SYLLABUS FOR M.Sc STATISTICS
(CBCS Pattern)
Effective from the Academic Year 2013-2014
Aim of the Course

The Degree of Master of Science in Statistics aims to train the students in the development and applications of Statistical techniques for analyzing data arising in the scientific investigation of problems in the various disciplines. It is also proposed to provide first hand practical experience in handling modern statistical softwares in the analysis of such data.

Eligibility for admission

Candidates for admission to the first year of the M.Sc. (Statistics) degree programme shall be required to have passed the B.Sc. degree examination of any Indian University recognized by the University Grants Commission with Statistics as the main subject or Mathematics as the main subject with Mathematical Statistics as one of the minor subject and a minimum of 55% marks in the main and allied subjects.

Duration of the Course

The course shall be of two years duration spread over four semesters. The maximum duration to complete the course shall not be more than 8 semesters.

Eligibility for admission to Examination

A candidate shall be permitted to appear for the M.Sc. examination in a subject of study only if he/she secures not less than 70% attendance in the subject concerned.

Medium

The medium of instruction shall be English.

Passing Minimum

As per the Choice Based Credit System regulations of the Pondicherry University.
M.Sc. (STATISTICS) – COURSE STRUCTURE
(With effect from 2013-14 onwards)

Objectives
The present course is intended to provide a platform for talented students to undergo higher studies in the subject as well as to train them to suit the needs of the society. Apart from teaching core Statistics subjects, the students are also trained to handle real life problems through practical classes. As part of the course, the students are taught some programming languages and also trained in various statistical softwares such as SPSS, SYSTAT, R language.

Eligibility
B.Sc. degree in Statistics or Mathematics with Mathematical Statistics as a minor subject with a minimum 55% of marks.

Duration of the Course
The course duration shall normally be of two years duration spread over four semesters.

Medium
The medium of instruction shall be English.

Choice Based Credit System (CBCS)
The M.Sc. Statistics program is offered through a unique CBCS. The salient feature of the CBCS is that the program is offered through credit based courses. Subjects are divided into Hard Core and Soft Core. Hard Core subjects are compulsory. The students have the choice to select from among the list of soft core subjects. Soft core subjects are similar to elective subjects.

A student is expected to complete a minimum of 72 credits within four semesters. Students are assessed and awarded letter grades based on their performances in the respective courses.

Weightage of marks
The weightage of marks for Continuous Internal Assessment (CIA) and end semester examinations shall be 40 and 60 respectively. A student is declared passed in a given subject when he/she secures
(a) a minimum of 40% in the end semester examination and
(b) a minimum of 50% marks in aggregate when Internal Assessment and End-Semester marks are added

Continuous Internal Assessment
The weightage of 40 marks for Continuous Internal Assessment component shall consist of the following:
a) Internal Assessment Tests (two) (2 x 15) = 30 marks
b) Seminars/Assignments/Presentations/Viva etc. (1 x 10) = 10 marks
  Internal Total = 40 marks

Supplementary Exam
(a) A failed student who meets the attendance requirement (70%) and has a minimum of 40% in the Internal Assessment marks may be permitted to register for the next End Semester examination in the following semester itself
(b) Students who have failed due to insufficient attendance and / or less than 40% in the Internal Assessment marks should repeat the course as and when it is offered.
PONDICHERRY UNIVERSITY
CHOICE BASED CREDIT SYSTEM
M.Sc. STATISTICS SYLLABUS
Effective from the Academic Year 2013 – 2014

<table>
<thead>
<tr>
<th>SEMESTER</th>
<th>COURSE CODE</th>
<th>TITLE OF THE COURSE</th>
<th>NATURE OF THE COURSE</th>
<th>NO. OF CREDITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>STAT 411</td>
<td>Linear Algebra and Matrix Theory</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 412</td>
<td>Probability Theory</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 413</td>
<td>Sampling Theory</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 414</td>
<td>Distribution Theory</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 415</td>
<td>Statistical Laboratory - I (Based on STAT 413 and SYSTAT)</td>
<td>Hard Core</td>
<td>3</td>
</tr>
<tr>
<td>II</td>
<td>STAT 421</td>
<td>Theory of Estimation</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 422</td>
<td>Statistical Quality Control and Total Quality Management</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 423</td>
<td>Stochastic Processes</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 424</td>
<td>Statistical Laboratory - II (Based on STAT 421, 422 and SPSS)</td>
<td>Hard Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Soft Core</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>STAT 531</td>
<td>Multivariate Statistical Analysis</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 532</td>
<td>Testing of Statistical Hypotheses</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 533</td>
<td>Linear Models and Regression Analysis</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 534</td>
<td>Statistical Laboratory - III (Based on STAT 531, 532, 533 and SPSS)</td>
<td>Hard Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Soft Core</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>STAT 541</td>
<td>Design and Analysis of Experiments</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT 542</td>
<td>Statistical Laboratory - IV (Based on STAT 541 and R-language)</td>
<td>Hard Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT 543</td>
<td>Project and Viva-Voce</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Soft Core</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Soft Core</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Soft Core Papers**

<table>
<thead>
<tr>
<th>Semester II</th>
<th>Semester III</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 425 Biostatistics</td>
<td>STAT 535 Reliability Theory</td>
</tr>
<tr>
<td>STAT 426 Decision Theory</td>
<td>STAT 536 Time Series Analysis</td>
</tr>
<tr>
<td>STAT 427 Econometrics</td>
<td>STAT 537 Actuarial Statistics</td>
</tr>
<tr>
<td>STAT 428 Bayesian Inference</td>
<td>STAT 538 Statistical Genetics</td>
</tr>
</tbody>
</table>

**Semester IV**

<table>
<thead>
<tr>
<th>STAT 544 Survival Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 545 Operations Research</td>
</tr>
<tr>
<td>STAT 546 Statistical Data Mining Methods</td>
</tr>
<tr>
<td>STAT 547 Demographic Techniques</td>
</tr>
</tbody>
</table>
Unit I
Vector Spaces, Sub-spaces, Basis of a vector space – Vector spaces with inner products - Gram-Schmidt orthogonalization.

Unit II
Linear transformation (LT) – Properties – Matrix of a linear transformation – Matrix of composite transformation – Matrix of an inverse transformation – Change of basis - Orthogonal transformation - Dual space.

Unit III
Linear equations – Solution space and null space – Sylvester’s law of nullity – Generalized inverse of a matrix – Moore – Penrose inverse

Unit IV
Eigen values and eigen vectors of an LT – left eigen vectors, right eigen vectors, Diagonalizable LT – Lambda matrix, Composition of lambda matrices, Operator polynomial, Cayley-Hamilton theorem and minimal polynomial for an LT – Eigen values of matrix polynomials.

Unit V
Bilinear forms - Canonical reduction – Sylvester’s law of inertia - Definitions of quadratic form - Lagrange’s reduction – Kronecker’s reduction - Reduction involving the eigen values of the matrix, Generalized eigen value problem.

Text books:
4. B.P.Parashar (1989), Linear Algebra, CBS Publishers and Distributors, Delhi

References:
2. Seymour Lipschutz and Marc Lipson (2003), Schaum’s Easy Outline of Linear Algebra, Tata Mc Graw Hill
4. Rao, A.R.and Bhimasankaran: Linear Algebra TMH.
STAT 412 - PROBABILITY THEORY  CREDITS: 4

Unit I

Unit II

Unit III
Convergence of a sequence of random variables - convergence in distribution, convergence in probability, almost sure convergence and convergence in quadratic mean - Weak and Complete convergence of distribution functions – Helly-Bray theorem.

Unit IV

Unit V

Books for Study

Books for Reference
Unit I

Unit II
Probability Proportional to size sampling- Inclusion Probabilities – Horvitz-Thompson estimator – Yates –Grundy form –Midzuno Sampling design - PPSWOR- Ordered estimator and unordered estimators

Unit III

Unit IV
Cluster Sampling: Equal cluster sampling – Estimators of mean and variance, optimum cluster size, Unequal cluster sampling – Estimators of mean and variance, varying probability cluster sampling - Two stage sampling – variance of the estimated mean – Double Sampling for stratification and Ratio estimation

Unit V
Randomized response methods – Warner’s, Simmon’s and Two Stage response methods - Sources of errors in Surveys - Mathematical model for the effects of call-backs and the errors of measurement – Official Statistical Systems in India – Role of NSSO and CSO and their activities – Organization of Large Scale Sample Surveys.

Books for Study

Books for Reference
Unit I
Brief review of distribution theory, distribution of functions of random variables - Laplace, Cauchy, Inverse Gaussian, Lognormal, Logarithmic series and Power series distributions - Multinomial distribution

Unit II
Bivariate Binomial – Bivariate Poisson – Bivariate Normal- Bivariate Exponential of Marshall and Olkin - Compound, truncated and mixture of distributions, Concept of convolution

Unit III
Multivariate normal distribution (Definition and Concept only) - Sampling distributions: Non-central chi-square, t and F distributions and their properties - Distributions of quadratic forms under normality-independence of quadratic form and a linear form - Cochran’s theorem

Unit IV
Order statistics, their distributions and properties- Joint and marginal distributions of order statistics - Distribution of range and mid range - Extreme values and their asymptotic distributions (concepts only)

Unit V

Books for Study

Books for Reference
I  Sampling Theory (20 marks) (Calculator based)
   2. PPSWR – Hurwitz Thompson estimator - Des Raj ordered estimator – Murthy’s unordered estimator – Midzuno scheme.
   3. Linear and circular systematic sampling.
   4. Stratified sampling – SRS, PPSWR, PPSWOR
   5. Cluster sampling of equal sizes.
   6. Ratio, Regression and Difference estimation.

II. Computations based on SYSTAT software (40 marks)
   1. Random number generation and Fitting of distributions.
      (i) Binomial, (ii) Poisson, (iii) Normal
   2. Computation of simple and multiple correlation coefficients
   3. Statistical Tests
      a. Large sample tests
      b. Tests for homogeneity of proportions, variances and correlation Coefficients
   4. Non – Parametric Tests
      a. Wilcoxon one sample test
      b. Mann – Whitney U test
   5. One-way ANOVA
Unit I
Parametric point estimation – properties of estimators – Consistency and its different forms
Sufficient condition for consistency- Unbiasedness – sufficient statistics – Factorization theorem - Distributions admitting sufficient statistic – Exponential and Pitman families procedure for finding minimal sufficient statistic.

Unit II
The information measure – Cramer - Rao (CR) inequality - Chapman - Robbins (KCR) inequality (single parameter case only) - Bhattacharya inequality (single parameter case only) - minimum variance bound estimator- Invariant (equivariant) estimators (concepts only)

Unit III
Uniformly minimum variance unbiased estimators (UMVUE)- condition for the existence of UMVUE- Completeness and Bounded completeness- Relation between complete statistic and minimal sufficient statistic- Rao - Blackwell Theorem- Lehmann – Scheffé’s theorem.

Unit IV

Unit V

Books for Study

Books for Reference
Unit I
Quality improvement: Meaning of quality and quality improvement – Control charts: Review of $\overline{X}$, R, p, c, d charts - Modified control charts for mean – CUSUM chart – technique of V-mask – Weighted Moving average charts – Slopping control charts and group control charts

Unit II
Process Capability analysis: Meaning, Estimation technique for capability of a process – Capability Indices: $C_p$, capability ratio and $C_{pk}$ index – Estimation of natural tolerance limit of a process. Acceptance Sampling plans for attributes: Single, double, multiple and continuous sampling plans for attributes (Dodge type)

Unit III
Acceptance Sampling plans for variables: one sided and two sided specification – Standardized plans (ANSI/ANSQ Z1.9) and MIL-STD-414
Plan-Do-Check-Act (PDCA), The Seven QC Tools, POKA YOKE, 5S, KAIZEN, Team Work and Quality circles

Unit IV
Need for TQM, TQM philosophy – Contributions of Deming, Juran, Crosby, Taguchi and Ishikawa, Concept of six sigma, Failure Mode and Effect Analysis (FMEA), Bench Marking and JIT

Unit V
Vision, Mission, Quality policy and objective, Planning and Organization for quality, Quality policy Deployment, Quality function deployment, Quality Audit, Need for ISO 9000 Systems, clauses, Documentation, Implementation, Introduction to QS 9000, Implementation of Quality Management System (QMS), Case Studies.

Books for Study

Books for Reference
Unit I

Unit II
Markov chains continuous in time – General pure birth processes and Poisson process, birth and death processes, finite state continuous time Markov chains.

Unit III
Branching processes discrete in time – Generating functions relations – Mean and variance – Extinction probabilities – Concept of Age dependent Branching process

Unit IV
Renewal processes – Definition and examples – key renewal theorem – Study of residual life time process

Unit V
Stationary process – weakly and strongly stationary process – Moving average and Autoregressive processes and their covariance functions - Brownian Motion process – Joint probabilities for Brownian motion process – Brownian motion as a limit of random walk

Books for Study

Books for Reference
I. Estimation (20 marks) (Calculator based)

1. MLE and Standard error of ML estimators.
2. MLE through the method of successive approximation.
3. MLE for truncated distribution.
4. Method of Moments
5. Method of Minimum Chi-square
6. Method of Least square
7. Interval estimation: Confidence interval for mean, difference of means and ratio of variances.

II. Statistical Quality Control (20 marks) (Calculator based)

Control charts:

   i. CUSUM chart
   ii. Modified Control chart
   iii. Moving Average Control chart
   iv. Exponentially Weighted Moving Average chart
   v. Sloping Control Chart

Acceptance sampling:

   i. Single sampling plans and double sampling plans (for attributes)
   ii. Variable Sampling plans (Single and double specifications)

III. Computations Based on SPSS software (20 marks)

2. Selection of cases, splitting and merging of files.
5. Fitting of curves – Linear, parabola, cubic and exponential.
6. Testing of Hypothesis – t, F, Chi square, one way and Two way ANOVA.
Unit I
Multivariate normal distribution– Marginal and conditional distributions – characteristic function. Maximum likelihood estimation of the parameters of Multivariate Normal and their sampling distributions – Inference concerning the mean vector when covariance matrix is known

Unit II
Total, Partial, Multiple correlation in the Multivariate setup – MLEs of Total, Partial and Multiple correlation coefficients. Sampling distributions of Total and Multiple Correlation in the null case. Hotelling $T^2$ statistic, derivation and its distribution - Uses of $T^2$ statistic - relation between $T^2$ and $D^2$ - Mahalanobis $D^2$ statistic and its distribution

Unit III
Generalized variance - Wishart distribution (statement only) – Properties of Wishart distribution - Test for covariance matrix – Test for equality of covariance matrices

Unit IV
Classification problems - Classification into one of two populations (known and unknown dispersion matrix) - Classification in to one of several populations – Fisher’s Linear discriminant function

Unit V

Books for Study

Books for Reference
Unit I
Randomized and non-randomized tests, Neyman – Pearson fundamental lemma, Most powerful tests, Uniformly most powerful test, Uniformly most powerful test for distributions with monotone likelihood ratio, Generalization of fundamental lemma and its applications

Unit II
Unbiasedness for hypothesis testing, Uniformly most powerful unbiased tests, Unbiased tests for one parameter exponential family, Similar test and complete sufficient statistics, Similar tests with Neyman structure, Locally most powerful tests.

Unit III
Invariant tests, maximal invariants, Uniformly most powerful invariant tests, Consistent tests, Likelihood ratio test, its properties and its asymptotic distribution, Applications of the LR method.

Unit IV
Non-parametric tests: Goodness of fit test : Chi-square and Kolmogorov Smirnov test - Test for randomness, Wilcoxon Signed rank test – Two sample problem: Kolmogrov-Smirnov test, Wald-Wolfowitz run test, Mann-Whitney U test, Median test

Unit V
Sequential tests: Basic Structure of Sequential tests – Sequential Probability Ratio Test (SPRT) and its applications – Determination of the boundary constants – Operating Characteristic and expected sample size of SPRT - Optimum properties of SPRT.

Books for Study

Books for Reference
Unit I
Full rank linear model – least square estimators of the parameters and their properties – Gauss-Markov theorem - Model in centered form – Estimators under normality assumption and their properties – Coefficient of determination – Generalized least squares – misspecification of the error structure and the model.

Unit II
Test for overall regression and for a subset of the parameters – test in terms of $R^2$ – General Linear Hypothesis testing – special cases – confidence region for the parameters and the mean – prediction intervals – likelihood ratio tests for the parameters – study of the residual outliers and influential observations

Unit III
Selection of input variables and model selection – Methods of obtaining the best fit - Stepwise regression, Forward selection and backward elimination – Multicollinearity – Collinearity diagnostics – Causes, Consequences and Remedy – Departure from normality

Unit IV
Introduction to general non-linear regression – Least squares in non-linear case – Estimating the parameters of a non-linear system – Reparametrisation of the model – Non-linear growth models – Concept of non-parametric regression

Unit V
Robust regression – Linear absolute deviation regression – M estimators – Robust regression with rank residuals – Resampling procedures for regression models – methods and its properties (without proof) - Jackknife techniques and least squares approach based on M-estimators.

Books for Study

Books for Reference
I Testing of hypothesis (20 marks) (Calculator based)

1. Construction of randomized and nonrandomized MP, UMP and UMPU tests of hypotheses and drawing the power curves.
2. Non parametric tests: Kolmogorov Smirnov test, Mann-Whitney U test, Median test for k-sample problem, Kruskal Wallis test and Friedman’s test

II Advanced Statistical Analysis using SPSS software (40 marks)

1. Hotelling’s $T^2$ test (Two sample)
2. Test for covariance matrix
3. Discriminant Analysis
4. One Way MANOVA.
5. Factor Analysis.
7. Test for Mean vector when dispersion matrix in known (Single and Two sample)
8. Fitting of Multiple linear regression- variable selection - diagnostics.
9. Residual Analysis for model adequacy, detection of outliers and influential observations
Unit I
Notion of design matrix- general analysis of design models (Inter and Intra Block analysis ) – C Matrix and its properties – Expected Mean Squares (EMS) and its uses, Algorithm for calculating EMS - Two way elimination of heterogeneity – Orthogonality – Connectedness and resolvability

Unit II
Principles of scientific experimentation – Basic Design: Completely Randomized Design (CRD), Randomized Block Design (RBD) and Latin Square Design (LSD) - Analysis of RBD (with one observation per cell, more than one but equal number of observations per cell) – Derivation of one and two missing values: Iterative and non-iterative methods – Loss of Efficiency due to missing values- Multiple comparison test: Least Significant Difference, Student Newman Kuel , Duncan’s Multiple Range, Tukey tests.

Unit III
Factorial experiments: 2^n and 3^n experiments and their analysis – Complete and Partial Confounding - Fractional Replication in Factorial Experiments – Split plot and strip plot design and their analysis.

Unit IV
Balanced Incomplete Block Design (BIBD) - Types of BIBD - Simple construction methods - Concept of connectedness and balancing – Intra Block analysis of BIBD – Recovery of InterBlock information – Partially Balanced Incomplete Block Design with two associate classes – intra block analysis only.

Unit V
Youden square and lattice design and their analysis – Analysis of Covariance with one concomitant variable – Analysis for CRD and RBD only

Books for Study
2. Douglas C. Montgomery (2009) : Design and Analysis of Experiments, 7/e, John Wiley and Sons, (Chapter 16 for Parts of Unit IV and Unit V)

Books for Reference
I. Design of Experiments (30 marks) (Calculator based)

1. Multiple Comparison tests
2. $2^4$, $3^2$, $3^3$ factorial experiment
3. Complete and partial confounding in $2^4$, $3^2$, $3^3$ factorial experiments
4. Fractional factorial
5. Split plot design and Strip plot design
6. Missing data analysis
7. Balanced Incomplete Block Design
8. Partially Balanced Incomplete Block Design with two associate classes
9. Youden Square Design
10. Analysis of Covariance – CRD and RBD

II. Computations based on R programming language (30 marks)

1. Plotting graphs
2. Creating objects, vectors, sequence, lists, arrays and matrices and performing basic operations.
3. Generating random numbers from Uniform, Binomial, Poisson, Normal, Multivariate Normal and Exponential distributions and fitting of the distributions.
5. Computation of descriptive statistics, correlation and regression coefficients.
6. One and two sample t tests, one way and two way ANOVA.
1. A project work is compulsory and shall be offered in semester IV. It will have 4 credits.

2. A project work may be taken individually or by a group of two students.

3. Project work shall be supervised by a faculty member assigned by the Head of the Department in the beginning of the semester.

4. The project work should be selected in such a way that there is enough scope to apply and demonstrate the statistical techniques learnt in the course.

5. At the end of the semester, a report on the work done should be submitted (two copies). If a team of two students jointly do a project work then they must submit individual reports separately (not copy of the same report).

6. The project report shall clearly state the selected problem, the statistical methodologies employed for data collection and analysis and the conclusions arrived at. Details of previous studies in the area and related references should also be given.

7. The project work will be assessed for a maximum of 100 marks. Each student will give a seminar before the end of the semester on their project work which will be evaluated internally for a maximum of 40 marks. There will be an external viva-voce examination for a maximum of 10 marks by an internal and an external examiner. The project report will be valued by the same external and internal examiner for a maximum of 50 marks.
Unit I
Statistical Methods in Clinical Trials: Introduction to clinical trial and it’s phases I, II, III and IV, statistical designs-fixed sample trials: simple randomized design, stratified randomized crossover design; Sequential design - open and close sequential design. Randomization-Dynamic randomization, Permuted block randomization; Blinding-Single, double and triple.

Unit II

Unit III
Data editing and transformations: Transformation in general - logarithmic, square root and power transformations; transformations for proportions – angular, probit and logit transformations. Outlying observations – box plot, M- estimators. Test for normality - p-p plot and q-q plot and Kolmogorov-Smirnov test.

Unit IV
Categorical Data Analysis: Categorical response data, logistic regression-odds ratio, Wald’s statistic, logistic regression and its diagnostics, ROC Curve analysis - Estimation of Binormal Model and the Area under the Curve, its applications, poison regression and its applications.

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

Books for Study

Books for Reference
Unit I
Basic elements of a decision problem - Randomized and non-randomized decision rules - Estimation and testing of hypothesis as decision problems - Bayes approach to inference and decision -

Unit II
Loss functions - Prior and Posterior distributions - Prior-Posterior analysis for Bernoulli, Poisson, and normal processes - Decision principles and Baye’s risk

Unit III
Utility theory - axioms, construction of utility functions, sufficiency, equivalence of classical and Bayesian sufficiency, complete and essentially complete classes of decision rules

Unit IV
Minimax analysis - Basic elements of game theory - General techniques of solving games - Finite games - Supporting and separating hyperplane theorems - Minimax theorem - Minimax estimation for normal and Poisson means

Unit V
Admissibility of Bayes and minimax rules, General theorems on admissibility, Robustness of Bayes rules, Invariant decision rules, Location parameter problems, Confidence and credible sets.

Books for Study:

Books for Reference:
1. Zellener (1971): An Introduction to Bayesian Inference in Econometrics, Willey
Unit I
Nature and Scope of Econometrics - Review of General Linear Model (GLM), Ordinary Least Squares (OLS), Generalized Least Squares (GLS) and Multicollinearity

Unit II
Heteroscedasticity - Autocorrelation, its consequences and tests - Ridge regression - Linear regression with stochastic regressors - Instrumental variable estimation - Errors in variables - Autoregressive linear regression - Distributed lag models.

Unit III
Simultaneous linear equations model - Identification problem - Restrictions on structural parameters - rank and order conditions - Restrictions on variances and covariances - Estimation in simultaneous equations model

Unit IV
Forecasting - Univariate forecasting methods - Forecasting in regression models - Forecasting with Simultaneous equations model - Evaluation of forecasts - Combination of forecasts

Unit V
Definition of causality - Granger causality - testing of causality - Cointegration, Bivariate cointegration tests - multivariate cointegration.

Books for Study

Books for Reference
Unit I
Subjective Interpretation of probability in terms of fair odds - Evaluation of Subjective probability of an event using a subjectively unbiased coin - Subjective prior distribution of a parameter - Bayes theorem and computation of the posterior distribution

Unit II
Natural Conjugate family of priors for a model - Hyper parameters of a prior from conjugate family - Conjugate families for exponential family models - models admitting sufficient statistics of fixed dimension - Enlarging the natural conjugate family by enlarging hyper parameter space - mixtures from conjugate family - choosing an appropriate member of conjugate prior family - Non informative, improper and invariant priors - Jeffrey’s invariant prior

Unit III
Bayesian point estimation: Prediction problem from posterior distribution - Bayes estimators for absolute error loss, squared error loss and 0 -1 loss - Generalization to convex loss functions - Evaluation of the estimate in terms of the posterior risk - Bayesian interval estimation : Credible intervals - Highest posterior density regions - Interpretation of the confidence coefficient of an interval

Unit IV
Bayesian Testing of Hypothesis: Prior and Posterior odds - Bayes factor for various types of testing hypothesis problems - Lindley’s paradox for testing a point hypothesis for normal mean

Unit V
Bayesian prediction problem: Prediction for Exponential family of Distributions - Predictive Distributions and Reliability Estimation - Predictive Interval - Ideas on Bayesian Robustness, Monte-Carlo Integration and Markov chain Monte Carlo techniques (without proof)

Books for Study

Books for Reference
SEMESTER III
STAT 535 – RELIABILITY THEORY CREDITS: 3

Unit I
Introduction to Reliability and its needs; Structural properties of coherent system: components and systems, coherent structures, representation of coherent systems in terms of paths and cuts, relevant & irrelevant structure; Modules of coherent systems; Reliability of a coherent systems; Reliability importance of components; Bounds on System Reliability.

Unit II
Life Distributions: Concept of distribution function, hazard function, Reliability function, MTTF, Bathtub failure rate; loss of memory property of Exponential distribution - parametric families of some common life distributions – Exponential, Weibull and Gamma and its characterization - Reliability estimation of parameters in these models.

Unit III
Notions of Ageing; Classes of life distributions and their duals - preservation of life distribution classes for reliability operation - Formation of coherent systems, convolutions and mixtures.

Unit IV
Univariate stock models and life distributions arising out of them: cumulative damage model, shock models leading to univariate IFR, Successive shock model; bivariate shock models; common bivariate exponential distributions due to shock and their properties. Maintenance and replacement policies; availability of reparable systems; modeling of a repairable system by a non-homogeneous Poisson process.

Unit V
Stress-Strength reliability - Concepts and its estimation; Reliability growth models; probability plotting techniques; Hollander –Prochan and Deshpande tests for exponentiality – Basic ideas of accelerated life testing.

Books for Study

Books for Reference
Unit I  
Exploratory Time Series Analysis: Forecasting trend and seasonality based on smoothing. Methods of Exponential and moving average smoothing; Types and implications of interventions; Outliers, additive and innovational outliers, procedure for detecting outliers.

Unit II  
Stationary Stochastic models: weak and strong stationarity, Deseasonalising and detrending an observed time series, Auto-covariance, autocorrelation function (ACF), partial autocorrelation function (PACF) and their properties, Conditions for stationarity and invertibility.

Unit III  
Models for Time Series: Time series data, Trend, seasonality, cycles and residuals, Stationary, White noise processes, Autoregressive (AR), Moving Average (MA), Autoregressive and Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) processes, Choice of AR and MA periods.

Unit IV  
Spectral analysis and decomposition: Spectral analysis of weakly stationary process, Periodogram and correlogram analysis, Spectral decomposition of weakly AR process and representation as a one-sided MA process – necessary and sufficient conditions, implication in prediction problems.

Unit V  
Modeling Seasonal Time Series: seasonal ARIMA models, estimation and forecasting, Fitting ARIMA models with Box-Jenkins procedure, Identification, Estimation, Verification, Test for white noise, Forecasting with ARMA models.

Books for Study  

Books for Reference  
Unit I
Basic deterministic model: Cash flows, discount function, interest and discount rates, balances and reserves, internal rate of return, The life table: Basic definitions, probabilities, construction of life tables, life expectancy, Life annuities: Introduction, calculating annuity premium, interest and survivorship discount function, guaranteed payments, deferred annuities.

Unit II
Life insurance: Introduction, calculation of life insurance premiums, types of life insurance, combined benefits, insurances viewed as annuities, Insurance and annuity reserves: The general pattern reserves, recursion, detailed analysis of an insurance, bases for reserves, non forfeiture values, policies involving a return of the reserve, premium difference and paid-up formula.

Unit III
Fractional durations: Life annuities paid monthly, immediate annuities, fractional period premium and reserves, reserves at fractional durations, Continuous payments: Continuous annuities, force of discount, force of mortality, Insurance payable at the moment of death, premiums and reserves. The general insurance – annuity identity, Select morality: Select an ultimate tables, Changed in formulas.

Unit IV
Multiple life contracts: Joint life status, joint annuities and insurances, last survivor annuities and insurances, moment of death insurances. The general two life annuity and insurance contracts, contingent insurances

Unit – V
Multiple decrement theory: Basic model, insurances, Determination of the models from the forces of decrement. Stochastic approach to insurance and annuities; Stochastic approach to insurance and annuity benefits, deferred contracts, Stochastic approach to reserves and premiums, variance formula.

Books for Study

Books for Reference
Unit I

Unit II

Unit III
Genetic components of variance: Relationship between phenotype and genotype, Different approaches, Genetic components of covariance between Traits; Linkage effects, Sex- linked genes, Maternal effect, Epistatic interaction, Genotype X Environment interaction.

Unit IV
Heritability, Estimation of Heritability, Precision of Heritability estimates, Repeatability, Estimates of Genetic correlation, Generalized Heritability

Unit V
Relation between phenotypic selection and genotypic selection, Intensity of selection correlated, Response to selection. Selection for improving several characters.

Books for Study

Books for Reference
Unit I
Concepts of time, order and random censoring and likelihood in these cases - Life distributions – Exponential, Gamma, Weibull, Lognormal, Pareto, Linear Failure rate
Parametric inference (Point estimation, Scores, MLE)

Unit II
Life tables, failure rate, mean residual life and their elementary properties - Ageing classes and their properties - Bathtub Failure rate

Unit III
Estimation of survival function – Actuarial Estimator - Kaplan- Meier Estimator - Estimation under the assumption of IFR / DFR - Tests of exponentiality against non-parametric classes – Total time on test, Deshpande test

Unit IV
Two sample problem: Gehan test, Log rank test. Mantel –Haenszel test, Tarone – Ware tests. Semi-parametric regression for failure rate – Cox’s proportional hazards model with one and several convariates - Rank test for the regression coefficients

Unit V
Competing risks model - parametric and non-parametric inference for this model - Multiple decrement life table

Books for Study

Books for Reference
STAT 545 – OPERATIONS RESEARCH

CREDITS: 3

Unit I
(Review of Linear Programming Problem (LPP) – Simplex, Big M and Two Phase methods) - Revised simplex method - Duality in LPP – Dual Simplex method – Some important theorems on duality - Sensitivity Analysis – Variation in cost vector and requirement vector – Addition and deletion of single variable and single constraint

Unit II
Integer Programming Problem (IPP) - Gomory’s cutting plane algorithm – Mixed IPP - Branch and Bound technique - Dynamic programming problem (DPP) - Bellman’s principle of optimality - General formulation - computation methods and application of DPP - Solving LPP through DPP approach

Unit III
Inventory models – Deterministic inventory models – Classic EOQ model – EOQ with price breaks – EOQ with storage limitations – Probabilistic Inventory models – Continuous review model – Single period model - No setup model – setup model (s-S policy)

Unit IV
Non-linear programming problem – Kuhn Tucker conditions – Quadratic Programming Problem (QPP) - Wolfe’s and Beale’s algorithms for solving QPP – Convex programming

Unit V
Queuing theory – Basic characteristics of queuing models – Arrival and service distribution – steady state solution of M/M/1 and M/M/C models with associated distribution of queue length and waiting time - M/G/1 queue-steady results using embedded Markov chain Methods - Pollazcek Khinchin formula.

Books for Study

Books for Reference
Unit I

Unit II

Unit III
Nearest Neighbor classifiers – kNN algorithm – Naïve Bayesian classifier – Binary logistic regression – odds ratio – Interpreting logistic regression coefficients – Multiple logistic regression

Unit IV

Unit V
Case studies based on k means clustering, fuzzy c means clustering, kNN classification, Binary logistic regression using R programming language or Excel Miner.

Books for study
1. Tan, T., Steinbach, M. and Kumar, V. (2006): Introduction to Data Mining, Pearson Education. (relevant portions of Chapters 1, 2, 4, 5 and 8).

Books for reference
Unit I

Unit II
Life tables: Construction of a life table, Graphs of \( l_x \), \( q_x \), \( d_x \), Functions \( L_x \), \( T_x \), and \( E_x \). Abridged life tables Mortality: Rates and Ratios, Infant mortality, Maternal mortality, Expected number of deaths, Direct and Indirect Standardization, Compound analysis, Morbidity.

Unit III
Fertility: Measures of Fertility, Reproductively formulae, Rates of natural increase, Fertility Schedules, Differential fertility, Stable Populations, Calculation of the age distribution of a stable population, Model Stable Populations.

Unit IV

Unit V
Ageing of the population, Estimation of demographic measures from incomplete date.

Books for Study

Reference Books
SOFT CORE COURSE FOR OTHER DEPARTMENTS

STAT 416 - STATISTICAL METHODS Credits: 3

Unit I

Unit II
Measures of central tendency – Mean, Median and Mode – Measures of dispersion – Range, Quartile deviation and Standard deviation – Coefficient of variation and skewness

Unit III

Unit IV
Elementary Probability theory: Addition theorem – Conditional probability and Multiplication theorem - Bayes’ Theorem – Random variables and probability distributions – Binomial, Poisson , Normal (simple applications of the distribution) – Sampling distributions: t, F and chi-square (definition only)

Unit V
Hypothesis testing: Basic concepts in Hypothesis Testing – Types of error – Tests for Mean and Proportion based on Normal and Student t-distribution - Chi-square test for independence of attributes – One-way and two-way Analysis of Variance

Books for Study

Books for Reference