M.Sc. BIOSTATISTICS –FRAMEWORK
(2013 – 14 onwards CBCS Pattern)

S.No	Study Components	Title of the Paper	Hours /Week	Marks CA	Marks ESE	Credits
		First Semester				
1.	Core	Probability and Distribution theory	6	25	75	4
2.	Core	Biostatistical Inference - I	6	25	75	4
3.	Core	Population Studies	6	25	75	4
4.	Core Elective	Statistical Genetics	6	25	75	3
5.	Core	Biostatistical Computing - I	6	40	60	4
		Second Semester				
6.	Core	Biostatistical Inference - II	5	25	75	4
7.	Core	Applied Regression Analysis	5	25	75	4
8.	Core	Research Designs and Analysis	5	25	75	4
9.	Core Elective	Categorical Data Analysis	5	25	75	3
10.	Non Major Elective	English for careers	5	25	75	3
11	Core	Biostatistical Computing – II	5	40	60	4
		Third Semester				
12.	Core	Applied Multivariate Analysis	5	25	75	4
13.	Core	Sample Survey Designs	5	25	75	4
14.	Core	Survival Analysis	5	25	75	4
15.	Core	Epidemiological Methods	5	25	75	4
16.	Non Major Elective	Basics of Bioinformatics	5	40	60	3
17	Core Elective	Exploratory Data Analysis using SYSTAT/SPSS	5	40	60	3
		Fourth Semester				
18	Core	Clinical Trials and Data mining	6	25	75	4
19	Core Elective	SAS Programming	6	25	75	3
20	Core Elective	Forecasting and decision making Techniques	6	25	75	3
21	Core	Statistical Analysis using R Language	6	40	60	4
22	Core	Project*	6	40	60	4
		TOTAL				81

* Internal – 40 marks   Project & Viva voce – 60 marks

Non Major Elective offered by the Dept. of Statistics:
1. DATA ANALYSIS USING SPSS for M.A HRM./M.Sc. PBPB

CBCS FOR PG

(2008 – 2009 onwards)

Study Components	No. of Papers	Credit Per Paper	Total Credit
Core:			
Major	11	4	44
Practical	3	4	12
Project		4	4
Elective::			
Major	5	3	15
Non Major	2	3	6
Soft Skills	4	2	8
Internship		2	2
Total			91

MAJOR PAPER I
PROBABILITY AND DISTRIBUTION THEORY (PBS/CT/1001)

6 hours /week
Credits: 4

UNIT 1:

Probability: Sample space - Probability axioms - discrete and continuous probability space - conditional probability – Bayes' theorem and its applications - screening tests, sensitivity, specificity, predictive value positive and negative - independence of events - simple problems.

UNIT 2:

Random variables - Probability distribution of a random variable - discrete and continuous random variables - functions of random variable -multiple random variable-independent random variable, Convolution, covariance, correlation, moments, conditional expectation-Classical Central Limit theorem and its applications.

UNIT 3:

Univariate distributions: Negative binomial, Hyper-geometric, Normal, Exponential, Lognormal, Cauchy, Beta, Gamma, Laplace, Logarithmic, Pareto and Weibull (mean & variance only)-Compound and mixture of distributions relating to Binomial, Poisson and Normal distributions -Truncated distributions-
Order statistics and their distributions.

UNIT 4:

Bivariate discrete distributions: Bivariate binomial, Bivariate Poisson, Multinomial distribution. (Mean, variance, covariance & correlation and applications of the above distributions). Bivariate Continuous distribution: Bivariate Exponential.

UNIT 5:

Bivariate Normal - Multivariate Normal Distributions – M.g.f - marginal and conditional distributions- independence of variables - linear transformation - distribution of sample mean vector.

BOOKS FOR STUDY:

1. Daniel, W.W (2005): A foundation for analysis in the Health Sciences, John Wiley & Sons.(chapter 3 and 4)
2. Hogg, R.V., Mckean, J. W. and Craig, A. T. (2006): Introduction to Mathematical Statistics, Sixth edition, Pearson Education. (for unit 5).
3. Meyer, P.L. (1970): Introductory probability and statistical Applications, 2ed, Oxford & IBH, New Delhi.
4. Rohatgi, V.K. and Saleh, A.K. (2001): An Introduction to Probability and Statistics, John Wiley & Sons. (Chapters 1,2 and 6 only)

BOOKS FOR REFERENCE:

1. Johnson, N.L., Kotz, S. and Balakrishnan, N. (2004): Continuous univariate distributions, Vols. II, Wiley series in Probability and Mathematical Statistics and I.
2. Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997), Discrete Multivariate Distribution, John Wiley & Sons, New York.

MAJOR PAPER II
BIOSTATISTICAL INFERENCE – I (PBS/CT/1002)

6 hours/week
Credits: 4

UNIT 1:

Point Estimation: Optimal properties of a point estimator - consistency and unbiasedness. Minimum variance unbiased estimators - Cramer-Rao lower bound - Chapman-Robbins bound. Examples

UNIT 2:

Sufficient Statistic - factorization criterion –One parameter exponential family (examples) - completeness - bounded completeness - invariance (definition only) -Rao-Blackwell theorem - Lehmann-Scheffe theorem

UNIT 3:

Methods of estimation: Method of Maximum likelihood (Statement of properties only) - maximum likelihood estimation based on grouped data, Censored data (type I, II and random censoring - MLE for exponential distribution only) - EM algorithm (steps only) - Method of minimum chi-square and method of moments.

UNIT 4:

Interval estimation – Based on small and large samples - sufficient statistic and confidence limits- fiducial interval - Confidence Interval for parameters of Normal distribution(s), Confidence Intervals based on large samples – examples.

UNIT 5:

Bayesian inference: Prior and posterior distributions, Baye's rule, Baye's estimator, conjugate prior, minimax estimator, admissibility rule, Bayesian Interval estimation

BOOKS FOR STUDY:

1. Daniel, W. W (2005): A foundation for analysis in the Health Sciences, John Wiley & Sons.
2. Hogg, R. V., McKean, J. W. and Craig, A. T. (2006): Introduction to Mathematical Statistics, Sixth edition, Pearson Education.
3. Rohatgi,V.K. and Saleh, A.K.(2001):An Introduction to Probability and Statistics, John Wiley & Sons. (Chapter 8,unit 8.1 to 8.7)

BOOKS FOR REFERENCE:

1. Casella, G. and Berger, R.L. (2002): Statistical Inference, Thomson Duxbury, Second edition.
2. Kale, B.K, (1999): A First Course on Parametric Inference, Narosa Publication , New Delhi.
3. Lehmann, E.L. and Casella, G. (2006): Theory of Point Estimation, Second edition, Springer Texts in Statistics
4. Pagano, M. and Gauvreau, K. (2000): Principles of Biostatistics, Second Edition, Duxbury.
5. Rao, C.R. (2002): Linear Statistical Inference and its applications, Second edition, Wiley series in Probability and Statistics.

MAJOR PAPER III
POPULATION STUDIES (PBS/CT/1003)

6 hours / week
Credits: 4

UNIT 1:

Introduction to Demography: Sources of Demographic data – completeness of vital registration – Chandrasekar Deming index - Nature, Scope and importance of demography – relationship with other disciplines. Analysis of age distribution - percent distribution – percent change by age – graphical representation of age data – population pyramid – sex ratio – aging of population – measures of aging of population.

UNIT 2:

Mortality and life tables: Crude and specific rates – infant mortality rate – standardized death rates – direct and indirect method of standardization. Life tables – constructions and uses – abridged life table – construction – Reed Merrell method – Greville's method – Chiang's method

UNIT 3:

Fertility: Crude and specific rates – General fertility rate – Total fertility rate – Age specific fertility rate - Gross reproduction rate – Net reproduction rate – parity progression ratio – child women ratio – fertility differential – determinants of fertility.

UNIT 4:

Migration: Definition and concepts – measures of migration – national growth rate method – residual method – balancing equation – survival ratio method – selectivity and differential – push and pull factors.

UNIT 5:

Population growth: Concept of stable and stationary population – measurement of population growth – arithmetic, geometric and exponential - population projection and estimation – different methods of projection –

logistic curve – Pearl-Reed method – method of selected points – component method population projection .Urbanism and Urbanization – urban population growth – components of urban population growth.

BOOKS FOR STUDY:
1. Pathak, K.B. and Ram, F. (1992): Techniques of demographic analysis, Himalaya publishing house.
2. Ramkumar, R. (1986): Technical Demography, Wiley Eastern Limited.

BOOKS FOR REFERENCE:

1. Cox, P. (1959): Demography, Cambridge University Press
2. Keyfitz, (1985): Applied mathematical Demography, Second edition, Springer-Verlag, New York.
3. Shrivastava, O.S. (1995): Demography and population Studies, Vikas Publishing house private limited.

6 hours/week
Credits: 3

UNIT 1:

Elements of genetics – Physical basis of heredity – cell structure, chromosomes and genes – interaction of genes – concept of genotype and phenotypes – linkage and crossing - over genetic maps.

UNIT 2:

Mendel's law of inheritance - Laws of segregation and independent assortment – concept of generation – Application of Chi-square tests

UNIT 3:

Sex – linked genes – concept of gene frequency, random mapping, Hardy Weinberg equilibrium, matrix theory of random mapping with applications – detection and estimation of linkage from back cross.

UNIT 4:

Multiple allelic system -elementary aspects of the study of the human blood group - statistical problems in human genetics.

UNIT 5:

Quantitative Genetics: Qualitative Vs Quantitative traits- Polygenic traits-normal distribution-Types of gene actions-Heritability -Selection methods-mating methods.

BOOKS FOR STUDY:

1. Arora, M.P and Sandhu, G.S. (1992): Genetics, Himalaya Publishing House, Delhi.
2. Elrod, S.L. (2006): Theory and Problems of Genetics, Fourth Edition, Tata McGraw Hill, New Delhi
3. Meyyan, R. (2002): Genetics and Biostatistics, First Edition, Saras Publications India.

BOOKS FOR REFERENCE:

1. Kowles, R. (2006): Solving Problems in Genetics, Springer (India), New Delhi.
2. Verma, P.S. and Agarwal, V.K.(1985): Genetics, S.Chand co Ltd.

MAJOR PAPER V
BIOSTATISTICAL COMPUTING –I (PBS/CP/1001)

6 hours/week
Credits: 4

1. Presentation of data – graphical and numerical
2. Calculation of simple, partial and multiple correlation coefficients (three variables only)
3. Fitting of Binomial, Poisson, Negative Binomial, Hyper geometric and normal distributions.
4. Estimation of parameters by the method of moments, maximum likelihood.
5. Confidence interval for proportion(s), mean(s), variance and ratio of variances.
6. Standard asymptotic and exact tests of significance with regard to proportion(s), mean(s), variance(s), coefficient of correlation(s), regression coefficient, one-way analysis of variance, Independence tests by contingency tables of order (p x q), (p, q <= 5), determination of sample sizes.
7. Rates and ratios, measures of fertility and mortality.
8. Construction of life tables and Abridged life table
9. Population estimation using logistic, Makhem and Gompertz curves.

Note: Emphasis is on Biostatistical applications

Max.   : 100 Marks
Internal   : 40 Marks
External   : 60 Marks

MAJOR PAPER VI
BIOSTATISTICAL INFERENCE –II (PBS/CT/2004)

5 hours /week
Credits: 4

UNIT 1:

Statistical hypotheses: formulation of hypotheses testing - Neyman -Pearson fundamental lemma: Most powerful tests - Uniformly most powerful (UMP) tests for distributions with monotone likelihood ratio (H_o: $\theta \leq \theta_0$ versus H_A: $\theta > \theta_0$ and it's dual).

UNIT 2:

Generalized Neyman Pearson Lemma - UMP test of two sided hypotheses (H_o: $\theta \leq \theta_1$ or $\theta \geq \theta_2$ versus H_A: $\theta_1 < \theta < \theta_2$) for one parameter exponential family of distributions - problems relating to normal distribution only. Unbiased tests: definition, UMPU test for one parameter exponential family (H_o: $\theta = \theta_0$ versus H_A: $\theta \neq \theta_0$).

UNIT 3:

Similar tests: uniformly most powerful similar test for one parameter exponential family (H_o: $\theta_1 \leq \theta \leq \theta_2$ versus H_A: $\theta < \theta_1$ or $\theta > \theta_2$) – Locally most powerful (LMP) tests: definition, statement of the theorem for construction of LMP tests only, problems relating to normal distribution only.

UNIT 4:

Likelihood ratio tests (LRT): Definition, LRT for Binomial ,LRT for Normal distributions (one and two populations) only.

Sequential probability ratio test (SPRT): Definition, Operating Characteristic function and Average sample number, optimum property of SPRT (statement only).

UNIT 5:

Non-parametric tests: Measurement scales, Sign test, Wilcoxon Signed – Rank test, Wald – Wolfowitz runs test, Median test, Mann-Whitney test, Kolmogorov-Smirnov Test (one sample and two samples) - Kruskal-Wallis test - Friedman test.

BOOKS FOR STUDY:

1. Conover, W. J. (2006): Practical Non-parametric Methods in Statistics, Second Edition, (Unit 5)
2. Daniel, W.W. (2006): Biostatistics: A foundation for analysis in the Health Sciences, 7th edition John Wiley & Sons. (Unit 5)
3. Rohatgi,V.K. (1984): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern. (Chapter 14-14.5 for SPRT)
4. Rohatgi, V.K. and Saleh, A.K. (2001): An Introduction to Probability and Statistics, John Wiley & Sons. (Chapters 8 - 8.3, 9, 10- 10.1,10.2,10.6, 11 - 11.3)

BOOKS FOR REFERENCE:

1. Gibbons, J.D. (1971): Nonparametric Statistical Inference, McGraw Hill.
2. Lehmann, E.L. (1986): Testing Statistical Hypothesis, John Wiley and Sons.

MAJOR PAPER VII
APPLIED REGRESSION ANALYSIS (PBS/CT/2005)

5 hours/week
Credits: 4

UNIT 1:

Simple linear regression-Assumptions, estimation of model parameters, standard error of estimators, testing of hypotheses on slope and intercept (β's), interval estimation of model parameters, prediction interval of a new observation, coefficient of determination, regression through origin, Tests of hypotheses about parallelism and equality of intercepts.

UNIT 2: Multiple linear regressions

Standard Gauss Markov setup, least square estimation of model parameters, variance covariance of least squares estimators, estimation of error variance

Tests of hypotheses – significance of regression (ANOVA, R^2 and adjusted R^2), individual regression coefficients, subset of regressor variables, general linear hypotheses- Confidence intervals and regions, prediction intervals, detecting hidden interpolation.

UNIT 3:

Model adequacy checking - residual plots for checking normality homoscedasticity and autocorrelation, detection of outliers. Test for Lack of fit of the model. Durbin – Watson test for autocorrelation. Analytical methods for selecting a transformation generalized and weighted least squares- Detection of influential observations – Cooks statistic, DFFITS, DFBETAS.

Variance stabilizing transforms and transforms to linearize the model, analytical methods for selecting a transform, generalized and weighted least squares.

Dummy (or indicator variables) – general concepts and their use

UNIT 4:

Multicollinearity – sources, effects, diagnostics, Methods of dealing with multicollinearity (collection of additional data, model respecification, Ridge regression).

Selection of Variables – forward selection, backward elimination and stepwise regression (algorithms only)

UNIT 5:

Nonlinear regression – transformation to a linear model, their use and limitations, initial estimates (starting values), parameter estimation using iterative procedures – Gauss-Newton, steepest Descent, Marquardt's compromise.

BOOKS FOR STUDY:

1. Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003): Introduction to Linear regression analysis, third edition, John Wiley and Sons, Inc. Chapters 2, 3, 4, 5, 6, 8 (8.1,8.2), 9, 10, 12 (12.1,12.3,12.4), 14 (14.1.2)
2. Zar, J.H. (2006): Biostatistical Analysis, fourth edition, Pearson education. Chapter 18 (18.1, 18.2,18.4,18.5)

BOOKS FOR REFERENCE:

1. Draper, N.R. and Smith, H. (2003): Applied Regression Analysis, third edition, John Wiley and Sons, Inc.
2. Johnston, J. (1984): Econometric methods, third edition, McGraw Hill International.

MAJOR PAPER VIII
RESEARCH DESIGNS AND ANALYSIS (PBS/CT/2006)

5 hours /week
Credits: 4

UNIT 1:

Single factor Analysis of variance-Two-factor ANOVA with unequal and equal replication (with/without interactions)- fixed, random and mixed effects model- Multiple comparison tests-Tukey, Newman-Keul, Scheffe test.

UNIT 2:

Randomized block designs-repeated measures designs with and without blocking, Latin square design, BIBD, ANOCOVA (for CRD and RBD).

UNIT 3:

General factorial design - blocking in a factorial design - 2^k Full factorial design: diagrammatic presentation of main effects and first order interactions, model, analysis of single as well as more than one replicates using ANOVA.

Total confounding of 2^k design in 2^p blocks, p = 1,2. Partial confounding. Fractional factorial design-the one-half fraction of the 2^k designs

UNIT 4:

Three-level factorial designs - 3^2 designs: contrasts for linear and quadratic effects, statistical analysis of 3^2 designs.
Nested designs- two stages, Split-plot, Split- Split plot design experiment.

UNIT 5:

Bioassay**:** Introduction, direct assays: the nature of direct assays, precision of estimates and the design of direct assays.

The quantitative dose Response relations: indirect assays, the dose response regression, standard curve estimation, slope estimation, and simultaneous trial estimation.
Concept of response surface methodology (RSM)

Note: Emphasis on design construction and inference.

BOOKS FOR STUDY:

1. Das, M.N. and Giri N.C. (2006): Design and Analysis of
   Delhi. (For unit 5)
   Experiments, Second Edition New Age International (P) Ltd., New
2. Montgomery D.C (2006): Design and Analysis of Experiments 5th ed., Wiley
   India Edition. (For Unit 1,2 3,4,5)
3. Zar, J.H. (2007): Biostatistical Analysis, fourth edition, Pearson Education.
   (for unit 1,2)

BOOKS FOR REFERENCE:

1. Berger, P.D. and Maurer, R.E. (2002): Experimental designs,
   ediion, S. Karger publications
2. Finney, D.J.(1951): Statistical Methods in Biological Assay, Charles
   Giffin & co limited, London .
3. Govindarajulu, Z. (2000): Statistical techniques in Bioassay, 2nd
   Thomson Duxbury.

MAJOR PAPER IX
(CORE ELECTIVE)
CATEGORICAL DATA ANALYSIS (PBS/CE/2002)

5 hours/week
Credits: 3

UNIT 1:

Nature of Categorical data - Contingency tables and their distribution: Binomial and Multinomial sampling. Comparing proportions in two-by-two tables: Difference of proportions - Relative risk - Odds Ratio - Properties of Odds Ratio - relationship between Odds Ratio and Relative Risk - Odds ratio in I x J tables - Nominal and Ordinal Measures of Association.

Inference for Contingency tables: Interval estimation for difference of proportions, odds ratio, log odds ratio and relative risk.

Testing Independence in Two-Way tables: Pearson and Likelihood-ratio chi-square tests - Yate's correction for continuity-Residuals for cells in a contingency table-Partitioning chi-squared.

UNIT 2:

Testing Independence for Ordinal Data-Trend tests for 2 x J tables-Exact Inference for small samples - Fisher Exact Test for 2 x 2 tables.

Association in Three-Way Tables: Partial Tables - Conditional and Marginal Odds Ratios - Homogeneous Association - Cochran-Mantel-Haenszel methods.

UNIT 3:

Components of Generalized Linear Models-Logit models for Binary data-Binomial GLM for 2 x 2 contingency tables

Logistic regression: parameter interpretation - maximum likelihood estimate - test of overall regression and goodness of fit - Wald test, deviance statistic, LR test, score test-Logistic regression diagnostics.

UNIT 4:

Logit model with Categorical predictors: Linear logit model for I x 2 tables - Cochran-Armitage Trend test. Logit models for multiway contingency tables.

Multinomial Logit models: Logit models for Nominal responses- Base-line category logits-estimating response probability

Logit models for ordinal responses: Cumulative logits - proportional odds model.

Probit models – complementary log-log link model (description only).

UNIT 5:

Poisson log-linear model for count data- ML estimation, goodness of fit test. Loglinear Models for Two-way Tables: Independence Model-Interpretation of parameters- Poisson regression for rates.

Models for matched pairs: Comparing dependent proportions-McNemar's test, small sample test for comparing matched proportions.
Measures of agreement-Kappa measure - Weighted Kappa - Bradley-Terry model for paired preferences. Generalized estimating equation methodology (basic idea).

BOOKS FOR STUDY:

1. Agresti, A. (2002): Categorical data analysis, John Wiley & Sons.
2. McCullagh, P. and Nelder, J.A. (1991): Generalized Linear Models, Second Edition, Chapman and hall, London.

BOOKS FOR REFERENCE:

1. Agresti, A. (1991): An Introduction to Categorical data analysis, John Wiley & Sons.
2. Armitage, P. and Berry, G. (1987): Statistical methods in Medical Research, Blackwell Scientific Publications, USA.
3. Deshpande, J.V., Gore, A.P. and Shanubhogue, A. (1995): Statistical Analysis of Non Normal Data, New Age International Publishers Ltd., New Delhi.
4. Hardin, J.W., and Hilbe, J.M. (1994):Generalized Estimating Equation, Chapman and Hall, London.
5. Hosmer, D.W. and Lemeshow, S.(1989): Applied Logistic Regression, John Wiley & Sons Inc.

MAJOR PAPER X
BIOSTATISTICAL COMPUTING –II (PBS/CP/2002)

5 hours/week
Credits: 4

1. Diagnostic checks for model adequacy
2. Two-way classification with equal / unequal number of observations per cell (model with interaction). Multiple comparison tests.
3. Repeated measures ANOVA.
4. Analysis of covariance in one-way and two-way data.
5. 2^n Factorial Experiments, Analysis of Single Replicate of 2^n, (n=2, 3, 4).
5. 2 x 3 Factorial experiment, 2 x 3 x 3 Factorial experiment with and without blocking.
6. Total and partial confounding
7. Fractional Replication.
8. 3^2 Factorial Experiment
9. Split plot design
10. Power curve
11. Nonparametric tests
12. Chi-square test for Independence
13. Interval estimation for odds ratio, log odds, relative risk
14. Cochran Armitage trend test
15. Cochran Mantel Haenzel test
16. Tests relating to simple and multiple regression models

Note: Emphasis is on Biostatistical applications

Max   : 100 Marks
Internal   : 40 Marks
External   : 60 Marks

MAJOR PAPER XI
APPLIED MULTIVARIATE ANALYSIS (MBA31)

5 hours /week
Credits: 4

UNIT 1:

Hotelling's T^2 statistic - applicatons of T^2 statistic - single sample, two samples, paired comparisons - confidence region and simultaneous comparison of component means. One-way and two-way Multivariate Analysis of Variance. Testing hypothesis for single covariance matrix, testing the equality of covariance matrices.

UNIT 2:

Principal component analysis: Extraction of components - properties and characteristics of components - total variation, relative importance, standardization of variables and components, special covariance structures - interpretation of principal components.

Factor analysis: orthogonal factor model-parameter estimation by principal component and maximum likelihood method-factor rotation - factor scores.

UNIT 3:

Canonical correlation analysis: Extraction of canonical correlations and their variable - testing the significance of canonical correlation - interpretation of canonical variables.

UNIT 4:

Discrimination and classification： Classification problem - standards of good classification - procedures of classification into one of two populations with known probability distributions - evaluation of classification function - Fisher's linear discriminant function.

UNIT 5:

Cluster analysis: Distance and similarity measures - hierarchical clustering techniques - agglomerative techniques - single linkage, complete linkage, average linkage methods - nonhierarchical clustering methods-K means method.

BOOKS FOR STUDY:

1. Johnson, R.A. and Wichern, D.W. (2002): Applied Multivariate Statistical Analysis, 5th edition, Pearson Education, Asia.

BOOKS FOR REFERENCE:

1. Anderson, T. W. (2003): An Introduction to Multivariate Statistical Analysis, 3rd edition, John Wiley & Sons.
2. Hair, J.F., Anderson, R.E., Tatham. R.L. and Black, W.C. (2006): Multivariate Data Analysis, 5th edition, Pearson Education, Asia.
3. Morrison, D.F. (1990): Multivariate Statistical Methods, 3rd edition, McGraw Hill.

MAJOR PAPER XII
SAMPLE SURVEY DESIGNS (MBA32)

5 hours/week
Credits: 4

UNIT 1:

Basic methods of sample selection: Inclusion probabilities, Simple random sampling with and without replacement, Probability proportional to size sampling with and without replacement – estimation of mean and its sampling variance, Horwitz-Thompson estimator.

UNIT 2:

Systematic sampling Introduction- Linear systematic sampling, schemes for population with linear trend. Stratified sampling-sample size allocations, comparison with simple random sampling and systematic sampling.

UNIT 3:

Ratio estimation-unbiased ratio type estimator, almost unbiased ratio estimator, Jackknife ratio estimator, bound for bias. Product estimator, Regression estimator, Difference estimator – bias and mean square error under srswor.

UNIT 4:

Cluster sampling: Estimate of mean per element and its variance, Optimum cluster size. Multistage sampling: Two stage sampling-units of equal size-variance of estimated mean under SRS in both the stages.

UNIT 5:

Sources of non-sampling errors, Randomized Response method- Warner's model, Simmons randomized response model-Adaptive sampling and Snowball sampling (procedure only).

N.B The emphasis is on the sampling procedure and estimation of mean, total and standard error

BOOKS FOR STUDY:

1. Mukopadhyay, P. (1998): Theory and Methods of Survey Sampling, Prentice-Hall India Ltd, New Delhi.
2. Sampath, S. (2000): Sampling Theory, Narosa Publications Ltd, New Delhi.

BOOKS FOR REFERENCE:

1. Cochran, W.G. (1989): Sampling Techniques, third edition, Wiley Eastern, Delhi.
2. Singh, D. and Choudhary, F.S. (1986): Theory and analysis of sample survey designs, Wiley Eastern Ltd.
3. Sukhatme, P.V. and Sukhatme, B.V. (1984): Sample Survey Methods and its Applications, Indian Society of Agricultural Statistics, India.
4. Thompson, S.K.(2002): Sampling, Wiley Series in Probability and Statistics, Second Edition.

MAJOR PAPER XIII
SURVIVAL ANALYSIS (MBA33)

5 hours/week
Credits: 4

UNIT 1:

Concept of Time and event. Order and Random Censoring. Censoring mechanism and truncations. Survival, hazard and density functions. Mean and median residual life and their elementary properties. Ageing classes - IFR, IFRA, NBU, NBUE, HNBUE, DMRL and their duals, Bathtub Failure rate

UNIT 2:

Life distributions - Exponential Gamma, Weibull, Lognormal, Pareto, Linear Failure rate-Parametric inference: Point estimation, Confidence Intervals, Scores, tests based on LR and MLE, Partial likelihood estimation-log logistic distribution.

UNIT 3:

Estimation of survival function–reduced sample method, actuarial estimator, Kaplan-Meier Estimator-life table estimation.

UNIT 4:

Nonparametric methods: Gehan Test, Log rank test. Mantel - Haentzel Test, Tarone - Ware tests, Efron Tests.

UNIT 5:

Semi-parametric regression for failure rate - PH assumptions - Cox's proportional hazards model with one and several covariates-goodness of fit. Extended Cox's model.

BOOKS FOR STUDY:

1. Deshpande, J.V., Gore, A.P. and Shanubhogue, A. (1995): Statistical Analysis of Non Normal Data, New Age International Publishers Ltd., New Delhi
2. Klein, J.P. and Moeschberger, M.L. (2003): Survival Analysis-Techniques for Censored and Truncated data
3. Miller, R.G. (1981): Survival Analysis, John Wiley and Sons.

BOOKS FOR REFERENCE:

1. Barlow, R. E. and Proschan, F. (1975): Statistical Theory of Reliability and Life testing, Holt, Rinehart and Winston, New York.
2. Johnson, E.R.E. and Johnson, N.L. (1980): Survival models and Data Analysis, John Wiley and Sons.
3. Lee, C.T. (1997): Applied survival analysis, John Wiley.

MAJOR PAPER XIV
EPIDEMIOLOGICAL METHODS (MBB41)

6 hours/week
Credits: 4

UNIT 1:

Introduction: Definition and Scope of Epidemiology – achievement in epidemiology – Epidemiology measures: Determinants of disease-Measures of disease frequency – Prevalence and incidence rates – cumulative incidence rate or risk – case fatality.

UNIT 2:

Causation and Prevention in Epidemiology: The concept of Cause – Single and Multiple causes – Factors in causation – Guidelines for causation –causation pie-Hill's Bradford model. The scope of Prevention – Level of Prevention: Primary prevention – secondary prevention and Tertiary prevention.

UNIT 3:

Epidemiology of Communicable and Non-Communicable diseases, environmental epidemiology- epidemics and epidemic investigation, surveillance, prevention, screening, the application of epidemiology in health services and research policy.

UNIT 4:

Types of Epidemiological studies: Observational and experimental studies – Descriptive studies – Ecological studies – Cross-sectional studies – Cohort studies - Random control trials –Field trials and Community trials – Potential errors in epidemiological studies (Random error, Sample size calculation, systematic error).

UNIT 5:

Planning an Epidemiological Project (Basic Concepts): Introduction – Choosing a project – Describing the population – Questionnaire preparation – Analysis and presentation of data – Final evaluation

Critical review of scientific literature-scientific writing; evidence based medicine-health research for policy decision

BOOKS FOR STUDY:

1. Beaglehole, R., Bonita, R. and Kjellstrom, T. (2005): Basic Epidemiology , World Health Organization. (WHO).
2. Park, K.(2005) : Textbook of Preventive and Social Medicine. 20th edn. Jabalpur: Bhanot Publishers.

BOOKS FOR REFERENCE:

1. Gerstman, B.B.(2003) : Epidemiology Kept Simple: An Introduction to traditional and Modern Epidemiology ,Second Edition, John Wiley and Sons.
2. Gordis, L.(2008): Epidemiology, Fourth Edition, Saunders.
3. Jekel, J.F., Katz, D.L. and Elmore, J.G. (2001): Epidemiology, Biostatistics and preventive Medicine, second edition, Saunders.
4. Timmreck, T.C.(2002): An Introduction to Epidemiology, Third Edition,Jones and Bartlett.

- Principles of Epidemiology- CDC Atlanta, Health Research Publication of WHO

MAJOR PAPER XV
(CORE ELECTIVE)
EXPLORATORY DATA ANALYSIS
(USING SYSTAT/SPSS) (EBS41)

6 hours/week
Credits: 3

1. Descriptive measures.
2. Simple, Partial and Multiple correlation
3. Simple Linear Regression
4. Multiple Regression
5. Regression Diagnostics: Examining Residuals, identifying and
   a. handling influential observations – Collinearity
6. Parametric and Non parametric tests
7. Power curve
8. ANOVA, MANOVA
9. Generalized linear models
10. Categorical Data Analysis: Tests of Association –Simple Logistic
11. Regression
12. Factor analysis
13. Cluster analysis
14. Discriminant analysis
15. Survival analysis: Actuarial and Kaplan-Meir method,
16. Non parametric tests-Gehan's test,Tarone Ware test,Log rank test

Note: Emphasis is on Biostatistical applications

Max   : 100 Marks
Internal   : 40 Marks
External   : 60 Marks

MAJOR PAPER XVI
CLINICAL TRIALS AND DATA MINING (MBA42)

6 hours/week
Credits: 4

UNIT 1:

Introduction to Clinical Trials: Historical background - The need and ethics of clinical trials. Organization and Planning: Main features of the study protocol - Selection of patients - Treatment schedule - Evaluation of patient response - follow-up studies-GCP/ICH guidelines.

UNIT 2:

Phases of clinical trials: Phase I, II, III and IV trials.
Basic study designs: Randomized control study, Nonrandomized concurrent control study, historical controls, cross-over design, withdrawal studies, group allocation design, hybrid designs, studies of equivalency (description only). Methods of Randomization: Fixed allocation randomization - Stratified randomization- Adaptive randomization - Unequal Randomization.
Blinding and Placebos: Unblinded, Single blind and Double-blind trials - conduct of double blind trials- Crossover trials- Design, analysis and interpretation.

UNIT 3:

Statistical methods for determining Trial size: Method for dichotomous response variable, continuous response variables, repeated measures, cluster randomization and equivalency of intervention. Multicenter trials.
Data management: Interim analysis - Case report form design - database design - data collection system for good clinical practice. Terminologies used in Clinical research.

UNIT 4:

Clinical data mining: Definition and importance – different tasks. Classification: Issues-general approach to solving a classification problem – Types of classification: Decision tree classifier – Hunt's algorithm – Rule based classifiers – Nearest Neighbor classifiers- Measures for accuracy of classifiers.

UNIT 5:

Association analysis: Association rule-support and confidence-rule mining problem-Apriori algorithm: Frequent item set generation. Evaluation of association

patterns. Cluster Analysis: Partitioning methods – Hierarchical methods - Genetic algorithm

BOOKS FOR STUDY:

1. Pocock, S.J. (1991): Clinical Trials – A Practical Approach, John Wiley and Sons.
2. Tan, P.N., Steinbach, M. and Kumar,V.(2008): Introduction to Data Mining, Pearson Education. (chapter 4 and 5)

BOOKS FOR REFERENCE:

1. Dunham, M.H. (2008): Data Mining: Introductory and Advanced Topics, Pearson Education.
2. Fleiss, J.L. (1986): The design and analysis of clinical experiments. John Wiley & Sons.
3. Han, J. and Kamber, M. (2002): Data Mining: Concepts and Techniques, Morgan Kaufman Publishers.
4. Meinert, C.L. (1986): Clinical trials: Design, Conduct and Analysis Oxford university press.
5. Piantadosi, S. (2005): Clinical Trials - A Methodological Perspective, Second Edition, Wiley series in probability and Statistics.

<div align="center">

MAJOR PAPER – XVII
(CORE ELECTIVE)
SAS PROGRAMMING (EBB34)

</div>

<div align="right">

5 hours/week
Credits: 3

</div>

UNIT 1: INTRODUCTION TO SAS

Introduction to SAS System - Components of SAS – Fundamental concepts of Programming –Introduction to SAS programs – Naming Conventions in SAS – SAS Library – Types of SAS Libraries – Creating and Maintaining SAS Libraries – SAS Datasets– Components of SAS Program and Execution – DATA STEP – PROC STEP - Diagnosing and Correcting syntax errors – Exploring SAS libraries and files – PROC PRINT - Producing list report – Sub-setting – PROC COMPARE - Logical and Special Operators – Enhancing Output -Formatting Data Values

UNIT 2: CREATING AND MANIPULATING SAS DATASETS

Creating Dataset using In stream Data - Reading form External raw data files -Examining Data errors and log errors – Assigning attributes to variables – Integration with Excel spreadsheets – Reading SAS data sets and Creating variables – SAS FORMATS & INFORMATS – Date Formats - Conditional processing – Dropping and keeping variables – Concatenating , Merging and Combining SAS data sets (one-to-one,one-to-many, many-to-many)- Match-Merging SAS datasets .

UNIT 3: BASIC DATA TRANSFORMATIONS

Iterative Data processing:DO Loop and SAS ARRAY processing.Producing Summary reports using PROC FREQ, PROC MEANS, PROC REPORT and PROC TABULATE procedures – Producing Graphical Charts (Bar, Pie etc.) using PROC CHART, PROC GPLOT and PROC GCHART Procedures – Enhancing Charts.

UNIT 4: ADVANCED STATISTICAL TRANSFORMATIONS

Descriptive Statistics and ANOVA: Selecting Random samples from SAS datasets - Performing Descriptive statistical analysis, Examining distributions, Estimating Confidence intervals for mean and Hypothesis testing of single Mean

- Two-Sample t- Tests using PROC CORR , PROC TTEST and PROC UNIVARIATE procedures –Paired t-Tests. Regression Analysis using PROC REG Procedure-PROC PHREG and PROC LIFE-One-Way ANOVA, Two-Way ANOVA with/ without interactions using GLM procedure –Multiple comparison tests - Non-Parametric ANOVA – Categorical data analysis.

UNIT 5: ADVANCED SAS PROGRAMMING CONCEPTS
        Connecting to Oracle and Other Databases and Creating Linked Libraries – PROC ACCESS Procedure for Mining Data from RDBMSs and Excel Files – PROC DBLOAD for Loading Data into RDBMSs and EXCEL files – Introduction to Datawarehousing Business Intelligence Concepts using SAS – Introduction to SAS Macros.

BOOKS FOR STUDY:
1. Delwiche, L.E and Slaughter, S.J. (2003): The Little SAS book, A primer 3rdEd. by SAS publishing
2. Everitt, B.SMand Der, G.(2004): Handbook of Statistical Analyses UsingSAS, 2nd edition by SAS publishing
3. Carpenter, Art (1998): Carpenter's Complete Guide to the SAS Macro Language. SAS Institute, Cary, NC.

BOOKS FOR REFERENCE :
1. O'Rourke, N, Hatcher, L and Stepanski, E (2005): A Step-by-Step Approach to Using SAS for Univariate and Multivariate Statistics,2ndEdition, SAS Press and John Wiley Sons Inc.
2. Michele M. Burlew (2006): SAS Macro Programming Made Easy, 2nd Edition by SAS Publishing [ISBN-10: 1-59047-882-7]
3. Ralph Kimball (1996): The Data Warehouse Toolkit, John Wiley and Sons.

MAJOR PAPER XVIII
(CORE ELECTIVE)
FORECASTING AND DECISION MAKING TECHNIQUES

6 hours/week
Credits: 3

UNIT 1:

Business forecasting – role of forecasting in Business – steps in forecasting – methods of forecasting, exploring data patterns- measuring forecasting errors

UNIT 2:

Moving average and smoothing methods: Naïve models-forecasting methods based on averages: simple, moving averages and double moving averages-Exponential smoothing methods: Holt's method, Winters' methods- forecasting based on smoothing, adaptive smoothing.

UNIT 3:

Exploratory Time Series Analysis: Components of time series- forecasting trend-seasonality-seasonally adjusted data-cyclical and irregular variations-forecasting a seasonal time series

UNIT 4:

Stationary processes: a) moving average) (MA), b) Auto Regressive (AR), c) ARMA and (d) AR integrated MA (ARIMA) models, Box-Jenkins methodology-steps only-choosing a forecasting technique.

UNIT 5:

Introduction, decision making-problem, process and environment, decision under uncertainty-Laplace criterion, Maximin or Minimax criterion, Maximax or Minimin criterion and Horwicz criterion, Decision under risk-Expected Monetary Value (EMV), Expected Opportunity Loss (EOL) and Expected Value of Perfect Information (EVPI), Decision tree analysis

BOOKS FOR STUDY

1. Hanke, J.E. and Wichern, D.W. (2007): Business Forecasting, eighth edition, Pearson education, Asia.
2. Kanthi Swarup, Gupta, P.K. and Manmohan (2003): Operations Reasearch . Sultan Chand & Sons
3. Taha, H.A. (2006): Operations Research, 7th edition, Collier MacMillan.

MAJOR PAPER XIX
STATISTICAL ANALYSIS USING R LANGUAGE (BSB31)

5 hours/week
Credits: 4

1. Descriptive measures.
2. Simple Linear Regression
3. Multiple Regression
4. Parametric and Non parametric tests
5. Estimation of sample mean and variance using SRS
6. Power curve
7. ANOVA
8. MANOVA
9. Categorical Data Analysis: Tests of Association –Simple Logistic Regression
10. Factor analysis
11. Cluster analysis
12. Discriminant analysis
13. Survival analysis: Actuarial and Kaplan-Meir method, Non parametric tests-Gehan's test,Tarone Ware test,Log rank test

Note: Emphasis is on Biostatistical applications

Max   : 100 Marks
Internal   : 40 Marks
External   : 60 Marks

MAJOR PAPER XX
(PROJECT & VIVA VOCE)

6 hours/week
Credit: 4

Student will undertake a project work by collecting primary data or secondary data related to biomedical field. The analysis will be done using a statistical package (SPSS 14.0/SYSTAT).

Max   : 100 Marks
Internal   :40 Marks
External   : 60 Marks

NON MAJOR ELECTIVE
(for postgraduate students of other departments)
DATA ANALYSIS USING SPSS (NBS2)

4 hours/week
Credits: 3

OBJECTIVE: To impart practical working knowledge of SPSS to the students.

Internal-40 marks Practical - 60 marks

1. Diagrams – Simple bar, Pie diagram, Multiple bar diagram(clustered), Subdivided (stacked bar).

2. Frequency distribution – Univariate (categorical data, quantitative data), Bivariate (cross tabulation).

3. Graphs – Histogram, Box- Whiskers plot.

4. Measures of location, dispersion, skewness and kurtosis – raw data, continuous data.

5. Correlation coefficient and scatter diagram – Karl-Pearson's and spearman's rank correlation.

6. Regression equations.

7. One sample t-test.

8. paired t-test (repeated measures t-test).

9. Two independent samples t-test.

10. Chi-square test for goodness of fit (one way Chi-square test).

11. Chi-square test for independent samples (two way Chi-square test).

12. ANOVA – one way and two way.

PATTERN OF THE QUESTION PAPER
PG –M.Sc (BIOSTATISTICS)

Time : 3 Hrs   Maximum marks : 75

PART – A
(10 x 2 = 20)
Answer any TEN questions out of TWELVE
Two questions from each unit

PART – B
(5 x 5 = 25)
Answer any FIVE questions out of SEVEN
At least one question from each unit

PART – C
(3 x 10 = 30)
Answer any THREE questions out of FIVE
One question from each unit