DEPARTMENT OF STATISTICS

M.PHIL STATISTICS

M.PHIL PROGRAMME

SYLLABUS

Effective from the Academic Year 2016-2017

Loyola College (Autonomous)

Chennai- 600 034

DEPARTMENT OF STATISTICS

M.Phil. Restructured Syllabus

(Effective from 2007-08)

Total Credits : 36

Semester	Nature	Code	Course Title	Credits
I	MC	ST 1117	Advanced Statistical Inference	6
I			(any one)	
	MC	ST 1118	Data Warehousing and Data Mining	6
	MC	ST 1119	Statistical Mathematics	
	MC	ST 1120	Stochastic Modelling	
II			(any one)	
	ES	ST 1257	Advanced Distribution Theory	6
	ES	ST 1258	Categorical Data Analysis	
	ES	ST 1259	Design of experiments	
	ES	ST 1260	Inference for Stochastic	
	ES	ST 1261	Inference for survey sampling	
	ES	ST 1262	Interactive computer graphics	
	ES	ST 1263	Mixtures of distributions	
	ES	ST 1264	Nonparametric statistical Inference	
	ES	ST 1265	Queuing Theory	
	ES	ST 1266	Reliability Theory	
	ES	ST 1267	Unbiased Estimating functions	
	MC	ST 1217	Dissertation & Viva Voce	18

*The Department will offer at most any two papers.

ST 1117 : ADVANCED STATISTICAL INFERENCE

Paper - I

Semester I — Credits 6

Category: Major Core (MC) — Hours/Week 6

Objective: To provide strong foundation in statistical inference with recent trends.

Unit 1: Unbiased estimation-Problem of point estimation-Simultaneous unbiased estimation-Convex loss function, Rao - Blackwell theorem, Lehmann - Scheffe theorem-Optimality criteria - Q,T,D and M criteria and their equivalence-Examples.

Unit 2: Equivariant estimation-Principles of invariance and equivariance-Estimation of parameter(s) in location, scale and location-scale models-Estimation of percentiles in location-scale models- Examples

Unit 3: Optimal tests-Monotone Likelihood Ratio-Generalised Neyman – Pearson theorem-Uniformly Most Powerful (UMP) and UMP Unbiased tests for one-parameter

exponential family-Similar tests and Neyman structure-Multiparameter exponential family-Conditional and unconditional tests-UMA confidence bounds-Examples.

Unit 4: Locally optimal and Asymptotic tests-Invariant tests-Locally optimal tests-Asymptotic properties of Likelihood ratio tests-Large sample Chi-square tests.

Unit 5: Order statistics and inference-Inference based on order statistics-Order statistics from life testing experiments-Types of censoring: Type I, Type II and progressive censoring schemes-Estimation based on censored observations from exponential distribution.

Books for study and Reference

1. Arnold, B.C, Balakrishnan, N. and Nagaraja,H.N - A first course on order statistics, John Wiley and Sons,NY., 1992
2. Balakrishnan, N and Aggarwala,R. - Progressive censoring: Theory and Methods. Birkhauser, Boston., 2000
3. Casella,G. and Berger,R.L. - Statistical Inference, Pacific Grove,Wadsworth,CA., 1990
4. David - Order statistics. John Wiley and Sons,NY., 1982
5. Kale, B.K. - A first course on parametric inference. Narosa Publishing House., 2005
6. Lawless,J.F. - Statistical Models and Methods for lifetime data. John Wiley and Sons, NY., 1982
7. Lehmann, E.L. and Casella, G - Theory of Point Estimation, Springer - Verlag, NY., 1998
8. Lehmann, E.L. - Testing Statistical Hypotheses, John Wiley and Sons, New York., 1986
9. Rao, C.R. - Linear statistical inference and its applications. Wiley Eastern Limited., 1973
10. Rohatgi, V.K. and Saleh, A.K.Md.E. - An Introduction to probability and Statistics. John Wiley and Sons, N.Y., 2002

ST 1118 : DATA WAREHOUSING AND DATA MINING

Paper - II

Semester I — Credits : 6

Category: Major Core (MC) — Hours/Week : 6

Objective: To bring out the hidden information in the database using Statistical and Data Mining Techniques.

Unit 1: Introduction: Definition-Motivation- Importance ; Relational databases -Data warehouses - Transactional databases ; Advanced data

and information systems and advanced applications. Data mining functionalities- Concepts - Mining frequent patterns, associations and correlations-Classifications and predictions - Basic Data Mining Tasks - Data Mining Versus Knowledge Discovery in Databases - The development of Data Mining - Data Mining Issues - Data Mining Metrics - Social Implications of Data Mining - Data Mining from a Database Perspective – Interfacing with MS Access, SQL Server and Oracle Databases.

Unit 2: Data processing - Descriptive data summarization - Data cleaning - Data integration and transformation - Data reduction - Data discrimination. Related Concepts: Database/OLTP Systems - Fuzzy Sets and Fuzzy Logic - Information Retrieval - Decision Support Systems - Dimensional Modeling - Multidimensional Schemes - Data Warehousing- ETL Process – OLAP - Machine Learning - Pattern Matching. Data Mining Techniques: A Statistical Perspective on Data Mining - Point Estimation - Models Based on Summarization - Bayes Theorem - Hypothesis Testing - Regression Techniques - Linear Regression Model - Discrete choice models - Logistic regression - Multinomial logit regression - Probit regression - Logit vs. Probit - Time series models - Classification and regression trees .

Unit 3: Data cube computation and date generalization - Efficient methods for data cube computations - Data cube and OLAP technology - Mining frequent patterns, associations and correlations - Constrain-Based association mining - Issues in Classification - Classification and precictions - Classification by decision tree induction — Statistical-Based Algorithms – Regression - Bayesian Classification - Distance-Based Algorithms - K Nearest Neighbors - Decision Tree-Based Algorithms - Neural Network-Based Algorithms - Rule-Based Algorithms - Classification by association rule analysis - Association Rules: Large Item sets - Basic Algorithms - Apriori Algorithm - Multiple-Level Association Rules - Quantitative Association Rules - Correlation Rules.

Unit 4: Cluster analysis - Type of data in cluster analysis - A Categorization of major custering methods - Partitioning methods - Hierarchical methods - Density based methods - Grid based methods - Model based clustering methods - Clustering-high dimensional data - K – Means Clustering - Clustering with Genetic Algorithms - Clustering with Neural Networks - Similarity and Distance Measures – Outliers - Mining stream, time series and sequence data- Data Parallelism

Unit 5: Types of predictive analytics - Predictive mocels - Descriptive models - Decision models - Predictive analytics - Web Mining: Web Content Mining – Crawlers - Harvest System - Virtual Web View - Personalization - Web Structure Mining - Web Usage Mining. Applications and Trends in Data Mining in Financial Data Analysis – Retail Industry – Tele Communications – Bio Logical Data and Intrusion Detection

Books for Study and References:

1. Jiawei Han and Micheline Kamber - Data Mining, Morgan Kaufmann Publishers., 2006
2. Margaret,H,Dunham and Sridhar,S - Data Mining,Dorling Kindersley India Pvt Ltd., 2006
3. Olivia Parr Rud - Data Mining Cook Book, John Wiley & Sons, NY., 2006
4. Ian H. Willy Frank - Data Mining ,Morgan Kaufmann Publishers., 2006
5. Bhavani Thuraisingam - Data Mining, CRC Publications, New Delhi., 2006

ST 1119 : STATISTICAL MATHEMATICS

Paper - II

Semester I — Credits 6

Category: Major Core (MC) — Hours/Week 6

Objective: To give strong mathematical foundations leading to advanced studies in statistics.

Unit 1 : Fundamentals of Measure and Integration theory. Introduction – field, sigma fields and measures – Extension of measures – Monotone Class Theorem – Caratheodory Extension Theorem – Approximation Theorem – Lebesgue - Satieties measures and distribution functions.

Unit 2: Measurable Functions and Integration.Basic Integration theorems – Monotone Convergence Theorem – Extended Monotone Convergence Theorem – Fatou's lemma – Dominated Convergence Theorem – Jordan – Hahn decomposition Theorem – Radon Nikodym Theorem and related results – Lebesque decomposition theorem.

Unit 3: Probability theory – probability spaces- random variables – Types of convergence of a sequence of random variables – convergence in probability – as. convergence – convergence in r^{th} mean – convergence in distribution – interrelationship among various types of convergence.

Unit 4: Conditional probability and Expectation. Introduction – general concept of conditional probability and conditional expectation- conditional expectation given a sigma field – properties of conditional expectations.

Unit 5: The Central Limit Theorem. Introduction – Inversion formula – truncation inequality - :Levy's theorem – Lindeberg's theorem – stable distributions – Infinitely divisible distributions – examples.

Book for study

Ash,R.B. - Real Analysis and Probability , Academic Press, NY., 1972

Books for reference

1. Chow,Y.S. and Teicher,H - Probability Theory. Independence, Interchangeability, Martingales, Springer-Verlag,NY., 1988
2. Loeve,M. - Probability Theory, Springer-Verlag, NY., 1977

ST 1120 : STOCHASTIC MODELLING

Paper - II

Semester I — Credits : 6

Category: Major Core (MC) — Hours/Week : 6

Objective: To lay strong foundation in Applied Probability and give an exposure to its applications.

Unit 1: Markov chains and Markov processes - Martingales: Definitions and examples – Super Martingales and Sub Martingales.

Unit 2: Renewal Processes: Definition of Renewal Process and related concepts – Some examples of Renewal Processes – Special Renewal Processes – Poisson Process viewed as a Renewal Process – Replacement Models – Renewal equations and the elementary Renewal Theorem – The renewal Theorem.

Unit 3 : Branching Processes : Discrete time Branching Processes – Generating function relations for Branching Processes – Extinction Probabilities – Continuous time Branching Processes – Extinction Probabilities for continuous time Branching processes.

Unit 4 : Queuing Models : Queues with Combined Arrivals and Departures – Queues with Priorities for Service – Tandem or Series Queues- Analysis of Queues by Imbedded Markov Chains – All derivations included.

Unit 5 : Simulation : Design Implementation – examples – Different Methods of generating random numbers – Congruential – Mixed – Multiplicative –

Quadratic Congruential method – Testing Procedures. Known distributions – generation of random numbers

Books for study and reference

1. Chow,Y.S. and Teicher,H - Probability Theory: Independence, Interchangeability, Martingales, Springer-Verlag, NY., 1988
2. Karlin,S and Taylor,H.M.-A First Course on Stochastic Processes, Academic Press., 1975
3. Ravindran,A.,Philips,D.T and Solberg, J.J. - Operations Research – Principles and Practice.John Wiley and sons., 2004
4. Ross,S.M. - Stochastic Processes. John Wiley and Sons, NY., 2004
5. Taha,H. - Operations Research – An Introduction (Fourth Edition), Macmillan Publishing Company., 1987