	Sampling Theory and Statistics for National Development
ST-114	Distribution Theory	ST-MJ-501	Distribution Theory
ST-115	Computer Programming in C++ and Numerical Methods	-	#
ST-116	Practicals- I	ST-MJP-504	Practicals- I
ST-121	Probability Theory	ST-MJ-551	Probability Theory
ST-122	Linear Models and Regression Analysis	ST-MJ-553	Linear Models and Regression Analysis
ST-123	Multivariate Analysis	-	#
ST-124	Parametric Inference	-	#
ST-125	Python Programming OR C++	-	#
ST-126	Practicals-II	ST-MJP-554	Practicals-II
		ST-EC-521	R Programming And Numerical Methods
		ST-EC-522	C++ Programming
		ST-RM-541	Research Methodology
		ST-EC-571	Python Programming
		ST-EC-572	SAS Programming
		ST-OJT-591	On job training: Internship

No equivalent courses available for these papers, so student may be allowed to appear by old courses.