M.Sc. Statistics Program

(Based on UGC – Learning Outcomes-Based Curriculum Framework)

School of Mathematics Statistics and Computational Sciences

Department of Statistics

Programme Objective:

The main objective of M.Sc. in Statistics programme is to enhance the theoretical, practical skills and concepts that students have been introduced to some extent at their undergraduate learning of Statistics. We teach higher level theoretical aspects and practicals as well which help them to formulate statistical problems and analyze them by using appropriate statistical methods and Statistical computing software. The course content and classroom teaching and evaluation system emphasis the applications of Statistics in different disciplines including Actuarial Science, Social Science by which they will be have better job opportunities and as well they are well prepared for competitive examinations..

Programme Learning Outcomes

Post Graduates of the M.Sc. Statistics program will be able to:

PLO1: Have a broad background in Statistics, an appreciation of how its various sub disciplines are inter-related, acquire an in-depth knowledge about topics chosen from those offered through the department.

PLO2: Develop the ability to effectively and aptly use techniques from different sub- disciplines in a broad range of real life problem solving.

PLO3: Have the versatility to work effectively in a broad range of companies (including R&D sectors of financial, pharmaceutical, market research, software development companies, consultancy, etc.), or analytic, scientific, government, financial, health, teaching and other positions or continue for higher education.

PLO4: Be able to independently read statistical literature including survey articles, scholarly books, and online sources.

PLO5: Be life-long learners able to independently expand their statistical expertise when needed, or out of own interest.

M.Sc./M.A. Statistics

I Semester

Course Code	Title	Credit	Hours per week		
			Lectures	Tutorial	Practical
STA 401	Probability Theory	4	3	1	0
STA 402	Distribution Theory	4	3	1	0
STA 403	Real Analysis and Linear Algebra	4	3	1	0
STA 404	Sampling Theory	4	3	1	0
STA 405	Practicals	4	0	0	8

II Semester

Course Code	Title	Credit	Hours per week		
			Lectures	Tutorial	Practical
STA 406	Estimation and Testing of Hypotheses	4	3	1	0
STA 407	Linear Models	4	3	1	0
STA 408	Stochastic Models	4	3	1	0
STA 409	Design of Experiments	4	3	1	0
STA 410	Practicals	4	0	0	8

III Semester

Course Code	Title	Credit	Hours per week		
			Lectures	Tutorial	Practical
STA 501	Time Series Analysis & Forecasting	4	3	1	0
STA 502	Multivariate Analysis	4	3	1	0
	Elective - 1	4	3	1	0
	Elective - 2	4	3	1	0
STA 503	Practicals	4	0	0	8

IV Semester

Course Code	Title	Credit	Hours per week		
			Lectures	Tutorial	Practical
	Elective-I	4	3	1	0
	Elective-1	4	3	1	0
STA 504	Practicals	4	0	0	8
STA 505	Project	12	-	-	-

Elective Courses for III-Semester

Course Code	Title	Credit	Hours per week		
			Lectures	Tutorial	Practical
STA 521	Financial Mathematics	4	3	1	0
STA 522	Data Mining	4	3	1	0
STA 523	National Development Statistics	4	3	1	0
STA 524	Population Studies	4	3	1	0
STA 525	Principal and Practices of Insurance	4	3	1	0
STA 526	Statistical Methods of Non-Life Insurance	4	3	1	0
STA 527	Statistical Quality Control	4	3	1	0
STA 528	Survival Analysis	4	3	1	0
STA 529	Statistical Methods for Bio-Computing	4	3	1	0
STA 530	Computer Intensive Statistical Methods	4	3	1	0
STA 531	Decision Theory and Non Parametric Inference	4	3	1	0

Course code from STA 521-STA 540 refer to elective courses for III semester (M.Sc./M.A. Statistics)

Elective Courses for IV-Semester

Course Code	Title	Credit	Hours per week		
			Lectures	Tutorial	Practical
STA 541	Contingencies	4	3	1	0
STA 542	Econometrics	4	3	1	0
STA 543	Extreme Value Theory	4	3	1	0
STA 544	Life and Health Insurance	4	3	1	0
STA 545	Statistical Methods for Reliability Theory	4	3	1	0
STA 546	Statistical Quality Management	4	3	1	0
STA 547	Stochastic Finance	4	3	1	0
STA 548	Machine Learning	4	3	1	0
STA 549	Statistical Analysis of Clinical Trials	4	3	1	0
STA 550	Bayesian Inference	4	3	1	0

Course code from STA 541-STA 560 refer to elective courses for IV Semester (M.A./ M.Sc. Statistics)

FIRST SEMESTER

Course Code	STA 401
Course Name	Probability Theory
Credits	04

Objective:
The main purpose is to introduce Probability Theory under Axiomatic approach and develop further theory and concepts including the limit behaviours.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

CLO-1: Recognize the concept of field, sigma field, probability space, probability measure.
CLO-2: Understand the concept of convergence of sequences of random variables.
CLO-3: Apply various inequalities to solve complex statistical problems.
CLO-4: List various types of random variables
CLO-5: Apply different probability theorems and laws for solving different mathematical problems.

Unit-1

Classes of sets, field and sigma fields, limit of sequences of subsets, sigma field generated by a class of subsets, Borel fields. Probability measure on a sigma field, probability space, continuity of a probability measure. Real and vector-valued random variables.

Unit-2

Distribution functions of discrete rvs, continuous and mixed type rv, decomposition of a df. Expectation of rv and its properties. Linear properties of Expectations, Inequalities: Jensen's, Chebychevs, Markov, Hölders and Lyapounov inequalities.

Unit-3

Independent of two events and n(>2) events, sequence of independent events, independent class of events π-systems and λ–systems of events, Dykin's theorem(without proof) independence of rvs of events. Borel zero-one law, Borel-Cantelli Lemma, Kolmogorov zero-one law.

Unit-4

Convergence of sequences of random variables. Convergence in distribution and in probability. Almost sure convergence and convergence in the r^{th} mean. Implication between modes of convergence. Slutsky's theorem. Monotonic convergence theorem and dominated convergence theorem. Fatous lemma. Law of large number: weak law of large number, Tchebychev and Khintchine theorem (with proof) and strong law of large number (without proof). Inversion, Continuity and Uniqueness theorems of Characteristics function. Demoivre-Laplace Central Limit Theorem, Liapounovs and Lindeberg's CLT (without proof).

References

1. Bhat, B. R. (1999). Modern Probability Theory, 2/e, New Age International, New Delhi.
2. Rao. B. L. S. Prakasa (2009). A First course in Probability and Statistics. World Scientific
3. Meyer, P.A. An Introduction to Probability and Its Applications. PHI
4. Rohatgi V.K & A.K. MD. EhsanesSaleh (2001): An Introduction to Probability Theory and Mathematical Statistics, 2nd. John Wiley and Sons.

Course Code	STA 401
Course Name	Probability Theory
Credits	04

CLOs – PLOs Mapping matrix for STA 401 Probability Theory
(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	1	0	2	1
CLO-2	2	1	0	2	1
CLO-3	2	3	1	2	1
CLO-4	1	1	0	2	1
CLO-5	1	2	0	2	1
Total	8	8	1	10	5
Average	1.6	1.6	0.2	2	1

Course Code	STA 402
Course Name	Distribution Theory
Credits	04

Objective:
The main objective is to know the genesis of important distributions, their properties. Introducing of bivariate distributions, conditional and marginal distributions and distributions of Order Statistics.

Learning Outcome:

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

CLO-1: Differentiate discrete and continues random variables and its distribution.
CLO-2: Recognize various probability models and their properties.
CLO-3: Understand concept of compound, mixture and truncated random variables.
CLO-4: To simulate the realizations of complex experiments.
CLO-5: Use various distributions for variety of real life situations.

Unit-1

Review of Discrete and Continuous distributions. Weibull, Pareto, lognormal, Laplace, Cauchy, logistic, Rayleigh distribution their properties and applications.

Unit-2

Discrete and continuous bivariate random variables: Definitions, Computation of probabilities of various events, marginal, conditional, product moments and correlations. Conditional expectation and conditional variance.
The p. d. f. of a bivariate normal distribution, Marginal and conditional distributions, conditional expectation and conditional variance, regression lines of Y on X and X on Y., independence and uncorrelated-ness imply each other, m. g. f and moments. Plotting of bivariate normal density function.

Unit-3

Functions of random variables and their distributions using Jacobian of transformation and other tools. Distribution of distribution function. Bivariate exponential distributions.
Concept of a sampling distribution. Sampling distributions of t, χ^2 and F (central and non central), their properties and applications. Cochran's theorem. Independence of quadratic forms.

Unit-4

Compound, truncated and mixture distributions. Convolutions of two distributions. Order statistics: their distributions and properties. Joint, marginal and conditional distribution of order statistics. The distribution of sample range and sample median. Extreme values and their asymptotic distribution (statement only) with applications.

References

1. Rohatgi V.K & A.K. MD. EhsanesSaleh: An Introduction to Probability Theory and Mathematical Statistics, 2nd. John Wiley and Sons, 2001.
2. Johnson, Kotz and Balakrishna, Continuous univariate distributions, Vol- 1 IInd Ed, John Wiley and Sons
3. Johnson, Kemp and Kotz,Univariate discrete distributions, IIInd Ed, John Wiley and Sons
4. Mukhopadhyay P. (1996): Mathematical Statistics, New central Book Agency (P) Ltd. Calcutta.
5. Goon, Gupta & Das Gupta (1991): An Outline of Statistical Theory, Vol. I, World Press.
6. David, H. A., &Nagaraja, H. N. (1970). Order statistics. John Wiley & Sons, Inc..

Course Code	STA 402
Course Name	Distribution Theory
Credits	04

CLOs – PLOs Mapping matrix for STA 402: Distribution Theory

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	2	2	3	1
CLO-2	2	2	2	3	1
CLO-3	2	2	2	3	1
CLO-4	1	3	2	1	1
CLO-5	2	2	2	1	1
Total	10	11	10	11	5
Average	2	2.2	2	2.2	1

Course Code	STA 403
Course Name	Real Analysis and Linear Algebra
Credits	04

Objective:
The main purpose is to provide mathematical foundation for statistics courses to enhance their knowledge in Real Analysis and Linear algebra.

Learning Outcome:

Upon successful completion of this course, the student will be able to:
CLO-1: Relate applicability of real analysis and linear algebra in the various disciplines of statistics.
CLO-2: Understand special matrices, their properties and applications in statistics.
CLO-3: Employ the results from real analysis to solve various problems of probability theory.
CLO-4: Apply matrix theory for solving advanced statistical problems.
CLO-5: Use theory of stationary values for optimizing complex objective functions.

Unit-1

Review of basic differential and integral calculus. Elementary set theory, finite, countable and uncountable sets, Real numbers, limit point, interior point, open and closed subsets of R, supremum, infimum. convergence, limsup, liminf, Bolzano-Weisstrass theorem, Heine Borel theorem, continuity, uniform continuity, differentiability, Riemann sums and Riemann integral, Improper Integrals. Mean value theorem. Riemann-Stieltjes (R-S) integral of a bounded real valued function. Necessary and sufficient condition for R-S integrability. Properties of R-S integrals. Integration by parts. Change of variables in R-S integrals.

Unit-2

Sequences and series of functions, uniform convergence, Weierstrass test. Monotonic functions, types of discontinuity, functions of bounded variation. Functions of several variables, partial derivative, derivative as a linear transformation. Maxima and minima of functions of several variables. Lagrangian multipliers.

Unit-3

Vector spaces, subspaces, linear dependence, basis, dimension, algebra of linear transformations. Algebra of matrices, rank and determinant of matrices, inverse matrices, generalized inverse of a matrix and its properties, linear equations, eigen values and eigenvectors and their applications. Cayley-Hamilton theorem. Spectral decomposition of a symmetric matrix.

Unit-4

Matrix representation of linear transformations. Orthogonal transformations. Orthogonal and idempotent matrices. Change of basis, inner product spaces, canonical forms, diagonal forms. Quadratic forms, reduction and classification of quadratic forms.

References

1. Searle, S. R. (1982). Matrix Algebra Useful for Statistics; John Wiley, New York.
2. RamachandraRao, A. and Bhimasankaram, P. (1992): Linear Algebra, Tata McGraw hill.
3. Trench William (2003). Introduction to Real Analysis , Pearson Education
4. Krishnamurthy V., Mainra V.P. and Arora J. L. (2009) An introduction to Linear Algebra, East-West Press Pvt Ltd.
5. Rudin, W. (1985). Principles of Mathematical Analysis, McGrawhill, New York.
6. Malik, S.C. and Arora, S. (1998). Mathematical Analysis, New Age, New Delhi.
7. Bartle,R.G.(1975) The Elements of Real Analysis, 2/e, John Wiley.

Course Code	STA 403
Course Name	Real Analysis and Linear Algebra
Credits	04

CLOs – PLOs Mapping matrix for STA 403: Real Analysis and Linear Algebra

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	1	1	2	2	1
CLO-2	1	1	1	2	1
CLO-3	1	2	1	2	1
CLO-4	1	1	3	1	1
CLO-5	2	2	2	2	1
Total	6	7	9	9	5
Average	1.2	1.4	1.8	1.8	1

Course Code	STA 404
Course Name	Sampling Theory
Credits	04

Objective:
The main objective is to provide the knowledge of concept of sample and population in statistics and also the various sampling schemes. Estimation of population parameters and their respective standard errors.

Learning Outcome:

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

CLO-1: Recall necessity of sampling techniques.

CLO-2: Understand various types of sampling schemes, their advantages, disadvantages, and estimation of population parameters.

CLO-3: Compare different sampling techniques.

CLO-4: Select appropriate sampling technique for different experimental scenario.

CLO-5: Conduct sample survey.

Unit-1

Fixed population and super-population approaches. Distinct features of finite population sampling, Probability sampling design and estimators along with basic statistical properties. Review of some important results in SRSWOR and SRSWR.

Unit-2

Estimation of population mean/Total in stratified population, Allocation problem in stratified random sampling in case of fixed cost and also for specified precision. Expression for variance of stratified sample mean in case of fixed cost, formation and construction of strata, Post stratification, Double sampling with post stratification, Deep stratification, Controlled sampling.

Unit-3

Unequal probability sampling: PPSWR/WOR methods (including Lahiri's scheme) and DesRaj estimator, Murthy estimator (for n=2). Horvitz Thompson Estimator of finite population total/mean, Expression for Variance (HTE) and its unbiased estimator, Issue of non-negative variance estimation.

Unit-4

Double sampling scheme, some double sampling estimators for mean using auxiliary character (Ratio, regression and product) method of estimation, Some unbiased ratio type estimators for population mean, Concept of cluster sampling, two stage sampling, Two phase sampling, Non-sampling error with special reference to non-response problems.

References

1. Cochran, W.G: Sampling Techniques, Wiley Eastern Ltd., New Delhi.
2. Sukhatme, P.V., Sukhatme, B.V. and Ashok A.: Sampling Theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi.
3. Murthy, M.N: Sampling Methods, Indian Statistical Institute, Kolkata.
4. Daroga Singh and Choudhary F.S.; Theory and Analysis of Sample Survey Designs,WileyEastern Ltd., New Delhi.
5. Mukhopadhay, Parimal: Theory and Methods of Survey Sampling, Prentice Hall.

Course Code	STA 404
Course Name	Sampling Theory
Credits	04

CLOs – PLOs Mapping matrix for STA 404: Sampling Theory

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	2	1	2	1
CLO-2	2	2	2	2	1
CLO-3	2	2	1	2	1
CLO-4	2	3	3	2	1
CLO-5	3	3	3	2	1
Total	11	12	10	10	5
Average	2.2	2.4	2	2	1

Course Code	STA 405
Course Name	Practicals
Credits	04

Objective:
The main objective is to enhance the practical knowledge of an individual in statistical problem solving using Computer Software.

Learning Outcome:

- Learning to perform Statistical Computation using Software.

Content	Practical based on IMST 411-414

Students will be required to do practicals using R-software based on opted theory papers

- Convergence of the random variable.
- Fitting of discrete and continuous distributions
- Sketching of p.m.f./ pdf of discrete/ continuous distributions
- Random variable generation for Weibull, Pareto, lognormal, Laplace, Cauchy, logistic, Rayleigh distribution and computation of distributional properties.
- R- program (User defined) for Matrix operations (Multiplication, determinate, inverse, Eigen values and vector)
- Simple random sampling with and without replacement.
- Stratified random sampling.
- Unequal probability sampling: PPSWR/WOR methods (including Lahiri's scheme)
- Horvitz-Thompson Method of Estimations
- Double sampling
- Ratio Method of Estimation.
- Regression Method of Estimation.
- Cluster sampling
- Two stage sampling, Two phase sampling, Non-sampling error

SECOND SEMESTER

Course Code	STA 406
Course Name	Estimation and Testing of Hypotheses
Credits	04

Objective:
The main purpose is to make an individual understand basic theoretical knowledge about fundamental principles of statistical inference.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

CLO-1: Recognize different methods of parameter estimation.
CLO-2: Recall various properties of estimators.
CLO-3: Apply different statistical test procedures for different testing of hypothesis problems.
CLO-4: Compare different statistical test through power comparison.
CLO-5: Analyze various real life data sets using tests of hypothesis.

Unit-1

Criteria of a good estimator: unbiasedness, consistency, efficiency and sufficiency. Concept of mean squared error. Fisher-Neyman factorization theorem, Family of distributions admitting sufficient Statistic.
Point estimation, Maximum likelihood method (MLE), moments, Least squares method. Method of minimum chi-square and percentiles. Properties of maximum likelihood estimator (with proof). Successive approximation to MLE, Method of scoring and Newton-Raphson method.

Unit-2

Cramer-Rao inequality and its attainment, Cramer-Huzurbazar theorem (statement only), Completeness and minimal sufficient statistic, Ancillary statistic, Basu theorem, Uniformly minimum variance unbiased estimator (UMVUE). Rao-Blackwell and Lehmann-Scheffe theorems and their applications, Review of convergences of random variables and their implications, Delta method and its application, Asymptotic efficiency and asymptotic estimator, consistent asymptotic normal (CAN) estimator.

Unit-3

Statistical Hypothesis, critical region, types of errors, level of significance, power of a test, Test function, Randomized and non-randomized tests, Most powerful test and Neyman-Pearson lemma. MLR family of distributions, unbiased test. Uniformly most powerful test. Uniformly most powerful unbiased test. Likelihood ratio test with its properties. SPRT, OC curve, ASN function, Wald's equation and problems.

Unit-4

Confidence interval, confidence level, construction of confidence intervals using pivots, Determination of confidence intervals based on large and small samples, uniformly most accurate one sided confidence interval and its relation to UMP test for one sided null against one sided alternative hypotheses.

References

1. George Casella, Roger L. Berger, Statistical Inference, 2nd ed., Thomson Learning.
2. Mukhopadhyay P.: Mathematical Statistics, New central Book Agency (P) Ltd. Calcutta.
3. Rao, C.R.: Linear Statistical Inference and its Applications, 2nd ed, Wiley Eastern.
4. Rohatgi, V.K.: An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.
5. Goon, Gupta & Das Gupta: An Outline of Statistical Theory, Vol. II, World Press.
6. Hogg, R.V. and Craig, A.T.: Introduction to Mathematical Statistics, McMillan.
7. Kale, B.K. : A First Course on Parametric Inference, Narosa Publishing House.
8. Lehmann, E.L. Testing Statistical Hypotheses, Student Editions.

Course Code	STA 406
Course Name	Estimation and Testing of Hypotheses
Credits	04

CLOs – PLOs Mapping matrix for STA 406: Estimation and Testing of Hypotheses

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	3	1	2	1
CLO-2	2	2	1	2	1
CLO-3	3	3	3	3	1
CLO-4	2	3	3	3	1
CLO-5	3	3	3	3	1
Total	12	14	11	13	5
Average	2.4	2.8	2.2	2.6	1

Course Code	STA 407
Course Name	Linear Models
Credits	04

Objective:
The main purpose is to provide the theoretical foundations for the Linear Estimation Theory and Regression Analysis.

Learning Outcome:
Upon successful completion of this course, the student will be able to:
CLO-1: Employ a multiple linear regression model for real life data sets.
CLO-2: Perform statistical tests and construct statistical intervals in a multiple linear regression set up.
CLO-3: Validate regression model using different diagnostic procedures.
CLO-4: Estimate regression parameters in the presence of multicollinearity using ridge regression.
CLO-5: Opt appropriate link function for building regression model.

Unit-1

Theory of linear estimation, Estimable function, Simple linear regression, multiple regression model, least squares estimation, variance and covariance of least squares estimator, Gauss-Markov theorem in linear estimation.

Unit-2

Interval Estimation for regression coefficients β_0,β_1 and σ^2, Interval estimation of the linear functions of β. Interval estimation of the mean response, simultaneous confidence intervals. The R^2 statistic. Hypothesis testing for model adequacy, testing of sub hypothesis. Test of hypothesis for a linear parametric function. Point and interval prediction.

Unit-3

Fundamental concept of generalized linear model (GLM), exponential family of random variables. Link functions such as Logit, Probit, binomial, inverse binomial, inverse Gaussian, gamma. Non linear models, ML estimation in non linear models.

Unit-4

Diagnostic checks for suitability and validation of a linear regression model, graphical techniques, tests for normality, linearity, uncorrelated ness, multi collinearity, lack of fit, C_p criterion. Ridge regression, outliers and influential observations. Stepwise, forward and backward procedures for selection of best sub-set of repressors.

References

1. Montgomery, Douglas C.; Peck, Elizabeth A.; Vining, G. Geoffrey: (2003) Introduction to Linear Regression Analysis. John Wiley and sons.
2. Draper, N. R. & Smith, H(1998) Applied Regression Analysis, 3rd Ed., John Wiley..
3. Dobson, A. McCullagh, P &Nelder, J. A. (1989) Generalized Linear Models, Chapman & Hall.
4. Ratkowsky, D.A. (1983) Nonlinear Regression Modelling (Marcel Dekker).
5. Hosmer, D.W. & Lemeshow, S. (1989) Applied Logistic Regression (John Wiley).
6. Seber, G.E.F. and Wild, C.J. (1989) Nonlinear Regression (Wiley)
7. Neter, J., Wasserman, W., Kutner, M.H. (1985) Applied Linear Statistical Models. (Richard D. Irwin).
8. Rao.C.R(1973).:Linear statistical Inference and its application.
9. Goon, A.M., Gupta, M.K. and Das Gupta, B. (1967): An Outline of Statistical Theory.

Course Code	STA 407
Course Name	Linear Models
Credits	04

CLOs – PLOs Mapping matrix for STA 407: Linear Models

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	3	3	2	1
CLO-2	3	3	3	2	1
CLO-3	3	3	3	2	1
CLO-4	2	2	3	2	1
CLO-5	2	2	3	2	1
Total	13	13	15	10	5
Average	2.6	2.6	3	2	1

Course Code	STA 408
Course Name	Stochastic Process
Credits	04

Objective:
The main objective of the paper is to provide theoretical foundations of Stochastic Processes and to introduce different Stochastic/Random Processes and their applications.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

CLO-1: Differentiate between various types of stochastic processes.
CLO-2: Understand Markovian property.
CLO-3: Evaluate higher transition probabilities.
CLO-4: Apply stochastics models in business problems.
CLO-5: Recognize applications of branching, Brownian and renewal processes.

Unit-1

Definition and examples of stochastic process: Classification of general stochastic processes into discrete/continuous time, discrete/continuous state spaces, elementary problems, Random walk and Gambler's ruin problems, Counting process.

Unit-2

Markov chains: Definition and examples of Markov Chain, Transition probability matrix, classification of states, communicating classes, recurrence: non-recurrence, Irreducibility, Stationary distribution and its interpretation. Chapman-Kolmogorov equation, Stationary probability distribution and its applications. Computation of n-step transition probability matrix by spectral representation. Absorption probability and mean time to absorption.

Unit-3

Continuous time Markov Chain: Poisson process and related inter-arrival time distribution, compound Poisson process, Pure birth process, pure death process, birth and death process, problems, Renewal processes, Elementary renewal theorem (statement only) and its applications.

Unit-4

Galton -Watson branching processes: Definition and examples of discrete time branching process, Probability generating function and its properties, Offspring mean and probability of extinction. Introduction to Brownian motion process and its basic properties.

References

1. Kulkarni, Vidyadhar: Modeling and Analysis of Stochastic systems, G. Thomson Science and Professional.
2. Bhat, B.R.:. Stochastic Models: Analysis and Applications, (2nd New Age International, India).
3. Medhi J. : Stochastic processes, new Age International (P) Ltd.
4. Karlin S. and Taylor H.M. : A First Course in Stochastic Process, Academic Press
5. Hoel P.G., Port S.C. and Stone C.J.: Introduction to Stochastic Process, Universal Book Stall.
6. Parzen E. : Stochastic Process, Holden-Day
7. Cinlar E. Introduction to Stochastic Processes, Prentice Hall.
8. Adke S.R. and ManjunathS.M.:An Introduction to Finite Markov Processes, Wiley Eastern.
9. Ross S.M.: Stochastic Process, John Wiley.
10. John G. Kemeny, J. Laurie Snell, Anthony W. Knapp: Denumerable Markov Chains.

Course Code	STA 408
Course Name	Stochastic Process
Credits	04

CLOs – PLOs Mapping matrix for STA 408: Stochastic Process

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	2	2	1	1
CLO-2	2	2	2	1	1
CLO-3	2	3	2	1	1
CLO-4	3	3	3	2	1
CLO-5	3	2	3	2	1
Total	12	12	12	7	5
Average	2.4	2.4	2.4	1.4	1

Course Code	STA 409
Course Name	Design of experiments
Credits	04

Objective:
The main objective is to provide the theoretical foundations for design and analysis of experiments.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

CLO-1: Understand basic principles of design of experiments.
CLO-2: Choose appropriate design for different experimental studies.
CLO-3: Analyze data by performing analysis of variance/covariance procedures.
CLO-4: Validate statistical models using different diagnostic tools.
CLO-5: Recognize need of confounding in factorial experiments.

Unit-1

Basic principle of experimental design, overview of RBD, CRD and LSD, Missing plot techniques in RBD with one and two missing observations, Analysis of LSD with one missing observation.

Unit-2

General theory of intra block analysis of block design, connectedness and balancing block design, incomplete block design, intra block analysis of BIBD and its properties.

Unit-3

Purpose of analysis of covariance. Practical situations where analysis of covariance is applicable. Model for analysis of covariance in CRD and RBD. Estimation of parameters (derivations are not expected).Preparation of analysis of covariance (ANOCOVA) table, test for $\beta = 0$, test for equality of treatment effects (computational technique only).

Unit-4

General description of factorial experiments, factorial effects, analysis of factorial experiment (2^n, 3^n), main and interaction effects, advantages and disadvantages, total and partial confounding, split plot experiment.

References

1. Goon, Gupta, Dasgupta: Fundamental of Statistics, Vol. I and II, The World Press Pvt. Ltd. Kolkata.
2. Montgomery, D.C.: Design and Analysis of Experiments, Wiley Eastern Ltd., New Delhi.
3. Cochran, W.G. and Cox, G.M.: Experimental Design, John Wiley and Sons, Inc., New York.
4. Gupta, S.C. and Kapoor, V.K. : Fundamentals of Applied Statistics, S. Chand & Sons, New Delhi.
5. Das, M.N. and Giri, N.C. : Design and Analysis of Experiments, Wiley Eastern Ltd., New Delhi.
6. Joshi, D. D.: Linear estimation and design of experiment.
7. Dey, Alok: Theory of block designs, Wiley Eastern.

Course Code	STA 409
Course Name	Design of experiments
Credits	04

CLOs – PLOs Mapping matrix for STA 409: Design of experiments

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	2	2	2	1
CLO-2	3	3	3	2	1
CLO-3	3	3	3	2	1
CLO-4	3	3	3	2	1
CLO-5	1	2	2	1	1
Total	12	13	13	9	5
Average	2.4	2.6	2.6	1.8	1

Course Code	**STA 410**
Course Name	**Practicals**
Credits	**04 (0-0-4)**
Objective: **The main objective is to enhance the practical knowledge of an individual in statistical problem solving using Computer Software.**	
Learning Outcome: - **Learning to perform Statistical Computation using software.**	
CONTENT	**Practical based on MST 421-424**
Students will be required to do practicals using R-software based on Course IMST 421-424 . There shall be minimum four practical assignments from each course of the semester.	

THIRD SEMESTER

PAPER CODE	STA 501
PAPER NAME	Time Series Analysis & Forecasting
CREDIT	04

Objective:
The main purpose is to teach the time series modelling and the concept of forecasting and future planning.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

CLO-1: Recall basic concepts of time series analysis.
CLO-2: Recognize different components of time series.
CLO-3: Select appropriate time series model to analyze the data.
CLO-4: Validate the time series models using different diagnostic tools.
CLO-5: Estimate the forecast values.

Unit-1

Basics of Time series: A model Building strategy, Time series and Stochastic process, stationarity, Auto correlation, meaning and definition–causes of auto correlation-consequence of autocorrelation–test for auto–correlation. Study of Time Series model and their properties using correlogram, ACF and PACF. Yule walker equations.

Unit-2

Time Series Models: White noise Process, Random walk, MA, AR, ARMA and ARIMA models, Box- Jenkins's Methodology fitting of AR(1), AR(2), MA(1), MA(2) and ARIMA(1,1) process. Unit root hypothesis, Co-integration, Dicky Fuller test unit root test, augmented Dickey – Fuller test.

Unit-3

Non-linear time series models, ARCH and GARCH Process, order identification, estimation and diagnostic tests and forecasting. Study of ARCH (1) properties. GARCH (Conception only) process for modelling volatility.

Unit-4

Multivariate Time series: Introduction, Cross covariance and correlation matrices, testing of zero cross correlation and model representation. Basic idea of Stationary vector Autoregressive Time Series with orders one: Model Structure, Granger Causality, stationarity condition, Estimation, Model checking.

References

1. Box, G. E. P. and Jenkins, G. M.: Time Series Analysis – Forecasting and Control, Holden – day, San Francisco.
2. Chatfield, C.: Analysis of Time Series, An Introduction, CRC Press.
3. Ruey S. Tsay :Analysis of Financial Time Series, Second Ed. Wiley& Sons.
4. Ruey S. Tsay :Multivariate Time series Analysis: with R and Financial Application, Wiley& Sons.
5. Montgeomery, D. C. and Johnson, L. A.:Forecasting and Time series Analysis, McGraw Hill.
6. Kendall, M. G. and Ord, J. K. :Time Series (Third edition), Edward Arnold.
7. Brockwell, P. J. and Davies, R. A. :Introduction to Time Series and Forecasting(second Edition – Indian Print). Springer.
8. Chatfield, C. :The Analysis of Time series: Theory and Practice. Fifth Ed. Chapman and Hall.
9. Hamilton Time Series Analysis
10. Jonathan, D. C. and Kung, S.C. :Time Series Analysis with R. Second Ed. Springer.

PAPER CODE	STA 501
PAPER NAME	Time Series Analysis & Forecasting
CREDIT	04

CLOs – PLOs Mapping matrix for STA 501: Time Series Analysis & Forecasting

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	1	3	2	1
CLO-2	2	1	3	2	1
CLO-3	3	3	3	2	1
CLO-4	3	3	3	2	1
CLO-5	2	3	2	1	1
Total	12	11	14	9	5
Average	2.4	2.2	2.8	1.8	1

PAPER CODE	STA 502
PAPER NAME	Multivariate Analysis
CREDIT	04

Objective:
The main objective is to introduce the concept of analysing multivariate data and to increase familiarity with the handling of multivariate data.

Learning Outcomes:

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

CLO-1: Understand structure of multivariate data.
CLO-2: Recognize different multivariate distributions.
CLO-3: Apply different data reduction techniques.
CLO-4: Implement statistical tests for testing different multivariate hypothesis problems.
CLO-5: Perform classification of multivariate data.

Unit-1

Concept of random vector and random matrix. Multivariate distribution function and marginal and conditional distribution. Review of Multivariate Normal Distribution (MVND) and its properties.
Distribution of sample mean vector and its independence. Estimation of parameters of MVND.
Multiple linear equations, Multiple correlation, partial correlation in multiple setup and Distribution of sample multiple and partial correlation in null case. Partial and multiple correlation coefficients, their maximum likelihood estimators (MLE).

Unit-2

Wishart distribution and its properties.Hotelling's T^2 and its applications.
Hotelling's T^2statistic as a generalization of square of Student's statistic. Distance between two populations, Mahalnobis D^2 statistic and its relation with Hotelling's T^2 statistic.

Unit-3

- Classification problem, discriminant analysis.
- Principle component analysis.
- Canonical correlation.

Unit-4

- Factor Analysis.
- Cluster Analysis

References

1. Kshirsagar A. M. : Multivariate Analysis. Maral-Dekker.
2. Johnosn, R.A. and Wichern. D.W.:Applied multivariate Analysis. 5thAd.Prentice –Hall.
3. Anderson T. W.: An introduction to Multivariate statistical Analysis2nd Ed. John Wiely.
4. Morrison D.F.: Multivariate Statistical Methods McGraw-Hill.
5. Giri, N. C. (2014). Multivariate statistical inference. Academic Press.

PAPER CODE	STA 502
PAPER NAME	Multivariate Analysis
CREDIT	04

CLOs – PLOs Mapping matrix for STA 502: Multivariate Analysis

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	2	3	2	1
CLO-2	3	2	3	2	1
CLO-3	3	3	3	2	2
CLO-4	3	3	3	3	2
CLO-5	3	3	2	3	2
Total	15	13	14	12	8
Average	3	2.6	2.8	2.4	1.6

Course Code	STA 503
Course Name	Practicals
Credits	04

Objective:
The main objective is to enhance the practical knowledge of an individual in statistical problem solving using Computer Software.

Learning Outcome:

- Learning to perform Statistical Computation using software.

CONTENT	Practicals

Students will be required to do practicals using R-software based on opted theory papers

1. Select a series and obtain Mean, Variance and auto covariance autocorrelationupto lag 5.
2. Compute and plot the empirical autocovariance function and the empirical autocorrelation
3. Generate and plot AR(3)-processes (Yt), t = 1, . . . , 500 where the roots of the characteristic polynomial have the following properties: (i) all roots are outside the unit disk, (ii) all roots are inside the unit disk, (iii) all roots are on the unit circle, (iv) two roots are outside, one root inside the unit disk, (v) one root is outside, one root is inside the unit disk and one root is on the unit circle, (vi) all roots are outside the unit disk but close to the unit circle.
4. Fit a time series using Box-Jenkins Methodology.
5. Establish Yule Walker equations of order 5.
6. Take a GDP series and test the unit root hypothesis using DF and ADF test.
7. Obtain the autocorrelation and cross correlation for a multivariate time series.
8. Sketch of posterior distribution with informative and non-informative priors.
9. Bayes estimation of parametric family of distributions.
10. Posterior predictive distribution.
11. Monte Carlo integration.
12. Acceptance reject method.

ELECTIVES offered in III-SEMESTER

Course Code	STA 522
Course Name	Data Mining (Elective)
Credits	04

Objective:
The main objective of this course is to introduce theoretical foundations of develop algorithms, and methods of deriving valuable insights from data which includes detection and identification of outliers and anomalies, understanding the sequential and temporal patterns.

Learning Outcome:

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

- CLO-1: Understand approach of data mining.
- CLO-2: Recognize the competency in the use of data mining to the decision-support level of organizations.
- CLO-3: Apply different data mining techniques to address real life problems.
- CLO-4: Design and Implement data-mining solutions for different applications.
- CLO-5: Proficiency in evaluating and comparing different models used for Data Mining.

Unit 1

Data Mining: Introduction, Techniques, Issues and challenges, applications, Data preprocessing, Knowledge representation
Association Rule Mining: Introduction, Methods to discover association rules, Association rules with item constraints

Unit 2

Decision Trees: Introduction, Tree construction principle, Decision tree construction algorithm, Pruning techniques, Integration of pruning and construction

Unit 3

Cluster analysis: Introduction, clustering paradigms, Similarity and distance, Density, Characteristics of clustering algorithms, Center based clustering techniques, Hierarchical clustering, Density based clustering, Other clustering techniques, Scalable clustering algorithms, Cluster evaluation

Rough set theory, use of rough set theory for classification & feature selection.
ROC Curves: Introduction, ROC Space, Curves, Efficient generation of Curves, Area under ROC Curve, Averaging ROC curves, Applications

Unit 4

Advanced techniques: Web mining - Introduction, Web content mining, Web structure mining, Web usage mining; Text mining- Unstructured text, Episode rule discovery from text, Text clustering; Temporal data mining – Temporal association rules, Sequence mining, Episode discovery, time series analysis; Spatial data mining – Spatial mining tasks, Spatial clustering, Spatial trends.

References

1. Data Mining Techniques: A.K. Pujari, Universities Press, 2001
2. Mastering Data Mining: M. Berry and G. Linoff, John Wiley & Sons., 2000

Course Code	STA 522
Course Name	Data Mining (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 522: Data Mining

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	1	1	2	1
CLO-2	3	1	2	2	1
CLO-3	3	2	3	2	1
CLO-4	3	2	3	3	1
CLO-5	2	1	2	3	1
Total	**14**	**7**	**11**	**12**	**5**
Average	**2.8**	**1.4**	**2.2**	**2.4**	**1**

Course Code	STA 523
Course Name	National Development Statistics(Elective)
Credits	04

Objective:
The main objective is to make individual understand the significance and role of statistics in national development.

Learning Outcome:

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

- CLO-1: Understand the concept of Economic development parameters.
- CLO-2: Differentiate the population growth of develop and developing countries.
- CLO-3: Apply different techniques of poverty measurement.
- CLO-4: Know various institutions responsible for the collection of data in India.
- CLO-5: Understand various issues in the measurement of poverty.

Unit-1

Economic development: Growth in per capital income and distributive justice, Indices of development, Human Development index, quality of life. Estimation of national income-product approach, income approach and expenditure approach.

Unit-2

Population growth in developing and developed countries, Population projection using Leslie matrix, Labour force projection

Unit-3

Poverty measurement-different issues, measures of incidence and intensity, combined measures e.q. indices due to Kakwani, Sen etc.

Unit-4

MOSPI- Statistical System of India: NSSO, CSO, NSSTA, NITI Ayoge, Different Institutions and committees are responsible for planning and execution of National Building.

References

1. Chatterjee, S.K.: Quality of life.
2. Chaubey, P. K.: Poverty Analysis, New Age International (P) Limited, Publishers. New Delhi.
3. Human Development Annual Report.
4. Sen, Amartya.: Poverty and Famines, Oxford University Press.
5. CSO. National Accounts Statistics- Sources and Health.
6. UNESCO: Principles of Vital Statistics Systems.

Course Code	STA 523
Course Name	National Development Statistics(Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 523: National Development Statistics

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	2	1	2	1
CLO-2	2	3	1	2	1
CLO-3	3	3	3	2	1
CLO-4	2	2	1	2	1
CLO-5	2	2	1	2	1
Total	**12**	**12**	**7**	**10**	**5**
Average	**2.4**	**2.4**	**1.4**	**2**	**1**

Course Code	STA 525
Course Name	Principles & Practice of Insurance (Elective)
Credits	04

Objective:
The main objective is to introduce the basics and concepts of insurance.

Learning Outcome:

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

- CLO-1: Understand the present status of insurance.
- CLO-2: Classify in between life and non-life insurance.
- CLO-3: Know different types of investing and saving schemes in various funds.
- CLO-4: Recollect the concepts from actuarial science.
- CLO-5: Understand the function of regulatory bodies like IRDA.

Unit-1

Origin, Development and Present Status of Insurance, Risk Management, List out the Benefit and Cost of Insurance, Fundamental Key Principles of Insurance, Types of Insurance Contracts, Classification of Insurance.

Unit-2

Classification of insurance in life and non-life insurance, micro insurance, social insurance and general insurance (motor, marine, fire, miscellaneous), Types of insurance plans: whole life, term, endowment.

Unit-3

Types of investments and saving, Insurance, Shares, Bonds, Annuities, Mutual and Pension Fund.

Unit-4

Basics of Under-writing, Claims Management, Reinsurance, Legal and Regulatory Aspects of Insurance.
Seminar/Assignments: Each student will have to prepare his/ her presentation/ making assignments based on any topic from Actuarial Science and presents it. The topics will cover cases studies covering various aspects of the principles of insurance including IRDA regulations, publications, the 1938 Act 2006 and accounting standards.

References

1. Principles and Practice if Life Insurance, ICAI, New Delhi
2. Black & Skipper: Life and Health Insurance, Pearson Education
3. Harrington, Scott E. & Gregory R. : Risk Management and Insurance: 2nd ed., Tata McGraw Hill Publicating Company Ltd. New Delhi

Course Code	STA 525
Course Name	Principles & Practice of Insurance (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 525: Principles & Practice of Insurance

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	1	2	1	2
CLO-2	2	2	2	1	2
CLO-3	2	2	3	1	2
CLO-4	3	1	2	1	2
CLO-5	3	1	2	2	2
Total	**12**	**7**	**11**	**6**	**10**
Average	**2.4**	**1.4**	**2.2**	**1.2**	**2**

Course Code	STA 527
Course Name	Statistical Quality Control (Elective)
Credits	04

Objective:
The main purpose of this paper is to introduce the most important field of applied statistics that contributes to quality control in almost all industries.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

- CLO-1: Understand the general theory and review of control charts.
- CLO-2: Apply statistical process control in auto-correlated process data.
- CLO-3: Perform various sampling plans to reduce consumer and producer risks.
- CLO-4: Recognize the need of multiple sampling plans.
- CLO-5: Apply the SQC techniques to analyse the industrial data.

Unit-1

General theory and review of control chart for attributes and variables, OC and ARL of control chart, Statistical process control short production runs, Modified and acceptance control charts.

Unit-2

Statistical process control with auto-correlated process data, Adaptive sampling procedures, Economic design of control chart, CUSUM charts, Control charts in health care monitoring and Public health surveillance.

Unit-3

Producer's risk, Consumer's risk, Acceptance sampling plan, Single and double sampling plans by attributes, OC, ASN (and ATI), LTPD, AOQ and AOQL curves, Single sampling plan for variables (one sided specification, known and unknown cases), use of IS plans and tables.

Unit-4

Multiple sampling plans, Sequential sampling plan, The Dodge-Roaming sampling plan, Designing a variables sampling plan with a specified OC curve, Other variables sampling procedures. Continuous sampling

References

1. D.C. Montgomery: Introduction to Statistical Quality Control. Wiley.
2. Wetherill, G.B. Brown, D.W.: Statistical Process Control Theory and Practice, Chapman & Hall.
3. Wetherill, G.B.: Sampling Inspection and Quality control, Halsteed Press.
4. Duncan A.J.: Quality Control and Industrial Statistics, IV Ed., Taraporewala and Sons.
5. Ott, E. R. : Process Quality Control (McGraw Hill)

Course Code	STA 527
Course Name	Statistical Quality Control (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 527: Statistical Quality Control

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	1	2	2	2
CLO-2	2	2	2	2	2
CLO-3	2	2	2	2	2
CLO-4	2	1	3	2	2
CLO-5	2	3	3	3	2
Total	**10**	**9**	**12**	**11**	**10**
Average	**2**	**1.8**	**2.4**	**2.2**	**2**

Course Code	STA 528
Course Name	Survival Analysis (Elective)
Credits	04

Objective:
The main objective of this paper is to introduce different concepts and applications of survival analysis.

Learning Outcome:

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

- CLO-1: Understand the need of life time distributions and their properties.
- CLO-2: Identify the different type of censoring.
- CLO-3: Estimate parameters in presence of censoring.
- CLO-4: Implement different parametric and nonparametric estimators for estimating survival function.
- CLO-5: Analyse the lifetime event data.

Unit-1

Survival Characteristics and Parametric Models: Survival function, quantiles, hazard rate, cumulative hazard function, and mean residual life, Parametric models for study of event time data: Exponential, Weibull, extreme value, gamma, Pareto, logistic, log-logistic, normal, log–normal and mixture models -their survival characteristics.

Parametric Inference: Longitudinal studies. Censoring mechanisms- type I, type II and left right and interval censoring. Likelihood function under censoring and estimation. Tests based on LR, MLE.

Unit-2

Nonparametric Inference: Actuarial and Kaplan–Meier estimators. Treatment of ties. Self-consistency property and asymptotic properties of K–M estimator (statement). Pointwise confidence interval for S(t). Nelson-Aalen estimator of cumulative hazard function and estimation of S(t) based on it. Two–sample methods. Comparison of survival functions: Log rank and Tarone-Ware tests.

Unit-3

Semi-parametric Inference: Explanatory variables- factors and variates. Cox proportional hazards model. The partial likelihood and estimation of regression coefficients and their standard errors. Breslow's estimator, Statement of asymptotic properties of the estimator. Confidence interval for regression coefficients. Wald, Rao and likelihood tests for β. Accelerated life model. Model selection criteria and comparison of nested models (-2logL, AIC, BIC). Using information on prognostic variables in a competing risks model.

Unit-4

Concept of frailty. Shared frailty models. Identifiability of frailty models. Various frailty models. Gamma, positive stable, inverse Gaussian, power variance function, compound Poisson and compound negative binomial shared frailty models. Frailty regression models. Bivariate and correlated frailty models. Additive frailty models. Reversed hazard rates, Cox's proportional reversed hazards model.

References

Books Recommended

1. Cox, D.R. and Oakes, D. (1984). Analysis of Survival Data, Chapman and Hall.
2. Deshpande, J.V. and Purohit S.G. (2005). Life Time Data: Statistical Models and Methods, Word Scientific.
3. Duchateau, L. and Johnson, P. (2008). The Frailty Model. Springer: New York.
4. Gross A.J. and Clark, V. A. (1975) Survival Distributions: Reliability Applications in the Biomedical Sciences, John Wiley and Sons.
5. Hanagal, D. D. (2011). Modeling Survival Data Using Frailty Models. CRC Press: New York.
6. Hougaard, P. (2000). Analysis of Multivariate Survival Data. Springer: New York.
7. Wienke, A. (2011). Frailty Models in Survival Analysis, CRC Press: New York.

Course Code	STA 528
Course Name	Survival Analysis (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 528: Survival Analysis

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	1	3	1	2
CLO-2	3	2	3	1	2
CLO-3	2	2	3	2	2
CLO-4	2	1	2	2	2
CLO-5	3	2	3	2	3
Total	**13**	**8**	**14**	**8**	**11**
Average	**2.6**	**1.6**	**2.8**	**1.6**	**2.2**

PAPER CODE	STA 529
PAPER NAME	Statistical Methods for Bio-Computing (Elective)
CREDIT	04

Objective:
The use of statistical methods and tools from applied probability to address problems in computational biology.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

- CLO-1: Understand the need of molecular and morphological data.
- CLO-2: Identify the alignment of biological sequences.
- CLO-3: Estimate a good distance function.
- CLO-4: Apply simulation techniques for the modelling of biological sequence.
- CLO-5: Design various clustering algorithms.

Unit-1	Lectures: 11

Type of genetic data: - Molecular and morphological data. Differences and advantages of molecular data on, morphological data, Character data and distance data, their relative merits and demerits. Concept of entropy, entropy as a measure of uncertainty, entropy of single and combined scheme/s, Measure of information content based on entropy. Relative entropy its similarity with likelihood ratio. Applications of these to biological sequences.

Unit-2	Lectures:11

(Alignment of biological sequences): Pairwise and local alignment of biological Sequences (DNA/protein sequences). How biological sequences are different from mathematical sequences? The scoring matrices for alignment algorithms PAM and BLOSUM matrices. Algorithm for global alignment (Needleman Wunch algorithm). Local alignment algorithms (Smith - Waterman) Gap Model, dynamic programming algorithms for alignment with gaps such as linear gap model, affine gap model. Introduction to heuristic alignment algorithms such as BLAST, FASTA..

Unit-3	Lectures: 11

Molecular phylogeny Analysis: Tree of life, gene and species tree. Distance based methods for reconstruction of phylogenetic tree such as UPGMA, weighted UPGMA, transformed distance method, nearest neighbor joining method. Comparison of trees generated using different distance function Requisites of a good distance function. Character based methods for molecular phylogeny, maximum likelihood method and maximum parsimony method. Assessing trees via bootstrap. Probabilistic approach to phylogeny. Probabilistic models of evolution, Felsenteins algorithm for likelihood computation. Juke Canter model and Kimura and other probabilistic models for evolution.

Unit-4	Lectures: 12

Applications of Markov and Hidden Markov models to biological sequence Analysis. Markov chain as a classifier, use of Markov chain Model for demarcation of a region in Biological sequence analysis. Application of these in genetic sequence analysis such as detection of CPG Island. Testing whether given stretch of sequence is coming from CPG Island (use of Markov model for discrimination) Markov model based classification clusterization, testing order of a Markov model, testing homogeneity of two Markov models, Use of these test to design clustering algorithm. Hidden Markov/chains. Difference between these and simple Markov chains. Analysis of Hidden Markov Models/chains. Verterb is algorithm, Forward and backward algorithm for hidden Markov model. Parameter estimation in hidden Markov model when path is known as well as unknown, BaumWelch algorithm.

References

1. Alexander Isaac: (2001). Introduction to Mathematical Methods Bioinformatics. Springer.
2. Durbin R., Eddy S. Krogh A. Michelson G. (1998). Biological Sequence Analysis, Cambridge University Press.
3. 3. Robin S., Rudolph F, Schboth S. (2003) DNA Words and models Statistics of Exceptional Words, Cambridge University Press.

PAPER CODE	STA 529
PAPER NAME	Statistical Methods for Bio-Computing (Elective)
CREDIT	04

CLOs – PLOs Mapping matrix for STA 529: Statistical Methods for Bio-Computing

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	3	3	2	1
CLO-2	2	3	3	2	1
CLO-3	2	3	3	3	1
CLO-4	2	2	2	3	1
CLO-5	2	3	3	3	2
Total	**10**	**14**	**14**	**13**	**6**
Average	**2**	**2.8**	**2.8**	**2.6**	**1.2**

PAPER CODE	STA 530
PAPER NAME	Computer Intensive Statistical Methods (Elective)
CREDIT	04

Objective:
The main objective of this paper is to make students understand computational intensive methods for doing statistical inference.

Learning Outcome:

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

- CLO-1: Understand the necessity of resampling methods.
- CLO-2: Recall missing data imputation techniques to impute missing values.
- CLO-3: Develop programing codes for computationally intensive methods.
- CLO-4: Evaluate different statistical algorithms using cross validation.
- CLO-5: Apply various Bayesian techniques to analyze real life problems.

Unit-1	Lectures: 11

Resampling Techniques: Re sampling paradigms, bias-variance trade-off. Bootstrap methods, estimation of sampling distribution, confidence interval, variance stabilizing transformation. Jackknife and cross-validation. Jackknife in sample surveys. Jackknife in regression under heteroscedasticity. Permutation tests.

Unit-2	Lectures:11

Missing Values and Imputations Techniques: Missing values and types of missingness, imputations methods for missing values, single and multiple imputations. EM Algorithm
and Applications: EM algorithm for incomplete data, EM algorithm for mixture models,EM algorithm for missing values, stochastic EM algorithm.

Unit-3	Lectures: 11

Smoothing techniques: Kernel estimators, nearest neighbor estimators, orthogonal and local polynomial estimators, wavelet estimators. Splines. Choice of bandwidth and other smoothing parameters.

Unit-4	Lectures: 12

Bayesian computing, Markov Chain Monte Carlo. Simulation using MCMC, Particle filtering, MCMC methods for missing values.

References

1. Buuren, Stef van (2012). Flexible Imputation of Missing Data. Chapman and Hall.
2. Chihara, L. and Hesterberg, T. (2011) Mathematical Statistics with Resampling andR. Wiley.
3. Davison, A.C. and Hinkley, D.V. (1997) Bootstrap methods and their Applications. Chapman and Hall.
4. Efron, B. and Tibshirani. R.J. (1994); An Introduction to the Bootstrap. Chapman
5. and Hall.
6. Christensen R, Johnson, W., Branscum A. and Fishman, G.S. (1996) Monte Carlo:Concepts, Algorithms, and Applications. Springer.
7. Gilks, W. R., Richardson, S., and Spiegelhalter, D. (eds.) (1995) Markov Chain MonteCarlo in Practice. Chapman and Hall.
8. Good, P. I. (2005) Resampling Methods: A Practical Guide to Data Analysis. BirkhauserBosel.
9. Hanson T. E. (2011). Bayesian Ideas and Data Analysis: An Introduction for Scientistsand Statisticians, Chapman Hall.
10. Jim, A. (2009). Bayesian Computation with R, 2nd Edn, Springer.
11. Kennedy W. J. Gentle J. E. (1980) Statistical computing. Marcel Dekker.
12. McLachlan, G.J. and Krishnan, T. (2008) The EM Algorithms and Extensions. Wiley.
13. Rubinstein, R.Y. (1981); Simulation and the Monte Carlo Method. Wiley.
14. Shao J. and Tu, D. (1995); The Jackknife and the Bootstrap. Springer Verlag.
15. 14. Tanner, M.A. (1996); Tools for Statistical Inference, Third edition. Springer.

PAPER CODE	STA 530
PAPER NAME	Computer Intensive Statistical Methods (Elective)
CREDIT	04

CLOs – PLOs Mapping matrix for STA 530: Computer Intensive Statistical Methods

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	3	3	2	1
CLO-2	3	3	3	2	1
CLO-3	3	3	3	1	1
CLO-4	3	2	3	1	1
CLO-5	3	3	3	2	1
Total	**15**	**14**	**15**	**8**	**5**
Average	**3**	**2.8**	**3**	**1.6**	**1**

FORTH SEMESTER

PAPER CODE	STA 504
PAPER NAME	Practicals
CREDIT	02 (0-0-4)
Total hours	30
	CONTENT
Practical based on elective papers opt by the students. There shall be at least five practicals exercises covered from each of the courses.	

Course Code	STA 505
Course Name	Project
Credits	10
	Guidelines for project

- **Project duration:** Students may start preliminary work related to their project after second semester.
- **Project Guide:** Teachers from the Department of Statistics and/or organization where student is going to visit for field work or training. Each project group will be guided by concerned teacher (guide) for 8 hour per week throughout the IV semester.
- **Project Topic:** Students in consultation with the guide will decide project topic. The modification on the title may be permitted after the pre-presentation as advised during the seminar in consultation with the supervisor. Project work may be carried out in a group of students depending upon the depth of fieldwork/problem involved.
- **Project report:** Project report should be submitted in typed form with binding within the time as stipulated be the Department.
- **Project evaluation:** Project evaluation will be based on
  - (i) Continuous evaluation of the work – 25 Marks awarded by supervisor
  - (ii) Project report and final presentation - 25 marks awarded by supervisor
  - (iii) Viva-voce and final presentation - 50 marks awarded by external expert

ELECTIVES for IV-SEMESTER

Course Code	**STA 542**
Course Name	**Econometrics**
Credits	**04**

Objective:
The main objective is to introduce branch which is an integration of mathematics, statistics, and economics used to deal with econometric models.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

- CLO-1: Understand the properties and problems of econometric models.
- CLO-2: Recall various estimation and testing of hypothesis procedures in econometric models.
- CLO-3: Understand the concept of panel data models.
- CLO-4: Identify the fixed and random effect models.
- CLO-5: Apply Simultaneous Equation Models to analyse the economic data.

Unit-1

Introduction of Econometrics, Multiple Linear Regression Model, Model with non-spherical disturbances, Test of Auto-correlation, restricted regression estimator, Errors in variables, Dummy variables, Logit and Probit Models

Unit-2

Seemingly unrelated regression equation (SURE) model and its Estimation, Simultaneous equations model, concept of structural and reduced forms problem of identification, rank and order condition of identifiability.

Unit-3

Methods of estimation of simultaneous equation model: indirect least squares, two stage least squares and limited information maximum likelihood estimation, idea of three stage least squares and full information maximum likelihood estimation, and prediction

Unit-4

Panel data models: Estimation in fixed and random effect models, Panel data unit root test

References

1. Apte, P.G.: Text books of Econometrics, Tata McGraw Hill.
2. Gujarathi, D.: Basic Econometrics; McGraw Hill.
3. Johnston, J.: Econometrics Methods. Third edition, McGraw Hill.
4. Srivastava, V.K. and Giles D. A. E.: Seemingly unrelated regression equations models, Marcel Dekker.
5. Ullah, A. and Vinod, H.D.: Recent advances in Regression Methods, Marcel Dekker.

Course Code	STA 542
Course Name	Econometrics
Credits	04

CLOs – PLOs Mapping matrix for STA 542: Econometrics

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	3	3	2	1
CLO-2	3	3	3	2	1
CLO-3	3	3	3	1	1
CLO-4	3	2	3	1	1
CLO-5	3	3	3	2	1
Total	**15**	**14**	**15**	**8**	**5**
Average	**3**	**2.8**	**3**	**1.6**	**1**

Course Code	STA 544
Course Name	Life & Health Insurance (Elective)
Credits	04

Objective:
The main objective of this paper is to make individuals aware about the mechanisms of life and health insurance.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

- CLO-1: Understand the different type of insurance.
- CLO-2: Recall the concept of conventional non-participating life insurance.
- CLO-3: Classify the insurance plans under different insurance schemes.
- CLO-4: Recollect the concepts from actuarial science.
- CLO-5: Understand the actuarial aspects of insurance plans.

Unit-1

Introduction to life and health insurance, various types of life and health insurance plans, available insurance policies in the Indian market

Unit-2

Conventional non-participating life insurance, Linked accumulating non-participating contracts , Non-linked Accumulating Non-participating Contracts Participating Life Insurance, Different Distribution Methods, Profit Distribution Strategies, With-profit polices, Dividends and Bonus Method

Unit-3

Health insurance data, pricing & reserving, Classification of group and individual insurance plan under life and health insurance, Social security schemes, Method of valuation, Analysis of surplus

Unit-4

The actuarial role in life office management: Introduction, product pricing, analysis of surplus, monitoring and uploading the assumptions in the control cycle. Further uses of models in Actuarial management.
Students are also expected to complete three assignments:

i. Each student is expected to write a brief report on an appropriate/ relevant real life problem related to life insurance/health insurance/ general insurance using statistical tools and techniques.
ii. Review one insurance existing policy in Indian market and advise change with comparative analysis.
iii. Review some case study reported to different insurance companies administrative or legal authorities of the University.

References

1. Black & Skipper: Life and health insurance, Pearson Education
2. Philip Booth et al.: Modern actuarial theory and practice, Second edition, Chapman and Hall/CRC

Course Code	STA 544
Course Name	Life & Health Insurance (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 544: Life & Health Insurance

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	2	2	1	1
CLO-2	2	2	2	1	1
CLO-3	2	1	2	1	1
CLO-4	2	1	2	1	1
CLO-5	2	2	2	1	1
Total	**10**	**8**	**10**	**5**	**5**
Average	**2**	**1.6**	**2**	**1**	**1**

Course Code	STA 546
Course Name	Statistical Quality Management (Elective)
Credits	04

Objective:
The main objective of this course is to understand the procedure which seeks to improve the quality of the output of a particular industrial process.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

- CLO-1: Understand the general theory and review of control charts.
- CLO-2: Apply statistical process control in auto-correlated process data.
- CLO-3: Perform various sampling plans to reduce consumer and producer risks.
- CLO-4: Recognize the need of multiple sampling plans.
- CLO-5: Use the various process capability indices to check the status of the process.

Unit-1

Moving average and exponentially weighted moving average charts, Cu-sum charts using V-masks and decision intervals. Economic design of $\bar{X}$–chart. Multivariate control charts.

Unit-2

Acceptance sampling plans for inspection by variables for two sided specifications. Millitary Standard 105E (ANSI/ASQC Z1.4, ISO 2859) plans.

Unit-3

Continuous Sampling plans of Dodge type and Wald-Wolfowitz type and their properties, Bulk and chain sampling plans, Bayesian sampling plans. Role of statistical techniques in quality management.

Unit-4

Process Capability Indices: their estimation, confidence intervals and test of hypotheses for normally distributed characteristics. Process capability analysis using control chart, Process capability analysis with attribute data. Gauge and Measurement System capability studies.

References

1. D.C. Montgomery: Introduction to Statistical Quality Control. Wiley.
2. Wetherill, G.B. Brown, D.W.: Statistical Process Control Theory and Practice, Chapman & Hall.
3. Wetherill, G.B.: Sampling Inspection and Quality control, Halsteed Press.
4. Duncan A.J.: Quality Control and Industrial Statistics, IV Edision, Taraporewala and Sons.
5. Ott, E. R.: Process Quality Control (McGraw Hill)

Course Code	STA 546
Course Name	Statistical Quality Management (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 546: Statistical Quality Management

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	1	2	2	1
CLO-2	2	2	2	2	1
CLO-3	2	2	2	2	1
CLO-4	2	1	3	2	1
CLO-5	2	3	3	3	1
Total	**10**	**9**	**12**	**11**	**5**
Average	**2**	**1.8**	**2.4**	**2.2**	**1**

Course Code	STA 548
Course Name	Machine Learning (Elective)
Credits	04

Objective:

- The objective is to familiarize the audience with some basic learning algorithms and techniques and their applications, as well as general questions related to analyzing and handling large data sets.
- Several libraries and data sets are publicly available, that will be used to illustrate the application of machine learning algorithms.
- The emphasis will be on machine learning algorithms and applications, with some broad explanation of the underlying principles.
- To develop the basic skills necessary to pursue research in machine learning.
- To develop the design and programming skills that will help you to build intelligent, adaptive artifacts.

Learning Outcome:

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

- CLO-1: Understand different types of learning methods.
- CLO-2: Recognize different prediction models.
- CLO-3: Apply different clustering algorithms to real life data.
- CLO-4: Compare different machine learning algorithms.
- CLO-5: Perform classification of massive data using appropriate machine learning algorithms.

Unit I

Basics: Introduction to Machine Learning - Different Forms of Learning
Classification: Classification tree, SVM, Instance Based Classification, LDA, Multiclass Classification.

Unit II

Clustering: Partitional Clustering - K-Means, K-Medoids, Hierarchical Clustering-Agglomerative, Divisive, Distance Measures, Density Based Clustering – DBscan, Spectral Clustering

Unit III

Ensemble Methods: Boosting - Adaboost, Gradient Boosting, Bagging - Simple Methods, Random Forest

Unit IV

Dimensionality Reduction: Multidimensional Scaling, and Manifold Learning
Reinforcement Learning: Q-Learning, Temporal Difference Learning

References

1. Pattern Recognition and Machine Learning. Christopher Bishop.
2. Machine Learning. Tom Mitchell.
3. Pattern Classification. R.O. Duda, P.E. Hart and D.G. Stork.
4. Data Mining: Tools and Techniques. Jiawei Han and Michelline Kamber.
5. Elements of Statistical Learning. Hastie, Tibshirani and Friedman. Springer.

Course Code	STA 548
Course Name	Machine Learning (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 548: Machine Learning

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	2	1	3	3	2
CLO-2	2	2	3	3	2
CLO-3	2	2	3	3	2
CLO-4	2	1	3	2	2
CLO-5	2	3	3	3	2
Total	**10**	**9**	**15**	**14**	**10**
Average	**2**	**1.8**	**3**	**2.8**	**2**

PAPER CODE	STA 550
PAPER NAME	Bayesian Inference (Elective)
CREDIT	04

Objective:
To know Bayesian approach to solve statistical decision problems and use Bayesian techniques for computation.

Learning Outcome:
Upon successful completion of this course, the student will be able to:

- CLO-1: Distinguish between frequentist and Bayesian approach.
- CLO-2: Employ prior information for analyse of real life data.
- CLO-3: Choose appropriate prior distributions.
- CLO-4: Compute Bayes estimates for the population parameters.
- CLO-5: Apply Bayesian theory in testing of hypothesis problems.

Unit-1

Basic elements of Statistical Decision Problem. Expected loss, decision rules (non-randomized and randomized). Overview of Classical and Bayesian Estimation. Advantage of Bayesian inference, Prior distribution, Posterior distribution, Subjective probability and its uses for determination of prior distribution. Importance of non-informative priors, improper priors, invariant priors. Conjugate priors, construction of conjugate families using sufficient statistics, hierarchical priors. Admissible and minimax rules and Bayes rules.

Unit-2

Point estimation, Concept of Loss functions, Bayes estimation under symmetric loss functions, Bayes credible intervals, highest posterior density intervals, testing of hypotheses. Comparison with classical procedures. Predictive inference. One- and two-sample predictive problems.

Unit-3 and 4

Bayesian approximation techniques: Normal approximation, T-K approximation, Monte-Carlo Integration, Accept-Reject Method, Idea of Markov chain Monte Carlo technique.

References

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

Course Code	STA 550
Course Name	Bayesian Inference (Elective)
Credits	04

CLOs – PLOs Mapping matrix for STA 550: Bayesian Inference

(1-low, 2-medium, 3-high)

PLO→ CLO ↓	PLO-1	PLO-2	PLO-3	PLO-4	PLO-5
CLO-1	3	1	3	2	1
CLO-2	2	3	3	2	1
CLO-3	3	2	3	2	1
CLO-4	2	3	3	1	1
CLO-5	1	3	3	2	1
Total	**11**	**12**	**15**	**9**	**5**
Average	**2.2**	**2.4**	**3**	**1.8**	**1**