ANDHRA KESARI UNIVERSITY

UNIVERSITY COLLEGE OF SCIENCE, ARTS, COMMERCE & LAW ONGOLE, PRAKASAM-523001, ANDHRA PRADESH

M.Sc. Statistics

Revised syllabus: 2023-24

(Revised Scheme of Instruction and Examination, Syllabus etc.,with effect from the Academic Year 2023-24)

DEPARTMENT OF STATISTICS
ANDHRA KESARI UNIVERSITY :: ONGOLE

ANNEXURE - I

(Agenda item No.1, P.G.B.O.S. meeting dated 05-08-2023)

M.Sc. Statistics

SYLLABUS:: SEMESTER-I

W.E.F. 2023-2024 ADMITTED BATCH OF STUDENTS

(Ten Questions are to be set TWO on each of Units I-V with the choice of ONE question from each unit)

ANDHRA KESARI UNIVERSITY

UNIVERSITY COLLEGE OF SCIENCES

DEPARTMENT OF STATISTICS

M.Sc. STATISTICS

TWO YEAR M.Sc. COURSE IN STATISTICS COURSE STRUCTURE AND EXAMINATION SCHEME

(with effect from 2023-2024 admitted batch of students)

SEMESTER-I

S. No.	Components of Study	Course Code	Title of the Course	Contact Hours / week	No. of Credits	Internal Assessment Marks	Semester End Examination Marks	Total Marks
1	**Core**	**ST 1.1**	Probability Theory and Distributions	6	4	30	70	100
2		**ST 1.2**	Statistical Computing Using R	6	4	30	70	100
3	**Compulsory Foundation**	**ST 1.3**	Estimation	6	4	30	70	100
4	***Elective Foundation**	**ST 1.4 (A)**	A) Sampling Theory	6	4	30	70	100
		ST 1.4 (B)	B) Linear Algebra					
5	**Practical-I**	**ST 1.5**	Papers on ST 1.1 & ST 1.2 - Statistical Software Practical using R	6	4	30	70	100
6	**Practical-II**	**ST 1.6**	Papers on ST 1.3 & ST 1.4	6	4	30	70	100
TOTAL				**36**	**24**	**180**	**420**	**600**

*** The student shall choose ONE PAPER from Elective Foundation**

SEMESTER-II

S. No.	Components of Study	Course Code	Title of the Course	Contact Hours / week	No. of Credits	Internal Assessment Marks	Semester End Examination Marks	Total Marks
1	**Core**	**ST 2.1**	Multivariate Analysis	6	4	30	70	100
2		**ST 2.2**	Testing of Hypothesis	6	4	30	70	100
3	**Compulsory Foundation**	**ST 2.3**	Theory of Linear Estimation and Analysis of Variance	6	4	30	70	100
4	***Elective Foundation**	**ST 2.4 (A)**	A) Stochastic Processes	6	4	30	70	100
		ST 2.4 (B)	B) Linear Models and AppliedRegression Analysis					
5	**Practical-I**	**ST 2.5**	Papers on ST 2.1 & ST 2.2 - Statistical Software Practical using SPSS and R	6	4	30	70	100
6	**Practical-II**	**ST 2.6**	Papers on ST 2.3 & ST 2.4	6	4	30	70	100
7	**Skill Development Course**	---	MOOCS Course	---	4	---	---	---
TOTAL				**36**	**28**	**180**	**420**	**600**

*** The student shall choose ONE PAPER from Elective Foundation**

SEMESTER-III

S. No.	Components of Study	Course Code	Title of the Course	Contact Hours / week	No. of Credits	Internal Assessment Marks	Semester End Examination Marks	Total Marks
1	**Core**	**ST 3.1**	Design of Experiments	6	4	30	70	100
2		**ST 3.2**	Statistical Quality Control	6	4	30	70	100
3	**Elective - I**	**ST 3.3 (A)**	A) Actuarial Statistics	6	4	30	70	100
		ST 3.3 (B)	B) Time Series Analysis					
4	**Elective - II**	**ST 3.4 (A)**	A) Reliability – I	6	4	30	70	100
		ST 3.4 (B)	B) Biostatistics					
5	**Practical-I**	**ST 3.5**	Papers on ST 3.1 & ST 3.2 - Statistical Software Practical using SPSS and R	6	4	30	70	100
6	**Practical-II**	**ST 3.6**	Papers on ST 3.3 & ST 3.4	6	4	30	70	100
7	**Skill Enhanced Course**	**---**	MOOCS Course	**---**	4	**---**	**---**	**---**
TOTAL				**36**	**28**	**180**	**420**	**600**

***The student shall choose <u>ONE PAPER</u> from each of Elective-I and Elective-II**

SEMESTER-IV

S. No.	Components of Study	Course Code	Title of the Course	Contact Hours / week	No. of Credits	Internal Assessment Marks	Semester End Examination Marks	Total Marks
1	**Core**	**ST 4.1**	Econometrics	6	4	30	70	100
2		**ST 4.2**	Acceptance Sampling Plans	6	4	30	70	100
3	**Elective - I**	**ST 4.3 (A)**	A) Operations Research	6	4	30	70	100
		ST 4.3 (B)	B) Machine Learning Using Python					
4	**Elective - II**	**ST 4.4 (A)**	A) Reliability – II	6	4	30	70	100
		ST 4.4 (B)	B) Knowledge Discovery and Data Mining					
5	**Practical-I**	**ST 4.5**	Papers on ST 4.1 & ST 4.2 - Statistical Software Practical using SPSS and R	6	4	30	70	100
6	**Practical-II**	**ST 4.6**	Papers on ST 4.3 & ST 4.4	6	4	30	70	100
7	**Project Work**			---	4	---	100	100
TOTAL				**36**	**28**	**180**	**420**	**700**
TOTAL SEMESTERS: I + II + III + IV								**2500**

The student shall choose ONE PAPER from each of Elective-I and Elective-II

Course - STA 1.1 :: PROBABILITY THEORY AND DISTRIBUTIONS

Course Code: STA 1.1

Course Objectives:

CO1: To explore the students to study the fundamental concepts of probability and probability measures.
CO2: To get expertise in analyzing the random phenomenon through convergence of random variable.
CO3: To be able to understand different types of inequalities associated with mathematical expectations.
CO4: To able to understand the concept of law of large numbers and their applications.
CO5: To get expertise in understanding the advanced probability distributions and utilizing them to model the random phenomenon.
CO6: To understand Discrete and Continuous distributions and its applications in data analysis.
CO7: To acquainted with distributions of order statistics of various random variable and their applications.

Learning Outcomes:

After successful completion of this course, student will be able to:

LO1: Aanalyse datasets with sound probability theory for scientific knowledge discovery.
LO2: Understand and analyse random phenomenon of different events.
LO3: Model stochastic nature through probability distributions.
LO4: Apply probability and distribution theory in solving practical problems.
LO5: Fit appropriate distribution for the given grouped data.

UNIT-I

Classes of sets, fields, σ fields, minimal σ fields, sequence of sets, limit supremum and limit infimum of sequence of sets, measure, probability measure, properties of measure, axiomatic definition of probability, continuity theorem of probability, conditional probability, statistical independence of events, , probability on finite sample spaces, geometrical probability.

UNIT-II

Measurable functions, notation of random variable, distribution function, properties of distribution, vector of random variables, statistical independence, concepts of joint, marginal and conditional distributions, mathematical expectation, conditional expectation, characteristic function, its properties. Inversion formula, characteristic functions and moments. Moments inequalities-Markov, Schwartz, Jensen, Holder, Minkowski, Kolmogrove's, Hajek-Renyi.

UNIT-III

Convergence of sequence of random variables-Type of convergence-in probability, almost sure, in mean square, in law- their interrelations. Law of large numbers-weak laws: Chebychevs's form of W.L.LN., Necessary and Sufficient Condition of W.L.L.N. Kintchines form of W.L.L.N., Kolmogrove's S.L.LN for i.i.d. random variables.

UNIT-IV

Discrete distributions - Compound Binomial, Compound Poisson, multinomial, truncated Binomial, truncated Poisson distributions and their properties. Continuous distributions- Laplace, Weibull, Logistic and Pareto distributions and their properties.

UNIT-V

Order statistics- distribution function, probability density function (p.d.f.) of single order statistic, joint p.d.f. of order statistics. Distribution of range with applications in rectangular and

exponential cases.

Books for study:

1. Modern probability theory by B.R. Bhat, Wiley Eastern Limited.
2. An introduction to probability theory and mathematical statistics by V.K.Rohatgi, John Wiley.
3. An Outline of statistics theory-1, by A.M.GOON, M.K. Gupta and B. Dasgupta, the World Press Private Limited, Calcutta.
4. The Theory of Probability by B.V. Gnedenko, MIR Publishers, Moscow.
5. Discrete distributions - N.L. Johnson and S. Kotz, John wiley & Sons.
6. Continuous Univariate distributions, vol.1&2- N.L.Johnson and S.Kotz, John Wiley & Sons.
7. Mathematical Statistics - Parimal Mukopadhyay, New Central Book Agency (P) Ltd., Calcutta.

Books for References:

1. Billingsley, P. (1986): Probability and Measure. Wiley.
2. Kingman, J F C and Taylor, S. J. (1966): Introduction to Measure and Probability. Cambridge University Press.
3. David, H.A (1981): Order Statistics, 2nd Ed, John Wiley.
4. David H. A. and Nagaraja H.N.(2003): Order Statistics, 3/e, John Wiley & Sons.
5. Feller, W (1966): Introduction to probability theory and its applications, Vol.II, Wiley.
6. Cramer H (1946): Mathematical Methods of Statistics, Princeton University Press.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	L	S	M	M
CO2	S	S	S	L	S	M	M
CO3	S	S	S	L	S	L	M
CO4	S	S	S	L	S	L	M
CO5	S	S	S	M	S	S	S
CO6	S	S	S	M	S	S	S
CO7	S	S	M	L	S	M	S

* S-Strong; M-Medium; L-Low

Course - STA 1.2 :: STATISTICAL COMPUTING USING R

Course Code: STA 1.2

Course Objectives:

CO1: To expose the students to R- statistical programming language developed by scientists that has open-source libraries for statistics, machine learning, and data science. It is a free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing.

CO2: The student will learn a large coherent and integrated collection of statistical tools available in R. Further, the student will be familiar with graphical facilities for data analysis available in R.

CO3: R is widely used by statisticians, data scientists and major corporations like Google, Airbnb, Facebook etc. for data analysis.

CO4: R is well-developed, simple and effective programming language for data handling and storage facility.

CO5: R is data analysis software includes a wide variety of functions, such as data manipulation, statistical modeling, and graphics. One really big advantage of R, however, is its extensibility. Developers can easily write their own software and distribute it in the form of add-on packages.

Learning Outcomes:

After learning the course the student

LO1: Will be able to handle the data analysis using the R-statistical tools and can also perform graphical presentation of the data.

LO2: Will be able to write own R-scripts for handling the data analysis in their own way.

LO3: Has better job potentiality to acquire opportunities as data analyst in software companies, clinical trial data analysis companies and etc.

UNIT – I

Introduction to R language: Objects (Atomics) -Basic types, modes and attributes, comments, constants. R–Data Types: character, numeric, integer, logical, complex and raw data types. R–Operators: arithmetic, relational, logical, assignment and miscellaneous operators. R–Variables: variable assignment, data type, finding variables using ls()function, deleting Variables using rm() function, R-I/O console functions-scan(), print(), cat(), format(), setwd() and getwd() functions. R-vectors: creating vectors, vector assignment, manipulating vectors, arithmetic, generating regular sequences, logical vectors, and character vectors, index vectors, selecting and modifying subsets of a vector. Manipulating character vectors using strsplit(), paste(), grep(), gsub() functions; R-factors: creating factor variables, handling factor data, generating factor levels using gl() function.

UNIT – II

R-Matrices: Creating matrices, arithmetic operators on matrices, matrix facilities, forming partitioned matrices, cbind() and rbind() functions, R-Lists: creating a list, naming, accessing and manipulating list elements, converting a list to a vector. R-Data frames: creation, adding rows and variables to data frame, attach() and detach(), working with data frames, data reshaping. Reading and getting data into R using files: reading data and writing data from / to files of type CSV, EXCEL, text and other data type files using the save(), load(), read.csv() and read.table(),write.csv() and write.table() functions. Retrieving files using file.choose(),function.

UNIT – III

R – Control Structures: Decision making-if, if-else, ladder if-else, nested if-else, and switch statements. Loops-repeat, while and for statements. Loop control statements -- break and next. R – Functions:

function definition, function components, built-in functions, user-defined function, syntax of a function, function arguments, arguments matching, scope and evaluation, calling a function, one-line functions, using default values in functions. Built in R-functions and writing own R-functions or R-codes for small standard statistical problems like finding summary statistics, correlation, one-sample t-test, two-sample t-test and paired samples t-test, etc. Group manipulation using apply family of functions - apply, sapply, lapply and tapply.

UNIT – IV

R-Probability Distributions: Computing values of pdf, cdf, quantile and generating samples for bionomial, poission, normal, exponential, Weibull and other prominent distributions using Built in R – functions. Plotting density and cumulative density curves for the distributions. Built in R-syntaxes for the Shapiro-Wilk test of normality, Kolmogorov-Smirnov test for one-sample and two-sample cases, Wilcoxon Mann-Whitney one-sample and two-sample U-tests, chi-square tests for association and goodness of fit. Writing own R-functions or R-codes: Fitting of binomial, Poisson, normal, exponential, Weibull and logistic distributions based on a given frequency data and test for goodness of fit. Solving a non-linear equation using Newton-Raphson method.

UNIT – V

R-Graphics: Use of high-level plotting functions for creating histograms, scatter plots, box-whisker plots, bar plot, dot plot, line charts using numeric data and categorical data, pie charts, bar Charts, Q-Q plot and curves. Controlling plot options using low-level plotting functions, adding lines, segments, points, polygon, grid to the plotting region; Add text using legend, text, mtex; and modify/add axes, putting multiple plots on a single page. Built in R – syntaxes for one-way ANOVA, two- way ANOVA.

Books for study:

1. Dr. Mark Gardener (2012): *Beginning R – The Statistical Programming Language*, Wiley India Pvt Ltd.
2. W.N. Venables and D.M. Smith (2016) : *An Introduction to R*
3. J.P. Lander (2014): *R for Everyone, Pearson Publications*
4. Garrrett Grolemund: *Hands-On Programming with R*

Books for References:

1. De Vries, A., and Meys, J. (2016). *R For Dummies*, Second Edition, John Wiley & Sons Private Ltd, NY
2. Crawley, M, J. (2007). *The R Book*, John Wiley and Sons Private Ltd., NY.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	S	M	S	S
CO2	S	S	S	S	L	S	S
CO3	S	S	S	S	M	S	S
CO4	S	S	S	M	S	L	M
CO5	S	S	S	M	L	M	M

* S-Strong; M-Medium; L-Low

Course - STA 1.3 :: ESTIMATION

Course Code: STA 1.3

Course Objectives :

CO1: To introduce the concepts of Estimation.

CO2: Characteristic features of good estimator, properties of good estimator with examples.

CO3: To introduce the role and importance of C. R. Rao Inequality, Rao – Blackwell theorem, Lehmann-Scheffe theorem.

CO4: Different methods of estimation in point estimation and internal estimation.

CO5: To introduce concepts of Baye's estimation.

Learning Outcomes:

LO1: The concepts in theory of estimation, properties of good estimator and Fisher-Neymann factorization theorem.

LO2: Upon completion of this unit the student will be able to understand unbiasedness, UMVUE, consistency, Efficiency with CAN, CAUN estimators.

LO3: By completing this unit the student will know the different methods of estimation for variables and attributes.

LO4: At the end of this unit the student will distinguish the difference between point estimation and internal estimation.

LO5: Construction of confidence intervals, also able to understand the concepts of Baye's estimation.

UNIT-I

Concepts of population, parameter (scalar, vector), parametric space, sample, sample space, statistic, estimator, estimate, sampling distribution, standard error, etc. Problem of point estimation, properties of good estimator, sufficiency- concept with examples, distinction between joint density and likelihood function. Fisher Neyman Factorization theorem. Complete sufficiency-examples, Exponential class, Minimal sufficiency.

UNIT-II

Unbiasedness-concept, examples, properties, LMVUE, UMVUE, regularity conditions, Cramer-Rao Inequality and condition(s) for existence equality, examples of construction of UMVUE using Cramer-Rao Inequality, Rao-Blackwell Theorem, Lehmann-Scheffe Theorem, Necessary and Sufficient condition for the existence of MVUE. Minimum Mean Square Error (MMSE) Estimation. Consistency-Concept and examples, necessary condition for the existence of consistent estimator, efficiency, asymptotic relative Efficiency (ARE), CAN, CAUN estimators.

UNIT-III

Moment method of Estimation, ML method of Estimation, Percentile estimation, Minimum Chi- square and Modified Minimum Chi- square.

UNIT-IV

Interval Estimation, Confidence level, Construction of Confidence intervals using pivots, shortest expected length, UMA, UMAU Confidence sets. Relationship between confidence estimation and testing of hypothesis. Priori and posteriori distributions, loss function, risk function, Minmax & Bayes Estimator.

UNIT-V

Censored and truncated distributions: Type 1 and Type 2 Censoring for normal and exponential distributions and their MLE's. Interval estimation: Confidence Intervals, using pivots; shortest expected length confidence intervals.

Books for Study:

1. Statistical Inference by H.C,. Saxena & Surendran.
2. An Introduction to Probability and Statistics by V.K.Rohatgi and A.K.Md.E.Saleh(2001).
3. Mathematical Statistics- Parimal Mukopadhyay(1996), New Central Book Agency (P) Ltd., Calcutta.

Books for references:

1. An Outline of Statistical Theory, Vol.II by A.M.Goom, M.K. Gupta and B. Dasgupta (1980), World Press, Calcutta.
2. Linear Statistical Inference and its Application by C.R. RAO (1973), John Wiley.
3. A First Course on Parametric Inference by B.K. kale(1999) Narosa Publishing Co.,
4. Lehman, E. L., and Cassella, G. (1998). Theory of Point Estimation, Second Edition, Springer, NY.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	M	M	S	L	M	S
CO2	S	S	S	S	L	S	M
CO3	S	S	S	S	M	S	S
CO4	S	S	S	L	M	L	M
CO5	S	S	M	M	L	M	L

* S-Strong; M-Medium; L-Low

Course - STA 1.4 (A) :: SAMPLING THEORY

Course Code: STA 1.4 (A)

Course Objectives:

CO 1: Able to understand the basic concepts of various probability sampling techniques.
CO 2: Understand the selection of PPS sampling and able to calculate the Des Raj, Murthy's estimator Horvitz Thompson estimator.
CO 3: Able to estimate population mean and variance of systematic sample when population exhibits linear trend.
CO 4: Understand Cluster sampling and their applications by estimation of population mean and variance.
CO 5: Able to apply two stage sampling technique and also Two-stage PPS sampling technique.
CO 6: Understand the concept of ratio and regression methods of estimation in stratified random sampling.
CO 7: Acquainted with multiphase and multistage sampling.

Learning Outcomes:

At the End of this Course Students will be able :

LO1: Acquire knowledge to assess different sampling methods and their applications
LO2: Students acquire the basic knowledge to understand different advanced techniques of sampling methods.
LO3: To explain and to compare various allocations using stratified random sampling.
LO4: To use practical applications of ratio and regression method of estimation.
LO5: Students acquire the theoretical as well as practical knowledge of field study to analyze the data, interpret the results and draw valuable conclusions.

UNIT-I

Systematic Sampling: Allocation problem in stratified sampling, gain in precision due to stratification, estimation of sample size with continuous data, stratified sampling for proportions, Methods of populations with linear trend: Yates end correction, Modified systematic sampling, balanced systematic sampling, centrally located sampling, circular systematic sampling.

UNIT-II

Varying probability and Cluster sampling: Cluster sampling with equal and unequal cluster sizes, optimum cluster size for fixed cost. PPS sampling with and without replacements, procedures of selection of a sample, estimator of population total and its sampling variance in PPS with replacement, Des Raj and Murthy's estimator (for sample size two), Horvitz-Thomson estimator, Grundy's estimator, Midzuno-Sen Sampling Scheme

UNIT-III

Two-stage sampling: Two-stage sampling with equal number of second stage units, estimation of population mean, its variance and estimation of variance. Double sampling (two phase sampling) for stratification, variance of the estimated mean, optimum allocation in double sampling.

UNIT-IV

Multiphase Sampling: Introduction, Double sampling for Difference estimation. Double sampling for ratio estimation. Double sampling for regression estimator, Optimum allocation varying probability sampling. Non sampling errors: Sources and types of non Sampling errors, Non response errors, techniques for adjustment of non response, Hansen and Hurwitz Technique, Deming's Model.

UNIT-V

Ratio Estimator: Introduction, Bias and Mean square error, Estimation of variance, confidence interval, comparison with mean per unit estimator, Ratio estimator in stratified random sampling. Difference estimator and Regression estimator: Introduction, Difference estimator, Difference estimator in stratified sampling. Regression estimator, Comparison of regression estimator with mean per unit estimator and ratio estimator. Regression estimator in stratified sampling.

Books for study:

1. Sampling techniques by W.G. Cochran, John Wiley
2. Sampling theory by Singh & Chaudhary
3. Sampling Theory, Narosa Publication by Des Raj and Chandok (1998)
4. Sampling Theory and Methods, Narosa Publishers by S. Sampath (2001)
5. Theory and Analysis of Sample Survey Designs, F.S. Chaudhary: New Age International Publishers, Delhi.

Books for references:

1. Sampling Theory & Methods by M.N. Murthy.
2. Sampling theory of surveys with Applications: P.V.Sukhatme & B.V. Sukhatme.
3. Theory and methods of survey sampling. Mukhopadhyay(1988).

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	M	L	S	M	S
CO2	S	S	M	L	S	S	S
CO3	S	S	M	L	S	S	L
CO4	S	S	M	S	S	L	M
CO5	S	S	S	M	S	M	M
CO6	S	S	M	S	S	L	M
CO7	S	S	S	M	S	M	M

* S-Strong; M-Medium; L-Low

Course - STA 1.4 (B) :: LINEAR ALGEBRA

Course Code: STA 1.4 (B)

Course Objectives:

CO 1: To prepare the students about algebra of matrices and vector spaces.
CO 2: To explain about characteristic roots and vectors and linear transformations with examples.
CO 3: To prepare the students on the concept of the orthonogonality and quadratic forms.
CO 4: To prepare the students to understand Cayley – Hamilton theorem.
CO 5: To Make the students to understand the concept of the spectral decomposition of the matrices.

Learning Outcomes:

At the End of this Course Students will be able :

LO 1: Students understand elementary transformations in a matrix and their solutions.
LO 2: Students learn characteristic roots and vectors with numerical examples. They also know theoretical proofs of theorems.
LO 3: Discriminate between diagonalizable and non-diagonalizable matrices; orthogonally diagonalizable symmetric matrices and quadratic forms
LO 4: Students understand and apply Cayley – Hamilton theorem.
LO 5: Understand the concept of the spectral decomposition of the matrices.

UNIT-I

Algebra of matrices; Elementary transformations; Rank and Inverse of a matrix; Nullity; Partitioned matrices; Kronecker product; Generalized inverse of matrix; Moore-Penrose generalized inverse; Solutions of simultaneous equations.

UNIT-II

Finite dimensional Vector Spaces; Vector Spaces and Subspaces; Linear dependence and independence; Basis and dimension of a vector space; Completion theorem.

UNIT-III

Inner product Spaces; Orthonormal basis and Gram-Schmidt orthogonalization process; Orthogonal projection of a vector. Linear transformations and properties; Orthogonal and unitary transformations; Real quadratic forms.

UNIT-IV

Reduction and classification of quadratic forms; Hermitian forms; Sylvesters law of inertia; Canonical reduction of quadratic form. Characteristic roots and vectors; Cayley – Hamilton theorem.

UNIT-V

Minimal polynomial; Similar matrices; Spectral decomposition of a real symmetric matrix; Reduction of a pair of real symmetric matrices; Hermitian matrices.

Books for study:

1. Campbell, H.G. (1980), Linear Algebra with Applications, 2nd Edition, Prentice-Hall, Englewood Cliffs (new Jersey), 1980.
2. Hadley, G. (1987), Linear Algebra, Narosa Publishing House.
3. Rao, A.R. and Bhimasankaram, P. (1992), Linear Algebra, Tata McGraw Hill Publishing Company Ltd.

Books for references:

1. Graybill, F.A. (1983). Matrices with applications in statistics, 2nd ed. Wadsworth, Belmont (California).

2. Rao, C. R. (1985). Linear statistical inference and its applications, Wiley Eastern Ltd., New Delhi.
3. Searle, S. R. (1982). Matrix Algebra useful for Statistics, John Wiley and Sons. Inc.
4. Bellman, R. (1970), Introduction to Matrix Analysis, 2nd ed. McGraw Hill, New York.
5. Biswas, S. (1984), Topics in Algebra of Matrices, Academic Publications.
6. Halmos, P.R. (1958), Finite-dimensional Vector Spaces 2nd ed. D.Van Nostrand Company, Inc.
7. Hoffman, K. and Kunze, R, (1971). Linear Algebra, 2nd ed., Prentice Hall
8. Rao, C.R. and Mitra, S.K. (1971), Generalized Inverse of Matrices and its Applications, John Wiley and Sons, Inc.
9. Narayan, S. (1970), Theory of Matrices, S. Chand &Company, New Delhi.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	M	S	S	L
CO2	S	M	S	M	S	S	M
CO3	S	S	S	M	S	S	M
CO4	S	M	S	S	S	L	L
CO5	M	M	S	M	S	M	L

* S-Strong; M-Medium; L-Low

M.Sc. STATISTICS

SYLLABUS :: SEMESTER-II

w.e.f. 2023-2024 Admitted Batch of Students

(Ten Questions are to be set TWO on each of Units I-V with the choice of ONE question from each unit)

Course – STA 2.1 :: MULTIVARIATE ANALYSIS

Course Code: STA 2.1

Course Objectives:

CO1: To acquaint the students the multivariate normal (MVN) distribution properties and estimation of the parameters of MVN populations.

CO2: To acquaint the students with the applications of Hotelling T^2 Statistic and Mahalanobis D^2 statistic in case of one sample, two samples and paired samples drawn from MVN populations.

CO3: To explore the students with application of Multivariate analysis of variances (MANOVA) for one way classification multivariate data.

CO4: To explore the students with application of principal component analysis (PCA), discriminant analysis, cluster analysis, factor analysis.

Learning Outcomes:

After successful completion of the course the student will be able

LO1: To apply Hotelling T^2 Statistic and Mahalanobis D^2 statistic for testing the equality of two MVN population mean vectors in two samples and paired samples drawn from MVN populations.

LO2: To carry out one-way (MANOVA) for one way classification multivariate data.

LO3: To explore the students with application of principal component analysis (PCA), discriminant analysis, cluster analysis, factor analysis and interpret results from multi-dimensional scaling.

LO4: Understand the various hierarchical and non-hierarchical clustering methods and their applications.

LO5: Use of popular statistical packages in analyzing the real data sets.

UNIT-I

The multivariate normal distribution and estimation: The multivariate normal distribution and its properties. Characteristic function of multivariate normal distribution. Sampling from multivariate normal distribution and maximum likelihood estimation, sampling distributions of Sample mean and sample covariance matrix.

UNIT-II

Inference: Wishart's distribution and it's properties. Definition of Hotelling's T^2- distribution (statistic). Invariance property of Hotelling's T^2-statistic. Application of T^2 statistic in tests of mean vector(s) in case of one and two multivariate normal populations. The likelihood ratio principle. Mahalanobis D^2-statistic and it's relation with T^2-statistic. Multivariate analysis of variances (MANOVA) for one way classification.

UNIT-III

Discriminant Analysis: Classification and discrimination procedures for discrimination between two multivariate normal populations, Fisher's discriminant function–separation of two multivariate populations. Classification with several multivariate normal populations. Fisher's method for discrimination among several multivariate populations.

UNIT-IV

Cluster Analysis: Similarity measures, Euclidian distance and Mahala Nobis squared distance- D^2 between two p-dimensional observations (items). Hierarchical Clustering methods - Single Linkage, Complete Linkage, Average Linkage, Ward's method and Centroid Linkage methods. Non-Hierarchical Clustering methods-K-Means method. Multidimensional scaling.

UNIT-V

Special topics: Principle components analysis - definition, derivation, properties and Computation. Canonical variates and canonical correlations - definition, derivation and computation. Factor Analysis - Orthogonal factor model, Methods of estimating factor loadings - the principal component method and maximum likelihood methods of estimation. Factor rotation: orthogonal factor rotation, varimax rotation.

Books for Study:

1. Anderson, T.W.(2000). *An Introduction to Multivariate Statistical Analysis*, 3rd Edition, Wiley Eastern
2. Johnson, A. and Wichern, D.W.(2001). *Applied Multivariate Statistical Analysis*, Prentice Hall and International
3. Mardia, K.V. *Multivariate Analysis*

Books for References:

1. Gin. N. C. (1977): *Multivariate Statistical Inference*. Academic Press
2. Seber, G. A. F. (1984): *Multivariate Observations*. Wiley
3. Kshirsagar, A. M. (1972): *Multivariate Analysis*, Marcel Dekker
4. Morrison. D. F. (1976): *Multivariate Statistical Methods*, 2nd Ed. McGraw Hill
5. Muirhead, R. J. (1982): *Aspects of Multivariate Statistical Theory*, J. Wiley
6. Rao, C. R. (1973): *Linear Statistical Inference and its Applications*, 2nd ed. Wiley
7. Sharma. S. (1996): *Applied Multivariate Techniques*, Wiley
8. Srivastava, M. S. and Khatri, C. G. (1979): *An Introduction to Multivariate Statistics*, North Holland.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	M	L	S	S
CO2	S	S	M	S	L	M	M
CO3	S	S	M	S	L	M	L
CO4	S	S	S	S	L	S	L

* S-Strong; M-Medium; L-Low

Course – STA 2.2 :: TESTING OF HYPOTHESIS

Course Code: STA 2.2

Course Objectives:

CO1: Draw inference about unknown population parameters based on random samples

CO2: Impart knowledge on statistical hypothesis.

CO3: Understand Neyman-Pearson fundamental lemma for testing statistical hypothesis.

CO4: Understand the test procedures MPT, UMPT, LMPT, LRT and SPRT.

CO5: Inculcate various parametric and non-parametric, sequential test procedures.

Learning Outcomes:

LO1: Upon completion of this unit the student will be able to understand the concepts and procedures of testing of hypotheses.

LO2: Generalization of Neymann Pearson-Lemma and different Uniformly Most Powerful Test will also be get acquainted.

LO3: After completing this unit, the student will understand the Neymann structure and Likelihood ratio test with properties.

LO4: This unit provides and understanding for the student to distinguish between parametric and non-parametric tests in this unit several non-parametric tests will be understand able to the student.

LO5: This unit provides and understanding for the student to distinguish between parametric and non-parametric tests in this unit several non-parametric tests will be understand able to the student.

LO6: At the end of this unit the student will understand the notion of SPRT, and its applications to different distributions.

UNIT-I

Tests of hypotheses, concept of critical region, critical function, two kinds of errors, power function, level of significance, MP and UMP tests, Neyman Pearson lemma, Randomized and Non Randomized tests.

UNIT-II

Generalized NP-lemma, UMP tests for simple null hypothesis against one sided alternatives, and for one sided null against one sided alternative in one parameter exponential family, extension of these results to distributions with MLR property, nonexistence of UMP test for simple null against two sided alternatives in one parameter exponential family.

UNIT-III

UMP unbiased tests and LMP tests. Similar regions, Neyman structure, Likelihood ratio test, properties of LR test, asymptotic distribution of LR test.

UNIT-IV

Chi-square and kolmogorov Smirnov tests for goodness of fit, Kendall's tau statistic, Kruskal-Wallis test, Friedman's two-way analysis of variance by ranks, Bartlett's test for homogeneity of variances, chi-square test for homogeneity of correlation coefficients, F-test for homogeneity of regression coefficients, variance stabilizing transformation and large sample tests.

UNIT-V

Notion of sequential tests, SPRT, Wald's fundamental identity, relation between the quantities A,B, alpha and beta, OC and ASN functions of SPRT, application to binomial, Poisson and normal distributions, efficiency of a sequential test.

Books for Study:

1. Statistical Inference by H.C,. Saxena & Surendran
2. An outline of Statistical Theory vol.2 by A.M. Goon and B. Das Gupta.
3. An Introduction to probability and Mathematical Statistics by V.K. Rohatgi.
4. Mathematical Statistics- Parimal Mukopadhyay(1996), New Central Book Agency (P)Ltd., Calcutta.

Books for references:

1. Advanced Theory of Statistics VOL.II by M.G. Kendall & A. Stuart.
2. Introduction to Mathematical Statistics by R.V. Hogg & A.T. Craig.
3. Linear Statistical Inference and applications by C.R. Rao.

pping of Program Outcomes with Course Outcomes							
COs	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	S	S	S	M	M	S	S
CO2	S	S	S	M	S	M	S
CO3	S	S	M	S	S	M	L
CO4	S	S	M	S	S	S	L
CO5	S	S	M	M	S	S	L

* S-Strong; M-Medium; L-Low

Course – STA 2.3 (23):: THEORY OF LINEAR ESTIMATION AND ANALYSIS OF VARIANCE

Course Code: STA 2.3

Course Objectives:

CO1: This course provides the students the ability to understand the design and conduct experiments, as well as to analyze and interpret data.

CO2: To learn the basic principles in the Design of simple experiments.

CO3: To learn different tests experiments of Analysis of variance and covariance

CO4: To learn about missing plot techniques of CRD, RBD, LSD.

Learning Outcomes:

After studying the course the students should be able to carry-out:

LO1: Important characteristics of matrices, such as determinant, rank, eigen values, eigen vectors and Quadratic forms.

LO2: Students will also be able to use characteristics of a matrix to solve a Theory of linear system of equations, BLUE and Gauss-Markov theorem etc.,

LO3: Apply different test experiments of analysis of variance for one – way, two –way classification and mixed effect models.

LO4: Understand make use of analysis of covariance-one way, two-way classification and Learn about missing plot techniques of RBD, LSD.

UNIT-I

Matrix algebra- Fundamental definitions, determinants, rank of a matrix, inverse of a matrix, orthogonal matrix, idempotent matrix, characteristic roots and vectors of a matrix. Numerical computation of characteristic roots and vectors for a positive definite matrix. Reduction of a positive definite matrix to a diagonal form using an Orthogonal matrix and non-singular matrix. Cauley-Hamilton theorem, trace of a matrix. Quadratic forms, reduction of quadratic forms using orthogonal transformation, statement of Cochran's theorem for quadratic forms.

UNIT-II

Theory of linear estimation, linear models, estimability of linear parametric function, best linear unbiased estimator, Gauss-Markov set-up, Gauss-Markov theorem, generalized linear model, generalized Gauss-Markov theorem (Atken's theorem).

UNIT-III

Decomposition of sum of squares in analysis of variance one way classification, two way classification with equal and unequal number of observations per cell. Multiple comparisons; Fisher's least significance difference test and Duncan's multiple range test, Fixed, random and mixed effect models.

UNIT-IV

Analysis of covariance of one way and two way classification, applications to standard designs- CRD,RBD missing plot technique- general theory and applications to RBD and LSD.

UNIT-V

Model Adequacy checking: Test for Normality, Test for equality of Variances (Bartlett test, Modified Levene Method). Multiple comparison tests: Turkey's test, The Fisher Least significant Difference (LSD) method, Duncan's Multiple range test.

Books for Study:

1. Montgomery, D.C,(1976), Design and Analysis of experiments., John Wiley & sons.
2. Joshi, D.D.(1987), Linear Estimation and Design of experiments., Wiley Eastern Ltd.
3. Das, M.N. and Giri, N.C.(1986), Design and An Analysis of Experiments, Wiley Eastern Ltd.

Books for References:

1. Datta, K.B. (2000)., Matrix and Linear Algebra
2. Rangaswamy, R,(1995), A text book of Agricultural Statistics., New Age international Publishers Limited.
3. Kempthorne,O,(1951)., The design and Analysis of Experiments., Wiley Eastern Private Limited.
4. Rao, C.R,(1983)., Linear Statistical inference and its applications., Wiley Eastern Ltd.
5. Raghavarao, D.(1987), statistical Techniques in Agricultural and Biological Research., Oxford & IBH publishing Company Private limited.
6. Federer, W.t(1967), Experimental Design Theory and Application, Oxford & IBH publishing company.
7. Biswas, S.(1984). Topics in Algebra of Matrices, Academic Publication.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	S	L	S	L
CO2	S	S	S	S	L	M	M
CO3	S	S	M	S	L	L	S
CO4	S	M	S	S	M	S	L

* S-Strong; M-Medium; L-Low

Course – STA 2.4 (A) :: STOCHASTIC PROCESSES

Course Code: STA 2.4 (A)

Course Objectives:

CO1: Understand the meaning of stochastic process along with classification of stochastic processes, Fine state Markov chains, classification of states of a Markov chain, limiting and stationary distribution of a Markov chain.

CO2: Understand homogeneous Poisson process, non homogeneous Poisson process, Compound Poisson process and important properties of these processes, pure birth process, pure death process, general birth-death process.

CO3: Understand renewal process, elementary renewal theorem, black well's and smith's renewal theorem, central limit theorem for renewals.

CO4: Understand Galton-Watson branching process and associated generating functions and their relations, probability of ultimate extinction of a branching process.

CO5: Understand the stochastic models of Brownian Motion, Limit of Random Walk, Martingale Processes applications and it's properties.

Learning Outcomes:

At the end of this Course Students will be able:

LO1: Acquired Knowledge about Different Types of Stochastic Processes

LO2: To use birth and death Poisson processes whenever necessary

LO3: To know the concept of renewal process and its applications

LO4: To know the concept of branching process and to compute extinction probabilities

LO5: To apply Browning motion in finance problems.

UNIT-I

Stochastic Processes Preliminaries: Introduction to stochastic processes (sp's); classification of sp's according to state space and time domain, countable-state Markov chains (mc's). Chapman-kolmogrov equation, calculation of n-step transition probability and its limit. Classification of states.

UNIT-II

Markov Process: Markov processes with discrete state space, Poisson process, postulates of Poisson process, properties of Poisson process, Poisson process and related distributions – interval time – further interesting properties of Poisson process, generalizations of Poisson process – Poisson process in higher dimensions-Poisson cluster process – pure birth process _ birth – immigration process – time dependent Poisson processes – random variation of the parameter λ , birth and death process – birth and death rates.

UNIT-III

Renewal Process: Renewal processes in discrete time - relation between F(s) and P(s), renewal interval, renewal theory in discrete time, renewal theorem.

Renewal process in continuous time- renewal function and renewal density, renewal equation. Stopping time, Wald's equation. Elementary renewal theorem, Black well's and smith's renewal theorem, Central limit theorem for renewals.

UNIT-IV

Branching Process: Branching processes, properties of generating functions of branching processes- moments of X_n; probability of extinction – asymptotic distribution of X_n; Distribution of the total number of progeny.

UNIT-V

Brownian Motion: Limit of Random Walk, Martingale Processes applications and it's properties. Brownian Motion Process applications and it's properties.

Books for study:

1. Medhi, J. (1982): *Stochastic processes*, second edition, new age international(p) Ltd.
2. Karlin, s and Taylor, H.M.(1975); *A first course in stochastic process*- vol.1. academic press
3. Bhat, U.N. (1984); *Elements of applied stochastic processes*, John Wiley and sons
4. Ross .(1995); *Stochastic processes*, John Wiley and sons
5. Basu, A.K.(2003), *Introduction to Stochastic Process*, Narosa Publishing House, New Delhi.

Books for references:

1. Adke, S.R. and Manjunath, S.M. (1984): *An Introduction to Finite Markov Processes*, Wiley Eastern
2. Cinlar, E. (1975): *Introduction to Stochastic Processes*, Prentice Hall
3. Feller, W. (1968): *Introduction to Probability and its Applications*, Vol. 1, Wiley Eastern
4. Hoel, P.G. Port, S.C. & Stone, C.J. (1972). *Introduction to Stochastic Processes*, Houghton Mifflin
5. Serfozo, R. (2009). *Basics of Applied Stochastic Processes*, Springer.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	M	L	S	M	S
CO2	S	S	M	L	S	M	S
CO3	S	S	M	L	S	M	S
CO4	S	S	M	L	S	S	S
CO5	S	S	M	L	S	S	S

* S-Strong; M-Medium; L-Low

Course – STA 2.4 (B): LINEAR MODELS AND APPLIED REGRESSION ANALYSIS

Course Code: STA 2.4 (B)

Course Outcomes:

CO1: Able to calculate least square estimates of Gauss-Markov model.
CO2: Able to obtain estimates in case of correlated observations.
CO3: Able to test hypothesis for linear parametric functions.
CO4: To be able to deal with multiple regression analysis.
CO5: Able to estimate regression coefficient under multicollinearly.

Learning Outcomes:

LO1: Able to get theoretical foundation for linear estimation theory and regression analysis.
LO2: Able to demonstrate the least square estimates of the parameters and their statistical significance.
LO3: Acquainted with different regression techniques that can be used in statistical analysis.
LO4: To perform Analysis of Variance and to construct confidence intervals and regions.
LO5: Expertise with different methods for estimating and testing the relationships between independent and dependent variables.

UNIT-I

Gauss-Markov set-up, Normal equations and Least squares estimates, variances and covariances of least squares estimates, estimation of error variance.

UNIT-II

Estimation with correlated observations, least squares estimates with restriction on parameters, simultaneous estimates of linear parametric functions.

UNIT-III

Tests of hypotheses for one and more than one linear parametric functions, confidence intervals and regions. Analysis of Variance.

UNIT-IV

Simple linear regression, multiple regression, fit of polynomials and use of orthogonal polynomials.

UNIT-V

Multicollinearity, Ridge regression and principal component regression, subset selection of explanatory variables.

Books for Study:

1. Graybill, F.A.(1983): Matrices with Applications in Statistics. Wadsworth.
2. Draper, N.R. and Smith, H (1998): Applied Regression Analysis. 3rdEdition. Wiley-Blackwell.
3. Douglas C. Montgomery, Elizabeth A. Peck and G. Geoffrey Vining(2012): Introduction to Linear Regression Analysis – 5th Edition. Wiley
4. Goon, Gupta and Das Gupta(2003): An outline of Statistical Theory. Volume II. The World Press Pvt. Ltd.

Books for References:

1. Bapat.R.B.(2012): Linear Algebra and Linear Models. 3rd Edition. Springer.
2. Cook, R.D. and Weisberg, S. (1983): Residual and Influence in Regression. 1st Edition. Chapman and Hall.
3. Johnson, J. (1996): Econometric Methods, 4th Edition. McGraw Hill.
4. Rao, C.R. (2002): Linear Statistical Inference and Its Applications. 2nd Edition. Wiley-Blackwell.
5. Weisberg, S. (2013): Applied Linear Regression. 4th Edition. Wiley.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	M	M	S	S
CO2	S	S	S	S	S	M	M
CO3	S	S	L	L	M	M	L
CO4	S	M	M	S	S	S	L
CO5	S	S	S	M	L	S	M

* S-Strong; M-Medium; L-Low

M.Sc. Statistics

SYLLABUS :: SEMESTER-III

W.E.F. 2023-2024 ADMITTED BATCH OF STUDENTS

(Ten Questions are to be set TWO on each of Units I-V with the choice of ONE question from each unit)

Course – STA 3.1 :: DESIGN OF EXPERIMENTS

Course Code: STA 3.1

Course Outcomes:

CO1: To understand ANOVA, ANCOVA, fixed and random effect models

CO2: To understand the concepts of CRD, RBD, LSD and their missing plot techniques

CO3: To construct the multiple comparison tests and split plot design

CO4: To summarize the analysis of 2^n and 3^2 factorial designs and able to test their Significance

CO5: To Familiarize with total and partial confounding

CO6: To construct BIBD and PBIBD and to perform their analysis

Learning Outcomes:

LO1: Acquire theoretical foundations for design and analysis of experiments.

LO2: Able to apply ANCOVA technique.

LO3: Expertized in analysis of experiments and perform the data analysis using CRD, RBD and LSD even in case of missing values and capable of testing the model adequacy.

LO4: Expertize in analyzing factorial designs and estimate factorial effects and test their significance. Experiment confounding techniques to real life problems.

LO5: Able to apply the Youden square design and intra block analysis for estimating the Parameters of BIBD and PBIBD.

LO6: Expertized in applying different analysis of variance techniques in agricultural business and industries.

UNIT-I

General factorial experiments, factorial effects, best estimates and testing the significance of factorial effects, estimation of main effects, interaction and analysis of 2^n factorial experiments in general with particular reference to n=2,3. 3^2 and 3^3 factorial experiments. Total and partial confounding in case of 2^n (for n=2,3), 3^2 and 3^3 factorial designs.

UNIT-II

Incomplete block designs; balanced Incomplete block designs (BIBD), parametric relations, intra block analysis, simple methods of constructions of BIBD, resolvable and affine resolvable designs, Partially Balanced Incomplete Block Designs (PBIBD) with two associate classes, parametric relations, intra block analysis.

UNIT-III

Youden square design, simple lattice design, split plot design, strip plot design and their analysis, Gracco latin square design.

UNIT-IV

Concept of response surface methodology (RSM), response surface designs, linear response surface designs, second order response surface designs, variance of estimated second order response surface, Rotatable designs; conditions for second order rotatable designs, construction of second order rotatable designs using central composite designs, Balanced incomplete block designs.

UNIT-V

Taguchi Method, Taguchi Philosophy, Loss Functions, Signal-to-Noise Ratio and Performance Measures, Critique of S/N Ratios. Experimental Design in the Taguchi Method. Parameter Design in the Taguchi Method.

Books for Study:

1. M.N. Das and N.C. Giri. (1986), Designs and Analysis of Experiments., Wiley Eastern Ltd.
2. Montgomery , D.C, (1976)., Design and Analysis of Experiments., John Wiley & sons.
3. D.D.Joshi. (1987), Linear Estimation and Design of Experiments., Wiley Eastern Ltd.
4. Taguchi Engineering by Philippe ross

Books for References:

1. Raghavarao, D. (1971), Constructions and Combinatorial Problems in Design of Experiments., John Wiley & Sons, Inc.
2. W.G. Cochran and G.M. Cox, (1957), Experimental designs., A wiley International Edition.
3. Box, G.E.P. and Draper, N.R. (1986), Empirical Model- Building and Response surfaces, John Wiley & Sons.
4. R.H. Myers. (1976), Response Surface Methodology, Allyn and Bacon, Boston
5. Aloke Dey. (1986), Theory of Block Designs, Wiley Eastern Limited.
6. Oscar Kempthorne (1951)., The Design and Analysis of Experiments., Wiley Eastern Private Limited.
7. Walter T.Federer.(1967)., Experimental Design Theory and Application., Oxford & IBH Publishing Company.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	M	S	S	M
CO2	S	M	S	S	M	S	L
CO3	S	S	S	S	S	M	L
CO4	S	S	S	S	M	M	L
CO5	S	S	L	S	L	S	S
CO6	S	L	M	S	S	M	S

* S-Strong; M-Medium; L-Low

Course – STA 3.2 :: STATISTICAL QUALITY CONTROL

Course Code: STA 3.2

Course Outcomes:

CO1: Able to understand basics of production process monitoring and apply concept of control charts to it.

CO2: Able to implement response surface methodology in quality control.

CO3: Acquainted with Six Sigma and lean thinking in industrial experimentation.

Learning Outcomes:

LO1: Can identify the cause of defects using statistical quality management techniques.

LO2: Able to apply statistical quality control techniques to minimize the variability in manufacturing and business process.

LO3: Expertise in the most import field of applied statistics that contributes to quality control in all most all industries.

UNIT - I

The $\overline{X}$ and S control charts with variable sample size. The S^2 control chart. The Shewart Control Chart for individual measurements. Control charts for multiple stream process and group control charts. Economic design of control charts: designing a control chart, process characteristics, cost parameters, early work & semi-economic designs and an economic model of the $\overline{X}$ control chart.

UNIT - II

The Cumulative Sum (CUSUM) control chart - Basic principles, the tabular or algorithmic CUSUM for monitoring the process mean, recommendations for CUSUM design. The standardized CUSUM, rational subgroups, one sided CUSUMs, CUSUM for monitoring process variability, CUSUMs for other sample statistics. The V-Mask procedure. The exponentially weighted moving average (EWMA) control chart - Design of a EWMA control chart, rational sub group, robustness of the EWMA to non-normality, extension of the EWMA. The moving average control chart.

UNIT - III

Statistical process control for short production runs - $\overline{X}$ and R charts for short production runs, attribute control charts for short production runs. Modified and acceptance control charts - modified control limits for $\overline{X}$ chart, acceptance control charts, control chart for a "Six-Sigma" process.

UNIT - IV

The multivariate process monitoring and control: description of multivariate data, the multivariate normal distribution, the sample mean vector and covariance matrix. The Hotelling T^2 control chart. The multivariate exponential weighted moving average (EWMA) control chart. Control chart for monitoring variability. Latest structure methods: principal component and partial least squares.

UNIT – V

Tools and Techniques of Total Quality Management (TQM), techniques for analyzing a quality process, SPC as a tool of quality management, Quality systems – ISO 9000 standards, QS-9000 standards, Bench marking practices and Quality Auditing Notion of Six – sigma and its uses.

Books for Study:

1. R.C. Guptha(2001): Statistical Quality Control. 9th Edition. Khanna Publishers.
2. Duncan Acheson (1986): Quality Control and Industrial Statistics. 5th Edition. Irvin.
3. Statistical Quality Control – 7th edition, E.L. Grant & R.S. Leavenworth; McGraw Hill, New York.

Books for references:

1. Cowden D J (1957): Statistical Methods in Quality Control. 1st Edition. Prentice-Hall Inc.
2. Mittag and Rinne (1993): Statistical Methods for Quality Assurance. 2nd Edition. Chapman and Hall Ltd.
3. Montgomerv. D.C (2012): Introduction to Statistical Quality Control. 7th Edition. John Wiley and Sons.

pping of Program Outcomes with Course Outcomes							
COs	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	S	S	S	S	M	M	S
CO2	S	S	L	S	S	M	S
CO3	S	M	S	S	L	L	S

* S-Strong; M-Medium; L-Low

Course – STA 3.3 (A) :: ACTUARIAL STATISTICS

Course Code: STA 3.3 (A)

Course Objectives:

CO1: To introduce and expose students to application of statistics in actuarial field.
CO2: To understand the basic concepts and statistical methods applicable in actuarial science
CO3: To provide an exposure to the basic models of insurance processes
CO4: To learn the concept of interest, different life insurance products, life annuities, net premiums etc.,

Learning Outcomes:

At the End of this Course Students will be able :

LO1: To equipped with the advanced concepts which is required for internship in research institutes working in financial domain
LO2: Get expertise in the with the role of insurance in society, basic economic theory, and the basics of how insurance is required for applications in real time financial data like stock market
LO3: When uncertainty involved, student can develop and analyze actuarial models with the help of probability theory, statistics and economic theories.
LO4: Have familiarity with several of the technical tools, computer languages or software packages used by actuaries.
LO5: Develop communication, leadership and teamwork skills, and understand their importance in the actuarial industry.
LO6: It gives the ideas of entrepreneurship and also allows the students to work as a statistical advisor in insurance company, policy makers, Banking Domain

UNIT-I

Effective Rate of interest - Nominal rate of interest - Force of interest Effective rate of interest corresponding to Nominal rate of interest and Force of interest - Accumulation of 1 unit corresponding to the given rate of interest present value of 1 unit due at the end of n years corresponding to the given rate of interest, effective rate of discount-Nominal rate of discount – Force of interest varying continuously.

UNIT-II

Definition of Annuity - Present value and Accumulation of an Immediate annuity - Annuity due - Present value and Accumulation of an Annuity – due - Present value and Accumulation of a Perpetuity and Perpetuity - due-Increasing and Decreasing Annuities. Annuities where payments are in A.P. Annuities where payments are made continuously.

UNIT-III

Analysis of Annuity payments - Loan installment - Principal and interest portions of a typical loan installment – General Expression for principal and interest portion – Purchase price of an annuity net of tax. Independent annual interest rates - Mean and Variance of the accumulation of a single investment, viz $E(S_n)$ and $V(S_n)$ – Mean and Variance of the accumulation of a series of investments, viz $E(A_n)$and $V(A_n)$ - Mean and Variance of a present value of unit due at the end of n years, viz $E(V_n)$ and $V(V_n)$. Dependent annual interest rates and simple problems.

UNIT-IV

Net present value (NPV) - Internal rate of return (IRR)- Interpretation of NPV and IRR- Comparison of two investment projects - Discounted Payback Period - The effect of inflation on IRR- Money weighted rate of return (MWRR) - Time weighted rate of return (WRR) - linked internal rate of return (LIRR).

UNIT-V

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 decrements, central force assumptions for multiple decrements. Uniform distribution assumption for multiple decrements.

Books for Study:

1. Fundamentals of Actuarial Mathematics, Promislow, S.D(2006): John Willey, Chapters 2-11 &14.
2. Newton L. Bowers, Jr, Hans U. Gerber, James C. Hickmann, Donald A. Jones and Cecil J. Nesbitt (1997): Actuarial Mathematics, The Society of Actuaries.
3. Compound Interest and Annuities certain by D.W.A. Donald, Heinemann, London.
4. An Introduction to Mathematics of Finance, Butter Worth & Heinemann by J. J. McCutcheon and W.F.Scott.
5. Gnana Deep Study Maeraials.

Books for References:

1. Actuarial mathematics by N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones and C.J. Nesbitt – Published by society of Actuaries, Ithaca, Illinois, U.S.A. Second Edition (1997) Chapters: 1,2,3,4,5,9 &10
2. Life Contingencies, by Spurgeon – Cambridge University Press.
3. Life Contingencies, Heinemann. Neill, A. (1977).
4. An Introduction to Actuarial Studies, Atkinson and Dickson (2011): Edward Publishing.
5. Life Contigencies by A. Neile – Published by Heineman.
6. Modern Actuarial Theory and Practice, Philip, M. et. al (2004): Chapman and Hall.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	S	S	S	S
CO2	S	S	S	S	M	L	L
CO3	S	S	S	S	L	L	M
CO4	S	S	M	S	M	M	L

* S-Strong; M-Medium; L-Low

Course – STA 3.3 (B) (23) :: TIME SERIES ANALYSIS

Course Code: STA 3.3 (B)

Course Objectives:

CO 1: To demonstrate advanced understanding of the concepts of time series analysis and their applications.

CO 2: To demonstrate decomposition of time series tests for trend and seasonality.

CO 3: To understand the various methods of smoothing and analysis of time series data.

CO 4: To understand the analysis of stationary models in time series.

CO 5: To develop stationary and non stationary models and their diagnostic check.

CO 6: To acquire scientific knowledge for forecasting and applying it to various sectors.

Learning Outcomes:

At the End of this Course Students will be able to:

LO 1: Understand and analyze components of time series.

LO 2: Explain time series with different structures.

LO 3: Perform trend analysis, develop seasonal indices, analysis cyclical variations and random components.

LO 4: Smooth the data and apply various smoothing methods for data analysis.

LO 5: Construct time series models and analyze.

LO 6: Forecast with different scientific methods and become an expert in time series analysis and applications.

UNIT-I

Time-series as discrete parameter stochastic process. Auto covariance and autocorrelation functions and their properties. Exploratory Time Series Analysis, Tests for trend.

UNIT-II

Exponential and Moving Average Smoothing. Holt and Winters' smoothing. Forecasting based on smoothing, Adaptive smoothing.

UNIT-III

Detailed study of the stationary process: Moving Average (MA), Auto Regressive (AR), ARMA and AR Integrated MA (ARIMA) models.

UNIT-IV

Box – Jenkins models. Discussion (without proof) of estimation of mean, auto covariance and autocorrelation functions under large sample theory. Choice of AR and MA periods. Estimation of ARIMA model parameters. Forecasting. Residual analysis and diagnostic checking.

UNIT-V

Spectral analysis of weakly stationary process. Periodgram and correlogram analysis. Computations based on Fourier transform.

Books for Study:

1. Box, G.E.P. and Jenkins, G.M. (1976): Time Series Analysis –Forecasting and Control. Holden Day, San Francisco.
2. Anderson, T.W. (1971): The Statistical Analysis of Time Series, Wiley, New York.
3. Makridakis, Wheelwright and McGee: Forecasting-Methods and Applications, John Willey and Sons.
4. Montgomery, D.C. and Johnson, L.A.(1977): Forecasting and Time Series Analysis, McGraw Hill.

Books for References:

1. Fuller, W.A. (1976): Introduction to Statistical Time Series, John Wiley.
2. N.V. Granger, C.W.J. and Newblod (1984): Forecasting Econometric Time Series, 3rd Edition, Academic Press.
3. Priestley, M.B. (1981): Spectral Analysis and Time Series, Griffin, London.
4. Kendall, Sir Maurice and Ord, J.K. (1990): Time Series Analysis. 3rd edition Edward
5. Kendall, M.G. and Stuart A. (1966): The Advanced Theory of Statistics, Volume 3, Charles Griffin. London.
6. Bloomfield, P. (1976): Fourier Analysis of Time Series – An Introduction, Wiley.
7. Granger, C.W.J. and Hatanka, M. (1964): Spectral Analysis of Economic Time Series, Princeton Univ. Press. New Jersey.
8. Koopmans, L.H.(1974): The Spectral Analysis of Time Series, Academic Press.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	M	S	S	M	M	M
CO2	S	M	S	L	M	S	M
CO3	S	S	S	S	M	M	M
CO4	S	S	S	S	L	M	L
CO5	S	M	M	L	S	L	S
CO6	S	L	S	M	M	S	L

*S-Strong; M-Medium; L-Low

Course – STA 3.4 (A) :: RELIABILITY – I

Course Code: STA 3.4 (A)

Course Objectives:

CO1: Understand the basic concepts of reliability, hazard function, bathtub curve and its applications.

CO2: To provide an exposure to the various lifetime failure models

CO3: Comprehend the importance of reliability theory in industries

CO4: Construct exponential components for system reliability involved in reliability analysis

CO5: Understand the importance of redundancy techniques and stand by models

CO6: Exhibit theoretical knowledge on some reliability in standard probability models using complete samples for estimating the parameter and testing validity.

Learning Outcomes:

At the End of this Course Students will be able :

LO1: To learn about the basic concepts of reliability, failure rate and hazard function etc., and analyze early failure and methods

LO2: To estimate different types of failure distributions

LO3: To know the concepts on system reliability with exponential components in Series, Parallel and r-out of-n systems

LO4: To evaluate the scope and with different models of each system component in system reliability

LO5: To predict the reliability of a component, system and of a finished product.

UNIT-I

Importance of reliability, definition of reliability and its measures, concept of failure. General provision of a reliability specification, Methods of achieving reliability, Broad functions of reliability. Bath tub curve, causes of early failure and methods to avoid them.

UNIT-II

Life distributions; reliability function, hazard rate, Common failure distributions: exponential, weibull, truncated normal, log normal - their properties and uses and Estimation of parameters and tests in these models.

UNIT-III

Series, parallel and r-out of n configurations; their block diagram, reliability graph and determination of reliability through combinatorial methods. Events space, cut set and tie set, Multistate models.

UNIT-IV

System reliability with exponential components in series, parallel and r- out of - n system. Usefulness of redundancy and improvement factor. MTTF, MTBF, Equivalents MTBF of series and parallel system. Cold and hot redundancy, reliability of stand-by system. Weakest link model, chain model, stress-strength model, non-parametric estimation of reliability.

UNIT-V

Problem of life testing, estimation of parameters and reliability in standard probability models (Exponential, Weibull, Normal) using complete samples. Probability plotting and graphical procedures for estimating the parameter and testing validity of model by some standard statistical tests.

Books for Study:

1. Probability Distributions Used in Reliability Engineering, Andrew N.O'Connor Mohammad Modarres, Ali Mosleh; Published by the Center for Risk and Reliability
2. Statistical Analysis of Reliability and Life-Testing Models, Bain, L.J, Dekker, New York
3. Statistical Models and Methods for Lifetime Data, Lawless, J.F., Wiley, New York
4. Bayesian Reliability Analysis, Martz, H.E. & Weller, A., Willey New York
5. Statistical Theory of Reliability and Life Testing Probability Models, Barlow R.E.& Proschan, F., Holt, Rinehart and Winston, New York
6. Reliability and Life Testing, Sinha, S.K., Wiley Eastern Limited

Books for References:

1. Applied Life Data Analysis, Nelson, W. (1982):, John Willey
2. Software Engineering: Design, Reliability and Management, Shooman, M.L., McGraw- Hill, New York
3. Reliability in Engineering Design, Kapur, K.C. and Lamberson, L.R., John Wiley, N.York
4. J.V. Deshpande and Sudha G. Purohit (2005) Life time data: Statistical Models and Methods World Scientific
5. Introduction to Reliability Analysis, Zacks S (1992), Springer - Verlag, New York
6. Applied Reliability 3rd Edition by Paul A. Tobias, CRC Press Taylor & Francis Group.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	S	L	M	M
CO2	S	S	S	S	M	L	S
CO3	S	S	S	S	M	S	S
CO4	S	S	L	S	L	S	M
CO5	S	S	L	S	S	M	L
CO6	S	S	M	S	S	L	M

*S-Strong; M-Medium; L-Low

Course – STA 3.4 (B) :: BIOSTATISTICS

Course Code: STA 3.4 (B)

Course Objectives:

CO1: Able to understand different statistical methods in clinical trials and their applications.

CO2: Demonstrate different types of statistical designs and perform randomization.

CO3: Summarize the biological assays such as parallel-line assay, slope-ratio assay and quantile-responses assay.

CO4: Able to carry out the Categorical Data Analysis, Logistic Regression Analysis and Poison Regression Analysis and their applications.

CO5: Able to measure the ANOVA for one way and two way classified data.

Learning Outcomes:

At the End of this Course Students will be able to :

LO1: Familiarize with clinical trials and its phases I, II, III and IV.

LO2: Students learn how to conduct analysis such as categorical analysis, ROC curve analysis and biological assays.

LO3: Students gain knowledge about various types of regression techniques used to apply medical data especially for count data.

LO4: Students can perform repeated measures for ANOVA one way and two way classified data.

LO5: Students can infer about disease frequency, incidence, prevalence and relative risk.

UNIT-I

Statistical Methods in Clinical Trials: Introduction to clinical trial and it's phases I, II, III and IV, statistical designs-fixed sample trials: simple randomized design, stratified randomized crossover design; Sequential design - open and close sequential design. Randomization Dynamic randomization, Permuted block randomization; Blinding-Single, double and triple.

UNIT-II

Biological Assays: Introduction, parallel-line assay, slope- ratio assays and quantile- response assay, Feller's theorem. Dose-response relationships-qualitative and quantitative response, dose response relation- estimation of median effective dose – PK-PD Analysis.

UNIT-III

Categorical Data Analysis: Categorical response data, logistic regression-odds ratio, Wald's statistic, logistic regression and its diagnostics, - Poisson regression – Estimation of relative risk and its applications.

UNIT-IV

ROC Curve analysis - Estimation of Binomial Model and the Area under the Curve, its applications – Properties of ROC curve - Kullback –Leibler Divergence (KLD)– definition – functional relationship between Kullback –Leibler Divergence and the slope of the ROC curve – derivations of KLD expressions for Bi-normal ROC model.

UNIT-V

Repeated Measures ANOVA – One Way and Two Classified Data –Measures of disease frequency – incidence – prevalence – relative risk – Epidemiological study designs – Cohort study design and its analysis – Case control study design and its analysis – concept of bias – information bias and selection bias.

Books for Study:

1. Elisa T. Lee and John Wenyu Wang (2003): Statistical Methods for Survival Data analysis, 3rd edition, John Wiley.
2. Jerrold H. Zar (1999): Biostatistician Analysis, 4thedition, Pearson.
3. Armitage, P, Berry G and Mathews J.N.S (2002): Statistical Methods in Medical Research, 4th edition, Blackwell Scientific Publications.
4. Krzanowski, W and Hand, D.J.(2009):ROC Curves for Continuous Data, Chapman and Hall.

Books for References:

1. Hosmer and Lemeshow (2000): Applied Logistic Regression, 2nd, Wiley Series.
2. Alan Agresti (2002): Categorical Data analysis, 2nd , John Wiley.
3. Sylvia Wasserthial and Smoller, (2004): Biostatistics and Epidemiology – A Primer for Health and Biomedical professionals, 3rd edition, Springer.
4. Rastogi, V.B. (2006): Fundamentals of Biostatistics, ANE Books, India.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	L	S	S	M	S
CO2	S	S	L	S	S	S	S
CO3	S	S	S	M	M	L	S
CO4	S	S	M	M	M	M	S
CO5	S	S	M	S	M	L	S

*S-Strong; M-Medium; L-Low

M.Sc. STATISTICS

SYLLABUS :: SEMESTER-IV

w.e.f. 2023-2024 Admitted Batch of Students

(Ten Questions are to be set TWO on each of Units I-V with the choice of ONE question from each unit)

Course – STA 4.1 :: ECONOMETRICS

Course Code: STA 4.1

Course Objectives:

CO1: To explore the students to the special subject Econometrics which consists of the construction of general linear regression models and non-linear regression models based on the given data for a given economic phenomenon using the well-known statistical tool multiple regression analysis.

CO2: To acquaint the students about multiple regression analysis including the case of simple regression analysis which involves the estimation and validation of general linear model (GLM), then making predictions with successfully validated model.

CO3: To explore the students about the diagnostics and remedies to the problems of multicollinearity, heteroscadesticity and autocorrelation which are resulting when one or some of the assumptions of the model are violated.

CO4: To acquaint the students with some popular Qualitative regression models namely logistic regression models and probit models.

CO5: To explore the students to Autoregressive and Distributed Lag models- Stock adjustment and partial adjustment models.

Course Outcomes:

After learning the course the student will be able

LO1: To carryout simple linear regression analysis of dependent variable with one independent variable which consists of the estimation of the model, validation of the model and prediction with validated model.

LO2: To carryout regression analysis with regard to some reputed non-linear regression models like semi- log, double log and reciprocal models.

LO3: To carryout multiple linear regression analysis with several independent variables if necessary including dummy variables.

LO4: To build a general linear model (GLM) in the presence of multicollenearity, heteroscadesticity and autocorrelation.

LO5: To construct and validate polynomial, logit and probit regression models based on a given appropriate data.

LO6: To estimate the Autoregressive and Distributed Lag models namely Stock adjustment and partial adjustment models.

UNIT-I

Simple Regression Analysis: What is Econometrics? Methodology of Econometrics. Simple linear model and assumptions, Least-Squares Estimators and their properties. ML estimation of the parameters. Statistical Inference and prediction with the simple regression model. Regression analysis versus Correlation analysis, Regression analysis and ANOVA. Other functional forms of regression models –Log-linear, Semilog, reciprocal and logarithmic reciprocal models.

UNIT-II

Multiple Regression Analysis: The general linear model (GLM) and assumptions of the model, ordinary least squares (OLS) and ML estimation, properties of OLS estimators (Gauss-

Markov theorem). The coefficient of determination R^2, and adjusted R^2 or $\bar{R}^2$. Inferences about regression model, problems of prediction. Linear restrictions-- restricted least squares.

UNIT-III

Testing the structural change in regression models, Chow test for testing the equality of two regression equations. The use of dummy variables in multiple regression. The problem of multicollenearity - nature, sources, consequences, diagnostics (variance inflation factors and condition index) and remedies (Ridge regression and principle component regression methods). Aitken's generalized least squares(GLS) method.

UNIT-IV

The problem of heteroscadesticity— nature, sources, consequences, detection (Glejser Test, Goldfeld-Quandt test, Breusch-Godfrey Test, Breusch-Pagan-Godfrey Test, White's test) and remedies. The problem of auto correlation— nature, sources, consequences, detection (Durbin-Watson d-test and Breusch-Godfrey (LM) Test,) and remedies (Cochrane-Orcutt iterative procedure, C-O two-step and D-W two-step methods).

UNIT-V

Special Topics: Non-Linear Regression Models, Polynomial regression models, Qualitative Response Regression Models-The LOGIT (Logistic regression) and the PROBIT models. Errors in Variables, Instrumental Variables, Autoregressive and Distributed Lag models-Koyak, Stock adjustment and partial adjustment models.

Books for study:

1. Gujarathi, D.N., Porter, D.C. and Gunasekar, S. (2011): ***Basic Econometrics***, 5th Edition, McGraw Hill
2. Johnston, J. and DiNardo, J.(1997). ***Econometric Methods***, 4th Ed., McGraw Hill
3. Montgomery, D.C., Peck, E.A. and Geoffrey Vininig, G. (2003). ***Introduction to Linear Regression Analysis,*** 3rd Ed., Wiley
4. Koutsoyiannis, A (1979): ***Theory of Econometrics***, Macmillan Press
5. Theil. H (1982): ***Introduction to the Theory and Practice of Econometrics***, John Wiley.

Books for References:

1. Apte P.G. (1990): *Textbook of Econometrics*. Tate McGraw Hill
2. Intrilligator, M.D. (1980): *Econometric models – Techniques and Applications*, Prentice Hall of India.
3. Klein, L.R.(1962): *An introduction to Econometrics*, Prentice Hall of India
4. Mai Invaud, E (1966): *Statistical Methods of Econometrics*, North Holland.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	M	S	M	S	S
CO2	S	S	M	S	M	M	S
CO3	S	S	M	S	L	L	S
CO4	S	S	S	S	S	M	M
CO5	S	S	S	S	L	S	S

* S-Strong; M-Medium; L-Low

Course – STA 4.2 :: ACCEPTANCE SAMPLING PLANS

Course Code: STA 4.2

Course Outcomes

CO1: Apply acceptance and continuous sampling plan in production process.
CO2: Able to construct sampling inspection plans for attributes and variables.
CO3: Able to learn some advanced control charts and capability indices.
CO4: Able to construct Six Sigma limits.

Learning Outcomes

LO1: Acquainted with Six Sigma and lean thinking in industrial experimentation.
LO2: Can identify the cause of defects using statistical quality management techniques.
LO3: Able to apply statistical quality control techniques to minimize the variability in manufacturing and business process.

UNIT – I

Attribute Sampling Plans: Introduction, advantages and disadvantages of sampling, Producer's risk and Consumer's risk, Operating Characteristic curve, evaluating sampling plans using average out going quality limit, average sample number. Lot-by-Lot attribute sampling plans: Single sampling plans – The OC curve, design of single sampling plans, double sampling plans, The OC curve , ASN curve, design of double sampling plan sequential sampling plan, standard sampling plans-Military Standard 105E and Dodge – Roming sampling plan.

UNIT – II

Acceptance sampling plans for variables, Introduction, advantages and disadvantages of sampling, variable sampling plans for a process parameter – Estimating process average – Single specification limit and known process standard deviation, estimating process average – double specification limits and known process standard deviation, estimating process average – Single specification limit and unknown process standard deviation. Standardized plans – MIL-STD-414, Sequential sampling plan by variables.

UNIT – III

Acceptance sampling procedures – importance, procedures, advantages and disadvantages of Chain sampling plan, Skip-lot sampling plan, Continuous sampling – CSP-1, CSP-2, CSP-3 and multi-level plans, Military standard sampling plan –MIL STD 1235b.

UNIT – IV

Industrial Experimentation, Fractional factorial experiments, Response surface methodology, Six sigma in process improvement and product development. Wald-Wolfwitz type and their properties.

UNIT – V

Rectifying inspection by Lot-By-Lot Sampling: Rectifying Inspection Plans Calling for 100 Percent inspection of Rejected Lots. Rectifying Inspection Plans with less than 100 Percent inspection of Rejected Lots.

Books for Study:

1. Introduction to Statistical Quality Control, Montgomery, D.C., John Wiley (Asia) 2001.
2. Modern Methods for Quality Improvement, H.M.Wadsworth, K.S.Stephens A.B.Godftrey, Second Edition; 2004, John Wiley and sons.
3. The Essence of TQM, John Bank, Printice,- Hall of India Pvt Ltd (1998).
4. Statistics of Quality Control – Sampling Inspection and Reliability by S.Biswas, New central book agency Pvt Ltd (2003).

Books for References:

1. R.C. Guptha (2001): Statistical Quality Control. 9th Edition. Khanna Publishers.
2. Duncan Acheson (1986): Quality Control and Industrial Statistics. 5th Edition. Irvin.

3. Statistical Quality Control – 7th edition, E.L. Grant & R.S. Leavenworth; McGraw Hill, New York.

pping of Program Outcomes with Course Outcomes							
COs	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	S	S	M	S	M	S	S
CO2	S	S	S	S	M	S	S
CO3	S	M	S	S	L	L	S
CO4	S	S	M	S	M	L	M

* S-Strong; M-Medium; L-Low

Course – STA 4.3 (A) :: OPERATIONS RESEARCH

Course Code: STA 4.3 (A)

Course Objectives:

CO1: To acquire knowledge of scope of Operations Research

CO2: Understand to develop the optimization techniques that will be useful in the personal and professional life

CO3: Can apply Game theory to analyze existing situations wherever there are limited resources, different decision options and different outcomes from different choices.

CO4: To find optimum solution and formulate LPP to solve problems using simplex methods

CO5: To apply with inventories of various goods with and without shortages to solve real life problems and derive study state solutions of Poisson queues

CO6: Using PERT and CPM able to find the minimum time required to complete project when it involves interdependent activities.

Learning Outcomes:

At the End of this Course Students will be able :

LO1: To Formulate the problem in operations research.

LO2: To Establish the relationship between the variables and constraints by constructing the model to analyze existing situations wherever there are limited resources, different decision options, different outcomes from different choices.

LO3: To provide the idea of formulating mathematical model and their optimum solution in the contest of practical problems belonging to Government /Private sectors.

LO4: To Learn the tools like Linear Programming Problems, replacement and operational gamming.

LO5: To Familiar with the queuing and different design and develop inventory models

LO6: To Obtain a firm foundation in advanced OR techniques for the real life problems.

UNIT-I

Games as decision problems, Two-person Zero-Sum game, Pure and mixed Strategies, maximum criterion, dominance, minimax theorem. Solutions of 2X2 , 2Xm, nX2 games and 3X3 games (using simplex algorithm).

UNIT-II

Inventory control, models of inventory-purchase model with instantaneous replenishment and without shortages, Manufacturing model without shortages, Purchase model with instantaneous Replenishment and with shortages. Manufacturing model with shortages, Operation of inventory system, Quantity Discount - Price breaks - Purchase Inventory Model with one price break, two price breaks and any number of price breaks.

UNIT-III

Queueing models - Characteristics of Queueing Systems, Classification of Queues. Steady-state solution of $M/M/1/\infty/FCFS$, $M/M/1/N/FCFS$, $M/E_k/1$ models and $M/G/1$ queue - length Pollazek - Khinchine result.

UNIT-IV

Replacement and maintenance analysis-Types of maintenance, Types of replacement problems, Determination of economic life of an asset, Basics of interest formulae – Present-worth factor (P/F, i, n), Equal payment series capital recovery factor (A/P, i, n). Simple probabilistic model for items which completely fail.

UNIT-V

Project management; CPM and PERT; probability of project completion, crashing.

Books for Study:

1. Hillier, F.S. and Leiberman(2017)G.J. ,Holden Dev, Introduction to Operations Research.
2. Sharma, J.K.,2003, Operations Research Theory and Applications,. Macmillan, India
3. Kantiswarup, Gupta, P.G. and Man Mohan, Operations Research, Sultan Chand & Sons
4. S.D. Sharma, Operations Research, Kedar Nath Ram Nath & Co, Meerut
5. Panneerselvam. R, Operations Research, Printice Hall of India, Pvt Ltd.
6. Churchman, C.W., Ackoff, R.L. and Amoff, E.L. (1957). Introduction to Operations Research.

Books for References:

1. Philps D.T. Ravindran A and Sal berg J., Operations Research, Principles and practice, John Wiley
2. Donald Gross and Carl M.Harrsis, Fundamentals of Queueing Theory, John Wiely
3. Leonard Kleinrock, Queueing System, volume 1. Wiley Inter science
4. Taha, HA. (1982). Operational Research
5. Philips, D.T., Ravindran, A. and Solberg, J. Operations Research, Principles and Practice

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	M	S	S	S	M
CO2	S	S	S	S	S	L	L
CO3	S	S	S	S	M	S	L
CO4	S	S	M	S	M	S	S
CO5	S	S	L	M	L	M	L
CO6	S	S	M	S	S	M	S

* S-Strong; M-Medium; L-Low

Course – STA 4.3 (B) :: MACHINE LEARNING USING PYTHON

Course Code: STA 4.3 (B)

Course Objectives:

CO 1: Comfortably Perform basics operations in Python
CO 2: Understand machine learning concepts
CO 3: Explore and execute the machine learning concepts for real time data using Python

Learning Outcomes:

At the End of this Course Students will be able :

LO 1: Perform basic operations and concepts in Python
LO 2: Understand and use the essential modules in Python
LO 3: Evaluate the scope and opportunities of machine learning
LO 4: Gain knowledge and hands-on training in machine learning techniques
LO 5: explore program skills for machine learning techniques

UNIT-I

Type of variables, data types, lists, control statements, functions, classes, files and exceptions.

UNIT-II

Jupyter Notebook, Numpy, Scipy, Matplotlib, Pandas, mglearn.

UNIT-III

Classification and Regression, k-Nearest Neighbors, Decision Trees, Neural Networks.

UNIT-IV

Preprocessing and Scaling, Scaling training, Dimensionality Reduction, Feature Extraction, and Manifold Learning.

UNIT-V

Clustering: k- Means clustering, Agglomerative Clustering, DBSCAN.

Books for study:

1. Introduction to Machine Learning with Python – A Guide for Data Scientists by Andreas C. Muller & Sarah Guido(2017), O'Reilly
2. Machine Learning in Python : Essential Techniques for Predictive Analysis by Micheal Bowles (2015), Wiley
3. Python Crash Course : A hands-on, Project-Based Introduction to Programming by Eric Matthes(2016), no starch press.

Books for references:

1. Python for Probability, Statistics and Machine Learning (second edition) (2019) by Jose Unpingco, Springer
2. Practical Statistics for Data Scientists(second edition)(2020) by Peter Bruce, Andrew Bruce & Peter Gedeck, O'Reilly.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	S	S	M	M	S
CO2	S	S	L	S	S	M	S
CO3	S	M	S	S	L	L	S

* S-Strong; M-Medium; L-Low

Course – STA 4.4 (A) :: RELIABILITY – II

Course Code: STA 4.4 (A)

Course Objectives:

CO1: Understand the basic concepts of Repairable and Non-Repairable systems and its applications

CO2: Know the concepts bounds on system reliability; structural and reliability importance of components

CO3: Learn reliability theory and analysis of survival data

CO4: Familiar with different concepts of reliability and lifetime models

CO5: Understand the special kinds of various Fault Trees construction used in reliability

CO6: To evaluate the concepts of IFR, IFRA, NBU, DMRL, and NBUE.

Learning Outcomes:

At the End of this Course Students will be able :

LO1: To learn about the basic concepts of one/two components in Repairable and Non-Repairable systems reliability

LO2: Recall to estimate the parameters of lifetime distributions

LO3: To explain and compute the structural properties of coherent system

LO4: To evaluate the structure function using modular decomposition

LO5: To expertise to use fundamentals and procedures of FTA formats and FMEAs etc.,

LO6: To familiar distinguish between the concepts of Notions of Ageing

LO7:Get theoretical knowledge of computing probability of survival of machines, models related to production in industries and can draw conclusions.

UNIT-I

(i) Non-Repairable System:

Single element-non repairable, two element-non-repairable system; solution through Laplace transform. Poisson process, Stand-by system.

(ii) Repairable System:

Reliability and availability function of one and two components system, up-time and down-time ratio, steady state probabilities.

UNIT-II

Coherent System and its Structural Properties :

Systems with independent components, coherent system, path sets and cut sets, reliability of coherent system, bounds on system reliability, Relative importance of components, Modular decomposition of coherent system and improved bounds for system reliability. Concept of associated random variables.

UNIT-III

Fault Tree Analysis:

Event tree, simple fault tree and its construction, Mathematics of FTA, Efficiency of FTA formats, FTA, Event space method, Monte-Carlo technique, Min-cut set algorithm, FMEA, Carrying out FMEA with practical example.

UNIT-IV

Life distributions – reliability function; hazard rate; common life distributions-Exponential, Gamma, Weibull, Lognormal, Pareto. Estimation of parameters and tests in these models. Life tables, mean residual life and their elementary properties.

UNIT-V

Notions of Ageing classes – IFR, IFRA, NBU, DMRL and NBUE and their duals. Estimation of survival function-Acturial Estimator, Kaplan–Meier Estimator, Semi–parametric regression for failure rate–Cox's proportional hazards model with one and several covariates.

Books for Study:

1. Reliability Engineering Theory and Practice, A. Birolini, Fourth Edition, Springer Int.
2. Bayesian Reliability Analysis, Martz, H.E. & Weller, A., Willey New York
3. Reliability and Life Testing, Sinha, S.K., Wiley Eastern Limited
4. Probability Distributions Used in Reliability Engineering, Andrew N.O'Connor Mohammad Modarres, Ali Mosleh; Published by the Center for Risk and Reliability
5. Statistical Analysis of Reliability and Life-Testing Models, Bain, L.J, Dekker, New York
6. Statistical Models and Methods for Lifetime Data, Lawless, J.F., Wiley, New York

Books for References:

1. Statistical Theory of Reliability and Life Testing Probability Models, Barlow R.E.& Proschan, F., Holt, Rinehart and Winston, New York.
2. Introduction to Reliability Analysis, S. Zacks, Springer Verlag, N.Y.
3. Gross A.J. and Clark, V.A.(1975) Survival Distributions: Reliability Applications in the Biomedical Sciences, John Wiley and sons.
4. Cox, D.R. and Oakes, D. (1984) Analysis of Survival Data, Chapman and Hall, New York
5. Elandt-Johanson, R.E.Johnson N.L. (1980) Survival models and Data Analysis, John Wiley and Sons.
6. Miller, R.G. (1981) Survival Analysis (Wiley).

pping of Program Outcomes with Course Outcomes							
COs	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	S	S	S	S	M	L	S
CO2	S	S	S	S	M	L	M
CO3	S	M	S	S	M	M	M
CO4	S	S	M	M	S	S	S
CO5	S	M	M	L	S	L	S
CO6	S	S	M	S	S	L	L

*S-Strong; M-Medium; L-Low

Course- STA 4.4 (B) :: KNOWLEDGE DISCOVERY AND DATA MINING

Course Code: STA 4.4 (B)

Course Objectives:

CO1: Able to outline classification methods.
CO2: To get expertise in different clustering methods.
CO3: Able to learn unsupervised learning techniques.
CO4: To acquaint with supervised learning techniques.
CO5: Able to get familiarity with analytical data processing.

Learning Outcomes:

At the End of this Course Students will be able:

LO1: Able to perform multivariate analysis with data sets.
LO2: Able to infer vector quantization methods.
LO3: Obtain a firm foundation in dimension reduction, feature selection and clustering techniques.
LO4: To appraise regression trees in predictive modeling.
LO5: To construct association rules on data sets.

UNIT-I

Review of classification methods from multivariate analysis; classification and decision trees.

UNIT-II

Clustering methods from both statistical and data mining viewpoints; vector quantization.

UNIT-III

Unsupervised learning from univariate and multivariate data; dimension reduction and feature selection.

UNIT-IV

Supervised learning from moderate to high dimensional input spaces; regression trees.

UNIT-V

Introduction to databases, including simple relational databases; data warehouses and introduction to online analytical data processing. Association rules and prediction; data attributes.

Books for study:

1. A.Berson and S.J. Smith (1997): Data Warehousing, Data Mining and OLAP. McGraw-Hill.
2. L.Breiman, J.H. Friedman, R.A. Olshen, and C.J.Stone (1984): Classification Regression Trees. Taylor Francis.
3. J.Han and M. Kamber (2006): Data Mining; Concepts and Techniques. 2nd Edition. Morgan Kaufmann.
4. T.M. Mitchell (2011): Machine Learning. Springer

Books for reference:

1. B.D.Ripley (2008): Pattern Recognition and Neural Networks. Cambridge University Press.

pping of Program Outcomes with Course Outcomes

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	S	S	L	S	S	S	S
CO2	S	S	S	S	S	S	S
CO3	S	S	L	S	M	S	S
CO4	S	S	M	S	M	M	S
CO5	S	S	S	S	L	S	S

*S-Strong; M-Medium; L-Low

ANNEXURE – IV

(Agenda item No.4 (iii), P.G.B.O.S. meeting dated 05-08-2023)

MODEL QUESTION PAPER **Course Code: ST 1.1**

M.Sc. Degree Examination

STATISTICS

FIRST Semester

Course - S 1.1 :: PROBABILITY THEORY AND DISTRIBUTIONS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT - I

1.(a) i) Given a class $\{A_i, i=1,2,...,n\}$ of n sets prove that there exists a class $\{B_i, i=1,2,...,n\}$ of n disjoint sets such that their unions are equal.

ii) Give the axiomatic definition of probability. State the various properties of probability.

(b) Define probability measure. Establish its simple properties.

2.(a) Write short notes on (i) σ fields (ii) conditional probability (iii) discrete probability space and give one example in each case.

(b) state and prove Borel- Cantelli lemma.

UNIT - II

3.(a) Define characteristic function. State and prove inversion formula.

(b) State and prove Holder's inequality. Hence, obtain Schwartz inequality.

4.(a) Define distribution function. State and prove its properties.

(b) State and prove i) Jensen and ii) Markov inequalities.

UNIT - III

5. a) State and prove Chebychev's form of weak law of large numbers.

b) Explain the types of convergence. Prove that almost sure convergence implies convergence in probability.

6. a) State and prove Kolmogorov's strong law of large numbers.

b) State and prove Kinchine's form of W.L.L.N.

UNIT - IV

7. a) Derive the distribution of compound binomial.

b) Obtain the m.g.f. of truncated Poisson distribution and hence, find its mean and variance.

8. a) Define Weibull distribution. Find its m.g.f., mean and variance.

b) Define Laplace distribution. Obtain its characteristic function, mean and variance.

UNIT - V

9. a) Define order statistics. Obtain the joint p.d.f. of $X_{(i)}$ and $X_{(j)}$ $1 \le i < j \le n$.

b) Derive the distribution of range.

10. a) Derive the distribution of the range in case of exponential distribution.

b) Let $X_{(1)}, X_{(2)},...,X_{(n)}$ be the order statistics of independent random variables $X_1, X_2,...,X_n$ from rectangual distribution U[a,b]. Derive the distribution function of r^{th} order statistic $X_{(r)}$, $1 \le r \le n$.

MODEL QUESTION PAPER **Course Code: ST 1.2**

M.Sc. Degree Examination

STATISTICS

FIRST Semester

Course - S 1.2 :: STATISTICAL COMPUTING USING R

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT – I

1. (a) Explain different types of data types and give an illustration for each type.
   (b) Explain various I/O console functions by means of illustrations.
2. (a) Explain how to create and manipulate vectors in R?
   (b) Explain the following functions with suitable illustrations.
   i) scan ii) print iii) format iv) paste and v) grep

UNIT – II

3. (a) Explain how to create and manipulate matrices in R with suitable illustrations. Also, explain various operators applicable on matrices.
   (b) Explain, in detail, the creation and manipulation of data frames.
4. (a) Explain how to read data from various types of files by means of illustrations. Further, explain write.csv (), write. table (), file. choose (), setwd() and getwd() functions with suitable illustrations.
   (b) Explain the creation and manipulation of lists.

UNIT – III

5. (a) Explain various control statements in R by writing their syntax. Give an illustration in each case.
   (b) Write R program to find mean and median of the given sample without using built-in R functions.
6. a) What are user-defined functions? Explain them in detail with suitable illustrations.
   (b) Write your own R function for two sample t-test.

UNIT – IV

7. (a) Write R-codes for generating samples of size n=1000 from each of the following probability distributions and for plotting the density functions for the respective distributions.
   i) Poisson(10) ii) N(10,100) iii) exp(5) and iv) logistic(100,20)
   (b) Write R- function for finding binomial probability and hence write R- code for fitting of binomial distribution based on a given frequency data and test for goodness of fit.
8. (a) Write down the built-in R-syntax for the following tests and explain them.
   i) Kolmogorov –Smirnov test for goodness of fit.
   ii) Wilcoxon Mann-Whitney two- sample U-test.
   iii) Chi-square test for goodness of fit.
   (b) Write R- code for solving the equation e^{2x}-x-6=0 using Newton-Raphson method.

UNIT – V

9. (a) Explain the following high-level plotting commands in details
   i) plot() ii) barplot() iii) pie() iv) hist()
   (b) Explain various low-level plotting commands available in R.
10. (a) Write down the built-in R-syntax for drawing bar chart and Q-Q plot.
    (b) Write down the built-in R-syntax for the following tests and explain them.
    i) CRD analysis ii) RBD analysis

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FIRST Semester

Course – S 1.3 :: ESTIMATION

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: Three Hours **Maximum: 70marks**

Answer **ONE** question from **each Unit**

(Each question carries equal marks)

UNIT-I

1. (a) Distinguish between estimator and estimate. Explain (i) Sampling distribution and (ii) Standard error and its utility.
   (b)Explain (i) Sufficiency and minimal sufficiency. Let X_i(i=1,2,…n) be a random Sample from an exponential distribution with p.d.f $f_\theta(x)=\exp[-(x-\theta)]$, $\theta<x<\infty$, $-\infty<\theta<\infty$ Obtain sufficient statistics for θ.
2. (a) State and prove Fisher- Neymann Factorization theorem.
   (b) Explain Complete sufficiency and Minimal Sufficiency with examples.

UNIT-II

3. (a) Construct UMVUE using Cramer- Rao Inequality.
   (b) Explain (i) Consistency (ii) relative efficiency and (iii) CAN and CAUN estimators with suitable examples.
4. (a) State and prove Lehman- Scheffe theorem.
   (b) State and prove necessary and sufficient conditions for the existence of MVUE.

UNIT-III

5. (a) Explain M.L method of estimation. Show that in sampling from the distribution with p.d.f $f_\theta(x)=\exp[-\theta x]$, $\theta<x<\infty$ $\frac{t}{\bar{x}}$ is the M.L estimator of θ and has a greater variance then the unbiased estimator $\frac{n-1}{n\bar{x}}$.
   (b) Explain minimum chi-square and modified minimum chi-square methods of Estimation.
6. (a) Prove that the maximum likelihood estimate of the parameter α of a population having density function:$\frac{2}{\alpha}(\alpha - x), 0 < x < \alpha$.
   (b) Describe Percentile Estimation.

UNIT-IV

7. (a) Define UMA and UMAU confidence sets. Describe the relationship between Confidence estimation and testing of hypothesis.
   (b) Let $X_1, X_2, \ldots, X_n$ be n independent N (μ, σ^2) variables when μ is unknown but σ^2 is known. Let prior distribution of μ be N (θ, σ^2). Find Bayes estimate of μ.
8. (a) Define (i) confidence bounds and (ii) UMA confidence bounds. State and prove a Sufficient condition for a pivot to yield a confidence interval for a real valued Parameter.
   (b) Obtain 100 (1-α)% confidence limits (for large samples) for the parameter λ of the Poisson distribution.

UNIT-V

9. (a) Define type I censoring and dicuss the ML estimation of the parameters of N (μ, σ^2).
   (b) Explain one parameter exponential distribution.
10. (a) Explain type II censoring and obtain the ML estimate of the parameter in the case of one parameter exponential distribution.
    (b) Let $X_1, X_2, \ldots, X_n$ be n independent N (μ, σ^2) variables when μ is unknown but σ^2 is Known. Obtain the shortest expected length confidence interval for θ.

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FIRST Semester

Course – S 1.4 (A) :: SAMPLING THEORY

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: Three Hours **Maximum: 70marks**

Answer **ONE** question from **each Unit**

(Each question carries equal marks)

UNIT – I

1. (a) Discuss the optimum allocation in stratified sampling. Obtain the variance of an estimate of population proportion with stratified random sampling.
   (b) Define systematic sampling. If a population consists of a linear trend, then prove that
   $Var(\bar{y}_{st}) \le Var(\bar{y}_{sys}) \le Var(\bar{y}_n)_R$
2. (a) Discuss the estimation of gain in precision due to stratification.
   (b) Explain circular systematic sampling. Obtain the method of estimation of sample size with continuous data.

UNIT – II

3. (a) What is cluster sampling? Obtain an unbiased estimator of population total based on cluster sampling, with clusters of equal size, and derive an expression for the sampling variance of this estimator.
   (b)How do you determine the optimum cluster size so as to minimize the variance for a fixed cost.
4. (a) Explain PPS sampling with replacement (wr-pps). Obtain an unbiased estimator of the population total and variance of the estimator under wr-pps. Also derive the estimator for the variance.
   (b) Define Horvitz – Thompson estimator of the population mean and derive the variance of this estimator.

UNIT – III

5. (a) Derive an expression for estimating the variance of population mean in two stage sampling where SRSWOR is adopted at both stages.
   (b) Obtain an estimator for the population mean under double sampling with SRSWR at the first stage and SRSWR at the second stage.
6. (a) Obtain the variance of an estimate for the population mean under double sampling with SRSWR at the first stage and SRSWR at the second stage.
   (b) Discuss the problem of optimal allocation in double sampling.

UNIT – IV

7. (a) What is double sampling? In case of double sampling for difference estimation, propose an estimator for the population mean and derive its variance, stating the necessary assumptions, If any.
   (b) Distinguish between multistage sampling and multiphase sampling.
8. (a) What are the various sources and types of non-sampling errors. Explain in detail?
   (b) Briefly explain the concepts i) Hansen and ii) Hurwitz Technique and Deming's Model.

UNIT – V

9. (a) Derive the bias and mean square error of regression estimator of the population total assuming SRSWR for the units.
   (b) Explain difference estimation. Define a separate difference estimator for population mean and obtain its variance.
10. (a) Explain ratio estimation. Obtain the variances of ratio estimates in stratified sampling.
    (b) Obtain the several of the regression estimation. Obtain the variance of regression coefficient with pre-assigned value.

Course Code: ST 1.4 (B)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FIRST Semester

Course – S 1.4 (B) :: LINEAR ALGEBRA

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: Three Hours **Maximum: 70 marks**

Answer **ONE** question from each unit

(Each question carries equal marks)

UNIT-1

1. a) State and prove rank – nullity theorem.
   b)Define a Linear transformation. Show that the mapping $T: R^3 \to R^2$ is defined by $T(x,y,z)=(x-y,x-z)$ is a linear transformation.
2. a) Show that $T: R^2 \to R^2$ is defined by $T(x_1,x_2)=(x_1+x_2,x_1-x_2+1)$ is not linear.
   b) Describe explicity the linear transformation $T: R^2 \to R^2$, such that $T(1,2)=(3,0)$ and $T(2,1)=(1,2)$.

UNIT-II

3. a) Show that every subspace of a finite dimensional vector space is finite dimensional.
   b) Prove that a subset W is a subspace of a vector space V of F $\Leftrightarrow$ $\alpha w_1+\beta w_2 \in W$ $\forall\ \alpha,\beta \in F$ and $w_1, w_2 \in W$.
4. Prove that if W_1 and W_2 are finite - dimensional subspace of a vector space V, then the subspace W_1+W_2 is finite dimensional and dim $(W_1+W_2)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)$.

UNIT-III

5. a) State and prove Bessel's inequality.
   b) State and prove Parseval's identity.
6. a) State and prove Cauchy's – Schwarz inequality.
   b) State and prove Gram – Schmidt orthogonalization process.

UNIT-IV

7. a) State and prove Cayley – Hamilton theorem.
   b) If $A=\begin{pmatrix}3&1&1\\1&2&1\\3&1&1\end{pmatrix}$, find two non–singular matrices P,Q . Such that I=PAQ and hence find A^{-1}.
8. a) Find the characteristic roots and characteristic vectors of the matrix $\begin{pmatrix}8&-6&2\\-6&7&-4\\2&-4&3\end{pmatrix}$.
   b) Show that the equations x+y+z=4, 2x+5y-2z=3 and x+7y-7z=5 are not consistent.

UNIT-V

9. a) Find the minimal polynomial of the matrix $\begin{pmatrix}5&-6&-6\\-1&4&2\\3&-6&-4\end{pmatrix}$ and show that it is derogatory.
   b) If A and B be n × n matrices and B be non – singular, then show that A and B^{-1} AB have the same minimal polynomial.
10. If A, is a real symmetric matrix, then show that there exists an orthogonal matrix P such that $P'AP=P^{-1}AP$ is a diagonal matrix with real elements.

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

SECOND Semester

Course – S 2.1 :: MULTIVARIATE ANALYSIS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours **Maximum: 70 marks**

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT - I

1. a) Define the p-variate normal distribution with mean vector μ and dispersion matrix Σ. Derive two important properties of the multivariate normal distribution.
   b) Prove that the marginal distribution obtained from the multivariate normal distribution is normal.
2. a) Define the characteristic function of a p-dimensional random variable. Obtain the characteristic function of multivariate normal distribution.
   b) In the p-variate normal case, show that the sample mean vector and the sample covariance matrix are independently distributed.

UNIT - II

3. a) Define Hotelling's T^2 statistic. Show that Hotelling's T^2 statistic can be used to test the equality of means of corresponding variables in two MVN populations having the same variance-covariance matrix.
   b) Explain in detail the likelihood ratio principle.
4. a) Stating the assumptions clearly, discuss the problem of comparing several multivariate normal population means.
   b) State and prove the invariance property of Hotelling's T^2 statistic.

UNIT - III

5. a) Describe the classification between two unknown multivariate normal populations.
   b) Explain the problem of classification. Distinguish between discrimination and classification.
6. a) Derive Fisher's linear discriminant function in case of two unknown p-variate populations.
   b) Describe the method of classification of an individual into one of several p-variate normal populations having a common dispersion matrix ξ, where all the parameters are known.

UNIT - IV

7. a) Distinguish between cluster analysis and discriminant analysis. Consider the hypothetical distance between pairs of five objects as follows.

$$\mathbf{D} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & & & & \\ 9 & 0 & & & \\ 3 & 7 & 0 & & \\ 6 & 5 & 9 & 0 & \\ 11 & 10 & 2 & 8 & 0 \end{bmatrix} \end{array}$$

Cluster the five objects using single linkage method

b) Explain the following methods of cluster analysis.

1) Centeroid Linkage method 2) K-means method.

(Contd…2)

8. a) Explain various similarity measures. Explain complete linkage method.
   b) Explain non-hierarchical methods. Describe Ward's method in cluster analysis.

UNIT -V

9. a) Define principal components. If $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then explain how you would compute various principal components.
   b) Define Canonical variables and Canonical correlations. Explain how you estimate canonical correlations.
10. a) Explain the orthogonal factor model. Explain the ML estimation method of factor loadings.
    b) State and prove two properties of principal components.

Course Code: ST 2.2

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

SECOND Semester

Course – S 2.2 :: TESTING OF HYPOTHESIS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours **Maximum: 70 marks**

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Write about most powerful test and uniformly most powerful tests.
   (b) State and prove Neyman Pearson Lemma.
2. (a) Explain the concept of distribution with monotone likelihood ratio (MLR) property. Examine whether the family of densities

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le 0 \\ 0, & otherwise \end{cases}$$

   admit MLR or not.
   (b) Write briefly about power function and level of significance.

UNIT-II

3. (a) Define unbiased critical region. Show that an MP critical region is necessarily unbiased.
   (b) Explain MLR property.
4. (a) Describe non existence of UMP test for one parameter exponential family.
   (b) Describe the asymptotic distribution of the likelihood ratio test criterion.

UNIT-III

5. a) Explain Kruskal-Wallis test.
   b) Write usual notations show that SPRT terminates with probability one.
6. a) Explain the concepts:
   i) Variance stabilizing transformation and ii) Asymptotic power of large sample tests.
   b) Explain the chi-square test for homogeneity of correlation coefficients.

UNIT-IV

7. a) Write about Friedman's test.
   b) What do you mean by non-parametric test? Explain its merits and demerits.
8. a) Explain the sign test and its importance in testing of hypothesis.
   b) Describe the median test procedure.

UNIT-V

9. a) Explain sequential probability ratio test procedure.
   b) Explain the chi-square test goodness of fit. Derive the asymptotic distribution of Chi-square.
10. a) Explain Kendal's test.
    b) With usual notation show that SPRT terminates with probability one.

Course Code: ST 2.3

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

SECOND Semester

Course – S 2.3: THEORY OF LINEAR ESTIMATION AND ANALYSIS OF VARIANCE

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours **Maximum: 70 marks**

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT – I

1. (a) Explain inverse matrix and idempotent matrix.
   (b) State and prove Cayley-Hamilton theorem.
2. (a) Explain (i) determinant (ii) rank of a matrix and (iii) Inverse of a matrix with suitable example.
   (b) State and prove a necessary and sufficient condition for a real matrix to be positive definite.

UNIT – II

3. (a) Explain (i) linear model (ii) Best linear unbiased estimate.
   (b) State and prove Gauss – Markov theorem.
4. (a) Describe Generalized linear model with suitable example.
   (b) State and prove Aitken's theorem.

UNIT – III

5. (a) Describe one-way classification for equal no. of observations per cell
   (b) Explain Duncan' multiple range test
6. (a) Explain Fisher's least significant difference method
   (b) Explain analysis of variance two-way classification with multiple observations per cell.

UNIT – IV

7. (a) Write the applications of CRD and RBD.
   (b) Explain analysis of covariance with a single concomitant variable .
8. (a) Explain the analysis of variance two-way classification.
   (b) Explain analysis of LSD with one missing value.

UNIT – V

9. (a) State and prove Bartlett's test.
   (b) Briefly explain test for normality difference of variances.
10. (a) What is multiple range test and it's properties.
    (b) State and prove Turkey's test.

Course Code: ST 2.4 (A)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

SECOND Semester

Course – S 2.4 (A) :: STOCHASTIC PROCESSESS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours **Maximum: 70 marks**

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Define stochastic process and give an example. Describe the classification of stochastic processes.
   (b) Show that the transition probability matrix together with the initial distribution completely specifies a Markov Chain $\{X_n, n = 0,1,2\}$
2. (a) Show that if i $\leftrightarrow$ j and i is recurrent, then j is also recurrent. Hence deduce that recurrence is a class property.
   (b) State and prove a necessary and sufficient condition for a state of Markov Chain to be persistent state.

UNIT-II

3. (a) Explain doubly Poission process. Explain birth and death process and give some limitations on it.
   (b) Explain pure birth process and derive explicit expression for $P_n(t)$.
4. Define Poisson process? Discuss the properties of Poisson process and prove any three of them.

UNIT-III

5. (a) Explain renewal process with a suitable example. State and prove Wald's equation.
   (b) State and prove the key renewal theorem.
6. (a) Define (i) Renewal function and (ii) renewal equation. Show that a renewal function satisfies the renewal equation.
   (b) State and prove elementary renewal theorem.

UNIT – IV

7. (a) Describe a discrete time branching process $\{X_n, n = 0,1,2,.....\}$ with $X(0) = 1$. Obtain its mean and variance.
   (b) Show that the probability of extinction in a branching process is unity under the condition to be specified by you.
8. (a) Define a Galton-Watson Branching process under the usual notations of p.g.f. for such a process in the usual notation, show that $P_n(s) = P_{n-1}(P(s)) = P(P_{n-1}(s))$.
   (b) Define total number of progeny and derive its distribution.

UNIT – V

9. (a) Explain Martingale process and its properties with application.
   (b) Define i) Brownian Motion ii) Limit of Random Walk with examples.
10. (a) Explain Brownian Process and its properties with application.
    (b) The hitting time T_a of a standard Brownian motion has pdf $\frac{|a|t}{\sqrt{2\pi t^3}} e^{-a^2/2t}$ t>0 for any $\alpha \neq 0$.

Course Code: ST 2.4 (B)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

SECOND Semester

Course – S 2.4 (B) :: LINEAR MODELS AND APPLIED REGRESSION ANALYSIS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours **Maximum: 70 marks**

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Explain the Gauss-Markov model along with underlying assumptions. Also give an example of the same.
   (b) Obtain the maximum likelihood estimator of error variance. Derive its mean and variance.
2. (a) Show that in a general linear model the least squares estimators are BLUE.
   (b) Derive an expression for the dispersion matrix of the BLUE for parametric vector of the general linear model.

UNIT-II

3. (a) Explain restricted least squares estimation for a general linear model.
   (b) State and prove Aitken's theorem.
4. (a) Give the Simultaneous estimates of linear parametric functions.
   (b) What is estimable function and obtain the estimation with correlated observations.

UNIT-III

5. (a) Describe a test procedure to test the significance of a single parametric function.
   (b) Explain the test procedure to test the significance of Multiple Hypothesis.
6. (a) Explain analysis of variance two-way classification with multiple observations per cell.
   (b) Obtain the confidence intervals for the least squares estimates in the case of two variable linear model.

UNIT-IV

7. (a) What is simple linear regression? Explain with suitable example and obtain the partial Correlation coefficient.
   (b) Explain the multiple regression in three variable case and obtain the coefficient of multiple determination.
8. (a)Define polynomial regression and explain the use an orthogonal polynomials in a detailed way.
   (b) Write down the properties of regression coefficient.

UNIT-V

9. (a) Explain multicollinearity with suitable examples. What are the consequences of multicollinearity and explain some methods of detecting it.
   (b) What are ridge regression estimators and give its properties? Explain their use with an example.
10. (a) What are principal components? Explain their use in regression analysis by a suitable example.
    (b) Explain the subset selection of explanatory variables.

Course Code: ST 3.1

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

THIRD Semester

Course – S 3.1 :: DESIGN AND ANALYSIS OF EXPERIMENTS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Explain 2^3 factorial experiment. Discuss the statistical analysis of 2^3 –design.
   (b) Explain Yates' method of computing factorial totals for 2^n factorial experiment.
2. (a) Explain the analysis of 3^2 design. Describe the linear and quadratic contrasts in this Design.
   (b) Distinguish between total and partial confounding. Explain the statistical analysis of confounding in 2^3 factorial design.

UNIT-II

3. (a) Explain BIBD. State and prove the parameter relations in a BIBD. Explain a method of constructing BIBD.
   (b) Explain PBIBD. Explain the intra block analysis of PBIBD.
4. (a) Describe the construction of BIBD and affine resolvable designs. Derive the parameter relations in an affine resolvable design.
   (b) Explain PBIBD with two associate classes. Explain the intra block analysis of PBIBD.

UNIT-III

5. (a) Explain spilt plot design. Discuss its analysis.
   (b) Describe the Gracco Latin square design. Explain the estimation and tests of hypothesis in the Model.
6. (a) Explain Youden square design and its analysis.
   (b) Explain simple lattice design and its analysis.

UNIT-IV

7. (a) Explain response surface methodology. Discuss the analysis of a second order response surface.
   (b) Explain the construction of a second order notable design using BIBD.
8. (a) Explain the construction of a second order notable design using central composite design V=3 factors.
   (b) Explain the properties of response surface designs. Obtain the variance of the estimated Second order response surface design.

UNIT-V

9. (a) Explain Taguchi Method.
   (b) Explain loss function and signal to ratio and their uses.
10. (a) Discuss the role of Taguchi philosophy in experimental design.
    (b) Explain Critique of S/N Ratios.

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

THIRD Semester

Course – S 3.2:: STATISTICAL QUALITY CONTROL

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT – I

1. a. Explain the construction and operation of $\bar{X}$ and S control charts.
   b. Describe She wart control charts for Individual measurements.
2. a. Explain the constructions of group control charts and multiple stream process.
   b. Explain process characteristics and cost parameters in economic design of control charts.

UNIT – II

3. a. Write the basic principles of CUSUM control chart for monitoring the process mean.
   b. Explain the V-mask procedure. What are its disadvantages?
4. a. Explain the constriction of moving average control chart. Comment on the equivalence of moving average and EWMA control chart.
   b. Explain Moving average control chart.

UNIT – III

5. a. Define statistical process control. Explain its importance in short production runs.
   b. Explain Modified control limts for $\bar{X}$ chart.
6. a. Discuss the constriction of attribute control chart for short production runs.
   b. Explain control chart for a "Six-Sixma process".

UNIT – IV

7. a. Explain the construction of Hotelling's T^2 Control chart for sub-grouped data.
   b. Discuss about the multivariate EWMA control chart.
8. a. Explain the control chart for monitoring variability with suitable example.
   b. Explain the method of partial least squares. Write its applications.

UNIT – V

9. a. Explain tools and techniques of Total Quality Management.
   b. Discuss statistical quality control as a tool of quality management.
10. a. What are Bench marking practices and Quality Auditing in Quality Systems.
    b. Discuss QS-9000 standards.

Course Code: ST 3.3 (A)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

THIRD Semester

Course – S 3.3 (A) :: ACTUARIAL STATISTICS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT - I

1. (a) Define the following terms:
   i. Effective rate of interest ii. Nominal rate of interest iii. Accumulation factor and iv. Discount rate. With suitable examples.
   (b) What is force of interest varying continuously? With the usual notation show that $\lim_{m\to\infty} d^{(m)} = \lim_{m\to\infty} i^{(m)} = \delta$

2. (a) With the normal motivation show that
   (i) d = 1-v and
   (ii) d = iv
   (b) Show that, if interest is paid at rate i, the amount at time t under simple interest is more than the amount at time t under compound interest provided $t < 1$. Show also that the inverse inequality holds if $t > 1$.

UNIT – II

3. (a) What is an annuity? Explain different types of annuities with examples.
   (b) With the usual notation show that $a_{\overline{n}|}^{(m)} = i\, a_{\overline{n}|} \,/\, i^{m}$.

4. (a) Derive the variance for the continuous annuity present value.
   (b) Obtain the expected value present value for arithmetically increasing annuities.

UNIT-III

5. (a) Stating the assumptions clearly derive the general expression for principal and interest portion of a loan.
   (b) Distinguish between loan and mortgage. Also explain how do you calculate APR for a given loan of amount L at force of interest $\delta(.)$ In a cash flow $\left((t_1, X_1),(t_2, X_2),...,(t_n, X_n)\right)$.

6. (a) Explain the stochastic nature of a present value of unit due at the end on n years. Also derive its mean and variance by stating the conditions clearly.
   (b) Explain different types of repayment schemes with dependent interest rates.

UNIT – IV

7. (a) Define NPV and IRR. Explain how do you compare two investment projects.
   (b) What is inflation? Explain its effect on IRR.

8. (a) Explain the concept and utility of LIRR in actuarial statistics with a suitable example.
   (b) Let C be a cash flow with yield y(c) and a constant inflation rate e. Then show that the real yield of c exists and is $y_e(c) = \frac{y(c) - e}{1 + e}$.

UNIT – V

9. (a) Define joint distributions of future life times and state its properties.
   (b) Show that $F_T(t) = 1 - S_{T(x)T(y)}(t, t)$.

10. (a) Under the assumption of uniform distribution of deaths over each year of age and i=0.06 Calculate the following for a 20-year endowment insurance issued to (50) with a unit benefit and true semiannual benefit premiums.
    (b) Discuss in detail (i) central rates of multiple decrements (ii) central force assumptions for multiple decrements.

Course Code: ST 3.3 (B)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

THIRD Semester

Course – S 3.3 (B) :: TIME SERIES ANALYSIS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT - I

1. (a) Define autocorrelation and auto covariance functions and show that autocorrelation function lies between -1 and +1.
   (b) Describe any two tests for seasonality.
2. (a) How time series can be viewed as a discrete parameter Stochastic Process?
   (b) Explain the variate difference method in estimating the variance of irregular component of a time series.

UNIT – II

3. (a) What is double exponential smoothing? Explain Holt's two parameter method.
   (b) Show that IMA (1, 1) is equivalent to single order exponential smoothing.
4. (a) Describe the Piegels classification of time series.
   (b) Explain moving averages for constant and linear trend process.

UNIT-III

5. (a) Explain Box-Jenkins mixed auto regressive average process.
   (b) What is an auto-regressive process? Explain any one method to choose the period of auto regressive process.
6. (a) How do you model the non-stationary time series? Explain with example.
   (b) Derive the auto correlation function of ARIMA process.

UNIT – IV

7. (a) Explain analysis of residuals as goodness of fit of a model.
   (b) How do you perform diagnostic checking of ARIMA model?
8. (a) Obtain the estimate of parameters of ARIMA model.
   (b) Explain estimation procedure of autocorrelation function under large sample theory.

UNIT – V

9. (a) Derive the spectral density of a ARMA process.
   (b) Discuss the role of correlogram analysis in fitting a model to the stationary process.
10. (a) Explain the periodigram. How it is useful in identifying periodicity of a time series?
    (b) What is a Fourier transformation? Explain the uses of Fourier transformations in time series analysis?

Course Code: ST 3.4 (A)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

THIRD Semester

Course – S 3.4 (A) :: RELIABILITY - I

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Define reliability. Explain the importance of reliability and concept of failure.
   (b) Discuss about the general provision of a reliability specification.
2. (a) Describe bath-tub curve and its characteristics.
   (b) Explain reliability measures. Obtain reliability in terms of hazard rate and failure density.

UNIT-II

3. (a) State and prove any three properties of exponential failure model.
   (b) Explain the importance of Weibull failure model. Give the applications of the model in reliability studies.
4. (a) Explain the importance of truncated normal distribution as a failure distribution. Obtain its reliability function.
   (b) Discuss any three types of log-normal distributions with their properties.

UNIT-III

5. (a) Describe: i) series ii) Parallel and iii) r out of n configuration and their block diagram with suitable examples.
   (b) Explain about the event space, cut set and tie set with suitable examples.
6. (a) Explain the determination of reliability through combinatorial methods with suitable illustrations.
   (b) Explain multi state models and write their applications.

UNIT-IV

7. (a) Describe system availability measures. Obtain the reliability of
   (i) Series (ii) Parallel and (ii) K and N systems with exponential components.
   (b) Explain: (i) MTBF (ii) MTTF (iii) Reliability of stand by System with examples.
8. (a) Explain chain model. Discuss the non-parametric estimation of reliability.
   (b) Explain (i) Weekest-link model and (ii) Stress-Strength model.

UNIT-V

9. (a) Obtain the M.L. estimate of the reliability in case of normal failure model.
   (b) Explain normal probability plot. Explain the importance of failure-rate plot in the analysis of life-test data.
10. (a) Obtain the Maximum Likelihood estimate of reliability in the case of exponential model.
    (b) Explain the probability plotting to fit a two-parameter weibull distribution.

Course Code: ST 3.4 (B)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

THIRD Semester

Course – S 3.4 (B) :: BIOSTATISTICS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT - I

1. (a) Explain the process in conducting Phase-I and Phase-II of a clinical trial.
   (b) Demonstrate the technique of sequential design.
2. (a) Discuss different phases of clinical trials.
   (b) Explain dynamic randomization.

UNIT – II

3. (a) Define parallel line assay. What is qualitative response?
   (b) Summarize the nature of direct assay. Define dose response relationship.
4. (a) Distinguish between qualitative and quantitative dose response relations.
   (b) State and prove Feller's theorem.

UNIT-III

5. (a) Describe logistic regression model and its diagnostics.
   (b) Explain the concept of Poisson regression model for count data.
6. (a) Explain Wald's statistics. Describe estimation of relative risk.
   (b) Define odds ratio. Explain binomial logit model for binary data.

UNIT – IV

7. (a) Explain ROC curve. What are the properties of ROC curve?
   (b) Derive Kullback-Leibler Divergence.
8. (a) Estimate the binomial model and area under the curve.
   (b) What are the functional relationships between Kullback-Leibler Divergence and slope of the ROC Curve?

UNIT – V

9. (a) Construct the analysis of variance of repeated measures for one way classified data.
   (b) Distinguish between information bias and selection bias.
10. (a) Explain cohort study design and its analysis.
    (b) Describe case control study design and give its analysis.

Course Code: ST 4.1

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FOURTH Semester

Course – S 4.1 :: ECONOMETRICS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT–I

1. a) Explain simple linear regression model and describe the justification of including stochastic error term in the model.
   b) Show that the OLS estimators in a simple linear regression model are BLUEs.
2. a) Describe the prediction problem in simple linear model.
   b) Explain the following models.
   (i) log-linear (ii) log-log and (iii) logarithmic Reciprocal

UNIT-II

3. a) In a general linear model $Y = X\beta + \varepsilon, \quad \varepsilon \sim (0, \sigma^2 I_n)$, derive the OLS estimators of β and σ^2.
   b) State and prove Gauss-Markov theorem.
4. a) Describe the general linear model stating the underlying assumptions.
   b) In general linear model, derive the test for significance of complete model.

UNIT-III

1. a) Explain the problem of multicollinearity and describe Ridge regression method as remedial measure of multicollinearity.
   b) State and prove Chow-test for comparison of the equality of two regression equations.
2. a) Derive GLS estimator in generalized linear model and show that GLS estimator is same as ML estimator.
   b) Discuss the sources and consequences of multicollinearity. Explain variance inflation factor test for detection of multicollinearity.

UNIT-IV

3. a) Discuss the sources and consequences of heteroscedasticity.
   b) Explain Cochrane-Orcutt's two-step method for estimation of the model under autocorrelation.
4. a) Explain the sources and consequences of autocorrelation. Explain Durbin-Watson test for autocorrelation.
   b) Explain Breusch–Pagan–Godfrey test for detection of heteroscedasticity.

UNIT-V

5. a) Explain the LOGIT Model and describe an estimation procedure of the model.
   b) What are instrumental variables and explain their use. Explain the Stock adjustment model.
6. a) Explain the need of qualitative response regression models and explain Polynomial regression model.
   b) What are errors in variables? Describe the estimation procedure of Koyak's distributed lag model.

Course Code: ST 4.2

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FOURTH Semester

Course – S 4.2 :: ACCEPTANCE SAMPLING PLANS

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT – I

1. a. Explain the advantages and disadvantages of sampling. What are the guidelines for using acceptance sampling?
   b. Explain single sampling plan. Obtain its OC and ASN functions.
2. a. Define Average Sampling Number. Write the useful steps in Military standard105E plans.
   b. Describe Dodge-Romig sampling plans.

UNIT – II

3. a. Define Acceptance Sampling Plans for variavles. Explain its advantages and disadvantages.
   b. Explain estimating process average in variable sampling plans.
4. a. Obtain Double specification limits for a known process standard deviation.
   b. Discuss Sequential sampling plans by variables.

UNIT – III

5. a. Explain Acceptance Sampling Procedures. Write its importance?
   b. Write the advantages and disadvantages of Chain sampling.
6. a. Discuss continuous sampling procedures (CSP)-III and its variations .
   b. Discuss MIL-STD-414 and its use.

UNIT – IV

7. a. Explain tools and techniques of Total Quality Management.
   b. Discuss Statistical Quality Control as a tool of quality management.
8. a. Discuss QS-ISO 9000 standards?
   b. Explain notion of six sigma and its use.

UNIT– V

9. a. Briefly explain lot-by-lot sampling technique.
   b. what is Accepting Sampling Plan? Briefly explain inspection of Accepting Sampling Plan.
10. a. Explain 100 percent inspection accepted plan. Give an example?
    b. Explain 100 percent inspection rejected plan. Give an example?

Course Code: ST 4.3 (A)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FOURTH Semester

Course – S 4.3 (A) :: OPERATIONS RESEARCH

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT–I

1. (a) Explain the graphical method of solving a 2 x N game with a suitable example.
   (b) Explain i) Two-person Zero sum game ii) Pure and Mixed strategies iii) Dominance property with suitable examples.
2. (a) State and prove Minimax theorem used in game theory.
   (b) Find the game value for the given payoff matrix using LPP Simplex method and also the optimal strategies for the players.

Player B

$$\text{Player A} \begin{bmatrix} 3 & -1 & -3 \\ -3 & 3 & -1 \\ -4 & -3 & 3 \end{bmatrix}$$

UNIT-II

3.(a) Explain the problem of Economic Order Quantity without shortages.
   (b) A contractor has to supply 10,000 bearings per day to an automobile manufacturer. He finds that, when he starts a modulation sum, he can produce 25,000 bearings per day. The cost of holding a bearing in stock for one year in 20 paise, and set-up cost of a production sum is Rs. 180.00. How frequently should production run be made?

4.(a) Explain purchase inventory problem with one price break. Write the algorithm for Purchase inventory problem with n price breaks.
   (b) Define the terms safety stock and EOQ with the help of ideal inventory model.

UNIT–III

5. (a) Explain the queueing system. What are its characteristics?
   (b) If a period of 2 hours in a day (8-10 A.M.) trains arrive at the yard every 20 minutes but the service time continuous to remain 36 minutes, then calculate for this period.
   i) the probability that the yard is empty ii) average queue length, on the assumption that the line capacity of the yard is limited to 4 trains only.
6. (a) Explain M/G/1 system. Derive the expected queue length.
   (b) Explain M/M/1:N/FIFO system. Obtain its study state solution. Find the expected queue length.

Contd...2

UNIT–IV

7. (a) Discuss about the replacement problem of items that deteriorate with time.
   (b) Explain the replacement of items that fail completely.
8. (a) Explain the types of maintenance. Explain (A/P, i, n) with suitable example.
   (b) A Truck owner finds his past records that the maintenance cost per year of a truck whose purchase price is Rs. 8000 are given below:

Year	1	2	3	4	5	6	7	8
Maintenance cost (Rs.)	1000	1300	1700	2200	2900	3800	4800	6000
Resale Price (Rs.)	4000	2000	1200	600	500	400	400	400

Determine at which time it is profitable to replace the truck.

UNIT–V

9. (a) Explain in brief: procedure of CPM, PERT, dummy activities and lead time with reference to project management.
   (b) The time and cost estimates and precedence relationship of the different activities constituting a project are given below:

Activity	Predecessor activity	Time (in weeks)		Cost (in rupees)	
		Normal	Crash	Normal	Crash
A	None	3	2	8,000	9,000
B	None	8	6	600	1,000
C	B	6	4	10,000	12,000
D	B	5	2	4,000	10,000
E	A	13	10	3,000	9,000
F	A	4	4	5,000	5,000
G	F	2	1	1,200	1,400
H	C,E,G	6	4	3,500	4,500
I	F	2	1	700	800

   (i) Draw a project network diagram and find the critical path.
   (ii) If a dead line of 17 weeks is imposed for the completion of the project, what activities will be crashed? What would be the additional cost and the critical activities after crashing the project?

10. (a) Briefly explain the following with examples in relation of network analysis:
    (i) Crashing (ii) Resource Allocations.
    (b) A project has the following characteristics

Activity	(1-2)	(2-3)	(2-4)	(3-5)	(4-5)	(4-6)	(5-7)	(6-7)	(7-8)	(7-9)	(8-10)	(9-10)
Most Optimistic Time (a)	1	1	1	3	2	3	4	6	2	5	1	3
Most Pessimistic Time (b)	5	3	5	5	4	7	6	8	6	8	3	7
Most likely Time (m)	1.5	2	3	4	3	5	5	7	4	6	2	5

Construct PERT network and compute: (i) Find critical path and variance for each event. Find the project duration at 95% probability.

Course Code: ST 4.3 (B)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FOURTH Semester

Course – S 4.3 (B) :: MACHINE LEARNING USING PYTHON

(w.e.f. 2023-2024 Academic Year admitted batch)

Time : Three Hours **Maximum: 70 marks**

Answer **ONE** question from each unit

(Each question carries equal marks)

UNIT-1

1. (a) Differentiate the list and dictionary data types of python by their characteristics along with example in brief.
   (b) What do you mean by slicing operation in string of python? Write an example of slicing to fetch first name and last name from full name of person and display it.
2. (a) Discuss the role of indentation in python.
   (b) Explain range() function with suitable examples.

UNIT-II

3. (a) List the type of plots that can be drawn using matplotlib.
   (b) Write a python program to read data from CSV files using pandas.
4. (a) Explain Series in Pandas with example.
   (b) Explain Numpy array with example.

UNIT-III

5. (a) Describe K-nearest Neighbour learning Algorithm for continues valued target function.
   (b) Greedily learn a decision tree using the ID3 algorithm and draw the tree
6. (a) Write in detail about Neural Networks
   (b) Explain about linear regression and logistic regression.

UNIT-IV

7. (a) What is dimensionality reduction? Justify the necessity for dimensionality reduction in the context of machine learning.
   (b) Explain in detail abut scaling training.
8. (a) Discuss the feature representation learning?
   (b) Explain about manifold learning

UNIT-V

9. (a) Explain about agglomerative clustering algorithm.
   (b) Explain the concepts of clustering approaches. How it differ from classification
10. (a) Cluster the following set of data using k-means algorithm with initial value of objects 2 and 5 with the coordinate values (4,6) and (12,4) as initial seeds.

Objects	X-coordinate	Y-coordinate
1	2	4
2	4	6
3	6	8
4	10	4
5	12	4

(b) Write DBSCAN algorithm.

Course Code: ST 4.4 (A)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FOURTH Semester

Course – S 4.4 (A) (22) :: RELIABILITY - II

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Explain two element non-repairable system and obtain its reliability.
   (b) Explain: (i) Poission process (ii) stand by system with suitable examples.
2. (a) Obtain the reliability of a two component repairable system.
   (b) Describe up-time and down-time ratio and steady state probabilities in repairable system.

UNIT-II

3. (a) Define modular decomposition. Explain its use with a suitable example. Obtain the improved bounds on system reliability stating the assumptions clearly.
   (b) Show that dual of a coherent structure is dual. Let $h(p)$ be the reliability function of a coherent structure. Then $h(p)$ is strictly increasing in each p_i, for 0< p_i <1.
4. (a) Explain (i) coherent system (ii) paths and cut sets with suitable examples.
   (b) Explain associated random variables. If $\chi_1, \chi_2, \ldots \chi_n$ are associated binary random variables, then show that

i) $P\left(\prod_{i=1}^{n} \mathrm{X}_i = 1\right) \geq \left(\prod_{i=1}^{n} P(\mathrm{X}_i = 1)\right)$ ii) $P\left(\coprod_{i=1}^{n} \mathrm{X}_i = 1\right) \leq P\left(\coprod_{i=1}^{n} \mathrm{X}_i = 1\right)$

UNIT-III

5. (a) Explain (i) Event tree technique and (ii) Fault tree and its construction with suitable examples.
   (b) The logical expression for a fault tree is given by $T = A \cap (B \cup C) \cap [E \cap F \cap G]$
   (i) Construct the corresponding fault tree (ii) Find the minimum cut sets
   (iii) construct and equivalent reliability block diagram.
6. (a) Describe Failure Modes and Effect Analysis (FMEA) with an example.
   (b) Explain minimum cut set formulation and its use. Explain min-cut set algorithms.

UNIT-IV

7. (a) Explain the importance of Exponential failure rate model in reliability analysis.
   (b) Explain Log normal and Pareto failure rate models.
8. (a) Describe the terms (i) Life tables and (ii) mean residual life with examples.
   (b) Explain Gamma and failure rate model. Explain the M.L. method of estimating the parameters in the model.

UNIT-V

9. (a) Explain (i) IFR (ii) IFRA (iii) NBU and (iv) NBUE classes with suitable examples.
   (b) Obtain Kaplan-Meier estimator. Explain a method of estimation it under the assumption of IFR / DFR.
10. (a) Define the terms (i) Estimation of Survival function (ii) Actuarial Estimator
    (b) Explain Cox's proportional hazard model several covariates and discuss the method of estimation.

Course Code: ST 4.4 (B)

MODEL QUESTION PAPER

M.Sc. Degree Examination

STATISTICS

FOURTH Semester

Course – S 4.4 (B) :: KNOWLEDGE DISCOVERY AND DATA MINING

(w.e.f. 2023-2024 Academic Year admitted batch)

Time: 3 hours Maximum: 70 marks

Answer ONE question from each unit.

(Each question carries equal marks)

UNIT-I

1. (a) Explain the classification problem involving two multivariate normal populations.
   (b) Explain how a classification tree is constructed.
2. (a) Explain ID3 algorithm. What are the merits and demerits of this algorithm?
   (b) Explain the pessimistic error pruning method.

UNIT-II

3. (a) Explain k-means clustering. What are its merits and demerits?
   (b) Describe vector quantization. Distinguish between clustering and vector quantization.
4. (a) Explain the density based clustering method DBSCAN. What are its strengths and weaknesses?
   (b) Explain vector quantization as a method of dimension reduction.

UNIT-III

5. (a) What are principal curves? Define principal curve of a p-dimensional distribution function. Give an algorithm for constructing principal curves.
   (b) What is feature selection? Distinguish between filter, wrapper and embedded feature selection methods.
6. (a) Give a brief account of projection pursuit?
   (b) Explain the following in the context of feature selection:
   (i) Subset Generation (ii) Evaluation of Subset
   (iii) Stopping Criteria (iv) Result Validation
7. (a) Describe the steps in building a regression tree. Explain the splitting criteria used in growing regression trees.
   (b) Explain the cost-complexity minimization method of pruning.

UNIT-IV

8. (a) Explain the stopping criteria used in growing a regression tree.
   (b) Explain the following pruning methods: (i) Reduced error pruning and (ii) Minimum error pruning.
9. (a) What are the differences between OLTP and OLAP?
   (b) What is association rule mining? Explain.

UNIT-V

10. (a) Describe the steps for the design and construction of data warehouses.
    (b) Describe Apriori Algorithm for mining frequent item sets for Boolen association rules.