University of Lucknow
Master of STATISTICS Programme
Regulations 2020

1. Applicability

These regulations shall apply to the Master in Statistics Programme from the session 2020-21.

2. Minimum Eligibility for admission

A three/four-year Bachelor's degree or equivalent degree awarded by a University or Institute established as per law and recognised as equivalent by this University with minimum 45% marks or equivalent grade, shall constitute the minimum requirement for admission to the Master in Statistics programme. (any other additional requirement may also be specified)

3. Programme Objectives

Preamble: M. A. / M. Sc. Statistics program is of minimum 96 credits spread over four semesters. The program emphasizes both theory and applications of statistics and is structured to provide knowledge and skills in depth necessary for the employability of students in industry, other organizations, as well as in academics. The program has some unique features such as independent projects, number of elective courses, extensive computer training of statistical computations including standard software packages such as STATA, R, and SPSS. The department has the academic autonomy and it has been utilized to add the new and need based elective courses. The independent project work (Master Thesis) is one of the important components of this program. In all semesters I, II, III & IV some courses are compulsory and others are elective. The syllabus has been framed to have a good balance of theory, methods and applications of statistics.

It is possible for the students to study basic courses from other disciplines such as economics, life sciences, computer science and mathematics in place of electives.

4. Programme Outcomes

The degree focuses on algorithms, designs and development to give solutions to real life problems. It is designed to develop student's knowledge in subjects like Probability, Bayesian Inference, Multivariate Analysis, Stochastic, Design of Experiments, Sampling, Linear Models etc. that will make them employable in Multi-National companies, Railways, MOSPI, etc. specializing in Data Analytics, Business insurance and finance. Extensive knowledge of a plethora of Statistical Software namely R, STATA, SPSS, TEXT ANALYTICS etc. are also necessitated under its preview. The Key Program Outcomes are:

1. To inculcate and develop aptitude to understand and apply statistical tools at a number of data generating fields in real life problems.
2. To train students to handle large data sets and carry out data analysis using software and programming language.
3. To teach a wide range of statistical skills, including problem-solving, project work and presentation so as enable students to take prominent roles in a wide spectrum of employment and research.

5. Specific Programme Outcomes

At the *Department of Statistics, University of Lucknow* the course matter has been specifically moulded towards targeted achievements in the advanced fields of Indian Statistical Services (ISS), Civil Services Examination, UPPSC, etc. Illustrious achievement of at least four to five students in these fields, from the Department every year, lend itself as veritable proof. On successful completion of the course a student will be able to:

1. Gain sound knowledge in theoretical and practical aspects of Statistics.
2. Describe complex statistical ideas to non-statisticians.
3. Handle and analyze large databases with computer skills and use their results and interpretations to make practical suggestions for improvement.
4. Get wide range of job opportunities in industry as well as in government sector.

Department of Statistics

University of Lucknow

Syllabus for M.A./M.Sc. Statistics Programme

(Proposed to be implemented from July 2020)

Course No.	Name of the Course	Credits	Remark
	Semester I		
STATCC-101	Real and Complex Analysis	04	Core Course
STATCC -102	Linear Algebra	04	Core Course
STATCC -103	Measure Theory & Probability	04	Core Course
STATCC -104	Sample Surveys	04	Core Course
STATCC -105	Practical	04	Core Course
STATVC-101	Statistical Methodology & Data Mining	04	Value added course (Credited)
	Semester Total	24	
	Semester II		
STATCC -201	Linear Models and Regression Analysis	04	Core Course
STATCC -202	Inference	04	Core Course
STATCC -203	Multivariate Analysis	04	Core Course
STATCC -204	Stochastic Processes	04	Core Course
STATCC -205	Sequential Analysis and Reliability Theory	04	Core Course
STATCC -206	Practical	04	Core Course
STATVNC-201	Official Statistics	00	Value added course (Non Credited)
	Semester Total	24	
	Semester III		
STATCC -301	Block Designs and Their Analysis	04	Core Course / MOOC
STATCC -302	Practical	04	Core Course
STATEL - 301A	Econometrics	04	Elective
STATEL - 301B	Actuarial Statistics		
STATEL-302A	Advanced Statistical Analysis Using SPSS	04	Elective
STATEL-302B	Qualitative Data Analysis		
STATIN-301	Summer Internship	04	Summer Internship
STATIER-301	Applied Statistics	04	Interdepartmental Course
	Semester Total	24	
	Semester IV		
STATCC -401	Decision Theory and Bayesian Analysis	04	Core Course
STATEL - 401A	Factorial Experiments and Response surfaces	04	Elective
STATEL - 401B	Ethics, Integrity and Aptitude		
STATEL-402A	Data Analysis Using R	04	Elective
STATEL-402B	Population Studies		
STATMT-401	Master Thesis	08	Master Thesis
STATIRA-401	Data Analysis Using STATA	04	Intradepartmental Course
	Semester Total	24	
	GRAND TOTAL	96	

Detailed Syllabus:

SYLLABUS OF M.A./M.Sc. (Statistics)

SEMESTER I

STATCC-101 : REAL AND COMPLEX ANALYSIS
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The main objective of this course is to introduce students the knowledge of real field and complex field with their properties and relativity between complex plane and real line. These properties and relations provide grounds for Probability Theory and help in theoretical research in Statistics.

Course Specific Outcomes:

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

1. Understand existence of integral and their evaluation.
2. Understand convergence of sequence and series of real valued function and complex valued functions.
3. Understand change of multiple integral into line integral.
4. Find maxima-minima of functions of several variables.
5. Understand complex region and relativity between complex plane and real line.
6. To solve contour integrals.
7. Find residue at singularity and infinity via definition and via Cauchy integral formula.

UNIT I

Introduction to real numbers, open and closed intervals, compact sets, Bolzano-Weirstrass theorem, open, closed sets, Heine - Borel theorem. Maxima and Minima of functions, constrained maxima and minima of functions of several variables.

UNIT II

Real valued functions, Continuous functions, Power series and Radius of convergence, Riemann integration, Mean value theorem of Integral Calculus, Multiple Integrals and their evaluation by repeated integration.

UNIT III

Improper integrals and their convergence, uniform convergence of improper integrals, convex functions and their properties, Jensen's, Minkowski's and Holder's inequalities.

UNIT IV

Functions of complex variables, Limit, continuity, Differentiation, Cauchy-Reimann equations, Analytic functions, Cauchy's theorem and Integral formula.

UNIT V

Taylors and Haurent's series, Residue theorem, Evaluation of standard integrals by Contours Integration. Laplace Transform and its properties, inverse Laplace transform, Convolution Theorem.

REFERENCES:

1. Apostol, T.M. (1985): Mathematical Analysis, Narosa, Indian Ed.
2. Courant, R. and John, F. (1965): Introduction to Calculus and Analysis, Wiley.
3. Miller, K.S. (1957): Advanced Real Calculus, Harper, New York.
4. Rudin, Walter (1976): Principles of Mathematical Analysis, McGraw Hill.

STATCC -102 : LINEAR ALGEBRA
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:
The main objective of this paper is to allow students to manipulate and understand multidimensional space.

Course Specific Outcomes:
After completing this course students will have a clear understanding of:
1. Whole system of equations with multiple dimensions/variables.
2. Importance of concept of linear algebra in multiple area of science.
3. Concepts of Generalized inverse theory and applications.
4. Concepts of Linear Transformations and inner product spaces.
5. Concepts and detailed theory of Eigen values and Eigen vectors.
6. Concepts of Quadratic equations.

UNIT - I
Fields, vector spaces, subspaces, linear dependence and independence, basis and dimension of a vector space, Finite dimensional vector spaces, examples of vector spaces over real and complex fields, Vector space with an inner product, Schwartz's inequality, Gram-Schmidt orthogonalization process, orthonormal basis and orthogonal projection of a vector

UNIT - II
Linear transformation, equality of linear transformations, Linear operator, algebra of matrices, row and column spaces of a matrix, elementary matrices, determinant of a square matrix, null space and nullity, symmetric matrices, Idempotent matrices, Hermitian matrices, Kronecker product.

UNIT - III
Rank and inverse of a matrix, partitioned matrices, their use in finding the inverse of matrix, generalized inverse, important results on generalized inverse, Solutions of simultaneous linear equations using matrices.

UNIT - IV
Characteristic roots and vectors, Caley-Hamilton theorem, minimal polynomial, similar matrices, algebraic and geometric multiplicity of a characteristic root, triangular reduction of a positive definite matrix, Reduction of a pair of matrices,
Real quadratic forms, reduction of quadratic forms, index and signature, classification of quadratic forms, Extrema of quadratic forms.

UNIT-V
Spectral decomposition of matrices: Eigen Value Decomposition, Singular values and singular value decomposition, Jordan decomposition. Vector differentiation, matrix differentiation

REFERENCES:

1. Graybill, F.A. (1983): Matrices with applications in Statistics, 2nd Ed. Wadsworth.
2. Rao C.R. (1973): Linear Statistical inference and its applications 2nd ed. John Wiley & Sons, Inc.

3. Searle S.R.(1982): Matrix Algebra Useful For Statistics, John Wiley & Sons, Inc

ADDITIONAL REFERENCES:

1. Bellman R. (1970): Introduction to Matrix Analysis, 2nd ed. Mc Graw Hill.
2. Biswas, S. (1984): Topics in Algebra of Matrices, Academic Publications.
3. Hadley G. (1987): Linear Algebra: Narosa Publishing House.
4. Halmos, P.R. (1958): Finite-dimensional Vector Spaces, 2nd ed. D. Van Nostrand Company, Inc.
5. Hoffman K. & Kunze, R. (1971): Linear Algebra, 2nd ed., Prentice Hall, Inc.
6. Rao A.R. and Bhimsankaran P. (1992): Linear Algebra, Tata McGraw Hill Publishing Company Ltd.
7. Rao CR and Mitra S.K. (1971): Generalized Inverse of Matrices and Its Applications, John Wiley & Sons, Inc.

STATCC -103: MEASURE THEORY AND PROBABILITY

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The aim of the course is to pay a special attention to applications of measure theory in the probability theory, understanding of Weak Law of Large Numbers, Strong Law of Large Numbers and the Central Limit Theorem with their applications.

Course Specific Outcomes:

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

1. Understand the concepts of random variables, sigma-fields generated by random variables, probability distributions and independence of random variables related to measurable functions.
2. Gain the ability to understand the concepts of measurable functions, sequence of random variables, convergence, modes of convergence.
3. Learn the concepts of weak and strong laws of large numbers and central limit theorem.

UNIT – I

Sets and sequences of sets, Rings, Fields, sigma Field and minimal sigma field. Monotone classes. Elements of Measure theory. Probability Measure and properties of measure. Measurable function and their properties.

UNIT II

Sequence of random variables, convergence in probability, almost sure convergence, convergence in distribution and r^{th} mean convergence. Interrelationship among different mode of convergences. Dominated convergence theorem.

UNIT III

Weak and Strong Law of large numbers for independent random variables. Kolmogrov's inequality and theorem. Central Limit Theorem. CLT for i.i.d. random variables, Khinchine's theorem. Central Limit theorem: Lindberg Levy and Liapunoff's form of CLT.

UNIT IV

Distribution function. Lebsegue and Lebsegue – Stiljes and Riemann Integral. Characteristic function. Inversion theorem, continuity theorem and its applications.

UNIT V

Product Space, Radon – Nikodymn Theorem (Without proof). Conditional Expectation. Borel – Cantelli Lemma, Borel Zero – One – Law. Signed Measure and Absolute Continuity.

REFERENCES:

1. Dudewicz, E.J. and Mishra, S.N. (1988): Modern Mathematical Statistics, Wiley, Int'l Students' Edition.
2. Rao C.R. (1973): Linear Statistical Inference and its Applications, 2/e, Wiley Eastern.
3. Rohatgi, V.K. (1984): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.

ADDITIONAL REFERENCES

1. Ash, Robert (1972): Real analysis and Probability, Academic Press.
2. Bhat, B.R. (1985): Modern Probability Theory, Wiley Eastern Limited.
3. Billingsley, P. (1986): Probability and Measure, Wiley.
4. Chow, Y.S. & Teicher, H. (1979) : Probability Theory, Narosa Publishing House, New Delhi
5. Cramer H. (1946): Mathematical Methods of Statistics, Princeton.
6. Dudley, R.M. (1989): Real Analysis and Probability, Wadsworth and Brooks/Cole.
7. Johnson, S. and Kotz, (1972): Distributions in Statistics, Vol. I, II and III, Houghton and Miffin.
8. Kingman, J.F.C. and Taylor, S.J. (1966): Introduction to Measure and Probability, Cambridge University Press.
9. Pitman, J. (1983): Probability, Narosa Publishing House.

STATCC -104: SAMPLE SURVEYS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The main objective of this course is to learn techniques in survey sampling with practical applications in daily life which would be beneficial for the students to their further research.

Course Specific Outcomes:

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

1. Understand the distinctive features of sampling schemes and its related estimation problems
2. Learn about various approaches (design based and model-based) to estimate admissible parameters; with and without replacement sampling scheme, sampling with varying probability of selection.
3. Learn about the methods of post–stratification (stratified sampling) and controlled sampling and also double sampling procedure with unequal probability of selection.
4. Learn about the applications of sampling methods; systematic, stratified and cluster sampling.
5. Understand the cluster and two stage sampling with varying sizes of clusters/first stage units.
6. Understand the super population approach to estimation.
7. Learn about the randomized response techniques.

UNIT I

Unequal probability sampling: PPS WR and WOR methods (including Lahiri's scheme) and related estimators of a finite population mean (Hansen-Hurwitz and Desraj estimators for general sample size and Murthy's estimator for a sample of size 2).

UNIT II

Horvitz-Thompson estimator, its variance and unbiased estimator of variance, IPPS schemes of sampling due to Midzuno-Sen, Rao-Hartley-Cochran and Samphord.

UNIT III

The Jackknife estimation: estimate of bias, estimate of variance; Ratio Estimation in reference to Jackknife

UNIT IV

The Bootstrap: estimate of bias, estimate of variance, Ratio estimation in reference to bootstraps, Relationship between the jackknife and the bootstrap. Interpenetrating sub sampling

UNIT V

Non-sampling errors. Randomized Response techniques (Warner's method: related and unrelated questionnaire methods).

REFERENCES:

1. Chaudhuri, A. and Mukerjee, R. (1988): Randomized Response: Theory and Techniques, New York: Marcel Dekker Inc.
2. Cochran, W.G.: Sampling Techniques (3rd Edition, 1977). Wiley.
3. Des Raj and Chandak (1998): Sampling Theory, Narosa.
4. Gray, H.L., and Schucany (1972): The generalized jackknife statistic. New York. Marcel Dekker, Inc.
5. Murthy, M. N. (1977): Sampling Theory & Methods, Statistical Publishing Society, Calcutta.
6. Singh, D. and Chaudhary, F.S. (1986): Theory and Analysis of Sample Survey Designs. New Age International Publishers.
7. Sukhatme et al (1984): Sampling Theory of Surveys with Applications. lowa state University Press & IARS.

STATCC -105 : PRACTICAL
(4 Credits – 6 hours of T/L per week)

1. Rank of matrix
2. Calculation of bases
3. Gram-Smidt orthonormalization
4. Inverse of matrix
5. Solution of a set of homogeneous and non-homogeneous equations
6. g-inverse of matrix
7. Characteristics roots and vectors
8. Reduction and classification of quadratic forms
9. Eigen value decomposition and singular value decomposition
10. Based on pps.
    a. to draw samples by cumulative total method/Lahiri method
    b. to estimate population mean/population total of the characteristics under study using ordered and unordered samples : Desraj, Murthy and H-T estimators.

STATVC-101: Credited: STATISTICAL METHODOLOGY AND DATA MINING
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This course will present statistical methods that have proven to be of value in the field of knowledge discovery in databases, with special attention to techniques that help managers to make intelligent use of these repositories by recognizing patterns and making predictions.

Course Specific Outcomes:

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

- Correctly plan a data mining process
- Choose the best suited methodology for the problem at hand
- Critically interpret the results

UNIT I

Truncated distributions, Compound distributions, Mixture of distributions, Generalized power series distributions, Exponential family of distributions, Non-central distributions of Beta, Chi-square, t and F and their properties.

UNIT II

Introduction to databases, tasks in building a data mining and database, goal of data mining, data warehouses, applications of data mining, supervised and unsupervised learning, data processing: representation, visualization, cleaning, reduction, transformation, outlier detection.

UNIT III

Clustering: Similarity and distance measures, squared error clustering, single linked clustering, centroid clustering, K-means clustering, Hierarchical clustering, Block clustering, Support vector machine (SVM) with linear class boundaries, multiclass SVM, Latent variable models for blind source separation: Independent component analysis (ICA) and its applications.

UNIT IV

Classification and Regression Trees (CART): Classification trees, Minimum spanning tree, node impurity function and entropy function, choosing the best split pruning algorithm for classification trees. Regression trees, terminal node value and splitting strategy, pruning the tree and best pruned subtree.

UNIT V

Artificial neural networks: Introduction, Rosenblatt's Single layer perceptron, single unit perceptron gradient descent learning algorithm, multilayer perceptron network, feed forward and back propogation learning algorithm, McCullon-Pitts Neuron (Threshold Logic Unit), self-organizing feature map (SOM) or Kohonen neural network, extensions of regression models.

REFERENCES:

1. Bishop, C.M. (1995): Neural Networks for pattern Recognition, Oxford University Press.

2. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984): Classification and Regression Trees. Wadsworth and Brooks.
3. Dunham, M. H. (2003). Data Mining: Introductory and Advanced Topics, Pearson Education.
4. Han, J. (2000): Data Mining: Concepts and Techniques. Morgan Kaufmann.
5. Han, J. and Kamber, M (2006). Data Mining: Concepts and Techniques, 2nd edition, Morgan Kaufmann.
6. Hand, D., Mannila, H., and Smyth, P. (2001). Principles of Data Mining, MIT Press.
7. Hastie T., Tibshirani R. and Friedman J. H., (2008). The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer.
8. Haykin, S. (1998) Neural Networks: A Comprehensive Foundation, 2nd ed., Prentice Hall.
9. James G., Witten, D., Hastie, T. Tibshirani, R. (2013). An Introduction to Statistical Learning: With Applications in R, Springer
10. Mohammad J. Zaki and Wagner Meira. (2014). Data Mining and Analysis. Fundamental Concepts and Algorithms. Cambridge University Press, New York.
11. Nisbet, R., Miner, G. and Elder, J (2009): Handbook of Statistical Analysis and Data Mining Applications. Academic Press.
12. Shmueli, G., Patel, N. Bruce, P. (2010). Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XL Miner, Wiley.
13. Yang, Xin-She (2019): Introduction to Algorithms for Data Mining and Machine Learning. Academic Press.

SEMESTER II

STATCC -201: LINEAR MODELS AND REGRESSION ANALYSIS
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:
The main objective of this course is to provide students the ability to learn and use linear and non-linear models for normal data, and generalized linear models for normal and non-normal responses.

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

1. Use linear and nonlinear models, apply data transformations, and appreciate the need and uses of generalized linear models.
2. Use logistic and Poisson regression models.
3. Understand deviance, analysis of deviance, Lack-of-Fit tests in Logistic and Poisson regression, and the concept of over dispersion.
4. Use Log linear models for contingency tables, and likelihood ratio tests for various hypotheses including independence, marginal and conditional independence, and partial association.
5. Understand graphical and non-graphical models.
6. Use the concepts of Generalized Linear Models in real life problems.
7. Understand and apply Quasi likelihood.

UNIT I

Generalized inverse, Moore-Penrose generalized inverse. Important results on g-inverse, Use of generalized inverse of matrices, Distribution of quadratic forms for multi-variate normal random vector, Cochran Theorem.

UNIT II

Linear models of full rank, Assumptions for linear models, Estimation of model parameters: least squares method, Likelihood method, Properties of the estimators, ANOVA for full rank model distribution of different sum of squares

UNIT III

Linear models not of full rank: Estimation for model parameters: least squares method, Likelihood method, Estimation and error spaces, estimable functions and their BLUE (Gauss-Markov Theorem) , ANOVA for not of full rank model, distribution of different sum of squares, test for general linear hypothesis.

UNIT IV

Analysis of Variance Models: One way classification, Two way classification, Multiple comparison tests due to Tukey and Scheffe, Simultaneous confidence intervals.

UNIT - V

Selecting the best regression equation: Forward selection method, backward elimination method and Stepwise regression Regression Diagnostics: Residuals and their plots, Tests for departure from assumptions of linear models: normality, homogeneity of variances, Multicollienarity and autocorrelation, Transformation: Box - Cox transformation.

REFERENCES:

1. Cook, R.D. and Weisberg, S. (1982): Residual and Influence in Regression. Chapman and Hall.
2. Draper, N.R.and Smith, H. (1998): Applied Regression Analysis, Third Edition Wiley.
3. Guest, R.F. and Mason, R.L. (1980): Regression analysis and its Applications - A Data Oriented Aproach. Marcel and Dekker.
4. Rao, C.R. (1973): Linear statistical inference and its Applications. Wiley Eastern.
5. Weisberg, S. (1985): Applied Linear Regression. Wiley.

STATCC -202: INFERENCE

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

To make aware the students of parametric, non-parametric and sequential estimation (point, as well as, interval) and testing (simple, as well as, composite hypotheses) procedures.

Course Specific Outcomes:

After successful completion of this course, student will be able to:
Apply various estimation techniques and testing procedures to deal with real life problems.
Understand consistency, CAN estimator, MLE.
Understand UMPU tests UMVU estimators.

UNIT – I

Consistency (mean squared and weak). Invariance property of consistency. Asymptotic properties of Consistent estimators: CAN estimators (single as well as multi-parameter cases), invariance of CAN estimators under differentiable transformations, generation of CAN estimators using central limit theorem.

UNIT – II

Sufficiency, Factorization Theorem, Minimal sufficient statistic, Completeness and bounded completeness. Distribution admitting sufficient Statistics, Extension of results to multi-parameter case. Bhattacharya's bounds, Rao-Blackwell Theorem, Lehman-Scheffe theorem and their applications.

UNIT – III

Minimum Chi square estimators and their modification and their asymptotically equivalence to maximum likelihood estimators. Computational routines: Newton – Raphson method, method of scoring, Consistency and inconsistency, Cramer Huzurbazar Theorem, Asymptotic efficiency of ML estimators, Best Asymptotically normal estimators.

UNIT – IV

Non-randomized and randomized tests. Generalized NP lemma. MLR families. Unbiased tests. UMPU tests against one-sided and two-sided alternatives. Locally most powerful unbiased test (type A test). Similar tests. Neymann structure.
Confidence set estimation, Relation with hypothesis testing, optimum parametric confidence sets.

UNIT – V

Large Sample tests: Likelihood ratio (LR) test, asymptotic distribution of LR statistic, Tests based on ML estimators: Pearson's chi-square test for goodness of fit and its relation to LR Test, Test consistency, Asymptotic power of test, Generalized likelihood ratio test, special cases such as multinomial distribution and Bartlett's test for homogeneity of variances.

REFERENCES

Kale, B. K.(1999) : A first course on parametric inference, Narosa Publishing House.
Rohatagi, V. (1988): An Introduction to probability and mathematical Statistics . Wiley Eastern Ltd. New Delhi (Student Edition)

ADDITIONAL REFERENCES

Lehmann, E.L.(1986) : Theory of point. Estimation (Student Edition)

Lehmann, E.L.(1986) : Testing statistical hypotheses (Student Edition).

Rao, C.R. (1973) : Linear Statistical inference.

Dudewicz, E.J. and Mishra, S.N. (1988) : Modern Mathematical Statistics. Wiley Series in Prob. Math. Stat., John Wiley and sons, New York (International Student Edition).

Ferguson T. S. (1967) : Mathematical Statistics. Academic press.

Zacks, S. (1971) : Theory of statistical Inference, John Wiley & Sons, New York.

STATCC -203: MULTIVARIATE ANALYSIS
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The main objective of this course is to introduce students to the analysis of observations on several correlated random variables for a number of individuals. Such analysis becomes necessary in Anthropology, Psychology, Biology, Medicine, Education, Agriculture and Economics when one deals with several variables simultaneously.

Course Specific Outcomes:

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

1. Account for important theorems and concepts in multivariate analysis.
2. Summarize and interpret multivariate data.
3. Appreciate the range of multivariate techniques available.
4. Understand the link between multivariate techniques and corresponding univariate techniques.
5. Conduct statistical inference about multivariate means including hypothesis testing, confidence region calculation, etc.
6. Use multivariate techniques appropriately, and draw appropriate conclusions.
7. Analyze multivariate data using the SPSS statistical software package.

UNIT I

Multivariate Normal distribution, its properties with proof: MGF, characteristic function, distribution of linear combination of multivariate normally distributed random variables, marginal and conditional distribution for partitioned multivariate normally distributed random variables.

UNIT II

Wishart distribution and its properties, Null and non-null distribution of simple correlation coefficient, Null distribution of partial and multiple correlation coefficient, Distribution of sample regression coefficients, Applications in testing and interval estimation.

UNIT III

Hotelling's T^2 statistic and its Null distribution, its Application: tests on mean vector for one and more multivariate normal populations, test for equality of the components of a mean vector in a multivariate normal population.

UNIT IV

Classification and discrimination procedures for discrimination between two multivariate normal populations-sample discriminant function, test associated with discriminant functions, probabilities of misclassification and their estimation, classification into more than two multivariate normal populations, Fisher Behrens Problem. Canonical variables and canonical correlations: definition, use, estimation and computation.

UNIT V

Dimension reduction, Principal components analysis, Factor analysis, Multivariate Analysis of variance (MANOVA) for one way classified data only,

REFERENCES

1. Anderson T.W. (1983): An Introduction to Multivariate Statistical Analysis (Second Edition) Wiley.
2. Giri, N.C. (1977): Multivariate Statistical Inference. Academic Press.
3. Khsirsagar A.M. (1972): Multivariate analysis. Marcel Dekker.
4. Morrison, D.F. (1976): Multivariate Statistical methods. 2nd. Ed. McGRAW Hill.
5. Muirhead, R.J. (1982) Aspects of multivariate statistical theory, J.Willey.
6. Rao CR (1973): Linear Statistical Inference and its Applications 2nd. Ed. Wiley.
7. Seber, G.A.F. (1984): Multivariate observations. Wiley
8. Sharma, S. (1996): Applied multivariate techniques. Wiley.
9. Srivastava M.S. & Khatri C.G. (1979): An Introduction to Multivariate Statistics. North Holland.
10. Johnson, R. and Wychern (1992): Applied multivariate Statistical analysis, prentice-Hall, 3rd. Ed.

STATCC -204: STOCHASTIC PROCESSES
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The main objective of this course is to develop awareness for the use of stochastic models for representing random phenomena evolving in time such as inventory or queueing situations or stock prices behavior.

Course Specific Outcomes:

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

1. Use notions of long-time behavior including transience, recurrence, and equilibrium in applied situations such as branching processes and random walk.
2. Construct transition matrices for Markov dependent behavior and summarize process information.
3. Use selected statistical distributions for modeling various phenomena.
4. Understand the principles and objectives of model building based on Markov chains, Poisson processes and Brownian motion.

UNIT I

Introduction to stochastic processes (sp's) : Classification of sp's according to state space and time domain, Countable state. Markov chains (MC's), Chapman-Kolmogorov equations, calculation of n-step transition probability and its limit

UNIT II

Stationary distribution, classification of states, transient MC, random walk and gambler's ruin problem, discrete state space continuous time Markov Chains: Kolmogorov-Feller differential equations

UNIT III

Poisson process, birth and death process, application to queues and storage problems, Wiener process as a limit of random walk, first-passage time and other problems

UNIT IV

Renewal theory: Elementary renewal theorem and applications, Statement and uses of key renewal theorem, study of residual life time process, Stationary process, weakly stationary and strongly stationary process, Moving average and auto regressive processes.

UNIT V

Branching process: Galton-Watson branching process, probability of ultimate extinction, distribution of population size, Martingale in discrete time, inequality, convergence and smoothing properties. Statistical inference in Markov Chains and Markov processes.

REFERENCES

1. Adke, S.R. and Munjunath, S.M. (1984): An Introduction to Finite Markov Processes, Wiley Eastern.
2. Bhat, B.R. (2000): Stochastic Models: Analysis and Applications, New Age International, India.
3. Cinlar, E. (1975): Introduction to Stochastic Process, Prentice Hall.
4. Feller, W. (1968): Introduction to probability and its Applications, Vol.1, Wiley Eastern.
5. Harris, T.E. (1963): The Theory of Branching Processes, Springer - Verlag.
6. Hoel, P.G... Port, S.C. and stone, C.J. (1972): Introduction to Stochastic Process, Houghton Miffin & Co.
7. Jagers, P. (1974): Branching Processes with Biological Applications, Wiley.
8. Karlin, S. and Taylor H.M. (1975): A First course in stochastic processes, Vol. I Academic press.
9. Medhi, J. (1982): Stochastic Processes, Wiley Eastern.
10. Parzen E. (1962): Stochastic Processes. Holden –Day.

STATCC -205: SEQUENTIAL ANALYSIS AND RELIABILITY THEORY
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The aim is to equip the students with the knowledge of the fundamentals of sequential estimation and Reliability concepts and measures

Course Specific Outcomes:

After the completion of this course the students

1. Would be able to understand apply the theory and applications of sequential procedures to the real data.
2. Get idea of important lifetime distributions such as for exponential, Weibull, gamma and lognormal distributions.
3. Understand the OC and ASN functions and their uses.

UNIT – I

Need for sequential procedures, SPRT and its properties. Wald's equation and identity, OC and ASN functions optimality of SPRT.

UNIT – II

Sequential estimation, Stein's two stage procedure.Anscombe Theorem. Chow-Robbin's procedure, Asymptotic consistency and risk efficiency, Estimation of normal mean.

UNIT – III

Reliability concepts and measures, components and systems, coherent systems, reliability of coherent systems cuts & paths, Bounds on system reliability.

UNIT – IV

Life distributions, reliability functions, hazard rate, Common life distributions : Exponential, gamma and Weibull, estimation of parameters and tests in these models.

UNIT-V

Notion of aging IFR, IFRA, NBU, DMRL and NBUE classes. Different types of redundancy and use of redundancy in reliability improvement, Problem of life testing, censored and truncated experiments.

REFERENCES

1. James O Berger (1985) : Statistical Decision Theory and Bayesian analysis. Springer
2. Fergusion T.S. (1967) : Mathematical Statistics - A decisions theoretic Approach. Academic Press.
3. Rohtagi, V.K. : An Introduction to Probability Theory and Mathematical Analysis, John Wiley and sons, New york.
4. Wald, A. : Sequential Analysis.
5. Whetherill, G.B. : Sequential Methods in Statistics, Methuen & Co. Ltd., New York, John Wiley & Sons.
6. DeGroot, M.H. : Optimal Statistical Decisions. McGraw Hill
7. Leonard T and Hsu J.S.J. : Bayesian Methods. Cambridge University Press.
8. Bernardo, J.M. and Smith AFM : Bayesian Theory. John willey.

9. Raiffa, h. &Schlaifer, r. (1961) : Applied Statistical Decision Theory.
10. Barlow R E and Proschen F (1985) : Statistical Theory of Reliability and Life Testing, Holt, Rinchart and Winston.
11. Lawless J F (1982) : Statistical models and methods of life Time Data. John Wiley.
12. Bain L J and Engelhardt (1991) : Statistical Analysis of Reliability and life testing models. Marcel Dekker.

STATCC -206 : PRACTICAL
(4 Credits – 6 hours of T/L per week)

1. Full rank model
2. Non full rank model
3. Tests for heteroscedasticity ; pure and mixed estimation.
4. Tests for autocorrelation.
5. Test for multicollinearity
6. Use of dummy variables (dummy variable trap) and seasonal adjustment.
7. Hotelling T^2/D^2 (Discriminant analysis)
   (a) To test $H_o : \mu = \mu_o$ from $N(\mu, \sum)$, $\sum$ unknown.
   (b) To test $H_0: \mu^{(1)} = \mu^{(2)}$ in $N_p(\mu^{(1)}, \Sigma)$, $N_p(\mu^{(2)}, \Sigma)$, Σ unknown.
   (c) Discriminant Analysis
   (d) Problem of Misclassification
8. Multivariate Analysis of variance (One way classified data only).
9. Principal components
10. Canonical correlations
11. Factor Analysis
12. Experiment based on system of Reliability
13. ASN & OC functions for SPRT
14. Experiments based on consistency, sufficiency and ML Estimation.
15. Experiments based on NP Lemma.

STATVNC-201: Non-Credited: OFFICIAL STATISTICS

Course Outcome:

This paper will help to know about different dimensions and issues related to our country through data and images such as graphs and visuals. It will help to provide basic information for decision making, evaluation related to administrative issues and policy making.

Course Specific Outcomes:

This course will provide students with an opportunity to develop:

1. Skill to visualize status of country.
2. To enhance basic information for decision making, evaluation, monitoring etc.
3. It will help in good policy and decision making.

UNIT-I

National and International official statistical system Official Statistics: (a) Need, Uses, Users, Reliability, Relevance, Limitations, Transparency, its visibility (b) Compilation, Collection, Processing, Analysis and Dissemination, Agencies Involved, Methods.

UNIT-II

National Statistical Organization: Vision and Mission, NSSO and CSO; roles and responsibilities; Important activities, Publications etc.

UNIT-III

National Statistical Commission: Need, Constitution, its role, functions etc.; Legal Acts/ Provisions/ Support for Official Statistics; Important Acts
Index Numbers: Different Types, Need, Data Collection Mechanism, Periodicity, Agencies Involved, Uses.

UNIT-IV

Sector Wise Statistics: Agriculture, Health, Education, Women and Child etc. Important Surveys & Census, Indicators, Agencies and Usages etc.

UNIT-V

National Accounts: Definition, Basic Concepts; issues; the Strategy, Collection of Data and Release. Population Census: Need, Data Collected, Periodicity, Methods of data collection, dissemination, Agencies involved. Misc.: Socio Economic Indicators, Gender Awareness/Statistics, Important Surveys and Censuses

References:

1. Goon A.M., Gupta M.K. and Dasgupta B. (2008): Fundamentals of Statistics (Vol.2), World Press.
2. Guide to current Indian Official Statistics, Central Statistical Office, GOI, and New Delhi.
3. C.S.O. (1984) : Statistical System in India
4. Mudgett B.D. (1951): Index Numbers, John Wiley
5. Allen R.G.D. (1975): Index Numbers in Theory and Practice, Macmillan
6. Mukhopadhyay P. (1999): Applied Statistics

SEMESTER III

STATCC -301: BLOCK DESIGNS AND THEIR ANALYSIS
(4 Credits – 4 hours of Theory Teaching per week)

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

Course Specific Outcomes:

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

1. Design and analyze incomplete block designs, understand the concepts of connectedness, balance, and orthogonality.
2. Understand the analysis of covariance, lattice design, and Split-plot designs and their analysis in practical situations.
3. Identify the effects of different factors and their interactions and analyze factorial experiments.
4. Understand the concepts of finite fields and finite geometries and apply them in construction of MOLS, balanced incomplete block designs, confounded factorial experiments.
5. Understand the concept of partially balanced incomplete block design.

UNIT – I
Fixed, mixed and random effects models; Variance components estimation: study of various methods, Tests for variance components.

UNIT – II
General block design and its information matrix (C), criteria for connectedness, balance design and orthogonality: Intrablock analysis (estimability, best point estimates/Interval estimates of estimable linear parametric functions and testing of linear hypotheses).

UNIT – III
BIBD - recovery of interblock information, Youden design - intrablock analysis, Lattice Design, Split plot design.

UNIT - IV
Analysis of covariance in a general Gauss-Markov model and its applications to standard designs, Missing plot technique - general theory and applications,

UNIT – V
Finite group and finite field, Finite geometry: projective and Euclidean, Construction of complete set of mutually orthogonal latin square (mols), Construction of BIBD's using mols and finite geometries, Symmetrically repeated differences, Steiner Triples and their use in construction of BIBD

REFERENCES

Raghava Rao D. (1971) : Construction and Combinatorial problems in Design of experiment. Wiley

Aloke Dey (1986) : Theory of Block Designs, Wiley Eastern.

Angela Dean and Daniel Voss (1999): Design and Analysis of Experiments, Springer.

Das, M.N. & Giri, N.(1979): Design and Analysis of experiments, Wiley Eastern.

Giri, N. (1986) : Analysis of Variance, South Asian Publishers.

John P.W.M.(1971): Statistical design and analysis of experiments, Mc Millan.

Joshi, D.D. (1987) : Linear Estimation and Design of Experiments, Wiley Eastern.

Montgomery, C.D.(1976): Design and analysis of experiments, Wiley, New York.

Meyer, R.H.(1971) : Response surface methodology. Allyn & Bacon.

Pearce, S.C.(1984) : Design of experiments Wiley, New York.

Rao, C.R. and Kleffe, J.(1988) : Estimation of Variance Components and applications, North Holland.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992) : Variance Components, Wiley.

Nigam, Puri & Gupta (1987-88) : Characterisation and Analysis of Block Design, Wiley Eastern.

V.K. Gupta & A.K. Nigam (1978-79) : Handbook an analysis of Agricutture Experiment, IASRI Publication.

OR

MOOC (Such as SWAYAM, etc.)

The candidates may opt any course from MOOC of the UGC/other websites and credit to be obtained is to be submitted in the office of CoE.

STATCC -302: PRACTICAL
(4 Credits – 6 hours of T/L per week)

Note: Practical will be performed on MS-Excel and SPSS.

1. Experiments based on BIBD
2. Experiment based on Lattice
3. Analysis of Covariance
4. Missing plot techniques
5. Split plot designs
6. Factorial Experiments
7. OLS estimation and prediction in GLM.
8. GLS estimation and prediction.
9. Ridge regression.
10. Estimation with lagged dependent variables.
11. Identification problems - checking rank and order conditions.
12. Estimation in recursive systems.
13. Two SLS estimation.
14. Simulation studies to compare OLS, 2SLS, LISE and FIML methods.

STATEL -301A: ECONOMETRICS
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The objective of this course is to study more advanced topics in econometrics and time series viz. MA, AR, ARMA, ARIMA Models, G.L.M. 2-SLS, 3-SLS estimators and Granger Causality Test.

Course Specific Outcomes:

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

1. Acquire knowledge of various advanced econometric models, estimation methods and related econometric theories.
2. Conduct econometric analysis of data.
3. Apply statistical techniques to model relationships between variables and make predictions.
4. Understand Auto-covariance, auto-correlation function and Vector Autoregression.
5. Understand Correlogram and Periodogram analysis and different Smoothing methods.

UNIT I

Nature of econometrics, review of general linear model (GLM), Generalized least squares (GLS) estimation and prediction, Heteroscedastic disturbances, Pure and mixed estimation, Grouping of observations and equations. Multicollinearity problem, its implications and tools for handling the problem.

UNIT II

Auto correlation, its consequences and tests, Linear regression with stochastic regressors, Instrumental variable estimation, Errors in variables, Autoregressive linear regression, Distributed lag models.

UNIT III

Stationary Processes: Moving average (MA) process, Auto-regressive (AR1 and AR2) process, ARMA and ARIMA models. Vector Auto Regression (VAR), the Granger Causality Test. Ridge regression. Introduction to GARCH model.

UNIT IV

Simultaneous linear equations model, Examples, Identification problem, Restrictions on structural parameters - rank and order conditions, Restrictions on variances and covariances.

UNIT V

Estimation in simultaneous equations model, Recursive systems, 2 SLS Estimators. Limited information estimators, k - class estimators. 3 SLS estimation, Full information maximum likelihood method, Prediction and simultaneous confidence intervals, Monte Carlo studies and simulation.

REFERENCES:

1. Apte, PG (1990): Text book of Econometrics. Tata McGraw Hill.
2. Cramer, J.S. (1971): Empirical Econometrics, North Holland.
3. Gujarathi, D. (1979): Basic Econometrics, McGraw hill.
4. Intrulligator, MD (1980): Econometric models - Techniques and applications, Prentice Hall of India.
5. Johnston, J (1984): Econometric methods, 3rd Ed. Mc Graw Hill.
6. Klein, L.R. (1962): An introduction to Econometrics, Prentice Hall of India.
7. Koutsoyiannis, A (1979): Theory of Econometrics, Macmillan Press.
8. Malinvaud, E (1966): Statistical methods of Econometrics, North Holland.
9. Srivastava, V.K. and Giles D.A.E. (1987): Seemingly unrelated regression equations models, Maicel Dekker.
10. Theil, H. (1982): Introduction to the theory and practice of Econometrics, john Wiley.
11. Walters, A (1970): An introduction to Econometrics, McMillan & Co.
12. Watherill, G.B. (1986): Regression analysis with applications, Chapman Hall.

STATEL -301B: ACTUARIAL STATISTICS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

In this course students learn about statistical models of transfers between multiple states, including processes with single or multiple decrements, and derive relationships between probabilities of transfer and transition intensities.

Course Specific Outcomes:

On successful completion of the course, one should be able to:

1. Explain the concept of survival models
2. Describe estimation procedures for lifetime distributions.
3. Describe the main methods of projecting/forecasting mortality rates.
4. Understand and discuss the ethical dimensions and implications of the modelling introduced in the course.

UNIT I

Introductory Statistics and Insurance Applications: Discrete, continuous and mixed probability distributions. Insurance applications, sum of random variables. Utility theory: Utility functions, expected utility criterion, types of utility function, insurance and utility theory, models for individual claims and their sums.

UNIT II

Survival function, Uncertainty of age at death, time until-death for a person, curate future lifetime, force of mortality. Life table and its relation with survival function, life table characteristics, assumptions for fractional ages, some analytical laws of mortality, select and ultimate life table.

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.
Principles of Premium Calculation: Properties of premium principles, examples of premium principles. Individual risk models: models for individual claims, the sum of independent claims, approximations and their applications.

UNIT IV

Life insurance: Insurance payable at the moment of death and at the end of the year of death-level benefit insurance, endowment insurance, deferred 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.

UNIT V

Net premiums: Continuous and discrete premiums, true monthly payment premiums, apportionable premiums, commutation functions, accumulation type benefits. A brief outline

of payment premiums and net premiums**,** Gross premiums and provisions. Profit testing-Determining provisions using profit testing. Factor affecting mortality and selections**.**

REFERENCES:

1. Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. (1997). Actuarial Mathematics. Society of Actuaries, Itasca, Illinois, U.S.A.
2. Daykin, C. D., Pentikainen, T. and Pesonen, M. (1993). Practical Risk Theory for Actuaries. Chapman & Hall/CRC.
3. Deshmukh, S.R. (2009). Actuarial Statistics: An Introduction Using R, University Press, India.
4. Dickson, C. M. D. (2005). Insurance Risk and Ruin (International Series no.1 Actuarial Science), Cambridge University Press.
5. Klugman,S. A., Panjer, H. H., and Willmotand, G. E. (2019). Loss Models: From Data to Decisions. Willy publication.
6. Neill, A. (1977). Life Contingencies, Heinemann.
7. Rotar, V.I. (2015). Actuarial Models: The Mathematics of Insurance, 2nd ed., CRC Press, New York.
8. Spurgeon, E.T. (1972). Life Contingencies, Cambridge University Press.

STATEL-302A: ADVANCED STATISTICAL ANALYSISS USING SPSS
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The objective of this course is to make students understand the significance of Data Preparation for Data Analysis and how to present and interpret data using statistical analysis software package SPSS. Discuss Data Analysis using Frequency Diagrams and Cross Tabulations. Introduce the visual representation of variables in graphs, bar charts and histograms. Understand the role and scope of Descriptive Statistics and Inferential Statistics

Course Specific Outcomes:

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

1. Discuss Data Analysis using Frequency Diagrams and Cross Tabulations
2. Introduce the visual representation of variables in graphs, bar charts and histograms.
3. Understand the role and scope of Descriptive Statistics and Inferential Statistics
4. Get an insight into use of EXCEL for Data Analysis

UNIT I

SPSS Environment & Interface, Data Preparation, Data Transformation: File Handling, File Transformation.
Exploratory Data Analysis: Frequencies, Descriptive Statistics, Explore, Cross-tabs, OLAP Cubes. Graphs.

UNIT II

Correlation and Partial Correlation. Simple and Multiple Linear Regression Models, Regression Diagnostics, Generalized Linear Regression Models: Binary Logistic, Ordinal Logistic, Poisson, Log Linear Regression Models.

UNIT III

Parametric and Nonparametric Tests: One and Two Sample problems. *ANOVA:* One-Way, Kruskal Wallis, General linear model: Two way ANOVA and ANCOVA, Repeated Measures. Basic experimental designs and factorial experiments.

UNIT IV

Factor Analysis, Discriminant Analysis, Nearest Neighbor Analysis, Choosing Procedures for Clustering, Two Step Cluster analysis, K-Means Cluster Analysis, Hierarchical Cluster Analysis. Control Charts.

UNIT V

Time Series Analysis: Creating and manipulating a time series, Components of a time series, autocorrelation and partial correlation function. *Developing Predictive Models*: Forecasting using exponential models, predictive accuracy measures for time-series forecast, testing for stationarity, Forecasting using ARIMA models.

REFERENCES:

1. Margan G A: SPSS for Introductory Statistics; Uses and Interpretation.
2. Practical Work Book by Bristol Information Services: Introduction to SPSS for Windows.
3. Rao A.R. and Bhimsankaran P. (1992): Linear Algebra, Tata McGraw Hill Publishing Company Ltd.
4. Rao CR and Mitra S.K.(1971): Generalized Inverse of Matrices and Its Applications, John Wiley & Sons, Inc.

STATEL - 302B: QUALITATIVE DATA ANALYSIS

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This course equips students with the necessary skills to analyse and interpret qualitative data to answer research and policy questions

Course Specific Outcomes:

1. Upon successful completion, students will have the knowledge and skills to:
2. Understand a range of qualitative analysis approaches that are commonly employed across disciplines;
3. Critically evaluate the advantages and disadvantages of a variety of qualitative analysis methods and select appropriate methods for application;
4. Develop and apply skills in thematic coding techniques;
5. Apply skills in qualitative data analysis using appropriate data management software; and
6. Assemble and present the results of qualitative research analyses in written and oral formats.

UNIT I

Categorical response variables: Nominal, ordinal, interval. Probability structure for contingency tables: joint, marginal and conditional probabilities, sensitivity and specificity, independence. Comparing proportions in 2x2 Tables: difference of proportions, relative risk. Odds Ratio: definitions and properties of odds ratio with examples, inference for odds ratio and log odds ratio, relationship between odds ratio and relative risk. Chi-square tests of independence: Pearson statistic, likelihood ratio statistic, tests of independence, partitioning Chi-squared.

UNIT II

Testing independence for ordinal data: linear trend alternative to independence, extra power with ordinal test, choice of score, trend tests for Ix2 and 2xJ tables, nominal-ordinal tables. Exact inference for small samples: Fisher's exact test for 2x2 table, p-values and conservatism for actual P(Type I error), small sample confidence interval for odds ratio. Association in three-way table: partial tables, conditional versus marginal associations, Simpson's paradox, conditional and marginal odds ratios, conditional independence versus marginal independence, homogeneous associations.

UNIT III

Models for binary response variables: logit, log linear, linear probability and logistic regression models. Logit models for categorical data, probit and extreme value models, models with log-log link, model diagnostics. Fitting logit models, conditional logistic regression, exact trend test. Log-linear models for two dimensions - independence model, saturated model and models for cell probabilities.Item Response Theory, Rasch Model.

UNIT IV

Loglinear models for two-way and three-way tables: log-linear model of independence for two-way table, saturated model for two-way tables, log-linear models for three-way tables. Inference for loglinear models: Chi-squared goodness of fit tests, log-linear cell residuals, tests about conditional associations, confidence intervals for conditional odds ratios, three

factor interactions, large samples and statistical versus practical significance.Fitting Log-linear models. Strategies in model selection, analysis of residuals, Cochran-Mantel-Haenszel test.

UNIT V

Models for Matched Pairs: Comparing Dependent Proportions, Conditional Logistic Regression for Binary Matched Pairs, Marginal Models for Square Contingency Tables, Symmetry, Quasi-symmetry, and Quasiindependence, Analyzing Repeated Categorical Response Data: Comparing Marginal Distributions: Multiple Responses, Marginal Modeling: Maximum Likelihood Approach, Marginal Modeling: Generalized Estimating Equations Approach, Quasi-likelihood and Its GEE Multivariate Extension, Markov Chains: Transitional Modeling.

REFERENCES:

1. Agresti, A. (2010): Analysis of ordinal categorical data, Wiley.
2. Agresti, A. (2013): Categorical Data Analysis, Third Edition, Wiley.
3. Bilder, C. R. and Loughin, T.M. (2013): Analysis of Categorical Data with R, CRC Press.
4. Bowerman, O. (2000): Linear Statistical models.
5. Congdon, P. (2005): Bayesian Models for Categorical Data, Willey.
6. Kleinbaum, D. G. (1994): Logistic Regression, Springer Verlag.
7. Sutradhar, B. C. (2014): Longitudinal Categorical Data Analysis, Springer.
8. Upton, G.J.G. (2017): Categorical Data Analysis by Example, Wiley.

STATIN-301 Summer Internship

Candidates/Students are required to attend an Internship Program during Summer Break. The Credits shall be provided by the Departmental Committee on submission of the report of internship.

STATIER-301 Applied Statistics

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This program is designed to improve the knowledge on application of statistics in Different applied fields. Our aim is to strengthen the students on application part with clear understanding of domain part.

Course Specific Outcomes:

This course will provide students with an opportunity to develop understanding of:

1. Statistical tools use for other discipline and their applications.

2. Critical thinking and problem solving in domain area.

3. Understanding the concepts of different sampling procedures.

4. Understand Inferential Statistics.

Note: The topics will be discussed with suitable examples and doesn't involve any kind of derivations. The use of non-programmable scientific calculator is permissible.

Unit I: Data and its representation

An introduction to Statistics, Data and its use in the field of Biology. classification of data: On the basis of collection, on the basis of measurement scales, on the basis of its nature. Tabular representation of data: frequency Distributions and contingency tables. Graphical Representation of data: Bar Charts of different types, Pie chart, Histograms, Frequency curve, Ogives.

Unit II: Descriptive Statistics:

Measures of Central Tendency: Arithmetic Mean, Median, Mode, Geometric Mean and Harmonic Mean and their properties. Measures of Dispersion: Range, Mean deviation, Standard Deviation and Quartile deviation and their properties, Relative measures of dispersion. Skewness and Kurtosis and their different measures.

Unit III: Survey Sampling:

Census and Sample Surveys, Steps involved in survey sampling, Sampling techniques: Probability and non probability samplings. Probability sampling techniques with equal probability of selection: Simple random sampling (with and without replacement), Stratified sampling, Systematic sampling, Cluster sampling, Multistage sampling. Non-Probability samplings: Purposive sampling, Judgement sampling, Quota sampling

Unit IV: Testing of hypothesis:

Inference, Parameter and Statistic. Testing of Hypothesis: Hypothesis, Null and alternative hypothesis, errors in testing of hypothesis, level of significance, power of test, p-value. Standard parametric tests: single sample t-test, Independent sample t-test, Paired sample t-

test, One-way ANOVA, F-test for equality of variances. Confidence Interval for mean and variance.

Non-parametric tests: Sign test, Wilcoxon signed rank test, Wilcoxon rank sum test, Median test, Mann-Whitney U test, Kruskal-Wallis test.

Unit V: Relation studies:

Multivariate data, For continuous data: Scatter plot, correlation: Carl Pearson's correlation coefficient, Spearman's rank correlation coefficient, Correlation matrix. Simple and Multiple Linear Regression, estimation of regression parameters, Inference regarding the parameters, ANOVA and t-tests for regression parameters, Selecting best regression equation: Forward and Backward selection methods, stepwise regression.

Attributes, 2x2 contingency table, RxC contingency table, association of attributes, Yule's coefficient of association, chi-square measure of association, Carl Pearson's coefficient of association, Tschuprow's coefficient of association, Chi square test for association.

References:

1. Hogg & Craig : Mathematical Statistics.
2. Mood, Graybill and Boes : Introduction to the theory of Statistics.
3. Goon, Gupta and Dasgupta : Fundamentals of Statistics Vol.1
4. Goon, Gupta and Dasgupta : Fundamentals of Statistics Vol.2
5. Cochran, W.G. : Sampling Techniques
6. Snedecor and Chochran: Statistical Methods

SEMESTER IV

STATCC -401: DECISION THEORY AND BAYESIAN ANALYSIS
(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

The objective of this course is to provide the understanding of the decision theory and fundamentals of Bayesian inference including concept of subjectivity and priors by examining some simple Bayesian framework.

Course Specific Outcomes:

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

1. To understand and utilize past experience along with present observation and improve the inferences.
2. Equip students with skills to carry out and interpret posterior data-based modeling and analyses.
3. Understand Decision theoretical concepts , game theory and their applications.
4. To understand the Bayesian estimation and testing procedures and compare them with classical inference.

UNIT – I

Decision problem and two person game, Utility theory, loss functions, Randomized and non-randomized decision rules, Essential completeness and completeness of class of rules based on sufficient statistic and the class of nonrandomized rules for convex loss , Optimal decision rules – unbiasedness, invariance, Bayes Rule, extended Bayes rule, Minimax rule, methods for finding minimax rules, admissibility of decision rules

UNIT - II

Generalized bayes and limit of bayes rule, Concept of admissibility and completeness Bayes rules, Admissibility of Bayes and minimax rules, Supporting and separating hyper plane theorems, complete class theorem, Minimax estimators of Normal and Poisson means

UNIT – III

Subjective interpretation of probability in terms of fair odds, Evaluation of (i) subjective probability of an event using a subjectively unbiased coin (ii) subjective prior distribution of a parameter, Bayes theorem and computation of the posterior distribution, Natural Conjugate family of priors for a model, Hyper parameters of a prior from conjugate family,

UNIT – IV

Bayesian point estimation as a prediction problem from posterior distribution, Bayes estimators for (i) absolute error loss (ii) squared error loss (iii) 0 -1 loss, Bayesian interval estimation : credible intervals, Highest posterior density regions,

UNIT – V

Interpretation of the confidence coefficient of an interval and its comparison with the interpretation of the confidence coefficient for a classical confidence interval, Bayesian Testing Hypothesis: Specification of the appropriate from of the prior distribution for a Bayesian testing of hypothesis problem, Prior odds, Posterior odds, Bayes factor.

REFERENCES

1. James O Berger (1985) : Statistical Decision Theory and Bayesian analysis. Springer.
2. Fergusion T.S. (1967) : Mathematical Statistics - A decisions theoretic Approach. Academic Press.
3. DeGroot. M.H. : Optimal Statistical Decisions. McGraw Hill.
4. Leonard T and Hsu J.S.J. : Bayesian Methods. Cambridge University Press.
5. Bernardo, J.M. and Smith AFM : Bayesian Theory. John Willey.
6. Rao, C. R. (1973): Linear Statistical Inference and its Applications, Wiley Eastern.
7. Robert, C. P.: The Bayesian Choice: A Decision Theoretic Motivation, Springer

STATEL -401A: FACTORIAL EXPERIMENTS AND RESPONSE SURFACES
(4 Credits – 4 hours of Theory Teaching per week)

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

Course Specific Outcomes:

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

1. Construct complete and partially confounded factorial designs and perform their analysis.
2. Understand concept of fractional factorial experiments and their analysis. Also construction of fractional factorial experiments by several methods.
3. Understand the applications area of designs of experiments such as response surface design, clinical trials, and treatment control design.
4. Understand the effects of independence or dependence of different factor under study.

UNIT – I

General factorial experiments, factorial effects, symmetric factorial experiments, best estimates and testing the significance of factorial effects; analysis of 2^n factorial experiment

UNIT – II

3^n factorial experiments in randomized blocks, analysis of 3^2, 3^3 and 3^n factorial experiments, Extension of Yates table for 3^n factorial experiments

UNIT – III

Complete and partial confounding in case of 2^n and 3^n factorial experiments, Fractional replication for symmetric factorials

UNIT – IV

Response surface experiments, first order designs and orthogonal designs.

UNIT – V

Clinical trials, longitudinal data, treatment- control designs, Model validation and use of transformation, Tukey's test for additivity.

REFERENCES

Raghava Rao D. (1971) : Construction and Combinatorial problems in Design of experiment. Wiley

Aloke Dey (1986) : Theory of Block Designs, Wiley Eastern.

Angela Dean and Daniel Voss (1999): Design and Analysis of Experiments, Springer.

Das, M.N. & Giri, N.(1979): Design and Analysis of experiments, Wiley Eastern.

Giri, N. (1986) : Analysis of Variance, South Asian Publishers

John P.W.M.(1971): Statistical design and analysis of experiments, Mc Millan.

Joshi, D.D. (1987) : Linear Estimation and Design of Experiments, Wiley Eastern.

Montgomery, C.D.(1976): Design and analysis of experiments, Wiley, New York.

Meyer, R.H.(1971) : Response surface methodology. Allyn & Bacon.

Pearce, S.C.(1984) : Design of experiments Wiley, New York.

Rao, C.R. and Kleffe, J.(1988) : Estimation of Variance Components and applications, North Holland.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992) : Variance Components, Wiley.

STATEL -401B: ETHICS, INTERGRITY AND APTITUDE

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This paper includes questions to develop the student's attitude and approach to issues relating to integrity, probity in public and social life and his/her approach to various issues and conflicts faced by him/her while dealing with society. Case study approach may be utilized to inculcate these values and appropriate positive aptitude for in depth understanding.

Course specific Outcome:

After completion of the course, a student will be able to:

Understand and have clarity of the human values, ethics and integrity in society and tackle various situations in life with positive attitude for common good.

UNIT- I

Ethics and Human Interface- Essence, determinants and consequences of Ethics in human actions; dimensions of ethics; ethics in private and public relationships.

Human Values - lessons from the lives and teachings of great leaders, reformers and administrators; role of family, society and educational institutions in inculcating values.

UNIT- II

Attitude- content, structure, function; its influence and relation with thought and behaviour; moral and political attitudes; social influence and persuasion.

Aptitude and foundational values for integrity, impartiality and non-partisanship, objectivity, dedication to social service, empathy, tolerance and compassion towards the weaker-sections.

UNIT- III

Emotional intelligence-concepts, and their utilities and application in academic, corporate sector and public services. Contributions of moral thinkers and philosophers from India and world. Values and Ethics in academic and pubic administration- Status and problems; ethical concerns and dilemmas in government and private institutions; laws, rules, regulations and conscience as sources of ethical guidance; strengthening of ethical and moral values in academic life.

UNIT- IV

Probity in Education and Research- Concept of Intellectual Property Rights (IPR) , Philosophical basis of teaching and probity; Information sharing and transparency in Education and Research, Codes of Ethics, Codes of Conduct, Citizen's Charters, Work culture, Quality of service delivery, challenges of corruption.

UNIT- V

Global Issues: Globalization and MNCs –Cross Culture Issues – Business Ethics – Media Ethics – Environmental Ethics – Endangering Lives – Bio Ethics – Computer Ethics – War Ethics – Research Ethics -Intellectual Property Rights.

REFERENCES

1. P.D.Sharma: Ethics, integrity and Aptitude, Rawat Publications, Jaipur.
2. G. Subba Rao and P.N. Roy Chowdhury: Ethics, Integrity and Aptitude, Access Publishing.
3. *Nanda Kishore Reddy and Santosh Ajmera:* Ethics, Integrity and Aptitude, Mcgraw Hill Education.
4. Professional Ethics by R. Subramaniam – Oxford Publications, New Delhi.
5. Ethics in Engineering by Mike W. Martin and Roland Schinzinger – Tata McGraw-Hill – 2003.

6. Professional Ethics and Morals by Prof.A.R.Aryasri, DharanikotaSuyodhana – Maruthi Publications.
7. Engineering Ethics by Harris, Pritchard, and Rabins, Cengage Specific, New Delhi.
8. Human Values & Professional Ethics by S. B. Gogate, Vikas Publishing House Pvt. Ltd., Noida.
9. Engineering Ethics & Human Values by M.Govindarajan, S.Natarajan and V.S.SenthilKumar-PHI Specific Pvt. Ltd – 2009.
10. Professional Ethics and Human Values by A. Alavudeen, R.Kalil Rahman and M. Jayakumaran – University Science Press.
11. Professional Ethics and Human Values by Prof.D.R.Kiran-Tata McGraw-Hill – 2013

STATEL-402A: DATA ANALYSIS USING R

(4 Credits – 4 hours of Theory Teaching per week)

Note: 3 credits for theory and 1 credit for practical.

Course Outcome:

The main objectives of this course are:

1. To learn the principles and methods of data analysis.
2. To provide a basic understanding of methods of analyzing data from different fields.
3. To learn R software.

Course Specific Outcomes:

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

1. Carry out data analysis using R software.
2. Effectively visualize and summarize the data.
3. Interpret the results of statistical analysis.

UNIT I

Introduction of R, R-Calculator, Vector and Matrices, Loading Packages, Data Entry and Exporting Data. Identify the components of R interface for Windows. Standard Arithmetic Calculation: Both Numerical and Matrix. Access R help, Load R Packages, Import and Export Data. Summary Statistics, Creating Summary for a Single Group and by different groups, Graphics in R.

UNIT II

Probability and Distributions. Parametric and Nonparametric tests. Karl Pearson's Correlation. Simple and Multiple Regressions, Regression Diagnostics. Tabular Data and Analysis of Categorical Data.

UNIT III

R Programming, Grouping, Loops and Conditional execution, functions, grouped expression and control statements, writing own functions.

Single and double integration, Gaussian integration. Monte Carlo Methods: Monte Carlo integration, Metropolis- Hastings and Gibbs sampler methods. Application of Monte Carlo methods to compute expected values of functions of random variables.

Random number generation by linear congruential, mixed congruential, inverse transforms, composition and acceptance-rejection methods. Statistical tests for pseudo random numbers.

UNIT IV

Bootstrap methods: resampling paradigms, bias and standard errors, confidence intervals, bootstrapping in regression. Jackknife and cross validation. Multivariate analyses.

UNIT V

EM algorithm: applications to missing and incomplete data problems, mixture models. Smoothing with kernels: density estimation, choice of kernels.

REFERENCES

1. McLachlan, G.J. and Krishnan, T. (1997). The EM Algorithms and Extensions, John Wiley & Sons.
2. Petra, Kuhnert and Bill, Venables: An Introduction to R; Software for Statistical Modeling and Computing.
3. Robert, J. Knell: Introductory R: A Beginner's Guide to Data Visualization and Analysis using R.
4. Robert, C.P. and Casella, G. (2004) Monte Carlo Statistical Methods, 2nd ed., Springer
5. Rubinstein, R.Y. (1981). Simulation and the Monte Carlo Method, John Wiley & Sons.

6. Sarah, S (2014): Using R for Statistics.
7. Sheskin, D. J. (2004): The Handbook of Parametric and Nonparametric Statistical Procedures, 3rd Edition, Chapman and Hall/CRC.
8. Everitt, Brian S. and Hothorn, Torsten (2006): A Handbook of Analyses Using R. Chapman & Hall.
9. Voss, J. (2014). An introduction to statistical computing: a simulation-based approach, Wiley series in computational statistics.

STATEL-402B: POPULATION STUDIES

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcomes

This subject introduces the main theories used to understand socio- economic and demographic patterns and changes in Indian context and all over the world. It covers a range of topics including the relationships between population size and available resources; social, biological and economic influences on population growth rates, fertility decline, population distribution and migration and gender statistics. The subject is very important to understand and develop statistical structures of economic policies and program implementation.

Course Specific Outcomes

Upon successful completion, students will have the knowledge and skills to:

1. Explain demographic changes in the world and their major determinants.
2. Apply demographic concepts and population theories to explain past and present population characteristics.
3. Evaluate the use of demographic concepts and population theories to understand contemporary socio-economic issues and current affairs.
4. Assess the relationship between demographic change and policy.

UNIT I

Demography- Sources of demographic data – census, vital statistics, sample surveys, official statistics. Demographic profiles of Indian Censuses. Age-Sex structure of population, factors affecting age sex structure.

Error and adjustments in collected data- Chandrasekaran—Deming formula to check completeness of registration data, adjustment of age data using Whipple, Myer and UN indices; Theory of population growth- Malthusian theory, Population transition theory.

UNIT II

Measures of Fertilty. Levels, trends and differentials of fertilty. Nuptiality and its indicators. Davis and Blake intermediate determinants of fertility, Bongaarts model of proximate determinants of fertility. Measures of Reproduction

Cohort rates; Parity and Birth Order; Child- Women ratio; Indirect estimation of fertility

UNIT III

Measures of Mortality; Levels, trends and determinants of mortality; Morbidity and its indicators; Reproductive morbidity; Life table- Construction of complete and abridged life tables; Model life table.

UNIT IV

Concept of Migration and its measures. Types of migration- Internal and International; Types of Internal migration; Trends and differentials of migration with special reference to India. Migration theories- Lee's theory, Ravenstein theory; Migration models- Zipf' model, Stouffer's model, gravity model, Harris todaro model; Concept of Urbanization and its association with migration; Trends and patterns of urbanization in India.

UNIT V

Stable and quasi-stable populations, -Intrinsic growth rate; doubling time, lexis diagram; Population growth models- Gompertz model, Logistic model; Methods od population projection- Algebraic methods and Cohort Component method.

Concept of gender and human development; Human Development Index; Inequality Adjusted Human Development Index; Gender Development Index; Women empowerment and its demographic consequences.

REFERENCES-

1. Kumar R. (1986): Technical Demography, Wiley Eastern Ltd.
2. Cox P. R. (1970): Demography, Cambridge University Press.
3. Keyfitz N., Caswell H. (2005): Applied Mathematical Demography Third Edition, Springer.
4. Bogue D. J.: Principles of Demography, John Wiley.
5. Techniques of Demographic Analysis (1998): by Ram F. & Pathak K.B. Himalaya Publishing House.

STATMT-401 Master Thesis

(Credits 08)

STATIRA-401 Data Analysis Using STATA

(4 Credits – 4 hours of Theory Teaching per week)

Course Outcome:

This program is designed to improve the knowledge on data analysis using sophisticated computer software called STATA. STATA is the most preferred computer software in developed nations. Our aim is to strengthen the students on data analysis using STATA.

Course Specific Outcomes:

This course will provide students with an opportunity to develop:

1. Deep discipline knowledge
2. Critical thinking and problem solving
3. Teamwork and communication skills

UNIT I

STATA Environment: STATA windows- command window and result window, Review window, variable window, data browser, data editor, do-file editor, viewer and Help. Files-log files, STATA data files; initial setup- memory allocation and setting setup file and other system parameters; functions, operators and expressions in STATA; STATA command syntax; working with STATA: basic unit of data, loading and saving data, working with data, value and variable labels, some elementary commands-commands for loading or importing and saving data in main memory, commands related to data manipulation, commands related to tabulation, command related to combining data, commands for reshaping/ re-structuring the datasets and commands which having `replace` options. other important commands-sorting, dealing with the variables and observations, and dealing with missing data.

UNIT II

Role of log files, concept of immediate commands, other immediate handy options-sample size and power estimation. Exploratory data analysis. Frequencies analysis, cross tabulations, descriptive Statistics, Three-way crosstabs, creation and editing of basic graphs. formatting graphs. Advanced Graphs-Scatter plots, Histograms, Catplot, Bars etc.

UNIT III

Analysis of continuous and binary outcomes, chi-squared test, t-test, one-way ANOVA, correlation, rank correlation, Simple linear regression, multiple linear regression .

UNIT IV

Multivariate logistic regression analysis, some epidemiological tests using STATA, Factor analysis, Principal component analysis.

UNIT V

Preparing data for advance statistical analysis- dealing with dates, setting time and time series variables. Advanced analysis- analysis of longitudinal data in STATA, Survival Analysis in STATA and Time series analysis

REFERENCES:

1. Rabe- Hesketh, S. and Everitt, Brain (2007): A handbook of statistical analysis using STATA, Chapman Hall/CRC Press.
2. Agresti, A. (2013): Categorical Data Analysis, Third Edition, Wiley.
3. Kothari, Prasad (2015): Data Analysis with STATA. Packt Publishing.
4. Longest, Kyle C. (2019). Using STATA for Quantitative Analysis. SAGE Publications.