BHARATHIAR UNIVERSITY, COIMBATORE - 641 046.

M.Sc. STATISTICS

Course Structure and Scheme of Examinations with effect from 2014-15 onwards

Sem	Title of the Subject	Inst. Hrs/ Week	Examination				Credits
			Dura-tion	CIA Marks	External Marks	Total Marks	
I	I Real Analysis	6	3	25	75	100	4
	II Distribution Theory	6	3	25	75	100	4
	III Sampling Theory	6	3	25	75	100	4
	IV Linear Algebra	6	3	25	75	100	4
	Statistics Practical I	3	-	-	-	-	-
	Elective I: Survival Analysis	6	3	25	75	100	4
II	V Probability Theory	6	3	25	75	100	4
	VI Statistical Inference I	6	3	25	75	100	4
	VII Demographic Methods	6	3	25	75	100	4
	VIII Operations Research	6	3	25	75	100	4
	Statistics Practical I	3	3	25	75	100	4
	Elective II: Financial Mathematics	6	3	25	75	100	4
III	IX Statistical Inference II	6	3	25	75	100	4
	X Multivariate Analysis	6	3	25	75	100	4
	XI Statistical Quality Control and Reliability	6	3	25	75	100	4
	XII Econometrics	6	3	25	75	100	4
	Statistics Practical II	3	-	-	-	-	-
	Elective III: Data Mining	6	3	25	75	100	4
IV	XIII Linear Models and Design of Experiments	6	3	25	75	100	4
	XIV Stochastic Processes	6	3	25	75	100	4
	XV Applied Regression Analysis	6	3	25	75	100	4
	XVI Computer Practical using R	6	3	25	75	100	4
	Statistics Practical II	3	3	25	75	100	4
	Project/Dissertation with Viva-voce*	9	-	50	100	150*	6
	Total	**135**	**-**	**575**	**1675**	**2250**	**90**

*** Project Report – 120 Marks, Viva voce- 30 Marks**

Note : The Syllabus for the above papers (**except the Core Paper II Distribution Theory, Elective Paper I - Survival Analysis and Elective Paper III - Data Mining**) be the same as prescribed for the academic year 2012-13. The syllabus for the Core Paper II Distribution Theory, Elective Paper I - Survival Analysis and Elective Paper III - Data Mining are furnished below:

Paper II. DISTRIBUTION THEORY

Unit I: Distribution function of a random variable and its decomposition into discrete and continuous parts. Standard discrete and continuous distributions. Distribution of a function of a random variables. Truncated, compound distributions. Non-central chi square, t, and F distributions (no derivations).

Unit II: Jointly distributed random variables. Joint, marginal, and conditional distributions. Marginal and expectations. Distributions of functions of several variables. Order statistics. Distributions of functions of order statistics.

Unit III: Characteristic function and its properties. Statements of inversion theorem and Levy's continuity theorem. Their applications.

Unit IV: Probability Distribution: Cauchy distribution –Laplace distribution –Pareto distribution–Log Normal distribution –Power Series distribution –Logarithmic Series distribution - Distribution of functions of random variables

Unit V: Distribution of quadratic forms. Cochran's theorem Independence of quadratic forms. Independence of linear and quadratic forms.

References

1. Anderson, T.W. (1984). An Introduction to Multivariate Analysis, 2/e, John Wiley, New York.
2. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley, New York.
3. Hogg, R.V. and Tanis, E.A. (1983). Probability and Statistical Inference, 3/e, McMillan, New York.
4. Hsu, Hwei. (2009). Probability, Random Variables, and Random Processes, Schaum's Outlines, Tata McGraw-Hill, New Delhi.
5. Johnson, N.L. and Kotz, S. (1969). Continuous Distributions, Vol. 1 and Vol. 2, Houghton and Miffin, New York.
6. Rao, C.R. (1973). Linear Statistical Inference and Its Applications, Wiley Eastern, New Delhi.
7. Rohatki, V.K. and Saleh, A.K.Md. E. (2002). An Introduction to Probability and Statistics, John Wiley, New York.
8. Ross, S.M. ( ). Probability Models,
9. Mukhopadhyay, P. (1996). Mathematical Statistics, New Central Book Agency, Calcutta.

Elective – I : SURVIVAL ANALYSIS

Unit -I

Introduction – censor data – outline of survival analysis-goals- basic data layout-examples Kaplan Meler survival curves – general features. The Log-Rank test-for two groups, several groups to the alternatives to the log rank test.

Unit -II

Cox PH model ML estimation of the Cox PH model Hazard Ratio adjusted survival curves

PH assumptions Cox likelihood.

Unit- III

Evaluating the proportional Hazards – Background – checking the PH Assumptions. Overview –graphical approach – log-log plots – observed versus expected plots – goodness of fit testing approach – time - dependent covariates.

Unit- IV

The stratified Cox procedure – general stratified Cox model – no interaction assumption - several stratification variables. Time dependent variables – extended Cox model for time dependent variables Hazard ratio formula – assessing time dependent variables – extended cox likelihood.

Unit- V

Parametric survival models – Probability density function in relation to the hazard and survival function- Accelerated failure time assumption - Exponential, Weibull and log logistic.

<u>Reference:</u>

1.David G. Kleinbacum and Michal Klein 2008 Survival analysis Springer.

2.Cox D. R and Oakes D. 1984 Analysis of Survival Data, Chapman and Hall, London.

3.Hosmer D.W and lemeshow.S 1989.Applied survival analysis. Newyork : wiley

4.Kalbfleisch J.D and Prentice R.L 1980. The statistical analysis of failure time data, Newyork : wiley

5.Lee E T 1980 statistical methods for survival data analysis . Belmont C A : Wadsworth

ELECTIVE III - DATA MINING

Unit I

Introduction to Data Warehousing - Characteristics. Data Mart – Types of data marts. Meta data – Data Model for a Data Mart. Online Analytical Processing - Introduction- OLTP and OLAP systems. Developing a Data Warehouse – How to build – Tools for Data Ware housing.

Unit II

Introduction to Data Mining – From data warehousing of Data mining – Steps of Data Mining. Data mining algorithm – Data base segmentation – Predictive modelling – Link analysis. Tools for data mining.Data mining applications.Data mining techniques.

Unit III

Classification – Introduction - decision tree – Building a decision tree.Over fitting and Pruning – Decision tree rules. Estimating predictive – Accuracy of classification method. Cluster analysis – Introduction – Measuring distance between two clusters – Hierarchical clustering – Non Hierarchical clustering – K-means algorithm

Unit IV

Association rules – Introduction - Discovering association rules in Transaction rule. Direct Hashing and Pruning (DHP) – Dynamic Itemset Counting (DIC). Performance Evaluation of Algorithm. Software for association rule mining .

Unit V

Web data Mining – Introduction .Web technology and characteristics.Locality and hierarchy in the web.Web content mining. Web document clustering – Web usage mining – Web mining software.

Books for References

1. C.S.R Prabhu, “Data warehousing ”.

2. G.K Gupta, “Introduction to Datamining with Case studies”.

3. GalitShmueli, Nitin R. Patel & Peter C. Bruce, “Data Mining for Business Intelligence”