PONDICHERRY UNIVERSITY

DEPARTMENT OF STATISTICS

SYLLABUS FOR M.Sc STATISTICS
(CBCS Pattern)
Effective from the Academic Year 2011-2012
PONDICHERRY UNIVERSITY
PUDUCHERRY 605 014

M.Sc. STATISTICS (CBCS - Semester Pattern)

REGULATIONS

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 Statistics as one of the minor subject with 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 examination in a subject of study only if he/she secures not less than 70% attendance in the subject concerned.

Medium

The medium of instructions 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 2011-12 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 exposed to various statistical softwares such as SPSS, SYSTAT, R language.

Eligibility

B.Sc. degree in Statistics or Mathematics with 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 minimum of 40% in the end semester examination in that subject.

Continuous Internal Assessment

The weightage of 40 marks for continuous internal assessment component shall consist of the following:

- Written test (best 2 of 3 class tests) = 30 marks
- Written assignments/ Seminar presentations = 10 marks

TOTAL = 40 marks
## M.Sc. STATISTICS SYLLABUS

Effective from the Academic Year 2010 – 2011

<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>Mathematical Methods for Statistics</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)</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 Operations Research</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></td>
<td>Soft Core</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT 424</td>
<td>Statistical Laboratory - II (Based on STAT 421, 422)</td>
<td>Hard Core</td>
<td>3</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></td>
<td>Soft Core</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT 534</td>
<td>Statistical Laboratory - III (Based on STAT 531, 532, 533)</td>
<td>Hard Core</td>
<td>3</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></td>
<td>Soft Core</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT 542</td>
<td>Statistical Laboratory - IV (Based on STAT 541)</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>
</tbody>
</table>

**Soft Core Papers**

<table>
<thead>
<tr>
<th>Semester II</th>
<th>Semester III</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 425</td>
<td>STAT 535</td>
</tr>
<tr>
<td>Econometrics</td>
<td>Reliability Theory</td>
</tr>
<tr>
<td>STAT 426</td>
<td>STAT 536</td>
</tr>
<tr>
<td>Demographic Techniques</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>STAT 427</td>
<td>STAT 537</td>
</tr>
<tr>
<td>Bayesian Inference</td>
<td>Actuarial Statistics</td>
</tr>
<tr>
<td>STAT 428</td>
<td>STAT 538</td>
</tr>
<tr>
<td>Survival Analysis</td>
<td>Total Quality Management</td>
</tr>
</tbody>
</table>

**Semester IV**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 544</td>
<td>Survival Analysis</td>
</tr>
<tr>
<td>STAT 545</td>
<td>Advanced Operations Research</td>
</tr>
<tr>
<td>STAT 546</td>
<td>Programming in C++</td>
</tr>
<tr>
<td>STAT 547</td>
<td>Time Series Analysis</td>
</tr>
<tr>
<td>STAT 548</td>
<td>Statistical Genetics</td>
</tr>
</tbody>
</table>
Unit I
Convergence of infinite numerical sequences and series (review only) – Absolute and conditional convergence – Sequences and series of functions – Pointwise and Uniform convergence – Tests for Uniform convergence – Properties of Uniform convergence

Unit II
Riemann - Stieltjes integral: Definition and properties – Integrals with step function and monotonic functions as integrators and their properties – Mean value theorem, Taylors theorem – Evaluation of Riemann - Stieltjes integral – Fundamental theorem

Unit III
Functions of several variables : Limits and continuity – Partial derivatives and Differentiability - Properties of differentiable functions – Higher order derivatives and differentials – Taylors theorem - Maxima and Minima – Extrema under constraints

Unit IV
Vector space and sub-space – Linear independence and orthogonality – Dimension and basis of a vector space – Orthonormal basis – Gram-Schmidt orthogonalization – Matrices: Rank, inverse, trace and their properties – Characteristic roots and vectors – Orthogonal Matrices and its properties - Idempotent and partitioned matrices

Unit V
G-inverse and Moore Penrose inverse - their properties – Reduction of a matrix into diagonal, echelon, canonical and triangular forms – Quadratic forms – reductions of different types – Definite quadratic forms – Cochran’s theorem

Books for Study

Books for Reference
Unit I

Unit II

Unit III
Convergence of a sequence of r.v.s. - convergence in distribution, convergence in probability, almost sure convergence and convergence in quadratic mean - Weak and Complete convergence of distribution functions – First Helly’s theorem

Unit IV

Unit V

Books for Study

Books for Reference
Unit I

Unit II

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 - 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
1. Cochran, W.G(1977): Sampling Techniques, 3/e, Wiley Eastern Ltd,. (Chapter 6 for Unit I, Chapter 7 for Unit II and Chapter 13 for Unit V)
2. Singh, D and Choudhary, F.S(1986): Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd,. (Chapter 5 for Unit III and Chapter 8 for Unit IV)

Books for Reference
STAT 414 – DISTRIBUTION THEORY

CREDITS: 4

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 and its properties – marginal and conditional distributions – characteristic function - Sampling distributions: Non-central chi-square, t and F distributions and their properties

Unit IV
Distributions of quadratic forms under normality-independence of quadratic forms and linear form- Cochran’s theorem - 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
Empirical distribution function and its properties, Kolmogorov Smirnov distributions, life time distributions - exponential and Weibull distributions - Mills ratio, distributions classified by hazard rate.

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. (i) Binomial, (ii) Poisson, (iii) Normal
2. Fitting of distributions – Binomial, Poisson and Normal
3. Computation of simple and multiple correlation coefficients
4. Statistical Tests
   (i) Z test and Confidence Interval
   (ii) One sample t-test and Confidence Interval
   (iii) Two sample t-test and Confidence Interval
   (iv) Chi-square test for independence of attributes
   (v) Tests for homogeneity of proportions, variances and correlation coefficients
5. Non – Parametric Tests
   (i) Run test
   (ii) Sign test
   (iii) Wilcoxon one sample test
   (iv) Mann – Whitney U test
6. Design of Experiments
   (i) One-way ANOVA
   (ii) Two-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 - Bhattacharya inequality - 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 

Unit IV
Methods of estimation – method of moments and its properties - method of maximum 
likelihood and its properties-Large sample properties of MLE - Method of minimum chi- 
square and its properties – Methods of least squares – Optimum properties of least square 
estimates in linear model.

Unit V
Interval estimation – Pivotal method of construction - shortest confidence intervals and their 
construction (minimum average width) - Construction of shortest confidence intervals in 
large samples. 
Notion of Bayes estimation – Concepts of prior, posterior and conjugate priors. Simple 
problems involving quadratic error loss function - Elementary notions of minimax estimation 
- Simple illustrations – Bayesian confidence intervals.

Books for Study

Books for Reference
USA. 
Kolkata.
STAT 422 – STATISTICAL QUALITY CONTROL AND OPERATIONS RESEARCH

CREDITS: 4

Unit I
Quality improvement: Meaning of quality and quality improvement – Different types of Quality costs and their management
Control charts: Review of $\bar{X}$, 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
Taguchi’s Loss function – Signal to Noise ratio – 5S concepts, Kaizen

Unit IV
Review of LPP – Simplex and revised simplex methods - Duality in LPP – Dual Simplex method – Some important theorems on duality - Sensitivity Analysis – Variation in cost vector ‘c’ – Variation in the requirement vector ‘b’ – Addition and deletion of single variable – Addition and deletion of single constraint

Unit V

Books for Study

Books for Reference
STAT 423 – STOCHASTIC PROCESSES

CREDITS: 4

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, variance 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. Standard plans

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 and one way ANOVA.
Unit I
Maximum likelihood estimation of the parameters of Multivariate Normal and their sampling distributions – Inference concerning the mean vector when covariance matrix is known - Total , Partial, Multiple correlation in the Multivariate setup – MLEs of Total, Partial and Multiple correlation coefficients and their sampling distributions in the null case

Unit II
Hotelling $T^2$ distribution and its applications - derivation of generalized $T^2$ statistic and its distribution - Uses of $T^2$ statistic - optimum properties of $T^2$ statistic - Mahalanobis $D^2$ statistic and its distribution - relation between $T^2$ and $D^2$ – Test based on $T^2$ statistic

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

Unit IV
Classification problems - Classification into one of two populations (known and unknown dispersion matrix) - Classification in to one of several populations – Linear discriminant function – Multivariate analysis of variance (MANOVA) – One-Way classification.

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, Uniformly most powerful unbiased tests, 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

Unit V
Sequential methods: Sequential unbiased estimation – Application to Normal distribution - Sequential test - 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

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  Multivariate analysis (20 marks) (Calculator based)

1. Maximum likelihood estimators – Mean vector and dispersion matrix, Test for Mean Vectors (\(\Sigma\) known and unknown)
2. Test for covariance matrix
3. Discriminant analysis
4. Principal Component Analysis
5. Canonical correlation and canonical variables

II  Testing of hypothesis (10 marks) (Calculator based)

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

III  Linear Models and Regression Analysis (10 marks) (Calculator based)

1. Fitting of Multiple linear regression model
2. Residual Analysis for model adequacy, detection of outliers and influential observations
3. Tests of General Linear Hypotheses

IV  Advanced Statistical Analysis using SPSS software (20 marks)

1. Computation of Multiple linear regression.
2. Computation of Hotelling \(T^2\) statistic.
3. Discriminant Analysis and MANOVA.
5. Factor Analysis.
Unit I
Notion of design matrix- general analysis of design models (Inter and Intra Block analysis ) – C Matrix and its properties – EMS and its uses, Algorithm for calculating EMS - Two way elimination of heterogeneity – Orthogonality – Connectedness and resolvability

Unit II
Principles of scientific experimentation – Pen and Plot techniques - Basic Design: CRD, RBD and LSD, Analysis of RBD (with one observation per cell, More then 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: LSD, SNK, DMR, 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
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 –Response Surface Designs – Method of Steepest Ascent-Taguchi Orthogonal Array Experiments

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. BIBD
8. PBIBD 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. Creating objects, vectors, sequence, lists, arrays and matrices and performing basic operations.
2. Generating random numbers from Uniform, Binomial, Poisson, Normal, Multivariate Normal and Exponential distributions and fitting of the distributions.
5. 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, before the last working day, 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 30 marks. There will be 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 60 marks.
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
STAT 426 – ECONOMETRICS

Unit I

Unit II
Auto correlation, its consequences and tests. Theil BLUS procedure, Estimation and prediction. Multicollinearity problem, its consequences, detection, implications and tools for handling the problem. Ridge regression.

Unit III

Unit IV

Unit V
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.

Books for Study

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
Unit I
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.

Unit II
Natural Conjugate family of priors for a model. Hyper parameters of a prior from conjugate family. Conjugate families for (i) exponential family models. (ii) models admitting sufficient statistics of fixed dimension. Enlarging the natural conjugate family by (i) enlarging hyper parameter space (ii) 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: as a prediction problem from posterior distribution. Bayes estimators for (i) absolute error loss (ii) squared error loss (iii) 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: Specification of the appropriate form of the prior distribution for a Bayesian testing of hypothesis problem - Prior odds, Posterior odds, Bayes factor for various types of testing hypothesis problems depending upon whether the null hypothesis and the alternative hypothesis are simple or composite. Specification of the Bayes tests in the above cases. Discussion of Lindley’s paradox for testing a point hypothesis for normal mean against the two sided alternative hypothesis.

Unit V
Bayesian prediction problem - Large sample approximations for the posterior distribution - Bayesian calculations for non conjugate priors: (i) Importance sampling, (ii) Obtaining a large sample of parameter values from the posterior distribution using Acceptance – Rejection methods, Markov Chain Monte Carlo methods and other computer simulation methods.

Books for Study

Books for Reference
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: IFR, IFRA, NBU, DMRL, NBUE, HNBUE (Duals: DFR, DFRA, NWU, IMRL, NWUE, HNWUE) ; 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.

Unit V

Reliability growth models; probability plotting techniques; Hollander –Proshan and Deshpande tests for exponentiality – Basic ideas of accelerated life testing.

Books for Study


Books for Reference

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, poison regression and its applications.

Unit V
One way ANOVA and Multiple comparisions- Tukey, Bonferroni, Scheffe’s, Dunnett’s test and Duncan Multiple range test. Confidence Interval for multiple comparisions, Non Parametric ANOVA and multiple comparision - contrasts, Multiple Comparisons among medians and variances

Books for Study

Books for Reference
STAT 537 - ACTUARIAL STATISTICS

CREDITS: 3

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
Need for TQM, evolution of quality, Definition of quality, TQM philosophy – Contributions of Deming, Juran, Crosby, Taguchi and Ishikawa.

Unit II
Vision, Mission, Quality policy and objective, Planning and Organization for quality, Quality policy Deployment, Quality function deployment, Analysis of Quality Costs.

Unit III

Unit IV
PDSA, The Seven QC Tools of Quality, New Seven management tools, Concept of six sigma, FMEA, Benchmarking, JIT, POKA YOKE, 5S, KAIZEN, Quality circles.

Unit V

Books for Study

Books for Reference
SOFT CORE PAPERS  
SEMESTER IV  

STAT 544- SURVIVAL ANALYSIS  

CREDITS: 3  

Unit I  
Concepts of time, Order and random Censoring, 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 – Acturial 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  
Unit I
Parameter Programming – Parameterization of the Cost Vector ‘c’ -Parameterization of requirement vector ‘b’ – All integer programming problem- Gomory’s cutting plane algorithm – Mixed integer programming problem – Branch and Bound technique.

Unit II
Inventory models with one or two price breaks - Multi item deterministic problem – Constraints on storage and investment – Probabilistic Inventory models – Periodic Review systems – Fixed order quantity system

Unit III
Non-linear programming problem – Kuhn Tucker conditions – Quadratic programming problem (QPP) - Wolfe’s and Beale’s algorithms for solving QPP – Geometric programming

Unit IV
Dynamic programming problem (DPP) - Bellman’s principle of optimality - General formulation - computation methods and application of DP - Solving LPP through DP approach - 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
Constants - Variables - Declaration of variables - Type conversions - Relational operators - Decision making, branching and looping - Functions - Simple functions - Passing arguments to functions - Returning values from functions - Reference arguments - Overloaded functions - Inline functions.

Unit II
Defining classes - Creating objects - Constructors - Accessing class members - Member functions - Overloaded constructors - Static class data - Arrays and strings.

Unit III
Operator overloading - Overloading unary and binary operators - Data conversion - Derived class - Class hierarchies - Public and private inheritance - Multiple inheritance.

Unit IV
Pointers in addresses - Arrays, functions and strings - Memory management - New and delete functions - Friend functions - Pointer to objects

Unit V
Files and streams - the fstream class - Exception handling - Class templates

Books for Study

Reference books
Unit I
Stochastic Time Series models – Classification of Stochastic Processes – The family of finite
dimensional distribution function - Stationary models and their autocorrelation properties –
Estimation of autocorrelation and partial auto correlation and their standard error
Deseasonalising and detrending an observed time series – Exponential and Moving average
smoothing

Unit II
General linear stationary models – stationarity and invertability – Autoregressive and moving
average processes and their autocorrelation functions – mixed autoregressive moving average
processes

Unit III
Model estimation – Likelihood and sum of squares functions – Nonlinear estimation –
estimation for special processes AR, MA, mixed processes – separation of linear and
nonlinear components in estimation – estimation using Bayes’ theorem

Unit IV
Forecasting: MMSE forecasts and their properties – Forecasts and their updating – Forecast
of functions and forecast of weights – examples

Unit V
ARIMA models – Box Jenkins methodology for fitting ARIMA models

Books for Study
   and Control, Pearson Education.
2. Brockwell P J and Davis R A (2002): Introduction to Time Series and Forecasting,
   Springer.

Books for Reference
1. Bovas Abraham and Johannes Ledolter(2005): Statistical Methods for Forecasting, 2/e,
   John Wiley & Sons.
   Chapman and Hall.
   McGraw Hill.
   Oxford University Press
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
SOFT CORE COURSES FOR OTHER DEPARTMENTS

STAT 416 - STATISTICAL METHODS FOR SOCIAL SCIENCES  

Unit I
Scientific research: Scientific methods and their characteristics – Various types and steps in scientific research – variable and types of variables – notion of hypothesis and its formulation – Research design and characteristics of a good design

Unit II

Unit III

Unit IV

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

Books for Study
5. Ram Ahuja : Research methods, Rawat Publications, New Delhi

Books for Reference
SOFT CORE COURSES FOR OTHER DEPARTMENTS

STAT 417 - STATISTICAL METHODS FOR BIOLOGICAL SCIENCES       Credits: 3

Unit I


Unit II

Study of relationship between variables – Quantitative: Correlation and Regression – Partial and Multiple correlation (three variables only) – Qualitative: Contingency tables – Measures of Association.

Unit III


Unit IV

Basic concepts in Hypothesis Testing – Types of error – p-value – Tests for Mean and Proportion based on Normal and Student t-distribution – Confidence Interval for large samples - Chi-square test for independence of attributes – One-way Analysis of Variance.

Unit V

Introduction to Design of Experiments – Completely Randomised Design, Randomised Block Design and Latin Square Design – Multiple comparison tests.

Books for Study

Books for Reference