M.Sc. Statistics Programme under CCSS
at the Department of Statistics, University of Calicut

(Under the Calicut University PG Regulations for the Choice-based Credit Semester System (CCSS)-2022 in the University Teaching Departments)

Programme Structure & Syllabi
(With effect from the academic year 2022-23 onwards)

Duration of programme: **Two years -** divided into **four semesters** of not less than **90** working days each.

Course Code	Type	Course Title	Credits
I SEMESTER(Total Credits: 20)			
STA1C01	Core	Mathematical Methods for Statistics – I	4
STA1C02	Core	Mathematical Methods for Statistics – II	4
STA1C03	Core	Probability Theory	4
STA1C04	Core	Distribution Theory	4
STA1C05	Core	Sampling Theory	4
STA1A01	Audit	Ability Enhancement Course	2(Not in CGPA)
II SEMESTER(Total Credits: 20)			
STA2C06	Core	Theory of Estimation	4
STA2C07	Core	Design and Analysis of Experiments	4
STA2C08	Core	Stochastic Processes	4
STA2C09	Core	Regression Methods	4
STA2C10	Core	Practical – I	4
STA2A02	Audit	Professional Competency Course	2(Not in CGPA)
III SEMESTER(Total Credits: 20)			
STA3C11	Core	Time Series Analysis	4
STA3C12	Core	Multivariate Analysis	4
STA3C13	Core	Testing of Statistical Hypotheses	4
- --	Open Elective	Elective-I (Offered by other departments)	4
STA3E --	Elective	Elective-II	4
IV SEMESTER(Total Credits: 20)			
STA4C14	Core	Project and Dissertation	8
STA4E--	Elective	Elective-III	4
STA4E--	Elective	Elective-IV	4
STA4C15	Core	Practical – II	4

Total Credits: **80** (Core courses-**56**, Project and Dissertation -**8** and Elective courses-**16**).
Elective –I is an open elective course offered by other departments; the courses Elective –II, Elective –III and Elective –IV shall be chosen from the following list of courses offered by the Department.

LIST OF ELECTIVES

Sl. No.	Course Title	Credits
E01	Models and Methods for Lifetime Data	4
E02	Operations Research	4
E03	Queueing Theory	4
E04	Analysis of Longitudinal Data	4
E05	Analysis of Clinical Trials	4
E06	Statistical Decision Theory	4
E07	Reliability Modelling	4
E08	Actuarial Statistics	4
E09	Statistical Quality Assurance	4
E10	Statistical Machine Learning	4
E11	Statistical Modelling and Data Mining Techniques	4
E12	Applied algorithms and Big Data Techniques	4
E13	Official Statistics	4
E14	Order Statistics	4
E15	Econometric Models	4
E16	Computer Oriented Statistical Methods	4
E17	Stochastic Finance	4

OPEN ELECTIVES: (For other P.G. Programmes under CCSS Scheme)

E18	Statistical Methods	4
E19	Topics in Probability	4

Question paper pattern:

For each course there shall be an external examination of duration **3 hours**. Each question paper will consists of two parts- **Part-A** consisting of **eight short answer type questions**, each of **2 marks**, in which **any five questions** are to be answered; **Part-B** consisting of **four essay type questions** each with two options A and B of **10 marks.** The candidate is required to answer all questions choosing **either Option-A or Option-B** of them. The questions are to be evenly distributed over the entire syllabus within each part.

MODEL QUESTION PAPER

I/II/III/IV SEMESTER M.Sc. DEGREE (CCSS) EXAMINATION, Month-Year

Branch: Statistics

Course Code: Course Title(___Credits)

Time: **3** hours Max. Marks: **50**

Section – A

(Answer any **FIVE**questions; each question carries **2**marks)

I (i) ….
(ii) …
(iii) …
(iv) …
(v) …
(vi) …
(vii) …
(viii) … (**5 x 2 = 10**)

Section – B

(Answer either Part-**A** or Part-**B** of all questions; each question carries **10** marks)

II A. a) …
b) …

(-+-)

OR

B. a) …
b) … **(-+-)**

III A. a) …
b) …

(-+-)

OR

B. a) …
b) …

(-+-)

IV A. a) …
b) …

(-+-)

OR

B. a) …
b) …

(-+-)

V A. a) …
b) …

(-+-)

OR

B. a) …
b) …

(-+-)

(4 x 10 = 40)

- - - - - - - - - - -

Objectives of the Programme

The present programme is intended to provide a platform for talented students to undergo higher studies in the subject as well as to train them to suit for the needs of the society. Apart from teaching core Statistics subjects, the students can choose electives depending upon their interests, under the choice based credit system. The students are also trained to handle real life problems through the practical classes and project work. As a part of the course the students are also exposed to various statistical software such as SPSS, MATLAB and R.

Programme Outcomes:

On successful completion of the programme, students will be able to:

PO-1: Gain sound knowledge in theoretical and practical aspects of Statistics;

PO-2: Acquire the knowledge on modern statistical techniques relevant for today's scientific community;

PO-3: Convince the need for systematic analysis of data in any scientific experiment;

PO-4: Acquire the working knowledge of various statistical softwares and programming languages;

PO-5: Acquire skills and competencies in statistical computing methods and develop algorithms and computer programmes for analyzing complex data sets;

PO-6: Communicate effectively complex statistical ideas to people working in diverse spheres of academics and organizational set ups;

PO-7: Handle and analyze large databases and make meaningful interpretations of the results;

PO-8: Become professionally inclined statistics teachers/statistician/data scientist who have sound knowledge of the subject matter and specialized in knowledge discovery through statistical methods;

PO-09: Acquire basic theoretical and applied principles of statistics with adequate preparation to pursue Doctoral (Ph.D.) degree or enter job force as an applied statistician;

PO-10: Make unique contribution for the development of discipline by addressing complex and challenging problems in emerging areas of the discipline;

PO-11: Continue to acquire relevant knowledge and skills appropriate to professional activities and demonstrate highest standards of ethical issues in Statistical Sciences;

PO-12: Create awareness to become an enlightened citizen with commitment to deliver one's responsibilities within the scope of bestowed rights and privileges.

SYLLABI OF CORE COURSES

STA1C01: Mathematical Methods for Statistics – I (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the concept of Riemann – Stieltjes Integral.
CO2: Evaluate the integrals of functions using Riemann-Stieltjes method.
CO2: Discuss the concepts of sequence and series of functions.
CO3: Illustrate the consequence of uniform convergence on continuity, integrability and differentiability.
CO4: Illustrate vector space, subspaces, independence of vectors, basis and dimension, with examples.
CO5: Solve systems of linear equations, diagonalize matrices and classify quadratic forms.
CO6: Compute g-inverse of matrices.
CO7: Compute algebraic and geometric multiplicity of characteristic roots.

Unit-I. Riemann – Stieltjes Integral - Definition, Linear properties.Integration by parts. Change of variable. Reduction to a Riemann integral. Step functions as integrators. Reduction to a finite sum. Monotonically increasing integrators. Riemanns conditions. Comparison theorems. Functions of bounded variations (concepts only). Sufficient conditions for the existence of Riemann Stieltjes integral. Necessary conditions for the existence of Riemann Stieltjes integral. Mean-value theorems.

Unit-II. Sequences and Series of Functions - Point wise convergence of sequence of functions. Examples of sequences of real valued functions. Definition of Uniform convergence. Uniform convergence and continuity. Cauchy condition for uniform convergence. Uniform convergence of infinite series of functions. Uniform convergence and Riemann-Stieltjes Integration. Uniform convergence and differentiation.

Unit-III. Algebra of vectors - Vector spaces (definition and examples). Subspaces. Linear independence. Basis and dimension. Linear equations. Vector spaces with an inner product: Properties. Gram-Schmidt orthogonalization. Orthogonal projection of a vector.

Unit-IV. Algebra of matrices- Theory of matrices and determinants: Matrix operations. Elementary matrices and diagonal reduction of a matrix. Determinants. Transformations. Generalized inverse of a matrix. Matrix representations of vector spaces, bases, etc. Idempotent matrices. Special products of matrices.
Eigen values and reduction of matrices: Classification and transformations of quadratic forms. Roots of determinantal equations. Canonical reduction of matrices. Projection operator.

Text Books

1. **Apostol, T.M.** (1974). Mathematical Analysis -Second Edition. Narosa Publishing House, New-Delhi.
2. **Rao, C.R. (2002).** Linear Statistical Inference & Its Applications- Second Edition. John Wiley & Sons, New York.
3. **Rao,A.R. and Bhimsankaram, P. (1992).** Linear Algebra.Tata McGraw Hill, New Delhi.
4. **Lewis, D.W. (1996).** Matrix Theory. Allied Publishers, Bangalore.
5. **Graybill, F. A. (1983).** Matrices with Applications in Statistics. John Wiley & Sons, New York.

References

1. **Malik, S.C. &Arora, S. (2006).** Mathematical Analysis- Second Edition. NewAge International, New Delhi.
2. **Beezer, R. A. (2004).** A First Course in Linear Algebra, Congruent Press, Washington.
3. **Bapat, R.B. (2011).** Linear Algebra and Linear Models. Springer and Hindustan Book Agency.

STA1C02: Mathematical Methods for Statistics –II (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Recollect basic concepts of set theory and explain the concepts of class of sets & set function.

CO2: Discuss measure and different types of measures such as Outer measure, Lebesgue measure, Lebesgue-Steiltjes measure together with their properties.

CO3: Describe Measurable function and related results.

CO4: Discuss the basic concepts of Product space and product measure.

CO5: Explain Multiple integral, Absolute continuity and singularity of measures.

CO6: Evaluate the limit and check the continuity of multivariable functions.

CO7: Compute the derivatives, partial and total derivatives and maxima and minima of multivariable functions.

CO8: Determine the derivatives and integrals of functions in the complex plane.

CO9: Determine the singularities and residues of complex functions.

Unit-I. Classes of Sets - Field of sets, sigma field, monotone class and minimal sigma field. Borel sigma field and Borelsets in R and Rp. Set functions. Additivityand sigma additivity – Measures - examples and properties. Outer measure. Lebesgue measure in R and Rp. Lebesgue-Steiltjes measure.

Unit-II. Measurable functions - Measurable sets and functions. Measures: Measure spaces. Almost everywhere.Charges. The integral: Simple functions and their integrals. The integral of a non-negative extended real valued measurable function. Monotone convergence theorem.Fatou's lemma.Properties of the integrals.Integrable real valued functions. Lebesgue dominated convergence theorem. Modes of convergence (Concepts only).Egoroff's theorem (without proof).

Product space and product measure. Multiple integral.Fubini's theorem (without proof). Absolute continuity and singularity of measures. Radon-Nikodym theorem (without proof) and its applications.

Unit-III. Multivariable functions - Limit and continuity of multivariable functions (Definitions and Problems). Derivatives. Directional derivatives and continuity. Total derivative in terms of partial derivatives. Taylor's theorem. Inverse and implicit functions. Optima of multivariable functions.

Unit-IV. Functions of complex variables - Limits. Continuity. Differentiability. Cauchy-Riemann (C-R) equations. Polar form of C-R equations. Analytic functions. Harmonic functions. Partial derivative in relation to $z$and $\bar{z}$. Milne Thomson method for constructiong analytic functions. Complex integration: Contour integral. Cauchy's theorem. Cauchy's integral formula. Cauchy's inequality. Liouville's theorem.Maximum modulus theorem. Taylor series and Laurent series (problems only). Classification of singularities. Residues. Residue theorem (statement and applications only).

Text Books

1. **Royden, H. L. (1995).** Real Analysis- Third Edition. Prentice Hall of India, New Delhi.
2. **Bartle, R.G. (1996).** The Elements of Integration. John Wiley & Sons, New York.
3. **Khuri, A. T. (1993).** Advanced Calculus with Applications in Statistics. John Wiley & Sons, New York.
4. **Kasana H (2001).** Complex variables: Theory and Applications. Prentice Hall of India.

References

1. **Kingman, J.F.C. and Taylor, S.J. (1973).** Introduction to Measure and Probability. Cambridge University Press, UK.
2. **Malik, S.C. &Arora, S. (2006).** Mathematical Analysis- Second Edition. NewAge International, New Delhi.

STA1C03: Probability Theory (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the concepts of probability measure, random variables and decomposition of distribution functions.
CO2: Describe induced probability space and notion of vector valued random variables.
CO3: Explain Expectation of simple, non-negative and arbitrary random variables.
CO4: Distinguish between different types of inequalities such as Cr-inequality, Basic inequality, etc.
CO5: Illustrate convergence of sequence of random variables with examples.
CO6: Discuss independence of events, Kolmogrov0-1 law and Borel 0-1 criterion.
CO7: Describe properties of Characteristic Function, inversion theorem and its applications.
CO8: Acquire knowledge about WLLN, SLLN, CLT and illustrate their applications.

Unit-I: Sets and classes of events - Events. Algebra of sets (set operations, sequences of sets and limits). Fields. Sigma fields. Minimal field. Partition. Borel field. Random variables: Functions and inverse functions. Limits of random variables: Sigma fields induced by random variables. Limits of random variables. Discrete Probability space. General Probability space. Induced probability space. Other measures (Concepts only).

Unit-II: Distribution functions of random variables. Decomposition of distribution functions. Distribution function of vector random variables. Correspondence theorem. Expectation and moments: Definition of expectation of simple, non-negative and arbitrary random variables. Properties of expectations. Moments and inequalities (Cr, Holder's, Jensen's, Basic and Markov inequalities).
Characteristic functions: Definition, Properties, Inversion formula (statement and problems only). Characteristic functions and moments. Bochner's theorem (Statement only),
Independence: Independence of events and classes of events. Independence of random variables. Multiplication properties. Kolmogorov 0-1 law. Borel 0-1 criterion. Borel Cantelli lemma.

Unit-III: Convergence of random variables: Convergence in probability. Convergence almost surely, Convergence in distribution, Convergence in r^{th} mean – their inter-relations- examples and counter examples. Convergence of distribution functions: Weak convergence, Helly-Bray lemma and Helly– Bray theorem.

Unit- IV: Laws of Large Numbers - Kolmogorov inequality, Kolmogorov three series theorem; Weak law of large numbers (both IID and Non-IID cases). Kolmogorov Strong Law of large numbers (Statement and problems only). Central Limit Theorem(CLT): Lindeberg-Levy theorem. Liapounov form of CLT. Lindeberg-Feller CLT (no proof required for these three CLT's). Association between Liapounov's condition and Lindeberg conditions. Simple applications of CLT.

Text Book

1. **Bhat, B.R. (1999).** Modern Probability Theory- Third Edition. New-Age International, New Delhi.

References

1. **Resnick, S.I. (1999).** Probability Paths. Birkhauser, Boston.
2. **Laha, R.G. and Rohatgi, V.K**. (1979). Probability Theory. John Wiley& Sons, New York.
3. **Billingsly, P. (1995).** Probability and Measure- Third Edition. John Wiley & Sons, New York.
4. **Basu, A.K. (1999).** Measure Theory and Probability. Printice Hall of India, New-Delhi.
5. **Rohatgi, V.K**. (1976). An Introduction to Probability Theory and Mathematical Statistics. John Wiley &Sons , New York.

STA1C04: Distribution Theory (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Distinguish different discrete distributions and illustrate their role in modeling count data.

CO2: Explain Pearson system and its different members and how they arise from the defining differential equation.

CO3: Describe different positive valued and real valued random variables along with their properties and role in modeling real life data.

CO4: Find marginal and conditional distributions, explain distribution of functions of random vectors, order statistics and their distributions.

CO5: Explain sampling distributions- t, Chisquare and F and their applications.

Unit-I. Discrete distributions - Bernoulli, Discrete Uniform, Binomial, Negative Binomial, Geometric, Hyper geometric, Poisson Logarithmic Series and multinomial distributions. Power series distribution and their properties.

Unit-II. Continuous distributions - Systems of Distributions- Pearson system and Transformed Distributions. Uniform, Exponential, Gamma, Beta, Cauchy, Normal, Pareto, Weibull, Laplace, lognormal distributions. Bivariate Normal Distributions and their properties.

Unit-III. Notion of vector of random variables. Distribution function marginal and joint distributions in the i.i.d. case. Functions of Random Vectors. Order Statistics and their Distributions.

Unit-IV. Sample moments and their distributions- Sample characteristics and their distributions, Chi-Square, t and F distributions (Central and Non-Central). Applications of Chi-square, t and F.

Text Books

1. **Rohatgi, V.K. (1976).** An Introduction to Probability Theory and Mathematical Statistics. John Wiley & Sons, New York. Chapter 4- Sections 2, 4 and 5; Chapter 5- Sections 2, 3 and 4; Chapter 7- Sections 3, 4 and 5.
2. **Krishnamoorthy, K. (2006).** Hand book of Statistical Distributions with Applications. Chapman & Hall, New York. Chapters- 8,14,20,23 and 24 Sections- 1,2 and 5.
3. **Johnson, N.L., Kotz. S. and Balakrishnan, N.(2004).** Continuous Univariate Distributions- Vol. I- Second Edition. John Wiley &Sons, New York. Chapter 12 -Sections 4.1, 4.3.

References

1. **Johnson, N.L., Kotz. S. and Balakrishnan,N**. (1995). Continuous Univariate Distributions- Vol. II- Second Edition. John Wiley & Sons, New York.
2. **Johnson, N.L., Kotz. S. and Kemp, A.W. (1992).** Univariate Discrete Distributions. John Wiley & Sons, New York.
3. **Kendall, M. and Stuart, A.** (1977). The Advanced Theory of Statistics- Vol.-I: Distribution Theory- Fourth Edition. Charles Griffin & Co. Ltd., USA.
4. **Goon, A.M. ,Gupta, M.K. and Das Gupta, B. (1991).** Fundamentals of Statistics- Vol. I and Vol. II (2001). World Press, Calcutta.

STA1C05: Sampling Theory (4 credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Distinguish between the concepts of probability and non-probability sampling.

CO2: Learn the principles underlying sampling as a means of making inferences about a population and the estimation methods for population mean, total and proportion under various sampling schemes.

CO3: Apply various sampling procedures like SRS, Stratified, systematic, Cluster etc., and estimate the population parameters for attributes and variables.

CO4: Describe the use of auxiliary information for the estimation of population characteristics using ratio, product and regression estimators.

CO5: Employ sampling strategies under varying probability sampling.

CO6: Explain various types of errors in surveys, and procedures to rectify them.

CO7: Analyse data from multi-stage and multiphase surveys.

CO8: Have an appreciation of the practical issues arising in sampling studies.

Unit-I. Census, Sampling, Probability sampling, and non-Probability sampling. SRSWOR and SRSWR. Estimation of population mean. Population total and population proportion.Variance of the estimates and standard error. Estimation of sample size. Stratified random sampling. Allocation problem. Various allocations. Construction of strata.

Unit-II. PPS sampling with and without replacement. Estimation of population mean, total and variance in PPS sampling with replacement. Desraj's ordered estimator. Murthy's unordered estimator. Horvitz – Thomson estimator. Their variances and standard error. Yates – Grundy estimator. Sen – Midzuno scheme of sampling. ΠPS sampling.

Unit-III. Ratio estimators and Regression estimators. Comparison with simple arithmetic mean estimator. Optimality properties of ratio and regression estimators. Hartly – Ross unbiased ratio type estimator.

Unit-IV. Circular, linear and balanced systematic sampling. Estimation of population mean and its variance. Cluster sampling with equal and unequal clusters. Multi stage and multiphase sampling. Comparison with simple random sampling and Stratified random sampling. Relative efficiency of cluster sampling. Two-stage sampling. Non-sampling errors.

Text Books

1. **Cochran, W.G. (1977).** Sampling Techniques. Wiley Eastern, New Delhi.
2. **Singh, D. and Chaudhury, F.S. (1986).** Theory and Analysis of Sample Survey Designs.Wiley Eastern, New Delhi.

References

1. **Des Raj (1976).** Sampling Theory. McGraw Hill, New York.
2. **Murthy, M. N. (1967).** Sampling Theory and Methods. Statistical Publishing Society, Calcutta.
3. **Mukhopadhyay, P. (1999).** Theory and Methods of Survey Sampling. Printice Hall India, New Delhi.

STA2C06: Theory of Estimation (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain Sufficiency, Minimal Sufficiency, Unbiasedness & BLUE.
CO2: Describe the ways of obtaining MUVE.
CO3: Check consistency of estimators and how to choose among consistent estimators.
CO4: Apply different methods of estimation such as method of percentiles, method of moments and method of maximum likelihood.
CO5: Describe various loss functions, Risk function and Bayesian estimation under squared error, absolute error and zero-one loss functions
CO6: Derive SELCI. Explain Bayesian and Fudicial intervals

Unit-I. Fisher Information- Sufficient statistic-Minimal sufficient statistic-Exponential family and minimal sufficient statistic.Unbiasedness–best Linear Unbiased estimator–MVUE – Cramer-Rao inequality and its application–Rao-Blackwell theorem-Completeness-Lehman-Scheffe theorem and its application.

Unit-II. Consistent estimator-Examples and properties-CAN estimator-invariance property-asymptotic variance-Multiparameter case-choosing between Consistent estimators.

Unit-III. Method of moments-method of percentiles-method of maximum likelihood- MLE in exponential family-Solution of likelihood equations-Bayesian method of estimation-Prior information-Loss functions (squared error absolute error and zero-one loss functions) – Posterior distribution-estimators under the above loss functions.

Unit-IV. Shortest expected length confidence interval-large sample confidence intervals-unbiased confidence intervals-examples-Bayesian and Fiducial intervals.

Text Books

1. **Kale, B.K. (2005).** A First Course on Parametric Inference- Second Edition, Narosa Publishing, New-Delhi.
2. **Casella, G. and Berger, R.L. (2002).** Statistical Inference- Second Edition. Duxbury, Australia.

References

1. **Rohatgi,V. K. (1976).** An Introduction to Probability Theory and Mathematical Statistics. John Wiley & Sons, New York.
2. **Rohatgi,V.K. (1984)**. Statistical Inference. John Wiley & Sons, New York.
3. **Lehman, E.L. (1983).** Theory of Point Estimation. John Wiley & Sons, New York
4. **Rao, C.R. (2002).** Linear Statistical Inference and Its Applications- Second Edition. John Wiley & Sons, New York.

STA2C07: Design and Analysis of Experiments (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain ANOVA, it'sregression approach, fixed and random effect models and their analysis.

CO2: Check model adequacy.

CO3: Explore designs like RBD, CRD, LSD, Greaco- LSD, BIBD, Youden square, Lattice design and Factorial designs.

CO4: Explain nested or hierarchical designs and response surfacemethods

CO5: Discuss ANCOVA.

Unit-I. Application, basic principles, guideline of design of experiments.Statistical techniques.Experiments with single factor.ANOVA. Analysis of fixed effect models – comparison of individual treatment means. Random effect models. Model adequacy checking. Choice of sample size.Regression approach to ANOVA.

Unit-II. Completely Randomized Block design. Randomized block design. Latin square design. Greaco- Latin square design.Incomplete block designs and approaches to analyzing them- BIBD – Recovering of intra block information in BIBD – PBIBD – Youden square – Lattice design.

Unit-III. Factorial designs – definition and principles. Two factor factorial design. Random and mixed models. The general factorial designs- 2^k-factorial experiments – confounding-two level fractional factorial design.

Unit-IV. Nested or hierarchical designs – response surface methods and design – ANCOVA

Text Books

1. **Montgomery, D.C. (2001).** Design and Analysis of Experiments- Fifth Edition. John Wiley & Sons, New York.

References

1. **Das, M. N. and Giri, N. C. (2002).** Design and Analysis of Experiments- SecondEdition. New Age International (P) Ltd., New Delhi.

STA2C08: Stochastic Processes (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the concepts of stochastic processes.
CO2: Classify stochastic processes.
CO3: Distinguish between strict stationarity and wide-sense stationary.
CO4: Describe the concepts of Discrete time Markov chains, classification of its states and limiting probabilities.
CO5: Describe and use the recurrence relation for generation sizes in a Branching Process and determine the probability of ultimate extinction.
CO6: Explain continuous time Markov chains, Poisson processes and its generalizations.
CO7: Employ birth-death methodology for solving queueing problems.
CO8: Explain the renewal processes and Brownian motion processes.

Unit-I. Concept of Stochastic processes, examples. Specifications. Markov chains- Chapman Kolmogorov equations – classification of states – mean time spent in transient states -limiting probabilities - Gamblers ruin problem— Discrete parameter branching processes - Hidden Markov chains.

Unit-II. Exponential distribution– counting process – inter arrival time and waiting time distributions. Properties of Poisson processes – Conditional distribution of arrival times. Generalization of Poisson processes – non –homogenous Poisson process, compound Poisson process, conditional mixed Poisson process. Continuous time Markov Chains – Birth and death processes – transition probability function-limiting probabilities.

Unit-III. Renewal processes- Renewal function- Elementary Renewal theorem and its applications. Renewal reward process. Regenerative processes, semi-Markov process. The inspection paradox- Insurers ruin problem.

Unit-IV. Basic characteristics of queues – Markovian models – network of queues. The M/G/I system. The G/M/I model, Multi server queues. Brownian motion process – hitting time – Maximum variable – variations on Brownian motion – Pricing stock options – Gaussian processes – stationary and weakly stationary processes.

Text Books

1. **Ross, S.M.(2007).** Introduction to Probability Models- Ninth Edition. Academic Press, New York.

References

1. **Medhi,J. (1996).** Stochastic Processes- Second Edition. Wiley Eastern, New Delhi**.**
2. **Karlin, S. and Taylor, H.M.** (1975). A First Course in Stochastic Processes- Second Edition. AcademicPress, New York.
3. **Cinlar, E.** (1975). Introduction to Stochastic Processes. Prentice Hall, New Jersey.
4. **Basu, A.K**.(2003). Introduction to Stochastic Processes. Narosa Publishing House, New Delhi.

STA2C09: Regression Methods (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Illustrate the concept of linear regression model.
CO2: Estimate regression parameters and explain the properties of estimators.
CO3: Describe the procedure for testing the significance of regression parameters and construct confidence intervals.
CO4: Build various Non parametric Regression models.
CO5: Examine model diagnostics techniques and remedies to overcome violating assumptions.

Unit-I. Least square estimation-properties of least square estimates-unbiased estimation of σ^2–distribution theory – maximum likelihood estimation – estimation with linear restrictions- design matrix of less that full rank-generalized least squares.

Unit-II. Hypothesis testing; Likelihood ratio test—F-test – multiple correlation coefficient-Confidence intervals and regions. Simultaneous interval estimation- confidence bands for the regression surface – prediction intervals and band for the response.

Unit-III. The straight line – weighted least squares for the straight line- Polynomials in one variable – piecewise polynomial fitting – Polynomial regression in several variables.

Unit-IV. Bias-incorrect variance matrix-effect of outliers-Diagnosis and remedies: residuals and hat matrix diagonals – nonconstant variance and serial Correlations-departures from normality – detecting and dealing with outliers- diagnosing collinearity, Ridge regression and principal component regression.

Text Books

1. **Seber, G. A. F. and Lee, A.J. (2003).** Linear Regression Analysis- Second Edition. John Wiley & Sons, New York.
2. **Draper, N.R. and Smith, H. (1988).** Applied Regression Analysis- Third Edition. John Wiley &Sons, New York.

References

1. **Searle, S.R.** (1997). Linear Models. Wiley Paperback Edition. Wiley Inter Science, New Jersey.
2. **Rao, C.R.**(1973). Linear Statistical Inference and Its Applications. Wiley Eastern, New Delhi.
3. **Abraham, B. and Ledolter, J.** (2005). Introduction to Regression Modeling. Duxbury Press, New York.
4. **Sengupta, D. and Jammalamadaka, S.R.** (2003). Linear Models: An Integrated Approach. World Scientific Press, New Jersey.
5. **Montgomery, D.C., Peck, F.A. and Vining, G**. (2001). Introduction to Linear Regression Analysis- Third Edition. John Wiley & Sons, New York.

STA2C10: Practical – I (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Apply the principles of Distribution theory, Sampling theory, Theory of estimation Design & Analysis of experiments and Regression methods using real data sets.
CO2: Know the formulas to be applied for the analysis.
CO3: Write the R codes for the analysis of the given data.
CO4: To install and load the packages required to run the R codes.
CO5: Enter the data given for analysis.
CO6: Explain how to make conclusions and write the inference for the data analysis based on the output obtained.

The practical is based on the following core papers in the first and the second semesters:

1. STA1C04: Distribution Theory
2. STA1C05: Sampling Theory
3. STA2C06: Theory of Estimation
4. STA2C07: Design and Analysis of Experiments
5. STA2C09: Regression Methods

Practicals are to be done using the R-programming language. The End Semester practical examination shall be conducted and evaluated by two or more examiners nominated by the Department Council.

STA3C11: Time Series Analysis (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Describe the basics of time series data, its auto-covariance, auto-correlation and stationarity.
CO2: Illustrate test for trend and seasonality.
CO3: Explain the smoothing methods for determining trend of the data.
CO4: Describe the properties of linear time series models.
CO5: Fit linear models for time series data sets.
CO6: Describe the maximum likelihood, Yule-Walker and least square estimation methods.
CO7: Learn to validate a model using residual analysis.
CO8: Define ARCH and GARCH models and derive their properties.
CO9: Analyze spectral density and periodogram.

Unit-I. Motivation, Time series as a discrete parameter stochastic process, Auto – Covariance, Auto- Correlation and spectral density and their properties. Exploratory time series analysis, Test for trend and seasonality, Exponential and moving average smoothing, Holt – Winter smoothing, forecasting based on smoothing, Adaptive smoothing.

Unit-II. Detailed study of the stationary process: Autoregressive, Moving Average, Autoregressive Moving Average and Autoregressive Integrated Moving Average Models. Choice of AR MA periods.

Unit-III. Estimation of ARMA models: Yule – Walker estimation for AR Processes, Maximum likelihood and least squares estimation for ARMA Processes, Discussion (without proof) of estimation of mean, Auto-covariance and auto-correlation function under large samples theory, Residual analysis and diagnostic checking. Forecasting using ARIMA models.

Unit-IV. Spectral analysis of weakly stationary process.Herglotzic Theorem. Periodogram and correlogram analysis. Introduction to no-linear time Series: ARCH and GARCH models.

Text Books

1. **Box G.E.P and Jenkins G.M.** (1970)**.** Time Series Analysis, Forecasting and Control. Holden-Day, SanFransisco.
2. **BrockwellP.J.and Davis R.A.** (1987). Time Series: Theory and Methods. Springer, New York.
3. **Abraham B and Ledolter J.C.** (1983). Statistical Methods for Forecasting, John Wiely& Sons, New York.

References

1. **Anderson T.W** (1971). Statistical Analysis of Time Series. John Wiley & Sons, New York.
2. **Fuller W.A.** (1978). Introduction to Statistical Time Series. John Wiley & Sons, New York.
3. **Kendall M.G.** (1978)**,** Time Series. Charles Graffin, London.
4. **K.Tanaka**(1996)**.** Time Series Analysis. John Wiely& Sons, New York.

STA3C12: Multivariate Analysis (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Describe the properties and applications of multivariate normal distribution.
CO2: Explain partial and multiple correlation coefficients.
CO3: Derive the ML Estimates of the mean vector and the dispersion matrix of multivariate normal.
CO4: Describe the genesis of Wishart distribution with its properties.
CO5: Define Hotelling T^2 and Mahalanobis D^2 statistics and able to apply them in testing problems.
CO6: Classify multivariate normal populations.
CO7: Be familiar with principal components and their analysis.

Unit-I. Multivariate Normal Distribution – Definition properties, conditional distribution, marginal distribution.Independence of a linear form and quadratic form, independence of two quadratic forms, distribution of quadratic form of a multivariate vector.Partial and multiple correlation coefficients, partial regression coefficients, Partial regression coefficient.

Unit-II. Estimation of mean vector and covariance vector – Maximum likelihood estimation of the mean vector and dispersion matrix. The distribution of sample mean vector inference concerning the man vector when the dispersion matrix is known for single and two populations. Wishart distribution– properties–generalized variance..

Unit-III. Testing Problems – Mahalnobis D^2 and Hotelling's T^2 Statistics Likelihood ratio tests – Testing the equality of mean vector, equality of dispersion matrices, testing the independence of sub vectors, sphericity test.

Unit-IV. The problem of classification – classification of one of two multivariate normal population when the parameters are known and unknown.Extension of this to several multivariate normal populations.Population principal components – Summarizing sample variation by principal components – Iterative procedure to calculate sample principal components.

Text Books

1. **Anderson, T.W. (1984).** Multivariate Analysis. John Wiley & Sons, New York.
2. **Rao, C.R.(2002).** Linear Statistical Inference and Its Applications- Second Edition. John Wiley & Sons, New York.

References

1. **Giri, N.C. (1996).** Multivariate Statistical Analysis. Marcel Dekker, New York.
2. **Kshirasagar, A.M. (1972).** Multivariate Analysis. Marcel Dekker**,** New York
3. **Rencher, A.C. (1998).** Multivariate Statistical Analysis. John Wiley & Sons, New York**.**

STA3C13: Testing of Statistical Hypotheses (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Compute MP and UMP tests corresponding to any given testing problem.
CO2: Formulate LR test, unbiased tests and similar tests corresponding to any given testing problem.
CO3: Apply different non parametric testing methods.
CO4: Construct SPRT corresponding to any given testing problem.

Unit-I. Tests of hypotheses – error probabilities – Most powerful tests – Neyman-Pearson Lemma – Generalized Neyman–Pearson Lemma.

Unit-II. Method of Finding Tests – Likelihood ratio tests – Bayesian tests – Union – intersection and intersection-union tests. Unbiased and invariant tests – Similar tests and locality most powerful tests.

Unit-III. Non-parametric Tests – Single sample tests – the Kolmogorov – Smirnov test– the sign test–the Wilcoxon signed rank test. Two sample tests – the chi-square test for homogeneity – the Kolmogorov – Smirnov test the median test – the Mann-Whitney-Wilcoxon test-Test for independence –Kendall's tau – Spearman's rank correlation coefficient – robustness.

Unit-IV. Sequential Inference – Some fundamental ideas of sequential sampling – sequential unbiased estimation – sequential estimation of mean of a normal population – the sequential probability tests (SPRT) – important properties – the fundamental identity of SPRT.

Text Books

1. **Casella, G. and Berger, R.L, (2002).** Statistical Inference -Second Edition. Duxbury, Australia..
2. **Rohatgi, V.K. (1976).** An Introduction to Probability Theory and Mathematical Statistics,. John Wiley & Sons, New York.

References

1. **Fraser, D.A. (1957).** Non – parametric Methods in Statistics. John Wiley & Sons, New York.
2. **Lehman, E.L. (1986).** Testing of Statistical Hypotheses. John Wiley & Sons, New York.
3. **Forguson, T.S**. (1967). Mathematical Statistics: A Decision – Theoretic Approach. Academic Press, New York.

STA4C14: Project and Dissertation (8 Credits)

Course Outcomes: On completion of the course, students will be able to:

CO1: Discuss the applications of various statistical techniques learned in the entire course in the form of project work.
CO2: Manage real practical situations where statistical analysis are sought.
CO3: Develop professional approach towards writing and presenting anacademic report.
CO4: Get more insight about the opportunities in research/career.
CO5: Know the works presented in various journals and current trends in their project/dissertation area.
CO6: Get an idea of how new developments in the topic have arose and why new computational techniques are needed.

As a part of the course work, during the fourth semester each student has to undertake a project work in a selected area of interest under a supervisor in the department. The topic could be a theoretical work or data analysis type. At the end of the fourth semester the student shall prepare a report/dissertation which summarizes the project work and submit to the H/D of the parent department positively before the deadline suggested in the Academic calendar. The project/ dissertation is of 8 credits for which the following evaluation will be followed:

The valuation shall be jointly done by the supervisor of the project in the department and an External/Internal Expert from the approved panel, based on a well defined scheme of valuation framed by them. The following break up of weightage is suggested for its valuation.

1. Review of literature, formulation of the problem and defining clearly the objective: 20%
2. Methodology and description of the techniques used: 15%
3. Analysis, programming/simulation and discussion of results: 30%
4. Presentation of the report, organization, linguistic style, reference etc.: 15%
5. Viva-voce examination based on project/dissertation: 20%.

STA4C15: Practical – II (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Apply the principles of Time series analysis, Multivariate analysis, Testing of Statistical Hypotheses and the electives offered in Semester III & IV using real data sets.
CO2: Know the formulas to be applied for the analysis.
CO3: Write the R codes for the analysis of the given data.
CO4: To install and load the packages required to run the R codes.
CO5: Enter the data given for analysis
CO6: Explain how to make conclusions and write the inference for the data analysis based on the output obtained.

The practical is based on the following courses in the third and fourth semesters**.**

1. STA3C11: Time Series Analysis
2. STA3C12: Multivariate Analysis
3. STA3C13: Testing of Statistical Hypotheses
4. Elective – II
5. Elective – III
6. Elective – IV

Practicals are to be done using the R-programming language. The End Semester practical examination shall be conducted and evaluated by two or more examiners nominated by the Department Council.

LIST OF ELECTIVES

Course Code	Course Title	Credits
STA-E01	Models and Methods for Lifetime Data	4
STA-E02	Operations Research	4
STA-E03	Queueing Theory	4
STA-E04	Analysis of Longitudinal Data	4
STA-E05	Analysis of Clinical Trials	4
STA-E06	Statistical Decision Theory	4
STA-E07	Reliability Modeling	4
STA-E08	Actuarial Statistics	4
STA-E09	Statistical Quality Assurance	4
STA-E10	Statistical Machine Learning	4
STA-E11	Statistical Modelling and Data Mining Techniques	4
STA-E12	Applied Algorithms and Big Data Techniques	4
STA-E13	Official Statistics	4
STA-E14	Order Statistics	4
STA- E15	Econometric Models	4
STA-E16	Computer Oriented Statistical Methods	4
STA-E17	Stochastic Finance	4

OPEN ELECTIVES (For other P.G. Programmes under CCSS Scheme)

E18	Statistical Methods	4
E19	Topics in Probability	4

STA-E01: Models and Methods for Lifetime Data (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Learn the basic concepts and standard techniques in the modeling and analysis of lifetime/survival data.

CO2: Examine the structural properties and methods for standard lifetime probability distributions.

CO3: Analyze complete and censored lifetime data with and without covariates.

CO4: Estimate survival functions using parametric and non-parametric methods.

CO5: Apply and interpret semi-parametric and parametric regression models for survival data.

CO6: Apply statistical techniques to model lifetime data and make predictions.

CO7: Use some key methods in system reliability modeling as well as survival analysis.

CO8: Use generalised linear models (GLMs) to analyse data with dependence on one or more explanatory variables.

Unit-I. Lifetime distributions-continuous and discrete models-important parametric models: Exponential Weibull, Log-normal, Log-logistic, Gamma, Inverse Gaussian distributions, Log location scale models and mixture models. Censoring and statistical methods.

Unit-II. Kaplan-Meier product-limit estimate and it properties. The Nelson-Aalen estimate, interval estimation of survival probabilities, asymptotic properties of estimators, descriptive and diagnostic plots, estimation of hazard function, methods for truncated and interval censored data, Life tables.

Unit-III. Inference Under exponential model – large sample theory, type-2 censored test plans, comparison of two distributions; inference procedures for Gamma distribution; models with threshold parameters, inference for log-location scale distribution: likelihood based methods: exact methods under type-2 censoring. Weibull regression and accelerated failure time models.

Unit-IV. Proportional hazard models and Cox regression: assumptions and interpretation, Partial likelihood,Rank test for comparing Distributions, Log-rank test, Generalized Wilcoxon test. A brief discussion on multivariate lifetime models and data.Generalized linear models: linear predictor, link function, canonical link, likelihood equation, the iterative reweighted least squares algorithm.

Text Books

1. **Lawless, J.F.(2003).** Statistical Methods for Lifetime –Second Edition. John Wiley & Sons, New Jersey.
2. **Kalbfiesche, J.D. and Prentice, R.L. (1980).** The statistical Analysis of Failure Time Data. John Wiley & Sons, New Jersey.`
3. **Dobson, A. J. and A. G. Barnett (2018)**. An Introduction to Generalized Linear Models, Chapman & Hall.
4. **A. McCullagh, P. and Nelder, J. A.(1990)**. Generalized Linear Models, Chapman & Hall.

References

1. **Lee, E. T. (1992).** Statistical Methods for Survival Data Analysis. John Wiley& Sons, New York.
2. **Miller, R.G.(1981)**. Survival Analysis. John Wiley & Sons, New York.
3. **A. Moore, D. F. F.(2016)**. Applied Survival Analysis using R, Springer.
4. **Cox, D.R and Oakes, D.(1984)**. Analysis of Survival Data. Chapman & Hall, New York.

STA-E02: Operations Research (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Describe linear programming.
CO2: Discuss simplex method, Big-M method, Two-phase method and Revised simplex method.
C03: Explain the concept of duality, related theorems and dual simplex method.
CO4: Discuss transportation problem, assignment problem.
CO5: Explore integer programming problem.
CO6: Describe game theory.
CO7: Examine deterministic and probabilistic inventory models.
CO8: Employ inventory models in real situations .
CO9: Explain the concept of Non-linear programming and discuss K-T theorems and conditions.

Unit-I. Linear programming:-graphical method, simplex computational procedure, artificial variablestechnique - big M method, two phase method; Revised simplex method. Integer linear programming- Brach and bound algorithm, cutting plane algorithm.

Unit-II. Duality concepts, duality theorems, dual simplex methods. Transportation problems:- general transportation problem, Finding initial basic feasible solution, test for optimality, degeneracy in transportation problem, unbalanced transportation problem, maximization transportation problem, Assignment problem:- mathematical formulation of the problem, the assignment method (Hungarian method).
Game theory: two person zero sum games, minimax theorem, game problem as a linearprogramming problem.

Unit-III. Non-linear programming problem (NLPP):- general non-linear programming problem, Constrained optimization with equality constraints - necessary conditions for a generalized NLPP, sufficient conditions for a general NLPP with one constraint, sufficient conditions for a general problem with m(<n) constraints, Constrained optimization with inequality constraints - Kuhn-Tucker conditions for general NLPP with m(<n) constraints, quadratic programming problem, convex programming problems.

Unit-IV. Inventory models:- Deterministic inventory models - general inventory model, Static economicorder quantity (EOQ) models - classic EOQ model, EOQ with price breaks, multi-item EOQ with storage limitation, Probabilistic inventory models:- Continuous review models - "probabilitized" EOQ model, probabilistic EOQ model, Single-period models - No setup model (Newsvendor model), setup model (s-S policy).

Text Books

1. **Gass S.I. (1985)**.Linear Programming - methods and applications, Fifth edition, McGraw Hill, USA.
2. **KantiSwarup, Gupta, P.K. and Man Mohan (2001)**.Operations Research, Ninth edition, Sultan Chand & Sons.
3. **Taha H.A. (2007).** Operations Research - An introduction, Eighth edition, Prentice-Hall of India Ltd.

References

4. **Ravindran A, Philips D.T and Soleberg J.J. (1997).** Operation Research - Principles and Practice, John Wiley & Sons.
5. **Sinha, S.M. (2006).** Mathematical programming theory and methods, Elsevier, a division of Reed Elsevier India Pvt. Ltd., New Delhi.
6. **Paneerselvam, R. (2008).**Operations Research, Second edition, Prentice Hall of India Pvt.Ltd., New Delhi.

STA-E03: Queueing Theory (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Have rigorous understanding of the theoretical background of queueing Systems.

CO2: Compute quantitative metrics of performance for queueing systems.

CO3: Apply and extend queueing models to analyze real world systems.

CO4: Describe various Markovian queueing models and their analysis.

CO5: Explain transient behaviour of queueing models and analysis of advanced Markovian models with bulk arrival and bulk service.

CO6: Describe various queueing networks and their extensions.

CO7: Explain various non Markovian queueing models and their analysis.

Unit-I. Introduction to queueing theory, Characteristics of queueing processes, Measures of effectiveness, Markovian queueing models, steady state solutions of the M/M/I model, waiting time distributions, Little's formula, queues with unlimited service, finite source queues.

Unit-II. Transient behavior of M/M/1 queues, transient behavior of M/M/∞. Busy period analysis for M/M/1 and M/M/c models.Advanced Markovian models. Bulk input $M^{[X]}$ /M/1 model, Bulk service $M/M^{[Y]}$ /1 model, Erlangian models, $M/E_k/1$and $E_k/M/1$. A brief discussion of priority queues.

Unit-III. Queueing networks-series queues, open Jackson networks, closed Jackson network, Cyclic queues, Extension of Jackson networks. Non Jackson networks.

Unit-IV. Models with general arrival pattern, The M/G/1 queueing model, The Pollaczek-khintchine formula, Departure point steady state systems size probabilities, ergodic theory, Special cases M/Ek/1 and M/D/1, waiting times, busy period analysis, general input and exponential service models, arrival point steady state system size probabilities.

References

1. **Gross, D. and Harris, C.M.(1985).** Fundamentals of Queueing Theory- Second Edition. John Wiley & Sons, New York.
2. **Kleinrock, L. (1976).** Queueing Systems, Vol. I &Vol.II. John Wiley & Sons, NewYork.
3. **Ross, S.M. (2007).** Introduction to Probability Models- Nineth Edition. Academic Press, New York.
4. **Bose, S.K. (2002).** An Introduction to Queueing Systems. Kluwer Academic/ Plenum Publishers, New York.

STA-E04: Analysis of Longitudinal Data (4 Credits)

Course Outcomes: After successful completion of the course, students will be able to:

CO 1: Describe the basic concepts of Linear Model in longitudinal data analysis.
CO 2: Analyze numerical methods to solve the problems in Linear Model.
CO 3: Explain basic concepts of Generalized Linear Model.
CO 4: Illustrate and study on missing data mechanism in longitudinal data analysis.
CO5: Describe Multivariate and Time-dependent covariates in longitudinal data analysis.

Unit-I. General Linear Model for Longitudinal Data.ML and REML estimation, EM algorithm: General linear mixed-effects model, Inference for ; the random effects, BLUPs, Empirical Bayes , Bayes, Shrinkage Model building and diagnostic, Relaxing parametric assumptions: generalized additive mixed model.

Unit-II. Generalized Linear Model for Longitudinal Data: Marginal models, for binary, ordinal, and count data: Random effects models for binary ordinal and count data: Transition models: Likelihood-based models for categorical data; GEE; Models for mixed discrete and continuous responses.

Unit-III. Dropouts and missing data:Classification missing data mechanism; Intermittent missing Values and dropouts; Weighted estimating equations; Modelling the dropout process (Selection and pattern mixture models).

Unit-IV. Time-dependent covariates and special topics:Dangers of time-dependent covariates: Lagged covariates; Marginal Structural models; Joint models for longitudinal and survival data; Multivariate longitudinal data; Design of randomized and observational longitudinal studies.

Text Books

1. **Diggle, P.J., Heagerty, P., Liang, K.Y and Zeger. S.L (2003).** Analysis of Longitudinal Data, 2nd Edn. Oxford University press, New York.
2. **Fitzmaurice, G.M., Laird, N.M and Ware, J.H.(2004).** Applied LongitudinalAnalysis, John Wiley & Sons, New York.

References

1. **Crowder, M.J. and Hand, D.J. (1990).** Analysis of Repeated Measures. Chapman and Hall/CRC Press, London .
2. **Davidian,M. and Giltinan, D.M. (1995**). Nonlinear Models for Repeated Measurement Data. Chapman and Hall/CRC Press, London.
3. **Hand,D and Crowder, M. (1996).** Practical Longitudinal Data Analysis. Chapman and Hall/CRC Press, New York. Lindsey, J.K. (1993) Models for Repeated Measurements. Oxford University Press, New York.
4. **Little, R.J.A, and Rubin, O.B. (2002).**Statistical Analysis with Missing Data, 2nd edition, Wiley, New York.
5. **McCullagh,P. and Nelder.J.A (1989).** Generalized Linear Models. 2nd edition, Chapman and Hall/CRC Press, London.
6. **Weiss, R.E. (2005).** Modeling Longitudinal Data. Springer, New York.

STA-E05 Analysis of Clinical Trials (4 Credits)

Course Outcomes: On successful completion of this course, student will be able to:

CO1:Appraisethe basic concepts of clinical trials.
CO2: Plan and develop the design of clinical trials.
CO3: Determine the sample size in clinical trials.
CO4: Conduct bioassays and assimilate the concepts of meta-analysisin clinical trials.

Unit-1. Basics of Clinical Trials: Introduction to clinical trials, the need and ethics of clinical trials, bias and random error in clinical studies, Protocols, conduct of clinical trials, over view of Phase I-IV trials, Data management-data definitions, standard operating procedure, informed consent form, case report forms, database design, data collection systems for good clinical practice.

Unit-2. Design of Clinical Trials: Design of clinical trials- Different phases, Comparative and controlled trials, Random allocation, Randomization, response adaptive methods and restricted randomization. Methods of Blinding, Parallel group designs, Crossover designs, Symmetric designs, Adaptive designs, Group sequential designs, Zelen's designs, design of bioequivalence trials. Outcome measures.

Unit-3.Sample Size Determination and Testing: Sample size determination in one and two sample cases, comparative trials, activity studies, testing and other purposes, unequal sample sizes and case of anova. Surrogate endpoints-selection and design of trials with surrogate endpoints, analysis of surrogate end point data. Reporting and Analysis-Interpretation of result, multi-center trials.

Unit-4. Meta-Analysis: Meta-analysis in clinical trials-concept and goals, fixed and random effect approaches. Bioassay: Direct and indirect assays, Quantal and quantitative assays, Parallel line and slope ratio assays, Design of bioassays.

Text Books

1. **Friedman, L. M., Furburg, C. D. Demets, L. (1998)**. Fundamentals of Clinical Trials, Springer Verlag.
2. **Jennison and B. W. Turnbull (1999).** Group Sequential Methods with Applications to Clinical Trials, CRC Press.
3. **Kulinskaya E, Morgeathaler S, Staudte R G (2008)**. Meta-analysis, Wiley.

References

1. **Fleiss, J. L. (1989).** The Design and Analysis of Clinical Experiments, Wiley.
2. **Marubeni, E. and M. G. Valsecchi (1994).** Analyzing Survival Data from Clinical Trials and Observational Studies, Wiley and Sons.
3. **Piantadosi S. (1997).** Clinical Trials: A Methodological Perspective. Wiley.
4. **W Rosenberger, J MLachin (2002).** Randomization in Clinical Trials Theory and Practice, Wiley

STA-E06: Statistical Decision Theory (4 Credits)

Course Outcomes: On successful completion of the course, students should be able to:

CO1: Explain different loss functions and decision principle.

CO2: Describe the use of priorinformation in decision making.

CO3: Compute Posterior distribution and check the admissibility of Bayes rules

CO4: Know general techniques for solving games.

Unit-I. Statistical decision Problem – Decision rule and loss-randomized decision rule. DecisionPrinciple – sufficient statistic and convexity. Utility and loss-loss functions-standard loss functions-vector valued loss functions.

Unit-II. Prior information-subjective determination of prior density-Non-informative priors-maximum entropy priors he marginal distribution to determine the prior-the ML-II approach to prior selection. Conjugate priors.

Unit-III. The posterior distribution-Bayesian inference-Bayesian decision theory-empirical Bayes analysis – Hierarchical Bayes analysis-Bayesian robustness Admissibility of Bayes rules.

Unit-IV. Game theory – basic concepts – general techniques for solving games Games with finite state of nature-the supporting and separating hyper plane theorems.The minimax theorem.Statistical games.

Text Book

1. **Berger, O,J.(1985).** Statistical Decision Theory and Bayesian Analysis - Second Edition.Springer, New York.

References

1. **Ferguson, T.S. (1967).** Mathematical Statistics-A Decision Theoretic Approach. Academic Press, New York.
2. **Lehman, E.L.(1983).** Theory of Point Estimation. John Wiley, New York.

STA-E07: Reliability Modelling (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the reliability concepts and measures.

CO2: Discover the system reliability using the concept of structure functions.

CO3: Explain various lifetime probability distributions and their structural properties.

CO4: Describe various concepts and different notions of ageing used in reliability analysis and their inter relations.

CO5: Estimate the reliability function for complete and censored samples.

CO6: Describe univariate and bivariate shock models and carry out reliability estimation based on failure times.

CO7: Describe Maintenance and Replacement Policies.

Unit-I. Reliability concepts and measures; components and systems; coherent systems; reliability of coherent systems; cuts and paths; modular decomposition; bounds on system reliability; structural and reliability importance of components.

Unit-II. Life distributions; reliability function; hazard rate; common life distributions-exponential, Weibull, Gamma etc.Estimation of parameters and tests in these models.Notions of ageing; IFR, IFRA, NBU, DMRL, and NBUE Classes and their duals; closures or these classes under formation of coherent systems, convolutions and mixtures.

Unit-III. Univariate shock models and life distributions arising out of them; bivariate shock models; common bivariate exponential distributions and their properties. Reliability estimation based on failure times in variously censored life tests and in tests with replacement of failed items; stress-strength reliability and its estimation.

Unit-IV. Maintenance and replacement policies; availability of repairable systems; modeling of a repairable system by a non-homogeneous Poisson process.Reliability growth models; probability plotting techniques; Hollander-Proschan and Deshpande tests for exponentiality; tests for HPP vs. NHPP with repairable systems.Basic ideas of accelerated life testing.

References

1. **Barlow, R.E. and Proschan, F.(1985).** Statistical Theory of Reliability and Life Testing. Holt, Rinehart and Winston, New York.
2. **Bain L.J. and Engelhardt**(1991). Statistical Analysis of Reliability and Life Testing Models. Marcel Dekker, New York.
3. **Aven, T**. and **Jensen, U.** (1999). Stochastic Models in Reliability. Springer, NewYork.
4. **Lawless, J.F.** (2003). Statistical Models and Methods for Lifetime -Second Edition. John Wiley & Sons, New York.
5. **Nelson, W** (1982). Applied Life Data analysis. John Wiley & Sons, New York.
6. **Zacks, S.** (1992). Introduction to Reliability Analysis- Probability Models and Statistics Methods. Springer, New York.

STA-E08: Actuarial Statistics (4 Credits)

Course Outcomes: On completion of the course, students will be able to:

CO1: Apply the elements of interest.
CO2: Discuss regular pattern of cash flows and related topics.
CO3: Illustrate and apply individual and collective risk models for a short period.
CO4: Discuss survival distributions and derive survival functions.
CO5: Explain and apply life insurance models.
CO6: Discuss and apply annuity models.

Unit-I. Utility theory, insurance and utility theory, models for individual claims and their sums, survival function, curate future lifetime, force or mortality. Life table and its relation with survival function, examples, assumptions, for fractional ages, some analytical laws of mortality, select and ultimate tables, Multiple life functions, joint life and last survivor status, insurance and annuity benefit through multiple life functions evaluation for special mortality laws.

Unit-II. Multiple decrement models, deterministic and random survivorship groups, associated single decrement tables central rates of multiple decrement, net single premiums and their numerical evaluations. Distribution of aggregate claims, compound Poisson distribution and its applications.

Unit-III. Principles of compound interest: Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding. Life insurance: ;Insurance payable at the moment of death and at the end of the year of death-level benefit insurance, endowment insurance, inferred insurance and varying benefit insurance, recursions, commutation functions. Life annuities: Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities-immediate and apportionable annuities-due.

Unit-IV. Net premiums: Continuous and discrete premiums, true monthly payment premiums, apporionable premiums, ums, commutation function accumulation type benefits. Payment premiums, apportionable premiums, commutations functions, accumulation type benefits. Net premium reserves; Continuous and discrete net premium reserve, reserves on a semi continuous basis, reserves based on true monthly premiums, reserves on an apportionable or discounted continuous basis, reserves at fractional durations, allocations of loss to policy years, recursive formulas and differential equations for reserves, commutation functions.

References

1. **Atkinson, M.E. and Dickson, D.C.M. (2000).** An Introduction toActuarial Studies. Elgar Publishing.
2. **Bedford, T. and Cooke, R. (2001).** Probabilistic Risk Analysis- Foundations and Methods. Cambridge University Press, UK.**Bowers, N. L., Gerber, H.U., Hickman,J.C., Jones D.A. and Nesbitt,C.J.(1986).** Actuarial Mathematics-Second Edition. Society of Actuaries, Ithaca, Illinois, U.S.A.
3. **Medina, P. K. and Merino, S. (2003).** Mathematical Finance and Probability-A Discrete Introduction. Birkhauser, Boston.
4. **Neill, A. (1977).** Life Contingencies. Heineman, London.
5. **Philip, B. et. al (1999).** Modern Actuarial Theory and Practice. Chapman & Hall, New York.
6. **Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1998).** Stochastic Process- es for Insurance and Finance. John Wiley & Sons, New York.
7. **Spurgeon, E.T. (1972).** Life Contingencies, Cambridge UniversityPress.
8. Relevant Publications of the Actuarial Education Co., 31, Bath Street,Abingdon, Oxford shire OX143FE (U.K.)

STA-E09: Statistical Quality Assurance (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Apply different statistical quality control techniques including various types sampling plans for attributes and measure the performance of these plans.

CO2: Explain and design various types of control charts, design control charts and distinguish between them.

CO3: Explain acceptance sampling by variables, Sampling Plans for a single and double specification limits with known and unknown variance, Sampling plans with double specification limits.

CO4: Compare sampling plans by variables and attributes and Continuous sampling plans I, II & III.

Unit-I. Quality and Quality assurance, Methods of Quality assurance, Introduction to TQM. Acceptance sampling for attributes, Single sampling, Double sampling. Multiple sampling and Sequential sampling plans.Measuring the performance of these sampling plans.

Unit-II. Acceptance sampling by variables, sampling plans for single specification limit with known and unknown and unknown variance, Sampling plans with double specification limits., comparison of sampling plans by variables and attributes, Continuous sampling plans I , II,III.

Unit-III. Control charts, Basic ideas, Designing of control charts for tshe number of non-conformities. Mean charts. Median charts. Extreme value charts, R-charts, and S-charts ARI, Economic design of control charts.

Unit-IV. Process capability studies, Control charts with memory – CUSUM charts, EWMA mean charts, OC and ARI for control charts, Statistical process control, Modeling and quality programming. Orthogonal arrays and robust quality.

Text Books

1. **Montgomory, R.C. (1985).** Introduction to Statistical Quality Control-Fourth Edition. John Wiley & Sons, New York.
2. **Mittage, H.J. and Rinne, H. (1993).**Statistical Methods for Quality Assurance. Chapman and Hall, New York. Chapters-13 and 14.
3. **Oakland, J.S. and Follorwel, R.F. (1990).** Statistical Process Control. East-West Press, New York. Chapters- 13 and 14.
4. **Schilling, E.G. (1982).**Acceptance Sampling in Quality Control. Marcel Dekker, New York.

References

1. **Duncan, A.J. (1886).** Quality Control and Industrial Statistics.
2. **Gerant, E.L. and Leaven Worth, R.S. (1980).** Statistical Quality Control. McGraw Hill, New York.
3. **Chin-Knei Chao (1987).** Quality Programming, John Wiley & Sons, New York.

STA-E10: Statistical Machine Learning (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: To explain the concept of supervised learning and its connections to Statistics.

CO2: To use the linear methods of classification, linear regression, logistic regression, Piecewise regression and lasso regression in machine learning.

CO3: To understand kernel smothers and related concepts for machine learning.

CO4: To learn various model assessment techniques and inferential procedures.

Unit-I. Introduction and overview of supervised learning:Variable Types and Terminology, Least Squares and Nearest Neighbours, Local Methods in HighDimension, Supervised Learning and Function Approximation, A Statistical Model for the JointDistribution of input and output vectors, Function Approximation, Structured RegressionModels, ClassesofRestrictedEstimators: Roughness Penalty and Bayesian Methods, Kernel Methods andLocal Regression, Basis Functions and Dictionary Methods. Model Selection and the Bias–VarianceTradeoff. Linear Methods for Regression: Least squares, Subset selection, Shrinkage Methods, Methodsusing derived input directions, Multiple outcome shrinkage and selection, Lasso and related pathalgorithms.

Unit-II. Linear methods for classification: Linear methods for classification using linear regression of an indicator matrix, linear discriminant analysis, logistic regression and separating hyperplanes. Basis expansions and regularizations: Piecewise polynomials and splines, Filtering and feature extraction, smoothing splines, Automatic Selection of theSmoothing Parameters, Nonparametric Logistic Regression, Multidimensional Splines, Regularization and Reproducing Kernel Hilbert Spaces, Wavelet smoothing.

Unit- III. Kernel smoothing: One-Dimensional Kernel Smoothers, Selecting the Width of the Kernel, Structured Local Regression Models in Rp, Local Likelihood and Other Models, Kernel Density Estimation and Classification: Kernel Density Estimation, Kernel Density classification and the Naïve Bayes classifier. Radial Basis Functions and Kernels, Mixture Models for Density Estimation and Classification.

Unit-IV. Model assessment, inference and averaging: Bias, Variance and Model Complexity, The Bias–Variance Decomposition, Optimism of the Training Error Rate, Estimates of In-Sample Prediction Error, The Effective Number of Parameters, The Bayesian Approach and BIC, Minimum Description Length, Vapnik– Chervonenk is Dimension, Cross-Validation, Bootstrap Methods, Conditional or Expected Test Error, introducing Model Inference and averaging: Local regression in IR, The Bootstrap and Maximum Likelihood Methods, Maximum Likelihood Inference, Bootstrap versus Maximum Likelihood, Bayesian Methods, Relationship Between the Bootstrap And Bayesian Inference, The EM Algorithm, MCMC for Sampling from the Posterior, Bagging, Model Averaging and Stacking, Stochastic Search: Bumping.

Text Books/ References

1. **Hastie, T., Tshibirai, R. and Friedman, J. (2017).** The Elements of StatisticalLearning : Data Mining, Inference and Prediction, 2nd edition. Springer, New York.
2. **James, G., Witten, D., Hastie, T. and Tibshirani, R.(2013).**An Introduction toStatistical Learning with Applications in R. Springer, New York.
3. Introduction to Machine Learning, The Wikipedia Guide.
4. **Vinoth, B, Rajarathian, A. and ManjuBargavi, S.K. (2016).**Nonlinear Regression andArtificial Neural Network Based Model for Forecasting Paddy (OryzaSativa) Production in Tamil Nadu. IOSR Journal of Mobile Computing & Application (IOSR-JMCA)e-ISSN: 2394-0050, P-ISSN: 2394-0042.Volume 3, Issue 3. (May. - Jun. 2016), pp. 01-06.

STA-E11: Statistical Modelling and Data Mining Techniques (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Acquire skills in advanced statistical modeling.
CO 2: Get acquainted with association analysis in data mining techniques.
CO3: Learn the data pre-processing steps in data mining.
CO 4: Understand various measures of patterns, data ware housing and the concepts of online transaction and analytical processes in data mining.

Unit-I. Statistical Modeling, Steps in Statistical Modeling, Regression Analysis: transformations and weighting to correct model inadequacies, Analytical methods for selecting a transformation, The Box-Cox method, Transformation on the regressor variables, Ridge regression, Basic form of ridge regression, Robust regression; Least absolute deviation regression, Least median of squares regression.

Unit-II. Data mining; Introduction, Data types for Data mining, Database and Data warehouse, Data mining functionalities -Concept/class description: characterization and discrimination, Association analysis, Classification and Prediction, Clustering analysis, Evolution and Deviation analysis, Data Pre-processing, Data cleaning, Data Integration and transformation, Data reduction, Discretization and concept hierarchy generation.

Unit-III. Measures of pattern interestingness; Objective measures of pattern- sup- port and confidence, Classification of data mining systems; Classifications according to databases, Knowledge and techniques, Major issues in data mining.

Unit IV. Data warehouse, On-line transaction process (OLTP) and On-line analytical processing (OLAP), Distinguishing features between OLTP and OLAP.

Text Books

1. **Montgomery, D.C., Peck, E.A. and Vining,G.G.(2007).** Introduction to Linear Regression Analysis.Wiley, New York.
2. **Han, J., Kamber, M. and Pei, J.(2000).** Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers.

References

1. **Draper, N. R. and Smith, H.(1998).** Applied Regression Analysis, 3rd Edition, Wiley.
2. **Berson, A. and Smith, S.J. (1997).** Data Warehousing, Data Mining, and OLAP. McGraw-Hill, New York.

STA-E12: Applied Algorithms and Big Data Techniques (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to**:**

CO1: Explain the concept of EM clustering algorithms.
CO2: Understand the classification techniques and the concept of supportvector machines.
CO3: Get acquainted with basic concepts related to big data.
CO4: Learn the multidimensional scaling techniques in unsupervisedlearning.

Unit-I: EM Algorithm: Two-Component Mixture Model, GaussianModels, The EM Algorithm in General, EM as aMaximization–Maximization Procedure.

Unit-II: Support Vector Machines: Maximal Margin Classifier,Support Vector Classifiers, Support Vector Machines,SVMs with More than Two Class- One- Versus-OneClassification and One-Versus-All Classification.

Unit-III: Big Data: Definition, Characteristics, Data Analytics,General Categories of Data Analytics, Structured,Unstructured and Semi Structured Data, Met data, BigData Analytics Life Cycle.

Unit-IV: Multi-Dimensional Scaling; Definition, Perceptual Map,Decision Frame- work for Perceptual Mapping, Non-metricversus Metric methods, Similarities Data, Preferences Data, Aggregate and Disaggregate Analysis, Decompositional and Compositional approaches, Interpreting the MDS results. (Practical problems using R).

Text Books/References

1. **Hastie, T., Tibshirani, R. and Friedman, J. (2017)**.The Elements of Statistical Learning, Data Mining,Inference and Prediction, 2nd edition. Springer, New York.
2. **James, G., Witten, D., Hastie, T. and Tibshirani,R.(2013)**. An Introduction to Statistical Learning withApplications in R. Springer, New York.
3. **Erl, T. and Khattak, W. (2016)**. Big DataFundamentals Concepts, Drivers & Techniques. Prentice Hall.
4. **Hair, J. F., Black, W. C., Babin, B. J. andAnderson, R. E.(2009)**. Multivariate Data Analysis, 7thedition. Prentice Hall, New York.

STA-E13: Official Statistics (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Describe the Indian and International Statistical systems.
CO2: Explain the nature of population growth in developed and developing countries.
CO3: Explain the concept of economic development, growth in per capita income and distributive justice.
CO4: Define the indices of development like Human development index etc.
CO5: Estimate national income through income and expenditure approaches.
CO6: Measure inequality in incomes, and measure poverty through measures of incidence and intensity combined.

Unit I: Introduction to Indian and International Statistical systems. Role, function and activities of Central and State Statistical organizations. Organization of large-scale sample surveys. Role of National Sample Survey Organization. General and special data dissemination systems. Scope and Contents of population census of India.

Unit II: Population growth in developed and developing countries, Evaluation of performance of family welfare programmes, projections of labor force and man power. Statistics related to Industries, foreign trade, balance of payment, cost of living, inflation, educational and other social statistics.

Unit III: Economic development: Growth in per capita income and distributive justice indices of development, human development index. National income estimation- Product approach, income approach and expenditure approach.

Unit IV: Measuring inequality in incomes: Gini Coefficient, Theil's measure; Poverty measurements: Different issues, measures of incidence and intensity;Combined Measures: Indices due to Kakwani, Sen etc.

Suggested Readings:

1. Basic Statistics Relating to Indian Economy (CSO) 1990
2. Guide to Official Statistics (CSO) 1999
3. Statistical System in India (CSO) 1995
4. Principles and Accommodation of National Population Census, UNEDCO.
5. **Panse, V.G**.: Estimation of Crop Yields (FAO)
6. Family Welfare Year Book. Annual Publication of D/O Family Welfare.
7. Monthly Statistics of Foreign Trade in India, DGCIS, Calcutta and other Govt. Publications.
8. CSO (1989)a: National Accounts Statistics- Sources and Methods.
9. **Keyfitz, N (1977).** Applied Mathematical Demography. Springer, New York.
10. **Sen, A (1977):** Poverty and Inequality.
11. UNESCO: Principles for Vital Statistics Systems, Series M-12.
12. CSO (1989)b: Statistical System in India
13. **Chubey, P.K (1995).** Poverty Measuremen. , New Age International, New Delhi.

STA–E14: Order Statistics (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Compute distribution free confidence intervals for quantiles.
CO2: Show whether order statistics possess Markov property.
CO3: Derive recurrence relations for moments of order statistics.
CO4: Explain Concomitants of Order Statistics.
CO5: Describe the role of order statistics in testing problems.

Unit-1: Basic distribution theory. Order statistics for a discrete parent.Distribution-free confidence intervals for quantiles and distribution-free tolerance intervals. Conditional distributions, Order Statistics as a Markov chain and characterizations. Order statistics for independently distributed variates.

Unit-2: Moments of order statistics. Large sample approximations to mean and variance of order statistics.Asymptotic distributions of order statistics.Recurrence relations & identities.Distibution-free bounds for moments of order statistics and of the range.

Unit-3: Order statistics for dependent variates, Bounds in the case of dependent variates. Random division of an interval.Concomitants.Application to estimation and hypothesis testing, Relation to Poisson Process.Order statistics from a sample containing a single outlier.

Unit-4: Rank order statistics related to the simple random walk. Dwass' technique.Ballot theorem, its generalization, extension and application to fluctuations of sums of random variables.Galton's rank test statistics. Statistics of Kolmogorov-Smirnov type for two samples.

References

1. **Arnold, B.C. and Balakrishnan, N. (1989).** Relations, Bounds and Approximations for Order Statistics, Vol. 53, Springer, New York.
2. **Arnold, B. C., Balakrishnan, N. and Nagaraja H. N. (1992).** A First Course inOrder Statistics. John Wiley & Sons, New York.
3. **David, H. A. and Nagaraja, H. N. (2003).** Order Statistics- Third Edition. John Wiley & Sons, New York.
4. **Dwass, M. (1967).** Simple random walk and rank order statistics. Ann. Math.Statist. 38, 1042-1053.
5. **Gibbons, J.D. and Chakraborti, S. (1992).** Nonparametric Statistical Inference - Third Edition. Marcel Dekker, New York.
6. **Takacs, L. (1967).** Combinatorial Methods in the Theory of Stochastic Processes, John Wiley & Sons, New York.

STA-E15: Econometric Models (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the meaning and methodology of econometrics.

CO2: Discuss the Leontief input output models and explain the optimization problems in Economics.

CO3: Explain the optimization problems with equality constraints and discuss various production functions like Cobb-Douglas production function and CES production function.

CO4: Discuss the Domar growth model, Solow growth model and Cobweb model.

CO5: Explain the meaning of Multi collinearity, Heteroscedasticity, Autocorrelation and discuss various dynamic econometric models.

CO6: Describe the Simultaneous equation models and Approaches to econometric forecasting.

Unit-I. Basic economic concepts: Demand, revenue, average revenue, marginal revenue, elasticity of demand, cost function, average cost, marginal cost. Equilibrium analysis: Partial market equilibrium- linear and nonlinear model, general market equilibrium, equilibrium in national income analysis. Leontief input output models. Optimization problems in economics, Optimization problems with more than one choice variable: multi product firm, price discrimination.

Unit-II. Optimization problems with equality constraints: utility maximization and consumer demand, homogeneous functions, Cobb-Duglas production function, least cost combination of inputs, elasticity of substitution, CES production function. Dynamic analysis: Domar growth model, Solow growth model, Cobweb model.

Unit-III. Meaning and methodology of econometrics, regression function, multiple regression model, assumptions, OLS and ML estimation, hypothesis testing, confidence interval and prediction. Multicollinearity, Heteroscedasticity, Autocorrelation: their nature, consequences, detection, remedial measures and estimation in the presence of them.
Dynamic econometric models: Auto regressive and distributed lag- models, estimation of distributed lag- models, Koyck approach to distributed lag- models, adaptive expectation model, stock adjustment or partial adjustment model, estimation of auto regressive models, method of instrumental variables, detecting autocorrelation in auto regressive models: Durbin- h test, polynomial distributed lag model.

Unit-IV. Simultaneous equation models: examples, inconsistency of OLS estimators, identification problem, rules for identification, method of indirect least squares, method of two stage least squares.
Time series econometrics: Some basic concepts, stochastic processes, unit root stochastic processes, trend stationary and difference stationary stochastic processes, integrated stochastic processes, tests of stationarity, unit root test, transforming non-stationary time series, cointegration. Approaches to economic forecasting, AR, MA, ARMA and ARIMA modeling of time series data, the Box- Jenkins methodology.

Text Books

1. **Chiang, A.C. (1984).** Fundamental Methods of Mathematical Economics –Third Edition. McGraw – Hill, New York.
2. **Gujarati, D.N. (2007).** Basic Econometrics -Fourth Edition. McGraw-Hill, New York.

References

1. **Johnston, J. (1984).** Econometric Methods -Third Edition. McGraw–Hill, New York.
2. **Koutsoyiannis, A (1973).** Theory of Econometrics, Harper & Row, New York.
3. **Maddala ,G.S. (2001).** Introduction to Econometrics - Third Edition. John Wiley & Sons, New York.
4. **Yamane, T. (1968).** Mathematics for Economists- An Elementary Survey- Second Edition.

STA-E16: Computer Oriented Statistical Methods (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the basic concepts of R software.

CO2**:** Create vectors and matrices and carry out basic matrix operations using R.

CO3: Visually display, analyze, clarify and interpret numerical data , functions and other quantitative structures.

CO4**:** Use various types of looping techniques.

CO5**:** Become familiar with different re-sampling techniques and use them for estimation of sampling distribution, creating confidence intervals etc.

CO6: Use EM algorithm method of finding maximum likelihood estimation when data contains missing values.

Note:-*The objective of the course is to enhance the programming skills and working knowledge of available numerical and statistical softwares. The primary need is to abreast them with the latest developments in the computing world thereby enabling them to perform data analysis effectively and efficiently in any specialized statistical software.*

Unit-I. Introduction to the statistical software R, Data objects in R, Creating vectors, Creating matrices, Manipulating data, Accessing elements of a vector or matrix, Lists, Addition, Multiplication, Subtraction, Transpose, Inverse of matrices. Read a file. Boolean operators.

Unit-II. R-Graphics- Histogram, Box-plot, Stem and leaf plot, Scatter plot, Matplot, Plot options; Multiple plots in a single graphic window, Adjusting graphical parameters. Looping- For loop, repeat loop, while loop, if command, if else command.

Unit-III. Bootstrap methods: re-sampling paradigms, bias and standard errors, Bootstrapping for estimation of sampling distribution, confidence intervals, variance stabilizing transformation, bootstrapping in regression and sampling from finite populations. Jackknife and cross- validation: jackknife in sample surveys, jack-knifing in regression with hetero-sedasticity cross-validation for tuning parameters.

Unit-IV. EM algorithm: applications to missing and incomplete data problems, mixture models. Applications to Bayesian analysis, Smoothing with kernels: density estimation, simple nonparametric regression.

Text Books / References

1. **Alain F. Zuur, Elena N. Ieno, and Erik Meesters**(2009): "A Beginner's Guide to R", Springer,
2. **Michael J. Crawley** (2005): "Statistics: An Introduction using R", Wiley, ISBN 0-470-02297-3.
3. **Phil Spector** (2008): "Data Manipulation with R", Springer, New York, ISBN 978-0-387-74730-9.
4. **Maria L. Rizzo** (2008): "Statistical computing with R", Chapman & Hall/CRC, Boca Raton, ISBN 1-584-88545-9.
5. **W. John Braun and Duncan J. Murdoch** (2007): "A first course in Statistical programming with R", Cambridge University Press, Cambridge, ISBN 978-0521872652.
6. **Fishman, G.S.** (1996): Monte Carlo: Concepts, Algorithms, and Applications.(Springer).

7. **Rubinstein, R.Y.** (1981): Simulation and the Monte Carlo Method. (Wiley).
8. **Tanner, M.A.** (1996): Tools for Statistical Inference, Third edition. (Springer.)
9. **Efron, B. and Tibshirani. R.J.** (1993): An Introduction to the Bootstrap.
10. **Davison, A.C. and Hinkley, D.V.** (1997): Bootstrap Methods and their applications , Chapman and Hall.
11. **Shao J. and Tu, D.** (1995): The Jackknife and the Bootstrap. Springer Verlag.
12. **McLachlan, G.J. and Krishnan, T**. (1997) : The EM Algorithms and Extensions. (Wiley.)
13. **Simonoff , J.S.** (1996) : Smoothing Methods in Statistics. (Springer).

STA-E17: Stochastic Finance (4 Credits)

Course Outcomes: After successful completion of the course, students will be able to:

CO1: Expain basic concepts of financial markets and market lines.
CO2: Learn the usage of Statistical models in modelingFinancial data.
CO3: Interpret and apply the black Scholes theorem and its properties.
CO4: Describe the pricing of European and American options by Monte-Carlo and finite difference methods.
CO5: Discuss on the modeling security market and price process models.
CO6: Describe the special features of the financial time series, their models and its estimation.

Unit-I. Basic concepts of financial markets. Forward contracts, futures contracts, options-call and put options, European option and American options. Hedgers, speculators, arbitrageurs. Interest rates, compounding, present value analysis, risk free interest rates. Returns, gross returns and log returns. Portfolio theory – trading off expected return and risk, one risky asset and one risk free asset. Two risky assets, estimated expected return. Optimal mix of portfolio CAPM, capital market line, betas and security market line.

Unit-II. Options, pricing via arbitrage, law of one price. Risk neutral valuation.Binomial model-single and multiperiod binomial model, martingale measure.Modelling returns: lognormal model, random walk model, geometric Brownian motion process. Ito lemma (without proof).Arbitrage theorem.The Black-Scholes formula. Properties of the Black-Scholes option cost, the delta hedging arbitrage strategy. Some derivatives, their interpretations and applications.

Unit-III. Volatility and estimating the volatility parameter.Impliedvolatility.Pricing American options.Pricing of a European option using Monte-Carlo and pricing an American option using finite difference methods. Call options on dividend-paying securities. Pricing American put options, Modeling the prices by adding jumps to geometric Brownian motion. Valuing investments by expected utility. Modeling security market: Self-financing portfolio and no arbitrage, price process models, division rule, product rule.

Unit-IV. Financial Time Series – Special features of financial series, Linear time series models: AR(1), AR(p), ARMA(p,q) processes, the first and second order moments, estimation and forecasting methods. Models for Conditional heteroscedasticity: ARCH(1), ARCH(p), GARCH(p,q) models and their estimation. Comparison of ARMA and GARCH processes.

Text Books/References

1. **Sheldon M. Ross (2003).** "An elementary introduction to Mathematical Finance",
2. **David Ruppert (2004).** "Statistics and Finance an Introduction" – Springer International Eddition.
3. **Masaaki Kijima (2003).** "Stochastic process with applications to finance", ChapmanHall.
4. **Ruey S. Tsay (2005).** "Analysis of Time Series III ed", John Wiley &Sons**Hull, J.C. (2008).** Options, Futures and other derivatives. Pearson Education India, New Delhi.
5. **Hull, J.C. (2008).** Options, Futures and other derivatives. Pearson Education India, New Delhi.
6. **Gourieroux, C. and Jasiak, J. (2005).** Financial Econometrics. New Age International (P) Ltd., New Delhi.
7. **Cuthbertson, K. and Nitzsche, D. (2001).** Financial Engineering - Derivatives and Risk Management. John Wiley & Sons, New York.
8. **BrockwellP.J.and Davis R.A. (2006).** Time Series: Theory and Methods, Springer – Verlag.
9. **Robert H Shumway and Davis S Stoffer(2016).** Time series analysis and its applications with R examples. Springer.

Elective Course for Other P.G. Programmes under CCSS Scheme

STA-E18: Statistical Methods (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Express data graphically.

CO2: Compute different measures of central tendency, dispersion and correlation coefficient.

CO3: Define Probability, conditional probability and independence of events.

CO4: Apply Simple Random Sampling with and without replacement.

CO5: Define Binomial, Poisson and Normal Distributions and apply them for real data modeling.

CO6: Develop confidence intervals under different set up.

CO7: Carry out one way and two way ANOVA.

Unit-I. Biostatistics – Definition, Applying Statistical Methods, Descriptive Methods – Tabular and Graphical Presentation of Data – Frequency Tables, Line Graphs, Bar Charts, Histograms, Stem and Leaf Plots, Dot Plots, Scatter Plots.Measures of Central Tendency – Mean, Median and Mode.Measures of Dispersion – Range and Percentiles, Box Plots, Variance.Correlation Coefficients – Pearson Correlation Coefficient, Spearman Rank Correlation Coefficient.

Unit-II. Probability – Definition, Conditional Probability, Independence of events, Baye's Theorem, Probability Sampling – Sampling With Replacement, Sampling Without Replacement. Designed Experiments – Single and Double Blind experiments. The Life Table – The First Four Columns in the Life Table – Uses of Life Table. Probability Distributions – Binomial , Poisson and Normal, Central Limit Theorem. Sampling Distributions-t , χ^2 and F distributions (Concepts and applications).

Unit-III. Interval Estimation – Confidence Intervals Based on the Normal Distribution, Confidence Intervals for the Difference of Two Means and Proportions. Tests of Hypotheses – Preliminaries.Testing Hypothesis about the Mean.Testing Hypothesis about the Difference of Two Means. Analysis of Categorical Data – Goodness of Fit Test, Contingency Table.

Unit-IV. Analysis of Variance – Assumptions for the Use of ANOVA – One-Way ANOVA, Two-Way ANOVA. Concept of Regression – Simple Linear Regression, multiple linear regression; Basic concept of multivariate distributions.

Text Books

1. **Parimal Mukhopadhyay (2016).** Mathematical Statistics. Books and Allied
2. **Parimal Mukhopadhyay (2012).** Theory of Probability. New Central Book Agency.
3. **S. P. Gupta (2021).** Statistical Methods. Sultan Chand & Sons. Edition 46
4. **Forthofer, R. (2012.N., Lee, E. S. and Hernandez, M.**(2007). Biostatistics – A Guide to Design, Analysis and Discovery- Second Edition. Elsevier, New Delhi.

References

1. **Rastogi, V. B.** (2006). Fundamentals of Biostatistics. Ane Books India, New Delhi.
2. **SundarRao, P. S. S. and Richard, J.** (1996). An Introduction to Biostatistics- Third Edition. Prentice Hall of India, New Delhi.

STA-E19: Topics in Probability (4 Credits)

Course Outcomes: On successful completion of the course, students will be able to:

CO1: Explain the concepts of probability measure, random variables and decomposition of distribution functions.
CO2: Describe induced probability space and notion of vector valued random variables
CO3: Explain Expectation of simple, non-negative and arbitrary random variables
CO4: Distinguish between different types of inequalities such as Cr-inequality, Basic inequality, etc.
CO5: Illustrate convergence of sequence of random variables with examples
CO6: Discuss independence of events, Borel Cantelli Lemma, Borel and Kolmogrov zero- one criteria.
CO7: Describe properties of Characteristic Function, inversion theorem and its applications.

Unit-I. Probability measure, measure, probability space, random variable. Inverse function and properties. Sequence of random variables and limit. Extension of probability measure - Caratheodory extension theorem (without proof). Distribution function. decomposition of distribution function. Vector valued random variables and its distribution function. Induced probability space of a random variable.

Unit-II. Mathematical expectation of simple, non-negative and arbitrary random variables - properties of expectation. Moment generating functions-moments. Inequalities. Cr- inequality, Jenson's inequality, Basic inequality, Markov inequality.

Unit-III. Different modes of convergence. Convergence in probability, convergence in distribution, r^{th} mean convergence, almost sure convergence and their mutual implications.

Unit-IV. Independence of events, classes of events. Independence of random variables. Kolmogorow's 0-1 law, Borel's 0-1 criteria. Borel-Cantelli Lemma. Characteristic Functions-definition, properties, inversion theorem, inversion formula for lattice distributions, Characteristic functions and moments, Taylor's series for characteristic functions, Bochner's theorem (without proof).

Text Books

1. **Bhat, B.R. (1999).** Modern Probability Theory- Third Edition. New-Age International, New Delhi.

References

1. **Resnick, S.I. (1999).** Probability Paths. Birkhauser, Boston.
2. **Laha, R.G. and Rohatgi, V.K**. (1979). Probability Theory. John Wiley& Sons, New York.
3. **Billingsly, P. (1995).** Probability and Measure- Third Edition. John Wiley & Sons, New York.
4. **Basu, A.K. (1999).** Measure Theory and Probability. Printice Hall of India, New-Delhi.
5. **Rohatgi, V.K**. (1976). An Introduction to Probability Theory and Mathematical Statistics. .John Wiley & Sons , New York.

M. Sc. Statistics Programme under CCSS
at the Department of Statistics, University of Calicut

Audit courses:

In addition to the core and elective courses of the programme there will be two Audit Courses (Ability Enhancement Course & Professional Competency Course) with 2 credits each. These have to be done one each in the first two semesters. These courses are mandatory for all programmes but their credits will not be counted for evaluating the overall SGPA & CGPA. The Department/College shall conduct examination for these courses and have to intimate /upload the results of the same to the University on the stipulated date during the Third Semester. Students have to obtain only minimum pass requirements in the Audit Courses. The details of Audit courses are given below.

STA1A01: Ability Enhancement Course (AEC) 2 Credits

The objective of this course is to enhance the ability and skill of students in the core and elective areas of statistics, through hands on experience, internship, industrial visits, case study, community linkage, book/research paper review, scientific word processing etc.

The faculty members in the department collectively or a particular faculty member shall be in charge of this course for students of the semester, which shall be decided by the Department council. The following are the requirements in this course:-

1. Short term internships at research institutions/R&D centre/Industry.
2. Seminar presentation on a topic in statistics or related fields that is not normally covered in the in the syllabi of the programme.
3. Case study and analysis on any relevant issues in the nearby society
4. Publication of articles in statistical magazines/journals
5. Interaction with Statistical Organizations/ Industries/ Research Institutions.
6. Any community linking programme relevant to the area of study
7. Book/paper review and summary.
8. English communication skills and technical writing with LATEX.

9. Survey methodology and Data collection- sampling frames and coverage error, non-response.
10. Developing a questionnaire, collect survey data pertaining to a research problem (such as gender discrimination in private vs government sector, unemployment rates, removal of subsidy, impact on service class). Formats and presentation of reports.

After conducting the AEC, the evaluation/examination should be done either common for all students of the semester or individually depending upon the AEC conducted. Evaluation/examination on AEC must contain the following components: MCQ type written examination, Report on study/investigation, Presentation, Viva voce etc. as decided by the Department council. Evaluation/examination must be conducted as in the theory courses and the GPA and overall grade of the AEC should be determined.

STA2A02 : Professional Competency Course (PCC) 2 Credits

The objective of this course is to get professional competency and exposure in the core areas of statistics. It particularly aims to improve the skill level of students, especially for using software useful in their respective field of study, both related to the core and elective subject area. Also it is a platform for the student community to undertake socially committed statistical investigations and thereby developing a method of learning process by doing through the involvement with society.

The faculty members in the department collectively or a particular faculty member shall be in charge of this course for students of the semester, which shall be decided by the Department council. The following are the requirements in this course:-

1. Working knowledge on different statistical software/utilities like SPSS (or GNU PSPP), R, Python. (Introduction of the software- Use of the software as a calculator, as a graphing (plotting) utility, for matrix operations and for problems on probability distributions)
2. Use of Internet and other technologies - Internet and www, applications, internet protocols.
3. E-commerce and financial statistics- Electronic fund transfer, payment portal, e-commerce security.

4. Mobile commerce, Bluetooth and Wi-Fi
5. Introduction to Data Science and Big-data issues.
6. Trend Analysis (elementary time series analysis) and Index numbers
7. Official Statistics: An outline of present official statistical systems in India, Methods of collection of official statistics, their reliability and limitations, Role of MoSPI, CSO, NSSO and NSC.
8. Monte Carlo methods: Brief look at some popular approaches- simulating a coin toss, a die roll and a card shuffle.
9. CDF inversion method- simulation of standard distributions
10. Monte Carlo Integration- Basic ideas of importance sampling.

After conducting the PCC, the evaluation/examination should be done either common for all students of the semester or individually depending upon the PCC conducted. Evaluation/examination on PCC must contain the following components: MCQ type written examination, Report on study/investigation, Presentation, Viva voce etc. as decided by the Department council. Evaluation/examination must be conducted as in the theory courses and the GPA and overall grade of the PCC should be determined.

Sd/-
Dr. K. Jayakumar
Professor & Chairman, BoS in Statistics (PG)
University of Calicut