STCT: 1.3: STATISTICAL COMPUTING

Unit 1: Programming in C / R. (The purpose of this unit is to introduce programming with the eventual aim of developing skills required to .write statistical software. Should there be previous exposure to programming, this unit can be replaced with a more advanced unit in object oriented programming in C++ or Java. Topics should include Simple syntax, loops, pointers and arrays, functions, input/output, and linking to databases.

20 Hrs

Unit 2: Numerical analysis and statistical applications. (The purpose of this unit is to apply programming skills in me1hods and algorithms useful in probability, statistics and data analysis. Topics should include numerical integration, root extraction, random number generation, Monte Carlo integration, matrix computations, drawing random samples : known univariate probability distributions -both discrete and continuous.

15 Hrs

Unit 3: Analysis of interesting data sets using known techniques on a suitable statistical package such as R / MINITAB / SAS / SPSS / JMPIN; Topics should include graphics, descriptive statistics, representation of multivariate data, hypotheses testing, analysis of variance and linear regression.

15 Hrs

Books for Reference:

Crawley, M. The R programming language, Shareware.
Keminghan, B. W. and Ritchie, D. M. (1988): The C Programming Language, Second Edition, Prentice Hall.
Press, W. H., Teukolsky, S. A., Vellering, W. T. and Flannery, B. P. (1993): Numerical recipes in C, Second Edition, Cambridge University Press.
Ryan, B. and Joiner, B. L. (2001):MINITAB Handbook, Fourth Edition, Duxbury.
Thisted, R. A. (1988): Elements of Statistical Computing. Chapman and Hall.

Practicals based on STCT: 1.3

STCT: 1.4: SAMPLE SURVEYS AND STATISTICS FOR NATIONAL DEVELOPMENT

A SAMPLE SURVEYS

Unit 1: Basic finite population sampling techniques (SRS WR/ WoR, stratified, systematic), related
problems of population mean estimation, allocation problems in stratified sampling.

10 Hrs

Unit 2: Unequal probability sampling: PPS WR / WoR methods (including Lahiri's scheme) and related estimators of a finite population mean (Hansen-Hurvitz and Desraj estimators for a general sample size and Murthy's estimator for a sample of size 2).

10 Hrs

Unit 3: Ratio and regression estimators based on SRS WoR method of sampling, two-stage sampling with equal. number of second stage units, double sampling, cluster sampling.

10 Hrs

B. STATISTICS'FOR NATIONAL DEVELOPMENT

Unit 4: Economic development: growth ip. per capita income distributive justice. Indices of development, Human Development Index. Estimation of National Income - product approach, income approach and expenditure approach. Population growth in developing and developed countries. Population projection using Leslie matrix. Labour force projection.

10 Hrs

Unit 5: Measuring inequality of incomes, Gini coefficient, Theil's measure. Poverty measurement- different issues, measures of incidence and intensity, combined measures, eg. Indices due to Kakwani, Sen. etc.

10 Hrs

Books for Reference:

Choudhary, A and Mukherjee, R (1989): Randomized Response techniques, Marcel Decker.
Cochran, W. G. (1977): Sampling techniques, Third Edition, Wiley.

Unit 4: Methods of Estimation: Method of moments, method of minimum chi-square, method of maximum likelihood and it's properties, Method of scoring and it's applications. Asymptotic effiency of MLE, CAN and BAN estimators.

15 Hrs

Books for Reference:

Kale B.K (1999) A first course on parametric inference. Narosa.
Lehmann E. L (1988) Theory of point estimation. John wiley & Sons
Rohatgi V.K (1984) An introduction to probability theory mathematical Statistics.Wiley eastern, New Delhi.
Zacks, S (1971) Theory of Statistical Inference. Wiley, Newyork.

Practicals based on STCT: 2.2

STCT: 2.3: LINEAR MODELS AND REGRESSION ANALYSIS

Unit 1: Gauss-Markov setup, normal equations and least squares estimates, error and estimation space, variance and covariance of least squares estimates, estimation of error variance, estimation with correlated observations, least squares estimates with restriction and parameters. Simultaneous estimates of linear parame1ric functions.

10 Hrs

Unit 2: Test of hypothesis for one and more than one parametric functions. Confidence intervals and regions, analysis of variance fable, power of F-test, multiple comparison test like Tukey and Scheffe, simultaneous confidence interval.

10 Hrs

Unit 3: One way linear models when parameters are random and estimation of variance components. Simple linear regression, multiple regression - estimation, testing linear hypotheses, confidence interval, confidence region, prediction of new observations, prediction interval, fit of polynomials and use of orthogonal polynomial Introduction to non-linear models.

10 Hrs

Unit 4: Model adequacy - residuals and their plot for examining the departure from assumptions such as fitness of the model, normality, homogeneity of variances and detection of outliers and remedies. Hat-matrix, leverages and detection of influential observations.

10 Hrs

Unit 5: Multicollinearity, ridge and principal component regression. Validation of regression model: analysis of estimated coefficients and predicted values, collecting fresh data, data splitting. Subset selection of explanatory variables and Mallow's C_p statistic, all possible regressions, stepwise, forward and backward regressions.

10 Hrs

Books for Reference:

Cook, RD. and Weisberg, S. (1982); Residual 'and Influence in Regression. Chapman and Hall, London.

Draper, N.R and Smith, H (1998): Applied Regression Analysis. Third Edition, Wiley, New York.

Gunst, R.F. and Mason, R.L. (1980) ; Regression Analysis and its Application - A Data Oriented Approach, Marcel-Dekker.

Montgomery,D.C, Peck, E.A. and Vining,G.G.(2003). Introduction to Linear Regression, John Wiley.

Rao, C. R (1913): Linear statistical Inference, Wiley Eastern.

Ryan,T.P.(1997). *Modern Regression Methods*, John Wiley, NY

Searle,S.R.(1971). *Linear Models*, John Wiley, NY.

Seber, G.A.F.(1997): *Linear Regression Analysis.* John Wiley, NY.

Seber, G.A.F. and Lee (2003): *Linear Regression Analysis.* 2/e John Wiley, NY

Weisberg, S. (1985): Applied Linear Regression, Wiley.

Practicals based on STCT: 2.3

STCT: 2.4: ELEMENTARY STOCHASTIC PROCESS

Unit 1: Introduction to stochastic processes (SP), classification of SP according to state space and time domain. Finite and countable state Markov chains (MC), Chapman – Kolmogorov's equations, calculation of n-step transition probabilities and their limits, stationary distribution, classification of states, transient MC, random walk and gambler's ruin problems.

16 Hrs

Unit 2: Continuous time Markov processes: Kolomgorov-Feller differential equation, Poisson process, pure birth process, Yule – Furry process, birth and death processes,

Gibbons J.D. (1985). Non Parametric Statistical inference. Marcel Dekkar, New York.
Randles R.H. and Wolfe D.A. (1979) Introduction to Theory of Non-Parametric Statistics. Wiley, New York.
Cramer H. (1957) Mathematical Methods of Statistics. Princeton University Press, New Jersey.

Practicals based on STCT: 3.2

STCT: 3.3: BIOSTATISTICS

Unit 1: Definition, agent, host and environment, mode of transmission; incubation period, spectrum of disease, herd immunity, classification of cause, of death, measures of mortality, studies of mortality. Measures of morbidity, I1lOIbidity surveys, issues and problems. Risk, cause and bias. Observational studies: retrospective, cross sectional and prospective studies.

10 Hrs

Unit 2: Clinical trials: Methods of randomization, ethical issues, cross over trials. Sequential and group sequential trials. Interim analysis, multiple testing and stopping rules. Equivalence trials.

10 Hrs

Unit 3: Clinical Epidemiology: Definition, reliability, validity, sensitivity, specificity, predictive values, likelihood ration test, selection and interpretation of diagnostic test. Deciding on the .best therapy.

10 Hrs

Unit 4: ROC curves, multiple and parallel test. Screening for disease, critical appraisal, Meta analysis. Epidemiologic Models - Epidemometric studies- Deterministic epidemic models: Simple, General, Recurrent.

10 Hrs

Unit 5: Time Series (Epidemic or others) Applications of Time series analysis in epidemiology - Simple descriptive techniques for detecting seasonal, Cyclical, secular and random variations

10 Hrs

Unit 4: System reliabilities: Series, parallel, k-out-of-n, standby redundant systems and their reliabilities.
Maintenance policies: Age replacement policy and Block replacement policies and their characteristics. Reliability modeling: Introduction to shock models, stress-strength models and proportional hazard models. 6 Hrs

Unit 5: Inference in Reliability: Type I and Type II Censoring schemes, likelihood functions based on these sampling schemes for exponential distribution. Reliability estimation (complete and censored samples) for exponential distribution, testing reliability hypotheses (exponential distribution). 12 Hrs

Books for Reference:

Montgomery D.C. (1996) Introduction to Statistical Quality Control, Wiley, New York.
Grant E.L. (1980) Statistical Quality Control McGraw Hill, New York.
Weetherhill G.B. and Brow D.W. (1991) Statistical Process Control. Chapman and Hall, London.
Barlow R.E. and Proschan F (1975) Statistical Theory of Reliability and Life Testing. Holt-Rinhart and Winston, New York.
Sinha S.K. and Kale B.K. (1990) Life Testing and Reliability Estimation. Wiley Eastern, New Delhi.
Mann N.R, Schaffer R.F and Singpurwalla N.D. (1974) Methods for Statistical Analysis of Reliability and Life Data. Wiley New York.
Zacks S (1992) Introduction to Reliability Analysis. Springer - Verlag, New York.
J.V. Deshpande and Sudha G. Purohit (2005) Life time data: Statistical Models and Methods. World Scientific.

Practicals based on STCT: 4.2

Elective Papers: ***Any two***

STET: 4.3: DEMOGRAPHY (*OPTIONAL*)

Unit 1: Demography and its interdisciplinary nature, sources of demographic data, Coverage and Content errors. The use of balancing equation, Chandrasekaran and Deming formula to check completeness of registration data. Use of Whipple's, Myers's and UN Indices.

10 Hrs

Unit 2: Measures of Mortality: Various measures of mortality, infant mortality rate, cause specific death rates and standardized death rates. Measures of Fertility: Period and cohort fertility measures, use of birth order statistics, child – women ratio, Brass P/F ratio to estimate current levels of fertility, Measures of reproduction and replacement. Sheps and Perrin stochastic human reproductive process.

15 Hrs

Unit 3: Life Tables: Types of life tables, inter – relationships between life table functions, construction of life tables using Reed – merrel and Greville's Method. Probability distribution of life table functions and their optimum properties. Population estimation and Projections: Mathematical, Statistical and Demographic Methods, Component method.

15 Hrs

Unit 4: Stable and Quasi – stable population: Derivation of Lotka's stable population model and its properties, Intrinsic growth rate and its derivation, age structure and birth rate of a stable population, mean length of generation, momentum of population growth, Quasi – stable population under changing fertility and mortality situations.

10 Hrs

Books for Reference:

Shryock, Henry S, Jacob S, Siegel and Associates (1964)- Methods and materials of demography (condensed edition) Academic press, London.

Barclay, George W. (1968) Techniques of population analysis, John Wiley and sons, New York.

Keyfitz N. (1968). Introduction to the Mathematics of Population. Addision-Wesley Publishing Co, Reading, Messachusetts.

Chiang C.L. (1968) Introduction to stochastic processes in Biostatistics, John Wiley and sons, New York.

R. Ramkumar (1986) Technical Demography, Wiley Eastern, New Delhi.

Sudhendu Biswas (1988), Stochastic Processes in Demography and Applications, Wiley Eastern, New Delhi.

Practicals based on STET: 4.3

Practicals based on STET: 4.4

STET: 4.5: SURVIVAL ANALYSIS (*OPTIONAL*)

Unit 1: Concepts of Time, Order and Random Censoring. Life distributions - Exponential Gamma, Weibull, Lognormal, Pareto, Linear Failure rate. Parametric inference Point estimation, Confidence Intervals, Scores, tests based on LR , MLE

10 Hrs

Unit 2: Life tables, Failure rate, mean residual life and their elementary properties. Ageing classes - IFR, IFRA, NBU, NBUE, HNBUE and their duals, Bathtub Failure rate.

10 Hrs

Unit 3: Estimation of survival function - Actuarial Estimator, Kaplan - Meier Estimator, Estimation under the assumption of IFR/DFR.

10 Hrs

Unit 4: Tests of exponentiality against non-parametric classes - Total time on test, Deshpande test. Two sample problem - Gehan Test, Log rank test. Mantel - Haenszel Test, Tarone - Ware tests.

10 Hrs

Unit 5: Semi-parametric regression for failure rate - Cox's proportional hazards model with one and several covariates.

10 Hrs

Books for Reference:

Cox, D.R. and Oakes, D. (1984). Analysis of Survival Data, Chapman and Hall, New York.

Gross, A. J. and Clark, V. A. (1975). Survival Distributions: Reliability Applications in the Biomedical Sciences, John Wiley and Sons.

Elandt - Johnson, R.E., Johnson, N.L. (1980). Survival models and Data Analysis, John Wiley and Sons.

Miller, R.G. (1981). Survival Analysis, Wiley.

Zacks, S. Reliability.

Practicals based on STET: 4.5

STET: 4.8: STATISTICAL GENETICS (*OPTIONAL*)

Unit 1: Basic biological concepts in genetics (relevant to this course)

05 Hrs

Unit 2: Mendel's law, Hardy Weinberg equilibrium. Mating tables, estimation of allele frequency
(dominant / co-dominant cases). Approach to equilibrium for X-linked gone, Natural selection,
mutation, genetic drift, equilibrium when both natural selection and mutation are operative.

20 Hrs

Unit 3: Non-random mating, inbreeding, phenotypic assortative mating.

10 Hrs

Unit 4: Analysis of family data (a) Relative pair data, I, T; 0 matrices, identity by descent, (b) family data - estimation of segretation ratio under ascertainment bias, (c) Pedigree data- Elston - Stewart algorithm for calculation of likelihoods. Linkage, Estimation of recombination fraction, inheritance of quantitative traits. Models and estimation of parameters.

10 Hrs

Unit 5: Sequence similarity, homology and alignment. Algorithms for (a) pairwise sequence alignment, (b) multiple sequence alignment, construction of phylogenetic trees, UPGMA, Neighbour joining, maximum parsimony and maximum likelihoodalgorithms.

05 Hrs

Books for Reference:

C. C. Li (1976). First course on population genetics. Boxwood Press, California.
W. J. Ewens (1979). Mathematical population genetics. Springer Verlag.
T. Nagylaki (992). Introduction to theoretical population genetics. Springer Verlag.
R. Durbin, S. R. Eddy, A. Krogh, G. Mitchison (998). Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids.

Practicals based on STET: 4.8

M.A / M.SC PROGRAMME IN STATISTICS UNDER CBCS SYSTEM FROM THE ACADEMIC YEAR 2018 – 2019

Xerox

SYLLABUS

SEMESTER I

Emp.

HCT1.1: LINEAR ALGEBRA (4 Credits – 4 hours of Theory teaching per week)

Unit 1: Fields, vector spaces, subspaces; linear dependence and independence; basis and dimension of a vector space, finite dimensional vector spaces completion theorem. Examples of vector spaces over real and complex fields. Linear equations. Vector spaces with an inner product, Gram-Schmidt orthogonalization process. Orthonormal basis and orthogonal projection of a vector.

Unit 2: Linear transformations, algebra of matrices, row and column spaces of a matrix. Elementary matrices, determinants, rank and inverse of a matrix. null space and nullity; partitioned matrices; Kronecker product. Hermite canonical form, generalized inverse, Moore- Penrose Inverse, Idempotent matrices. Solutions of matrix equations.

Unit 3: Triangular reduction of a positive definite matrix. Characteristic roots and vectors, Cayley- Hamilton thoorem, minimal polynomial, similar matrices. Algebraic and geometric multiplicity of characteristic roots, spectral decomposition of a real symmetric matrix, reduction of a pair of real symmetric matrices, Hermitian matrices.

Unit 4: Real quadratic forms, reduction and classification of quadratic forms, index and signature. Singular value and singular decomposition, Jordan decomposition, extrema of quadratic forms. Vector and matrix differentiation.

Book[illegible] Reference:

Bell[illegible] (1970): Introduction to Matrix Analysis, Second Edition, McGraw Hill.
Bis[illegible] 984): Topics in Algebra of Matrices, Academic Publications.
Gra[illegible] (1983): Matrices with Applications in Statistics, Second Edition, Wadsworth..
Ha[illegible] 87): Linear algebra, Narosa.
He[illegible] (1958): Finite Dimensional Vector Spaces, Second Edition, D. Van Nostrand Company.
Ho[illegible] nd Kunze, R. (1971): Linear Algebra, Second Edition, Prentice Hall.
Ra[illegible] Bhimasankaram, P.(1992): Linear Algebra, Tata McGraw Hill.
Rg[illegible]): Linear Statistical Inference and its Applications, Second Edition, Wiley.
R[illegible] Mitra, S. K (1971): Genera1ized Inverse of Matrices and its Applications, Wiley.

Xerox

Unit 5: Markov, Chebyshev, Hoelder, Minkowski, Jensen and Liapunov inequalities. Relationship between tail of distributions and moments.

10 Hrs

Books for Reference:

Cramer, H. (1946): Mathematical Methods of Statistics, Princeton.
Jolmson, S. and Kotz. (1972); Distributions in Statistics, Vols.I, II and III, Houghton and Miffin.
Mukhopadhyaya, P. (1996): Mathematical Statistics, Calcutta Publishing House.
Pitman, J. (1993): Probability, Narosa.
Lukacs C (1970)' Characteristic functions, Griffin Publications.

EMP

2.2

SCT: 1.1: SAMPLE SURVEYS AND STATISTICS FOR NATIONAL DEVELOPMENT, WOMEN'S STATISTICS

A. SAMPLE SURVEYS

Unit 1: Basic finite population sampling techniques (SRS WR/ WoR, stratified, systematic), related problems of population mean estimation, allocation problems in stratified sampling.

10 Hrs

Unit 2: Unequal probability sampling: PPS WR / WoR methods (including Lahiri's scheme) and related estimators of a finite population mean (Hansen-Hurvitz and Desraj estimators for a general sample size and Murthy's estimator for a sample of size 2).

10 Hrs

Unit 3: Ratio and regression estimators based on SRS WoR method of sampling, two-stage sampling with equal. number of second stage units, double sampling, cluster sampling.

10 Hrs

B. STATISTICS'FOR NATIONAL DEVELOPMENT

Unit 4: Economic development: growth ip. per capita income distributive justice. Indices of development, Human Development Index. Estimation of National Income - product approach, income approach and expenditure approach. Population growth in developing and developed countries. Population projection using Leslie matrix. Labour force projection.

10 Hrs

Unit 5: Measuring inequality of incomes, Gini coefficient, Theil's measure. Poverty measurement- different issues, measures of incidence and intensity, combined measures, eg. Indices due to Kakwani, Sen. etc.

10 Hrs

C. WOMEN'S STATISTICS

Women's Empowerment Index: 1) Determinants of empowerment like education, employment, access to health, access to media, access to finance, involvement in decision making process at family and at office/organization . 2) Various measures of women's empowerment index based on sex, geographical region and time etc.

Books for Reference:

Choudhary, A and Mukherjee, R (1989): Randomized Response techniques, Marcel Decker.
Cochran, W. G. (1977): Sampling techniques, Third Edition, Wiley.
Des Raj and Chandok (1998): Sampling Theory, Narosa.
Murthy, M. N. (1977): Sampling Theory and Methods, Statistical Publishing Society, Calcutta.
Singh,D. and Choudhary, F. S. (1986): Theory and Analysis of Sample Survey Designs, New Age International.
Sukhatme et al. (1984): Sampling Theory of Surveys with Applications, Iowa State University Press.

C.S.O. (1980): National Accounts Statistics- Sources and Health.
Keyfitz, N. Mathematical Demography.
UNESCO: Principles of Vital Statistics Systems, Series M-12.
Sen, A(1997): Poverty and Inequality.

HCP 1.1: STATISTICAL COMPUTING (PRACTICAL PAPER - I)

Unit 1: A general overview of R . Numeric/Character/logical; real /integer/complex string and the paste command matrices, dataframes, lists, setwd,read.table,read.csv.write. matrix, write.csv, creation of new variables, catergorization cut, factor,round, apply, creation of pattern variables, saving output files; source;print saving work space/ history. R-help command; help. search(), r mailing list, contributed documentation on cran. Descriptive statistics and Graphics in R: summary statistics for single group, the plot-command, histogram, box plot bar plot, lines, points, segments, arrows, paste, inserting math. symbols in a plot, pie diagram, customization of plot- setting graphical parameters text and m-text the pairs command. Graphical parameters such as mar/mai/mfrow/xlab/ylab/las/xaxp/xlim/ylim/cex/axis/tck/srt/main/title/legend/locator/ identity.

Unit 2: One and two sample t-tests, chi-squared tests, F- test for equality of variance, nonparametric tests, regression analysis, checking the assumptions of normality, Q-Q plots, P-P plots. ANOVA. Matrix operations, addition, subtraction, multiplication, linear equation and eigenvalues, finding rank, inverse, g-inverse, determinant. R –functions; some useful built in r functions, attach, detach, sort, order, rank, ceiling, floor, round, trunc, signif, apply, lapply, by. Programming in R;fro/while/loops, functions, the source command.

Unit 3: Numerical analysis and statistical applications. Numerical integration, root extraction, random number generation, Monte Carlo integration, matrix computations, drawing random samples from known univariate probability distributions -both discrete and continuous and bivariate normal distribution - the inverse method, the accept- rejection method, decomposition of discrete mixtures, Classical Monte Carlo integration. R-functions for generating random variables and simulations; rnorm, rbinom, rpiosson, runif,rchisq,rt, etc.,; sample, set.seed.

Unit 4: Writing / performing programs using MATLAB /MINITAB/ SPSS/ Excel on problems from the following topics: Descriptive statistics and Graphs, One and two sample parametric and nonparametric tests, Chi-square tests, Regression and correlation analysis, Analysis of Variance and Kruskal-Wallis Test.

References:

1. Dalgaard, P. (2002). Introductory Statistics with R. Springer Verlag, New York.
2. Kerns, G.J. (2010). Introduction to Probability and Statistics Using R. Free Software Foundation.
3. Kunte, Sudhakar (1999). Statistical Computing: 1. Understanding Randomness and Random Numbers,Resonance, Vol.4, No.10, pp.16-21.
4. Kunte. Sudhakar (2000). Statistical Computing: 2. Technique of Statistical *Simulation, Resonance,* Vo1.5, No.4, pp.18-27.
5. Robert, C. and Casella, G. (2010). Introducing Monte Carlo Methods with R. Springer Verlag, New York.
6. Sudha G Purohit, Sharad D Gore, Shailaja R Deshmukh (2010) Statistics Using R, Alpha Science Intl. Publ.
7. Verzani, J. (2005). Using R for Introductory Statistics. Taylor & Francis

HCP 1.2: (PRACTICAL PAPER – II) Practical's based on HCT: 1.1 AND SCT 1.1

Unit 4: Methods of Estimation: Method of moments, method of minimum chi-square, method of maximum likelihood and it's properties, Method of scoring and it's applications. Asymptotic effiency of MLE, CAN and BAN estimators.

15 Hrs

Books for Reference:

Kale B.K (1999) A first course on parametric inference. Narosa.
Lehmann E. L (1988) Theory of point estimation. John wiley & Sons
Rohatgi V.K (1984) An introduction to probability theory mathematical Statistics. Wiley eastern, New Delhi.
Zacks, S (1971) Theory of Statistical Inference. Wiley, Newyork.

HCT: 2.3: LINEAR MODELS AND REGRESSION ANALYSIS

Unit 1: Gauss-Markov setup, normal equations and least squares estimates, error and estimation space, variance and covariance of least squares estimates, estimation of error variance, estimation with correlated observations, least squares estimates with restriction and parameters. Simultaneous estimates of linear parame1ric functions.

10 Hrs

Unit 2: Test of hypothesis for one and more than one parametric functions. Confidence intervals and regions, analysis of variance fable, power of F-test, multiple comparison test like Tukey and Scheffe, simultaneous confidence interval.

10 Hrs

Unit 3: One way linear models when parameters are random and estimation of variance components. Simple linear regression, multiple regression - estimation, testing linear hypotheses, confidence interval, confidence region, prediction of new observations, prediction interval, fit of polynomials and use of orthogonal polynomial Introduction to non-linear models.

10 Hrs

Unit 4: Model adequacy - residuals and their plot for examining the departure from assumptions such as fitness of the model, normality, homogeneity of variances and detection of outliers and remedies. Hat-matrix, leverages and detection of influential observations.

10 Hrs

Unit 5: Multicollinearity, ridge and principal component regression. Validation of regression model: analysis of estimated coefficients and predicted values, collecting fresh data, data splitting. Subset selection of explanatory variables and Mallow's C_p statistic, all possible regressions, stepwise, forward and backward regressions.

10 Hrs

Books for Reference:

Cook, RD. and Weisberg, S. (1982); Residual 'and Influence in Regression. Chapman and Hall, London.
Draper, N.R and Smith, H (1998): Applied Regression Analysis. Third Edition, Wiley, New York.
Gunst, R.F. and Mason, R.L. (1980) ; Regression Analysis and its Application - A Data Oriented Approach, Marcel-Dekker.
Montgomery,D.C, Peck, E.A. and Vining,G.G.(2003). Introduction to Linear Regression, John Wiley.
Rao, C. R (1913): Linear statistical Inference, Wiley Eastern.
Ryan,T.P.(1997). *Modern Regression Methods*, John Wiley, NY
Searle,S.R.(1971). *Linear Models*, John Wiley, NY.
Seber, G.A.F.(1997): *Linear Regression Analysis*. John Wiley, NY.
Seber, G.A.F. and Lee (2003): *Linear Regression Analysis*. 2/e John Wiley, NY
Weisberg, S. (1985): Applied Linear Regression, Wiley.

two binomial and Poisson populations.

10 Hrs

Unit 3: Confidence intervals and their connection with the tests of hypotheses. UMA, UMAU confidence intervals, shortest length confidence intervals.

Likelihood ratio tests, large sample properties. Chi-square goodness-of-fit tests for simple and composite hypothesis.

10 Hrs

Unit 4: Nonparametric methods-run test, sign test, signed-rank test, median test, Wilcoxon-Mann-Whitney test, Kolmogorov – Smirnov tests, Tests involving rank correlation, Linear rank statistics, Large sample properties and applications.

10 Hrs

Unit 5: Sequential analysis, need for sequential tests, SPRT and its properties, termination property, fundamental identity and Wald's equation, OC and ASN functions. SPRT for testing hypothesis in binomial, Poisson, normal and exponential distribution-computation of OC and ASN functions.

10 Hrs

Books for Reference:

Lehmann E.L. (1986)Testing Statistical Hypothesis, Wiley, New York.

Rohatgi V.K. (1984). An Introduction to Probability Theory and Mathematical Statistics. Wiley Eastern, New Delhi.

Dudewicz E.J. and Mishra S.N. (1988) Modern Mathematical Statistics, Wiley and Sons, New York.

Ferguson T.S. (1967), Mathematical Statistics- Decision Theoretic Approach. Academic Press, New York.

Kendall M.G. and Stuart A (1968) Advanced Theory of Statistics, Vol II, Charles Griffin and Co., London.

Rao C.R (1973). Linear Statistical inference. Wiley Eastern, New Delhi.

Wald A (1947) Sequential Analysis, Wiley New York.

Gibbons J.D. (1985). Non Parametric Statistical inference. Marcel Dekkar, New York.

Randles R.H. and Wolfe D.A. (1979) Introduction to Theory of Non-Parametric Statistics. Wiley, New York.

Cramer H. (1957) Mathematical Methods of Statistics. Princeton University Press, New Jersey.

HCT: 3.3: DESIGN AND ANALYSIS OF EXPERIMENTS

Unit 1: Gauss-Markov setup, estimability of linear parametric functions, normal equations and least squares estimation. Error and estimation spaces, variance and covariance of least square estimates. Estimation of Error variance, Linear Estimation in the correlated setup. Least squares Estimates with restriction on the parameters, simultaneous estimates of linear parametric functions.

14 Hrs

Unit 2: Tests of hypotheses for one and more than one linear parametric functions. confidence intervals and regions, ANOVA table, power of F – Test, multiple compararion procedures of Tukey and Scheffe, simultaneous confidence intervals.

8 Hrs

Unit 3: Application of Gauss – Markov theory to the analysis of two – way and three – way classification models. CRD, RCBD, LSD and Missing Plot techniques.

8 Hrs

Unit 4: Incomplete bock designs – BIBD and PBIBD. Balance, connectedness and orthogonality in relation to two – way designs. Analysis of Covariance.

8 Hrs

Unit 5: Factorial experiments, confounding in 2^n factorial experiments, Split – Plot design, Random effects models: one – way and two – way classification.

12 Hrs

Books for Reference:

Chakravarthy M.C. (1971) Mathematics of Design and Analysis of Experiments. Asia Publishing House.
Joshi, D.D. (1987) Linear Estimation and Design of experiments. Wiley Eastern.
Kshirsagar, A.M. (1983) Linear Models, Marcel Dekker.
Das M.N. and Giri, N.C. (1988) Design and Analysis of experiments. Wiley Eastern. Ltd.
Montgomery D.C. (1991) Design and Analysis of experiments, John Wiley and sons.
Ogawa, J (1974) Statistical Theory of the analysis of the experimental design. Marcel Dekker.
Rao C.R. (1985) Linear Statistical Inference and its applications. Wiley Eastern.
Searle S.R. (1971) Linear Models. John Wiley & Sons.

SCT 3.1: BIO-STATISTICS

Unit 1: Definition, agent, host and environment, mode of transmission; incubation period, spectrum of disease, herd immunity, classification of cause, of death, measures of mortality, studies of mortality. Measures of morbidity, I1lOIbidity surveys, issues and problems. Risk, cause and bias. Observational studies: retrospective, cross sectional and prospective studies.

10 Hrs

Unit 2: Clinical trials: Methods of randomization, ethical issues, cross over trials. Sequential and group sequential trials. Interim analysis, multiple testing and stopping rules. Equivalence trials.

10 Hrs

Unit 3: Clinical Epidemiology: Definition, reliability, validity, sensitivity, specificity, predictive values, likelihood ration test, selection and interpretation of diagnostic test. Deciding on the .best therapy.

10 Hrs

Unit 4: ROC curves, multiple and parallel test. Screening for disease, critical appraisal, Meta analysis. Epidemiologic Models - Epidemometric studies- Deterministic epidemic models: Simple, General, Recurrent.

10 Hrs

Unit 5: Time Series (Epidemic or others) Applications of Time series analysis in epidemiology - Simple descriptive techniques for detecting seasonal, Cyclical, secular and random variations

10 Hrs

Books for Reference:

Lilienfeld, A.M. and Lilienfeld, D.C. Foundations of epidemiology, Second Edition, Oxford Univ. Press, New York, 1980.
Fletcher, R.H., Fletcher, S. W. and Wagner, E.H. Clinical Epidemiology - the essentials, Second Edition, 1982.
Harold A Hahn, Christopher T. Sempos. Statistical Methods in Epidemiology, Oxford Univ. Press, New York, 1989.
David G.Kleinbaum, Lawrence L.Kupper and Hall Morgenstem. Epidemiologic Research, Van Nostrand, USA, 1982.
Chatfield, C. The Analysis of Time Series - An Introduction. Third Edition, Chapman and Hall, London, 1984.
Bailey N.T.J. The Mathematical Approach to Biology and Medicine. - Chapters 1, 2, and 9, John Wiley, 1967.

HCP 3.1: (PRACTICAL PAPER – I) Practicals based on HCT: 3.1 AND HCT 3.2

HCP 3.2: (PRACTICAL PAPER – II) Practicals based on HCT: 3.3 AND SCT 3.1

OEP 3.1 QUANTITAIVCE TECHNIQUES FOR REASERACH

Unit 1: Descriptive Statistics: Data presentation by charts and tables, measures of central tendency, Measures of dispersion, Skewness and kurtosis.

Unit 2: Correlation and regression Analysis , Scatter plot, Karl-pearson's coefficient of correlation, Spearman's rank correlation, Simple regression Analysis.

Unit 3: Test of Significance

Basic s of testing of hypothesis, Test for proportions (one sample and two sample problems), Test for mean, test for variance (one sample and two sample problems), Chi-Square test of independence, ANOVA.

Unit 4: Operations Research

Linear programming problems(LPP), Solution by graphical method, Transportation , Assignment and sequencing (Feasible solutions only).

Books for Reference:

Medhi J (1994), Stochastic Processes, 2nd edn., Wiley Eastern Ltd., New Delhi.

Bhattacharya, G.K. and Johnson, R.A. Statistical concepts and methods. Wiley Eastern. Calcutta, Bombay and Delhi.

Goon A.M., Gupta, M.K. and Das Gupta B: Fundamental of Statistics Vol.1 and II. World Press Pvt. ltd., Calcutta.

Levin, R.I. : Statistics for Management, Prentice Hall of India, New Delhi.

Unit 1: Introduction to stochastic processes (SP), classification of SP according to state space and time domain. Finite and countable state Markov chains (MC), Chapman – Kolmogorov's equations, calculation of n-step transition probabilities and their limits, stationary distribution, classification of states, transient MC, random walk and gambler's ruin problems.

16 Hrs

Unit 2: Continuous time Markov processes: Kolomgorov-Feller differential equation, Poisson process, pure birth process, Yule – Furry process, birth and death processes, Weiner process as a limit of random walks, first passage time and other problems, diffusion process.

14 Hrs

Unit 3: Renewal Theory: Elementary renewal theorem and applications, key renewal theorem and its uses, study of residual life time process, discrete time renewal theory. Stationary process: weakly stationary and strongly stationary processes, spectral decomposition, moving average and auto regressive processes.

14 Hrs

Unit 4: Branching process: Galton-Watson branching process, probability of ultimate extinction, distribution of population size, and statistical inference in MC and Markov process.

6 Hrs

Books for Reference:

Medhi J (1994), Stochastic Processes, 2nd edn., Wiley Eastern Ltd., New Delhi.
Bhat U.N. (1984), Elements of Applied Stochastic processes, 2nd edn., Wiley, New York.
Basawa I.V. and Prakash Rao B.L.S. (1980) Statistical Inference for stochastic processes, Academic press, New York.
Karlin S and Taylor H.M. (1975), A first course in stochastic processes, 2nd edn., Academic press, New York.

Elective Papers:

Any two

SCT: 4.1(a): DEMOGRAPHY (*OPTIONAL*) Emp

Unit 1: Demography and its interdisciplinary nature, sources of demographic data, Coverage and Content errors. The use of balancing equation, Chandrasekaran and Deming formula to check completeness of registration data. Use of Whipple's, Myers's and UN Indices.

10 Hrs

Unit 2: Measures of Mortality: Various measures of mortality, infant mortality rate, cause specific death rates and standardized death rates. Measures of Fertility: Period and cohort fertility measures,

use of birth order statistics, child – women ratio, Brass P/F ratio to estimate current levels of fertility, Measures of reproduction and replacement. Sheps and Perrin stochastic human reproductive process.

15 Hrs

Unit 3: Life Tables: Types of life tables, inter – relationships between life table functions, construction of life tables using Reed – merrel and Greville's Method. Probability distribution of life table functions and their optimum properties. Population estimation and Projections: Mathematical, Statistical and Demographic Methods, Component method.

15 Hrs

Unit 4: Stable and Quasi – stable population: Derivation of Lotka's stable population model and its properties, Intrinsic growth rate and its derivation, age structure and birth rate of a stable population, mean length of generation, momentum of population growth, Quasi – stable population under changing fertility and mortality situations.

10 Hrs

Books for Reference:

Shryock, Henry S, Jacob S, Siegel and Associates (1964)- Methods and materials of demography (condensed edition) Academic press, London.

Barclay, George W. (1968) Techniques of population analysis, John Wiley and sons, New York.

Keyfitz N. (1968). Introduction to the Mathematics of Population. Addision- Wesley Publishing Co, Reading, Messachusetts.

Chiang C.L. (1968) Introduction to stochastic processes in Biostatistics, John Wiley and sons, New York.

R. Ramkumar (1986) Technical Demography, Wiley Eastern, New Delhi.

Sudhendu Biswas (1988), Stochastic Processes in Demography and Applications, Wiley Eastern, New Delhi.

SCT: 4.1(b): TIME SERIES ANALYSIS *(OPTIONAL)*

Unit 1: Time series as discrete parameter stochastic process, auto-covariance and auto-correlation functions and ,their properties.

Unit 2: Detailed study of the stationary processes: (i) moving average (MA), (ii) auto-regressive (AR), (iii) ARMA, and, (iv) AR integrated MA (ARIMA) models. Box-Jenkins models. Discussion (without proof) of estimation of mean, auto-covariance and auto-correlation functions under large sample theory.

Unit 3: Choice of AR and MA orders. Estimation of ARIMA model parameters. Forecasting. Residual analysis and diagnostic checking.

Unit 4: Spectral analysis of weakly stationary process, periodogram and correlogram analysis, computation based on Fourier transforms, Spectral decomposition of weakly AR process and representation as a one-sided MA process -necessary and sufficient conditions.

Unit 5: Implication of spectral decomposition in prediction problems. State space representation of time series. Kalman filter techniques.

Books for Reference:

Anderson. T.W. (1971). The Statistical Analysis of Time Series. Wiley.
Bloomfield, P. (2000). Fourier Analysis of Time Series: An Introduction. Second Edition, Wiley.
Box, G.E.P., Jenkins, G. W. and Reinsel, G.C. (1994). Time Series Analysis:Forecasting and Control Prentice Hall.
Box, G.E.P. and Jenkins, G.M (1976). Time Series Analysis - Forecasting and Control Holden-day, San Francisco.
Chatfield, C. Analysis of Time Series - Theory and Practice, Chapman and Hall.
Chow, C.G. (1985). Econometrics.. Mc Graw Hill.
Findley, D.F..ed., (1981). Applied Time Series Analysis II. Academic. Press.
Fuller, W.A. (1976). Introduction to Statistical Time series. Wiley.
Granger, C W.J. and Newbold (1984). Forecasting Econometric Time Series, Third Edition, Academic Press.
Granger, C.W.J. and Hatanka, M. (1964). Spectral Analysis of Economic Time Series, Princeton University Press.
Hannan. E.J. (1960). Time Series Analysis, Metheun, London.
Kendall, MG. (1974). Time Series, CnMles Griffin, London.
Kendall, MG> and Sroan, A. (1966). The Advanced Theory of Statistics, Vol. 3, Charles Griffin, London.
Koopmans, L.H. (1974). The Spectral Analysis of Time Series, Academic Press.
Montgomery, D.C. and Johnson, L.A (1977) Forecasting and Time Series Analysis, McGraw Hill.
Nelson, C.R (1973). Applied Time Series for managerial forecasting. Holden-day.
Priestly, MB. (1981). Spectral Analysis and Time Series. Griffin, London.

SCT: 4.1 (c): ACTUARIAL STATISTICS (*OPTIONAL*)

Unit 1: Basic deterministic model: Cash flows, discount function, Interest and discount rates, balances and reserves, internal rate of return, The life table: Basic definitions, probabilities, construction of life tables, life expectancy, Life annuities: Introduction, calculating annuity premiums, interest and survivorship discount function, guaranteed payments, deferred annuities

15 Hrs

Unit 2: Life insurance: Introduction, calculation of life insurance premiums, types of life insurance, combined benefits, insurances viewed as annuities, Insurance and annuity reserves: The general pattern of reserves, recursion, detailed analysis of an insurance, bases for reserves, non forfeiture values, policies involving a return of the reserve, premium difference and paid-up formula.

10 Hrs

Unit 3: Fractional durations: Life annuities paid monthly, immediate annuities, fractional period premium and reserves, reserves at fractional durations, Continuous payments: Continuous annuities, force of discount, force of mortality, Insurance payable at the moment of death, premiums and reserves. The general insurance - annuity identity, Select mortality: Select an ultimate tables, Changes in formulas.