PONDICHERRY UNIVERSITY

DEPARTMENT OF STATISTICS

SYLLABUS FOR M.Sc. STATISTICS
(CBCS Pattern)
Effective from the Academic Year 2018-2019
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 various disciplines. It is also
proposed to provide first hand practical experience in handling modern
statistical software 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 and Weight age of marks
The weight age of marks for Continuous Internal Assessment (CIA) and end
semester examinations shall be 40 and 60 respectively. As per the Choice
Based Credit System regulations of the Pondicherry University, a student is
declared as pass in a given subject he / she secures
(a) A minimum of 40% marks in end-semester exam and
(b) A minimum of 50% marks in aggregate when Internal assessment and End-Semester marks are added together

**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.

**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

**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.
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, R Programming, MINITAB.

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 shall be of two years duration spread over four semesters. The maximum duration to complete the course shall not be more than 8 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.
# PONDICHERRY UNIVERSITY
## CHOICE BASED CREDIT SYSTEM
### M.Sc. STATISTICS SYLLABUS

**Effective from the Academic Year 2018 – 2019**

<table>
<thead>
<tr>
<th>Semester</th>
<th>Course Code</th>
<th>Title of the Course</th>
<th>Course Type</th>
<th>No. of Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Semester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I Semester (4 Hard core) &amp; 2 Soft Core</td>
<td>STAT-411</td>
<td>Optimization Techniques</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>Distribution Theory</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT-414</td>
<td>Programming in R (Lab Based)</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT-415 to STAT 418</td>
<td>Any Two Soft Core Courses among set of Four are allowed</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td><strong>Second Semester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II Semester (4- Thy. H.C., 1-Pract.(HC) (2-Thy. S.C)</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>Sampling Theory</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>Regression Analysis</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT-425</td>
<td>Stat Practical Lab-1 Based on Stat-421, Stat-422 and Stat-424 (based on Calculator, Excel &amp; R Programming as per suitability)</td>
<td>Hard Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT-426 to STAT 429</td>
<td>Any Two Soft Core Courses among set of Four are allowed</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td><strong>Third Semester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III Semester (Any 2 from 4 Soft core papers)</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 Design of Experiments</td>
<td>Hard Core</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>STAT-534</td>
<td>Stat Practical Lab-2 Based on Stat-531,Stat-532 and Stat-533 (based on Calculator, Excel &amp; R Programming as per suitability)</td>
<td>Hard Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT-535 to STAT 538</td>
<td>Any Two Soft Core Courses among set of Four are allowed</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>STAT-530#</td>
<td>Internship on Statistical Works</td>
<td>Soft Core</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>STAT-539#</td>
<td>Internship on Data Analysis</td>
<td>Soft Core</td>
<td>2</td>
</tr>
<tr>
<td><strong>Fourth Semester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV Semester</td>
<td>STAT-541#</td>
<td>Project Work &amp; Comprehensive Viva</td>
<td>Soft Core</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>STAT-542 to STAT 545</td>
<td>Any Two relevant soft core courses either within or outside the department</td>
<td>Soft Core</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Soft Core</td>
<td>3</td>
</tr>
</tbody>
</table>

Total Credits: 72-Mandatory (+4-Non mandatory):

Hard Core:50-Mandatory; Soft Core:24 Mandatory (+4: Non mandatory)

@ Students can choose either STAT-541 or any Two courses from STAT-542 to STAT-545 as soft core courses in the Final Semester

# Selecting these soft core courses is optional, not a mandatory. Minimum duration of internship shall be 4 weeks,

Evaluation and grading will be based on the internship report submission and satisfactory performance in viva voce examination during the third semester end examination

<table>
<thead>
<tr>
<th>Soft Cores Courses in 1st Semester</th>
<th>Soft Cores Courses in 2nd Semester</th>
<th>Soft Core Courses in 3rd Semester</th>
<th>Soft Cores Courses in 4th Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT-416: Statistical Quality Control</td>
<td>STAT-427: Decision Theory</td>
<td>STAT-536: Bayesian Inference</td>
<td>STAT-543: Data Analysis Using Statistical Packages</td>
</tr>
</tbody>
</table>
Unit I
Mathematical Programming - Solving of LPP by graphical method - Linear Programming Problem (LPP) - Simplex, Big M and Two Phase methods – Revised simplex method – Solving LPP using Duality - Dual Simplex method

Unit-II
Post Optimality and Sensitivity Analysis – Variation in cost vector and requirement vector – Addition and deletion of single variable and single constraint - Integer Programming Problem (IPP) - Gomory’s cutting plane algorithm – Mixed IPP – Branch and Bound technique

Unit III
Dynamic programming problem (DPP) - Bellman’s principle of optimality - General formulation - computation methods and application of DPP - Solving LPP through DPP approach

Unit IV
Non Linear Programming: Constrained and Unconstrained Problems of Maxima and minima, Constraints in the form of equations (Lagrangean Method) and in equations (Kuhn-Tucker conditions), Quadratic programming: Beale’s and wolf’s methods simplex method for quadratic programming.

Unit- V
PERT - CPM: Applications, Basic Steps in PERT/CPM techniques; Time estimates and Critical Path in Network Analysis; Optimum and minimum duration cost, PERT, Resource Allocations.

Text Books

Reference Books
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 convergence of distribution functions – Slustky theorem - Helly-Bray theorem.

Unit IV

Unit V

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
Unit I
R language Essentials: Expressions and objects, Assignments, creating vectors, vectorized arithmetic, creating matrices, operations on matrices, lists, data frames – creation, indexing, sorting and conditional selection; examples.

Unit II
R Programming: conditional statements – if and if else; loops – for, while, do-while; functions – built-in and user defined; Data entry – reading from text file, data editor; examples.

Unit III
Descriptive Statistics and Graphics: Obtaining summary statistics; generating tables; Bar plots, Pie charts, Box plots, Histogram; exercises.

Unit IV
Probability and Distributions: Random sampling and combinatorial; obtaining density, cumulative density and quantile values for discrete and continuous distributions; generating samples from discrete and continuous distributions; Plotting density and cumulative density curves; Q-Q plot.

Unit V
Correlation: Pearson, Spearman and Kendall’s correlation; Regression – fitting, obtaining residuals and fitted values; one and two sample tests for mean and variance – one way and two way ANOVA.

Books for Study
1. Michael J.Crawley (2007), The R Book, John Wiley and Sons Ltd.

Lab Exercises:
1. Operations on vectors and matrices
2. Creating and manipulating data frames.
3. Writing user defined functions for finding arithmetic mean, median, factorial, matrix addition and multiplication.
4. Bar and Pie charts.
5. Box plots for single and multiple groups.
6. Density and cumulative density plots for Binomial, Poisson, Normal and exponential distributions.
7. Checking Normality using Histogram and Q-Q plot.
8. Correlation coefficient – Pearson’s, Spearman and Kendall’s Tau.
10. One sample and two sample t test.
11. One way and 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 (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 – Scheffe’s theorem.

Unit IV

Unit V

Books for Study

Books for Reference
5. Srivastava, Khan and Srivastava (2014), Statistical Inference: Theory of Estimation, PHI, India
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 – 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

Books for Study

Books for Reference
Unit I

Unit II

Unit III
Weiner Process: Brownian Motion, Joint probabilities, Wiener process, Differential equations, Kolmogorov equations, First passage time distribution; Branching Process: properties of generating functions of branching processes, Probability of Ultimate extinction, Distribution of Total number of Progeny, Age dependent Branching process

Unit IV
Renewal processes: Definition, examples and relationships between terms – renewal interval, delayed recurrent event, Renewal Processes in continuous time, Renewal Function and renewal density, renewal equation, renewal theorems – Study of residual life time process

Unit V

Books for Study

Books for Reference
Unit I

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 – Non-linear growth models – Concept of non-parametric regression

Unit V
Robust regression – Linear absolute deviation regression – M estimators – least squares approach based on M-estimators – Re-sampling procedures for regression models – Bootstrap and Jacknife methods and its properties (without proof).

Books for Study

Books for Reference
I. **Estimation (20 marks)**

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,
8. Interval estimation - difference of means,

II. **Sampling Theory (20 marks)**

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 estimators.

III. **Linear Models and Regression Analysis (20 marks)**

1. Fitting of Multiple linear regression model
2. Residual Analysis for model adequacy, detection of outliers and influential observations
3. Variable Selection procedures
4. Collinearity Diagnostics
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, Kruskal Wallis test and Friedman’s 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
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 III
Principles of scientific experimentation – Basic Designs: Overview of 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 IV
Balanced Incomplete Block Design (BIBD)– Types of BIBD – Simple construction methods – Concept of connectedness and balancing – Intra Block analysis of BIBD – Recovery of Inter Block information – Partially Balanced Incomplete Block Design with two associate classes – intra block analysis only - Split plot and strip plot design and their analysis.

Unit V
Factorial experiments: $2^2$, $2^3$, $2^4$ and $3^2$, $3^3$ experiments and their analysis – Complete and Partial Confounding - Fractional Replication in Factorial Experiments

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

Books for Reference
I Testing of Hypotheses (20 marks)

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

II Multivariate Statistical Analysis

1. Test for equality of mean vectors when covariance matrix is unknown (Hotelling’s $T^2$ test)
2. Test for Two Covariance matrices
3. Discriminant Analysis
4. Canonical correlation and canonical variables
5. One Way MANOVA with Post hoc tests (DMRT and Tukey’s).
6. Principal Component Analysis
7. Factor Analysis

III Design of Experiments (20 marks)

1. Multiple Comparison tests (Least Significant Difference (LSD) test, Bonferonni’s test)
2. Missing Data Analysis- one and two observations in RBD
3. Missing Data Analysis- one and two observations in LSD
4. $2^4, 3^2$ factorial experiments
5. Fractional factorial experiments
6. Complete confounding in $2^4, 3^2$ factorial experiments
7. Partial confounding in $2^4, 3^2$ factorial experiments
8. Split plot design
9. BIBD
10. Youden Square Design
11. Analysis of Covariance – CRD – One Concomitant Variable
12. Analysis of Covariance – RBD – One Concomitant Variable
SOFT CORE PAPERS

SEMESTER I

STAT415 - LINEAR ALGEBRA AND MATRIX THEORY  CREDITS: 3

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.

Books for Study

New Delhi.
Age International (P) Ltd, New Delhi.
Wiley and Sons

Books for Reference

3. Lipschutz S. and Lipson M. (2009), Schaum’s outlines, Linear Algebra, Fourth Edition,
McGraw Hill Company.
John Wiley and Sons, New Jersey.
Sons, New Jersey.
Unit I
Modified control charts for mean – CUSUM chart – technique of V-mask – Weighted Moving average charts – multivariate control charts – Hotelling’s $T^2$ control charts and Economic design of X-bar chart

Unit II
Process Capability analysis: Meaning, Estimation technique for capability of a process – Capability Indices: Process capability ratios $C_p$, $C_{pk}$, $C_{pm}$, $C_{pk}$, $C_{pm}$, $C_{pk}$ – Process capability analysis using a control chart – Process capability analysis using design of experiments

Unit III
Acceptance sampling – Terminologies – Attribute sampling plan by attributes – Single sampling plan and Double sampling plan – OC, ASN, AOQ, AOQL and ATI curves – MILSTD -105E Tables

Unit IV
Acceptance sampling variables for process parameter – Sequential plans for process parameter ($\sigma$ known and unknown) – Sampling variables for proportion non-conforming - $\bar{X}$ method, K method –

Unit V
Double specification limits – M-method, Double sampling by variables - MILSTD -414 Tables – Continuous Sampling plan – CSP-1, CSP-2, CSP-3, Wald and Wolfowitz SP-A and SP- B

Text Books

Reference Books
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 - Co integration, Bivariate co integration tests - multivariate co integration.

Text Books:

Reference Books:
STAT 418 - TIME SERIES ANALYSIS  

CREDITS: 3

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.

Text Books:

Reference Books:
SEMESTER – II

STAT 426 - 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. Concept of Ageing, Types of Ageing classes and their properties and relationship between them, Bathtub Failure rate, Concept of Inverse Hazard rate.

Unit III

Unit IV
Two sample problem- Gehan test, Log rank test. Mantel Haenszel test, Tarone Ware tests. Introduction to Semi-parametric regression for failure rate, Cox’s proportional hazards (PH) model with one and several covariates and estimation problems in Cox’s PH Model. Rank test for the regression coefficients.

Unit V
Introduction to Competing risks analysis and estimation problems in competing risk model for parametric and non-parametric semi parametric set up. Ideas of Multiple decrement life table and its applications.

Books for Study:


Books for Reference:

STAT 427–DECISION THEORY

CREDITS: 3

Unit I
Basic elements of a decision problem - Randomized and non-randomized decision rules - Estimation and testing of hypothesis as decision problems - Baye's approach to inference and decision -

Unit II
Loss functions - Prior and Posterior distributions and its 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 hyper plane theorems - Minimax theorem - Minimax estimation for normal and Poisson means

UNIT V
Admissibility of Baye's and minimax rules, General theorems on admissibility, Robustness of Baye's rules, Invariant decision rules, Location parameter problems, Confidence and credible sets.

Text Books:

Reference Books:
1. Zellener (1971): An Introduction to Bayesian Inference in Econometrics, Willey
Unit I
Statistical Methods in Clinical Trials: Introduction to clinical trial and its 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

Unit IV
ROC Curve analysis - Estimation of Binomial Model and the Area under the Curve, its applications – Properties of ROC curve - Kullback –Leibler Divergence (KLD) – definition – functional relationship between Kullback –Leibler Divergence and the slope of the ROC curve – derivations of KLD expressions for Bi-normal ROC model

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

Text Books

Reference Books
STAT 429-TOTAL QUALITY MANAGEMENT CREDITS: 3

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, Bench Marking, JIT, POKA YOKE, 5S, KAIZEN, Quality circles.

Unit V

Text Books

Reference Books
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 repairable systems; modeling of a repairable system by a non-homogeneous Poisson process.

Unit V
Stress-Strength reliability - Concepts and its estimation for exponential, Weibull and gamma distributions; Reliability growth models; probability plotting techniques; Hollander – Proschan and Despande tests for exponentiality – Basic ideas of accelerated life testing.

Text Books:

Reference Books:
Unit I
Introduction about Thomas Baye's - Motivations and Contributions - Evaluation of Subjective probability of an event using a subjectively unbiased coin - Subjective prior distribution of a parameter – Baye's theorem and computation of the posterior distribution.

Unit II
Introduction of Prior Distributions, Types of Prior Distributions, Proper Prior - 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 - Baye's estimators for absolute error loss, squared error loss, linear loss function, Jeffrey's and 0 -1 loss - Generalization to convex loss functions - Evaluation of the estimate in terms of the posterior risk

Unit IV
Bayesian interval estimation : Credible intervals - Highest posterior density regions - Interpretation of the confidence coefficient of an interval.

Unit V
Bayesian Testing of Hypothesis: Prior and Posterior odds - Baye's factor for various types of testing hypothesis problems - Monte-Carlo Integration and Basic Concepts on Markov chain Monte Carlo techniques (MCMC)(without proof).

Text Books

Reference Books
Unit I
Poisson Queuing Models with single server: Descriptions of queuing models, Generalized Birth and Death Processes, steady state Birth and death processes- Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/M/1): (∞/FIFO) and (M/M/1): (N/FIFO) Models, simple numerical problems

Unit II
Poisson Queuing Models with multiple server: Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/M/C): (∞/FIFO), (M/M/C): (N/FIFO) and (M/M/C): (C/FIFO) Models, simple numerical problems

Unit III
Non Poison Queuing Models (Erlangian): Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, Derivation of characteristics on (M/Ek/1), (Ek/M/1), simple numerical problems

UNIT – IV
Scope and notion of Inventory, Terminology, overview on probabilistic& Deterministic Models, optimality issues with Inventory; Deterministic Inventory models with shortages and without shortage. Finding EOQ and other characteristics, Simple Problem

UNIT – V
Deterministic Inventory models with simultaneous replenishment and stock clearance, with shortages and without shortages, finding EOQ and other parameters, Simple Numerical Problems; Deterministic Inventory models with Single and multiple price Breaks and numerical examples.

Books for Study:
1. KantiSwarup et al.:Operations Research, Sultan Chand and Sons,New Delhi
2. S.D Sharma: Operations Research

Books for Reference:
3. J.Medhi (2009), Stochastic Processes, 3/e, New Age International
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.

Text Books
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).

Reference Books
STAT 530  INTERNSHIP ON STATISTICAL WORKS  CREDITS: 2

**Time duration:** Student who wish to claim these credits should undergo the training or internship for a duration of at least 4 weeks

**Eligible Agencies:** The organization of internship shall be the governing body of either State or Central Government/ Quasi Government/ autonomous research body/ Finance or Banking organization/ Any official statistics of State and central governments.

**Eligible Slots:** The internship duration shall be during any of the following three slots of their course period of two years (M.Sc. Regular two years / M.Sc. Integrated 4th & 5th Years)
(i) Winter vacation after first semester (during second semester)
(ii) summer vacation after second semester
(iii) winter vacation after 3rd semester (during fourth semester)

**Material of internship for Assessment:** Students have to submit the internship/ apparent ship in black and white reflecting any statistical works such as Fieldwork, Primary/secondary Data Collection, Data compilation, Data processing, designing of structured questionnaires, Statistical tool designing, sampling frame and design of surveys, etc.

**Faculty In-charge:** A faculty in-charge from the department will be allotted to such students during the third semester work allotment. Students should register themselves under one faculty in-charge by applying the request to the HOD during the beginning of second year/third semester. The allocated faculty in-charge will be the responsible for the evaluation of internal and external examinations along with the end semester examinations of third semester.

**Evaluation:**
Student performance will be evaluated for 100 marks will be split in to two parts.

**Internal examination:** 40 Marks based on the certificate of assessment issued by the in-charge of industry/ office/ agency where the student got the statistical work internship.

**External/ End Semester Examination:** 60 Marks (40 marks for the submitted report of the student; 20 marks for comprehensive Viva Voce examination)

**Other Information:** Each Faculty in-charge may be allotted a maximum of 10 internship students. The registrations with PUSAMS has to be done along with the other courses of third semester. Non registered courses are not considered for any grade of this type of non mandatory credits.
Time duration: Student who wish to claim these credits should undergo the training or internship for a duration of at least 4 weeks

Eligible Agencies: The organization of internship shall be the governing body of either State or Central Government/ Quasi Government/ autonomous research body/ Finance or Banking organization/ Any official statistics of State and central governments. The Multinational companies, Data practicing agencies of private, government are also considered.

Eligible Slots: The internship duration shall be during any of the following three slots of their course period of two years(M.Sc. Regular two years / M.Sc. Integrated 4th & 5th Years)
(i) Winter vacation after first semester (during second semester)
(ii) summer vacation after second semester
(iii) winter vacation after 3rd semester (during fourth semester)

Material of internship for Assessment: Students have to submit the internship/ apparent ship in black and white reflecting any Data Analysis activity such as Data compilation, Data processing, Data Cleansing, Data Mining, Data Analysis Planning, Statistical Data Programming, Data Modeling, Prediction, Forecasting, Database Management, Structured Query Language Programming, Database Administration, Data Security, Data Warehousing, Data Mining, Data Networking and Cloud Computing, Online Data Analytical Processing, etc.

Faculty In-charge: A faculty in-charge from the department will be allotted to such students during the third semester work allotment. Students should register themselves under one faculty in-charge by applying the request to the HOD during the beginning of second year/third semester. The allocated faculty in charge will be the responsible for the evaluation of internal and external examinations along with the end semester examinations of third semester.

Evaluation:
Student performance will be evaluated for 100 marks will be splited in to two parts.

Internal examination: 40 Marks based on the certificate of assessment issued by the in-charge of industry/ office/ agency where the student got the statistical work internship.

External/ End Semester Examination: 60 Marks (40 marks for the submitted report of the student; 20 marks for comprehensive Viva Voce examination)

Other Information: Each Faculty in-charge may be allotted a maximum of 10 internship students. The registrations with PUSAMS has to be done along with the other courses of third semester. Non registered courses are not considered for any grade of this type of non mandatory credits.
SEMESTER IV  
STAT 541 - PROJECT WORK & COMPREHENSIVE VIVA  CREDITS: 6

1. It is an individual project work offered in IV semester with 6 credits.
2. The Project work shall be guided and supervised by a faculty member assigned in the beginning of the semester.
3. The project work should be undertaken in a reputed and relevant organization and topics are to be selected in such a way that there is enough scope to apply and demonstrate the statistical techniques learnt in the course.
4. At the end of the semester, before the last working day, project report should be submitted (two copies) with a certificate from industrial guide.
5. The project report shall contain the statement of problem, Methodology adopted, statistical tools used for analysis, findings, conclusions, suggestions and references.
6. The project work will be assessed for 6 credits. Students have to give a seminar of their project report at the end of the semester and which will be evaluated internally.
7. There will be viva-voce examination by an internal and an external examiner during end semester examination in 4th semester.
8. Report shall have the following format: Chapter I for Introduction for providing conceptual clarity, Chapter II for Review of Literature, Chapter III for Methodology, Chapter IV, V & VI for analysis and interpretations of each objectives (Number of chapter can be reduced or increased depending upon the number of objectives), chapter VII for findings and suggestions.
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.

Text Books

Reference Books
STAT 543 DATA ANALYSIS USING STATISTICAL PACKAGES CREDITS:3

Unit – I

Unit – II

Unit – III
Testing of Hypotheses – two sample and paired samples t – test; F-test for two sample variances; Chi-square test for independence of attributes – One way and Two Way Analysis of Variance – Multiple Comparison tests: Tukey’s test, Duncan’s Multiple range test and Dunnett’s test.

Unit – IV
Non-Parametric tests: One sample and Two sample Kolmogorov – Smirnov test, Kruskal – Wallis test, Friedman test, Median Test – One Way MANOVA – Hotelling’s $T^2$ two sample test – Test for two Covariance matrices – One way Repeated Measures ANOVA.

Unit - V
Factor Analysis : Identification of Principle Component, Varimax rotation – Discriminant Analysis – Enter and Stepwise procedures, discriminant scores – Logistic regression – variable selection procedures (Backward and Forward with conditional and wald methods), Odds ratio, Classification matrix – $2^2$, $2^3$, $3^2$ and $3^3$ factorial designs – Split Plot designs.

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 data.

Text Books:

Reference Books:
Unit I

Unit II
Control chart for variables – $\bar{X}$ and R chart – simple problems - Control charts for attributes – p, np, c charts – simple problems

Unit III
Basics of Experimental design - Principles of design of experiments: Randomization, Replication and local control - determination of experimental units and notion of experimental error – Completely Randomized Design (CRD) – Randomized Block Design (RBD) – Concepts and Simple problems

Unit IV
Latin Square Design (LSD) – Concepts and simple problems – Estimating a missing value in RBD and LSD - Multiple comparison tests : Duncan’s , Tukey’s and Least Significant Difference test

Unit V
Factorial Experiments – Concepts - $2^2$, $2^3$ and $3^2$ designs – Simple Problems

Books for Study
5. Peter W.M.John (1998), Statistical Design and analysis of experiment, SIAM publications

Books for Reference
SOFT CORE COURSE FOR OTHER DEPARTMENTS

STAT 419  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 - Baye’s 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

Text Books

Reference Books