Shri Swami Vivekanand Shikshan Sanstha's

Vivekanand College Kolhapur
(Empowered Autonomous)

Department of Statistics

M. Sc. I-Applied Statistics

Semester I & II

Syllabus to be implemented from
Academic year 2025-26
(As per NEP 2020)

Teaching and Evaluation scheme

One/Two- Years PG Programme

Department/Subject Specific Core or Major (DSC)

Semester	Course Code	Course Title	No. of Credits
I	DSC18STA11	Distribution Theory	**04**
	DSC18STA12	Estimation Theory	**04**
	DSC18STA13	Statistical Computing	**02**
	RMD18STA11	Research Methodology	**04**
	DSC18STA19	Practical I	**04**
	DSE18STA11	Data Base Management System (DBMS)	**04**
	DSE18STA12	Statistical Analysis Using Minitab	**04**
		Total Credits	**22**
II	DSC18STA21	Probability Theory	**04**
	DSC18STA22	Theory of Testing of Hypothesis	**04**
	DSC18STA23	Multivariate Analysis	**02**
	DSC18STA29	Practical II	**04**
	FPR18STA21	FP/OJT	**04**
	DSE18STA21	Linear model and Regression Analysis	**04**
	DSE18STA22	Statistical Analysis System (SAS)	**04**
	DSE18STA23	Power BI and Tableau	**04**
		Total Credits	**22**

First Year Semester I & II (M. Sc. Applied Statistics)

Objectives:

1.The students are expected to understand the principles, concepts, and recent developments in the Statistics.

2. To enhance student sense of enthusiasm for Statistics and to involve them in an intellectually stimulating experience of learning in a supportive environment.

3. The practical course is framed in relevance with the theory courses to improve the understanding of the various concepts in Statistics.

Program Outcomes (PO):

On successful completion of the program students will able to:

PO1: Understand the principles and concepts in the statistical theory at an advanced level which take into account recent advances in the subject.

PO2: Acquire the strong foundation of statistical concepts which will benefit them to become good Statistician.

PO3: Use acquired statistical methodologies and modelling techniques to address real-life problems.

PO4: Gain the knowledge of software which has the wide range of opportunities in the Quality control, Planning and development, IT sector, industries, Business, Government and private sector etc.

PO5: Qualify various National / State level competitive exams like ISS, DSO, CSIR-UGC NET, SET, GATE, MPSC, UPSC, Banking etc.

Program Specific Outcomes (PSO):

On successful completion of the program students will able to:

PSO1: Enhance sense of enthusiasm for Statistics and to involve them in an intellectually stimulating experience of learning in a supportive environment.

PSO2: Handle and analyse small as well as large databases with statistical software and computing environment including C, R, Python, MS-EXCEL.

PSO3: Understand, implement and develop statistical models.

PSO4: Describe complex statistical ideas to non-statisticians and to present the results of their analyses in written, oral forms and can make practical suggestions for improvement.

PSO5: Apply statistical techniques to optimize and monitor real life phenomena related to industry and business analytics etc.

1) Nature of the theory question papers (4 credits):

a) There shall be 7 questions each carrying 16 marks.

b) Question No.1 is compulsory.

i) It consists of 8 multiple choice questions for 1 mark each. (8 marks)

ii) short note type questions (attempt any 2 out of 3) (8 marks)

c) Students have to attempt any 4 questions from question No. 2 to 7.

d) Question No. 2 to 7 shall contain 2 to 4 sub-questions.

2) Nature of the theory question papers (2 credits):

a) There shall be 4 questions.

b) Question No.1 is compulsory.

i) It consists of 4 multiple choice questions for 1 mark each. (4 marks)

ii) short note type questions (attempt any 1 out of 2) (4 marks)

c) Question No. 2 to 4 shall be of 16 marks each.

d) Students have to attempt any 2 questions from question No. 2 to 4.

e) Question No. 2 to 4 shall contain 2 to 4 sub-questions.

3)Nature of Practical examination: -

Sr. No.	Component	Max. marks
1	Practical examination: Examination will be of 3-hour duration. There shall be 8 questions each of 12 marks, of which a student has to attempt any 5 questions.	60
2	Day-to-day practical performance and journal	20
3	Viva: Viva will be based on all practical's	20
	Total	100

4) Each of the following courses have the same question paper in all examinations of M.Sc. Statistics and M.Sc. Applied Statistics.

Semester I	Semester II
Distribution Theory	Probability Theory
Estimation Theory	Theory of Testing of Hypothesis
Statistical Computing	Multivariate Analysis
Research Methodology	

M.Sc., I Applied Statistics (Sem I)
DSC18STA11: DISTRIBUTION THEORY

Course Outcomes: At the end of the course students will be able to:
CO1: Recognize and learn concept of mixture of distribution and their decomposition.
CO2: Execute transformation of univariate random variables and different moment inequalities.
CO3: Describe the concept of central and non-central distributions.
CO4: Learn the concept of order statistics.

Unit	DSC18STA11: DISTRIBUTION THEORY	No. of hours
I	Review of Random experiment and its sample space, events, random variables, discrete random variables, continuous random variables. Cumulative distribution function (CDF), properties of CDF, computation of probabilities of events using CDF, quantiles, absolutely continuous and discrete distributions, mixtures of probability distributions, decomposition of mixture CDF into discrete and continuous CDFs, expectation and variance of mixture distributions. Transformations of univariate random variables, probability integral transformation	15
II	Concepts of location, scale and shape parameters of distributions with examples. Symmetric distributions and their properties. Moment inequalities (with proof): Basic, Holder, Markov, Minkowski, Jensen, Tchebysheff and their applications, Random vectors, joint distributions, Independence, variance-covariance matrix, joint MGF. Conditional expectation and variances, Transformations of bivariate random variables, Convolutions, compound distributions.	15
III	Bivariate Normal distribution, Multivariate normal distribution: two definitions and their equivalence, singular and nonsingular normal distribution, characteristic function, moments, marginal and conditional distributions, Independence of Random vectors.	15
IV	Sampling distributions of statistics from univariate normal random samples: central and non-central chi-square, t and F distributions. Distributions of linear and quadratic forms involving normal random variables, Fisher Cochran and related theorems: only statements and applications. Order Statistics: Distribution of an order statistics, joint distributions of two order statistics, distribution of spacings, normalized spacings with illustration to exponential case, distribution of sample median and sample range.	15

Books Recommended:

1 Rohatagi V. K. & Saleh A. K. Md. E. (2001): Introduction to Probability Theory and Mathematical Statistics- John Wiley and sons Inc.
2 Johnson N. L. &Kotz. S. (1996): Distributions in Statistics Vol-I, II and III, JohnWiley and Sons New york.
3 S. Kotz, N. Balakrishnan, N. L. Johnson: Continuous Multivariate Distributions - Second Edition, Wiley.
4 Casella & Berger (2002): Statistical Inference - Duxbury advanced series.IInd edition
5 C. R. Rao (1995): Linear Statistical Inference and Its Applications (Wiley Eastern) Second Edition
6 Dasgupta, A. (2010): Fundamentals of Probability: A First Course (Springer)

DSC18STA12: ESTIMATION THEORY

Course Outcomes: At the end of the course students will be able to:

CO1: Describe the notion of a parametric models, point estimation of the parameters of those models.

CO2: Construct the sufficient statistic, minimal sufficient statistic, m.l.e., moment estimator of the parameter.

CO3: Discuss the concept of MVUE, MVBUE, UMVUE.

CO4: Describe the concept of Bayesian inference and their real-life applications.

Unit	DSC18STA12: ESTIMATION THEORY	No. of hours
I	Sufficiency principle, factorization theorem, minimal sufficiency, minimal sufficient partition, construction of minimal sufficient statistics, minimal sufficient statistic for exponential family, power series family, curved exponential family, Pitman family. Completeness, bounded completeness, ancillary statistics, Basu's theorem and applications.	15
II	Problem of point estimation, unbiased estimators, minimum variance unbiased estimator, Rao- Blackwell theorem and Lehmann-Scheffe theorem and their uses. Necessary and sufficient condition for MVUE and their applications. Fisher information and information matrix, Cramer- Rao inequality, Chapmann-Robinson bounds, Bhattacharya bounds, their applications.	15
III	Method of maximum likelihood (MLE) and small sample properties of MLE, method of scoring and application to estimation in multinomial distribution. MLE in non-regular families. Other methods of estimation: method of moments, minimum Chi square. U-Statistics: one and two sample; U-Statistics theorem for one sample and two sample (statements only).	15
IV	Consistency of an estimator, weak and strong consistency, joint and Marginal consistency, invariance property under continuous transformations, methods of Constructing consistent estimators, asymptotic relative efficiency. Consistent and Asymptotic Normal (CAN) Estimators: Definition of CAN estimator for real and vector valued parameters, invariance of CAN property under non-vanishing differentiable transformation. Methods of constructing CAN estimators: Method of Moments, method of percentiles, comparison of CAN estimators. CAN and BAN estimators in one parameter exponential family of distributions, BAN estimators	15

Books Recommended:

1. V. K. Rohatgi, and A. K. MD. E. Saleh (2015): Introduction to Probability Theory and MathematicalStatistics, John Wiley & sons, 3rd Edition.
2. E. L. Lehmann (1983): Theory of Point Estimation, John Wiley & sons.
3. C. R. Rao (1973): Linear Statistical Inference and its Applications, wiley, 2nd Edition.
4. B. K. Kale, and K. Muralidharan (2015): Parametric Inference: An Introduction, Alpha Science InternationalLtd.
5. P. Mukhopadhyay (2015): Mathematical Statistics, Books and Allied (p)Ltd.
6. E. J. Dudewicz and S. N. Mishra (1988): Modern Mathematical Statistics, John Wiley and Sons.
7. Casella and Berger (2002): Statistical Inference, Duxbury advanced series, IInd edition.

DSC18STA13: STATISTICAL COMPUTING

Course Outcomes: At the end of the course students will be able to:

CO1: Construct formulas, including the use of built-in functions and analysis tool pack.

CO2: Learn Lookup functions, pivot table and pivot chart

CO3*:* Develop the fundamentals of statistical analysis in R environment.

CO4: Learn different control statements and test procedures.

Unit	DSC18STA13: STATISTICAL COMPUTING	No. of hours
I	MSEXCEL: Introduction to MSEXCEL. Cell formatting, conditional formatting, Data manipulation using EXCEL: sort and filter, find and replace, text to columns, remove duplicate, data validation, consolidate, what-if-analysis. Working with Multiple Worksheets and Workbooks. Built-in mathematical and statistical functions for obtaining descriptive statistic, computing PMF/PDF, CDF and quantiles of the well-known distributions, rand and randbetween function, Logical functions: if, and, or, not. Lookup functions: hlookup, vlookup, Formula Errors, Creating and Working with Charts, Database functions, Text functions, Date and time functions, Excel add-ins: analysis tool pack, Pivot tables and charts, Simple Tasks with Macros.	15
II	R-software: Introduction to R, data types and objects, operators, data input, data import and export, built in functions for descriptive statistics, random sampling and computation of pdf, cdf and quantiles of well-known distribution. Strings and Dates in R. apply family of functions. Saving work in R. Matrix algebra, graphical procedures, frequencies and cross tabulation, built in functions: lm, t.test, prop.test, wilcox.test, ks.test, var.test, chisq.test, aov. Control statements. Programming, user defined functions, Rpackages. R-studio.	15

Books Recommended:

1.Gardener, M. (2012). Beginning R: the statistical programming language. John Wiley & Sons.

2. Held, B., Moriarty, B., & Richardson, T. (2019). Microsoft Excel Functions and Formulas with Excel 2019/Office 365. Mercury Learning and Information.

3. Herkenhoff, L., & Fogli, J. (2013). Applied statistics for business and management using Microsoft Excel. New York: Springer.

4. Purohit, S. G., Gore, S. D., & Deshmukh, S. R. (2015). Statistics using R. Alpha Science International.

DSE18STA11: Data Base management System (DBMS)

Course Outcomes: At the end of the course students will be able to:
CO1: Basic knowledge in database systems.
CO2: Enable the students to be well placed in leading business organizations anywhere in the world.
CO3: To acquaint the students with the use of current relational database systems.
CO4: Ability to build a solid foundation for advanced studies in database area.

Unit	DSE18STA11: Data Base management System (DBMS)	No. of hours
I	**Introduction to Databases and Data Models** Concept of database system, purpose of database system, view of data, different sources of databases, relational databases, database architecture, importance of data models, basic building blocks, business rules, the evolution of data models, data abstraction, database users and administrators.	15
II	**Introduction to Data Models and Normalization Database design and ER Model:** Overview, ER-Model, Constraints, ER-Diagrams, ERD issues, weak entity sets, Codd's rules, relational schema, introduction to UML relational database model: logical view of data, keys, and integrity rules. Relational Database design: features of good relational database design, atomic domain and normalization (1NF, 2NF, 3NF, BCNF).	15
III	**Introduction to SQL** **Constraints:** Concept and types of constrains. **Views:** Introduction to views, data independence, security, updates on views, comparison between tables and views. **SQL:** Basics of SQL, DDL, DML, DCL, TCL, DQL commands. **Defining constraints:** primary key, foreign key, unique, not null, check, **Operators in SQL**: Arithmetic operator, Comparison operator, Logical operator, Special operators. **Functions:** aggregate functions, built-in functions, numeric, date, string functions, set operations, sub-queries, correlated sub- queries, **Clauses:** use of group by, having, order by. **Join**: join and its types, exist, any, all, view and its types joined relations. Triggers.	15
IV	**Transaction management:** ACID properties, serializability and concurrency control, lock-based concurrency control (2PL, Deadlocks), time stamping methods, optimistic methods, database recovery management, data dictionary. **NoSQL:** Overview, and history of NoSQL databases, definition of the four types of NoSQL database, introduction to Big Data.	15

Books Recommended:

1.Abraham Silderschutz, H. Korth and S. Sudarshan: "Database systems concepts ",6th Edition, McGraw Hill Education
2. Peter Rob, Carlos Coronel: Database Systems: Design, Implementation, &Management.
3. C.J.Date, Longman, Introduction To Database Systems, Pearson Education
4. Raghu Ramakrishnan, Johannes Gehrke, Database Management Systems, McGraw-Hill

DSE18STA12: Statistical Analysis Using Minitab

Course Outcomes: At the end of the course students will be able to:

CO1: Enter and manipulate data efficiently.

CO2: Learn diagrammatical and graphical representation of data.

CO3: Compute descriptive statistics and probabilities.

CO4: Model and test hypothesis using Minitab.

Unit	DSE18STA12: Statistical Analysis Using Minitab	No. of hours
I	**Introduction** To Minitab How to Use this Guide , Assumptions, Register as a MINITAB User, Global Support, Customer Support, MINITAB on the Internet, About the Documentation, **Basic Statistics** Basic Statistics Overview, Descriptive Statistics Available for Display or Storage, Display Descriptive Statistics, Store Descriptive Statistics, One-Sample Z-Test and Confidence Interval, One-Sample t-Test and Confidence Interval, two-Sample t-Test and Confidence Interval, Paired t-Test and Confidence Interval, Test and Confidence Interval of a Proportion, Test and Confidence Interval of Two Proportions, Test for Equal Variances, Correlation, Covariance, Normality Test. **Regression** Regression Overview, Regression, Stepwise Regression, Best Subsets Regression, Fitted Line Plot, Residual Plots, Logistic Regression Overview, Binary Logistic Regression, Ordinal Logistic Regression, Nominal Logistic Regression.	15
II	**Analysis of Variance** Analysis of Variance Overview, One-Way Analysis of Variance, Two-Way Analysis of Variance, Analysis of Means, Overview of Balanced ANOVA and GLM, Balanced ANOVA, General Linear Model, Fully Nested ANOVA, Balanced MANOVA, General MANOVA, Test for Equal Variances, Interval Plot for Mean, Main Effects Plot, Interactions Plot. **Multivariate Analysis** Multivariate Analysis Overview, Principal Components Analysis, Discriminant Analysis. **Nonparametric Test** Overview of Nonparametric test,One-Sample Sign Test ,One-Sample Wilcoxon Test, Two-Sample Mann-Whitney Test, Kruskal -Wallis Test for a One-Way Design,Mood's Median Test for a One-Way Design, Friedman Test for a Randomized Block Design, Runs Test, Pairwise Averages, Pairwise Differences ,Pairwise Slopes.	15
III	**Tables** Tables Overview ,Arrangement of Input Data, Cross Tabulation, Tally Unique Values, Chi-Square Test for Association ,Chi-Square Goodness-of-Fit Test, Simple Correspondence Analysis **Time Series** Overview of Time Series ,Trend Analysis, Decomposition,Moving Average, Single Exponential Smoothing,Double Exponential Smoothing,Winters' Method,Differences,Lag,Autocorrelation,Partial Autocorrelation,Cross Correlation,ARIMA.	15

	Exploratory Data Analysis Overview of Exploratory Data Analysis,Letter Values,Median Polish,Resistant Line,Resistant Smooth,Rootogram. **Power and Sample Size** Power and Sample Size Overview,Z-Test and t-Tests,Tests of Proportions,One-Way Analysis Of Variance,Two-Level Factorial and Plackett-Burman Designs	
IV	**Quality Planning Tools** Quality Planning Tools Overview, Run Chart ,Pareto Chart, Cause-and-Effect Diagram, Multi- Vari Chart ,Symmetry Plot ,Variables Control Charts, Variables Control Charts Overview, Defining Tests for Special Causes ,Box-Cox Transformation for Non-Normal Data, Control Charts for Data in Subgroups, X bar Chart ,R Chart, S Chart ,X bar and R Chart ,X bar and S Chart, Moving Range Chart, Control Charts Using Subgroup Combinations ,EWMA Chart ,Moving Average Chart, CUSUM Chart Control Charts for Short Runs Attributes Control Charts Overview ,P Chart1,NP Chart ,C Chart,U Chart Options for Attributes Control Charts **Process Capability** Process Capability Overview, Capability Analysis (Normal Distribution) ,Capability Analysis (Between/Within) ,Capability Analysis (Weibull Distribution) ,Capability Six pack (Normal Distribution) ,Capability Six pack (Between/Within) ,Capability Six pack (Weibull Distribution) ,Capability Analysis (Binomial) ,Capability Analysis (Poisson)	15

Books Recommended:

1. Robert B. Miller, Minitab Handbook for Business and Economics (Duxbury series in statistics & decision sciences) Second edition.PWS
2. Ruth K. Meyer, David Krueger A Minitab Guide to statistics, Second edition, Pearson

3.Rehman M. Khan Problem Solving and Data Analysis Using Minitab, Wiley

4. McLaughlin, Kathleen; Wakefield, Dorothy, Introduction to Data Analysis Using Minitab for Windows, An – Softcover ,Third Edition Pearson.

RMD18STA11: Research Methodology

Course Outcomes: At the end of the course students will be able to:
CO1: Understand the concept of research, research process, and research ethics.
CO2: Understand and apply various sampling methods for data collection and estimate the parameters.
CO3: Understand the concept of simulation and able to simulate real life processes
CO4: Apply numerical methods to solve systems of linear equations and definite integrals.

Unit	RMD18STA11: Research Methodology	No. of hours
I	Meaning of research, objectives of research, motivation in research, types of research, research approaches, significance of research, research methods vs. methodology, research and Scientific method, research process, criteria of good research, defining research problem, Research Ethics, publication of research, Plagiarism, Intellectual property rights, Patents and its filing procedures.	15
II	Role of statistics in scientific research, research design, statistical research project Types of statistical research: empirical, field experiments, laboratory experiments, and secondary sources of data, exploratory and confirmatory research, planned and ad-hoc methods of data collection, non-response and methods of recovering the missing response.	15
III	Concept and need of simulation, random number generator, true random number and pseudo random number generators, requisites of a good random number generator. Tests for randomness. Congruential method of generating uniform random numbers. Algorithms for generating random numbers from well-known univariate discrete and continuous distributions, generating random vectors from multinomial, bivariate normal, and bivariate exponential distributions, generating random numbers from mixture of distributions (related results without proofs). Acceptance-Rejection Technique.	15
IV	Resampling methods: Bootstrap methods, estimation of bias and standard errors, estimation of sampling distribution, confidence intervals. Jackknife method: estimation of bias and standard errors, bias reduction method. Numerical methods for solution to system of linear equations: Jacobi and Gauss-Seidel methods with convergence analysis. Numerical methods for finding roots of nonlinear equation: Newton-Raphson method, bisection method; Newton-Raphson for system of non- linear equations.	15

Books Recommended:

1. Atkinson, K. E. (1989). *An introduction to numerical analysis*, John Wiley and Sons.
2. Chaudhuri, A., & Stenger, H. (2005). *Survey sampling: theory and methods*. CRC Press.
3. Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.
4. Kennedy, W. J., & Gentle, J. E. (2021). *Statistical computing*. Routledge.
5. Kothari, C. R. (2004). *Research methodology: Methods and techniques*. New Age International.
6. Morgan, B. J. (1984). *Elements of simulation* (Vol. 4). CRC Press.
7. Mukhopadhyay, P. (2008). *Theory and methods of survey sampling*. PHI Learning Pvt. Ltd..
8. Ross, S. M. (2022). *Simulation*. Academic Press.

10. Sukhatme P. V., Sukhatme S. & Ashok C (1984). *Sampling Theory of surveys and applications.* Iowa university press and Indian society of agricultural statistics, New Delhi.

DSC18STA19: PRACTICAL -I

Practical Number	Practical Name
1	Sketching of pdf and CDF for Discrete distribution
2	Sketching of pdf and CDF for Continuous distribution
3	Mixture of distribution
4	Applications of Bivariate Normal distribution & Multivariate Normal distribution.
5	Sufficient, minimal sufficient, and complete sufficient statistics
6	UMVUE and lower bunds for variances of unbiased estimators
7	Maximum likelihood and method of moments estimation
8	Method of Scoring and method of minimum chi-square estimation
9	Construction of Consistent/CAN & BAN Estimators.
10	Confidence interval based on CAN &BAN.
11	Practical on MSEXCEL
12	Practical on R- Software
13	Applications of Simulation techniques I
14	Applications of Simulation techniques II
15	Numerical Methods
16	Resampling Techniques
17	2 to 3 Practicals on elective Paper

(Each practical should consist of problems to be solved using at least two of the following software: EXCEL/ R/python)

Semester II

DSC18STA21: PROBABILITY THEORY

Course Outcomes: At the end of the course students will be able to:

CO1: Memorize the basic concepts of Sets, Sequence, Measurable function and limit.

CO2: Recognize the measure theory, random variable, distribution function, limit of sequencevariables.

CO3: Explain the concept of convergence and applications with example.

CO4: Implement the central limit theorem and large-sample approximations for common statistics

Unit	DSC18STA21: PROBABILITY THEORY	No. of hours
I	Classes of sets: Sequence of sets: limsup, liminf and limit of sequence of sets field, σ- field, σ- field generated by a class of sets, Borel σ- field. Probability measure, Probability space, properties of a probability measure, continuity, mixture of probability measures. Lebesgue and Lebesgue- Steltjes measures on R. Independence of events.	15
II	Measurable function, random variable, distribution function of a random variable, simple random variable, elementary random variable, liminf, limsup and limit of sequence of random variables. Method of obtaining a random variable as a limit of sequence of simple random variables. Integration of a measurable function with respect to a measure, expectation of a random variable, independence. Characteristic function, simple properties. Inversion theorem and uniqueness property (Statement only).	15
III	Monotone convergence theorem, Fatous Lemma, Dominated Convergence theorem, Borel- Cantelli Lemma, (Statements only), and their applications. Convergence of sequence of random variables, Convergence in distribution, Almost sure convergence, a characterizing property, convergence in probability, uniqueness of limit, Yule Slutsky results and preservation under continuous transform. Convergence in r^{th} mean, interrelationships (Statements only), their illustration with examples	15
IV	Characteristic function: definition and properties of characteristic function, inversion formula (without proof), characteristic function and moments. Laws of large numbers: Convergence of distribution functions, convergence of series of independent random variables, Kolmogorov inequalities and almost sure convergence, weak law of large numbers (without Proof) for iid and non-iid random variables, Strong law of large numbers (without Proof). Central limit Theorem (CLT) (without proof): Lindeberg-Levy, Liaponove's and Lindeberg-Feller forms and applications.	15

Books Recommended:

1. Bhat B. R. (1981): Modern Probability Theory –3^{rd} edition: New age international limited.
2. Alan Karr, (1993) : Probability Theory – Springer Verlag.

3. Billingsley P. (1986): Probability & Measure –John Wiley and sons
4. Athreya K. B. and Lahiri S. (2006). Probability Theory, (Hindustan Book Agency).
5. Feller, W. (1969). Introduction to Probability and its Applications vol.II (Wiley Eastern Ltd.)
6. Loeve, M. (1978). Probability Theory (Springer Verlag). Fourth edition
7. Rohatgi, V.K. and Saleh, A. K. MD. E. (2015). Introduction to Probability Theory andMathematical Statistics -3rd Edition, John Wiley & sons.

DSC18STA22: THEORY OF TESTING OF HYPOTHESIS

Course Outcomes: At the end of the course students will be able to:

CO1: Formulate null and alternative hypotheses, compute probabilities of types of error, MP tests and MLR property.

CO2: Understand UMP and UMPU test with their applications.

CO3: Construct asymptotic confidence interval of a parameter and its relation with testing of hypothesis problem.

CO4: Execute small, large sample size tests and non-parametric tests in real life problems.

Unit	DSC18STA22: THEORY OF TESTING OF HYPOTHESIS	No. of hours
I	Problem of testing of Hypothesis, Simple and composite hypotheses. Randomized and non- randomized tests, Most powerful test, Neyman-Pearson Lemma and its applications. Determination of minimum sample size to achieve the desired strengths. Monotone likelihood ratio property, UMP test, power function of a test, existence of UMP, Tests for one-sided alternatives. Concept of p- value.	15
II	UMP tests for two sided alternatives examples, their existence and non-existence. Generalized Neyman Pearson lemma, unbiased test, UMPU test and their existence in the case of exponential families (Statements of the theorems only). Similar tests, test with Neyman structure.	15
III	Problem of confidence intervals, relation with testing of hypotheses problem, shortest length confidence intervals, UMA and UMAU confidence intervals. Asymptotic Confidence Intervals based on CAN estimators, Variance stabilizing transformations (VST), confidence interval based on VST, Asymptotic Confidence regions.	15
IV	Likelihood ratio test and its application to standard distribution. Goodness of fit tests based on Chi-square distribution and application to contingency tables. Spearman's Rank Correlation Test; Kendall's Rank Correlation Test; Kruskal-Wallis Test; Friedman's Two-way analysis of variance by ranks. Wald test, Rao's Score test, Bartlett's test for homogeneity of variances, Consistent test	15

Books Recommended:

1. V. K. Rohatgi, and A. K. MD. E. Saleh (2015): Introduction to Probability Theory and Mathematical Statistics,John Wiley & sons, 3rd Edition.
2. B. K. Kale, and K. Muralidharan (2015), Parametric Inference: An Introduction, Alpha Science InternationalLtd.
3. E. J. Dudewicz and S. N. Mishra (1988): Modern Mathematical Statistics, John Wiley and Sons.
4. E. L. Lehmann (1983): Theory of Point Estimation, John Wiley & sons.
5. T. S. Ferguson, Mathematical Statistics (1967): A decision theoretical approach, Academic press.
7. S. Zacks (1971): Theory of Statistical Inference, John Wileyand Sons, New York.
8. R. H. Randles (1979): and D. A. Wolfe, Introduction to theory of nonparametric Statistics, Wiley.

9. J. D. Gibbons and S. Chakraborti(2010): Nonparametric Statistical Inference, CRC Press, Fifth Edition.

DSC18STA23: Multivariate Analysis

Course Outcomes: At the end of the course students will be able to:

CO1: Review of multivariate normal distribution and their real-life applications.

CO2: Understand Wishart distribution, Hotelling T2 and Mahalanobis D2 statistic.

CO3: Implement dimension reduction techniques using software on real life problems.

CO4: Demonstrate knowledge of the basic ideas behind discriminant and clustering analysis techniques with applications.

Unit	DSC18STA23: Multivariate Analysis	No. of hours
I	Review of Multivariate Normal distribution, Hotelling's T^2 Statistic and its null distribution. Applications of T^2 statistics and its relationship with Mahalanobis' D^2 statistic. Confidence region for the mean vector, Maximum likelihood estimators of the parameters of the multivariate normal distribution and their sampling distributions, Wishart matrix and its distribution (Without Proof), properties of Wishart distribution, distribution of generalized variance.	15
II	Discrimination and classification. Fisher's discriminant function and likelihood ratio procedure, minimum ECM rule, Rao's U statistics and its use in tests associated with discriminant function, classification with three populations. Cluster analysis, Hierarchical methods: Single, Complete, average linkage method and non-hierarchical clustering method-k means clustering. Canonical correlation analysis, Introduction to principal component analysis and related results, Introduction to factor analysis and estimation.	15

Books Recommended:

1.Kshirsagar A. M. (1972): Multivariate Analysis. Marcel-Dekker.

2. Johnson, R.A. and Wichern. D.W (2002): Applied multivariate Analysis. 5thAd.Prentice – Hall.

3. Anderson T. W. (1984): An introduction to Multivariate statistical Analysis2nd Ed. John Wiely.

4. Morrison D.F. (1976): Multivariate Statistical Methods McGraw-Hill.

DSE18STA21: Linear Models and Regression Analysis

Course Outcomes: At the end of the course students will be able to:

CO1: Understand General linear model, Gauss Markov theorem, variances and covariance's of BLUEs.

CO2: Understand and apply multiple regression models in real life situations.

CO3: Understand concept of multicollinearity and non-linear regression.

CO4: Understand concept of Robust regression, Logistic regression and Poisson regression.

Unit	DSE18STA21: Linear Models and Regression Analysis	No. of hours
I	General linear model: definition, assumptions, concept of estimability, least squares estimation, BLUE, estimation space, error space, Gauss Markov theorem, variances and covariances of BLUEs, Distribution of quadratic forms for normal variables: related theorems (without proof), Tests of hypotheses in general linear models. Description of the ANOVA and linear regression models as the particular cases of the general linear model.	15
II	Multiple regression model, Least squares estimate (LSE), Properties of LSE, Hypothesis testing, confidence and prediction intervals, General linear hypothesis testing. Dummy variables and their use in regression analysis. Model adequacy checking. Transformations to correct model inadequacies: VST and Box-Cox power transformation.	15
III	Multicollinearity: Sources, Consequences, detection and remedies, ridge regression. Autocorrelation: sources, consequences, detection (Durbin-Watson test) and remedies. Parameter estimation using Cochrane-Orcutt method. Variable Selection Procedures: Rsquare, adjusted R-square, Mallows' Cp, forward, backward and stepwise selection methods, AIC, BIC. Robust Regression: need for robust regression, M-estimators, properties of robust estimators: breakdown and efficiency. Asymptotic distribution of M-estimator (Statement only).	15
IV	Generalized linear models: concept of generalized linear model, Link function, ML estimation, large sample tests about parameters, goodness of fit, analysis of deviance. Residual analysis, types of residuals: raw, Pearson, deviance, Anscombe, quantile; residual plots. Variable selection: AIC and BIC. Logistic regression: logit, probit and cloglog models for dichotomous data, ML estimation, Odds ratio and its interpretation, hypothesis tests about model parameters. Hosmer-Lemeshow test, multilevel logistic regression, Logistic regression for Nominal response. Poisson regression.	15

Books Recommended:

1. Birkes, D., & Dodge, Y. (2011). Alternative methods of regression. John Wiley & Sons.
2. Cook, R. D., & Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman and Hall.
3. Draper, N. R., & Smith, H. (1998). Applied regression analysis. John Wiley & Sons.
4. Huber,P.J.andRonchetti,E.M (2011)Robust Statistics,Wiley,2ndEdition.

5. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Wasserman, W. (2004). Applied linear regression models. New York: McGraw-Hill/Irwin.

6. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to linear regression analysis. 5th Ed. John Wiley & Sons.
7. Seber,G.A.,Wild,C.J.(2003).Non linear Regression,Wiley.

DSE18STA22: Statistical Analysis System (SAS)

Course Outcomes: At the end of the course students will be able to:

CO1: Understand the SAS environment, its components, and basic programming.

CO2: Perform data handling, preprocessing, and management using SAS procedures.

CO3: Conduct statistical analysis, hypothesis testing, regression, and ANOVA using SAS.

CO4: Use SAS for visualization, report generation, and automation of statistical workflows.

Unit	DSE18STA22: Statistical Analysis System (SAS)	No. of hours
I	**Introduction to SAS and Data Handling:** Overview of SAS and its applications, SAS interface and programming structure, data import/export (Excel, CSV, TXT, SQL), creating and modifying datasets (SET, MERGE, APPEND), handling missing values, applying arithmetic, character, and date functions, sorting, filtering, and subsetting data.	15
II	**Descriptive Statistics and Data Visualization:** Measures of central tendency and dispersion using PROC MEANS, PROC SUMMARY, and PROC UNIVARIATE, frequency analysis (PROC FREQ), data visualization techniques (bar charts, histograms, box plots, scatter plots) using PROC SGPLOT and PROC GPLOT, probability distributions (Normal, Binomial, Poisson) with PROC RAND, data transformation methods, and introduction to macro variables.	15
III	**Inferential Statistics and Regression Analysis:** Hypothesis testing (t-tests, chi-square, ANOVA, non-parametric tests) using PROC TTEST, PROC FREQ, PROC ANOVA, and PROC NPAR1WAY, correlation analysis (Pearson & Spearman using PROC CORR), regression techniques (simple, multiple, and logistic regression using PROC REG and PROC LOGISTIC), model selection and validation (R-squared, AIC, BIC), residual analysis, assumption checking, and introduction to time series analysis.	15
IV	**Advanced Data Processing, Macros, and Report Generation:** SQL programming in SAS (PROC SQL), data selection, filtering, grouping, and joins, automation with SAS Macros, macro variables and functions, handling large datasets, user-defined formats, report generation using PROC REPORT, exporting results to Excel, PDF, and RTF, and case studies in healthcare, finance, and market research.	15

Books Recommended:

1. Delwiche, L. D., & Slaughter, S. J. (2012). The Little SAS Book: A Primer. SAS Institute Inc.
2. Cody, R. P., & Smith, J. K. (2005). Applied Statistics and the SAS Programming Language. Pearson Prentice Hall.
3. Elliott, A. C., & Woodward, W. A. (2014). SAS Essentials: Mastering SAS for Data Analytics. Wiley.
4. Marasinghe, M. G., & Kennedy, W. J. (2007). SAS for Data Analysis. Wiley.
5. Bailer, A. J. (2004). Statistical Programming in SAS. Wiley-Interscience.

DSE18STA23: Power BI and Tableau

Course Outcomes: At the end of the course students will be able to:
CO1: Understand data preparation, transformation, and integration in Power BI and Tableau.
CO2: Apply statistical functions, predictive analytics, and trend analysis in both tools.
CO3: Create interactive dashboards and reports for effective data visualization.
CO4: Utilize advanced analytics, automation, and storytelling techniques for decision-making.

Unit	DSE18STA23: Power BI and Tableau	No. of hours
I	**Data Preparation and Transformation in Power BI** data connections, transformations, and integration with sources like Excel and SQL. Students will use Power Query for data cleaning, filtering, and merging datasets. It introduces DAX functions for statistical calculations, including averages, variance, and time-based aggregations like YTD and MTD.	15
II	**Advanced Visualization and Predictive Analytics in Power BI** creating interactive dashboards using slicers, filters, and advanced charting techniques. Students will integrate Python and R for predictive modeling, trend forecasting, and machine learning. Report automation and sharing through Power BI Service are also covered.	15
III	**Data Processing and Statistical Functions in Tableau** data handling, calculated fields, and statistical functions like correlation, regression, and forecasting. The unit covers table calculations, Level of Detail (LOD) expressions, and predictive analytics for trend analysis.	15
IV	**Dashboard Development and Data Storytelling in Tableau** interactive dashboard design using filters, tooltips, and parameter controls. Students will explore best practices in visualization, storytelling with Tableau Storyboards, and publishing reports via Tableau Public and Server.	15

Books Recommended:

1. Powell, B. (2017). Mastering Microsoft Power BI. Wiley.
2. Ferrari, A., & Russo, M. (2015). Analyzing Data with Power BI and Power Pivot for Excel. Microsoft Press.
3. Ferrari, A., & Russo, M. (2015). The Definitive Guide to DAX: Business Intelligence with Microsoft Power BI. Apress.
4. Murray, D. G. (2013). Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software. Wiley.
5. Milligan, J. N. (2013). Learning Tableau: Business Intelligence and Data Visualization. Wiley.

DSCPR12: PRACTICAL -II

Course Objective: Students should to understand and implement theory in real life problems.

Practical Number	Practical Name
1	Probability Measure and Independence of Events
2	Calculation of Expectation and Distribution Function of a Random Variable
3	Convergence of Sequence of Random Variables and Monotone Convergence Theorem
4	Application of Characteristic Function and Laws of Large Numbers
5	MP, UMP, and UMPU Tests
6	Likelihood ratio tests
7	Confidence Intervals
8	Non-parametric Tests
9	Exploratory data analysis.
10	Application of Hotelling's T^2 statistics
11	Discriminant Analysis
12	Principle component analysis and Factor Analysis.
13	At least two practical's on elective paper

(Each practical should consist of problems to be solved using at least two of the following software: EXCEL/ R/python)

VIVEKANAND COLLEGE KOLHAPUR ESTD JUNE 1964

HEAD
DEPARTMENT OF STATISTICS
VIVEKANAND COLLEGE, KOLHAPUR
(EMPOWERED AUTONOMOUS)