Curriculum for Master's in Statistics Programme (2016-18) Program Code : 74

Department of Statistics
Indira Gandhi National Tribal University
Amarkantak, India

Contents

1 The Curriculum

1.1 The Program

The four semester programme leading to degree of Masters in Statistics, shall, henceforth be termed as the *Program*.

1.2 Courses

Courses refer to various papers that are offered in the program.

1.3 Credit

A Credit is equivalent to an hour of teaching (Lecture (L) / Tutorial (T) per week or two hours of Practicals (P) / Field Work (F) per week. Each course in the program carries credits. The credits attributed to a course indicate the relative weight / importance of the respective course among the set of courses offered in the program.

1.4 Choice Based Credit System (CBCS)

Under the CBCS the learner may opt for courses and earn the required number of *credits* so as to qualify for the award of the degree.

The constituent courses of the program consist of the following types:

1. **Core Courses** (CC): Core Courses are a compulsory requirement for completion of the program.
2. **Elective Courses**: Elective Courses are mandatory courses that have to be chosen by the learner so as to enhance her / his generic proficiency while undergoing the program.

    (a) **Discipline Specific Electives** (DSE):These are courses from Statistics that help in enhancing generic proficiency in specialized areas.

(b) **Open Electives** (OE): These are courses from disciplines other than Statistics and are offered by other departments. Henceforth, OE courses shall be termed as **Generic Elective** (GE) **Courses**.

The learner has to choose the (DSE) courses in consultation with their mentors for semesters - III and IV. The learner is also advised to explore all GE courses available in various departments in Semesters - I and II. Based upon her / his choices she / he may enrol for the chosen GE course(s) for the respective semester.

The program shall be a four semester program with the following course and credit components

1. CC : 70 Credits
2. DSE Courses : 8 Credits
3. GE Courses : 6 Credits

In order to successfully complete the program the student shall earn at least 84 Credits from among CC, DSE and GE courses as mentioned above.

1.5 Admission to the Program

Eligibility Criteria A Bachelors degree from a recognized Indian or foreign university (as per the AIU foreign equivalence list) and secured a minimum of 50% aggregate of marks (45% for SC/ST/PWD/ Transgender and Kashmiri Migrant Students) and having studied Mathematics or Statistics as a major/optional subject at undergraduate level are eligible to apply. To pass the entrance examination conducted by the University the candidates belonging to open category have to secure a minimum of 40% marks [35% for SC/ST and Kashmiri Migrant students/ Transgender].

Number of Seats The number of students to be admitted to the Program are 20.

Reservation The reservation of seats shall be as per Government of India norms.

1.6 An Overview of the Syllabus

Introduction The programme for Masters in Statistics, henceforth termed as the Program, aims to impart domain specific knowledge to the learner. Additionally it attempts to inculcate skills into her/him that are beneficial not only for the individual's personal growth but also for the development of the society at large. Whence, the learner as a potential domain expert may grow into efficient appliers and creators of the domain knowledge. Further, it is expected that these experts should be able to work with experts of other domains as tackling the challenges of a growing complex society requires interdisciplinary approaches to problems.

Expected Learning Outcomes The learner at the post graduate level is expected to be well equipped with all necessary skills that make out of him a competent statistician. These include a firm base in mathematical underpinnings, the statistical inference and computational expertise to apply theory for extracting information from large data sets. A specialization in one of the applied fields like Biostatistics, Quantitative Social Sciences, Environmental Sciences, Actuarial Sciences, Industrial Statistics and Operations Management or Economic and Financial Statistics gives an additional advantage.

Computational Expertise As mentioned above computational skills are an added advantage for a practitioner. Therefore data centric programming skills with a knowledge of C/C++/R are essential tools for a statistician. Keeping this in view the curriculum emphasizes R - programming as acquaintance with *R* and its IDE *RStudio* have an edge over other statistical software due to the following reasons:

1. Open source virtually comes without any cost.
2. In tune to STATISTICAL THINKING must for learners of statistics.
3. Plenty of help available online. 5000+ packages.
4. Can be learned easily by self - learning.
5. Most powerful. Whatever is in Statistics is available in R.
6. Updated regularly.

7. Everyone can work with it after an initial training [even non-programmers].
8. Conducive to development of statistical-analytical abilities among learners.
9. Most of the proprietary software provide R interface.

Outline of the Syllabus

The core courses offered in *Semester-I* pertaining to *Mathematical Analysis*, *Probability Theory* and *Distribution Theory* aim at firming the foundation for the study of courses that follow in the subsequent semesters.

The courses mentioned above along with the core courses offered in *Semester-II* namely *Linear Algebra* and *Sampling Distributions* aim at laying firm foundations for taking up courses in *Statistical Inference* , *Design of Sample Surveys / Experiments*, Inferring from data arising out of *Sample Surveys / Designed Experiments*, courses based on *Linear Models*, *Multivariate Statistical Analysis* and *Time Series Analysis*.

The computational aspects of Statistics have been addressed through a series of *Statistical Computing and Data Analysis* courses numbered *I-IV* amounting to 16 credits. The *Statistical Computing and Data Analysis* courses also take care of the practical aspects of the theory courses offered in the respective semesters. In addition, the learner shall opt for minimum two courses in specialized areas in the form of *Discipline Specific Electives*.

The Dissertation to be submitted at the end of the Semester-IV shall be treated as an ambiance to further research.

1.7 Core Courses

S. No.	Title of the Course	Code	Credits
1	Mathematical Analysis	MSTAT-CC-01	3
2	Multi-variable Calculus	MSTAT-CC-02	2
3	Probability Theory	MSTAT-CC-03	4
4	Distribution Theory	MSTAT-CC-04	4
5	Probability Models	MSTAT-CC-05	4
6	Statistical Computing and Data Analysis - I	MSTAT-CC-06	3
7	Linear Algebra	MSTAT-CC-07	2
8	Inference - I	MSTAT-CC-08	4
9	Inference - II	MSTAT-CC-09	4
10	Linear Models and Inference	MSTAT-CC-10	4
11	Design of Sample Surveys	MSTAT-CC-11	4
12	Statistical Computing and Data Analysis - II	MSTAT-CC-12	3
13	Design and Analysis of Experiments	MSTAT-CC-13	4
14	Regression Analysis	MSTAT-CC-14	4
15	Statistical Computing and Data Analysis - III	MSTAT-CC-15	4
16	Multivariate Statistical Analysis	MSTAT-CC-16	4
17	Time Series Analysis	MSTAT-CC-17	4
18	Statistical Computing and Data Analysis - IV	MSTAT-CC-18	4
19	Masters Dissertation	MSTAT-CC-19	4
20	Bridge Course (non - credit)	MSTAT-CC-00	-

1.8 Elective Courses

1.9 Discipline Specific Elective Courses

S. No.	Title of the Course	Code	L-T-P-C
1	Statistical Methods in Demography	MSTAT-DE-01	3-1-0-4
2	Survival Analysis	MSTAT-DE-02	3-1-0-4

1.10 Semester wise Distribution of Generic Elective Courses Offered Under CBCS Exclusively for Students Pursuing Masters Programmes Other Than Statistics

S. No.	Semester	Code	Title	L-T-P-C
1	Odd	MSTAT-GE-01	Epidemiology and Vital Statistics	2-1-0-3
2	Odd	MSTAT-GE-02	Probability Theory and Models	2-1-0-3
3	Even	MSTAT-GE-03	Designing Scientific Studies	2-1-0-3
4	Even	MSTAT-GE-04	Scientific Computing	2-0-1-3

1.11 Scheme of Evaluation

Students undergoing a course shall be examined on a continuous basis by the course instructor during the semester in which the course is conducted. Broadly there shall be *Continuous Internal Evaluations (CIE)* during the semester and an *Semester-End Examination (SEE)* at the completion of the course. The distribution of marks over the CIE and SEE for CC / DSE Courses is as follows:

S. No.	Type of Assessment	Marks
1	CIE	40
2	SEE	60

Aiming to assess values, skills and knowledge imbibed by students, *CIE* and the designing of *SEE* is to be done by the respective course instructor. *CIE* may have different components for internal evaluation with marks distribution given as follows:

S. No.	Mode of Assessment	Marks
1	Term Paper / Sessional Tests	20
2	Other modes	20

The pattern and schedule of Continuous Internal Assessment and evaluation need to be decided by the concerned Faculty and made known to all students.

The distribution of marks for Summer Internship / Masters Dissertation is as follows:

S. No.	Type of Assessment	Marks
1	Periodical Presentation	60
2	Concise Dissertation	20
3	Viva - voce	20
	Total Marks	100

Dissertation / project report will be valued jointly by the supervisor and an examiner external to from a neighboring institution within the city/state/country.

1.12 Semester-wise Structure of the Programme

Semester	Code	Title of the Course	Credits
I	Mathematical Analysis	MSTAT-CC-01	3
	Multi-variable Calculus	MSTAT-CC-02	2
	Probability Theory	MSTAT-CC-03	4
	Distribution Theory	MSTAT-CC-04	4
	Probability Models	MSTAT-CC-05	4
	Statistical Computing and Data Analysis - I	MSTAT-CC-06	3
	Generic Elective - I		3
II	Linear Algebra	MSTAT-CC-07	2
	Inference - I	MSTAT-CC-08	4
	Inference - II	MSTAT-CC-09	4
	Linear Models and Inference	MSTAT-CC-10	4
	Design of Sample Surveys	MSTAT-CC-11	4
	Statistical Computing and Data Analysis - II	MSTAT-CC-12	4
	Generic Elective - II		3
III	Design and Analysis of Experiments	MSTAT-CC-13	4
	Regression Analysis	MSTAT-CC-14	4
	Statistical Computing and Data Analysis - III	MSTAT-CC-15	4
	Discipline Specific Elective - I		4
IV	Multivariate Statistical Analysis	MSTAT-CC-16	4
	Time Series Analysis	MSTAT-CC-17	4
	Statistical Computing and Data Analysis - IV	MSTAT-CC-18	4
	Discipline Specific Elective - II		4
IV	Masters Dissertation	MSTAT-CC-19	4

2 Detailed Syllabus

2.1 Core Courses (CC)

2.1.1 Mathematical Analysis

Functions Sets, metric spaces, functions on metric spaces, limits, continuity, uniform continuity and derivatives.

Sequences Sequences as functions, convergence of sequences; series and sequence of partial sums, convergence of series;

Sequence of Functions Uniform and point-wise convergence, applications related to uniform convergence.

Integration Riemann - Stiltjes integral; fundamental theorem of calculus; improper integrals, beta function, gamma function; Differentiation under the sign of integral - Leibnitz rule.

Text

1. Apostol, T. M. (1985). Mathematical Analysis, Second Edition, Narosa Publishing House, New Delhi.
2. Rudin, Walter. (1976). Principles of Mathematical Analysis, McGraw Hill.

References

1. Bartle, R. G. and Sherbert, D. R. (2000). Introduction to Real Analysis, 3rd edition, John Wiley & Sons, Inc., New York.
2. Tao, T. Analysis, Vol. I, Hindustan Book Agency, Delhi, India.
3. Tao, T. Analysis, Vol. II, Hindustan Book Agency, Delhi, India.

2.1.2 Multi-variable Calculus

Functions of several variables Limit, continuity and constrained maxima-minima.

Integration Evaluation of multiple integrals; transformation of variables and integration.

Text

1. Apostol, T. M. (2007). Calculus Vol. 2, John Wiley & Sons.

References

1. Apostol, T. M. (2007). Calculus Vol. 1, John Wiley & Sons.

2. Ghorpade, Sudhir R. and Limaye, Balmohan V. (2006). A Course in Calculus and Real Analysis, Undergraduate Texts in Mathematics, Springer (SIE), Indian reprint.

2.1.3 Probability Theory

Random Experiment to Random Variables Uncertainty and randomness, random experiment, classes of sets, field, sigma field, minimal sigma field, Borel sigma field and induced sigma field, sequence of sets, lim sup, lim inf of a sequence of sets.

Measure Probability measure and induced probability measure, Carther-dory extension theorem (statement only), Lbesgue and Lbesgue-Stieltjes measures. Measurable and Borel measurable functions, random variables, Integration of a measurable function with respect to a measure, Monotone convergence theorem, Fatous lemma, Dominated convergence theorem.

Sequence of Random Variables Various measures of convergence of sequence of random variables and their interrelationships. Laws of large numbers and central limit theorems. Convergence of rational functions of random variables.

Generating functions Probability generating function, moment generating function, characteristic function and their utility.

Text

1. Ash, R. (1972). Real Analysis and Probability. Academic Press.

2. Rohatgi, V.K. and Saleh, A.K. Md. E. (2001). An Introduction to Probability and Statistics. Wiley.

References

1. Billingsley, P. Probability and Measure.

2. Dudley, R. M. (1989). Real Analysis and Probability, Wadsworth and Brooks/Cole.

3. Kingman, J F C and Taylor, S. J. (1966). Introduction to Measure and Probability. Cambridge University Press.

4. Medhi, J. (2009). Stochastic Processes, New Age International Publishers.

2.1.4 Distribution Theory

Probability Distribution Random variables, distribution function, properties of the distribution function and its relation to pmf/pdf; expected value of a random variable; distribution of a function of random variable. Compound, truncated and mixture distributions.

Order Statistics The distributions and properties of order statistics. Joint and marginal distributions of order statistics. Extreme values and their asymptotic distributions (Statement only) with applications.

Inequalities Chebyshev, Markov, Holder, Jensen, Liapunov, Cauchy-Schwartz, Minkowski, Kolmogrov and Hajek Renyi inequalities.

Multiple Random Variables Random vectors, expectation of a random vector; joint distributions, joint m.g.f., mixed moments, variance covariance matrix and its properties; conditional distribution and independence, conditional expectation and variances, conditional expectation and regression, correlation, multiple correlation and partial correlation; marginal distributions.

Text

1. Hogg, R. V. and Craig, A. T. (2004). Introduction to Mathematical Statistics, Pearson Edward (Indian Print).

2. Rao, C.R.(1973). Linear Statistical Inference and Its Applications, 2/e, Wiley Eastern.

3. Rohatgi, V.K. and Saleh, A.K. Md. E. (2001). An Introduction to Probability and Statistics. Wiley.

References

1. Anderson, T.W. (1987). An Introduction to Multivariate Statistical Analysis, 2^{nd} edn., Wiley.

2. Dudewicz, E.J. and Mishra, S.N. (1988) : Modern Mathematical Statistics, Wiley, International Students Edition.

3. Feller, W. An Introduction to Probability Theory and Its Applications, Wiley.

2.1.5 Probability Models

Exponential Family of Distributions General expression for the PMF/PDF and properties.[Univariate probability models namely, Bernoulli, binomial, Poisson, geometric, exponential, gamma, negative exponential and normal may be illustrated].

Other Univariate Probability Models hypergeometric, negative binomial, beta, lognormal, Pareto, Weibull, Laplace, Cauchy and logistic.

Distribution of statistics when a random sample is drawn out of a normal distribution $N(\mu, \sigma^2)$ Non - central and central chi-square, non - central and central F, non-central and central t distributions. Distributions of linear and quadratic forms, conditions for independence of a quadratic form and linear/quadratic form and related distribution theory.

Multivariate Probability Models Multinomial distribution and multivariate normal distribution; conditional distribution and regression; multiple correlation and partial correlation; marginal distributions.

Text

1. Rao, C.R. (1973). Linear Statistical Inference and Its Applications, 2/e, Wiley Eastern.
2. Rohatgi, V.K. and Saleh, A.K. Md. E. (2001). An Introduction to Probability and Statistics. Wiley.
3. Searle, S.R. Linear Models, John Wiley & Sons.

References

1. Anderson, T.W. (1987). An Introduction to Multivariate Statistical Analysis, 2^{nd} edn., Wiley.
2. Kotz, S., N. Balakrishnan and N. L. Johnson. (2005). Univariate Discrete Distributions 3rd Edition, John Wiley & Sons.
3. Kotz, S., N. Balakrishnan and N. L. Johnson. (2005). Continuous Univariate Distributions 2nd Edition, Volume 1, John Wiley & Sons.
4. Kotz, S., N. Balakrishnan and N. L. Johnson. (2005). Continuous Univariate Distributions 2nd Edition, Volume 2, John Wiley & Sons.
5. Kotz, S., N. Balakrishnan and N. L. Johnson. Discrete Multivariate Distributions, John Wiley & Sons.
6. Kotz, S., N. Balakrishnan and N. L. Johnson. (2005). Continuous Multivariate Distributions, Volume 1, Models and Applications, John Wiley & Sons.
7. McCullagh, P. and Nelder, J. A. Generalized Linear Models.

2.1.6 Statistical Computing and Data Analysis - I

R language Essentials Expressions and objects, Assignments, creating vectors, vectorized arithmetic, creating matrices, operations on matrices, lists, data frames - creation, indexing, sorting and conditional selection ; examples.

R Programming conditional statements - if and if else; loops - for, while, do-while; functions - built-in and user defined; Data entry - reading from text file, data editor; examples.

Descriptive Statistics and Graphics Obtaining summary statistics; generating tables; Bar plots, Pie charts, Box plots, Histogram; exercises.

Probability and Distributions Random sampling and combinatory; obtaining density, cumulative density and quantile values for discrete and continuous distributions; generating samples from discrete and continuous distributions; Plotting density and cumulative density curves; Q-Q plot.

Correlation Pearson, Spearman and Kendalls correlation; Regression fitting, obtaining residuals and fitted values; one and two sample tests for mean and variance one way and two way ANOVA.

Lab Exercises

1. Operations on vectors and matrices
2. Creating and manipulating data frames.
3. Writing user defined functions for finding arithmetic mean, median, factorial, matrix addition and multiplication.
4. Bar and Pie charts.
5. Box plots for single and multiple groups.
6. Density and cumulative density plots for Binomial, Poisson, Normal and exponential distributions.
7. Checking Normality using Histogram and Q-Q plot.
8. Correlation coefficient Pearsons, Spearman and Kendalls Tau.
9. Fitting simple linear and multiple linear regressions.
10. One sample and two sample t test.
11. One way and Two way ANOVA.

Texts

1. Crawley, J.(2007) *The R Book by Michael*, John Wiley and Sons.
2. Goon, A.M., Gupta, M.K. and Das Gupta, B. (2005). *Fundamentals of Statistics, Vol. I*, The World Press Pvt. Ltd., Calcutta.
3. Dalgaard, P. *Introductory Statistics with R* , Springer, 2nd edition, 2008.
4. Ross, S. M. *Simulation*.

References

1. Chambers, J. M. and Hastie, T. J (Editors).(1997). *Statistical Models in S*. Chapman and Hall.
2. Chambers, J. M.(1998). *Programming with Data: A Guide to S Language*. Springer.
3. Venables, W. N. and Ripley, B. D. (2000). *S Programming*, Springer.
4. *www.r-project.org*

2.1.7 Linear Algebra

Vector Spaces 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, Vector spaces with an inner product, Gram-Schmidt orthogonalization process, orthonormal basis and orthogonal projection of a vector.

Linear Transformations and Matrices 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.

Some Special Topics on Matrices

Positive Definite Matrices triangular reduction of a positive definite matrix.

Characteristic Roots and Vectors Cayley - Hamilton theorem, minimal polynomial, similar matrices, algebraic and geometric multiplicity of a characteristic root, spectral decomposition of a real symmetric matrix, reduction of a pair of real symmetric matrices, Hermitian matrices.

Generalised Inverse of a Matrix Hermite canonical form, generalized inverse, Moore-Penrose generalized inverse, Idempotent matrices, Solutions of matrix equations.

Real Quadratic Forms Reduction and classification of quadratic forms, index and signature, extrema of quadratic forms

Singular Value Decomposition Singular values and singular value decomposition, Jordan decomposition.

Calculus Vector and matrix differentiation.

Text

1. Hoffman, K. and Kunze, R. (1971). Linear Algebra, 2nd ed., Prentice Hall, Inc.

2. Graybill, F.A. (1983): Matrices with Applications in Statistics, 2nd Ed. Wadsworth.

3. Rao, A. R.and Bhimsankaram,P. Linear Algebra, Hindustan Book Agency, Delhi, India.

4. Searle, S.R. (1982): Matrix Algebra for Statistical Applications, John Wiley and Sons inc.,

References

1. Bellman, R. (1970). Introduction to Matrix Analysis, 2nd ed. McGraw Hill.

2. Halmos, P.R. (1958): Finite Dimensional Vector Spaces, 2nd ed. van. Nortrand Company Inc.

3. Rao, C. R. (1973). Linear Statistical Inference and its Applications, Wiley.

4. Rao, C. R. and Mitra, S. K. (1971). Generalized Inverse of Matrices and its Applications, John Wiley and Sons, Inc.

5. Shanti Narain: A text book of matrices, S. Chand and Company (Pvt.) Ltd.

2.1.8 Inference - I

Decision Problems Parametric models, point estimation, tests of hypotheses and interval estimation viewed as decision problems with given loss functions, joint distribution of a sample and induced sampling distribution of a statistic.

Likelihood Functions and Sufficiency information in data about the parameters as variation in likelihood functions, concept of no information, sufficiency, Neyman factorizability criterion, likelihood equivalence, minimal sufficient statistic, ancillary statistics, completeness, exponential families and Pitman families, invariance property of sufficiency under one-one transformation of sample space and parameter space. Fisher's information for one and several parameters models, Basu's theorem, Jensen's inequality.

Choice of Estimators Based on Unbiasedness Minimum variance, mean squared error, minimum variance unbiased estimators, Rao - Blackwell theorem, completeness, Lehmann Scheffe theorem, necessary and sufficient conditions for MVUE, Cramer Rao lower bound approach;

Methods of Estimation Maximum likelihood method, consistency of maximum likelihood estimates (Bahadur-Wald), large sample properties of maximum likelihood estimates; methods of moments and percentiles;

Consistent Estimators Consistent Estimation of real and vector valued parameter. Invariance of Consistent estimator under continuous transformation, Consistentancy of estimators by method of moments, and method of percentiles, Mean squared error criterion, Asymptotic relative efficiency,

Error probabilities and their rates of convergence, Minimum sample size required to attain given level of accuracy.

Text

1. Casella, G and Berger, R. L. (2002). Statistical Inference, 2nd edition. Duxbury Press

2. Kale, B.K. (1999). A first course in Parametric Inference, Narosa Publishing House.

References

1. Lehmann, E .L. and Casella, G. (1998).Theory of Point Estimation, Springer International.

2. Rao, C. R. (1973). Linear Statistical Inference and its Applications, Wiley.

3. Rohatgi, V.K. and Saleh, A.K. Md. E. (2001). An Introduction to Probability and Statistics. Wiley.

2.1.9 Inference - II

Tests of Hypotheses

Basic Concepts Test , critical regions, test functions, two kinds of errors, size function, power function, level.

Developing Tests Neyman - Pearson Lemma, MP test for simple null against simple alternative hypothesis. UMP tests for simple null hypothesis against one sided alternatives and for one sided null against one sided alternatives in one parameter exponential family. Extension of these results to Pitman family when only upper or lower end depends on the parameter and to distributions with MLR property, non-existence of UMP test for simple null against two sided alternatives in one parameter exponential family; UMP and UMP unbiased tests. Detailed analysis in exponential models; LMP tests, Neyman structure, P-values and its examples, Sequential Probability Ratio test, Walds SPRT with prescribed errors of two types, Stein's two-stage procedure.

Likelihood Ratio Test (LRT) Asymptotic distribution of LRT statistic.

Interval Estimation Confidence level, construction of confidence intervals using pivots, shortest expected length confidence interval, uniformly most accurate one sided confidence interval and its relation to UMP test for one sided null against one sided alternative hypotheses; non-parametric confidence intervals for percentiles.

Text

1. Casella, G., and Berger, R. L. (2002). Statistical Inference, 2^{nd} edition. Duxbury Press.

2. Gibbons, J.D. (2003). Nonparametric Statistical Inference, 4^{th} Edition, Marcel Dekker.

3. Kale, B.K. (1999). A first course in Parametric Inference, Narosa Publishing House.

References

1. Lehmann, E L. (1986). Testing Statistical Hypotheses, Springer.

2. Rao, C. R. (1973). Linear Statistical Inference and its Applications, Wiley.

3. Randles, R.H. and Wolfe, D.A. (1979). Introduction to the Theory of Nonparametric Statistics, Wiley.

4. Rohatgi, V.K. and Saleh, A.K. Md. E. (2001). An Introduction to Probability and Statistics. Wiley.

2.1.10 Linear Models and Inference

Prerequisites *Inference - I, Inference - II, Probability Models and Linear Algebra*

Gauss-Markov Models Linear statistical models: illustrations of Gauss-Markov models.

Estimation under Linear Model Setup Method of least squares, normal equations and least square estimators; full rank and non-full rank models; estimable linear parametric functions, g-inverse and solution of normal equations; error space and estimation space; estimation with correlated observations; least squares estimates with restriction on parameters; simultaneous estimates of linear parametric functions; confidence intervals for least squares estimates, BLUEs, variances and covariances of BLUEs; Estimation of error variance.

Testing of Hypotheses under Linear Model Setup Fundamental theorems of least squares and applications to the tests of linear hypotheses; Fisher-Cochran theorem; tests of hypotheses for one and more than one linear parametric functions; Analysis of Variance, Power of F-test, Multiple comparison tests due to Tukey and Scheffe, simultaneous confidence intervals; analysis of covariance in a general Gauss-Markov model.

Random and Mixed Effects Linear Models Linear statistical models: random effects linear models and estimation of variance components, tests for variance components.

Text

1. Searle, S. R. Linear Models.

References

1. Bapat, R. B. Linear Algebra and Linear Models, Hindustan Book Agency, Delhi, India.

2. Goon, A.M., Gupta, M.K. and Das Gupta, B. (2005). An Outline of Statistical Theory, Vol. II, The World Press Pvt. Ltd., Calcutta.

3. Giri,N. (1986). Analysis of Variance, South Asian Publishers

4. Graybill, I.A. (1961). An Introduction to Linear Statistical Models, Vol. 1, McGraw Hill Book Co. Inc.

5. Rao, C. R. (1973). Linear Statistical Inference and its Applications, Wiley.

6. Renchner, A. C. and Schaalje, G. B. (2008). Linear Models in Statistics (Second edition), John Wiley and Sons.

7. Searle, S.R., Casella, G. and McCulloch, C.E. (1992). Variance Components. John Wiley, New York.

2.1.11 Design of Sample Surveys

Introduction to Unified Theory of Sampling sampling designs and sampling schemes, correspondence;

Basic Finite Population Sampling Techniques srs wr/wor, stratified, systematic sampling; related results on estimation of population mean/total. Allocation problem in stratified sampling; Ratio estimators; Regression estimators.

Double Sampling Double sampling on successive occasions, double sampling for stratification; cost and variance functions;

Two-Stage Sampling Two-stage sampling with equal number of second stage units. Two-stage sampling with unequal number of second stage units, multistage sampling;

Cluster Sampling Single stage cluster sampling: clusters of equal sizes, comparisons of precision made from survey data, variance, cost function; Cluster units of unequal sizes - SRS of clusters: unbiased estimate, ratio to size estimate.

Text

1. Des Raj and Chandhok (1998). Sampling Theory. Narosa.

2. Mukhopadhay, P. (2009) Theory and methods of survey sampling, Second edition, PHI Learning Pvt Ltd., New Delhi.

References

1. Cochran, W.G. (1999) Sampling Techniques, Third edition, John Wiley & Sons.

2. Chaudhuri, A. and J.W.E. Vos (1988) Unified Theory and Strategies of Survey Sampling. North-Holland, Amsterdam.

3. A. and R. Mukerjee (1988) Randomized Response : Theory and Techniques, New York : Marcel Dekker Inc.

4. Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953). Sample Survey Methods and Theory, Volume II, John Wiley.

5. Heyday, A. S. and Sinha, B. K. (1991). Design and inference sampling in finite population. Wiley.

6. Mukhopadhyay, P. (1996). Inferential problems in survey sampling. New Age International (P).

7. Mukhopadhyay, P. (1998). Small area estimation in survey sampling. Narosa.

8. Murthy, M. N. (1977). Sampling Theory & Methods. Statistical Publishing Society, Calcutta.

9. Sampath, S. (2001). Sampling theory and methods, Alpha Science International Ltd., India.

10. Singh, D. and Chaudhary, F. S. (1986). Theory and Analysis of Sample Survey Designs. New Age International Publishers.

11. Sukhatme, P. V. et al. (1984). Sampling theory of surveys with applications. Iowa State Univ. Press.

12. Wolter, K. M. (1985). Introduction to variance estimation. Springer-Verlag.

2.1.12 Statistical Computing and Data Analysis - II

Prerequisites *Statistical Computing and Data Analysis - I*

Theory of Estimation

1. MLE and Standard error of ML estimators.
2. MLE through the method of successive approximation.
3. Method of Moments.
4. Method of Least square
5. Interval Estimation: Confidence interval for mean, difference of means and ratio of variances.

Design of Sample Survey

1. Simple random sampling without replacement - Estimation of the population total and its variance.
2. Ratio, Regression and Difference estimation.
3. Stratified sampling SRS- Equal, Proportional allocations.
4. Linear and circular systematic sampling.
5. Cluster sampling of equal sizes

Texts

1. James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R, Springer.
2. Press, W. H., Teukolsky, S. A., Vellering, W. T. and Flannery, B. P. (1993). Numerical Recipes in C, II edition, Cambridge University Press.
3. Ross, S. M. Simulation.

References

1. Chambers, J. M. and Hastie, T. J (Editors).(1997). Statistical Models in S. Chapman and Hall.
2. Venables, W. N. and Ripley, B. D. (2000). S Programming, Springer.
3. www.r-project.org

2.1.13 Design and Analysis of Experiments

Prerequisites *Inference - I, Inference - II, Linear Inference*

Designed Experiments and Block Designs Two way classification with equal number of observations per cell (with and without interaction), BIBD intra block analysis, incidence matrix, connectedness balanced, orthogonality for two way classification with unequal number of observations per cell, random effect models for one factor, estimation of variance components.

2^k **Full Factorial Experiments** Diagrammatic presentation of main effects and first and second order interactions, model, analysis of single as well as more than one replicates using ANOVA, total confounding of 2^k design in 2p blocks p 2, partial confounding in 2p blocks; p = 2, 3, fractional factorial experiments, resolution of a design (III, IV & V), aberration of a design.

3^2 **designs** contrasts for linear and quadratic effects , statistical analysis of 3^2 design. Response surface methodology (RSM): linear and quadratic model, stationary point, canonical analysis, central composite designs (CCD), ridge 24 systems, multiple responses, concept of rotatable designs.

Taguchi Methods Concept of loss function, S/N ratio, orthogonal arrays, triangular tables, linear graphs, inner and outer arrays; nested and split plot designs.

Texts

1. Das, M.N. and Giri, N. (1979). Design and Analysis of Experiments, Wiley Eastern.

2. Dey, A. Theory of Block Designs.

3. Dean, A. and Voss, D. (1999). Design and Analysis of Experiments, Springer.

4. Ogawa J. (1974). Statistical Theory of the Analysis of Experimental Design, Marcel Dekker.

References

1. Atkinson, A. C. and Donev, A. N. (1992). Optimal Experimental Designs. Oxford University Press.

2. George E. P. Box, Draper N.R. (1987). Empirical Model-Building and Response Surfaces, Wiley.

3. Hicks, C.R., Kenneth V. and Turner, Jr. (1999). Fundamental Concepts in the Design of Experiments, Oxford University Press.

4. John P.W.M. (1971). Linear Models, Wiley.

5. Kshirsagar A.M. (1983). Linear Models, Marcel Dekker .

6. John, P. W. M. (1971). Statistical Design and Analysis of Experiments. MacMillan.

7. Joshi, D. D. Linear Estimation and Design of Experiments.

8. Montgomery, C.D. (1976). Design and Analysis of Experiments, Wiley, New York.

9. Pukelsheim, F. (1993). Optimal Design of Experiments. Wiley.

10. Shah, K. R. and Sinha, B. K. (1989). Theory of Optimal Designs. Springer-Verlag.

2.1.14 Regression Analysis

Prerequisites *Linear Inference, Statistical Computing and Data Analysis - I*

Simple Linear Regression Assumptions, least square (LS) estimators of parameters, standard error of estimators, testing of hypothesis for coefficient of regression, s.e. of prediction, testing of hypotheses about parallelism (Slopes), equality of intercepts, congruence, extrapolation, optimal choice of independent variables, diagnostic checks and correction: graphical technique, tests for normality , uncorrelatedness, homoscadasticity, lack of fit. modifications like polynomial regression, transformations of dependent or independent variables , weighted LS, inverse regression.

Multiple Regression Standard Gauss-Markov setup, least square estimation, error and estimation spaces, variance and covariance of LS estimators, properties of LS estimators, estimation of error variance, case with correlated observation, LS estimation with restriction on parameters, simultaneous estimation of linear parametric functions, testing of hypothesis for one and more than one linear parametric functions, confidence intervals and regions. Mallows Cp, forward, backward selection method.

Regression Diagnostics Multicollinearity: consequences, detection and remedies, autocorrelation consequences, Durbin Watson test, estimation of parameters in autocorrelation. Multiple correlation, adjusted multiple correlation coefficient, null distribution of simple correlation and multiple correlation coefficient, partial correlation coefficient and its relation with multiple correlation coefficient, test for significance of simple , multiple and partial correlation coefficients, variable selection procedures. Residual and residual diagnostics, transformation of variables: Box- Cox power Transformation, generalized weighted least sequence.

Non-linear Regression Non-linear least squares transformation to a linear model, statistical inference in non-linear regression; Logistic regression: Logit transform, ML estimation, tests of hypothesis, Wald test, LR test, score test, test for overall regression, introduction to link functions such as binomial, inverse binomial, inverse Gaussian and Gamma.

Texts

1. Draper, N.R. and Smith H. (1998). Applied Regression Analysis, 3^{rd} Ed. Wiley.

2. Hosmer, D. W. and Lemeshow, S. (1989) Applied logistic regression, John Wiley

3. McCullagh, P. and Nelder, J. A.(1989) Generalized linear models, Chapman and Hall

4. Neter, J.; Wasserman, W. and Kutner, M.H.(1985) Applied linear statistical models

5. Ratkowsky, D. A.(1983) Nonlinear regression modeling (Marcel Dekker

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

References

1. Cook, R.D. and Weisberg, S. (1982). Residuals and Inference in Regression, Chapman and Hall.
2. Montogomery, D. C., Peck, E. A. and Vining, G. G. Introduction to Linear Regression Analysis.
3. James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R, Springer.
4. Ryan, T. P. Modern Regression Methods.
5. Searle, S.R. Linear Models, John Wiley & Sons.

2.1.15 Statistical Computing and Data Analysis - III

Prerequisites *Statistical Computing and Data Analysis - I, Statistical Computing and Data Analysis - II*

Testing of Statistical Hypothesis

1. Construction of randomized and nonrandomized MP, UMP and UMPU tests of hypotheses and drawing the power curves.
2. Construction of SPRT and its OC and ASN curves. Based on R
3. Non parametric tests: Runs Test, Kolmogorov - Smirnov test, Mann - Whitney U test, Wilcoxon Sign Test, Chi-Square Test for independence of Attributes, Median test, Kruskal Wallis test and Friedmans test.

Regression

1. Simple Regression
2. Multiple Regression
3. Regression Diagnostics
4. Logistic Regression
5. Poisson Regression

Texts

1. Johnson, R.A. and Wichern, D.W. (1990). Applied Multivariate Statistical Analysis, Prentice Hall.

2. Montogomery, D. C., Peck, E. A. and Vining, G. G. Introduction to Linear Regression Analysis.

References

1. Chambers, J. M. and Hastie, T. J (Editors).(1997). Statistical Models in S. Chapman and Hall.

2. Venables, W. N. and Ripley, B. D. (2000). S Programming, Springer.

3. www.r-project.org

2.1.16 Multivariate Statistical Analysis

Prerequisites *Inference - I and Inference - II*

Multivariate normal distribution Marginal and conditional distributions - characteristic function. Maximum likelihood estimation of the parameters of Multivariate Normal and their sampling distributions - Inference concerning the mean vector when covariance matrix is known.

Total, Partial, Multiple Correlation in the Multivariate Setup MLEs of Total, Partial and Multiple correlation coefficients. Sampling distributions of Total and Multiple Correlation in the null case. Hotelling T2 statistic, derivation and its distribution - Uses of T2 statistic - relation between T2 and D2 - Mahalanobis D2 statistic and its distribution.

Generalized Variance Wishart distribution (statement only) Properties of Wishart distribution - Test for covariance matrix - Test for equality of covariance matrices.

Classification Problems Classification into one of two populations (known and unknown dispersion matrix) - Classification in to one of several populations - Fishers Linear discriminant function.

Principal Components Properties, Extraction of Principal components and their variances Canonical correlation - Estimation of canonical correlation and variates. Factor analysis - Mathematical model- Estimation of Factor Loadings - Concept of factor rotation - Varimax criterion.

Texts

1. Anderson, T.W. (1984) An Introduction to Multivariate Statistical Analysis, John Wiley.
2. Johnson, R. A and. Wichern D.W (2007): Applied Multivariate Statistical Analysis, 6 /e, Prentice-Hall of India Private Ltd., New Delhi.

References

1. Alvin C. Rencher(2002): Methods of Multivariate Analysis, 2/e, Wiley Interscience.
2. Giri, N. Multivariate Statistical Inference, Academic Publishers.
3. Jolliffe I.T.(2002): Principal Component Analysis, 2/e, Springer.
4. Ksheersagar, A. M. Multivariate Analysis. Marcell Dekkar.
5. Morrison, D.F. Multivariate Analysis.
6. Rao, C.R(1998): Linear Statistical Inference and its Applications, Wiley Eastern Ltd.
7. Seber, G.A.F. (1977) Multivariate Observations, Wiley.
8. Srivastava M.S. and Khatri C.G.(1979):Introduction to Multivariate Analysis, Elsevier.

2.1.17 Time Series Analysis

Time Series Exploratory time Series analysis, tests for trend and seasonality. exponential and Moving average smoothing. Holt -Winters smoothing. Forecasting based on smoothing, adaptive smoothing. Time - series as a discrete parameter stochastic process. Auto covariance and autocorrelation functions and their properties, Portmanteau tests for noise sequences, transformation to obtain Gaussian series.

Stationary Processes General linear processes, moving average (MA), auto regressive (AR), and autoregressive moving average (ARMA). Stationarity and inevitability conditions. nonstationary and seasonal time series models: Auto regressive integrated moving average (ARIMA) models, Seasonal ARIMA (SARIMA) models, Transfer function models (Time series regression)

Forecasting in Time Series Models Durbin-Levinson algorithm, innovation algorithm (without proof). Estimation of mean, auto covariance and autocorrelation functions, Yule-Walker estimation, Estimation of ARIMA models parameters, maximum likelihood method, large sample theory (without proofs). Choice of AR and MA periods, FPE, AIC, BIC, residual analysis and diagnostic checking. Unit-root non stationarity, unit-root tests.

Multivariate Time Series Models VAR models, vector ARMA models. Conditional heteroschedastic models, ARCH and GARCH, properties, examples, estimation & forecasting, extensions of ARCH & GARCH.

Texts

1. Box, G.E.P and Jenkins G.M. (1970), Time Series Analysis, Forecasting and Control, Holden-Day.

2. Brockwell, P.J. and Davis, R. A. Introduction to Time Series Analysis, Springer

3. Hamilton N. Y. (1994). Time Series Analysis. Princeton University press. Princeton

4. Lutkepohl, H. and Kratzing, M. (Ed.) (2004). Applied Time Series Econometrics, Cambridge University Press, Cambridge

5. Shumway, R. H.and Stoffer D. S. (2010) Time Series Analysis & Its Applications, Springer, New York. 8. Tsay, R. S. (2010). Analysis of Financial Time

6. Kendall, M.G. (1978). Time Series, Charler Graffin.

References

1. Brockwell, P.J and Davis R.A. (1987). Time Series: Theory and Methods, Springer-Verlag.

2. Chatfield, C. (2004). The Analysis of Time Series - An Introduction, Sixth edition, Chapman and Hall.

3. Fuller, W. A. Introduction to statistical time series.

2.1.18 Statistical Computing and Data Analysis - IV

Multivariate Statistical Analysis

1. Test for Mean vector when dispersion matrix in known (Single and Two sample).

2. Hotelling T^2 test (One and two sample).

3. Test for covariance matrices.

4. Discriminant Analysis.

5. One way MANOVA.

6. Principal Component Analysis.

7. Canonical Correlation Analysis.

8. Factor Analysis

Design and Analysis of Experiments

1. Design and analysis of a BIBD.

2. Design and analysis of a 2^k factorial experiment.

3. Design analysis of 3^2 factorial experiment.

4. Response surface designs.

Time Series Analysis

1. Exploratory time Series analysis.
2. Auto regressive integrated moving average (ARIMA) models
3. Seasonal ARIMA (SARIMA) models.
4. Forecasting with time series models.

Texts

1. Das, M.N. and Giri, N. (1979). Design and Analysis of Experiments, Wiley Eastern.
2. Johnson, R. A and. Wichern D.W (2007): Applied Multivariate Statistical Analysis, 6 /e, Prentice-Hall of India Private Ltd., New Delhi.
3. Kendall, M.G. (1978). Time Series, Charler Graffin.

References

1. Box, G.E.P and Jenkins G.M. (1970), Time Series Analysis, Forecasting and Control, Holden-Day.
2. Montgomery, C.D. (1976). Design and Analysis of Experiments, Wiley, New York.

2.1.19 Masters Dissertation

The student pursuing the program shall work for a dissertation under the supervision of a faculty member of the Department of Statistics, IGNTU. The work may be started after the Semester -II and be submitted before the culmination of Semester - IV.

2.1.20 Bridge Course

Note This is a non-credit course. The course instructor may recommended this course for students who have not studied statistics courses at the under graduate level.

Data and Random Samples Uncertainty, randomness and variation; Concepts of experiments: Deterministic and Probabilistic (concept only); Data generated in controlled conditions and that generated freely in nature; Data generated by random mechanisms (for ex sampling); Prospective studies and Retrospective Studies; Data Types (concepts only with real life examples):Linear Data: Cross sectional, Time series, Longitudinal and Panel Data. Circular Data. Merits of a sample surveys over census surveys. Various useful sampling designs and how these are practically applied in field works. Examples on Sampling Designs like Simple Random Sampling, Proportional Allocation in Stratified Random Sampling, Systematic Sampling and Probability Proportional to Size Sampling [When to use what], weighting the observation and using the weights in the analysis of sample based data.

Descriptive Statistics Presentation of Data and Graphical Tools: Univariate frequency distribution, Tabulation of data and Graphical representation of linear and circular data. Summarizing a frequency distribution in terms of summary measures for Central tendency, Variance; Utility of third and fourth central moments; The concept of change in origin and scale; Standardization of a variable; The concept of association in simultaneous study of two or more variables; Various measures of association depending on the nature of variables involved.

Probability Models Concept of a probability distribution, Normal distribution as a model for data arising in social sciences, probability curve for a standardized normal variable: its area properties and uses.

Statistical Inference Inferences about Population Parameters based on Sample Data. Concept of STATISTIC and Examples on how sampling generates a SAMPLING DISTRIBUTION of a statistic and STANDARD ERROR. Estimates of population mean, variance, proportion and coefficient of correlation (with confidence regions). Tests of hypothesis regarding population mean, variance, proportion and coefficient of correlation.

References

1. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I, 8th Edn. The World Press, Kolkata.

2. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. II, 8th Edn. The World Press, Kolkata.

2.2 Discipline Specific Elective Courses (DE)

2.2.1 Statistical Methods in Demography

Demographic Data and Measures Sources of demographic Statistics, Basic demographic measures: Ratios, Proportions and percentages, Population Pyramids, Sex ratio Crude rates, Specific rates, Labour force participation rates, Density of population, Probability of dying.

Life tables Construction of a life table, Graphs of l_x, q_x , d_x , Functions L_x , T_x and E_x. Abridged life tables Mortality: Rates and Ratios, Infant mortality, Maternal mortality, Expected number of deaths, Direct and Indirect Standardization, Compound analysis, Morbidity.

Fertility Measures of Fertility, Reproductivity formulae, Rates of natural increase, Fertility Schedules, Differential fertility, Stable populations, Calculation of the age distribution of a stable population, Model Stable Populations.

Population Projections Component method, Mortality basis for projections, Fertility basis for projections, Migration basis for projections, Ageing of the population, Estimation of demographic measures from incomplete data.

Text

1. Ram, F. and Pathak, K. B. (1998): Techniques of Demographic Analysis, 2^{nd} Ed, Himalaya Publishing House, Bombay.

2. Alho, J., and Spencer, B. (2005). Statistical Demography and Forecasting. Springer-Verlag, New York.

References

1. Bhende, A. A. and Kanitkar,T. (2003). Principles of Population Studies, Sixteenth Revised Edition, Himalaya Publishing House, Mumbai.

2. Keyfilz, N. and Caswell, H. (2005) Applied Mathematical Demography, Third edition, Springer.

3. Pollard, A.H. Yusuf, F. and Pollard, G.N. (1990). Demographic Techniques, Pergamon Press.

4. Retherford, R. D. and Choe, M. K. (1993). Statistical Models for Causal Analysis, John Wiley & Sons, Inc.

5. Siegel, J. S. and Swanson, D. A. (2004). The Methods and Materials of Demography, Second Edition, Elsevier Science, USA.

6. United Nations Manuals on Demography.

7. Weeks, J. R. (2005). Population: An Introduction to Concepts and Issues, Nineth Edition, Wadsworth Publishing Company, Belmont, California.

2.2.2 Survival Analysis

Basic concepts Concepts of time, order and random censoring and likelihood in these cases - Life distributions Exponential, Gamma, Weibull, Lognormal, Pareto, Linear Failure rate Parametric inference (Point estimation, Scores, MLE).

Life tables failure rate, mean residual life and their elementary properties - Ageing classes and their properties - Bathtub Failure rate.

Estimation Estimation of survival function Actuarial Estimator - Kaplan-Meier Estimator - Estimation under the assumption of IFR / DFR - Tests of exponentiality against non- parametric classes Total time on test, Deshpande test.

Two Sample Problems Two sample problem: Gehan test, Log rank test. Mantel Haenszel test, Tarone Ware tests. Semi- parametric regression for failure rate Coxs proportional hazards model with one and several convariates - Rank test for the regression coefficients.

Texts

1. Miller, R.G. (1981) : Survival analysis, John Wiley

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

3. Elisa T.Lee, John Wenyu Wang and Timothy Wenyu Patt (2003): Statistical Methods for Survival Data Analysis, 3/e, Wiley Inter Science.

4. Klein P. John and Moeschberger(2003): Survival Analysis: Techniques for Censored and Truncated Data, 2/e, Springer.

References

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

2. Elandt Johnson, R.E. Johnson N.L.(1999): Survival Models and Data Analysis, John Wiley and sons.

3. Kalbfleisch J.D. and Prentice R.L.(2003), The Statistical Analysis of Failure Time Data, John Wiley.

4. Lawless J.F. (2002): Statistical Models and Methods for Life Time Data, 2/e, John Wiley & Sons.

5. Xian Liu(2012): Survival Analysis Models and Applications- John Wiley & Sons.

2.3 Generic Electives (GE)

2.3.1 Epidemiology and Vital Statistics

Prerequisite High School Algebra

Population Composition and Distribution Population density, percentage distribution, population potential, Lorenz curve, Gini concentration ratio; Age distribution: per cent distribution, per cent change in distribution, index of relative difference, index of dissimilarity; Population pyramid.

Epidemiology Definition and scope, health and disease; Population at risk, prevalence rate, incidence rate, risk, case-fatality; Morbidity; Disability; Comparing disease occurrence: risk difference, attributable fraction, population attributable risk, relative comparison.

Vital Statistics Rates of vital events: measurement of mortality, measurement of fertility, measurement of population growth.

Life Tables Life Tables: description and construction; Abridged life tables.

References

1. Bonita, R., Beaglehole, R., and Kjellström, T. (1993). Basic epidemiology. World Health Organization, Geneva.

2. Goon A.M., Gupta M.K. and Dasgupta B. (2002). Fundamentals of Statistics, Vol. II, 8th Edn. The World Press, Kolkata.

3. Ram, F. and Pathak, K. B. (1998). Techniques of Demographic Analysis, 2^{nd} Ed, Himalaya Publishing House, Bombay.

2.3.2 Probability Theory and Models

Prerequisite High School Algebra

Elementary Probability Theory Sample spaces, events, probability; Axioms of Probability, conditional probability and independence, Bayes Theorem.

Random Variables Discrete and continuous random variables; distribution function; Random vectors, joint, conditional and marginal distributions, functions of random vectors.

Expectation Moments of a probability distribution, moment generating functions and characteristic functions, conditional expectation.

Convergence of Random Variables Modes of convergence, weak and strong laws of large numbers, central limit theorem.

References

1. P. Billingsley: Probability and Measure (2 nd edition), John Wiley & Sons.

2. S M Ross: Introduction to Probability Models, Academic Press. H M Taylor and S Karlin: An Introduction to Stochastic Modelling, Academic Press.

3. P.G. Hoel, S.C. Port and C.J. Stone: Introduction to Probability, Universal Book Stall, New Delhi.

4. J.S. Rosenthal: A First Look at Rigourous Probability Theory, World Scientific.

5. M. Woodroofe: Probability with Applications, McGraw-Hill.

2.3.3 Designing Scientific Studies

Prerequisite High School Algebra and Elementary Probability

Introduction to the Design of Experiments Basic terminology and principles; Construction and analyses of Completely Randomized Design, Randomized Block Design, Latin Square Design.

Design of Factorial Experiments Construction and analyses of 2^2 and 2^3 factorial design.

Design of Sample Surveys Sampling designs and sampling schemes; basic finite population sampling techniques [SRS WR/WOR, stratified, systematic] and related results on estimation of population mean/total. Allocation problem in stratified sampling.

Designing Epidemiological Studies Descriptive studies: ecological studies, cross-sectional studies, case-control studies, cohort studies; Experimental studies: randomized controlled trials, community trials, field trials; Potential errors in epidemiological studies: random error, systematic error, bias and confounding; Validity and reliability.

References

1. Bonita, R., Beaglehole, R., and Kjellström, T. (1993). Basic epidemiology. World Health Organization, Geneva.

2. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. II, 8th Edn. The World Press, Kolkata.

2.3.4 Scientific Computing

Prerequisite High School Algebra

Programming in Matlab/Scilab Executing a function, Global variables, Loops, branches and control flow, Multidimensional matrices.

Applications in Linear Algebra Solving linear system of equations, Gaussian elimination, matrix factorization.

Matrices and Vectors Matrix and array operations, character strings, Finding the determinant, Eigen values and Eigen vectors of a matrix.

Other Applications Curve fitting and interpolation; Numerical integration; Finding the roots of polynomial equations by various numerical methods.

References

1. Rudra Pratap: Getting started with Matlab, Oxford University Press.

2. Brian R Hunt, Ronald L Lipsman, Jonathan M Rosenberg: A Guide to Matlab for beginners and experienced users, Cambridge University Press.

3. S R Otto and J P Denier: An Introduction to programming and Numerical Methods in Matlab, Springer Inc.

4. Won Young Yang, Wenwu Cao, Tae-Sang Chung, and John Morris: Applied Numerical Methods using Matlab, John Wiley & Sons Inc.