NATIONAL UNIVERSITY

Fourth Year Syllabus
Department of Statistics

Four Year B.Sc Honours Course
Effective from the Session : 2013–2014
National University

Subject: Statistics
Syllabus for Four Year B. Sc Honours Course
Effective from the Session: 2013-2014

Year wise Papers and marks distribution

FOURTH YEAR

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>Paper Title</th>
<th>Marks</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>243601</td>
<td>Multivariate Analysis</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243603</td>
<td>Advance Sampling Technique</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243605</td>
<td>Demography</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243607</td>
<td>Computer Programming (FORTRAN &amp; R)</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243609</td>
<td>Stochastic Process</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Optional Papers (Any two)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>243611</td>
<td>Econometrics</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243613</td>
<td>Operational Research and Research Methodology</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243615</td>
<td>Epidemiology &amp; Biostatistics</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>243616</td>
<td>Lab: 9 Multivariate Analysis</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>243618</td>
<td>Lab: 10 Advance Sampling Technique</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><strong>Optional Practical Papers (Any two)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>243620</td>
<td>Lab: 11 Econometrics</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>243622</td>
<td>Lab: 12 Operational Research &amp; Demography</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>243624</td>
<td>Lab: 13 Epidemiology &amp; Biostatistics</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>243626</td>
<td>Viva Voce</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td>1000</td>
<td>40</td>
</tr>
</tbody>
</table>
Detailed Syllabus

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>243601</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Title:</td>
<td>Multivariate Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Distribution of quadratic forms in normally distributed random variables. The Fisher-Cochran’s theorems. The idea of multivariate data and multidimensional scaling. Multivariate normal distribution-importance and properties, the distribution of linear combinations of normally distributed variables. Estimation of mean vector of multivariate normal distribution (MND), Marginal and conditional distribution of MND. Distribution of partial and multiple correlation coefficients, the characteristic function, moments, maximum likelihood estimates of the mean vector and the co-variance matrix in MND. Multinomial distribution, Hotelling T², its distribution, properties & uses. Wishart’s distribution and its properties.

Discriminant function & analysis, Mahalanabis D², Classification-general classification problem, optimal classification rules for two populations, classification with several populations. Population principal components, summarizing sample variation by principal components, estimates of the principal components and their variances, large sample properties applications, Testing for the equal correlation structures.

Latent variable models, the orthogonal factor models, method of estimation MLE of factor loading and specific variances, adequacy of the model and choice of the number of factors, Factor rotation, factor scores, applications.

Books Recommended:
5. Srivestava, M.S.: An Introduction to Applied Multivariate Carter

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>243603</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Title:</td>
<td>Advance Sampling Technique</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ratio Estimation:** Definition, examples, properties, Bias Mean square error, approximate variance, condition for unbiased ratio estimation, Unbiased ratio type estimate.

**Regression Estimation:** properties and limitations, bias and approximate variance, estimate of the variance, Bias of the linear regression estimate. Comparative merits and demerits of ratio and regression estimate.


Stratified cluster sampling, its advantage, estimate, bias , standard error and efficiency. Comparison with simple random sampling, Systematic sampling and usual stratified random sampling.

Sampling of unequal clusters with unequal probability with replacement. The Horvitz-Thompson estimator and its properties. The Hansen-Hurwitz estimator and its properties. Selection of samples with probability proportional to size (PPS) with replacement –use, application, estimate and standard error.
Sampling of unequal clusters with unequal probability without replacement—different selection methods. Brewer’s, Durbin’s, Stamford’s, PPS systematic, Raj’s, Murthy’s and Rao-Hartley-Cochran methods of selection. Detailed study on the related formulae estimates, variances, estimates of variances for these methods.


Books Recommended:

2. Hansen, Hurwitz, Madow : Sample Survey Methods and Theory
4. Sukhatne, P.V. : Sampling Theory of Surveys with Applications
5. Islam, M.N. : An Introduction to Sampling Methods, Mullick & Brothers
7. Kish, L : Survey Sampling

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>243605</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Title:</td>
<td>Demography</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Source and types of errors in Demographic data, age heaping, digital preference, application of different indices (Such as whiples and Myre’s) in detection errors, Dual record system, Chandra sekhar Deming formula, methods of graduating age data.


Life Table: Definition, use and functions of life tables, different types of life table, construction of life tables. Actuarial life tables, its construction and application.


**Migration:** Concept of migration, Internal and international migration.

**Books Recommended:**

1. Barely, G.W.: Technique of Population Analysis
3. Gullaume, J, Hansch and Others: Introduction to Demographic Analysis
6. William Brass: Demography of Tropical Africa

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>243607</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Paper Title:** Computer Programming (FORTRAN & R)

**Programming Basics:** Programming concepts and preliminaries, steps involved in computer programming algorithm, decision table, flowchart, solution of statistical and mathematical problems using flowchart, program data files and database.

**FORTRAN:** constants and variables, arithmetic expression, input/ output statement, arithmetic assignment statement, control statements, type statement, intrinsic function, repetition program structure, subscripted variables. DO loops, logical expressions, FORTRAN subprograms (function and subroutine), use of files on external storage, FORTRAN programs of different statistical and mathematical problems.

**R Programming:**
Introduction: the S language, R versus S-plus, R and the window system, downloading R, installing R, using R: a simple R session with some basic R commands, case-sensitivity, recall and correction of previous commands, assignments and expressions, simple manipulations of numbers, getting help with existing R functions and features, executing commands from a file, diverting output to a file.

**Vectors:** Generating regular sequences, creating vectors, vector arithmetic, logical vectors, character vectors, missing values, selecting and modifying subsets of a vector, combining elements of two vectors.

**Objects:** Their modes and attributes, class of an object.

**Matrices:** Creating matrices, matrix operations, selecting and modifying subsets of a matrix, linear equations and inversions of matrices, Eigen values and Eigen vectors, combining rows (cases ) and columns (variables) of two matrices.

**List and Data Frames:** Making list and data frames, attaching and detaching data frames.

**Probability Distributions:** R names for different distributions, obtaining densities, cumulative probabilities, quantiles and random samples from different distributions.
**Control Statements in R:** Conditional execution with if statement, repetitive execution with for, repeat and while.

**Writing R functions:** simple examples, arguments and defaults, assignments within functions, returning multiple objects as output.

**Numerical Mathematics:** Newton-Raphson method, numerical integration and solution of equations.

**Importing Data in R:** Reading text files with read. Table and scan functions; importing data from other systems like SAS, SPSS, S-Plus, Excel, Stata; editing data.

**Statistical Analyses:** Basic statistical techniques, correlation and regression, estimation of parameters of multiple regression model, inference in multiple regression, partial correlation, multiple correlation and related tests, model selection, fitting polynomial regression, orthogonal polynomials, splines, examination of residuals, outliers, influential points.

**S-PLUS:** Limitations and advantages compared to R, important differences with R.

**References**

1. R reference manual
2. S-Plus reference manual

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>243609</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Title</td>
<td>Stochastic Process</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Stochastic Processes:** Definition of different type of stochastic process, Recurrent events, Renewal equation, Delayed recurrent events, Number of occurrences of a recurrent event, random walk and ruin problem. Markov chains: Transition matrix, Higher transition Probabilities, Classification of States and chains, Ergodic Properties, Evaluation of P^n.


**Reference:**

6 Ross, S.M. : Stochastic Process
7 Bartlett, M.S. : An Introduction to Stochastic Process

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Title:</td>
<td>Econometrics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heteroscedasticity: Basic concept, estimation of parameters, (GLS and WLS and other problems).
Model Misspecification: Basic concept and consequence, Omitted variable, irrelevant variable and measurements errors.

Errors in Variables: Basic concept, causes, parameter estimates and their properties.

Lag variable: Definition, Mean lag, median lag, Lag operator, different types of lag model and their estimation.

Multicollinearity and Autocorrelation: Detection, tests of their presence, estimation of parameters in those cases.

Identification: Different types of identifications, estimation issue: Indirect least square, (ILS), two-stage least square, (2SLS).

Simultaneous Equation: Concept of endogenous, exogenous, predetermine, variable, Structural model, reduced form of a structural model and their relations.

Model specification: Basic concepts and consequences, Omitted variables, Irrelevant variables and Measurement errors.

Dummy Variables with Regression Model: Basic concepts, ANOVA models, caution in the use of dummy variables, ANOVA models with two qualitative variables, regression with a combination of quantitative and qualitative repressors. The ANCOVA models, testing the structural stability of regression model and comparing two regression by dummy variable, comparison with chow test, use of dummy variables in seasonal and piecewise linear regression, panel data regression models.

Books Recommended:
1 Gujratli, O : Basic Econometrics
3 Kontsoyiannis, A : Theory of Econometrics.
4 Bidges, J.L. : Applied Econometrics
5 Desai, M : Applied Econometrics

<table>
<thead>
<tr>
<th>Paper Code</th>
<th>Marks:100</th>
<th>Credits: 4</th>
<th>Class Hours: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Title:</td>
<td>Operational Research and Research Methodology</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Operational Research:


**Integer Linear Programming:** Formulation of integer linear programming by cutting plane and branch and Bound methods, Solution of mixed integer. Integer programming problem by cutting plane method.

**Transportation Problem:** Basic feasible solution of transportation problem, Optimality test, Degeneracy, Unleal Lane variations and Least time transportation problem, Transshipment problem, Assignment problem.

**Research Methodology:**

**Concept and Idea of Research:** Basic research. Action research. Operations research. Social research. Stages of research. Types of Study: Exploratory study, Descriptive study and explanatory study,

**Research Problem:** Research problem and non-research problem, criteria for choosing a research problem. Identifying, defining and justifying a research problem. Objective, hypothesis and variables: General objective. Specific objective. Immediate objective and Ultimate objective. Formulating research hypothesis. Identifying variables. Study design: Reliability and validity. Validity theats, Types of study design. Issues relating to the choice of study design.

**Sampling:** A brief review of the sampling methods appropriate for social research. Sampling and non-sampling error, Determination of sample size.

**Data Collection:** Collection of qualitative and quantitative data, construction of questionnaire and schedule. Interviewing technique.

**Data Management:** Editing and coding and preparing tabulation plan, Errors associated with these processes. Analysis of data: Types of analytical problems such as variable transformation. Univariate, Bivariate and Multivariate analysis. Choice of the variables, Determination of the number of variables. Choice of statistical model for analysis. Use of dummy variables and their significance. Presentations of research findings. Report writing: Steps of report writing, with special reference to conclusion, Summary and recommendation. Appendix Bibliography, Footnotes.

**Books Recommended:**

1. Dantzig : Linear Programming and Extension
2. Hadly : Linear Programming
3. Vajda, S : Mathematical Programming
4. Gdass, S. I. : Linear Programming
8. Bailey : Social Research Methods
**Epidemiology concepts: Definition and Scope of Epidemiology:** Key Issue in Epidemiology: Sources of Data of Community Health, Vital Statistics and Morbidity Data. **Descriptive Epidemiology:** Person, Place and time, Analytic Epidemiology: causality.

**Types of Epidemiological Research:** Experiment, Quasi Experiments and Observational Studies; Study types in Population Health Research: Randomized Controlled Trials, Cohort Study, Case-Control study, Cross-sectional Study, Ecological, Before and After Study. Measures of disease

**Frequency:** Incidence and prevalence, Estimation of Risk and Rate, Age, period and cohort effects.

**Measures of Association:** Relative Risk, Odds Ratios; Relative Protection and Relative Risk Reduction, Risk Difference, Number Needed to Treat, Attributable Risk, Comparison of proportions from several samples, Standard Error of estimators, Test of hypothesis.

**Bio-statistics**

**Basic concepts of lifetime Distributions:** Interrelationships, Survival functions and Hazard Rate, Rate, Mean Residual Life Function and Median Life Time Incomplete Data: Truncation and Censoring, Right and Left Censoring, Type I and Type II censoring, Random censoring, Progressive type II censoring.

**Parametric Survival Distribution:** Likelihood Functions, Joint Density Function of Failure time Data for Different Censoring Mechanisms, Exponential, Two Parameter Exponential, Weibull, Lognormal, Normal and Gamma distribution, Inference Procedures (Estimation and Test for Small and Large Samples).


**Comparison of Two Survival Curves:** Log-Rank Test, Hazard Ratio, Confidence Interval for Hazard Ratio, Mantel-Heanszel Test, Stratified Log-Rank Test median survival, Non Proportional Hazards, Comparing Follow-up in Two Groups, comparison of More than Two Groups.

**Books Recommended:**

1. Kenneth, J. Rothman, Sander Greenland - Modern Epidemiology
2. Newman, S. - Biostatistical Method in Epidemiology
3. Sahai, H. and A. Khurshid - Statistics in Epidemiology
4. Lawless, J.F - Statistical Models and Method for Life Time Data
5. Lee, Elisa, T - Statistical Methods for Survival Data Analysis

**Lab-9: Multivariate Analysis**

**Lab-10: Advance Sampling Technique**

**Sampling:** Ratio & regression estimation. Sampling with unequal prob. PPS sampling, Stratified cluster sampling, estimate, standard, error and efficiency, different selection methods of sampling of unequal cluster with unequal probability, 2-stage, 3-stage sampling, double sampling estimates standard errors, etc.

**Lab-11: Econometrics**

Estimation and tests of parameters of multiple regression model. Estimation of lag regression model. Estimation Parameters in generalized linear model assuming the presence of auto Correlation, heteroscedasticity and autocorrelation coefficient, test of heteroscedasticity and auto correlation, test of multicollinearity. Identification and estimation of simultaneous equation model. (1LS, 2SLS, method)

**Lab-12: Operational Research & Demography**

Solution of linear Programming Problems by Simplex and Revised simplex method. Dual Simplex Algorithm, Solution of transportation Problems through linear programming method Graphical solution of (2×2), (2×n), and (m×2) Games and solution of (m×2) Games and Solution of (m×n) Games by simplex method. Age-sex composition, Population Pyramid, detection of errors in age data, Population growth, Calculation of various rates and ratios of fertility, mortality, nuptiality, Standardization of rates and ratios Construction of life tables. Population projection.

**Lab-13: Epidemiology & Biostatistics**

**Epidemiology:** Measures of incidence and prevalence. Relative risk, odds ratios. Standard errors, Confidence intervals, Attributable risk.

**Biostatistics**

Non Parametric Estimation of survival probabilities and their standard error for ungroup and grouped data, Construction of survival curves and their confidence belts. Graduation of life data by plotting procedures. Fitting appropriate parametric model to observed data and testing goodness of fit to fitted models, (kolmogrove Smirnov test, L.R. test) Construction of confidence limits of life parameters for the fitted models.