CHRIST COLLEGE (AUTONOMOUS), IRINJALAKUDA

DEGREE OF M. Sc. Statistics

MASTER OF SCENCE IN STATISTICS

(CHOICE BASED CREDIT AND SEMESTER SYSTEM FOR POSTGRADUATE CURRICULUM)

UNDER THE FACULTY OF SCIENCE

SYLLABUS

(FOR THE STUDENTS ADMITTED FROM THE ACADEMIC YEAR 2022 – '23 ONWARDS)

BOARD OF STUDIES IN STATISTICS (PG)

CHRIST COLLEGE (AUTONOMOUS), IRINJALAKUDA - 680125, KERALA, INDIA

Members of Board of Studies

Sl. No.	Name	Designation
1	Dr. Davis Antony Mundassery (Chairman)	Associate Professor, Christ College (Autonomous) Irinjalakuda
2	Dr. K. Jayakumar(University Nominee)	Associate Professor, University of Calicut
3	Dr. S. M. Sunoj	Professor, Cochin University of Science and Technology
4	Dr. V. M. Chacko	Assistant Professor, St. Thomas (Autonomous) College Thrissur
5	Dr. Vinu C.T.	Assistant Professor, Indian Institute of Management, Trichy
6	Mr. Ravikumar K.	Assistant Professor, K.K.T.M. Govt. College, Pullut
7	Dr. Divya P.R.	Assistant Professor, Vimala college (Autonomous) Thrissur
8	Mr. Nijesh M.A.	Accenture India Pvt. Ltd. Bengaluru
9	Dr. Sr. Mariyamma K.D.	Assistant Professor, Christ College (Autonomous) Irinjalakuda
10	Ms. Megha C.M.	Assistant Professor (Adhoc) Christ College (Autonomous) Irinjalakuda
11	Ms. Jiji M.B.	Assistant Professor (Adhoc) Christ college (Autonomous) Irinjalakuda
12	Ms. Geethu Gopinath	Assistant Professor (Adhoc) Christ college (Autonomous) Irinjalakuda
13	Ms. Mary Priya	Assistant Professor(Adhoc) Christ college (Autonomous) Irinjalakuda
14	Ms. Sreedevi P.N.	Assistant Professor(Adhoc) Christ college (Autonomous) Irinjalakuda
15	Ms. Linett George	Assistant Professor(Adhoc) Christ college (Autonomous) Irinjalakuda
16	Ms. Sruthi Mohan	Assistant Professor(Adhoc) Christ college (Autonomous) Irinjalakuda

SYLLABI of M.Sc. Statistics

M.Sc. (Statistics) Degree Programme under the Credit Semester System (CSS) (2019 admissions) Duration of programme: Two years – divided into four semesters of not less than 90 working days each.

Course Code	Type	Course Title	Credits
1st SEMESTER (Total Credits: 18)			
ST1C01	Core	Analytical Tools for Statistics -1	4
ST1C02	Core	Analytical Tools for Statistics -2	4
ST1C03	Core	Distribution Theory	4
ST1C04	Core	Probability Theory	4
ST1C05	Core	Statistical Computing – 1	2
2nd SEMESTER (Total Credits: 18)			
ST2C06	Core	Design and Analysis of Experiments	4
ST2C07	Core	Estimation Theory	4
ST2C08	Core	Sampling Theory	4
ST2C09	Core	Testing of Statistical Hypotheses	4
ST2C10	Core	Statistical Computing – 2	2
3rd SEMESTER (Total Credits: 18)			
ST3C11	Core	Applied Regression Analysis	4
ST3C12	Core	Stochastic Processes	4
ST3E--	Elective	Elective-I	4
ST3E--	Elective	Elective-II	4
ST3C13	Core	Statistical Computing – 3	2
4th SEMESTER (Total Credits: 18)			
ST4C14	Core	Multivariate Analysis	4
ST4E--	Elective	Elective-III	4
ST4C15	Core	Project/Dissertation and External Viva-Voce	8
ST4C16	Core	Statistical Computing – 4	2

The courses Elective –I, Elective –II and Elective –III shall be chosen from the following list.		
Course Code	**Course Title**	**Credits**
01	Operations Research-I	4
02	Time Series Analysis	4
03	Operations Research – II	4
04	Queueing Theory	4
05	Lifetime Data Analysis	4
06	Advanced Distribution Theory	4
07	Statistical Decision Theory	4
08	Reliability Theory	4
09	Actuarial Statistics	4
10	Statistical Quality Control	4
11	Advanced Probability Theory	4
12	Official Statistics	4

13	Biostatistics	4
14	Econometric Models	4
15	Demographic Techniques	4
16	Statistical Computing	4
17	Stochastic Finance	4
18	Longitudinal Data Analysis Computer Oriented	4
19*	Statistical Machine Learning	4
20*	Advanced Statistical Machine Learning Techniques	4

* Courses with Code No. 19 and 20 are newly approved courses

SEMESTER - I
STA1C01 – ANALYTICAL TOOLS FOR STATISTICS - I

Number of Credits: 4

Course Outline

Unit-I- Multivariable Functions

Limits and continuity of multivariable functions. Derivatives, directional derivatives and continuity. Total derivative in terms of partial derivatives, Taylor's theorem. Inverse and implicit functions. Optima of multivariable functions. Method of Lagrangian multipliers, Riemann integral of a multivariable function.

Unit-II- Analytic functions and complex integration

Analytical functions, Harmonic functions, Necessary condition for a function to be analytic, Sufficient condition for a function to be analytic, Polar form of Cauchy- Riemann equation, Construction of analytic function. Complex integral, Cauchy's theorem, Cauchy's integral formula and its generalized form. Poisson integral formula, Morera's theorem. Cauchy's inequality, Lioville's theorem, Taylor's theorem, Laurent's theorem.

Unit-III- Singularities and calculus of residue

Zeroes of a function, singular point, different types of singularities. Residue at a pole, residue at infinity, Cauchy's residue theorem, Jordan's lemma, Integration around a unit circle. Poles on the real axis, Integration involving many valued functions.

Unit-IV- Laplace transform and Fourier Transform

Laplace transform, Inverse laplace transform. Applications to differential equations, The infinite Fourier transform, Fourier integral theorem. Different forms of Fourier integral formula, Fourier series.

Text Book:

1. Andre's I. Khuri(1993) Advanced Calculus with applications in statistics. Wiley & sons (Chapter7)

2. Pandey, H.D, Goyal, J. K & Gupta K.P (2003) Complex variables and integral transforms Pragathi Prakashan, Meerut.

3. Churchill Ruel. V. (1975), Complex variables and applications. Mc Graw Hill.

References:

1. Apsostol T. M. (1974): Mathematical Analysis, Second edition Norosa, New Delhi.
2. Malik, S.C & Arora. S (2006): Mathematical analysis, second edition, New age international

STA1C02 – ANALYTICAL TOOLS FOR STATISTICS - II

Number of Credits: 4

Course Outline

Unit-I- Basics of linear algebra

Definition of vector space, sub spaces, linear dependence and independence, basis and dimensions, direct sum and compliment of a subspace, quotient space, Inner product and orthogonality.

Unit-II- Algebra of Matrices

Linear transformations and matrices, operations on matrices, properties of matrix operations, Matrices with special structures-triangular matrix, idempotent matrix, Nilpotent matrix, symmetric, Hermitian and skew Hermitian matrices, unitary matrix. Row and column space of matrix, inverse of a matrix. Rank of product of matrix, rank factorization of a matrix, rank of a sum and projections, Inverse of a partitioned matrix, Rank of real and complex matrix, Elementary operations and reduced forms.

Unit-III- Eigen values, spectral representation and singular value decomposition

Characteristic roots, Cayley-Hamilton theorem, minimal polynomial, eigen values and eigen spaces, spectral representation of a semi simple matrix, algebraic and geometric multiplicities, Jordan canonical form, spectral representation of a real symmetric, Hermitian and normal matrices, singular value decomposition.

Unit- IV- Linear equations generalized inverses and quadratic forms

Homogenous system, general system, Rank Nullity Theorem, generalized inverse, properties of ginverse, Moore-Penrose inverse, properties, computation of g-inverse, definition of quadratic forms, classification of quadratic forms, rank and signature, positive definite and non negative definite matrices, extreme of quadratic forms, simultaneous diagonalization of matrices.

Text Books:

1. Ramachandra Rao and Bhimashankaran (1992). Linear Algebra Tata McGraw hill
2. Lewis D.W (1995) Matrix theory, Allied publishers, Bangalore.
3. Walter Rudin (1976). Principles of Mathematical Analysis, third edition, McGraw –hill international book company New Delhi.

References:

1. Suddhendu Biswas (1997) A text book of linear algebra, New age international.
2. Rao C.R (2002) Linear statistical inference and its applications, Second edition, John Wiley and Sons, New York.
3. Graybill F.A (1983) Matrices with applications in statistics.

STA1C03 – DISTRIBUTION THEORY

Number of Credits: 4

Course Outline

Unit-1:

Discrete distributions Random variables, Moments and Moment generating functions, Probability generating functions, Discrete uniform, Binomial, Poisson, Geometric, Negative binomial, Hyper geometric and Multinomial distributions, Power series distributions.

Unit-2:

Continuous distributions: Uniform, Normal, Exponential, Weibull, Pareto, Beta, Gama, Laplace, Cauchy and Log-normal distributions. Pearsonian system of distributions, location and scale families.

Unit-3:

Functions of random variables. Joint and marginal distributions, Conditional distributions and independence, Bivariate transformations, Covariance and Correlations, Bivariate normal distributions,

Hierarchical models and Mixture distributions, Multivariate distributions, Inequalities and Identities. Order statistics.

Unit-4:

Sampling distributions Basic concept of random sampling, Sampling from normal distributions, Properties of sample mean and variance. Chi-square distribution and its applications, t-distribution and its applications. F-distributions- properties and applications. Noncentral Chi-square, t, and F- distributions.

Text Books:

1. Rohatgi, V. K. (1976). Introduction to probability theory and mathematical statistics. John Wiley and sons.
2. George Casella and Roger L. Berger (2003). Statistical Inference. Wodsworth & brooks Pacefic Grove, California.

References:

1. Johnson, N. L., Kotz. S. and Balakrishnan N. (1995). Continuous univariate distributions, Vol. I &Vol. II, John Wiley and Sons, New York.
2. Johnson, N. L., Kotz. S. and Kemp. A.W. (1992). Univariate Discrete distributions, John Wiley and Sons, New York
3. Kendall, M. and Stuart, A. (1977). The Advanced Theory of Statistics Vol I: Distribution Theory, 4th Edition

STA1C04 – PROBABILITY THEORY

Number of Credits: 4

Course Outline

Unit-I: Sets and classes of events

Sequences of sets and their limits – Fields, Sigma fields, Borel field. Random variables, Sigma fields induced by random variables, Vector random variables, limits of sequence of random variables, Probability space, General Probability space, Induced probability space, Concepts of other measures.

Unit-II: Distribution functions of random variables.

Decomposition of distribution functions, Distribution function of vector random variables, Correspondence theorem, Expectation and moments, Properties of expectations, Moments and inequalities, Characteristic functions, Properties, Inversion theorem, Characteristic functions and moments, Bochner's theorem (No proof required), Independence of classes of events; Independence of random variables; Kolmogorov 0-1 law; Borel 0-1 law.

Unit-III: Convergence of random variables

Convergence in probability, Convergence almost surely, Convergence in distribution, Convergence in rth mean – their inter-relations- examples and counterexamples. Convergence of distribution functions; Weak convergence, Helly-Bray Lemma and Helly – Bray theorem, Levy continuity theorem.

Unit- IV: Law of Large Numbers

Kolmogorov inequality, Kolmogorov three series theorem; Weak law of large numbers (both IID and Non-IID cases). Strong Law of large numbers (Law of iterated logarithm not included), Central Limit Theorem (CLT), Lindeberg-Levy theorem, Liapounov form of CLT. Lindeberg-Feller CLT (no proof required). Association between Liapounov's condition and Lindeberg conditions; Simple applications of CLT

Text book:

1. B.R Bhat (1999), Modern Probability theory, Wiley Eastern
2. Laha & Rohatgi (19 79), Probability theory, Wiley New York

References:

1.   Patrick Billingsley (1995), Probability and measure, Wiley New York
2.   Galambos (1988), Advanced probability theory, Marcel Dekar, New York.

STA1C05 – STATISTICAL COMPUTING - I

Contact Hours per Week: 6 Hrs.

Number of Credits: 2

Course Outline

Statistical Computing-I is a practical course. Its objectives are to develop scientific and experimental skills of the students and to correlate the theoretical principles with application-based studies. The practical is based on the following THREE courses of the first semester.

1. ST1C01: Analytical Tools for Statistics – I
2. ST1C02: Analytical Tools for Statistics – II
3. ST1C03: Distribution theory

Practical is to be done using R & Phython Programming. At least five statistical data oriented/supported problems should be done from each course. Practical Record shall be maintained by each student and the same shall be submitted for verification at the time of external examination. Students are expected to acquire working knowledge of the statistical package – EXCEL.

The Board of Examiners (BoE) shall decide the pattern of question paper and the duration of the external examination. The external examination at each centre shall be conducted and evaluated on the same day jointly by two examiners – one external and one internal, appointed at the centre of the examination by the University on the recommendation of the Chairman, BoE. The question paper for the external examination at the centre will be set by the external examiner in consultation with the Chairman, BoE and the H/Ds of the centre. The questions are to be evenly distributed over the entire syllabus. Evaluation shall be done by assessing each candidate on the scientific and experimental skills, the efficiency of the algorithm/program implemented, the presentation and interpretation of the results. The valuation shall be done by the direct grading system and grades will be finalized on the same day.

SEMESTER - II

STA2C06 – DESIGN AND ANALYSIS EXPERIMENTS

Number of Credits: 4

Course Outline

Unit-I: Randomization, Replication and local control.

One way and two-way classifications with equal and unequal number of observations per cell with and without interaction, Fixed effects and Random effects model. Model adequacy checking, CRD, RBD and Latin Square designs, Analysis of co-variance for completely randomized and randomized block designs. Analysis of experiments with missing observations.

Unit-II: Incomplete Block Designs

Balanced Incomplete Block designs, Construction of BIB Designs, Analysis with recovery of inter-block information and intra-block information. Partially balanced incomplete block designs, Analysis of partially balanced incomplete block designs with two associate classes, Lattice designs.

Unit-III: 2n Factorial experiments.

Analysis of 2n factorial experiments. Total confounding of 2n designs in 2n blocks. Partial confounding in 2n blocks. Fractional factorial designs, Resolution of a design, 3n factorial designs. Split plot design and strip plot design (outline only).

Unit-IV:

Response surface designs, Orthogonality, Rotatability blocking and analysis - Method of Steepest accent, Models properties and Analysis.

Text Books:

1. Montgomery, D.C. (2001)) Design and Analysis of Experiments, John Wiley.
2. Das M N and Giri N.C. (1979) Design and Analysis of Experiments, second edition, Wiley.
3. Hinkleman and Kempthrone, C. (1994) Design and Analysis of Experiments Volume I, John Wiley.

Reference Books:

1. Joshi D.D. (1987) Linear Estimation and Design of Experiments, Wiley Eastern.
2. Chakrabarti, M.C. (1964) Design of experiments, ISI, Calcutta.

STA2C07 – ESTIMATION THEORY

Number of Credits: 4

Course Outline

Unit-I: Sufficient statistics and minimum variance unbiased estimators.

Sufficient statistics, Factorisation theorem for sufficiency, Joint sufficient statistics, Exponential family, Pitman family, Minimal sufficient statistics (MSS). Criteria to find the MSS, Ancillary statistics, Complete statistics, Basu's theorem, Unbiasedness, Best Linear Unbiased estimator (BLUE), Minimum variance

unbiased estimator (MVUE), Rao-Blackwell theorem, Lehman-Scheffe theorem, Necessary and sufficient condition for MVUE, Fisher Information, Cramer Rao inequality and its applications.

Unit-II: Consistent estimator and Consistent asymptotically normal estimators

Consistent estimator, Invariance property of consistent estimator, Method of moments-method of percentiles to determine consistent estimators, Choosing between Consistent estimators. CAN estimators.

Unit-III: Methods of estimation

Method of moments-method of percentiles-method of maximum likelihood-MLE in exponential family-Cramer family, Cramer Huzurbazar Theorem, 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: Interval estimation

Definition - 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 in parametric inference, Second Edition, Narosa Publishing House, New Delhi.
2. George Casella and Roger L Berger (2002). Statistical inference, Second Edition, Duxbury, Australia.

Reference books:

1. Lehmann, E.L (1983). Theory of point estimation, John Wiley and sons, New York.
2. Rohatgi, V.K (1976). An introduction to Probability Theory and Mathematical Statistics, John Wiley and sons, New York.
3. Rohatgi, V.K (1984). Statistical Inference, John Wiley and sons, New York.
4. Rao, C.R (2002). Linear Statistical Inference and its applications, Second Edition, John Wiley and sons, New York.

STA2C08 – SAMPLING THEORY

Number of Credits: 4

Course Outline

Unit-I: Census and Sampling

Basic concepts, probability sampling and non-probability sampling, simple random sampling with and without replacement- estimation of population mean and total estimation of sample size- estimation of proportions.

Systematic sampling- linear and circular systematic sampling-estimation of mean and its variance- estimation of mean in populations with linear and periodic trends.

Unit-II: Stratification and stratified random sampling.

Optimum allocations, comparisons of variance under various allocations. Auxiliary variable techniques. Ratio method of estimation-estimation of ratio, mean and total. Bias and relative bias of ratio estimator. Mean square error of ratio estimator. Unbiased ratio type estimator. Regression methods of estimation. Comparison of ratio and regression estimators with simple mean per unit method. Ratio and regression method of estimation in stratified population.

Unit-III: Varying probability sampling

PPS sampling with and without replacements. Des- Raj ordered estimators, Murthy's unordered estimator, Horwitz-Thompson estimators, Yates and Grundy forms of variance and its estimators, Zen-Midzuno scheme of sampling, πPS sampling.

Unit-IV: Cluster sampling with equal and unequal clusters.

Estimation of mean and variance, relative efficiency, optimum cluster size, varying probability cluster sampling. Multi stage and multiphase sampling. Non-sampling errors.

Text books / References

1. Cochran W.G (1992): Sampling Techniques, Wiley Eastern, New York.
2. D. Singh and F.S. Chowdhary (1986): Wiley Eastern (New Age International), New Delhi.
3. P. V. Sukhatme et.al. (1984): Sampling Theory of Surveys with Applications. IOWA State University Press, USA.
4. Des Raj (1976): Sampling Theory. McGraw Hill
5. Mukhopadhyay. P. (1999). Theory and Methods of Survey Sampling. Printice-Hall India, New-Delhi.

STA2C09 – TESTING OF STATISTICAL HYPOTHESIS

Number of Credits: 4

Course Outline

Unit-I: Tests of hypotheses & Most Powerful Tests

Simple versus simple hypothesis testing problem – Error probabilities, p-value and choice of level of significance – Most powerful tests – Neyman Pearson Lemma – Generalized Neyman–Pearson Lemma, One-sided UMP tests, two- sided UMP tests and UMP unbiased tests.

Unit-II: UMP test for multi-parameter case

UMP unbiased test, α-similar tests and α-similar tests with Neyman structure, construction of α-similar tests with Neyman structure. Principle of invariance in testing of hypotheses, locally most powerful tests – Likelihood ratio tests – Bayesian tests.

Unit-III: Non-parametric Tests: Single sample tests

Testing goodness of fit, Chi-square tests- Kolmogorov– Smirnov test – sign test – Wilcoxon signed rank test. Two sample tests – the chi-square test for homogeneity – Kolmogorov – Smirnov test; the median test – Mann-Whitney-Wilcoxon test - Test for independence – Kendall's tau – Spearman's rank correlation coefficient – robustness.

Unit-IV: Sequential Tests

Some fundamental ideas of sequential sampling – Sequential Probability Ratio Test (SPRT) – important properties, termination of SPRT – the fundamental identity of SPRT – Operating Characteristic (OC) function and Average Sample Number (ASN) of SPRT – Developing SPRT for different problems.

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.
3. Manojkumar Srivastava and Namita Srivstava (2009): Statistical Inference: Testing of Hypothesis, Eastern Economy Edition, PHI Learning Pvt. Ltd., New Delhi.

Reference books:

1. Rohatgi,V.K. (1984). Statistical Inference, John-Wiley and Sons, New-York.
2. Lehman, E.L. (1983). Theory of Point Estimation, John-Wiley and Sons, New-York
3. Kale, B.K. (2005). A First Course on Parametric Inference. Second Edition, Narosa Publishing, New-Delhi.
4. Lehman E.L. and Romano, Joseph P. (2005). Testing Statistical Hypotheses. Third Edition, Springer, New York.

STA2C10 – STATISTICAL COMPUTING - II

Contact Hours per Week: 6 Hrs.

Number of Credits: 2

Course Outline

Statistical Computing-II is a practical course. Its objectives are to develop scientific and experimental skills

of the students and to correlate the theoretical principles with application based studies. The practical is based on the following FOUR courses of the second semester.

1. ST2C06: Design and Analysis of Experiments
2. ST2C07: Estimation Theory
3. ST2C08: Sampling Theory
4. ST2C09: Testing of Hypotheses

Practical is to be done by using R & Python. At least five statistical data oriented/supported problems should be done from each course. Practical Record shall be maintained by each student and the same shall be submitted for verification at the time of external examination.

The Board of Examiners (BoE) shall decide the pattern of question paper and the duration of the external examination. The external examination at each centre shall be conducted and evaluated on the same day jointly by two examiners – one external and one internal, appointed at the centre of the examination by the University on the recommendation of the Chairman, BoE. The question paper for the external examination at the centre will be set by the external examiner in consultation with the Chairman, BoE and the H/Ds of the centre. The questions are to be evenly distributed over the entire syllabus. Evaluation shall be done by assessing each candidate on the scientific and experimental skills, the efficiency of the algorithm/program implemented, the presentation and interpretation of the results. The valuation shall be done by the direct grading system and grades will be finalized on the same day.

SEMESTER - III

STA3C11 – APPLIED REGRESSION ANALYSIS

Number of Credits: 4

Course Outline

Unit-I:

Linear Regression Model, Least squares estimation, Gauss Markov Theorem, Properties of the estimates, Distribution Theory, Maximum likelihood estimation, Estimation with linear restrictions, Generalised least squares; Hypothesis testing - likelihood ratio test, F-test; Confidence intervals.

Unit-II:

Residual analysis, Departures from underlying assumptions, Effect of outliers, Collinearity, Non-constant variance and serial correlation, Departures from normality, Diagnostics and remedies.

Unit-III:

Polynomial regression in one and several variables, Orthogonal polynomials, Indicator variables, Subset selection of explanatory variables, stepwise regression and Mallows Cp -statistics, Introduction to non-parametric regression.

Unit-IV:

Introduction to nonlinear regression, Least squares in the nonlinear case and estimation of parameters, Models for binary response variables, estimation and diagnosis methods for logistic and Poisson regressions. Prediction and residual analysis, Generalized Linear Models – estimation and diagnostics.

Text Books:

1. Seber A.F. and Lee, A.J. (2003) Linear Regression Analysis, John Wiley, Relevant sections from chapters 3, 4, 5, 6, 7, 9, 10.
2. Montegomery D.C., Peck E.A. and Vining G.G. (2001) Introduction to Regression Analysis, Third edition. Wiley.
3. B. Abraham and Ledotter J. (1983) Statistical Methods for Forecasting, John Wiley & Sons.

Reference Books

1. Searle, S.R. (1971) Linear models, John Wiley & Sons, Inc.
2. N.Draper and H. Smith (1986) Applied Regression Analysis – John Wiley & Sons.
3. Fox, J. (1984) Linear Statistical Models and Related methods, John Wiley, Chapter 5.
4. Christensen, R. (2001) Advanced Linear Modeling, Chapter 7.

STA3C12 – STOCHASTIC PROCESS

Number of Credits: 4

Course Outline

Unit-I: Concept of Stochastic processes, examples, Specifications

Markov chains- Chapman Kolmogorov equations – classification of states – limiting probabilities; Gamblers ruin problem and Random Walk – Mean time spent in transient states – Branching processes (discrete time), 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

Limit theorems and their 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. Ixth Edition, Academic Press.
2. Medhi, J. (1996): Stochastic Processes. Second Editions. Wiley Eastern, New-Delhi.

References:

1. Karlin, S. and Taylor, H.M. (1975): A First Course in Stochastic Processes. Second Edition Academic Press. New-York.
2. Cinlar, E. (1975): Introduction to Stochastic Processes. Prentice Hall. New Jersey.
3. Basu, A.K. (2003): Introduction to Stochastic Processes. Narosa, New-Delhi.

STA3C13 – STATISTICAL COMPUTING - III

Contact Hours per Week: 6 Hrs.

Number of Credits: 4

Course Outline

Statistical Computing-I is a practical course. Its objectives are to develop scientific and experimental skills of the students and to correlate the theoretical principles with application based studies. The practical is based on the following THREE courses of the third semester.

1. STA3C11: Applied Regression Analysis
2. STA3E-- : Elective -I
3. STA3E-- : Elective -II

Practical is to be done by using R & Python. At least five statistical data oriented/supported problems should be done from each course. Practical Record shall be maintained by each student and the same shall be submitted for verification at the time of external examination.

The Board of Examiners (BoE) shall decide the pattern of question paper and the duration of the external examination. The external examination at each centre shall be conducted and evaluated on the same day jointly by two examiners – one external and one internal, appointed at the centre of the examination by the University on the recommendation of the Chairman, BoE. The question paper for the external examination at the centre

will be set by the external examiner in consultation with the Chairman, BoE and the H/Ds of the centre. The questions are to be evenly distributed over the entire syllabus. Evaluation shall be done by assessing each candidate on the scientific and experimental skills, the efficiency of the algorithm/program implemented, the presentation and interpretation of the results. The valuation shall be done by the direct grading system and grades will be finalized on the same day.

SEMESTER - IV

STA4C14 – MULTIVARIABLE ANALYSIS

Number of Credits: 4

Course Outline

Unit-I: Multivariate Normal Distribution

Definition and 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 mean vector when the dispersion matrix is known for single and two populations. Distribution of simple, partial and multiple (null-case only) correlation coefficients; canonical correlation. Wishart distribution – properties – generalized variance.

Unit-III: Testing Problems

Mahalanobis D2 and Hotelling's T2 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; Factor analysis.

Text Books:

1. Anderson T.W. (1984): Multivariate Analysis. John – Wiley, New York.
2. Johnson R.A. and Wichern, D.W. (2001): Applied multivariate statistical analysis, 3rd Edn., Prentice Hall of India, New Delhi.

3. Rao C. R. (2002): Linear Statistical Inference and Its Applications, Second Edition, John Wiley and Sons, New York.

References:

1. Giri, N.C. (1996): Multivariate Statistical Analysis. Marcel Dekker. Inc., New York.
2. Kshirasagar, A.M. (1972): Multivariate Analysis. Marcel Dekker. New-York
3. Rencher, A.C. (1998): Multivariate Statistical Analysis. Jon Wiley, New York.
4. Morrison, D.F. (1976): Multivariate statistical methods, McGraw Hill, New York.

STA4C15 – PROJECT/DISSERTATIONAND VIVA-VOCE

Number of Credits: 4

Course Outline

(5 credits for Project/Dissertation and 3 credits for Viva-Voce)

In partial fulfillment of the M.Sc. programme, 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 5 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 Expert appointed by the University, 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: 20%
3. Analysis, programming/simulation and discussion of results: 20%
4. Presentation of the report, organization, linguistic style, reference etc.: 20%
5. Viva-voce examination based on project/dissertation: 20%.

The Viva-Voce shall be conducted a Board of Examiners, consisting of at least two external experts. The viva-voce shall cover all the courses undergone in the two-year programme and carries 3 credits.

STA4C16 – STATISTICAL COMPUTINDG - IV

Contact Hours per Week: 6 Hrs.

Number of Credits: 2

Course Outline

Statistical Computing-I is a practical course. Its objectives are to develop scientific and experimental skills of the students and to correlate the theoretical principles with application based studies. The practical is based on the following TWO courses of the fourth semester.

1. ST4C14: Multivariate Analysis
2. ST4E-- : Elective -III

Practical is to be done by using R & Python. At least five statistical data oriented/supported problems should be done from each course. Practical Record shall be maintained by each student and the same shall be submitted for verification at the time of external examination.

The Board of Examiners (BoE) shall decide the pattern of question paper and the duration of the external examination. The external examination at each centre shall be conducted and evaluated on the same day jointly by two examiners – one external and one internal, appointed at the centre of the examination by the University on the recommendation of the Chairman, BoE. The question paper for the external examination at the centre will be set by the external examiner in consultation with the Chairman, BoE and the H/Ds of the centre. The questions are to be evenly distributed over the entire syllabus. Evaluation shall be done by assessing each candidate on the scientific and experimental skills, the efficiency of the algorithm/program implemented, the presentation and interpretation of the results. The valuation shall be done by the direct grading system and grades will be finalized on the same day.

ELECTIVE COURSES
E02 – TIME SERIES ANALYSIS

Number of Credits: 4

Course Outline

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, Use of computer packages like SPSS.

Unit-IV. Spectral analysis of weakly stationary process.

Herglotzic Theorem. Periodogram and correlogram analysis. Introduction to non-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
2. Brockwell P.J.and Davis R.A. (1987). Time Series: Theory and Methods, Springer – Verlag.
3. Abraham B and Ledolter J.C. (1983). Statistical Methods for Forecasting, Wiely

References:

1. Anderson T.W (1971). Statistical Analysis of Time Series, Wiely.
2. Fuller W.A. (1978). Introduction to Statistical Time Series, John Wiley.
3. Kendall M.G. (1978), Time Series, Charler Graffin
4. K.Tanaka (1996). Time Series Analysis – Wiely Series.

E08 – RELIABILITY MODELING

Number of Credits: 4

Course Outline

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.

Text Books / References:

1. Barlow R.E. and Proschan F. (1985). Statistical Theory of Reliability and Life Testing; Holt, Rinehart and Winston.

2. Bain L.J. and Engelhardt (1991). Statistical Analysis of Reliability and Life Testing Models; Marcel Dekker.

3. Aven T. and Jensen U. (1999). Stochastic Models in Reliability, Springer-Verlag, New York, Inc. Lawless, J.F. (2003). Statistical Models and Methods for Lifetime (Second Edition), John Wiley & Sons Inc., New Jersey.

4. Nelson, W (1982) Applied Life Data analysis; John Wiley.

5. Zacks, S. (1992). Introduction to Reliability Analysis: Probability Models and Statistics Methods. New York: Springer-Verlag.

E10 – STATISTICAL QUALITY CONTROL

Number of Credits: 4

Course Outline

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 the number of non- conformities. Mean charts. Median charts. Extreme value charts, R-charts, and S-charts ARI, Economic design of control charts.

Unit-IV.

Basic concepts of process monitoring and control; process capability and process optimization. Control charts with memory – CUSUM charts, EWMA mean charts, OC and ARl 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. 4th edition. Wiley, New-York.
2. Mittage H. J. and Rinne H. (1993). Statistical Methods for Quality Assurance. Chapman and Hall. Chapters 13 and 14.
3. Oakland J. S. and Follorwel R. F. (1990). Statistical Process Control. East-West Press. Chapters 13 and 14.
4. Schilling E.G. (1982). Acceptance Sampling in Quality Control. Marcel Dekker.
5. Duncan A. J. (1986). Quality Control and Industrial Statistics.

References:

1. Gerant E. L. and Leaven Worth R.S. (1980). Statistical Quality Control. Mc-Graw Hill
2. Chin-Knei Chao (1987). Quality Programming, John Wiley.
3. Ott E.R. (1975): Process Quality Control; McGraw Hill.
4. Wetherill G.B. and Brown D.W ( ).: Statistical Process Control: Theory and Practice.

No. 25539/GA - IV - J2/2013/CU

Calicut University.P.O

Dated: 06.07.2023

From

The Registrar

To

The Principal,

Christ College (Autonomous),

Irinjalakkuda

Sir/Madam,

Sub:- Christ College (Autonomous), Irinjalakkuda - Syllabus of two elective courses for M.Sc Statistics Programme from 2022 admission onwards - Approval - Intimating - Reg

Ref:- 1) Syllabus of two elective courses for M.Sc Statistics Programme from 2022 admission onwards submitted by your college

2) Item No. 1 in the minutes of the meeting of Board of Studies in Statistics (PG) held on 22.06.2023

With reference to the above, I am to inform you that the Board of Studies in Statistics (PG) at its meeting held on 22.06.2023, considered the syllabus of two elective courses for M.Sc Statistics Programme submitted from your college (2022 admission onwards) and approved the same vide item No.1.

The matter is intimated to you for favour of information and necessary action.

Encl : Syllabus under reference

Yours faithfully

Ajitha P.P

Joint Registrar

(For The Registrar)

E 19 STATISTICAL MACHINE LEARNING

Number of Credits 4

Module I Introduction to Python

Variables-Control statements-generating sequence numbers-control flow statements-functions-Collections(List, Tuples, Set, Dictionary)-strings-Functional Mapping(Map, Filter)-Modules and Pacages-Working with Data Frames in Python-Features of Pandas Data Frame-Handling missing values-Exploring of Data using visualization   (10 Hours)

Module II Data Preprocessing and Exploratory Data Analysis

Data Preprocessing: Data Cleaning-Handling Missing Data-Identifying Misclassifications-Graphical Methods for Identifying Outliers-Data Transformation-Normalization-Numerical Methods for Identifying Outliers-Flag variables-Transforming Categorical Variables into Numerical Variables - Binning Numerical Variables - Reclassifying Categorical Variables-Exploratory Data Analysis: Exploring Categorical Variables - Exploring Numeric Variables - Exploring Multivariate Relationships -Selecting Interesting Subsets of the Data for Further Investigation-Uncover Anomalous fields-Binning based on predictive value   (20 Hours)

Module III: Statistical Learning

What Is Statistical Learning? - The Trade-Off Between Prediction Accuracy and Model Interpretability - Supervised Versus Unsupervised Learning - Regression Versus Classification - Assessing Model Accuracy   (15 Hours)

Module IV: Linear Regression and Classification and Resampling Methods

Simple Linear Regression - Estimating the Coefficients- Assessing the Accuracy of the Coefficient Estimates - Multiple Linear Regression - Estimating Coefficients – Classification -KNN Algorithm-Decision Trees-Logistic Regression - Estimating Coefficients – Predictions-Multiple Logistic Regression- Logistic Regression for more than two   Response Classes Linear   Discriminant Analysis- Quadratic Discriminant Analysis Generative Models for Classification - Generalized Linear Models -Cross-Validation - k-Fold Cross-Validation - Bias-Variance Trade-Off for k-Fold Cross-Validation - Cross-Validation on Classification Problems -The Bootstrap   (30 Hours

Module V: Survival Analysis and Censored Data

Survival and Censoring times - Kaplan-Meier Survival Curve - Regression Models With a Survival Response   (15 Hours)

Textbooks
1. Larose, D.T., 2015. Data Mining and Predictive Analytics. John Wiley & Sons.
2. James, G., Witten, D., Hastie, T., Tibshirani, R., James, G., Witten, D., Hastie, T. and Tibshirani, R., 2021. An introduction to statistical learning: with applications
3. Manaranjan Pradhan, Dinesh Kumar U 2019 Machine Learning using Python, Wiley

Reference Books
1. Hastie, T., Tshibirai, R. and Friedman, J. (2017) The Elements of Statistical Learning: Data Mining, Inference and Prediction,
2. Mitchell, T.M., 2007. *Machine learning* (Vol. 1). New York: McGraw-hill.

E 20 ADVANCED STATISTICAL MACHINE LEARNING TECHNIQUES

Number of Credits 4

Module 1: Model Selection, Regularization and Nonlinearity

Subset Selection – Best subset selection-stepwise selection-optimum model-Shrinkage Methods – Ridge regression-The Lasso-selection of Tuning parameter-Polynomial Regression – step function-basic function-Regression Splines- Smoothing Splines - Generalized Additive Models   (15 Hours)

Module 2: Tree-based Methods and Support Vector Machines

The Basics of Decision Trees - Regression Trees - Classification Trees - Trees Versus Linear Models - Advantages and Disadvantages of Trees – Bagging-Random Forests- Boosting-Bayesian additive regression trees-Summary of tree ensemble methods- Maximal Margin Classifier-hyperplane-classification using a separating hyperplane-construction of maximal Margin Classifier-support vector classifier-Support vector machine-Classification with non linear decision boundaries-SVM with more than two classes   (25 Hours)

Module 3: Neural Networks and Deep Learning

Input and Output Encoding-Neural Networks for Estimation, Prediction-Simple Example of a Neural Network Single Layer Neural Networks-Multilayer Neural Networks - Convolutional Neural Networks – Document classification-Recurrent Neural networks-When to Use Deep Learning - Fitting a Neural Network- Interpolation and Double Descent   (20 Hours)

Module 4: Unsupervised Learning: Clustering

The Challenge of Unsupervised Learning - Principal Components Analysis – Missing values and Matrix completion-Clustering Methods - K-Means Clustering - Hierarchical Clustering Single-Linkage Clustering -Complete-Linkage Clustering methods   (15 Hours)

Module 5: Association Rules

Affinity Analysis - Market Basket Analysis - Data Representation for Market Basket Analysis - Support, Confidence, Frequent Item sets, and the a Priori Property – working of Priori Algorithm in Generating Frequent Itemset and in Generating Association Rules - Measure the Usefulness of Association Rules   (15 Hours)

Textbooks

1. Larose, D.T., 2015. Data mining and predictive analytics. John Wiley & Sons.
2. James, G., Witten, D., Hastie, T., Tibshirani, R., James, G., Witten, D., Hastie, T. and Tibshirani, R., 2021. An introduction to statistical learning: with applications

Reference Books

1. Hastie, T., Tshibirai, R. and Friedman, J. (2017) The Elements of Statistical Learning: Data Mining, Inference and Prediction,
2. Mitchell, T.M., 2007. Machine learning (Vol. 1). New York: McGraw-hill.
3. Shmueli, G., Bruce, P.C., Yahav, I., Patel, N.R. and Lichtendahl Jr, K.C., 2017. Data mining for business analytics: concepts, techniques, and applications in R. John Wiley & Sons.
4. Feldman, R. and Sanger, J., 2007. The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge university press.