

Presidency University, Kolkata

Department of Statistics

SYLLABUS IN STATISTICS

**FOR PG COURSE [Semesters 7 – 10]
(With effect from 2016 – 2017 ACADEMIC SESSION)**

OUTLINE :

SEMESTER - 7				
SL. No.	Paper	Code	Title	Marks
1	Major-17	STAT 0701	Real Analysis & Probability Theory-I	35+15
2	Major-18	STAT 0702	Sample Survey and Linear Models	35+15
3	Major-19	STAT 0703	Sampling Distributions, Inference-I (Estimation)	35+15
4	Sessional-9	STAT 0791	Sample Survey and Regression Diagnostics	50
5	Practical / Sessional-10	STAT 0792	Statistical Computing-I and Linear Models	50
SEMESTER - 8				
SL. No.	Paper	Code	Title	Marks
1	Major-20	STAT 0801	Probability Theory-II, Advanced Regression Analysis and Missing Data Analysis	35+15
2	Major-21	STAT 0802	Inference-II (Hypothesis Testing) & Asymptotic Theory	35+15
3	Major-22	STAT 0803	Design of Experiments & Advanced Multivariate Analysis	35+15
4	Practical / Sessional-11	STAT 0891	Based on STAT-0801, 802 and Statistical Computing-II	50
5	Practical / Sessional-12	STAT 0892	Based on STAT-0803	50

Department of Statistics, Presidency University, Kolkata

SEMESTER - 9				
SL. No.	Paper	Code	Title	Marks
1	Major-23	STAT 0901	Elective Paper : Bayesian Analysis and Decision Theory	35+15
2	Major-24	STAT 0902	Elective Paper : Stochastic Process and Time-Series Analysis	35+15
3	Major-25	STAT 0903	Elements of Research Methodology : Applied Multivariate Analysis & Resampling	35+15
4	Practical / Sessional-13	STAT 0991	Based on STAT-0901 & 0902	50
5	Practical / Sessional-14	STAT 0992	Based on STAT-0903	50

SEMESTER - 10				
SL. No.	Paper	Code	Title	Marks
1	Major-26	STAT 1001	Special Paper	35+15
2	Major-27	STAT 1002	Special Paper	35+15
3	Major-28	STAT 1003	Research Project : Nonparametric Inference and Sequential Analysis	35+15
4	Practical / Sessional-15	STAT 1091	Based on STAT-1001, 1002, 1003	50
5	Practical / Sessional-16	STAT 1092	Project	50

DETAILED SYLLABUS

SEMESTER - 7

SL. No.	Paper	Code	Title	Marks
1	Major-17	STAT 0701	Real Analysis & Probability Theory-I	35+15
2	Major-18	STAT 0702	Sample Survey and Linear Models	35+15

Department of Statistics, Presidency University, Kolkata

3	Major-19	STAT 0703	Sampling Distributions, Inference-I (Estimation)	35+15
4	Sessional-9	STAT 0791	Regression Diagnostics and Sample Survey	50
5	Practical / Sessional-10	STAT 0792	Linear Models and Statistical Computing-I	50

Major 17- Paper Code - STAT 0701 : Real Analysis & Probability Theory-I 35+15

Real Analysis (20)

Real Number System, Property of rational numbers, Cantor's diagonalization argument;

Bounded and unbounded subsets of the line, intervals, closed and open sets, characterizations, limit points, closures, interiors, Heine-Borel Theorem;

Sequence: Some preliminary results, limits superior and inferior; Subsequence and Bolzano-Weierstrass Theorem;

Functions: Continuous functions, Uniform continuity, Sequence of functions and convergence, uniform convergence;

Preliminary introduction to Metric Spaces.

References:

- T.M.Apostol : Mathematical Analysis
W.Rudin : Principles of Mathematical Analysis
R.R.Goldberg : Methods of Real Analysis
J.C.Burkill : First Course of Mathematical Analysis
J.C.Burkill & H.Burkill : Second Course of Mathematical Analysis
R.G.Bartle & D.R.Sherbert : Introduction to Real Analysis
H.L. Royden &
P.M. Fitzpatrick : Real Analysis, 4th edition.

Probability Theory – I (15)

Axioms of Measure; Classes of Sets; Borel sigma-algebra; Good Set principle, Pi-lambda theorem and Monotone class theorem; Measurable functions (random variables) and discussion of Riemann and Lebesgue integrals; Uniform Integrability;

Modes of convergence: almost sure, measure and in L_p .

References:

K. Athreya & S. Lahiri : Measure Theory and Probability Theory
P. Billingsley : Probability and Measure
Robert Ash & C.A. Doleans : Probability and Measure Theory, 2nd edition.
Dade : Probability and Measure Theory, 2nd edition.

Major 18- Paper Code - STAT 0702 Sample Survey and Linear Models 35+15

Sample Survey (20)

Probability sampling from a finite population – Notions of sampling design, sampling scheme, inclusion probabilities, Horvitz-Thompson estimator of a population total

Basic sampling schemes, Unequal probability sampling with and without replacement, Systematic sampling, Related estimators of population total / mean, their variances and variance estimators – Mean per distinct unit in simple random with replacement sampling. Hansen-Hurwitz estimator in unequal probability sampling with replacement. Des Raj and Murthy's estimator (for sample of size two) in unequal probability sampling without replacement

Unbiased Ratio estimators – Probability proportional to aggregate size sampling, Hartley-Ross estimator in simple random sampling.

Sampling and sub-sampling of clusters. Two-stage sampling with unequal number of second stage units and simple random sampling without replacement / unequal probability sampling with replacement at first stage, Ratio estimation in two-stage sampling.

Double sampling for stratification. Double sampling ratio and regression estimators. Sampling on successive occasions.

References :

W. G. Cochran : Sampling Techniques, 3rd edition, Wiley.
Des Raj & Chandak : Sampling Theory, Narosa.
A. S. Hedayat & B. K. Sinha : Design and inference in finite population sampling, Wiley.
P. Mukhopadhyay : Theory & Methods of Survey Sampling, Prentice-Hall, India.
M. N. Murthy : Sampling Theory and Methods, Statistical

Linear Models (15)

Gauss Markov Model : Estimable function, error function, BLUE, Gauss Markov theorem, Correlated set-up, least squares estimate with restriction on parameters.

Linear set & associated sum of squares, General linear hypothesis-related sampling distribution, Multiple comparison techniques due to Scheffe and Tukey.

Inference on Variance components

Analysis of covariance (general approach)

References :

- H. Scheffe : The Analysis of Variance.
S. R. Searle : Linear Models.
G. A. F. Seber : Linear Regression Analysis.
N. Giri : Analysis of Variance.
D. D. Joshi : Linear Estimation & Design of Experiments.
C. R. Rao : Linear Statistical Inference.

Major 19- Paper Code : STAT 0703 : Sampling Distributions & Inference-I 35 +15

Sampling Distributions (20)

Non-central χ^2 , t & F distributions – definitions and selected properties.

Distribution of quadratic forms – Cochran's theorem.

Sampling from Multivariate normal distribution – independence of sample mean vector and variance-covariance matrix. Wishart distribution.

Distributions of partial and multiple correlation coefficients and regression coefficients, distribution of intraclass correlation coefficient.

Hotelling T^2 and Mahalanobis's D^2 application in testing and confidence set construction.

Multivariate linear model: estimation of parameters, tests of linear hypotheses, Multivariate Analysis of variance of one- and two- way classified data (only LR test).

References:

- C.R.Rao : Linear Statistical Inference and its Applications
T.W.Anderson : Introduction to Multivariate Analysis
A.M.Khirsagar : Multivariate Analysis
S.S.Wilks : Mathematical Statistics

Department of Statistics, Presidency University, Kolkata

M.S.Srivastava & C.G.Khatri	:	Introduction to Multivariate Statistics
R.J.Muirhead	:	Aspects of Multivariate Statistical Theory

Inference-I (Estimation) (15)

Sufficiency & Minimal Sufficiency. Complete and boundedly complete families, Exponential family. Rao-Blackwell and Lehmann-Scheffe Theorems, Minimum Variance Unbiased Estimators and related problems

Interval estimation : Confidence sets based on pivot, relation with hypothesis testing, Optimum parametric confidence interval

References:

E.L.Lehmann, G. Casella	:	Theory of Point Estimation
S.Zacks	:	The Theory of Statistical Inference
C.R.Rao	:	Linear Statistical Inference and its Applications
T.S.Ferguson	:	Mathematical Statistics
J.O.Berger	:	Statistical Decision Theory and Bayesian Analysis
Jun Shao	:	Mathematical Statistics

Sessional -9 STAT 0791 : Marks – 50

Sample Survey (Practical : 15 M)
Regression Diagnostics (Theory : 20 M + Practical : 15 M)

Regression Diagnostics

Building a regression model : Transformations – Box-Cox model, Stepwise regression, Model selection (adjusted R^2 , Cross Validation and C_p criteria, AIC), Multicollinearity.

Detection of outliers and influential observations : residuals and leverages, DFBETA, DFFIT and Cook's Distance.

Checking for normality : Q-Q plots, Normal Probability plot, Shapiro-Wilks test.

Departures from the Gauss-Markov set-up : Heteroscedasticity and Autocorrelation –detection and remedies.

References:

N.R.Draper & H.Smith	:	Applied Regression Analysis
D.A.Belsley, Kuh & Welsch	:	Regression Diagnostics – identifying influential data & sources of collinearity
R.D.Cook & S.Weisberg	:	Residual and its Influence in Regression

Practical / Sessional -10 STAT 0792 : Marks – 50

Based on Linear Models and Statistical Computing-I

Statistical Computing I

C- Language

Overview of C-language: Simple Syntax, loops, pointers, arrays, functions, files, dynamic memory allocation. Sorting and Searching,

Generation of Random Numbers and Test of Randomness

Generating random permutations and drawing of random samples.

Generating samples from discrete and continuous distributions including bivariate and multivariate normal and multinomial distributions.

Generating samples from mixture distributions.

Spearman and Kendall rank correlation coefficient (both tie and no-tie cases)

R-Package

Topic	Details
A general overview	
Data types in R	numeric/character/logical; real/integer/complex strings and the paste command matrices, dataframes,lists, setwd, read.table, read.csv, write.matrix, write.csv, creation of new variables, categorisation cut, factor; round, apply creation of patterned variables saving output to a file; source; print saving workspace/history
Graphics in R	the plot command, histogram, barplot, boxplot points, lines, segments, arrows, paste inserting mathematical symbols in a plot, pie diagram, customisation of plot- setting graphical parameters text and mtext the pairs command, colours and palettes saving to a file; graphical parameters such as mar/mai/mfrow xlab/ylab/las/xaxp/yaxp/xlim/ylim/cex/axis/tck/srt

Department of Statistics, Presidency University, Kolkata

	main/title/legend/locator identify
Basic statistics	R help-command help, help.search(), r mailing list contributed documentation on cran one and two sample t tests, bartlett's test for variance, f test for equality of variances, multi sample means, non parametric tests, chi squared tests - randomness, homogeneity, independence exact tests and confidence intervals checking the assumptions, distribution fitting
Vector matrix operations	matrix operations addition, subtraction, multiplication, linear equations and eigenvalues matrix decomposition - lu, qr and svd and inverse the linear model and qr decomposition determinant g inverse finding a basis, orthonormalisation, finding rank
Linear models	the lm function; fitting a linear model; anova /ancova/regression models, the summary function, goodness of fit measures, predicted values and residuals; residual plots the anova table, testing nested models, confidence intervals and confidence ellipsoids, multiple testing manipulating data; creating factors
Regression	case study from regression analysis
R functions	some useful inbuilt r functions - sort, order, rank, ceiling, floor, round, trunc, signif, apply, lapply, by programming in r- for/while/if loops, functions, the source command
Random no. generation &	rmnorm, rchisq, rt, rbinom etc; sample; set.seed

Department of Statistics, Presidency University, Kolkata

P.Billingsley	:	Probability and Measure
J.F.C. Kingman & S.J.Taylor	:	Introduction to Measure and Probability
R.G.Laha & V.K.Rohatgi	:	Probability theory
R.Ash	:	Real Analysis and Probability

Advanced Regression Analysis (15)

Generalized Linear Models : Introduction, Components of a GLM, Goodness of fit – deviance, Residuals, Maximum likelihood estimation.

Binary data and Count data: ungrouped and grouped. Polytomous data

Overdispersion, Quasi-likelihood.

Models with constant coefficient of variation, joint modeling of mean and variance, Generalized additive models.

References:

N.Draper & H.Smith	:	Applied Regression Analysis
H.D.Vinod & A.Ullah	:	Recent Advances in Regression Methods
P.McCullagh & A.J.Nelder	:	Generalized Linear Models
McCullough C. E. & S.R Searle	:	Generalized, Linear and Mixed Models, 2nd Edition
J.Rousseeuw & A.M.Leroy	:	Robust Regression & Outlier Detection
T. Hastie and R. Tibshirani	:	Generalized Additive Models

Missing Data Analysis (5)

Missing Data Mechanisms, Complete Case and Available Case Analysis, Imputation Based Measures, Likelihood Based Inference, EM Algorithm and its applications. Convergence property of EM algorithm

References:

D.Rubin & R.J.A. Little	:	Statistical Analysis with Missing Data
G.McLachlan & T.Krishnan	:	The EM Algorithm and Extensions

Major 21- Paper Code : STAT 0802 : Inference-II & Asymptotic Theory 35+15

Inference – II (Hypothesis Testing) (18)

Department of Statistics, Presidency University, Kolkata

Review of notions of nonrandomized and randomized tests, level, size, power function, Fundamental Neyman-Pearson Lemma (Proof: Existence and Sufficiency parts) and its generalization (Sufficiency part only), UMP Tests and related problems . Monotone Likelihood ratio.

UMPU Tests: One parameter exponential family(without derivation) , Locally best tests(Concepts only)

Similar tests, Neyman structure, UMPU tests for composite hypotheses and related problems

References:

- E.L.Lehmann : Testing Statistical Hypotheses
S.Zacks : The Theory of Statistical Inference
C.R.Rao : Linear Statistical Inference and its Applications
T.S.Ferguson : Mathematical Statistics
J.O.Berger : Statistical Decision Theory and Bayesian Analysis
Jun Shao : Mathematical Statistics

Asymptotic Theory (17)

Basic convergence concept : Various modes of convergence (in distribution, in probability, L_p)

Moment convergence and uniform integrability

Scheffe's theorem, Slutsky's theorem. Asymptotic normality, multivariate CLTs, delta method. Glivenko-Cantelli Lemma (Statement)

Consistency and Asymptotic Efficiency of Estimators, Large sample properties of Maximum Likelihood estimators.

Asymptotic distributions and properties of Likelihood ratio tests, Rao's score test and Wald's tests in the simple hypothesis case.

References:

- R.J.Serfling : Approximation Theorems of Mathematical Statistics
E.L.Lehmann : Large Sample Theory
C.R.Rao : Linear Statistical Inference and its Applications
Jun Shao : Mathematical Statistics

Major 22- Paper Code STAT 0803 : Design of Experiments & Advanced Multivariate Analysis 35+15

Design Of Experiments(20)

Block Designs : Connectedness, Orthogonality, Variance Balance

Construction of mutually orthogonal Latin Squares (MOLS)

BIB designs – Properties, Intra-block analysis, Construction through mols, Bose's difference method and block section and intersection, Resolvable BIB designs

Factorial designs : Analysis, confounding and balancing in symmetric factorials (prime-power case)

References :

- M. C. Chakraborty : Mathematics of Design and Analysis of Experiments.
A. Dey : Theory of Block Designs.
D. Raghavarao : Constructions & Combinatorial Problems in Designs of Experiments.
R. C. Bose : Mathematical Theory of Symmetric Factorial Design
(Sankhya – Vol. 8).
R. C. Bose : On the Construction of Balanced Incomplete Block Design
(Annals of Eugenics – Vol. 9).
R. C. Bose : On Application of Galois fields to the problem of construction of hyper- graeco-
latin square -Sankhya Vol.3, page 328 – 338.
D. C. Montgomery : Design and Analysis of Experiments.

Advanced Multivariate Analysis:(15)

Estimation of mean vector and dispersion matrix: Maximum Likelihood Estimate of mean and variance under multivariate normal setup, Shrinkage Estimators of mean- James Stein's Estimate.

Tests involving mean vector: Hotelling T^2 test, Application to one sample and two sample problems. Union-Intersection Principle, Wilk's Λ statistic

Multivariate Linear Model, MANOVA

References :

- C.R.Rao : Linear Statistical Inference and its Applications
T.W. Anderson : An Introduction to Multivariate Analysis.
G.A.F. Seber : Multivariate Observations.
N.C. Giri : Multivariate Statistical Analysis.
R. J. Muirhead : Aspects of Multivariate Statistical Theory

Practical / Sessional -11 : STAT 0891 : Marks – 50

Based on STAT-0801, 802 and Statistical Computing-II

Statistical Computing II

C- Language

Department of Statistics, Presidency University, Kolkata

Acceptance Rejection Sampling, Monte Carlo Integration, Importance Sampling.

Missing Data Analysis: Imputation Based Measures, EM Algorithm.

Univariate Sign Test, Wilcoxon Rank Sum test, Kolmogrov-Smirnov Test for goodness of fit including simulation of p-values

R package

Computational problems:

- 1) Monte Carlo Techniques
- 2) EM Algorithm
- 3) Density Estimation

Practical / Sessional -12 : STAT 0892 : Marks – 50

Based on STAT 0803

Semester-9

1	Major-23	STAT 0901	Elective Paper : Bayesian Analysis and Decision Theory	35+15
2	Major-24	STAT 0902	Elective Paper : Stochastic Process and Time-Series Analysis	35+15
3	Major-25	STAT 0903	Elements of Research Methodology : Applied Multivariate Analysis & Resampling	35+15
4	Practical / Sessional-13	STAT 0991	Based on STAT-0901 & 0902	50
5	Practical / Sessional-14	STAT 0992	Based on STAT-0903	50

Major 23- Paper Code - STAT 0901 - Elective Paper : Bayesian Analysis and Decision Theory Marks – 35+15

Bayesian Analysis (20)

Overview and comparison of two paradigms-Classical statistical analysis and Bayesian analysis. Relative advantages and disadvantages, Motivation for choice of different priors, Details on Conjugate Priors and Posteriors.

Department of Statistics, Presidency University, Kolkata

Bayesian Inference- point estimation and credible sets. Predictive distributions. Illustration with examples of one-parameter and multiparameter models using conjugate and noninformative priors.

Bayesian testing and Model selection. BIC. Objective Bayes factors. Intrinsic priors. Bayesian variable selection, Comparison of p-value and posterior probability of H_0 as measures of evidence. Bayesian p-value.

Bayesian approaches to some common problems in Inference including Linear Regression.

Hierarchical and Empirical Bayes

Brief discussions on Bayesian computational techniques and their applications.

References:

- J.O.Berger : Statistical Decision Theory and Bayesian Analysis
- C.P.Robert : The Bayesian Choice
- J.K.Ghosh, M.Delampady & T.Samanta : An Introduction to Bayesian Analysis: Theory and Methods.
- P.Lee : Bayesian Statistics – An Introduction
- P. Congdon : Bayesian Statistical Modelling (2nd Edition)
- Gelman, et.al. : Bayesian Data Analysis (3rd Edition).

Decision Theory (15)

Decision Problem and two-person game, Nonrandomized and randomized rules,

Risk function, Admissibility of decision rules, Complete, essentially complete, minimal complete and minimal essentially complete classes. Essential completeness and completeness of class of rules based on sufficient statistic and the class of nonrandomized rules for convex loss

Bayes rules, Extended Bayes, Generalized Bayes and Limit of Bayes rules, Admissibility of Bayes rule. Minimax rules, Method for finding minimax rules

References:

- T.S.Ferguson : Mathematical Statistics
- J.O.Berger : Statistical Decision Theory and Bayesian Analysis
- E.L.Lehmann : Theory of Point Estimation

Major 24- Paper Code - STAT 0902 - Elective Paper : Stochastic Process & Time Series Analysis
Marks – 35+15

Stochastic Processes (20)

Conditional Expectations;

Introduction to discrete time discrete space Markov Chains: Markov property, Transition probability matrix , Kolmogorov –Chapman equations, definition of irreducibility, recurrence and aperiodicity;

Stationary Distribution, Basic Theorem of Markov chains, Strong Markov property;

Infinite State space processes, birth-death processes, branching processes;

Continuous-time Markov Chains: Rate matrix, Kolmogorov first and second equations, Poisson process and Renewal process;

Basics of Queuing Theory.

Applications in Simulation: Markov Chain Monte Carlo (MCMC) techniques, Gibbs Sampling.

References:

- S.Karlin and H.M.Taylor : A First Course in Stochastic Process, Vol-1
J Medhi : Stochastic Process
D.R.Cox: : Renewal Theory
S.Ross : Stochastic Process
Basu A.K. : Stochastic Process
Hoel P.G., Port S.C. and Stone C.J : An Introduction to Stochastic Process
Bosq, Denis & Hung T. Nguyen : A course in Stochastic Process.
Bhattacharyya R.N and Waymire, E : Stochastic Processes and Applications

Time Series Analysis (15)

Recapitulation of Classical decomposition models, AR and MA models, estimation of parameters;

Hilbert Space approach: projection, best linear predictor;

ARMA processes, conditions for causality and invertibility;

Time Series Analysis in the spectral domain: Herglotz theorem, spectral distribution function, spectral density function, spectral density of AR, MA and ARMA processes, Periodogram Analysis

Maximum Likelihood methods for estimation;

State Space model: Kalman Filters.

Introduction to Non-Linear Time Series.

References :

- C.Chatfield : The Analysis of Time Series – An Introduction
Brockwell, Peter J. & Davis, Richard A. : Time Series - Theory and Methods
G.E.P.Box ,G.M.Jenkins & G.C.Reinsel : Time Series Analysis – Forecasting and Control
A.Pankratz : Forecasting with Univariate Box-Jenkins Model
G. Jancek and L. Swift : Time Series – Forecasting, Simulation, Applications
Robert H. Shumway & David S. Stoffer : Time Series Analysis and Its Applications

Major 25- Paper Code - STAT 0903 - Elements of Research Methodology: Applied Multivariate Analysis & Resampling Marks – 35+15

Applied Multivariate Analysis (20):

Population and sample principal component and their uses, Plotting Techniques, large sample inferences.

Canonical variables and canonical correlations and their interpretations. Large sample inferences

The orthogonal factor model, Estimation of factor loading, Factor rotation, Estimation of Factor scores, Interpretation of Factor Analysis.

Hierarchical clustering for continuous and categorical variables, Use of Different proximity measures, Agglomerative and Divisive algorithms, Non-hierarchical clustering methods: K-means clustering, Within and Between Cluster Variations, Optimal choice of the number of clusters, Introduction to other related methods: K-mediod, Partitioning Around Mediod(PAM), Fuzzy Clustering.

Classification and discrimination procedures for discrimination between two known populations – Bayes, Minimax and Likelihood Ratio procedures. Discrimination between two multivariate normal populations. Sample discriminant function. Probabilities of misclassification and their estimation. Classification of several populations

References :

- R. A. Johnson & D. W. Wichern : Applied Multivariate Statistical Analysis.
Léopold Simar & Wolfgang Härdle : Applied Multivariate Statistical Analysis.
K. V. Mardia, J. T. Kent, J. M. Bibby: Multivariate Analysis
N. C. Giri : Multivariate Statistical Inference
S. C. Sharma : Applied Multivariate Techniques.
T.W. Anderson : An Introduction to Multivariate Analysis.
G.A.F. Seber : Multivariate Observations.

Resampling (15):

Introduction to the Jackknife, bias reduction using Jackknife, Jackknife Bias estimate of standard statistics, Tukey's Jackknife variance estimate, Introduction to deleted jackknife.

Introduction to the bootstrap method, Bootstrap sampling distribution of a statistic, Bootstrap in regression models, Bootstrap consistency measures.

References :

- B.Efron : The Jackknife, the Bootstrap and other Sampling Plans
B.Efron : Bootstrap methods – another look at jackknife
B.Efron & R.J.Tibshirani : An Introduction to the Bootstrap

J.Shao & D.Tu : The Jackknife and Bootstrap

[Elective Papers may be selected from the following tentative list.]

Practical / Sessional -13 - STAT 0991 – Based on Elective Papers: Marks – 50
Practical / Sessional -14 - STAT 0992 – Based on Elective Papers: Marks – 50

Semester-10

1	Major-26	STAT 1001	Special Paper	35+15
2	Major-27	STAT 1002	Special Paper	35+15
3	Major-28	STAT 1003	Nonparametric Inference and Sequential Analysis	35+15
4	Practical -15	STAT 1091	Based on STAT-1001, 1002, 1003	50
5	Sessional-16	STAT 1092	Project	50

Major 28- Paper Code - STAT 1003 - : Nonparametric Inference and Sequential Analysis
Marks – 35+15

Nonparametric Inference (20):

U-Statistics, Linear Rank Statistic and its asymptotic distribution, Consistency and Asymptotic Relative efficiency. Optimality of tests. Bivariate Sign Test, Hodges-Lehmann Estimators and their properties
M, L and R – estimators, Projection Principle, Density Estimation, Influence function
Simple Regression under nonparametric set up

References

J.Hajek&Z.Sidek : Theory of Rank Tests
R.H.Randles&D.A.Wolfe : Introduction to the theory of nonparametric statistics
T.P.Hettmansperger : Statistical Inference based on ranks
E.L.Lehmann : Theory of Point Estimation
Jun Shao : Mathematical Statistics
D.A.S.Fraser : Nonparametric methods in Statistics
J.D.Gibbons : Nonparametric Inference

Sequential Analysis (15):

Concept Of Sequential Procedures, Wald's Fundamental and Generalized Identity, Random Walk Procedure, Concept of OC and ASN, SPRT and its properties, Optimality of SPRT (under usual approximation), Sequential Estimation.

References

- A. Wald : Sequential Analysis
N. Mukhopadhyay and
Basil M. de Silva : Sequential Methods and their applications
Gobindarajulu : Sequential Statistics

Practical / Sessional -15 - STAT 1091 – Based on Special Papers; Marks – 50

Practical / Sessional -16 - STAT 1092 – Based on Special Papers; Marks – 50

[Special Papers may be selected from the following tentative list.]

**TENTATIVE LIST OF TOPICS FOR ELECTIVE/
SPECIAL PAPERS ****

1. Clinical Trials
2. Survival Analysis
3. Applied Probability in Finance
4. Applied Probability in Genetics
5. Data Mining
6. Econometrics
7. Astrostatistics
8. Advanced Design of Experiments
9. Advanced Inference
10. Operations Research
11. Optimization Technique
12. Functional Data Analysis

**** Particular Elective / Special papers and the corresponding detailed syllabus will be decided by the department in due time.**

The detailed syllabus of some of the elective groups from the above list are given below:

1. Syllabus for Clinical Trial :

Preliminaries: Definition, Different phases, FDA, Controlled study, Bias, Variability, Confounding, Target population, Selection of controls, Ethics of Randomization, Blinding/masking. **6L**

Randomization for balancing treatment assignments: Complete randomization, Random allocation rule, Truncated binomial design, Permuted block design, Urn design and Biased coin design. **8L**

Balancing on known covariates: Stratified randomization, Several covariate adaptive randomization. **4L**

Effect of unobserved covariates: Probability bound on covariate imbalance, Accidental bias. **2L**

Response adaptive randomization: Generalized Friedman's urn model, Play-the-Winner (PW) rule, Randomized Play-the-Winner (RPW) rule, Ternary urn models, Ethical Allocation Designs. **5L**

References:

1. Randomization in Clinical Trials: Theory and Practice.
By *William F. Rosenberger, John M. Lachin.*
2. Statistical aspects of the design and analysis of clinical trials.
By *Brian S. Everitt, Andrew Pickles*
3. Clinical Trials: A Methodologic Perspective.
By *Steven Piantadosi*
4. The Theory of Response-Adaptive Randomization in Clinical Trials.
By *Feifang Hu, William F. Rosenberger*

2. Syllabus for Survival Analysis :

Reliability concepts and measures; components and systems, coherent systems; reliability of coherent systems. Life-distributions, reliability function, hazard rate, common univariate life distributions.

Notions of aging – IFR, IFRA, NBU, and NBUE classes and their duals. Loss of memory property of the exponential distribution; closures of these classes under formation of coherent systems, convolutions and mixtures.

Concepts of Time, order, random censoring and truncation.

Life distributions- Exponential, Gamma, Weibull, Lognormal, Pareto; Linear failure rate, parametric inference, point estimation, confidence intervals, scores, tests based on LR, MLE (Rao-Willks-Wald).

Life tables, failure rate, mean residual life and their elementary properties. Bathtub Failure rate.

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.

Two sample problem- Gehan test, Log rank test, Mantel-Haenszel test, Taron-Ware tests.

Semi Parametric regression for failure rate- Cox's proportional hazards model with one and several convariates.

Rank test for the regression coefficients.

REFERENCES

1. Barlow R E and Proschan F : Statistical Theory of Reliability and Life Testing .
2. Lawless J F : Statistical Models and Methods of Life Time Data.
3. Bain L and Engelhardt J : Statistical Analysis of Reliability and Life Testing Model.
4. Zacks S : Reliability Theory.
5. Cox, D R and Oakes, D : Analysis of Survival Data, Chapman and Hall, New York.
6. Gross A J and Clark VA : Survival Distrubutions : Reliability Applications in Biomedical Sciences, John Wiley & Sons.
7. Elandt Johnson, RE Johnson NL : Survival models and Data Analysis, John Wiley & Sons.
8. Miller, RG : Survival Analysis.
9. Klein, JP & Moeschberger ML : Survival Analysis : Techniques for Censored and Truncated Data.
10. Smith, PJ : Analysis of Failure and Survival Data.

11. Kalbfleisch, JD & Prentice, RL : The Statistical Analysis of Failure Time Data.

3. Syllabus for Applied Probability in Finance :

Introduction – Random walk-Brownian Motion-Martingales-Present Value - Continuous compounding with varying interest rate – Returns - Random walk – Option pricing - One-step Binomial Model - Two-step Binomial Model - General Binomial Tree Model - Black-Scholes formula - Implied Volatility - Properties of the European Call Option Price Given by the Black-Scholes Formula - Properties of the European Call Option Price under General Price Process - Put Options - Portfolio Optimization - Efficient frontier and Tangency portfolio - Efficient portfolio with N risky assets and one risk-free asset - Estimation of Volatility and Value-at-Risk Conditional Value-at-Risk (CVaR) - Capital Asset Pricing Model - Capital Market Line (CML) - Security Market Line (SML) - Security Characteristic Line (SCL) - Testing for CAPM .

References:

1. An Elementary Introduction to Statistics in Finance – S M Ross
2. Statistics in Finance – David Ruppert
3. Statistics of Financial Markets – J Franker, C M Hafner
4. Options, Futures and other derivatives – John C Hull.
5. Stochastic Processes-S.M.Ross

4. Syllabus for Applied Probability in Genetics :

1. Basic Models: Basic Genetics, Wright -Fisher Models, Infinite alleles model, Moran Model;
2. Estimation and Hypothesis testing: Site frequency, Tajma's D, Fu and Li's D, Fay and Wu's H, McDonald-Kreitman's test;
3. Natural Selection;
4. Diffusion process: Infinitesimal mean and variance, Transition probabilities, Hitting probabilities, Stationary measures, Green's function;
5. Multi-dimensional Model: K-allele model, Recombination;

References:

1. Probability models for DNA sequence evolution – Rick Durrett – Springer;
2. Statistical models for QTL mapping – Zehua Chen – CRC Press;

5. Syllabus for Econometrics :

Dummy Variable Regression Models, Dynamic Econometric Models: Distributed-Lag Models, Panel Data Regression Models, Simultaneous-Equation Models, Identification Problems

References

J.Johnston : Econometric Methods

G.G.Judge, et.al.	:	The Theory and Practice of Econometrics (2nded.)
W.Greene	:	Econometric Analysis
E.Malinvaud	:	Statistical Methods in Econometrics
A.Pagan&A.Ullah	:	Non-parametric Econometrics
B.M. Baltagi	:	Econometric Analysis of Panel Data

6. Syllabus for Data Mining:

Generalization of Linear Regression- Ridge Regression, Partial least squares, LASSO and Least angle regression, Principal Components Regression.

Splines, Generalized Additive Models

Generalization of Principal Component Analysis- Kernel Principal Components, Sparse Principal Component Analysis, Independent Component Analysis (ICA). Multidimensional Scaling.

Tree based methods- Classification and Regression Trees (CART), Patient rule induction method (PRIM), Multivariate Adaptive Regression Splines (MARS).

Generalization of Linear Discriminant Analysis- Flexible Discriminant Analysis, Penalized Discriminant Analysis, Mixture Discriminant Analysis. Support Vector Machines.

Concept of Bagging and Boosting

References:

Hastie, T. Tibshirani, R & Friedman, J.	:	The Elements of Statistical Learning
Alan Julian Izenman	:	Modern Multivariate Statistical Techniques
Friedman, B.L et al.	:	Classification and Regression Trees
Hyvarinen, A et al.	:	Independent Component Analysis
Stephen R. and Richard E	:	Independent Component Analysis: Principles and Practice
Tan, Steinbach & Kumar	:	Introduction to Data Mining
Johnson R.A, & Wichern D.W.	:	Applied Multivariate Statistical Analysis

7. Syllabus for Advanced Design of Experiments:

Hadamard matrices and its relationship with BIBD's.

PBIB(2) designs based on group divisible, triangular and L_2 association schemes.

Intrablock analysis of PBIB(2) designs.

Fractional Factorial designs and their analyses. Orthogonal Arrays and their properties.

Response Surface Designs and their analyses. Rotatable Designs.

Change over Designs and their analyses.

References:

D. Raghavarao	:	Constructions & Combinatorial Problems in Designs of Experiments.
---------------	---	---

Department of Statistics, Presidency University, Kolkata

- A. Dey & R. Mukherjee : Fractional Factorial Plans
- A.S. Hedayat et.al. : Orthogonal Arrays : Theory and Applications
- A.I. Khuri & J.A. Cornell : Response Surfaces: Design and Analyses

8. Syllabus for Astrostatistics Special: Total Marks: 50

Group A: (Astrophysics: basic concepts)

20 Marks

Astronomical Instruments, Black body radiation, Doppler effect.
Magnitude, Motion and Distance of Star
Spectra and Spectral Classification of Stars: HR diagram
Saha's equation (only statement)
Radiative transfer equation: Solar limb darkening
Structure and evolution of stars (preliminary ideas)
Preliminary concept of stellar population
Our Galaxy and external galaxies

Group-B (Statistics for Astronomy and Astrophysics)

30 Marks

Generalization of linear regression, unknown cause and effect relation, Measurement errors, Partial least squares
Generalization of Principal Component Analysis- Independent Component Analysis
Clustering and Classification of galaxies
Theory of Spatial Statistics, Spatial cross correlation,
Spatial distribution of galaxies

References:

T.Hastie,R.Tibshirani&J.Friedman : The Elements of Statistical Learning
B.L.Friedman, et al. : Classification and Regression Trees
A.Hyvarinen, et al. : Independent Component Analysis
R.A.Johnson&D.W.Wichern : Applied Multivariate Statistical Analysis
G.J.Babu&e.d.Feigelson : Astrostatistics
B.Basu et al. : An Introduction to Astrophysics